

**An Empirical Analysis of Income Dynamics among Men
in the PSID: 1968–1989**

John Geweke
Department of Economics
University of Minnesota
and Federal Reserve Bank of Minneapolis

Michael Keane
Department of Economics
University of Minnesota
and Federal Reserve Bank of Minneapolis

March 1997

We thank Garrett TeSelle and Susumu Imai for assistance in preparation of the data, and Dan Hauser 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.

Tables and Figures for this paper follow the text.

IRP publications (discussion papers, special reports, and the newsletter *Focus*) are now available electronically. The IRP Web Site can be accessed at the following address:
<http://www.ssc.wisc.edu/IRP/>

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. A Gibbs sampling-data augmentation algorithm facilitates use of the entire sample and provides numerical approximations to the exact posterior distribution of properties of earnings paths. This algorithm copes with the complex distribution of endogenous variables that are observed for short segments of an individual's work history, not including the initial period.

The study reaches several firm conclusions about life cycle earnings mobility. Incorporating non-Gaussian shocks makes it possible to account for transitions between low and higher earnings states, a heretofore unresolved problem. The non-Gaussian distribution substantially increases the lifetime return to postsecondary education, and substantially reduces differences in lifetime wages 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.

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 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 Moffitt (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 percent of the variance in log earnings, the random effect accounts for 61 percent 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 percent, while the actual sample probability is only 37 percent. 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 percent, 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 percent. The actual sample probabilities of the person being in poverty in 1969 given these past histories are 23.5 percent and 21.1 percent, respectively. This again suggests that the model overstates short-run persistence.

A number of possible reasons may explain why the normally distributed random effect in a 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 (1993), 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 1986–87 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.

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 years. Given the objective of distinguishing among alternative serial correlation specifications for the error term, tests based on more than 20 years of data should have much greater power than ones that use only 7 or 10 years 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 years from now, based on point estimates of the fraction of variance due to a random effect and the parameters of a complex autoregressive-moving average (ARMA) process, all estimated on only 20 years of data (not to mention 7 to 10 years 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 5,000 households that were interviewed in 1968. Of these, about 3,000 were sampled to be representative of the nation as a whole and about 2,000 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 who can be identified clearly as household heads. 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, we drop all observations for that person from that point onward. This last screen is convenient, but not essential, because methods like those in Appendix C could be used to treat the missing observations as latent variables assuming an independent censoring process. The resulting sample for analysis contains 4,766 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 5,267 such individuals in the PSID. The various missing data screens only eliminate 501 of these.

Table 1 reports personal characteristics within the earnings distribution of the analysis sample. We define earnings quintiles based on the full sample. In 1967 dollars these are \$3,817, \$5,786, \$7,798 and \$10,454 (to convert to 1995 dollars multiply by 4.44). 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. Another important issue is that the PSID does not distinguish between missing earnings data and zero earnings. Both are represented by zero. We assume that all zeros represent missing earnings, since annual earnings that are truly zero for a male household head should be unusual.

In our model of the stochastic process for earnings, described in Section 3, we treat the process as beginning at age 25. Thus, if we do not observe an individual’s earnings

until an age later than 25, we face an initial conditions problem. Of the individuals in the sample, only 1,728 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 a difficult 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.”

We treat marital status as endogenous, because in previous studies marital status appears to have 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 the life cycle it is important to forecast his marital status as well. This requires us to 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.

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 Ω_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 ($i = 1, \dots, n$);

¹The 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.

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

The model of individual earnings is

$$y_{i1} = \tilde{\mathbf{b}}' \mathbf{x}_{i1} + \mathbf{e}_{i1},$$

$$y_{it} = \mathbf{g} y_{i,t-1} + (1 - \mathbf{g}) \mathbf{b}' \mathbf{x}_{it} + (1 - \mathbf{g}) \mathbf{t}_i + (1 - \mathbf{g}) \mathbf{f} \mathbf{e}_{i1} + \mathbf{e}_{it} \quad (t = 2, \dots, T), \quad (1)$$

$$\mathbf{e}_{i2} = \tilde{\mathbf{r}} \mathbf{e}_{i1} + \mathbf{h}_{i2}, \quad (2)$$

$$\mathbf{e}_{it} = \mathbf{r} \mathbf{e}_{i,t-1} + \mathbf{h}_{it} \quad (t = 3, \dots, T),$$

$$\mathbf{t}_i \stackrel{iid}{\sim} \mathbf{N}(0, \mathbf{s}_t^2).$$

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 \mathbf{b}_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 $\mathbf{e}_{i1}, \mathbf{h}_{i2}, \mathbf{h}_{it}$ ($t \geq 3$) are mutually independent across both time and individuals. The shocks \mathbf{e}_{i1} are identically distributed, as are the shocks \mathbf{h}_{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 \mathbf{f} in (1)). The mean level of earnings in the dynamic equation (1) is also heterogeneous, by virtue of the shock \mathbf{t}_i . The variance of the disturbance vector $\mathbf{e}'_i = (\mathbf{e}_{i1}, \dots, \mathbf{e}_{iT})$ is a function of the six terms $\text{var}(\mathbf{e}_{i1}), \text{var}(\mathbf{h}_{it}), \mathbf{s}_t^2, \mathbf{r}, \tilde{\mathbf{r}}$, and \mathbf{f} : 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 \mathbf{t}_i and \mathbf{e}_{i1}) can range from zero to one and can change smoothly from year to year.

In the mixture model the distributions of \mathbf{e}_{i1} and \mathbf{h}_{it} are each mixtures of three normal distributions; e.g.,

$$\mathbf{h}_{it} \sim N(\mathbf{a}_{2(j)}, h_{2(j)}^{-1}) \text{ with probability } p_{2j},$$

where $\mathbf{a}_{2(1)} < \mathbf{a}_{2(2)} = 0 < \mathbf{a}_{2(3)}$; $0 < h_{2(j)} < \infty$ ($j = 1, 2, 3$); $p_{2j} \geq 0$ ($j = 1, 2, 3$); and $p_{21} + p_{22} + p_{23} = 1$. (Similarly, $\mathbf{e}_{i1} \sim N(\mathbf{a}_{1(j)}, h_{1(j)}^{-1})$ 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.² 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 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 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 E.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.

²For a discussion of these models and a generalization to multiprocess models, see West and Harrison (1989), Section 12.3.4.

3.2 Marital Status Model

We adopt a dynamicprobit specification for marital status. Denote

$d_{it} = 1$ if individual i is married in period t

and $d_{it} = 0$ if not ($t = 1, \dots, T_i; i = 1, \dots, n$);

$\mathbf{s}_{i1} = (p_1 \times 1)$ vector of period 1 explanatory variables for individual i
($i = 1, \dots, n$);

$\mathbf{s}_{it} = (p_2 \times 1)$ vector of period t explanatory variables for individual i
($t = 2, \dots, T_i; i \in \Omega_2$);

$m_{it}^* = \text{Probit (latent) that determines } d_{it}$ ($t = 1, \dots, T_i; i = 1, \dots, n$).

The model for marital status is

$$\begin{aligned} m_{i1}^* &= \tilde{\mathbf{q}}' \mathbf{s}_{i1} + \mathbf{x}_{i1}, \\ \mathbf{x}_{i1} &\stackrel{iid}{\sim} \text{N}\left[0, (1 - I^2)^{-1}\right] \\ m_{it}^* &= \mathbf{q}' \mathbf{s}_{it} + \mathbf{x}_{it} \quad (t = 2, \dots, T), \\ \mathbf{x}_{it} &= \mathbf{I} \mathbf{x}_{i,t-1} + \mathbf{y}_{it} \quad (t = 2, \dots, T), \\ \mathbf{y}_{it} &\stackrel{iid}{\sim} \text{N}(0, 1) \quad (t = 2, \dots, T), \\ d_{it} &= \begin{cases} 1 & \text{if } m_{it}^* \geq 0 \\ 0 & \text{if } m_{it}^* < 0 \end{cases}. \end{aligned}$$

The vector \mathbf{s}_{it} used in this study is described in Table 2; $p_1 = 3$ and $p_2 = 9$. The vector \mathbf{s}_{i1} contains an intercept, the individual's education, and a race indicator. The vector \mathbf{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 E.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. An accessible introduction to both topics for economists is Geweke (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.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 \mathbf{e}_{i1} (if $t = 1$) or \mathbf{h}_{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 \mathbf{q}_E denote the 45×1 vector of parameters in the earnings model, and \mathbf{q}_M the 13×1 vector of parameters in the marital status model.

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

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

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

$$p_M(m^*_{it} | \mathbf{Y}_{i,t-1}, \mathbf{M}^*_{i,t-1}, \mathbf{z}_i, \mathbf{q}_M) \quad p_M(d_{it} | m^*_{it}).$$

The corresponding prior distributions for each model provide, respectively, $p_E(\mathbf{q}_E)$ and $p_M(\mathbf{q}_M)$.

By the standard definition of conditional probability,

$$\begin{aligned} & p\left\{\mathbf{q}_M, \mathbf{q}_E, \left[\mathbf{t}_i, (y_{it}, d_{it})_{t=1}^{S_i-1}, (L_{it}, m_{it}^*)_{t=1}^{T_i}\right]_{i=1}^n \middle| \left[\mathbf{z}_i, (y_{it}, d_{it})_{t=S_i}^{T_i}\right]_{i=1}^n\right\} \\ & \propto p_E(\mathbf{q}_E) \prod_{i=1}^n \left\{ p_E(\mathbf{t}_i | \mathbf{q}_E) \prod_{t=1}^{T_i} \left[p_E(L_{it} | \mathbf{q}_E) p_E(y_{it} | \mathbf{Y}_{i,t-1}, \mathbf{D}_{it}, \mathbf{z}_i, \mathbf{t}_i, L_{it}, \mathbf{q}_E) \right] \right\} \\ & \quad \cdot p_M(\mathbf{q}_M) \prod_{i=1}^n \left[\prod_{t=1}^{T_i} p_M(m_{it}^* | \mathbf{Y}_{i,t-1}, \mathbf{D}_{i,t-1}, \mathbf{M}_{i,t-1}^*, \mathbf{z}_i, \mathbf{q}_E) p_M(d_{it} | m_{it}^*) \right] \end{aligned}$$

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 \mathbf{q}_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 \mathbf{t}_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 $(\mathbf{Y}_{i,T_i}, \mathbf{D}_{i,T_i})_{i=1}^n$. 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 \mathbf{q}_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 drawing 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 $(\mathbf{Y}_{i,T_i}, \mathbf{D}_{i,T_i})_{i=1}^n$.

1. In the third group of steps, first the unobserved earnings $(\mathbf{Y}_{i,S_i-1})_{i=1}^n$ are drawn. These are conditionally independent across individuals and jointly normally distributed. Then, the unobserved probits and marital statuses $(\mathbf{D}_{i,S_i-1}, \mathbf{M}_{i,S_i-1}^*)_{i=1}^n$ 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.

It is straightforward, though somewhat tedious, to verify that the likelihood function for the earnings and marital status models is a bounded function of the 58 parameters of the models. Since the prior distribution of the 58 parameters is proper, the posterior density kernel is finitely integrable and therefore the posterior distribution exists.

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 $E \left\{ g(\mathbf{q}_E, \mathbf{q}_M) \left[\mathbf{z}_i, (y_{it}, d_{it})_{t=S_i}^T \right]_{i=1}^n \right\}$ exists, then the corresponding sample average of $g(\mathbf{q}_E, \mathbf{q}_M)$ from the posterior simulator converges almost surely to this posterior moment.

It is always necessary to verify the existence of a posterior moment analytically, before approximating it in this way. All of the moments reported in this study are one of two kinds. In the most common case, $g(\mathbf{q}_E, \mathbf{q}_M)$ is an indicator function or corresponds to a probability, so it is bounded below by 0 and above by 1. In some other cases, the prior moment $E[g(\mathbf{q}_E, \mathbf{q}_M)]$ exists, and since the likelihood function is bounded, the corresponding posterior moment also exists.

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 r in the earnings model. Another is the difference in unconditional expected log real earnings at ages 35 and 25, given 16 years 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 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 2,000 iterations which were discarded. These computations were undertaken on a Sun Model 20 workstation, and required about 25 seconds per iteration for each model. For the mixture model based on the full sample, all results are based on 2,500 iterations of the Gibbs sampler following an initial 294 iterations which were discarded. These computations required about 332 seconds per iteration. For the normal model based on the full sample all results are based on 1,500 iterations of the Gibbs sampler following an initial 276 iterations which were discarded. These computations required about 325 seconds 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.

5. RESULTS

Table 3 and Figures 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 $-.195$, implying that first-period earnings are roughly 20 percent lower for blacks. All four sets of results indicate that those with missing values for father's education tend to have lower initial earnings, but there is little evidence of any other relation between parents' education and initial earnings.

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 percent increase in initial earnings, while the normal model based on the full sample indicates a 12 percent 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 to 9 percent 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 percent lower for blacks than whites, *ceteris paribus*. And the four models imply that married men have earnings that range from 4 to 10 percent greater than single men. The parents' education variables show no clear pattern across the models, and most are within two posterior standard deviations of zero.

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 .231, implying earnings growth of roughly 23 percent from age 25 to 35 for those with 12 years of education. For age 45 vs. 35 the growth is 8 percent, whereas for 55 vs. 45 it is -4 percent. 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 .469, implying that college graduates earn roughly 47 percent more than high school graduates at age 35, *ceteris paribus*.

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 years 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 percent to 20 percent of prior standard deviations. This reflects the deliberate weakness of the prior (as discussed fully in Appendix E) 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 E), 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.

Parameters of these distributions are tightly estimated. Posterior standard deviations are considerably smaller than their prior standard deviations in the case of the mean and standard deviation parameters. In the case of the probability parameters our prior distributions were (in retrospect) rather informative, but observe that the posterior means are up to several prior standard deviations from the prior mean, and (especially in the case of \mathbf{h}_t) posterior standard deviations for the probabilities are very small.

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 \mathbf{h}_t , posterior means for the full sample show the probability of a shock that cuts wages by 50 percent or more is 5 percent, while the probability of a shock that more than doubles wages is 2.7 percent; posterior standard deviations are negligible.

The implied p.d.f.'s are shown in Figures 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 normal distributions are of course symmetric. The mode is around $\log(1.18)$ for the first-period shock and around $\log(1.09)$ for the t 'th period shock. 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 $\log(.325)$ 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 \mathbf{g} on lagged earnings is $-.121$. 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, \mathbf{r} having a posterior mean of $.655$. 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 \mathbf{e}_{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 $.366$. Thus, a person with a one standard deviation above the average \mathbf{t}_i value of zero would have earnings about 37 percent above average, given his personal characteristics. Finally, the posterior mean for \mathbf{f} in the mixture model based on the full sample is $.240$. This implies that the first-period shock could be decomposed into independent permanent and transitory components, with the permanent component having about one-third 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 $(1 - \mathbf{g})^2 [\mathbf{s}_t^2 + \mathbf{f}^2 \text{var}(\mathbf{e}_{it})]$, this is accounted for by $\text{var}(\mathbf{e}_{it}) > \text{var}(\mathbf{h}_{it}) / (1 - \mathbf{r}^2)$. The fraction of variance accounted for by the transitory shock \mathbf{h}_{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 \mathbf{t}_i and \mathbf{e}_{it} ; consistent with the fraction of variance due to the transitory component, these values are about one-third.

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 $-.232$, 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 $-.092$, 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 $.177$, 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 these two 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 .152 for whites and .342 for blacks. And if we look at the set of all two year sequences, the frequency of “- -” is .089 for whites and .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 .181 and .538, respectively. The mixture model predicts frequencies of .181 and .532, while the normal model predicts .278 and .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.

It is also interesting to examine how the models fit the cross-sectional log wage distribution at various ages. Figure 3.1 reports kernel density estimates for log wages 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 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 wage 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 wage distribution.

Figure 3.2 reports the same kernel density estimates for the mixture model. Clearly, this model captures the shape of the wage density much better than the normal model. Figures 4.1 and 4.2 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.

We now turn to evaluation of the wage distribution in models based on the full sample. Kernel density estimates for log wages at ages 25, 30, 45, and 60 in the full sample, using the normal model and the mixture model, are reported in Figures 5 through 8. The interesting pattern in these figures is that, based on the full sample, the mixture model fits the log wage 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 Figure 7-2, 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 apparent (see Figure 8), but discrepancies between its predictions and the observed wage density are still apparent.

We conjecture that these problems arise because at age 45 a greater fraction of the variation in log wages 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 wage 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 wages with age, and (2) to allow the variance of the shocks to vary with age (so that it may rise in the middle of the life cycle).

Finally, we explore the fit of the models to conditional log wage distributions. Figure 9-1 reports kernel density estimates for log wages 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. Figure 9-2 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.

Figure 10 reports similar age 35 wage density estimates for the young men data and models, but now conditioning on not being in the bottom quintile at age 34. Comparing Figures 10-1 and 10-2, it is apparent that the mixture model fits the shape of the conditional wage density quite closely, while the normal model does not. Figures 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 wage 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 25–40), and again at older ages (near 60). At age 45 it fails to capture important features of the wage distribution. We conjecture that this is due either to failure to allow for a sufficiently flexible pattern of changing effects of covariates on wages with age, or failure to allow the variance of the shocks to vary 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 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.

Table 5 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 earnings mobility. For example, in Table 5.1, 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 years of education the simulated probabilities of these sequences are .122 and .047 respectively. But Table 5.2, which contains results based on the full sample, shows that the same simulated probabilities in the normal mixture model are .102 and .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 predicted wage 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 percent 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 contrast, the normal model estimated from the full sample implies a higher fraction of years in the bottom quintile (51.9 percent), a substantially larger expected number of spells (4.90), and a substantially shorter mean spell length (4.28 years). 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.

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 percent when education is less than 12 years but only 15.1 percent when education is 16 years or more). The normal model predicts a less sharp decline (e.g., 68.1 percent for blacks with education less than 12 years and 28.9 percent when education is 16 years 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 percent of the remaining years in the bottom quintile. If he is not in the bottom quintile at age 30, the fraction is only 32.0 percent. The corresponding figures for the normal model are 65.6 and 36.4 percent. 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 than 12 years. 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.

Table 8 reports simulations of year-to-year transition probabilities at age 35. Table 8.1 contains results for the models based on the young men data, and Table 8.2 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 mixture model predicts he has an 81.2 percent 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 percent. The corresponding figures for the normal model are 73.4 and 17.5 percent. Thus, the mixture model predicts more persistence in low-earnings status.

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 percent for whites with education less than 12 years, while the normal model predicts 36.7 and 55.7 percent 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.

6.3 Present Value of Lifetime Earnings

In this section we use the various models to simulate lifetime wage streams and construct present values of lifetime earnings. Table 9.1 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 .95. Table 9.2 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 to 10 percent) 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 wage growth beyond that age are based entirely on prior growth combined with extrapolations beyond age 46. These lead it to predict modest wage growth from age 45 to 65. On the other hand, the models based on the full sample predict wage declines at older ages. For this reason, we view models based on the full sample as providing more reliable simulations of lifetime wage paths.

Turning to the full sample results in Table 9.2, 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

wealth 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 percent) less than that for whites. Much but not nearly all of this difference is accounted for by the education differences between blacks and whites. For example, among those with education in the 12–15 year range, the mean present value of lifetime earnings for blacks is predicted to be \$26,600 (or 18 percent) less than that for whites.

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

The mixture model predicts greater differences in wealth across education classes than does the normal model. For example, comparing whites in the 16+ vs. 12–15 years of education categories, it predicts a \$82,200 (or 56 percent) greater present value of lifetime earnings for members of the college-educated group. The normal model predicts only \$63,200 (or 42 percent) greater earnings for members of the college-educated group.

The mixture model predicts that the increase in mean present value of lifetime wealth in moving from the 12-15 to 16+ years 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 Figures 13 and 14, which report kernel density estimates for the earnings distributions within race and education classes. Figure 13-1 reports results for whites based on the normal model, while Figure 13-2 reports results for whites based on the mixture model, both using the full sample. Comparison of Figures 13-1 and 13-2 shows clearly that the mixture model predicts a greater shift right in the mode with increasing education than does the normal model. Figure 14 shows that this difference is even more apparent for blacks.

Table 10 reports the distributions of lifetime earnings over ages 31–65, conditional on earnings quintile status at age 30, using the full sample results; Tables 10.1 is based on the normal model and Table 10.2 is based on the mixture model. The mixture model implies that mean present value of lifetime earnings over ages 31–65 is \$159,000 if a male is not in the bottom quintile at age 30, and \$70,700 if he is in the bottom quintile. This is \$88,300, or 56 percent less. Interestingly, conditioning on race and education class does little to eliminate this difference. For example, for whites in the 12–15 years of education range, the difference is \$161,900 vs. \$92,500. This is still \$69,400, or 43 percent 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 10.1, the normal model implies a mean 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 percent 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 expected present value of lifetime earnings that is \$19,214 less than whites, *ceteris paribus*. This is 14.6 percent 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 percent 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 percent of the mean). The normal model predicts a college premium of only \$41,112 (31.6 percent of the mean).

Largely due to the greater wage differences across education classes predicted by the mixture model, the covariates explain 33.7 percent of the variation in present value of earnings in that model, while in the univariate normal model they explain only 26.3 percent. The mixture model attributes 38.7 percent of the variation in lifetime earnings to unobserved individual effects (t_i and e_{it}) while the univariate normal model attributes 44.6 percent. The fact that a larger percentage of variance is due to unobserved heterogeneity than due to race, education, and family background explains why, in Table

10, 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 percent to 70 percent of the variation in the *log* of earnings is accounted for by transitory components whose serial correlation is relatively weak, about .7 from year to year. But over a lifetime transitory components (by definition) average out. The posterior distributions show that about 60 percent 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 \$228,536 greater for those with a college degree than for those with only a high school degree (in 1995 dollars). The normal model implies a college premium of only \$180,893. And while the normal model predicts that present value of lifetime earnings is \$135,186 less for blacks in 1995 dollars (*ceteris paribus*), the mixture model predicts a substantially smaller black/white differential of \$84,542. 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.

APPENDICES

A. Posterior and Conditional Posterior Distributions 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 Ω_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 \mathbf{e}_i has distribution $\mathbf{e}_i \sim N(\mathbf{a}_i, h_i^{-1})$, and the distribution of \mathbf{a}_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 \mathbf{a}_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 ($t = 2, \dots, T$).

Conditional on \mathbf{e}_{it} ($t = 1, \dots, T; i = 1, \dots, n$) for each i the model for individual earnings is,

$$y_{i1} = \tilde{\mathbf{b}}' \mathbf{x}_{i1} + \mathbf{e}_{i1}; \quad (\text{A1})$$

$$y_{it} = \mathbf{g}' y_{i,t-1} + (1 - \mathbf{g}) \mathbf{b}' \mathbf{x}_{it} + (1 - \mathbf{g}) \mathbf{t}_i + (1 - \mathbf{g}) \mathbf{f} \mathbf{e}_{i1} + \mathbf{e}_{it} \quad (t = 2, \dots, T); \quad (\text{A2})$$

$$\mathbf{e}_{i2} = \tilde{\mathbf{r}} \mathbf{e}_{i1} + \mathbf{h}_{i2}; \quad (\text{A3})$$

$$\mathbf{e}_{it} = \mathbf{r} \mathbf{e}_{i,t-1} + \mathbf{h}_{it} \quad (t = 3, \dots, T); \quad (\text{A4})$$

$$\mathbf{e}_{i1} = \mathbf{a}_{i1} + h_{i1}^{-1/2} \mathbf{z}_{i1} \quad (\text{A5})$$

$$\mathbf{h}_{it} = \mathbf{a}_{it} + h_{it}^{-1/2} \mathbf{z}_{it} \quad (t = 2, \dots, T); \quad (\text{A6})$$

$$\mathbf{t}_i = \mathbf{s}_i \mathbf{z}_{i0}. \quad (\text{A7})$$

$$\mathbf{z}_{it} \quad (t = 0, \dots, T; i = 1, \dots, n) \stackrel{iid}{\sim} N(0, 1). \quad (\text{A8})$$

The parameters $\{\mathbf{a}_{it}, h_{it}\}$ arise from a discrete mixture of normals model. In this model, the disturbance \mathbf{e}_{i1} is a discrete mixture of three normals, with one set of parameters; the distribution of \mathbf{h}_{it} is a discrete mixture of three normals with another set of parameters.

The prior distribution is composed of the following independent components:

$$\mathbf{d} = (\tilde{\mathbf{b}}', \mathbf{b}')' \sim N(\mathbf{d}, \mathbf{H}_d^{-1}), \quad \mathbf{H}_d \text{ p.d.}; \quad (\text{A9})$$

$$\mathbf{p} = (\tilde{\mathbf{r}}, \mathbf{r})' \sim \mathbf{N}(\underline{\mathbf{p}}, \underline{\mathbf{H}}_p) \quad \underline{\mathbf{H}}_p \text{ p.d.}, \quad -1 < \mathbf{r} < 1; \quad (\text{A10})$$

$$\mathbf{f} \sim \mathbf{N}(\underline{\mathbf{f}}, \underline{h}_f^{-1}); \quad (\text{A11})$$

$$\mathbf{g} \sim \mathbf{N}(\underline{\mathbf{g}}, \underline{h}_g^{-1}) \quad -1 < \mathbf{g} < 1; \quad (\text{A12})$$

$$\underline{\mathbf{n}}_t \underline{s}_t^2 / \mathbf{s}_t^2 \sim \mathbf{c}^2(\underline{\mathbf{n}}_t) \quad \underline{s}_t^2 > 0, \underline{\mathbf{n}}_t > 0; \quad (\text{A13})$$

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

$$\begin{aligned} & \mathbf{p}\left(\tilde{\mathbf{b}}, \mathbf{b}, \tilde{\mathbf{r}}, \mathbf{r}, \mathbf{f}, \mathbf{g}, \mathbf{s}_t^2, (\mathbf{t}_i)_{i=1}^n \left| \left\{ \left\{ y_{it}, \mathbf{x}_{it}, \mathbf{a}_{it}, h_{it} \right\}_{t=1}^{T_i} \right\}_{i=1}^n, \underline{\mathbf{d}}, \underline{\mathbf{H}}_d, \underline{\mathbf{p}}, \underline{\mathbf{H}}_p, \underline{\mathbf{f}}, \underline{h}_f, \underline{\mathbf{g}}, \underline{h}_g, \underline{s}_t^2, \underline{\mathbf{n}}_t \right)\right) \\ & \propto \prod_{i=1}^n \left[\mathbf{p}\left(\left\{ y_{it} \right\}_{t=1}^{T_i} \left| \left\{ \mathbf{x}_{it}, \mathbf{a}_{it}, h_{it} \right\}_{t=1}^{T_i}, \tilde{\mathbf{b}}, \mathbf{b}, \tilde{\mathbf{r}}, \mathbf{r}, \mathbf{f}, \mathbf{g}, \mathbf{s}_t^2, (\mathbf{t}_i)_{i=1}^n \right) \mathbf{p}(\mathbf{t}_i | \mathbf{s}_t^2) \right. \\ & \quad \cdot \mathbf{p}(\underline{\mathbf{d}}, \underline{\mathbf{p}}, \underline{\mathbf{f}}, \underline{\mathbf{g}}, \underline{\mathbf{s}}_t \left| \underline{\mathbf{d}}, \underline{\mathbf{H}}_d, \underline{\mathbf{p}}, \underline{\mathbf{H}}_p, \underline{\mathbf{f}}, \underline{h}_f, \underline{\mathbf{g}}, \underline{h}_g, \underline{\mathbf{n}}_t, \underline{s}_t^2 \right) \\ & \quad \left. \prod_{i=1}^n h_{i1}^{1/2} \exp\left[-\frac{1}{2} \sum_{i=1}^n h_{i1} (y_{i1} - \mathbf{a}_{i1} - \tilde{\mathbf{b}}' \mathbf{x}_{i1})^2\right] \right] \quad (\text{A14}) \end{aligned}$$

$$\begin{aligned} & \cdot \prod_{i \in \Omega_2} h_{i2}^{1/2} \exp\left\{ \frac{1}{2} \sum_{i \in \Omega_2} h_{i2} [y_{i2} - \mathbf{a}_{i2} - \mathbf{g}_1 y_{i1} - \mathbf{g}_1 \mathbf{b}' \mathbf{x}_{i2} - \mathbf{g}_1 \mathbf{t}_i \right. \\ & \quad \left. - (\tilde{\mathbf{r}} + \mathbf{g}_1 \mathbf{f})(y_{i1} - \tilde{\mathbf{b}}' \mathbf{x}_{i1}) \right\} \quad (\text{A15}) \end{aligned}$$

$$\begin{aligned} & \cdot \prod_{i \in \Omega_3} \prod_{t=3}^{T_i} h_{it}^{1/2} \exp\left\{ \frac{1}{2} \sum_{i \in \Omega_3} \sum_{t=3}^{T_i} h_{it} [y_{it} - \mathbf{a}_{it} - \mathbf{g}_1 y_{i,t-1} - \mathbf{g}_1 \mathbf{b}' \mathbf{x}_{it} - \mathbf{g}_1 \mathbf{t}_i \right. \\ & \quad \left. - \mathbf{g}_1 \mathbf{f}(y_{i1} - \tilde{\mathbf{b}}' \mathbf{x}_{i1}) - \mathbf{r}(y_{i,t-1} - \mathbf{g}_1 y_{i,t-2} - \mathbf{g}_1 \mathbf{b}' \mathbf{x}_{i,t-1} - \mathbf{g}_1 \mathbf{t}_i - \mathbf{g}_1 \mathbf{f}(y_{i1} - \tilde{\mathbf{b}}' \mathbf{x}_{i1})) \right\} \quad (\text{A16}) \end{aligned}$$

$$\cdot \mathbf{s}_t^{-n} \exp\left(-\frac{1}{2} \mathbf{s}_t^{-2} \sum_{i=1}^n \mathbf{t}_i^2\right) \quad (\text{A17})$$

$$\cdot \exp\left[-\frac{1}{2} (\underline{\mathbf{d}} - \underline{\mathbf{d}})' \underline{\mathbf{H}}_d (\underline{\mathbf{d}} - \underline{\mathbf{d}})\right] \quad (\text{A18})$$

$$\cdot \exp\left[-\frac{1}{2} (\underline{\mathbf{p}} - \underline{\mathbf{p}})' \underline{\mathbf{H}}_p (\underline{\mathbf{p}} - \underline{\mathbf{p}})\right] c_{(-1,1)}(\mathbf{r}) \quad (\text{A19})$$

$$\cdot \exp\left[-\frac{1}{2} \underline{h}_f (\underline{\mathbf{f}} - \underline{\mathbf{f}})^2\right] \quad (\text{A20})$$

$$\cdot \exp\left[-\frac{1}{2} \underline{h}_g (\underline{\mathbf{g}} - \underline{\mathbf{g}})^2\right] c_{(-1,1)}(\mathbf{g}) \quad (\text{A21})$$

$$\cdot \mathbf{s}_t^{-(\underline{\mathbf{n}}_t + 2)} \exp\left(-\frac{1}{2} \underline{\mathbf{n}}_t \underline{s}_t^2 \mathbf{s}_t^{-2}\right). \quad (\text{A22})$$

The component (A14) of the posterior distribution arises from expressions (A1), (A5) and (A8); (A15) from (A1), (A2), (A3), (A6) and (A8); (A16) from (A2), (A4), (A6) and

(A8); (A17) from (A7) and (A8); and (A18) through (A22) from (A9) through (A13) respectively.

Conditional distribution of $\mathbf{d} = (\tilde{\mathbf{b}}', \mathbf{b}')'$. The kernel from (A14), (A15), (A16) and (A18) is

$$\exp\left[-\frac{1}{2}\sum_{i=1}^n\sum_{t=1}^{T_i}h_{it}(y_{it}^d - \mathbf{d}'\mathbf{z}_{it}^d)^2\right] \quad (\text{A23})$$

$$\cdot \exp\left[-\frac{1}{2}(\mathbf{d} - \underline{\mathbf{d}})' \underline{\mathbf{H}}_d (\mathbf{d} - \underline{\mathbf{d}})\right]. \quad (\text{A24})$$

Expression (A23) as a function of \mathbf{d} is proportional to the product of (A14), (A15) and (A16), with

$$\mathbf{z}_{i1}^d = \begin{pmatrix} \mathbf{x}'_{i1}, & \mathbf{0}' \\ & 1 \times k_2 \end{pmatrix}, \quad y_{i1}^d = y_{i1} - \mathbf{a}_{i1} \quad (i = 1, \dots, n)$$

from (A14);

$$\mathbf{z}_{i2}^d = [-\mathbf{x}'_{i1}(\tilde{\mathbf{r}} + \mathbf{g}_1\mathbf{f}), \mathbf{x}'_{i2}\mathbf{g}_1] \quad y_{i2}^d = y_{i2} - \mathbf{a}_{i2} - \mathbf{g}_{y_{i1}} - \mathbf{g}_1\mathbf{t}_i - (\tilde{\mathbf{r}} + \mathbf{g}_1\mathbf{f})y_{i1} \quad (i \in \Omega_2)$$

from (A15);

$$\mathbf{z}_{it}^d = [\mathbf{x}'_{it}\mathbf{g}_1\mathbf{f}(\mathbf{r}-1), (\mathbf{x}'_{it} - \mathbf{x}'_{i,t-1}\mathbf{r})\mathbf{g}_1]$$

$$y_{it}^d = y_{it} - \mathbf{a}_{it} - (\mathbf{g} + \mathbf{r})y_{i,t-1} + \mathbf{r}\mathbf{g}_{y_{i,t-2}} + \mathbf{f}\mathbf{g}_1(\mathbf{r}-1)y_{i1} + \mathbf{g}_1(\mathbf{r}-1)\mathbf{t}_i \quad (t = 3, \dots, T_i; i \in \Omega_3)$$

from (A16). Expression (A23) is in standard form for the likelihood function of a regression model with heteroscedastic disturbances, and (A24) is standard form for a multivariate normal prior distribution. Applying *Result 1* and *Result 2* of Appendix G,

$$\mathbf{d} \sim \mathbf{N}(\bar{\mathbf{d}}, \bar{\mathbf{H}}_d^{-1}),$$

with

$$\bar{\mathbf{H}}_d = \sum_{i=1}^n \sum_{t=1}^{T_i} h_{it} \mathbf{z}_{it}^d \mathbf{z}_{it}^d + \underline{\mathbf{H}}_d, \quad \bar{\mathbf{d}} = \bar{\mathbf{H}}_d^{-1} \left[\sum_{i=1}^n \sum_{t=1}^{T_i} h_{it} \mathbf{z}_{it}^d y_{it}^d + \underline{\mathbf{H}}_d \right].$$

$$\text{For the balance, let } \mathbf{e}_{i1} = y_{i1} - \tilde{\mathbf{b}}'\mathbf{x}_{i1}.$$

Conditional distribution of $\mathbf{p} = (\tilde{\mathbf{r}}, \mathbf{r})'$. The kernel from (A15), (A16) and (A19) is

$$\exp\left[-\frac{1}{2}\sum_{i \in \Omega_2} \sum_{t=1}^{T_i-1} h_{it}^p (y_{it}^p - \mathbf{p}'\mathbf{z}_{it}^p)^2\right] \quad (\text{A25})$$

$$\cdot \exp\left[-\frac{1}{2}(\mathbf{p} - \underline{\mathbf{p}})' \underline{\mathbf{H}}_p (\mathbf{p} - \underline{\mathbf{p}})\right] c_{(-1,1)}(\mathbf{r}), \quad (\text{A26})$$

with

$$\mathbf{z}_{i1}^p = (\mathbf{e}_{i1}, 0), \quad y_{i1}^p = y_{i2} - \mathbf{a}_{i2} - \mathbf{g}_{y_{i1}} - \mathbf{g}_1\mathbf{b}'\mathbf{x}_{i2} - \mathbf{g}_1\mathbf{t}_i - \mathbf{g}_1\mathbf{f}\mathbf{e}_{i1}, \quad h_{i1}^p = h_{i2} \quad (i \in \Omega_2)$$

from (A15) and

$$\mathbf{z}_{it}^p = \left(0, y_{it} - \mathbf{g}'\mathbf{y}_{i,t-1} - \mathbf{g}_1\mathbf{b}'\mathbf{x}_{it} - \mathbf{g}_1\mathbf{t}_i - \mathbf{g}_1\mathbf{f}\mathbf{e}_{i1} \right)$$

$$y_{it}^p = y_{i,t+1} - \mathbf{a}_{i,t+1} - \mathbf{g}'\mathbf{y}_{it} - \mathbf{g}_1\mathbf{b}'\mathbf{x}_{i,t+1} - \mathbf{g}_1\mathbf{t}_i - \mathbf{g}_1\mathbf{f}\mathbf{e}_{i1}, \quad h_{it}^p = h_{i,t+1} \quad (t = 2, \dots, T_i - 1; i \in \Omega_3)$$

from (A22). Expression (A25) is in standard form for the likelihood function of a regression model with heteroscedastic disturbances, and (A26) is the standard form for a multivariate normal prior distribution, except for the truncation of \mathbf{r} . Applying *Result 1* and *Result 2* of Appendix G,

$$\mathbf{p} \sim \mathbf{N}(\bar{\mathbf{p}}, \bar{\mathbf{H}}_p^{-1})$$

subject to the constraint $-1 < \mathbf{r} < 1$, with

$$\bar{\mathbf{H}}_p = \sum_{i \in \Omega_2} \sum_{t=1}^{T_i-1} h_{it}^p \mathbf{z}_{it}^p \mathbf{z}_{it}^p + \underline{\mathbf{H}}_p, \quad \bar{\mathbf{p}} = \bar{\mathbf{H}}_p^{-1} \left[\sum_{i \in \Omega_2} \sum_{t=1}^{T_i-1} h_{it}^p \mathbf{z}_{it}^p y_{it}^p + \underline{\mathbf{H}}_p \right]$$

Conditional distribution of \mathbf{f} . The kernel from (A15), (A16), and (A20) is

$$\exp \left[-\frac{1}{2} \sum_{i \in \Omega_2} \sum_{t=1}^{T_i-1} h_{it}^f (y_{it}^f - \mathbf{f}'\mathbf{z}_{it}^f)^2 \right] \exp \left[\underline{h}_f (\mathbf{f} - \underline{\mathbf{f}})' \right] \quad (\text{A27})$$

with

$$z_{i1}^f = \mathbf{g}_1\mathbf{e}_{i1}, \quad y_{i1}^f = y_{i2} - \mathbf{a}_{i2} - \mathbf{g}'\mathbf{y}_{i1} - \mathbf{g}_1\mathbf{b}'\mathbf{x}_{i2} - \mathbf{g}_1\mathbf{t}_i - \tilde{\mathbf{r}}\mathbf{e}_{i1}, \quad h_{i1}^f = h_{i2} \quad (i \in \Omega_2)$$

from (A15) and

$$\begin{aligned} z_{it}^f &= \mathbf{g}_1(1 - \mathbf{r})\mathbf{e}_{i1}, \\ y_{it}^f &= y_{i,t+1} - \mathbf{a}_{i,t+1} - \mathbf{g}'\mathbf{y}_{it} - \mathbf{g}_1\mathbf{b}'\mathbf{x}_{i,t+1} - \mathbf{g}_1\mathbf{t}_i - \mathbf{r}(y_{it} - \mathbf{g}'\mathbf{y}_{i,t-1} - \mathbf{g}_1\mathbf{b}'\mathbf{x}_{it} - \mathbf{g}_1\mathbf{t}_i) \\ h_{it}^f &= h_{i,t+1} \quad (t = 2, \dots, T_i - 1; i \in \Omega_3) \end{aligned}$$

from (A16). Expression (A27) is the standard form for a simple regression model with heteroscedastic disturbances and a normal prior for the coefficient. From *Result 1* and *Result 2* in Appendix G the conditional distribution of \mathbf{f} is

$$\mathbf{f} \sim \mathbf{N}(\bar{\mathbf{f}}, \bar{h}_f^{-1}),$$

with

$$\bar{h}_f = \sum_{i \in \Omega_2} \sum_{t=1}^{T_i-1} h_{it}^f (z_{it}^f)^2 + \underline{h}_f, \quad \bar{\mathbf{f}} = \bar{h}_f^{-1} \left(\sum_{i \in \Omega_2} \sum_{t=1}^{T_i-1} h_{it}^f z_{it}^f y_{it}^f + \underline{h}_f \underline{\mathbf{f}} \right)$$

Conditional distribution of \mathbf{g} . The kernel from (A15), (A16), and (A21) is

$$\exp \left[-\frac{1}{2} \sum_{i \in \Omega_2} \sum_{t=1}^{T_i-1} h_{it}^g (y_{it}^g - \mathbf{g}'\mathbf{z}_{it}^g)^2 \right] \exp \left[\underline{h}_g (\mathbf{g} - \underline{\mathbf{g}})' \right] c_{(-1,1)}(\mathbf{g}) \quad (\text{A28})$$

with

$$z_{i1}^g = y_{i1} - \mathbf{b}'\mathbf{x}_{i2} - \mathbf{t}_i - \mathbf{f}\mathbf{e}_{i1}, \quad y_{i1}^g = y_{i2} - \mathbf{a}_{i2} - \mathbf{b}'\mathbf{x}_{i2} - \mathbf{t}_i - (\tilde{\mathbf{r}} + \mathbf{f})\mathbf{e}_{i1}, \quad h_{i1}^g = h_{i2}$$

from (A15) and

$$\begin{aligned}
z_{it}^g &= y_{it} - \mathbf{b}'\mathbf{x}_{i,t+1} - \mathbf{t}_i - \mathbf{f}\mathbf{e}_{i1} - \mathbf{r}(y_{i,t-1} - \mathbf{b}'\mathbf{x}_{it} - \mathbf{t}_i - \mathbf{f}\mathbf{e}_{i1}), \\
y_{it}^g &= y_{i,t+1} - \mathbf{a}_{i,t+1} - \mathbf{b}'\mathbf{x}_{i,t+1} - \mathbf{t}_i - \mathbf{f}\mathbf{e}_{i1} - \mathbf{r}(y_{it} - \mathbf{b}'\mathbf{x}_{it} - \mathbf{t}_i - \mathbf{f}\mathbf{e}_{i1}), \\
h_{it}^g &= h_{i,t+1} \quad (t = 2, \dots, T_i - 1; i \in \Omega_3)
\end{aligned}$$

from (A16). Expression (A28) is the standard form for a simple regression model with heteroscedastic disturbances and a normal prior for the coefficient. From *Result 1* and *Result 2* in Appendix G the conditional distribution of \mathbf{g} is

$$\mathbf{g} \sim \mathbf{N}(\bar{\mathbf{g}}, \bar{h}_g^{-1}),$$

subject to the constraint $-1 < \mathbf{g} < 1$, with

$$\bar{h}_g = \sum_{i \in \Omega_2} \sum_{t=1}^{T_i-1} h_{it}^g (z_{it}^g)^2 + \underline{h}_g, \quad \bar{\mathbf{g}} = \bar{h}_g^{-1} \left(\sum_{i \in \Omega_2} \sum_{t=1}^{T_i-1} h_{it}^g z_{it}^g y_{it}^g + \underline{h}_g \underline{\mathbf{g}} \right)$$

Conditional distribution of \mathbf{s}_t^2 . The kernel from (A23) and (A32) is

$$\mathbf{s}_t^{-(n+\mathbf{u}_t+2)} \exp \left[-\frac{1}{2} \left(\sum_{i=1}^n \mathbf{t}_i^2 + \mathbf{u}_t \mathbf{s}_t^2 \right) \mathbf{s}_t^{-2} \right]$$

Hence the conditional distribution of \mathbf{s}_t^2 is given by

$$\left(\sum_{i=1}^n \mathbf{t}_i^2 + \mathbf{u}_t \mathbf{s}_t^2 \right) / \mathbf{s}_t^2 \sim c^2(n + \mathbf{u}_t)$$

Conditional distribution of \mathbf{t}_i ($i = 1, \dots, n$). The kernel from (A15), (A16) and (A17)

is

$$\exp \left[-\frac{1}{2} \sum_{t=1}^{T_i} h_{it}^t (y_{it}^t - z_{it}^t \mathbf{t}_i)^2 \right]$$

with

$$z_{i1}^t = \mathbf{g}_1, \quad y_{i1}^t = y_{i2} - \mathbf{a}_{i2} - \mathbf{g}_{y_{i1}} - \mathbf{g}_1 \mathbf{b}'\mathbf{x}_{i2} - (\tilde{\mathbf{r}} + \mathbf{g}_1 \mathbf{f})\mathbf{e}_{i1}, \quad h_{i1}^t = h_{i2} \quad (i \in \Omega_2)$$

from (A15);

$$\begin{aligned}
z_{it}^t &= (1 - \mathbf{r})\mathbf{g}_1, \quad y_{it}^t = y_{i,t+1} - \mathbf{a}_{i,t+1} - \mathbf{g}_{y_{it}} - \mathbf{g}_1 \mathbf{b}'\mathbf{x}_{i,t+1} - \mathbf{g}_1 \mathbf{f}\mathbf{e}_{i1} - \mathbf{r}(y_{it} - \mathbf{g}_{y_{i,t-1}} - \mathbf{g}_1 \mathbf{b}'\mathbf{x}_{it} - \mathbf{g}_1 \mathbf{f}\mathbf{e}_{i1}) \\
h_{it}^t &= h_{i,t+1} \quad (t = 2, \dots, T_i - 1; i \in \Omega_3)
\end{aligned}$$

from (A16); and

$$z_{i,T_i}^t = 1, \quad y_{i,T_i}^t = 0, \quad h_{i,T_i}^t = \mathbf{s}_t^{-2} \quad (i \in \Omega_2)$$

from (A17). Applying *Result 1* of Appendix G,

$$\mathbf{t}_i \sim \mathbf{N} \left\{ \sum_{t=1}^{T_i} h_{it}^t z_{it}^t y_{it}^t / \sum_{t=1}^{T_i} h_{it}^t (z_{it}^t)^2, \left[\sum_{t=1}^{T_i} h_{it}^t (z_{it}^t)^2 \right]^{-1} \right\}.$$

Parameters of the disturbance normal mixture distributions. The shocks \mathbf{e}_{i1} ($i = 1, \dots, n$) are independent and identically distributed, with a discrete mixture of

normals distribution. The model and methods of Appendix F are applied to

$$\mathbf{e}_{i1} = y_{i1} - \tilde{\mathbf{b}}' \mathbf{x}_{i1} \quad (i = 1, \dots, n).$$

The shocks \mathbf{h}_t ($t = 2, \dots, T_i, i \in \Omega_2$) are independent and identically distributed, with a discrete mixture of normals distribution. The model and methods of Appendix F are applied to

$$\mathbf{h}_{i2} = y_{i2} - \mathbf{g}' y_{i1} - \mathbf{g}_1 \mathbf{b}' \mathbf{x}_{i2} - \mathbf{g}_1 (\mathbf{t}_i + \mathbf{f} \mathbf{e}_{i1}) - \tilde{\mathbf{r}} \mathbf{e}_{i1} \quad (i \in \Omega_2),$$

$$\mathbf{h}_t = y_{it} - \mathbf{g}' y_{i,t-1} - \mathbf{g}_1 \mathbf{b}' \mathbf{x}_{it} - \mathbf{g}_1 (\mathbf{t}_i + \mathbf{f} \mathbf{e}_{i1})$$

$$- \mathbf{r} [y_{i,t-1} - \mathbf{g}' y_{i,t-2} - \mathbf{g}_1 \mathbf{b}' \mathbf{x}_{i,t-1} - \mathbf{g}_1 (\mathbf{t}_i + \mathbf{f} \mathbf{e}_{i1})] \quad (t = 3, \dots, T_i; i \in \Omega_3).$$

B. Posterior and Conditional Posterior Distributions 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 Ω_j . The total number of observations is $N = \sum_{i=1}^n T_i = \sum_{t=1}^T N_t$. For $(t = 1, \dots, T_i; t = 1, \dots, n)$ further denote

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

The model for marital status is

$$m_{i1}^* = \tilde{\mathbf{q}}' \mathbf{s}_{i1} + \mathbf{x}_{i1}, \quad (\text{B1})$$

$$\mathbf{x}_{i1} \stackrel{iid}{\sim} N\left[\mathbf{0}, (1 - I^2)^{-1}\right], \quad (\text{B2})$$

$$m_{it}^* = \mathbf{q}' \mathbf{s}_{it} + \mathbf{x}_{it} \quad (t = 2, \dots, T), \quad (\text{B3})$$

$$\mathbf{x}_{it} = I \mathbf{x}_{i,t-1} + \mathbf{y}_{it} \quad (t = 2, \dots, T), \quad (\text{B4})$$

$$\mathbf{y}_{it} \stackrel{iid}{\sim} N(0, 1) \quad (t = 2, \dots, T), \quad (\text{B5})$$

$$d_{it} = \begin{cases} 1 & \text{if } m_{it}^* \geq 0 \\ 0 & \text{if } m_{it}^* < 0 \end{cases}. \quad (\text{B6})$$

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

$$\mathbf{m} = (\tilde{\mathbf{q}}, \mathbf{q}') \sim N(\underline{\mathbf{m}}, \underline{\mathbf{H}}_m^{-1}) \quad \underline{\mathbf{H}}_m \text{ p.d.}; \quad (\text{B7})$$

$$I \sim N(\underline{I}, \underline{h}_I^{-1}), \quad \underline{h}_I > 0, \quad -1 < I < 1. \quad (\text{B8})$$

The joint posterior density of the parameters and latent variables is

$$p\left(\tilde{\mathbf{q}}, \mathbf{q}, I, \left\{m_{it}^*\right\}_{t=1}^T \mid \left\{d_{it}, \mathbf{s}_{it}\right\}_{t=1}^T, \underline{\mathbf{m}}, \underline{\mathbf{H}}_m, \underline{I}, \underline{h}_I\right) \\ \propto (1 - I^2)^{n/2} \exp\left[-\frac{1}{2}(1 - I^2) \sum_{i=1}^n (m_{i1}^* - \tilde{\mathbf{q}}' \mathbf{s}_{i1})^2\right] \quad (\text{B9})$$

$$\cdot \exp\left\{\frac{1}{2} \sum_{i \in \Omega_2} \left[m_{i2}^* - \mathbf{q}' \mathbf{s}_{i2} - I(m_{i1}^* - \tilde{\mathbf{q}}' \mathbf{s}_{i1})\right]^2\right\} \quad (\text{B10})$$

$$\cdot \exp\left\{\frac{1}{2} \sum_{i \in \Omega_3} \sum_{t=3}^{T_i} \left[m_{it}^* - \mathbf{q}' \mathbf{s}_{it} - I(m_{i,t-1}^* - \mathbf{q}' \mathbf{s}_{i,t-1})\right]^2\right\} \quad (\text{B11})$$

$$\cdot \prod_{i=1}^n \prod_{t=1}^{T_i} \mathbf{I}_{S(d_i)}(m_{it}^*) \quad (\text{B12})$$

$$\cdot \exp\left[-\frac{1}{2}(\underline{\mathbf{m}} - \underline{\mathbf{m}}) \underline{\mathbf{H}}_q (\underline{\mathbf{m}} - \underline{\mathbf{m}})'\right] \quad (\text{B13})$$

$$\cdot \exp\left[-\frac{1}{2} \underline{h}_l (\underline{\mathbf{l}} - \underline{\mathbf{l}})^2\right] c_{(-1,1)}(\underline{\mathbf{l}}), \quad (\text{B14})$$

where $S(0) = (-\infty, 0)$ and $S(1) = [0, \infty)$ in (B12). The component (B9) of the posterior distribution arises from (B1)-(B4); (B10) from (B1), (B3) and (B4); (B11) from (B3) and (B4); (B12) from (B6); (B13) from (B7); and (B14) from (B8).

Conditional distribution of $\mathbf{m} = (\tilde{\mathbf{q}}', \mathbf{q}')'$. The kernel from (B9), (B10), (B11) and (B13) is

$$\exp\left[-\frac{1}{2} \sum_{i=1}^n \sum_{t=1}^{T_i} h_{it}^m (m_{it}^{*m} - \underline{\mathbf{m}}' s_{it}^m)^2\right] \quad (\text{B15})$$

$$\cdot \exp\left[-\frac{1}{2}(\underline{\mathbf{m}} - \underline{\mathbf{m}}) \underline{\mathbf{H}}_q (\underline{\mathbf{m}} - \underline{\mathbf{m}})'\right].$$

Expression (B15) as a function of \mathbf{m} is proportional to the product of (B9), (B10) and (B11), with

$$\mathbf{s}'_{i1}{}^m = \left(\mathbf{s}'_{i1}, \mathbf{0}' \right)_{1 \times p_2}, \quad m_{i1}^{*m} = m_{i1}^*, \quad h_{i1}^m = 1 - I^2 \quad (i = 1, \dots, n)$$

from (B9);

$$\mathbf{s}'_{i2}{}^m = \left(-\mathbf{s}'_{i1} \underline{\mathbf{l}}, \mathbf{s}'_{i2} \right), \quad m_{i2}^{*m} = m_{i2}^* - \underline{\mathbf{l}} m_{i1}^*, \quad h_{i2}^m = 1 \quad (i \in \Omega_2)$$

from (B10); and

$$\mathbf{s}'_{it}{}^m = \left(\mathbf{0}'_{1 \times p_1}, \mathbf{s}'_{it} - \mathbf{s}'_{i,t-1} \underline{\mathbf{l}} \right), \quad m_{it}^{*m} = m_{it}^* - \underline{\mathbf{l}} m_{i,t-1}^*, \quad h_{it}^m = 1 \quad (t = 3, \dots, T_i, i \in \Omega_3)$$

from (B11). Applying *Result 2* of Appendix G,

$$\mathbf{m} \sim \mathbf{N}(\underline{\mathbf{q}}, \underline{\mathbf{H}}_m^{-1})$$

with

$$\underline{\mathbf{H}}_m = \sum_{i=1}^n \sum_{t=1}^{T_i} \mathbf{s}_{it}^{*m} \mathbf{s}_{it}^{*m} + \underline{\mathbf{H}}_m, \quad \underline{\mathbf{m}} = \underline{\mathbf{H}}_m^{-1} \left[\sum_{i=1}^n \sum_{t=1}^{T_i} \mathbf{s}_{it}^{*m} m_{it}^{*m} + \underline{\mathbf{H}}_m \underline{\mathbf{m}} \right].$$

Conditional distribution of $\underline{\mathbf{l}}$. The kernel from (B9), (B10), (B11) and (B14) is

$$(1 - I^2)^{n/2} \exp\left\{\frac{1}{2} I^2 \left[\sum_{i \notin \Omega_2} (m_{i1}^* - \tilde{\mathbf{q}}' \mathbf{s}_{i1})^2 + \sum_{i \in \Omega_3} \sum_{t=3}^{T_i} (m_{i,t-1}^* - \mathbf{q}' \mathbf{s}_{i,t-1})^2 \right]\right\}$$

$$\cdot \exp\left\{\left[\sum_{i \in \Omega_2} (m_{i2}^* - \mathbf{q}' \mathbf{s}_{i2}) (m_{i1}^* - \tilde{\mathbf{q}}' \mathbf{s}_{i1})^2 + \sum_{i \in \Omega_3} \sum_{t=3}^{T_i} (m_{it}^* - \mathbf{q}' \mathbf{s}_{it}) (m_{i,t-1}^* - \mathbf{q}' \mathbf{s}_{i,t-1}) \right]\right\}$$

$$\cdot \exp\left[-\frac{1}{2} \underline{h}_l (\underline{\mathbf{l}} - \underline{\mathbf{l}})^2\right] c_{(-1,1)}(\underline{\mathbf{l}}),$$

$$\propto (1 - I^2)^{n/2} \exp\left[-\tilde{h}_l (\mathbf{1} - \tilde{I})^2 / 2\right] \quad (\text{B16})$$

with

$$\begin{aligned} \tilde{h}_l &= \underline{h}_l \mathbf{1} - \sum_{i \in \Omega_2} (m_{i1}^* - \tilde{\mathbf{q}}' \mathbf{s}_{i1})^2 + \sum_{i \in \Omega_3} \sum_{t=3}^{T_i} (m_{i,t-1}^* - \mathbf{q}' \mathbf{s}_{i,t-1})^2, \\ \tilde{I} &= \tilde{h}_l^{-1} \left[\underline{h}_l \mathbf{1} + \sum_{i \in \Omega_2} (m_{i2}^* - \mathbf{q}' \mathbf{s}_{i2})(m_{i1}^* - \tilde{\mathbf{q}}' \mathbf{s}_{i1}) + \sum_{i \in \Omega_3} \sum_{t=3}^{T_i} (m_{it}^* - \mathbf{q}' \mathbf{s}_{it})(m_{i,t-1}^* - \mathbf{q}' \mathbf{s}_{i,t-1}) \right] \end{aligned}$$

The kernel (B16) is log-concave. Hence drawings can be made from the conditional distribution using the algorithm of Gilks and Wild (1992).

Conditional distribution of $\left\{m_{it}^\right\}_{t=1}^{T_i}$.* Consider the distribution of each m_{it}^* conditional on all other parameters and latent variables. To simplify the notation, define

$$u_{it}^* = \begin{cases} m_{i1}^* - \tilde{\mathbf{q}}' \mathbf{s}_{i1} & \text{if } t = 1; \\ m_{it}^* - \mathbf{q}' \mathbf{s}_{it} & \text{if } t \geq 2 \end{cases}.$$

In each case the conditional distribution of u_{it}^* is univariate normal, subject to the constraint

$$\begin{aligned} u_{i1}^* &\geq -\tilde{\mathbf{q}}' \mathbf{s}_{i1} \text{ if } d_{i1} = 1, \quad u_{i1}^* < -\mathbf{q}' \mathbf{s}_{i1} \text{ if } d_{i1} = 0; \\ u_{it}^* &\geq -\mathbf{q}' \mathbf{s}_{it} \text{ if } d_{it} = 1, \quad u_{it}^* < -\mathbf{q}' \mathbf{s}_{it} \text{ if } d_{it} = 0 \quad (t \geq 2). \end{aligned}$$

If $T_i = 1$ the univariate normal kernel for u_{i1}^* is

$$\exp\left[-\frac{1}{2}(1 - I^2)u_{i1}^{*2}\right]$$

from (B9), whence

$$u_{i1}^* \sim \text{N}\left[0, (1 - I^2)^{-1}\right]$$

If $T_i = 2$ the univariate normal kernel for u_{i1}^* and u_{i2}^* is

$$\exp\left[-\frac{1}{2}(1 - I^2)u_{i1}^{*2} + (u_{i2}^* - Iu_{i1}^*)^2\right] \quad (\text{B17})$$

from (B9) and (B10), whence

$$u_{i1}^* \sim \text{N}(Iu_{i2}^*, 1), \quad u_{i2}^* \sim \text{N}(Iu_{i1}^*, 1). \quad (\text{B18})$$

If $T_i \geq 3$ the univariate normal kernel for u_{i1}^* is given by (B17) and its conditional distribution by (B18). For $t = 2, \dots, T_i - 1$ the univariate normal kernel is

$$\exp\left\{\frac{1}{2}\left[(u_{it}^* - Iu_{i,t-1}^*)^2 + (u_{i,t+1}^* - Iu_{it}^*)^2\right]\right\}$$

from (B11), which is proportional to

$$\exp\left\{-\frac{1}{2}(1 + I^2)\left[u_{it}^* - \frac{I}{1 + I^2}(u_{i,t-1}^* + u_{i,t+1}^*)\right]^2\right\}.$$

Hence

$$u_{it}^* \sim N\left[\frac{\mathbf{I}}{1 + \mathbf{I}^2} (u_{i,t-1}^* + u_{i,t+1}^*), (1 + \mathbf{I}^2)^{-1}\right] \quad (t = 2, \dots, T_i - 1).$$

When $t = T_i$ the univariate normal kernel is

$$\exp\left\{-\frac{1}{2}\left[u_{i,T_i}^* - \mathbf{I}u_{i,T_i-1}^*\right]^2\right\}$$

implying

$$u_{i,T_i}^* \sim N(\mathbf{I}u_{i,T_i-1}^*, 1).$$

C. Conditional Distributions of Early Earnings and Marital Status

In the PSID data pertaining to a segment of an individual's working career are available. While this segment includes period 1 for some individuals, for most individuals it does not. The distribution of the stochastic variables in the earnings and marital status models is nonstationary. The covariate vectors \mathbf{x}_{it} and \mathbf{s}_{it} generally include age and education. Conditional on covariates, distributions are nonstationary because the first period's conditional distribution of earnings (A1) differs from subsequent periods (A2) and the first period's conditional marital status (B1) differs from later periods (B3). Therefore analytical expressions for the joint distribution of the observed stochastic variables conditional on deterministic covariates are quite difficult to derive and would probably be impractical to use in any event. It is more direct to derive conditional distributions involving an individual's earnings and marital status from period 1 to the period immediately preceding the first observed period. These distributions provide the foundation of the Gibbs sampling - data augmentation algorithm described in the text.

In what follows, S_i denotes the first period individual i is observed. Let Ψ_t denote the set of individuals for whom $S_i > t$. Since earnings in periods $1, \dots, S_i - 1$ are unobserved, denote them by $y_{i1}^*, \dots, y_{i,S_i-1}^*$, respectively, and let $\mathbf{y}_i^* = (y_{i1}^*, \dots, y_{i,S_i-1}^*)'$. Similarly denote marital status in these periods by $d_{i1}^*, \dots, d_{i,S_i-1}^*$ and let $\mathbf{d}_i^* = (d_{i1}^*, \dots, d_{i,S_i-1}^*)'$; let $\mathbf{m}_i^* = (m_{i1}^*, \dots, m_{i,S_i-1}^*)'$. If $S_i = 1$ (equivalently, $i \notin \Psi_1$) then these vectors are null.

C.1. Conditional distribution of \mathbf{y}_i^*

To isolate the presence of earnings in the marital status model rewrite (B3) as

$$u_{it} = \mathbf{q}_{it}^y y_{i,t-1} + \mathbf{x}_{it} \quad (t = 1, \dots, S_i + 1).$$

The coefficient \mathbf{q}_{it}^y can depend on both i and t because \mathbf{s}_{it} in (B3) may include multiplicative interactions of $y_{i,t-1}$ with covariates.

The kernel density for \mathbf{y}_i^* ($i = 1, \dots, n$) is given by the product of (A14), (A15), (A16), (B10) and (B11). It factors into n separate components, one for each \mathbf{y}_i^* : i.e., the \mathbf{y}_i^* are mutually conditionally independent. The posterior density kernel $f_{\mathbf{y}_i^*}$ may be expressed

$$\exp \left[-\frac{1}{2} \sum_{t=1}^{T_i} h_{it} \left(p_{it} - \mathbf{y}_i^* \mathbf{a}_t \right)^2 - \frac{1}{2} \sum_{t=2}^{S_i+1} \left(q_{it} - \mathbf{y}_i^* \mathbf{b}_t \right)^2 \right]. \quad (\text{C1})$$

Expression (C1) as a function of \mathbf{y}_i^* is proportional to the product of the i 'th multiplicative components of (A14), (A15), (A16), (B10) and (B11). The vector \mathbf{a}_t consists of the first $S_i - 1$ elements of row t of the matrix $\mathbf{A} = [a_{ts}]$ in which

$$\begin{aligned} a_{11} &= -1; \\ a_{21} &= \mathbf{g} + \tilde{\mathbf{r}} + \mathbf{g}_1 \mathbf{f}, \quad a_{22} = -1; \\ a_{31} &= (1 - \mathbf{r}) \mathbf{g}_1 \mathbf{f} - \mathbf{r} \mathbf{g}, \quad a_{32} = \mathbf{r} + \mathbf{g}, \quad a_{33} = -1; \\ a_{t1} &= (1 - \mathbf{r}) \mathbf{g}_1 \mathbf{f}, \quad a_{t,t-2} = -\mathbf{r} \mathbf{g}, \quad a_{t,t-1} = \mathbf{r} + \mathbf{g}, \quad a_{tt} = -1 \quad (t \geq 4); \end{aligned}$$

and all other elements are 0. The scalars $\varphi_{i1}, \dots, \varphi_{i,S_i+1}$ are

$$\begin{aligned} p_{i1} &= -\mathbf{a}_{i1} - \tilde{\mathbf{b}}' \mathbf{x}_{i1}; \\ p_{i2} &= -\mathbf{a}_{i2} + (\tilde{\mathbf{r}} + \mathbf{g}_1 \mathbf{f}) \tilde{\mathbf{b}}' \mathbf{x}_{i1} - \mathbf{g}_1 \mathbf{b}' \mathbf{x}_{i2} - \mathbf{g}_1 \mathbf{t}_i + \tilde{y}_{i2}; \\ p_{it} &= -\mathbf{a}_{it} + \mathbf{g}_1 \mathbf{f} (1 - \mathbf{r}) \tilde{\mathbf{b}}' \mathbf{x}_{i1} + \mathbf{r} \mathbf{g}_1 \mathbf{b}' \mathbf{x}_{i,t-1} - \mathbf{g}_1 \mathbf{b}' \mathbf{x}_{it} - \mathbf{g}_1 (1 - \mathbf{r}) \mathbf{t}_i \\ &\quad + \tilde{y}_{it} - (\mathbf{g} + \mathbf{r}) \tilde{y}_{i,t-1} + \mathbf{g} \mathbf{r} \tilde{y}_{i,t-2} \quad (t \geq 3); \end{aligned}$$

where $\tilde{y}_{it} = 0$ if $t < S_i$ and $\tilde{y}_{it} = y_{it}$ if $t \geq S_i$.

Similarly the vector \mathbf{b}_t^i consists of the first $S_i - 1$ elements of row t of the matrix $\mathbf{B} = [b_{ts}^i]$ in which

$$\begin{aligned} b_{21}^i &= \mathbf{q}_{i2}^{y^*}; \\ b_{t,t-2}^i &= -\mathbf{I} \mathbf{q}_{i,t-1}^{y^*}, \quad b_{t,t-1}^i = \mathbf{q}_{it}^{y^*} \quad (t \geq 3). \end{aligned}$$

The scalars are $q_{it} = u_{it} - \mathbf{I} u_{i,t-1} - \mathbf{q}_{it}^{y^*} \tilde{y}_{i,t-1}$.

Applying *Result 2* of Appendix G,

$$\mathbf{y}_i^* \sim \mathbf{N}(\bar{\mathbf{y}}_i^*, \bar{\mathbf{H}}_{\mathbf{y}_i^*}^{-1}),$$

with

$$\bar{\mathbf{H}}_{\mathbf{y}_i^*} = \sum_{t=1}^{T_i} h_{it} \mathbf{a}_t \mathbf{a}_t' + \sum_{t=2}^{S_i+1} \mathbf{b}_t^i \mathbf{b}_t^{i'}, \quad \bar{\mathbf{y}}_i^* = \bar{\mathbf{H}}_{\mathbf{y}_i^*}^{-1} \left[\sum_{t=1}^{T_i} h_{it} \mathbf{a}_t p_{it} + \sum_{t=2}^{S_i+1} \mathbf{b}_t^i q_{it} \right]$$

In the special case $T_i = S_i$ the second and fourth summations extend only to S_i , not $S_i + 1$.

C.2. Conditional distributions for (m_{it}^*, d_{it}^*)

The kernel density for $(\mathbf{d}_i^*, \mathbf{m}_i^*)$ ($i = 1, \dots, n$) is given by the product of (A14), (A15), (A16), (B10), (B11) and (B12). It factors into separate components, one for each $(\mathbf{d}_i^*, \mathbf{m}_i^*)$, so the $(\mathbf{d}_i^*, \mathbf{m}_i^*)$ are mutually conditionally independent. For given i the joint conditional distribution of $\{m_{it}^*, d_{it}^*\}_{t=1}^{S_i-1}$ is complicated by the truncations of the m_{it}^* implied by the d_{it}^* . However the distribution of a single (m_{it}^*, d_{it}^*) pair conditional on all the parameters and latent variables is more tractable.

Let $f_i(\mathbf{d}_i^*)$ denote the kernel of the posterior of the earnings model in \mathbf{d}_i^* : it is proportional to the product of (A14), (A15), and (A16). As in Appendix B, define

$$u_{i1}^* = m_{i1}^* - \tilde{\mathbf{q}}' \mathbf{s}_{i1},$$

$$u_{it}^* = u_{it}^*(d_{i,t-1}) = m_{it}^* - \mathbf{q}' \mathbf{s}_{it} \quad (t = 2, \dots, S_i + 1).$$

The notation $u_{it}^*(d_{i,t-1})$ reflects the possible presence of $d_{i,t-1}$ in \mathbf{s}_{it} .

The posterior kernel in (m_{it}^*, d_{it}) alone ($t = 2, \dots, S_i - 1$) is

$$f_i(\mathbf{d}_i^*) \exp\left[-\frac{1}{2}(u_{it}^* - I u_{i,t-1}^*)^2\right] \exp\left\{\frac{1}{2}[u_{i,t+1}^*(d_{it}) - I u_{it}^*]^2\right\} \exp\left\{\frac{1}{2}[u_{i,t+2}^* - I u_{i,t+1}^*(d_{it})]^2\right\} \quad (\text{C2})$$

$$\propto f_i(\mathbf{d}_i^*) \exp\left\{-\frac{1}{2}(1 + I^2)[u_{it}^* - (1 + I^2)^{-1} I (u_{i,t-1}^* + u_{i,t+1}^*(d_{it}))]^2\right\}$$

$$\cdot \exp\left\{\frac{1}{2}\left[(u_{i,t+2}^* - I u_{i,t+1}^*(d_{it}))^2 - (1 + I^2)^{-1} I^2 (u_{i,t-1}^* + u_{i,t+1}^*(d_{it}))^2 + I^2 u_{i,t-1}^{*2} + u_{i,t+1}^*(d_{it})^2\right]\right\}$$

The posterior kernel of d_{it} alone, conditional on all other parameters and latent variables except m_{it}^* , is found by integrating u_{it}^* from this expression:

$$p(d_{it}^* = j | \cdot) \propto f_i(d_{i1}, \dots, d_{i,t-1}, j, d_{i,t+1}, \dots, d_{i,T_i})$$

$$\cdot \Phi\left\{-1\right\}^{(j+1)} (1 + I^2)^{1/2} \left[\mathbf{q}' \mathbf{s}_{it} + (1 + I^2)^{-1} I (u_{i,t-1}^* + u_{i,t+1}^*(j))\right]$$

$$\cdot \exp\left\{\frac{1}{2}\left[(u_{i,t+2}^* - I u_{i,t+1}^*(d_{it}))^2 - (1 + I^2)^{-1} I^2 (u_{i,t-1}^* + u_{i,t+1}^*(d_{it}))^2 + u_{i,t+1}^*(d_{it})^2\right]\right\} \quad (\text{C3})$$

Conditional on $d_{it} = j$ and all other parameters and latent variables,

$$u_{it}^* \sim N\left[(1 + I^2)^{-1} I (u_{i,t-1}^* + u_{i,t+1}^*(j)), (1 + I^2)^{-1}\right]$$

subject to the constraint $u_{it}^* < -\mathbf{q}' \mathbf{s}_{it}$ if $j = 0$ and $u_{it}^* \geq -\mathbf{q}' \mathbf{s}_{it}$ if $j = 1$.

The posterior kernel in (m_{i1}^*, d_{i1}) alone is

$$f_i(\mathbf{d}_i^*) \exp\left[-\frac{1}{2}(1 - I^2) u_{i1}^{*2}\right] \exp\left\{\frac{1}{2}[u_{i2}^*(d_{i1}) - I u_{i1}^*]^2\right\} \exp\left\{\frac{1}{2}[u_{i3}^* - I u_{i2}^*(d_{i1})]^2\right\} \quad (\text{C4})$$

$$\propto f_i(\mathbf{d}_i^*) \exp\left\{\frac{1}{2}[u_{i1}^* - I u_{i2}^*(d_{i1})]^2\right\} \exp\left\{\frac{1}{2}[u_{i3}^* - I u_{i2}^*(d_{i1})]^2\right\} \exp\left[-\frac{1}{2}(1 - I^2) u_{i2}^*(d_{i1})\right]$$

Hence

$$p(d_{i1} = j | \cdot) \propto f_i(j, d_{i2}, \dots, d_{i,T_i}) \Phi\left\{-1\right\}^{(j+1)} [\tilde{\mathbf{q}}' \mathbf{s}_{i1} + I u_{i2}^*(j)]$$

$$\cdot \exp\left\{\frac{1}{2}[u_{i3}^* - I u_{i2}^*(j)]^2\right\} \exp\left[-\frac{1}{2}(1 - I^2) u_{i2}^*(j)^2\right] \quad (\text{C5})$$

Given $d_{i1} = j$ and all other parameters and latent variables,

$$u_{i1}^* \sim N[I u_{i2}^*(j), 1]$$

subject to the constraint $u_{i1}^* < -\tilde{\mathbf{q}}' \mathbf{s}_{i1}$ if $j = 0$ and $u_{i1}^* \geq -\tilde{\mathbf{q}}' \mathbf{s}_{i1}$ if $j = 1$.

To implement drawing from these conditional distributions some further elaboration on $\mathbf{u}_{it}^*(d_{i,t-1})$ and $f_i(\mathbf{d}_i^*)$ is useful. To isolate the presence of lagged marital status in the marital status model rewrite (B1) and (B3) as

$$m_{it}^* = v_{it} + \mathbf{q}_{it}^{\mathbf{d}^*} d_{i,t-1} + \mathbf{x}_{it} \quad (t = 1, \dots, S_i + 1; \mathbf{q}_{i1}^{\mathbf{d}^*} = 0).$$

The coefficient $\mathbf{q}_{it}^{\mathbf{d}^*}$ can depend on both i and t because \mathbf{s}_{it} in (B3) may include multiplicative interaction of $d_{i,t-1}$ with covariates. Then,

$$\mathbf{u}_{it}^*(d_{i,t-1}) = m_{it}^* - v_{it} - \mathbf{q}_{it}^{\mathbf{d}^*} d_{i,t-1}.$$

To characterize $f_i(\mathbf{d}_i^*)$, first write (A1) as

$$w_{i1} = \mathbf{b}_{i1}^{\mathbf{d}^*} d_{i1} + \mathbf{e}_{i1}.$$

and (A2) as

$$w_{it} = \mathbf{b}_{it}^{\mathbf{d}^*} d_{it} + \mathbf{b}_{2it}^{\mathbf{d}^*} d_{i,t-1} + \mathbf{g}_1 \mathbf{f}(w_{i1} - \mathbf{b}_{i1}^{\mathbf{d}^*} d_{i1}) + \mathbf{e}_{it} \quad (t = 2, \dots, T_i).$$

With this notation the kernel $f_i(\mathbf{d}_i^*)$ expressed by (A14)-(A16) is

$$\begin{aligned} & \exp\left[-\frac{1}{2} h_{i1} (w_{i1} - \mathbf{a}_{i1} - \mathbf{b}_{i1}^{\mathbf{d}^*} d_{i1})^2\right] \\ & \cdot \exp\left\{-\frac{1}{2} h_{i2} \left[w_{i2} - \mathbf{a}_{i2} - \mathbf{b}_{i2}^{\mathbf{d}^*} d_{i2} - \mathbf{b}_{2i2}^{\mathbf{d}^*} d_{i1} - (\tilde{\mathbf{r}} + \mathbf{g}_1 \mathbf{f})(w_{i1} - \mathbf{b}_{i1}^{\mathbf{d}^*} d_{i1})\right]^2\right\} \\ & \cdot \prod_{t=3}^{T_i} \exp\left\{\frac{1}{2} h_{it} \left[w_{it} - \mathbf{a}_{it} - \mathbf{b}_{it}^{\mathbf{d}^*} d_{it} - \mathbf{b}_{2it}^{\mathbf{d}^*} d_{i,t-1} - \mathbf{r}(w_{i,t-1} - \mathbf{b}_{i,t-1}^{\mathbf{d}^*} d_{i,t-1} - \mathbf{b}_{2i,t-1}^{\mathbf{d}^*} d_{i,t-2})\right.\right. \\ & \quad \left.\left.+ (\mathbf{r} - 1) \mathbf{g}_1 \mathbf{f}(w_{i1} - \mathbf{b}_{i1}^{\mathbf{d}^*} d_{i1})\right]^2\right\} \\ & = \exp\left[-\frac{1}{2} \sum_{t=1}^{T_i} h_{it} (r_{it} - \mathbf{b}_{it}^{\mathbf{d}^*} \mathbf{d}_i)^2\right] \end{aligned}$$

The vector \mathbf{b}_{it} consists of the first $S_i - 1$ elements of row t of the matrix $\mathbf{B}_i = [b_{its}]$ in which

$$\begin{aligned} b_{i11} &= \mathbf{b}_{i1}^{\mathbf{d}^*}; \\ b_{i21} &= \mathbf{b}_{2i2}^{\mathbf{d}^*} - (\tilde{\mathbf{r}} + \mathbf{g}_1 \mathbf{f}) \mathbf{b}_{i1}^{\mathbf{d}^*}, \quad b_{i22} = \mathbf{b}_{i2}^{\mathbf{d}^*}; \\ b_{i31} &= (\mathbf{r} - 1) \mathbf{g}_1 \mathbf{f} \mathbf{b}_{i1}^{\mathbf{d}^*} - \mathbf{r} \mathbf{b}_{2i2}^{\mathbf{d}^*}, \quad b_{i32} = \mathbf{b}_{2i3}^{\mathbf{d}^*} - \mathbf{r} \mathbf{b}_{i2}^{\mathbf{d}^*}, \quad b_{i33} = \mathbf{b}_{i3}^{\mathbf{d}^*}; \\ b_{it1} &= (\mathbf{r} - 1) \mathbf{g}_1 \mathbf{f} \mathbf{b}_{i1}^{\mathbf{d}^*}, \quad b_{i,t-2} = -\mathbf{r} \mathbf{b}_{2i,t-1}^{\mathbf{d}^*}, \quad b_{i,t-1} = \mathbf{b}_{2it}^{\mathbf{d}^*} - \mathbf{r} \mathbf{b}_{i,t-1}^{\mathbf{d}^*}, \quad b_{it} = \mathbf{b}_{it}^{\mathbf{d}^*} \quad (t = 4, \dots, T_i) \end{aligned}$$

The scalars r_{it} are

$$\begin{aligned} r_{i1} &= w_{i1} - \mathbf{a}_{i1}; \\ r_{i2} &= w_{i2} - \mathbf{a}_{i2} - \mathbf{b}_{i2}^{\mathbf{d}^*} \tilde{d}_{i2} - (\tilde{\mathbf{r}} + \mathbf{g}_1 \mathbf{f}) w_{i1}; \\ r_{it} &= w_{it} - \mathbf{a}_{it} - \mathbf{b}_{it}^{\mathbf{d}^*} \tilde{d}_{it} - \mathbf{b}_{2it}^{\mathbf{d}^*} \tilde{d}_{i,t-1} - \mathbf{r}(w_{i,t-1} - \mathbf{b}_{i,t-1}^{\mathbf{d}^*} \tilde{d}_{i,t-1} - \mathbf{b}_{2i,t-1}^{\mathbf{d}^*} \tilde{d}_{i,t-2}) \\ & \quad + (\mathbf{r} - 1) \mathbf{g}_1 \mathbf{f} w_{i1} \quad (t = 3, \dots, T_i) \end{aligned}$$

where $\tilde{d}_{it} = 0$ if $t < S_i$ and $\tilde{d}_{it} = d_{it}$ if $t \geq S_i$.

In the special case $T_i = S_i \geq 3$, the posterior kernel (C2) must be modified by eliminating the term $\exp\left\{\frac{1}{2}\left[u_{i,t+2}^* - I u_{i,t+1}^*(d_{it})\right]^2\right\}$. Carrying the elimination through, the term $\left(u_{i,t+2}^* - I u_{i,t+1}^*(d_{it})\right)^2$ is then eliminated from (C3). The conditional distribution of u_{it}^* is unaffected.

In the special case $T_i = S_i = 2$ the posterior kernel (C4) must be modified by eliminating the term $\exp\left\{\frac{1}{2}\left[u_{i3}^* - I u_{i2}^*(d_{i1})\right]^2\right\}$. Carrying this elimination through, the term $\exp\left\{\frac{1}{2}\left[u_{i3}^* - I u_{i2}^*(j)\right]^2\right\}$ is then dropped from (C5). The conditional distribution of u_{i1}^* is unaffected.

D. Simulation of the Earnings and Marital Status Models

For purposes of describing the simulation some minor elaborations on the notation of Appendices A and B are useful. Let

\mathbf{z}_i = Vector of time-invariant personal characteristics of individual i

\mathbf{z}_{it} = Vector of time-varying personal characteristics of individual i at time t , possibly including age, current and lagged marital status, lagged earnings, and interactions of these variables with each other and time-invariant personal characteristics.

We may then write the state equations of the marital status model

$$m_{i1}^* = \tilde{\mathbf{q}}_1' \mathbf{z}_i + \tilde{\mathbf{q}}_2' \mathbf{z}_{i1} + \mathbf{x}_{i1}, \quad m_{it}^* = \mathbf{q}_1' \mathbf{z}_i + \mathbf{q}_2' \mathbf{z}_{it} + \mathbf{x}_{it} \quad (t > 1),$$

where $\tilde{\mathbf{q}}_1$ is conformable with \mathbf{z}_i and with $\tilde{\mathbf{q}}$ in (B1), and \mathbf{q}_1 and \mathbf{q}_2 are conformable with \mathbf{z}_i and \mathbf{z}_{it} and with \mathbf{q} in (B3). We may write the earnings equations

$$y_{i1} = \tilde{\mathbf{b}}_1' \mathbf{z}_i + \tilde{\mathbf{b}}_2' \mathbf{z}_{i1} + \mathbf{e}_{i1}, \quad y_{it} = \mathbf{b}_1' \mathbf{z}_i + \mathbf{b}_2' \mathbf{z}_{it} + \mathbf{f} \mathbf{e}_{i1} + \mathbf{t}_i + \mathbf{e}_{it} \quad (t > 1),$$

where $\tilde{\mathbf{b}}_1$ and $\tilde{\mathbf{b}}_2$ are conformable with \mathbf{z}_i and \mathbf{z}_{i1} and with $\tilde{\mathbf{b}}$ in (A1), and \mathbf{b}_1 and \mathbf{b}_2 are conformable with \mathbf{z}_i and \mathbf{z}_{it} and with \mathbf{b} and \mathbf{g} in (A2).

For a given individual i , a single simulation proceeds as follows.

$$\mathbf{x}_{i1} \sim N\left[0, (1 - I^2)^{-1}\right]$$

$$m_{i1}^* = \tilde{\mathbf{q}}_1' \mathbf{z}_i + \tilde{\mathbf{q}}_2' \mathbf{z}_{i1} + \mathbf{x}_{i1}$$

$$d_{i1} = \begin{cases} 1 & \text{if } m_{i1}^* \geq 0 \\ 0 & \text{if } m_{i1}^* < 0 \end{cases}$$

$$\mathbf{t}_i \sim N(0, \mathbf{s}_t^2)$$

$$\mathbf{e}_{ij} \sim N(\mathbf{a}_{1(j)}, h_{1(j)}^{-1}) \text{ with probability } p_{1(j)} \quad (j = 1, 2, 3)$$

$$y_{i1} = \tilde{\mathbf{b}}_1' \mathbf{z}_i + \tilde{\mathbf{b}}_2' \mathbf{z}_{i1} + \mathbf{e}_{i1}$$

Set $t = 2$

(A) Update \mathbf{z}_{it} except for d_{it}

$$\mathbf{y}_{it} \sim N(0, 1)$$

$$\mathbf{x}_{it} = I \mathbf{x}_{i,t-1} + \mathbf{y}_{it}$$

$$m_{it}^* = \mathbf{q}_1' \mathbf{z}_i + \mathbf{q}_2' \mathbf{z}_{it} + \mathbf{x}_{it}$$

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

Update \mathbf{z}_{it} to reflect d_{it}

$\mathbf{h}_t \sim N(\mathbf{a}_{2(j)}, h_{2(j)}^{-1})$ with probability $p_{2(j)}$ ($j = 1, 2, 3$)

If $t = 2$, $\mathbf{e}_{it} = \mathbf{r}_1 \mathbf{e}_{i,t-1} + \mathbf{h}_t$

If $t > 2$, $\mathbf{e}_{it} = \mathbf{r}_2 \mathbf{e}_{i,t-1} + \mathbf{h}_t$

$y_{it} = \mathbf{b}'_1 \mathbf{z}_i + \mathbf{b}'_2 \mathbf{z}_{it} + \mathbf{e}_{it}$

If $\text{Age}(t) \leq 65$ increment t by 1 and go to (A)

Done

E. 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.

E.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 $d' = (\tilde{b}', b')$ is multivariate normal.

The first 8 covariates are the same for all periods. The covariates and their prior means and standard deviations are as follows.

Covariate	Prior mean	Prior stan dev
Father education missing dummy (1 if missing, 0 if present)	.05	.20
Father high school dummy (1 if known high school grad, else 0)	.05	.10
Father college dummy (1 if known college grad, else 0)	.10	.10
Mother education missing dummy (1 if missing, 0 if present)	.05	.20
Mother high school dummy (1 if known high school grad, else 0)	.05	.10
Mother college dummy (1 if known college grad, else 0)	.10	.10
Race dummy (1 if nonwhite, 0 if white)	-.10	.10
Lagged marital status dummy (1 if married, 0 if not married)	.20	.20

Prior means reflected our best subjective judgment 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 years of education. In our model, the only covariates that affect expected earnings

growth are age and education. Denote the impact of age and education and earnings on expected log earnings growth from 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 as follows.

Age/education linear combination	Prior mean	Prior stan dev
G(25, 35; 12, 12)	.15	.10
G(35, 45; 12, 12)	.10	.10
G(45, 55; 12, 12)	.05	.10
G(25, 25; 12, 16)	.26	.15
G(35, 35; 12, 16)	.34	.20
G(45, 45; 12, 16)	.37	.225
G(55, 55; 12, 16)	.40	.25

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.

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

\mathbf{g} , \mathbf{r}_1 , \mathbf{r}_2 . Normal priors for these serial persistence parameters were each centered at .5, with a standard deviation of .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.

f . The value .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 .5. This prior is independent of priors for all other parameters.

s_t^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 $[\log(2)]^2=.480$. We assigned the prior $4/s_t^2 \sim \mathcal{C}^2(1)$, implying that *a priori* $P(.104 \leq s_t^2 \leq 101.7)=.9$.

Shock distributions. The disturbances e_{it} and h_{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

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

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

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

$$\mathbf{a}_{(1)} \sim N(-3, 1)$$

$$\mathbf{a}_{(2)} = 0$$

$$\mathbf{a}_{(3)} \sim N(.1, .1)$$

(The restriction $\mathbf{a}_{(2)} = 0$ is a normalization, since an intercept term, $\tilde{\mathbf{h}}_0$ or \mathbf{h}_0 , is already included as a covariate.)

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

$$200h_{(1)} \sim \mathcal{C}^2(100) \Rightarrow P[1.268 \leq h_{(1)}^{-1/2} \leq 1.602]=.9;$$

$$125h_{(2)} \sim \mathcal{C}^2(100) \Rightarrow P[1.002 \leq h_{(2)}^{-1/2} \leq 1.266]=.9;$$

$$10h_{(3)} \sim \mathcal{C}^2(100) \Rightarrow P[0.284 \leq h_{(3)}^{-1/2} \leq 0.358]=.9.$$

The prior distribution for \mathbf{p} is multivariate beta, with parameters 2, 48, and 50, implying modal values of .0103, .4895, and .5052 for the prior distribution of \mathbf{p} . The marginal prior distributions are univariate beta, with

$$P(.0036 \leq p_1 \leq .0470)=.90;$$

$$P(.3984 \leq p_2 \leq .5621)=.90;$$

$$P(.4181 \leq p_3 \leq .5819)=.90.$$

E.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 $\mathbf{m}' = (\tilde{\mathbf{q}}', \mathbf{q}')$ 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 .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 as follows.

Age/education linear combination	Prior mean	Prior stan dev
$\Delta(25, 40; 12, 12)$	0	$\Phi^{-1}(.6) - \Phi^{-1}(.5) = .255$
$\Delta(40, 55; 12, 12)$	0	$\Phi^{-1}(.6) - \Phi^{-1}(.5) = .255$
$\Delta(25, 25; 12, 16)$	0	$\Phi^{-1}(.6) - \Phi^{-1}(.5) = .255$
$\Delta(40, 40; 12, 16)$	0	$\Phi^{-1}(.6) - \Phi^{-1}(.5) = .255$
$\Delta(55, 55; 12, 16)$	0	$\Phi^{-1}(.6) - \Phi^{-1}(.5) = .255$

A sixth prior distribution, independent of these five and the other coefficients, assigns a mean of $\Phi^{-1}(.5) = 0$ and standard deviation of $[\Phi^{-1}(.6) - \Phi^{-1}(.5)]/2 = .255$ to the probability of being married at age 25 and education 12.

Of the remaining three covariates for periods beyond the first, the non-white dummy is assigned a mean of 0 and a standard deviation of $\Phi^{-1}(.6) - \Phi^{-1}(.5) = .255$. Lagged marital status is given both a mean and standard deviation of $\Phi^{-1}(.8) - \Phi^{-1}(.2) = .68$. 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 .2 when that probability was about .5 to start: so the standard deviation was chosen to be $\Phi^{-1}(.6) - \Phi^{-1}(.4) = .510$. 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 ρ , with a mean and standard deviation of .7.

F. Prior and Conditional Posterior Distributions in a Discrete Normal Mixtures Model

In the earnings model the shocks \mathbf{e}_{it} and \mathbf{h}_{it} are each distributed as mixtures of three normals. This appendix describes a general approach to Bayesian inference for a set of i.i.d. random variables (x_1, \dots, x_n) with a mixture of normals distribution. It is assumed that (x_1, \dots, x_n) is observed. In the context of the Gibbs sampling algorithm described in Appendix A, this assumption is appropriate.

Suppose that the random variables $\mathbf{x} = (x_1, \dots, x_n)'$ are independently and identically distributed, with the common distribution being a discrete normal mixtures model. Then we may write

$$x_i = \sum_{j=1}^m e_{ij} (\mathbf{a}_{(j)} + h_{(j)}^{-1/2} \mathbf{z}_i)$$

where $\mathbf{z}_i \stackrel{iid}{\sim} N(0, 1)$. The vectors $\mathbf{e}'_i = (e_{i1}, \dots, e_{im})$ are i.i.d multinomial,

$$P(e_{ij} = 1) = p_j.$$

Further denote $\mathbf{a}' = (\mathbf{a}_{(1)}, \dots, \mathbf{a}_{(m)})$, $\mathbf{h} = (h_{(1)}, \dots, h_{(m)})$, $\mathbf{e}' = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$, $L_i = (j: e_{ij} = 1)$, and $N_j = \sum_{i=1}^n e_{ij}$. Then

$$\begin{aligned} p(\mathbf{x}, \mathbf{e} | \mathbf{a}, \mathbf{h}, \mathbf{p}) &= p(\mathbf{e} | \mathbf{p}) p(\mathbf{x} | \mathbf{e}, \mathbf{a}, \mathbf{h}) \\ &\propto \prod_{i=1}^n p_j^{e_{ij}} \prod_{i=1}^n \left\{ h_{(L_i)}^{1/2} \exp \left[-\frac{1}{2} h_{(L_i)} (x_i - \mathbf{a}_{(L_i)})^2 \right] \right\} \\ &= \prod_{j=1}^m p_j^{N_j} h_{(j)}^{N_j/2} \exp \left[\frac{1}{2} \sum_{i=1}^n h_{(L_i)} (x_i - \mathbf{a}' \mathbf{e}_i)^2 \right] \end{aligned}$$

The prior distribution is composed of the following independent components:

$$\mathbf{a} \sim N(\underline{\mathbf{a}}, \underline{\mathbf{H}}_a^{-1}); \quad (\text{F1})$$

$$\underline{\mathbf{n}}_j \underline{s}_j^2 h_{(j)} \sim \mathbf{c}^2(\underline{\mathbf{n}}_j) \quad \underline{s}_j^2 > 0, \underline{\mathbf{n}}_j > 0 \quad (j = 1, 2, 3); \quad (\text{F2})$$

$$\mathbf{p} = (p_1, p_2, p_3) \sim \text{Beta}(r_1, r_2, r_3).$$

The joint posterior density of the parameters and latent variables is

$$\begin{aligned} p(\mathbf{a}, \mathbf{h}, \mathbf{p}, \{\mathbf{e}_i\}_{i=1}^n \mid \{x_i\}_{i=1}^n, \underline{\mathbf{a}}, \underline{\mathbf{H}}_a, \{\underline{s}_j^2, \underline{\mathbf{n}}_j, r_j\}_{j=1}^3) \\ \propto \prod_{i=1}^n [p(x_i | \mathbf{e}_i, \mathbf{a}, \mathbf{h}) p(\mathbf{e}_i | \mathbf{p})] p(\mathbf{a}, \mathbf{h}, \mathbf{p} \mid \underline{\mathbf{a}}, \underline{\mathbf{H}}_a, \{\underline{s}_j^2, \underline{\mathbf{n}}_j, r_j\}_{j=1}^3) \\ \propto \prod_{j=1}^m p_j^{N_j} h_{(j)}^{N_j/2} \exp \left[\frac{1}{2} \sum_{i=1}^n h_{(L_i)} (x_i - \mathbf{a}' \mathbf{e}_i)^2 \right] \quad (\text{F3}) \end{aligned}$$

$$\cdot \exp\left[-\frac{1}{2}(\mathbf{a} - \underline{\mathbf{a}})' \underline{\mathbf{H}}_a (\mathbf{a} - \underline{\mathbf{a}})\right] \quad (\text{F4})$$

$$\cdot \prod_{j=1}^m h_{(j)}^{\left(\frac{\underline{\mathbf{n}}_j - 2}{2}\right)} \exp\left(-\frac{1}{2} \underline{\mathbf{n}}_j \underline{\mathbf{s}}_j^2 h_{(j)}\right) \quad (\text{F5})$$

$$\cdot \prod_{j=1}^m p_j^{r_j - 1} \quad (\text{F6})$$

In application, \mathbf{x} typically is not observed directly but is derived from a parent model in which a discrete mixture of normals distribution for an unobservable has been specified; e.g., \mathbf{x} might be the vector of disturbances in a regression equation. Conditional on the other parameters of the parent model for the observables we observe \mathbf{x} , however. Given the usual Gibbs sampling data augmentation procedures it suffices to know the posterior distributions of each of \mathbf{a} , \mathbf{h} , \mathbf{p} and $\{\mathbf{e}_i\}_{i=1}^n$ conditional on each other and \mathbf{x} .

Conditional distribution of \mathbf{a} . The kernel from (F3)-(F4) implies that

$$\mathbf{a} \sim \text{N}(\bar{\mathbf{a}}, \bar{\mathbf{H}}_a), \text{ with } \bar{\mathbf{H}}_a = \sum_{i=1}^n h_{(L_i)} \mathbf{e}_i \mathbf{e}_i' + \underline{\mathbf{H}}_a, \quad \bar{\mathbf{a}} = \bar{\mathbf{H}}_a^{-1} \left(\sum_{i=1}^n h_{(L_i)} \mathbf{e}_i x_i + \underline{\mathbf{H}}_a \underline{\mathbf{a}} \right) \quad (\text{F7})$$

Conditional distribution of \mathbf{h} . The kernel from (F3) and (F5) implies that the $h_{(j)}$ are conditionally independent,

$$\left(\sum_{i=1}^n e_{ij} (x_i - \mathbf{a}_{(j)})^2 + \underline{\mathbf{n}}_j \underline{\mathbf{s}}_j^2 \right) h_{(j)} \sim \mathbf{c}^2(N_j + \underline{\mathbf{n}}_j). \quad (\text{F8})$$

Conditional distribution of \mathbf{p} . The kernel from (F3) and (F6) is $\prod_{\ell=1}^m p_j^{(N_j + r_j - 1)}$, and hence $\mathbf{p} \sim \text{Beta}[N_j + r_j, (j = 1, \dots, m)]$.

Conditional distribution of \mathbf{e}_i . From (F3), the \mathbf{e}_i are conditionally mutually independent multinomial, with

$$\text{P}(e_{ij} = 1) \propto p_j h_{(j)}^{1/2} \exp\left[-\frac{1}{2} h_{(j)} (x_i - \mathbf{a}_{(j)})^2\right].$$

Further restrictions of interest.

(1) $\mathbf{a}_{(1)} = \dots = \mathbf{a}_{(m)} = \mathbf{a}$; $\mathbf{a} \sim \text{N}(\underline{\mathbf{a}}, \underline{h}_a^{-1})$. (The distribution is a scale mixture of normals.)

Then $\mathbf{a} \sim \text{N}(\bar{\mathbf{a}}, \bar{h}_a^{-1})$; $\bar{h}_a = \underline{h}_a + \sum_{i=1}^n \sum_{j=1}^m e_{ij} h_{(j)}$, $\bar{\mathbf{a}} = \bar{h}_a^{-1} \left(\underline{h}_a \underline{\mathbf{a}} + \sum_{i=1}^n x_i \sum_{j=1}^m e_{ij} h_{(j)} \right)$

For many parent models, this restriction can be imposed by restricting $\mathbf{a} = 0$ and incorporating the appropriate free parameter in the parent model.

(2) $h_{(1)} = \dots = h_{(m)} = h$; $\underline{\mathbf{n}} \underline{s}^2 h \sim \mathbf{c}^2(\underline{\mathbf{n}})$. (The distribution is a mean mixture of normals.)

Then $\left(\sum_{i=1}^n e_{ij} \left(x_i - \mathbf{a}_{(j)} \right)^2 + \underline{\mathbf{n}} \underline{s}^2 \right) h \sim \mathbf{c}^2(N + \underline{\mathbf{n}})$.

(3) $\mathbf{a}_{(1)} \leq \dots \leq \mathbf{a}_{(m)}$, in addition to (F1). (Means are monotonically nondecreasing.) The kernel in (F4) is multiplied by $\mathbf{c}_{\mathbf{a}: \mathbf{a}_{(1)} < \mathbf{a}_{(2)} < \mathbf{a}_{(3)}}(\mathbf{a})$, and the same inequality constraint is imposed on the conditional distribution (F7). The $\mathbf{a}_{(j)}$ can be drawn in succession as described in Geweke (1991). (This restriction is used in the earnings model.)

(4) $h_{(1)} \leq \dots \leq h_{(m)}$, in addition to (F2). (Precisions are monotonically nondecreasing.) The kernel in (F5) is multiplied by $\mathbf{c}_{h_{(1)} \leq \dots \leq h_{(m)}}(\mathbf{h})$. The $h_{(j)}$ are drawn in succession, each having the posterior distribution (F8) but subject to the constraint $h_{(j-1)} < h_{(j)} < h_{(j+1)}$ ($h_{(0)} = 0, h_{(m+1)} = \infty$).

For a given draw, denote $\mathbf{a}_i = \mathbf{a}_{(L_i)}$ and $h_i = h_{(L_i)}$.

G. Priors and Posteriors for Some Normal and Inverted Gamma Likelihood Functions

Result 1. (Likelihood function for the coefficients of a normal linear regression model with heteroscedastic disturbances.)

$$\exp\left[-\frac{1}{2}\sum_{t=1}^T h_t (y_t - \mathbf{b}'\mathbf{x}_t)^2\right] \propto \exp\left[-\frac{1}{2}(\mathbf{b} - \hat{\mathbf{b}})' \mathbf{H}(\mathbf{b} - \hat{\mathbf{b}})\right] \quad (\text{G1})$$

with $\mathbf{H} = \sum_{t=1}^T h_t \mathbf{x}_t \mathbf{x}_t'$ and $\hat{\mathbf{b}} = \mathbf{H}^{-1} \sum_{t=1}^T h_t \mathbf{x}_t y_t$.

Proof.

$$\begin{aligned} (\mathbf{b} - \hat{\mathbf{b}})' \mathbf{H}(\mathbf{b} - \hat{\mathbf{b}}) &= \mathbf{b}' \mathbf{H} \mathbf{b} - 2\mathbf{b}' \mathbf{H} \hat{\mathbf{b}} + \hat{\mathbf{b}}' \mathbf{H} \hat{\mathbf{b}} \\ &= \mathbf{b}' \left(\sum_{t=1}^T h_t \mathbf{x}_t \mathbf{x}_t' \right) \mathbf{b} - 2\mathbf{b}' \left(\sum_{t=1}^T h_t \mathbf{x}_t y_t \right) \hat{\mathbf{b}} + \hat{\mathbf{b}}' \mathbf{H} \hat{\mathbf{b}} \\ &= \sum_{t=1}^T h_t (\mathbf{b}' \mathbf{x}_t)^2 - 2 \sum_{t=1}^T h_t \mathbf{b}' \mathbf{x}_t y_t + \hat{\mathbf{b}}' \mathbf{H} \hat{\mathbf{b}} \\ &= \sum_{t=1}^T h_t (y_t - \mathbf{b}' \mathbf{x}_t)^2 - \sum_{t=1}^T h_t y_t^2 + \hat{\mathbf{b}}' \mathbf{H} \hat{\mathbf{b}}. \end{aligned}$$

Result 2. (Posterior distribution of the coefficients of a normal linear regression model with heteroscedastic disturbances, given a normal prior.) Given the prior distribution $\mathbf{b} \sim \mathbf{N}(\underline{\mathbf{b}}, \underline{\mathbf{H}}^{-1})$ and the likelihood function (G1), the posterior distribution is

$$\mathbf{b} \sim \mathbf{N}(\bar{\mathbf{b}}, \bar{\mathbf{H}}^{-1})$$

with $\bar{\mathbf{H}} = \mathbf{H} + \underline{\mathbf{H}}$ and $\bar{\mathbf{b}} = \bar{\mathbf{H}}(\mathbf{H}\hat{\mathbf{b}} + \underline{\mathbf{H}}\underline{\mathbf{b}})$.

Proof.

$$\begin{aligned} (\mathbf{b} - \bar{\mathbf{b}})' \bar{\mathbf{H}}(\mathbf{b} - \bar{\mathbf{b}}) &= \mathbf{b}'(\mathbf{H} + \underline{\mathbf{H}})\mathbf{b} - 2\mathbf{b}'(\mathbf{H}\hat{\mathbf{b}} + \underline{\mathbf{H}}\underline{\mathbf{b}}) + \bar{\mathbf{b}}' \bar{\mathbf{H}} \bar{\mathbf{b}} \\ &= (\mathbf{b} - \hat{\mathbf{b}})' \mathbf{H}(\mathbf{b} - \hat{\mathbf{b}}) + (\mathbf{b} - \underline{\mathbf{b}})' \underline{\mathbf{H}}(\mathbf{b} - \underline{\mathbf{b}}) - \hat{\mathbf{b}}' \mathbf{H} \hat{\mathbf{b}} - \underline{\mathbf{b}}' \underline{\mathbf{H}} \underline{\mathbf{b}} + \bar{\mathbf{b}}' \bar{\mathbf{H}} \bar{\mathbf{b}}. \end{aligned}$$

Result 3. (Likelihood function for the variance parameter in a normal linear regression model with heteroscedastic disturbances.)

$$(\mathbf{s}^2)^{-T/2} \exp\left[-\frac{1}{2}\mathbf{s}^{-2} \sum_{t=1}^T h_t (y_t - \mathbf{b}'\mathbf{x}_t)^2\right] \propto (\mathbf{s}^2)^{-T/2} \exp[-s^2/2\mathbf{s}^2], \quad (\text{G2})$$

with $s^2 = \sum_{t=1}^T h_t (y_t - \mathbf{b}'\mathbf{x}_t)^2$.

Proof. Immediate.

Result 4. (Posterior distribution of the variance parameter in a normal linear regression model with heteroscedastic disturbances, given an inverted gamma prior)

distribution.) Given the prior $\underline{s}^2/\mathbf{s}^2 \sim \mathbf{c}^2(\underline{\mathbf{u}}_s)$ and the likelihood function (G2), the posterior distribution of \mathbf{s}^2 is given by

$$(\underline{s}^2 + \underline{\mathbf{s}}^2)/\mathbf{s}^2 \sim \mathbf{c}^2(T + \underline{\mathbf{u}}_s).$$

Proof. If $u = \underline{s}^2/\mathbf{s}^2 \sim \mathbf{c}^2(\underline{\mathbf{u}}_s)$, then

$$f(u) = 2^{-u_s/2} [\Gamma(\underline{\mathbf{u}}_s/2)]^1 u^{u_s/2-1} \exp(-u/2).$$

Making the change of variable, the prior density kernel for \mathbf{s}^2 is

$$\begin{aligned} f(\mathbf{s}^2) &= 2^{-u_s/2} [\Gamma(\underline{\mathbf{u}}_s/2)]^1 (\underline{s}^2/\mathbf{s}^2)^{u_s/2-1} \exp(-\underline{s}^2/2\mathbf{s}^2) (\underline{s}^2/\mathbf{s}^4) \\ &\propto (\underline{s}^2/\mathbf{s}^2)^{u_s/2+1} \exp(-\underline{s}^2/2\mathbf{s}^2). \end{aligned} \quad (\text{G3})$$

Hence the posterior density kernel is

$$\left[(\underline{s}^2 + \underline{\mathbf{s}}^2)/\mathbf{s}^2 \right]^{(T+\underline{\mathbf{u}}_s)/2+1} \exp\left[-(\underline{s}^2 + \underline{\mathbf{s}}^2)/2\right]$$

which via (G3) implies the result.

TABLE 1
Some Sample Properties of Earnings Data (Full sample)

Personal Characteristics			Number in Earnings Quintile					Total
Race	Education	Age	1st	2nd	3rd	4th	5th	
White	<12	25–34	516	624	461	290	111	2,002
White	<12	35–44	464	562	557	472	276	2,331
White	<12	45–54	582	677	622	536	449	2,866
White	<12	55–65	883	601	454	397	269	2,604
White	12–15	25–34	993	1,504	1,812	1,682	891	6,882
White	12–15	35–44	457	682	1065	1301	1284	4,789
White	12–15	45–54	294	518	703	1063	1052	3,630
White	12–15	55–65	374	403	481	580	545	2,383
White	≥16	25–34	332	319	513	812	843	2,819
White	≥16	35–44	95	129	230	502	1448	2,404
White	≥16	45–54	45	90	105	300	1136	1,676
White	≥16	55–65	83	71	87	211	598	1,050
Nonwhite	<12	25–34	944	524	274	94	23	1,859
Nonwhite	<12	35–44	747	577	420	183	81	2,008
Nonwhite	<12	45–54	881	655	369	201	94	2,200
Nonwhite	<12	55–65	817	405	217	96	59	1,594
Nonwhite	12–15	25–34	859	917	716	446	169	3,107
Nonwhite	12–15	35–44	156	254	314	283	157	1,164
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			9,747	9,749	9,747	9,747	9,748	48,738

(table continues)

TABLE 1, continued

Personal Characteristics			Proportion in Earnings Quintile				
Race	Education	Age	1st	2nd	3rd	4th	5th
White	<12	25–34	.258	.312	.230	.145	.055
White	<12	35–44	.199	.241	.239	.202	.118
White	<12	45–54	.203	.236	.217	.187	.157
White	<12	55–65	.339	.231	.174	.152	.103
White	12–15	25–34	.144	.219	.263	.244	.129
White	12–15	35–44	.095	.142	.222	.272	.268
White	12–15	45–54	.081	.143	.194	.293	.290
White	12–15	55–65	.157	.169	.202	.243	.229
White	≥16	25–34	.118	.113	.182	.288	.299
White	≥16	35–44	.040	.054	.096	.209	.602
White	≥16	45–54	.027	.054	.063	.179	.678
White	≥16	55–65	.079	.068	.083	.201	.570
Nonwhite	<12	25–34	.508	.282	.147	.051	.012
Nonwhite	<12	35–44	.372	.287	.209	.091	.040
Nonwhite	<12	45–54	.400	.298	.168	.091	.043
Nonwhite	<12	55–65	.513	.254	.136	.060	.037
Nonwhite	12–15	25–34	.276	.295	.230	.144	.054
Nonwhite	12–15	35–44	.134	.218	.270	.243	.135
Nonwhite	12–15	45–54	.164	.222	.253	.218	.144
Nonwhite	12–15	55–65	.282	.155	.206	.171	.187
Nonwhite	≥16	25–34	.130	.186	.295	.223	.165
Nonwhite	≥16	35–44	.009	.027	.309	.264	.391
Nonwhite	≥16	45–54	.106	.148	.148	.393	.295
Nonwhite	≥16	55–65	.300	.071	.200	.129	.300

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 (in years)	10	10	3	3
Age/100	-	11	-	4
Education x (Age/100)	-	12	-	5
(Age/100) ²	-	13	-	6
Education x (Age/100) ²	-	14	-	7
(Age/100) ³	-	15	-	-
Education x (Age/100) ³	-	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	-	-

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	
<i>Earnings period 1covariates:</i>										
Father ed missing	.050	(.200)	-.163	(.056)	-.237	(.080)	-.124	(.067)	-.175	(.075)
Father ed 12+	.050	(.100)	.014	(.023)	.025	(.036)	-.011	(.027)	.013	(.034)
Father ed 16+	.100	(.100)	.001	(.036)	-.019	(.051)	-.042	(.041)	-.062	(.051)
Mother ed missing	.050	(.100)	.059	(.065)	.081	(.096)	.133	(.105)	-.189	(.091)
Mother ed 12+	.050	(.200)	-.021	(.021)	-.006	(.032)	-.030	(.028)	-.011	(.030)
Mother ed 16+	.100	(.100)	-.032	(.046)	-.029	(.061)	.023	(.052)	.003	(.058)
Nonwhite indicator	-.100	(.100)	-.191	(.024)	-.261	(.035)	-.195	(.030)	-.203	(.034)
Marital status current	.200	(.200)	.072	(.022)	.083	(.029)	-.006	(.048)	.092	(.028)
Intercept	7.22	(4.00)	7.85	(.079)	7.86	(.116)	8.09	(.158)	6.88	(.215)
Education	.065	(.075)	.050	(.006)	.045	(.009)	.032	(.011)	.123	(.017)
<i>Earnings period tcovariates:</i>										
Father ed missing	.050	(.200)	-.196	(.050)	-.139	(.065)	-.040	(.100)	-.079	(.040)
Father ed 12+	.050	(.100)	-.011	(.021)	.045	(.028)	.085	(.042)	.069	(.020)
Father ed 16+	.100	(.100)	.016	(.033)	.010	(.042)	-.060	(.076)	-.013	(.033)
Mother ed missing	.050	(.100)	.093	(.067)	.122	(.093)	-.103	(.066)	-.052	(.026)
Mother ed 12+	.050	(.200)	-.006	(.018)	-.009	(.025)	-.019	(.040)	-.017	(.017)
Mother ed 16+	.100	(.100)	-.086	(.042)	-.046	(.050)	.005	(.089)	.024	(.036)
Nonwhite indicator	-.100	(.100)	-.208	(.021)	-.268	(.026)	-.164	(.051)	-.267	(.019)
Marital status current	.200	(.200)	.036	(.008)	.081	(.013)	.050	(.009)	.100	(.009)
Earnings age 25, Ed 12	8.00	(4.00)	8.54	(.024)	8.41	(.030)	8.51	(.037)	8.41	(.023)
Earnings age 35 vs. 25, Ed=12	.150	(.100)	.236	(.117)	.245	(.024)	.231	(.029)	.261	(.019)
Earnings age 45 vs. 35, Ed=12	.100	(.100)	.109	(.023)	.125	(.034)	.081	(.008)	.155	(.010)
Earnings age 55 vs. 45, Ed=12	.050	(.100)	.157	(.116)	.099	(.095)	-.043	(.008)	-.113	(.009)
Earningsed 16 vs. 12, Age=25	.260	(.150)	.195	(.026)	.173	(.038)	.294	(.034)	.287	(.041)
Earningsed 16 vs. 12, Age=35	.340	(.200)	.341	(.021)	.450	(.029)	.469	(.038)	.362	(.016)
Earningsed 16 vs. 12, Age=45	.370	(.225)	.374	(.033)	.389	(.055)	.483	(.040)	.352	(.013)
Earningsed 16 vs. 12, Age=55	.400	(.250)	.284	(.197)	.400	(.230)	.447	(.037)	.318	(.014)

(table continues)

TABLE 3, continued

	Young Men					Full Sample				
	Prior		Mixed Model	Normal Model		Mixed Model	Normal Model			
<i>Properties of first period shock:^a</i>										
Mean 1	-3.00	(1.00)	-1.966	(.233)	—	-2.625	(.271)	—		
Mean 2	0.00	(0.00)	0.00	(0.00)	—	0.00	(0.00)	—		
Mean 3	.100	(.100)	.147	(.029)	—	.199	(.004)	—		
Standard deviation 1	1.425	(.102)	1.345	(.074)	—	1.525	(.114)	—		
Standard deviation 2	1.127	(.081)	.791	(.035)	.774	(.013)	.776	(.033)	.809	(.016)
Standard deviation 3	.319	(.022)	.329	(.013)	—	.321	(.013)	—		
Probability 1	.020	(.014)	.058	(.011)	—	.046	(.007)	—		
Probability 2	.480	(.050)	.337	(.033)	—	.385	(.029)	—		
Probability 3	.500	(.050)	.606	(.031)	—	.568	(.028)	—		
P[<log(.2)]	.052	(.012)	.041	(.004)	.019	(.002)	.041	(.003)	.023	(.002)
P[<log(.5)]	.149	(.016)	.111	(.005)	.185	(.004)	.113	(.004)	.195	(.005)
P[<log(.8)]	.296	(.023)	.250	(.006)	.387	(.002)	.242	(.006)	.391	(.002)
P[<log(.9)]	.367	(.029)	.321	(.009)	.446	(.001)	.309	(.008)	.448	(.001)
P[<0]	.442	(.035)	.401	(.010)	.500	(.000)	.383	(.009)	.500	(.000)
P[>log(1.111)]	.476	(.038)	.508	(.012)	.446	(.001)	.529	(.010)	.448	(.001)
P[>log(1.25)]	.385	(.039)	.402	(.012)	.387	(.002)	.424	(.010)	.391	(.002)
P[>log(2)]	.149	(.020)	.103	(.006)	.185	(.004)	.110	(.006)	.195	(.005)
P[>log(5)]	.038	(.009)	.008	(.001)	.019	(.002)	.008	(.001)	.023	(.002)
<i>Properties of tth period shock:^a</i>										
Mean 1	-3.00	(1.00)	-.955	(.088)	—	-.899	(.043)	—		
Mean 2	0.00	(0.00)	0.00	(0.00)	—	0.00	(0.00)	—		
Mean 3	.100	(.100)	.064	(.046)	—	.066	(.014)	—		
Standard deviation 1	1.425	(.102)	1.313	(.058)	—	1.284	(.029)	—		
Standard deviation 2	1.127	(.081)	.574	(.015)	.466	(.003)	.462	(.008)	.448	(.002)
Standard deviation 3	.319	(.022)	.146	(.003)	—	.117	(.001)	—		

(table continues)

^a Prior moments shown are for the mixed normal model, not the normal model.

TABLE 3, continued

	Young Men					Full Sample				
	Prior		Mixed Model		Normal Model		Mixed Model		Normal Model	
<i>Properties of t' th period shock, continued:</i>										
Probability 1	.020	(.014)	.044	(.006)	—		.051	(.003)	—	
Probability 2	.480	(.050)	.272	(.009)	—		.315	(.005)	—	
Probability 3	.500	(.050)	.684	(.009)	—		.634	(.007)	—	
P[<log(.2)]	.052	(.012)	.041	(.004)	.019	(.002)	.041	(.003)	.023	(.002)
P[<log(.5)]	.149	(.016)	.111	(.005)	.185	(.004)	.113	(.004)	.195	(.005)
P[<log(.8)]	.296	(.023)	.144	(.002)	.316	(.001)	.138	(.002)	.309	(.001)
P[<log(.9)]	.367	(.029)	.236	(.006)	.410	(.001)	.208	(.002)	.407	(<.001)
P[<0]	.442	(.035)	.400	(.013)	.500	(.000)	.369	(.005)	.500	(.000)
P[>log(1.111)]	.476	(.038)	.386	(.013)	.410	(.001)	.384	(.006)	.407	(<.001)
P[>log(1.25)]	.385	(.039)	.806	(.006)	.316	(.001)	.171	(.003)	.309	(.001)
P[>log(2)]	.149	(.020)	.035	(.003)	.068	(.001)	.027	(.001)	.061	(.001)
P[>log(5)]	.038	(.009)	.002	(.0003)	<.001	(<.001)	.001	(.0001)	<.001	(<.001)
<i>Other earnings model parameters:</i>										
g (lagged earnings)	.500	(.500)	-.090	(.015)	-.201	(.014)	-.121	(.007)	-.213	(.008)
r (autocorrelation period 2)	.500	(.500)	.375	(.032)	.529	(.038)	.344	(.027)	.398	(.032)
r (autocorrelation period t)	.500	(.500)	.652	(.029)	.737	(.016)	.655	(.008)	.739	(.100)
f (first period perm. effect)	.400	(.500)	.209	(.017)	.225	(.025)	.240	(.018)	.320	(.028)
s_t (s.d. individual shock)	^b	^b	.260	(.012)	.215	(.017)	.366	(.016)	.302	(.017)
<i>Variances and decompositions:</i>										
Disturbance variance, age 25	.914	(.236)	.614	(.045)	.599	(.020)	.738	(.057)	.655	(.027)
Disturbance variance, age 30	^b	^b	.455	(.021)	.473	(.013)	.505	(.015)	.528	(.010)
Disturbance variance, age 45	^b	^b	.442	(.020)	.445	(.011)	.488	(.014)	.498	(.007)
Disturbance variance, age 60	^b	^b	.442	(.020)	.445	(.011)	.488	(.014)	.497	(.008)

(table continues)

^b Prior moments that do not exist.

TABLE 3, continued

	Young Men					Full Sample				
	Prior		Mixed Model		Normal Model		Mixed Model		Normal Model	
<i>Variances and decompositions, continued:</i>										
Fraction var. transitory, age 30	.461	(.290)	.752	(.019)	.748	(.019)	.608	(.017)	.615	(.013)
Fraction var. transitory, age 45	.478	(.299)	.785	(.024)	.826	(.025)	.638	(.017)	.679	(.017)
Fraction var. transitory, age 60	.476	(.299)	.785	(.024)	.826	(.025)	.638	(.017)	.680	(.017)
Correlation, ages 25 and 30	.186	(.271)	.313	(.019)	.392	(.018)	.354	(.020)	.459	(.023)
Correlation, ages 30 and 45	.526	(.305)	.230	(.021)	.212	(.023)	.376	(.017)	.355	(.015)
Correlation, ages 45 and 60	.522	(.316)	.216	(.023)	.182	(.023)	.363	(.017)	.326	(.016)
Correlation, ages 25 and 45	.186	(.257)	.246	(.024)	.262	(.028)	.295	(.022)	.368	(.030)
Correlation, ages 30 and 60	.534	(.291)	.229	(.022)	.204	(.024)	.375	(.017)	.348	(.016)
Correlation, ages 25 and 60	.183	(.258)	.246	(.024)	.261	(.029)	.295	(.022)	.367	(.031)
<i>Marital status period I covariates:</i>										
Nonwhite indicator	.000	(.255)	-.442	(.119)	-.444	(.123)	-.593	(.116)	-.606	(.118)
Intercept	.000	(.255)	2.046	(.390)	2.049	(.390)	4.567	(.335)	4.549	(.328)
Education	.000	(.128)	-.075	(.030)	-.074	(.030)	-.232	(.025)	-.226	(.027)
<i>Marital status period tcovariates:</i>										
Nonwhite indicator	.000	(.255)	-.771	(.110)	-.772	(.113)	-.961	(.092)	-1.015	(.023)
Probit age 25, Ed 12	.000	(.255)	.028	(.204)	.028	(.208)	.078	(.219)	.138	(.193)
Probit age 40 vs. 25, Ed=12	.000	(.255)	.545	(.124)	.554	(.133)	.898	(.116)	.906	(.073)
Probit age 55 vs. 40, Ed=12	.000	(.255)	-.101	(.236)	-.109	(.240)	.294	(.068)	.293	(.071)
Probit ed 16 vs. 12, Age=25	.000	(.255)	-.207	(.133)	-.207	(.132)	-.662	(.103)	-.667	(.085)
Probit ed 16 vs. 12, Age=40	.000	(.255)	.210	(.145)	.218	(.144)	-.258	(.076)	-.283	(.060)
Probit ed 16 vs. 12, Age=55	.000	(.255)	-.066	(.249)	-.072	(.250)	-.092	(.064)	-.099	(.075)
Lagged marital status	.680	(.680)	.546	(.056)	.536	(.055)	.668	(.052)	.645	(.040)
Lagged earnings	.000	(.510)	.125	(.025)	.127	(.025)	.177	(.023)	.182	(.227)
<i>Other marital status model parameters:</i>										
I (autocorrelation)	.700	(.700)	.907	(.006)	.909	(.006)	.926	(.008)	.931	(.002)

TABLE 4
Comparison of Model Fit: Probabilities of Income Quintile Sequences

Sequence	White			Black		
	PSID	Mixture	Normal	PSID	Mixture	Normal
<i>Sample of Young Men:</i>						
-	.152	.179	.255	.342	.327	.469
+	.848	.821	.745	.658	.673	.531
- -	.089	.110	.157	.240	.232	.346
- +	.058	.071	.101	.091	.097	.124
+ -	.053	.062	.084	.087	.083	.111
+ +	.800	.758	.659	.582	.588	.420
- - -	.058	.075	.109	.181	.181	.278
- - +	.028	.036	.049	.050	.054	.069
- + -	.014	.010	.028	.031	.019	.046
- + +	.043	.061	.074	.060	.077	.078
+ - -	.025	.033	.039	.046	.049	.061
+ - +	.026	.029	.045	.040	.033	.049
+ + -	.037	.047	.054	.054	.056	.063
+ + +	.769	.708	.601	.538	.532	.355
<i>Full Sample:</i>						
-	.144	.206	.249	.348	.422	.491
+	.856	.821	.751	.652	.673	.509
- -	.088	.144	.162	.248	.349	.381
- +	.044	.060	.083	.084	.076	.110
+ -	.051	.057	.081	.089	.072	.105
+ +	.817	.739	.674	.578	.504	.404
- - -	.061	.112	.118	.196	.306	.318
- - +	.020	.031	.040	.042	.045	.064
- + -	.012	.011	.026	.033	.019	.043
- + +	.031	.048	.057	.051	.057	.067
+ - -	.023	.031	.040	.044	.045	.061
+ - +	.022	.026	.041	.038	.027	.044
+ + -	.039	.044	.054	.056	.049	.061
+ + +	.792	.696	.625	.539	.452	.343

Notes: The income 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 he is 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 are observed for the person in the relevant PSID sample (young men or full).

TABLE 5.1
 Simulations of Income Mobility Based on Young Men Sample:
 Probabilities of Income 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>						
- - -	.023	.007	.002	.053	.022	.003
- - +	.061	.029	.007	.116	.064	.013
- + -	.023	.016	.011	.040	.030	.017
- + +	.155	.122	.084	.177	.161	.120
+ - -	.017	.007	.003	.018	.012	.010
+ - +	.096	.070	.039	.118	.096	.053
+ + -	.052	.047	.066	.053	.060	.073
+ + +	.573	.701	.787	.424	.554	.710
<i>Normal Distribution for Transitory Error Component:</i>						
- - -	.065	.014	.002	.143	.052	.007
- - +	.079	.049	.012	.135	.108	.047
- + -	.071	.029	.008	.104	.065	.027
- + +	.178	.162	.107	.170	.196	.153
+ - -	.043	.013	.003	.051	.034	.013
+ - +	.106	.084	.045	.102	.106	.050
+ + -	.091	.055	.039	.087	.066	.060
+ + +	.367	.594	.784	.208	.373	.643

Notes: Based on simulations of income from age 25 through 65. The sequence refers to status of being in the bottom income 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.

TABLE 5.2
 Simulations of Income Mobility Based on Full Sample:
 Probabilities of Income 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>						
- - -	.201	.037	.006	.334	.068	.019
- - +	.058	.024	.008	.064	.036	.017
- + -	.078	.034	.010	.083	.057	.019
- + +	.097	.102	.068	.083	.120	.058
+ - -	.049	.017	.004	.057	.024	.007
+ - +	.071	.058	.035	.059	.064	.045
+ + -	.088	.086	.046	.081	.097	.062
+ + +	.358	.643	.824	.238	.534	.772
<i>Normal Distribution for Transitory Error Component:</i>						
- - -	.150	.041	.010	.296	.117	.029
- - +	.057	.032	.012	.070	.054	.028
- + -	.134	.080	.033	.165	.130	.104
- + +	.121	.117	.087	.093	.125	.112
+ - -	.058	.031	.011	.078	.056	.023
+ - +	.058	.048	.031	.047	.052	.043
+ + -	.129	.140	.097	.109	.142	.119
+ + +	.293	.512	.720	.142	.325	.542

Notes: Based on simulations of income from age 25 through 65. The sequence refers to status of being in the bottom income 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.

TABLE 6
 Simulations of Income 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	.264	.179	.121	.364	.271	.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	.422	.254	.128	.571	.422	.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	.476	.221	.107	.609	.307	.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 Sample:</i>						
Fraction	.519	.315	.170	.681	.477	.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

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

TABLE 7
 Simulations of Income Mobility:
 Expected Fraction of Remaining Lifetime in Bottom Quintile
 Conditional on Income 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>						
-	.398	.288	.213	.471	.391	.294
+	.190	.134	.095	.249	.190	.138
<i>Normal Distribution, Young Men Sample:</i>						
-	.525	.355	.197	.633	.497	.291
+	.320	.181	.083	.446	.315	.148
<i>Mixture Distribution, Full Sample:</i>						
-	.705	.437	.250	.796	.524	.387
+	.292	.152	.079	.357	.198	.104
<i>Normal Distribution, Full Sample:</i>						
-	.656	.492	.351	.772	.616	.476
+	.364	.225	.123	.491	.339	.192

Notes: Based on simulations of income 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 8.1
 Simulations of Income Mobility Based on Young Men Sample:
 Probabilities of Bottom Quintile at Age 35
 Conditional on Income 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>						
? -	.646	.588	.409	.722	.649	.571
? +	.113	.073	.042	.147	.095	.060
- -	.698	.641	.422	.763	.698	.667
- +	.239	.151	.051	.234	.187	.050
+ -	.521	.498	.398	.599	.522	.429
+ +	.098	.066	.041	.132	.085	.061
<i>Normal Distribution for Transitory Error Component:</i>						
? -	.707	.580	.409	.747	.684	.477
? +	.166	.113	.046	.278	.170	.082
- -	.770	.647	.506	.796	.748	.611
- +	.293	.269	.200	.429	.354	.208
+ -	.566	.478	.327	.578	.539	.385
+ +	.133	.091	.038	.219	.126	.069

Notes: Based on simulations of income 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 8.2
 Simulations of Income Mobility Based on Full Sample:
 Probabilities of Bottom Quintile at Age 35
 Conditional on Income 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>						
? -	.812	.639	.469	.872	.707	.552
? +	.132	.067	.033	.159	.081	.029
- -	.854	.702	.547	.906	.765	.639
- +	.280	.160	.103	.307	.170	.120
+ -	.623	.507	.385	.646	.546	.400
+ +	.110	.060	.030	.129	.072	.025
<i>Normal Distribution for Transitory Error Component</i>						
? -	.734	.608	.503	.820	.697	.541
? +	.175	.098	.050	.246	.162	.086
- -	.792	.689	.615	.862	.753	.569
- +	.367	.274	.224	.400	.366	.200
+ -	.557	.461	.378	.618	.548	.455
+ +	.129	.075	.038	.190	.114	.071

Notes: Based on simulations of income 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 9.1

Present Value of Lifetime Wages at Age 25 Based on Young Men Sample

	Mean	Stan Dev	Q(.05)	Q(.25)	Median	Q(.75)	Q(.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

Notes: Values are in thousands of 1967 dollars. Annual discount factor is .95. “Q(p)” is the p’th quantile.

TABLE 9.2
Present Value of Lifetime Wages at Age 25 Based on Full Sample

	Mean	Stan Dev	Q(.05)	Q(.25)	Median	Q(.75)	Q(.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

Notes: Values are in thousands of 1967 dollars. Annual discount factor is .95. “Q(p)” is the p’th quantile.

TABLE 10.1
Present Value of Lifetime Wages, Ages 31-65, Based on Full Sample:
Normal Distribution for Transitory Error Component

	Mean	Stan Dev	Q(.05)	Q(.25)	Median	Q(.75)	Q(.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

Notes: Values are in thousands of 1967 dollars. Annual discount factor is .95. "Q(p)" is the p'th quantile.

TABLE 10.2
Present Value of Lifetime Wages, Ages 31-65 Based on Full Sample:
Mixture Distribution for Transitory Error Component

	Mean	Stan Dev	Q(.05)	Q(.25)	Median	Q(.75)	Q(.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

Notes: Values are in thousands of 1967 dollars. Annual discount factor is .95. "Q(p)" is the p'th quantile.

TABLE 11
 Present Value of Lifetime Wages at Age 25:
 The Regression on Race, Education, Parents Education
 and Unobserved Individual Effects t_i and e_{it}

	Sample of Young Men				Full Sample			
	Mixture Model		Normal Model		Mixture Model		Normal Model	
Education	11.282	(.237)	15.248	(.255)	12.985	(.108)	102.78	(.113)
Black	-28.497	(1.031)	-35.305	(1.111)	-19.214	(.720)	-30.724	(.752)
<i>Father's Ed.:</i>								
Missing	-22.454	(2.384)	-14.910	(2.566)	-4.472	(1.593)	-8.458	(1.663)
HS	-.477	(1.057)	7.408	(1.138)	13.607	(.834)	11.209	(.871)
College	1.461	(1.713)	5.678	(1.844)	-1.233	(1.414)	4.120	(1.476)
<i>Mother's Ed.:</i>								
Missing	10.372	(2.881)	10.110	(3.102)	-8.989	(.975)	-6.534	(1.018)
HS	-2.149	(.943)	-2.808	(1.015)	-2.682	(.714)	-2.705	(.745)
College	-12.138	(2.190)	-5.668	(2.358)	5.756	(1.589)	7.780	(1.660)
t_i	139.224	(1.137)	134.483	(1.573)	126.337	(.561)	124.374	(.690)
e_{it}	34.301	(.383)	45.146	(.436)	30.954	(.245)	52.113	(.259)
Constant	17.224	(2.993)	-37.521	(3.222)	-12.586	(1.361)	20.548	(1.421)
-Regressors	.2115		.3019		.3367		.2630	
- t_i , e_{it}	.4492		.3565		.3866		.4457	
Unexplained	.3393		.3416		.2767		.2913	
<i>Dependent Variable:</i>								
Mean	146.501		143.139		131.586		130.066	
S.D.	66.810		76.435		85.413		84.607	

Notes: The dependent variable is the present value of lifetime wages from a simulated lifetime wage path, using a discount factor of .95 from ages 25 to 65. Wages 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.

Figures 1 through 14.2 were not available for the online web version of this paper. The figures occupied pages 76-89.

References

- Geweke, John. 1991. "Efficient Simulation from the Multivariate Normal and Student-*t* Distributions Subject to Linear Constraints." In E.M. Keramidas, ed., *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, 571–578.
- Geweke, John. 1996. "Posterior Simulators in Econometrics." In D. Kreps and K.F Wallis (eds.), *Advances in Economics and Econometrics: Theory and Applications*. Cambridge: Cambridge University Press, forthcoming. (Invited symposium paper, Econometric Society Seventh World Congress) Also Federal Reserve Bank of Minneapolis, Working Paper No. 555, September 1995.
- Gilks, Walter R., and Pascal Wild. 1992. "Adaptive Rejection Sampling for Gibbs Sampling." *Applied Statistics (JRSS Series C)* **41**:337–348.
- Gottschalk, Peter. 1982. "Earnings Mobility: Permanent Change or Transitory Fluctuations?" *Review of Economics and Statistics* **64**:450–456.
- Gottschalk, Peter and Robert Moffit. 1994. "Trends in the Autocovariance Structure of Earnings in the U.S.: 1969–1987." Brown University, Department of Economics, working paper.
- Horowitz, Joel L. and Marianthi Markatou. 1993. "Semiparametric Estimation of Regression Models for Panel Data." Working Paper #93-14, Department of Economics, University of Iowa.
- Lillard, Lee, and Robert Willis. 1978. "Dynamic Aspects of Earnings Mobility." *Econometrica* **46**:985–1012.
- MaCurdy, Thomas. 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, John J. 1973. *Income Mobility, Racial Discrimination, and Economic Growth*. Lexington, Mass.: Lexington Books.
- Shorrocks, Anthony. 1976. "Income Mobility and the Markov Assumption." *Economic Journal* **86**:566–578.
- Tierney, Luke. 1994. "Markov Chains for Exploring Posterior Distributions" (with discussion and rejoinder) *Annals of Statistics* **22**:1701–1762.
- West, Mike and Jeff Harrison. 1989. *Bayesian Forecasting and Dynamic Models*. New York: Springer-Verlag.