

Using Semiparametric Methods in an Analysis of Earnings Mobility

Shawn W. Ulrick¹

U.S. Federal Trade Commission

600 Pennsylvania Ave., NW

Washington, DC 20580

sulrick@ftc.gov

January 24, 2007

Abstract

This paper describes a dynamic random effects econometric model from which inferences on earnings mobility may be made. It answers questions such as, given some initial level of observed earnings, what is the probability that an agent with certain characteristics will remain below a specified level of earnings (for example the poverty level) for a specified number of time periods? Existing research assumes that the distributions of the unobserved permanent and transitory shocks in the model are known up to finitely many parameters. However, predictions of earnings mobility are highly sensitive to assumptions about these distributions. The present paper estimates the distributions of the random effects nonparametrically. The results are used to predict the probabilities of remaining in a low state of earnings. The results from the nonparametric distributions are contrasted to those obtained under a normality assumption. Using the nonparametrically estimated distributions gives estimated probabilities that are smaller than those obtained under the normality assumption. Through a Monte Carlo experiment and by examining unconditional predicted earnings distributions, it is demonstrated that the nonparametric method is likely to be considerably more accurate, and that assuming normality may give quite misleading results.

¹ I gratefully acknowledge the support of Joel Horowitz. I also wish to thank Brendand Cunningham, John Geweke, George Deltas, and David Schmidt for their helpful suggestions. I also wish to thank two anonymous referees. This research was supported in part by NSF grant SBR-9617925. The views in this article are those of the author and do not necessarily reflect those of the Federal Trade Commission or any individual Commissioner.

I. Introduction

It is well documented that the wage gap between high and low earners has been growing in past years, and anyone living in America today witnesses the regularity with which politicians and the media reference this phenomenon. However, as pointed out by Moffitt and Gottschalk (1998), this phenomenon can be caused by changing trends in either permanent or temporary fluctuations in income. The dispersion of wages is less troubling if low earners are only temporarily poor, or, said another way, “if individuals are able to climb up the earnings ladder, then changes in the dispersion of annual earnings are less informative” (Daly and Valetta 2003). Thus, an accurate analysis of earnings mobility can be useful in putting into perspective the implications of the growing wage gap between the “rich and poor.” To this end, several papers have examined the extent of earnings mobility. These papers generally make predictions of mobility based on a regression model and some rigid assumption about the distributions of permanent and temporary fluctuations in earnings (*i.e.*, usually normality). The purpose of this paper is to use a recent semiparametric technique to relax the rigid assumptions about the distributions. As will be seen, relaxing these assumptions can have serious consequences on the estimated mobility.

This and many past papers examining mobility use the basic framework of Lillard and Willis’s (L&W) 1978 paper, in which expected earnings are estimated via a regression model, and the deviation from expected earnings is estimated using the distributions of temporary and permanent fluctuations in workers’ earnings.² Specifically, L&W use panel data to estimate a linear, random effects model of earnings. L&W regress logged annual earnings on several variables including education, experience in the labor force, race, and time effects. The residual structure includes a random, permanent component and a random, serially correlated transitory component.

² Others, such as Geweke and Keane (2000), and Datcher–Laoury (1986), have implemented similar studies using the L&W techniques.

Together, these components represent unmeasured variables affecting earnings. The permanent component allows the model to capture influences on earnings that are specific to an individual. The serially correlated transitory component allows the model to capture effects of time persistent shocks. With their estimated model, the assumed structure of the random effects, and a further assumption about the distribution of the random effects, L&W are able to estimate the probability that an individual's earnings will fall into a specific income bracket continuously for some period of time. L&W and most past authors assumed that the distributions of the random error components are normal. This is a serious assumption which often proves dubious (see, *e.g.*, Geweke and Keane 2000 (G&K), Horowitz and Markatou 1996 (H&M), White and MacDonald 1980). If the assumption is incorrect, the predicted mobilities may be wrong. Using the wrong distribution in such an analysis can be compared to a quality control expert assuming the lifespan of an electronics component is normally distributed when in fact it follows the exponential.

The assumption about the distributions of the errors is so central to the predictions, it seems wrong not to evaluate the correctness of it and whether it matters in application. To this end, the present paper uses a recently developed semiparametric method to estimate the individual distributions of the error components and uses these estimates in predicting mobility. Although nonparametric methods for estimating distribution functions have been around for years (see, *e.g.*, Silverman 1985), most past authors have not had the luxury of separately estimating the distributions of the permanent and transitory error components via non- or semiparametric techniques due to the fact that the errors are only observed when convoluted. The semiparametric method used in this paper (developed in H&M) uses the multiplicative relationship between characteristics functions of convoluted random variables to deconvolute the permanent and transitory components.³

³ G&K use a mixture of normals, which shows a substantial improvement over the basic normality assumption. Robin and Bonhomme (2004) provide another approach in dealing with the unobservable permanent and transitory components, using a statistical technique based on copulas.

This paper contrasts the resulting transition probabilities to those obtained under a normality assumption. As will be seen, the differences in transition probabilities found by using the nonparametric method as opposed to assuming normality are, at times, substantial. Specifically, assuming normality gives probabilities of upward mobility that are, in some cases, twice as large as those obtained by using the nonparametrically estimated distributions. The result of a Monte Carlo experiment shows that the nonparametric method is likely to be considerably more accurate than is assuming normality. Further evidence that the nonparametric densities provide more accurate results is that the nonparametric densities do a much better job at predicting the unconditional distribution of earnings.

The organization of the rest of this paper is as follows. Section II discusses the data used in the analysis. Section III presents the earnings model to be used and the details of its estimation. This section also discusses estimation of the transition probabilities. Section IV presents the estimation results, including the estimates of the parameters of the earnings model, the densities, and the transition probabilities. This section also provides some informal tests suggesting that using the estimated densities rather than assuming normality provides a better fitting model. Section V presents the results of some Monte Carlo experiments aimed at investigating the properties of the estimators used. Concluding remarks appear in Section VI.

II. Data

This analysis is based on the same subset of the Panel Study of Income Dynamics (PSID) that G&K used. G&K used 22 continuous years of the PSID from 1967 onward. They applied several screens to the data. First, they considered only men between the ages of 25 and 65 who are household heads. Second, they removed individuals for whom race and education data are missing. Third, they dropped the first observation for each household head, since often the first year's observation is only for a partial year's income. Finally, if a person has a missing observation of earnings or marital status for a

given year, G&K dropped all observations for that person from that point onward. This leaves an unbalanced panel of 4766 male household heads in their prime earning years.

Means and standard deviations of the data appear in Table 1.

III. Statistical model

This section introduces the structure of the earnings model, develops a method of estimation, and discusses estimating the earnings mobility probability.

III (a). Earnings model

The earnings model for individual i at time t is:

$$y_{it} = \alpha y_{i,t-1} + X_{it}\beta + V_i\gamma + \eta_i + \varepsilon_{it} \quad (1)$$

with

$$\varepsilon_{it} = \rho\varepsilon_{i,t-1} + \xi_{it}. \quad (2)$$

The variable y_{it} denotes log annual earnings, X_{it} is a vector of seven time varying independent variables, and V_i is a vector of eight time invariant independent variables, including an intercept. A complete listing of these variables may be found in Table 1. Let there be $T_i > 3$ observations on individual i . Let there be N individuals. The error components consist of η_i , an iid individual effect; ε_{it} , a serially correlated transitory component; and ξ_{it} , an iid error. The individual effect represents unmeasured time invariant variables that affect earnings and are unique to an individual. The serially correlated transitory effect captures unmeasured, serially correlated variables and time persistent shocks. This model was also used by G&K. Here, however, in contrast to

G&K, the distributions of η , ξ , and ε are not assumed to belong to known parametric families. Instead, their distributions are estimated nonparametrically.

III (b). Estimation

Hsiao (1986) and Chamberlain (1983) summarize methods for estimation of panel data models, including those with a lagged dependent variable (LDV) or a serially correlated transitory component. Blundell and Bond (1997) and Arellano and Bond (1991) discuss improved, more efficient GMM methods for estimating a model with a lagged dependent variable. However, there has been relatively little research on estimation of models like (1 – 2), which have both an LDV and a serially correlated transitory effect. OLS would not provide a consistent estimate of (1 – 2), due to the correlation of the LDV with the individual effect and the AR(1) transitory effect. Instead, I use an instrumental variables approach.

To deal with the correlation between the LDV and the individual effect, remove the latter by subtracting the group mean, that is, the mean of individual i 's observations. This yields

$$y_{it} - \bar{y}_i = \alpha(y_{i,t-1} - \bar{y}_i) + (X_{it} - \bar{X}_i)\beta + \varepsilon_{it} - \bar{\varepsilon}_i, \quad (3)$$

where

$$\bar{y}_i = \frac{1}{T_i} \sum_{t=1}^{T_i} y_{it}; \quad \bar{X}_i = \frac{1}{T_i} \sum_{t=1}^{T_i} X_{it}; \quad \bar{\varepsilon}_i = \frac{1}{T_i} \sum_{t=1}^{T_i} \varepsilon_{it}.$$

(Note that $\bar{\varepsilon}_i$ is the unobservable within-group mean of the serially correlated error.) Equation (3) does not contain the time invariant variables, so γ cannot be estimated from (3). Methods for estimating γ are discussed at the end of this subsection. Note that the LDV term $(y_{i,t-1} - \bar{y}_i)$ is still correlated with the transitory effect, but an instrumental

variables method will estimate (3) consistently. I use $(X_{i,t-1} - \bar{X}_{i,-1})$ as instruments for $(y_{i,t-1} - \bar{y}_i)$, where

$$\bar{X}_{i,-1} = \frac{1}{T_i-1} \sum_{t=1}^{T_i-1} X_{it}.$$

The structure of (1) suggests that $(X_{i,t-1} - \bar{X}_{i,-1})$ is correlated with $(y_{i,t-1} - \bar{y}_i)$ but uncorrelated with the error.

To estimate γ substitute estimates of α and β into

$$\bar{y}_i - \alpha \bar{y}_{i,-1} - \bar{X}_i \beta = V_i \gamma + \eta_i + \bar{\varepsilon}_i,$$

and apply OLS (Hsiao, 1986, pp. 50 - 51). To estimate ρ , substitute the estimated residuals $(\varepsilon_{it} - \bar{\varepsilon}_i)$ and $(\varepsilon_{i,t-1} - \bar{\varepsilon}_i)$ (obtainable from the estimated equation 3) into

$$(1/N) \sum_i (1/T_i) \sum_t (\varepsilon_{it} - \bar{\varepsilon}_i) (\varepsilon_{i,t-1} - \bar{\varepsilon}_i) / (\varepsilon_{i,t-1} - \bar{\varepsilon}_i)^2.^4$$

III (c). Earnings mobility probability

This subsection discusses a method to estimate the probability that an individual will remain below a specified level of log earnings, y^* , for θ time periods, given that he initially had log earnings less than or equal to a specified value, y_{i0} . This probability may be written

$$p(Y_{i1} < y^*, \dots, Y_{i\theta} < y^* | Y_{i0} \leq y_{i0}). \quad (4)$$

⁴ This is the same procedure as in Hsiao (1986), p. 55. However, because of the lagged dependent variable, for Hsiao's Step 2, I used the aforementioned instrumental variables estimator instead of OLS.

The expression for (4) is analytically very complicated and involves several iterated integrals. The integrals are evaluated by simulation. For several thousand individuals with initial log earnings less than or equal to y_{i0} and specified characteristics (*i.e.*, specified values of X_{i0} , V_i), generate earnings paths out to θ time periods according to an estimated version of (1 – 2). Calculate the proportion of those whose simulated log earnings lie below y^* for all θ periods. This number is an estimate of the value of (4).

To generate the earnings paths, one needs estimates of the parameters of the model (1 – 2), a method of sampling the initial conditions y_{i0} and ε_{i0} , and the appropriate values of X_{it} for $t > 1$. (V_i does not change; ε_{it} is known for $t \geq 1$, given ρ , ε_{i0} , and ξ_{it} .) Finally, to sample the initial conditions and subsequent shocks, one needs estimates of the distributions of η_i and ξ_{it} . (The distribution of ε_{it} is not needed in this exercise, since, as written in Appendix 1, ε_{it} may be written as a function of anterior ξ_{it} 's.)

The method to estimate the parameters was described in subsection III (b) of this paper. Details on generating the initial conditions are discussed in Appendix 1. As shown, this may be done using the estimated density functions of ξ_{it} and η_i , though care must be taken, since the error components are correlated with each other and with y_0 . To calculate X_{it} for $t > 1$, note that with the exception of marriage, every element of X_{it} is a function of only age and education level. Education level is time invariant. If age is known at time t , age is known at every time period. If one makes an assumption about marriage (I assume that it is unchanging), all the elements of X_t are known at every t .

Lastly, the distributions of the error components are estimated nonparametrically. Standard kernel density estimators (such as those in, *e.g.*, Silverman 1985) are not applicable to this model, because we never observe estimates of η_i , ε_{it} , or ξ_{it} . Instead we only observe estimates of the errors convoluted together in some form. For example, we can observe estimates of

$$\hat{\varepsilon} + \hat{\eta} = y_{it} - [\hat{\alpha}y_{i,t-1} + X_{it}\hat{\beta} + V_i\hat{\gamma}]; \quad (5)$$

$$\begin{aligned} \hat{\xi}_{it} - \hat{\xi}_{i,t-1} = & \hspace{15em} (6) \\ (\Delta y_{it} - \hat{\rho} \Delta y_{i,t-1}) - (\Delta y_{i,t-1} - \hat{\rho} \Delta y_{i,t-2}) \hat{\alpha} - (\Delta X_{it} - \hat{\rho} \Delta X_{i,t-1}) \hat{\beta} - (1 - \hat{\rho}) V_i \hat{\gamma}, \end{aligned}$$

where the “hat” notation represents an estimated value.

I hence use a nonparametric deconvolution estimator developed in Horowitz (1997) and H&M. In essence, H&M use the fact that the characteristic function for the sum of two independent random variables is the product of the characteristic functions of the individual variables. The authors use this relationship to write the characteristic functions of η , ε , and ξ in terms of the characteristic functions of the convoluted residuals whose estimates are observable. The characteristic functions can then be estimated from the residuals and smoothed with a kernel function, and the inversion formula can be applied to the smoothed estimates to obtain estimates of the densities.⁵ The basic estimator presented in H&M assumes that η and ε are iid. The authors present extensions that allow for asymmetry or serial correlation in ε . Horowitz (1997, pp. 125 - 127) combines the two extensions to allow for both asymmetry and serial correlation in ε . I allow for asymmetry and serial correlation and hence implement this version of the estimator.⁶ There has been little research in the best way to choose the tuning parameters of the estimator. See Appendix 2 for details on the choices made in this paper.

IV. Estimation results

This section presents the results of estimating the parameters of the earnings model, the densities of the random effects, and the transition probabilities. Parameter estimates are presented in subsection (a). Subsection (b) presents the estimated density functions and

⁵ Note that the estimator does not function well in the tails of the distributions. Thus, the tails are estimated by assuming they belong to a mixture of two normals having the same first three moments as the actual distribution in the data. Few observations are in the tails, so this should have minimal effect.

⁶ I do not implement the bias correction of H&M.

the results of some informal specification tests which suggest that the error components are not normal. The estimated transition probabilities are given in subsection (c).

IV (a). Estimated parameters

Table 2 shows the estimates of the coefficients of the earnings model and their standard errors. Estimates of ρ , the variances of ξ and η , and the third through fifth moments of ξ are also given.⁷ The negative third moment suggests the distribution of ξ is skewed to the left, and therefore asymmetrical. Under normality, the value of $\mu_3/\mu_2^2 - 3$ is zero. The data give this value as 35.1, which is evidence that the distribution is thick tailed and not normal. The estimate of ρ is .43; the estimate of α , the coefficient to the LDV, is negative, but insignificant. These values are similar to what has been found in the literature. Lillard and Willis found $\hat{\rho}$ to be in the neighborhood of .40. Geweke and Keane found, for the mixture model, $\hat{\rho}$ to be .655 and $\hat{\alpha}$ to be -.121 – negative, but small in magnitude.

As expected, marriage has an upward influence on earnings, as does being white. In general, greater parental education predicts higher earnings, although not all variables pertaining to parent's education are significant.

The effects of age and education on earnings are not obvious from the table, due to the fact they are formed from a polynomial. A simple plot of earnings vs. age, holding education and all other variables at a constant value, such as their means, reveals that initially earnings increase with age, and later, around the age of 45, decrease – a characteristic commonly found in the literature (Freeman, 1972). A year's increase in education level has a net effect of 5 to 9 percent higher earnings, depending the specific age level. Again, this is a number similar to what has been found in the literature (Psacharopoulos, 1992). To help provide interpretation to the age/education coefficients,

⁷ Methods for estimating the variances of ξ and η are in Hsiao (1986), pp. 55 - 56. The higher moments of ξ were estimated by a relatively straightforward extension (involving higher powers of $\xi_{it} - \xi_{i,t-1}$), available from the author upon request.

Table 3 lists the predicted percentage change in earnings for changes in age and education.

IV (b). Estimated density functions

To motivate the need for estimating the densities, some informal graphical tests of normality were carried out on the data. These graphs reconfirm the non-normality revealed by the moments. Were ξ normally distributed, $\xi_{it} - \xi_{i,t-1}$ would be normally distributed, and hence a normal probability plot of the estimated $\xi_{it} - \xi_{i,t-1}$ would be a straight line, up to random sampling error. Figure 1 depicts this normal probability plot, where the estimated residuals are obtained from (6) above. The plot is S-shaped, suggesting that the distribution of ξ_{it} is not normal; its tails are too thick.⁸

The evidence that η is not normally distributed is less strong. Because we only observe η when convoluted with some form of ε or ξ , we cannot check for normality in the η 's by creating normal probability plot similar to that used above. However, we can perform another informal graphical test. If η is normally distributed, then its characteristic function is $e^{-c\tau^2}$, where c is a constant. Hence a plot of $\ln[\hat{h}_\eta(\tau)]$ against $-\tau^2$ would be straight line. Figure 2 depicts this plot. Any departure in the distribution of η from normality is small, since the curvature in Figure 2 is only slight.

Further evidence that the transitory component may not be normally distributed comes from plotting kernel density estimates of simulated earnings obtained both by assuming normality and by using the nonparametric method. These plots are overlaid with a kernel density estimate of the actual earnings found in the PSID sample. Figure 3 displays this plot. The simulated earnings were generated from the same distribution of covariates found in the data. The predicted distribution of earnings obtained by using the

⁸ A more formal normality test for time-dependent error terms in complicated estimators is in Bai (2003), though it is not implemented here for two reasons. First, it is applicable to $\xi_{it} - \xi_{i,t-1}$ but not η , since η is

nonparametric densities closely mimics that of the sample. The predicted distribution obtained from the normality assumption does not fit well. Simulated quantiles of earnings are examined, as well. Table 4 shows various quantiles of log earnings in the PSID sample. It also shows the percentage of individuals in the simulated data that lie below the given quantiles. As can be seen, the percentage of individuals who lie below each quantile in the simulated earnings data obtained from the nonparametric densities is quite close to the truth. This is not the case under the normality assumption. In fact, the simulated earnings data under the normality assumption places 19% of individuals in the bottom of decile earnings, and 36% in the bottom quartile.

IV (c) Earnings mobility probability

The previous subsection presented evidence that the true distribution of the transitory error effect is not normal, and that assuming normality gives poor prediction of unconditional earnings distribution. This subsection shows the implications of assuming normality on the transition probabilities. Table 5 presents the estimated probabilities that a person has earnings in the bottom quintile of the sample continuously for θ time periods, conditional on that he was initially there. The 20th percentile in the PSID sample is log annual earnings of 8.2. The transition probabilities were estimated for black and white men who are married and initially 30 years old, with both parents possessing only a high school diploma. The probabilities are calculated under the assumption of normality of the error components and by using the nonparametrically estimated densities. The table shows that those with more education tend to have a lower probability of remaining in the bottom quintile. Similarly, whites have a lower probability of staying in the low-earnings state than do blacks. More importantly, the probabilities of remaining in the bottom quintile obtained by using the estimated densities are, in most cases, higher than those obtained under the assumption of normality – sometimes nearly twice as high. Monte Carlo experiments in Section V suggest that these differences are real features of

not observable. Second, the primary goal is to see whether the normality assumption affects the final results, not proving whether these random variables are normal.

the data and not artifacts of the estimation procedure. This evidence suggests that assuming normality usually overstates upward mobility.

For low earnings categories, *e.g.*, uneducated blacks, the nonparametric distribution gives less upward mobility than does assuming normality. Table 5 shows that, for example, under normality, blacks with 8 years of education have about a 23% chance of remaining in the bottom quintile for 11 years, but when using the empirical densities, this number is a larger 30%. Other groups show similar results. This difference is reasonable. Investigating the simulated data reveals that for all agent-types displayed, except college educated whites, the conditional mean level of earnings (conditional on being initially poor) for the first few time periods is in the bottom quintile. To move out of this low-earnings state, an individual would need a large, positive transitory shock to his earnings. This event is more likely with the normal density than it is with the nonparametric one. Figure 4 presents a graph of the nonparametric density function overlaid with the normal of the same variance. As can be seen, the nonparametric density has much more mass near zero, and hence, provides less probability for moderately large transitory shocks and thus less probability of upward mobility.

V. Monte Carlo experiments

This section presents results of Monte Carlo experiments aimed at providing information about the accuracy of the methods used in estimating the parameters of the model, the density of f_{ξ} , and the mobility probabilities.

V(a). Estimates of the parameters

This subsection reports the results of a Monte Carlo experiment that was carried out to check the finite sample performance of the IV estimator of the coefficients described in Section III. This experiment consisted of repeatedly applying the estimator to data generated by simulation from the estimated model.

A balanced panel dataset, with the assumption that each individual has observations for 14 years, was used. The distribution for X_{i1} and V_i found in the real data was used, excluding those who are initially over 51 years of age. After 14 years, these men would exceed 65, and would be at an age outside the typical working years. This yielded a sample size of 57,400, roughly the same as that of the real dataset, with approximately the same distribution of covariates. For each i , I generated ε_{i0} , η_i , and y_{i0} by sampling error terms from the nonparametrically estimated densities. The sampling was done by the method discussed in Appendix 1, which was also used in estimating the earnings mobility probability, except that here I kept all y_{i0} 's. Next, for each individual, for $t = 1, \dots, 14$, I sampled ξ_{it} from its estimated distribution, used the V_i and the appropriate X_{it} , and generated the corresponding y_{it} . With the exception of the marriage variable, the appropriate values for the elements of X_{it} for each individual are predetermined, as discussed in Section IV. To deal with the uncertain path of marriage, I introduced an element of randomness: I changed the value of marriage with a probability equal to the proportion of observations in which marital status changed in the real data.

The coefficients α and β were estimated by the methods discussed in Section III. This process was repeated 1000 times. The means and standard deviations of the estimates for α and each element in β , as well as the actual values used in the data generation process, are reported in Table 6. All the means are very close to the true values, suggesting that the finite sample performance of the IV estimation method is satisfactory.

V (b). Estimates of the density functions

This subsection discusses the results of a Monte Carlo experiment designed to investigate the accuracy of the nonparametric method used to estimate the distributions of the error components. The experiment consisted of sampling error components from a known distribution and generating simulated earnings data from the model described by

(1 – 2). The steps of the nonparametric method were implemented on the simulated earnings data, to obtain an estimate of the density of the transitory shock.

To generate simulated earnings, I sampled ξ from a mixture of two normals with the first 5 moments equal to those estimated from the actual data. Since the evidence suggests that η is not highly non-normal, I sampled it from a zero-mean normal distribution with standard deviation equal to that found in the data. I used the parameters estimated from the real data and the same distribution of covariates as in Section V (a). When executing the steps of the nonparametric estimation method, I used the parameters by which the data were generated, instead of using consistent estimates of them. The rate of convergence of the estimated parameters is much faster than that of the estimated distributions, so this should make little difference and simplifies computation.

A plot of the estimated density of ξ superimposed on the real density used in the generation process appears in Figure 5. The plot shows that the nonparametric method works well for a model with an AR(1) and an asymmetrical transitory error component.

V (c). Estimates of the mobility probability

This subsection presents the results of a Monte Carlo experiment that was designed to check the finite sample properties of the method used in section IV to estimate the earnings mobility probabilities. By sampling the permanent and transitory error terms from a known “true” distribution (specifically, the same distributions used in Section V (b)), earnings paths for black and white married individuals, initially 30 years of age, with 8, 12, and 16 years of education were generated. From the simulated earnings paths, the “true” transition probabilities were calculated. Next, the densities of ξ and η were estimated and the methods of Section IV were used to estimate the transition probabilities. In this experiment, the parameters of the model (α , β , γ and ρ) were held fixed and assumed known, for the same reason discussed in Section V (b).

The results of this experiment are in Table 7. The table presents the probability that the specified individual has log annual earnings of 8.2 or less for θ time periods, conditional on initially having log annual earnings of 8.2 or less. This level of earnings was chosen since it is the 20th percentile in the PSID sample used in this paper. The “true” probabilities from the generated data are listed, as are the estimated probabilities obtained by assuming normality and using the nonparametric densities. The probabilities obtained from nonparametric densities are considerably closer to the truth than the results obtained by assuming normality. The probabilities calculated from the nonparametric distributions are always within 10% of the truth. In some cases, these probabilities remain within 2% of the truth for 11 periods. The probabilities obtained from assuming normality are not even close to the truth. In the later years, the probabilities obtained from the normal distributions are less than half the true probabilities—in all cases displayed. Hence, this experiment suggests that using the nonparametric method provides more accurate results.

VI. Concluding remarks

This paper uses a random-effects panel data model to estimate a model of earnings mobility. This paper uses a recently developed nonparametric method to estimate the distribution of the error components of the model. The results reveal that the transitory component is not normally distributed. Moreover, assuming normality of the error components leads to estimates of upward mobility that are often much higher than those obtained by using the estimated densities and understates the effects of factors such as race and education on mobility. It was demonstrated through a Monte Carlo experiment that the results obtained by using the estimated densities are likely to be more accurate than those obtained by assuming normality.

Appendix 1

Sampling η and the initial conditions ε_0 and y_0 .

To motivate the process for sampling η and the initial conditions ε_0 and y_0 , observe that ε_0 is a function of the independent lagged ξ 's and that y_0 is a function of the same ξ 's and η . Hence, write y_0 as a function of η and anterior ξ 's; write ε_0 as a function of the same lagged ξ 's. To write y_0 in terms of ξ 's and η , note that (i subscript dropped):

$$Y_t = \alpha Y_{t-1} + X_t \beta + Vp + \eta + \varepsilon_t, \quad (\text{A1})$$

and

$$\varepsilon_t = \rho \varepsilon_{t-1} + \xi_t. \quad (\text{A2})$$

Repeatedly substituting (A1) into itself yields:

$$y_t = \sum_{j=0}^{\infty} \alpha^j (X_{t-j} \beta + Vp) + \frac{1}{1-\alpha} \eta + \varepsilon_t + \alpha \varepsilon_{t-1} + \alpha^2 \varepsilon_{t-2} + \dots, \quad (\text{A3})$$

if $|\alpha| < 1$. Repeatedly substituting (A2) into itself, gives:

$$\varepsilon_t = \xi_t + \rho \xi_{t-1} + \rho^2 \xi_{t-2} + \dots \quad (\text{A4})$$

Substituting (A4) into (A3), results in:

$$\begin{aligned} y_t &= \sum_{j=1}^{\infty} \alpha^j (X_{t-j} \beta + Vp) + \frac{1}{1-\alpha} \eta + [\xi_t + \rho \xi_{t-1} + \rho^2 \xi_{t-2} + \rho^3 \xi_{t-3} + \rho^4 \xi_{t-4} + \rho^5 \xi_{t-5} + \dots] \\ &\quad + \alpha [\xi_{t-1} + \rho \xi_{t-2} + \rho^2 \xi_{t-3} + \rho^3 \xi_{t-4} + \rho^4 \xi_{t-5} + \dots] \\ &\quad \quad + \alpha^2 [\xi_{t-2} + \rho \xi_{t-3} + \rho^2 \xi_{t-4} + \rho^3 \xi_{t-5} + \dots] \\ &\quad \quad \quad + \alpha^3 [\xi_{t-3} + \rho \xi_{t-4} + \rho^2 \xi_{t-5} + \dots] + \alpha^5 [\xi_{t-5} + \dots] + \dots \\ &= \sum_{j=1}^{\infty} \alpha^j (X_{t-j} \beta + Vp) + \frac{1}{1-\alpha} \eta + \xi_t + (\alpha + \rho) \xi_{t-1} + (\rho^2 + \alpha \rho + \alpha^2) \xi_{t-2} + \dots \end{aligned}$$

Note that, for sufficiently small α and ρ ,

$$\begin{aligned}
y_t &= \sum_{j=1}^{\infty} \alpha^j (X_{t-j} \beta + V_j p) + \frac{1}{1-\alpha} \eta + \xi_t + (\alpha + \rho) \xi_{t-1} + (\rho^2 + \alpha \rho + \alpha^2) \xi_{t-2} + \dots \\
&\approx A_t + \frac{1}{1-\alpha} \eta + \xi_t + c_1 \xi_{t-1} + c_2 \xi_{t-2} + c_3 \xi_{t-3} + c_4 \xi_{t-4} + c_5 \xi_{t-5}, \quad (\text{A5})
\end{aligned}$$

with

$$\begin{aligned}
A_t &= \sum_{j=1}^5 \alpha^j (X_{t,j-1} \beta + V_j p) \\
c_1 &= (\alpha + \rho) \\
c_2 &= (\rho^2 + \alpha \rho + \alpha^2) \\
c_3 &= (\rho^3 + \rho^2 \alpha + \rho \alpha^2 + \alpha^3) \\
c_4 &= (\rho^4 + \rho^3 \alpha + \rho^2 \alpha^2 + \rho \alpha^3 + \alpha^4) \\
c_5 &= (\rho^5 + \rho^4 \alpha + \rho^3 \alpha^2 + \rho^2 \alpha^3 + \rho \alpha^4 + \alpha^5)
\end{aligned}$$

Thus, y_0 may be approximated by a function of η , ξ_0 , ξ_{-1} , ξ_{-2} , ξ_{-3} , ξ_{-4} , ξ_{-5} , and exogenous variables. The ε_0 is also a function of these same variables, by equation (A4). The initial conditions may be sampled by the following steps:

1. Generate y_0 according to equation (A5) by sampling η and the five ξ 's from their estimated empirical distributions.
2. When conditioning on $y_0 < y^*$, proceed to step 3 only if the y_0 resulting from step 1 is less than y^* . Otherwise repeat step 1.
3. Generate the corresponding ε_0 's by using equation (A4), and the same ξ 's as in step 1.
4. Use the resulting data in the simulation used to calculate (4).

This algorithm quickly generates a set of y_0 , η , and ε_0 .

Appendix 2

Implementing the nonparametric density estimators and choosing the tuning parameters

This appendix discusses the tuning parameters used to estimate the densities of ξ and η , since suggestions for many of these choices are not presented in H&M or Horowitz (1997).⁹ It is necessary to introduce some notation to discuss the tuning parameters. The density of ξ may be estimated via the extensions in H&M (p. 162, Section 6b) and in Horowitz (1997, pp. 135 - 136):

$$\hat{f}_\xi(z) = \int_{-\infty}^{\infty} e^{-jz} \hat{\phi}_\xi(t) g(\lambda_\xi t) dt, \quad (\text{A6})$$

where \hat{f}_ξ is the density; $j = \sqrt{-1}$; $\hat{\phi}_\xi(t)$ is the estimated characteristic function (CF) of ξ ; $g(\cdot)$, a smoothing function, is a CF with support $[-1, 1]$; and λ_ξ is a bandwidth. H&M show that $\hat{\phi}_\xi(t)$ may be estimated by

$$\hat{\phi}_\xi(t) = |\hat{\phi}_{\Delta\xi}(t)|^{1/2} \exp[j\hat{\omega}(t)],$$

where $\hat{\omega}(t)$ is an estimate of the argument or phase to the complex variable $\phi_\xi(t)$, and $\hat{\phi}_{\Delta\xi}(t)$ is the estimated CF of $\xi_{it} - \xi_{i,t-1}$. Evaluating $\hat{\phi}_{\Delta\xi}(t)$ is straight-forward, given residuals $\hat{\xi}_{it} - \hat{\xi}_{i,t-1}$ (obtainable from equation 6). Estimating $\hat{\omega}(t)$ is more complicated. H&M suggest obtaining it by estimating a power series approximation

$$\hat{\omega}(\tau) = \sum_{i=1}^{K_n} \hat{a}_i \tau^i,$$

for some $K_n \rightarrow \infty$ as $n \rightarrow \infty$. The \hat{a}_i 's are obtained as follows: Let \hat{q} be the empirical CF of $(\xi_{i2} - \xi_{i1}), (\xi_{i3} - \xi_{i1}), \dots, (\xi_{iT} - \xi_{i1})$, for some T . That is,

$$\hat{q}(\tau_2, \dots, \tau_T) = \frac{1}{n} \sum_{i=1}^n \exp\{j[(\xi_{i2} - \xi_{i1})\tau_2 + \dots + (\xi_{iT} - \xi_{i1})\tau_T]\}.$$

Then, the \hat{a}_i 's can be obtained by the OLS regression of

$$\text{Im}\{\log[\hat{q}(\tau, \dots, \tau) / |\hat{q}(\tau, \dots, \tau)|]\} = \sum_{i=2}^{K_n} a_i [(T-1) - (T-1)^i] \tau^i. \quad (\text{A7})$$

Once $\hat{\phi}_\xi(t)$ is obtained, it is relatively straight-forward to obtain the empirical CF's of ε and η and their corresponding densities. Methods for estimating these are presented in the H&M and Horowitz papers.

The smoothing function g and several tuning parameters needed to be selected: λ_ξ , T , K_n , τ_{\min} , and τ_{\max} , where τ_{\min} and τ_{\max} are the minimum and maximum values, respectively, of τ in (A7). Finally, a bandwidth λ_η needs to be selected to estimate the density of η . I now discuss each of these choices; none of them are obvious, and future work would clearly benefit by more rigorous selections. In this paper, the

⁹ There is no need to estimate the density of ε , since, as shown in Appendix 1, the initial conditions used in calculating the earnings mobility probabilities can be written as a function of ξ (rather than ε) and since, given these initial conditions, only ξ is required to simulate earnings for subsequent time periods.

smoothing function g was chosen to be the same as in H&M (p. 156), *i.e.*, the fourfold convolution of the uniform distribution with itself (this CF has corresponding density of $c[\text{Sin}(x)/x]^4$, where c is a normalization constant). The bandwidths were chosen in the informal graphical manner put forth by H&M (page 156, last paragraph). Specifically, my estimated $\hat{\phi}_\xi(t)$ is nearly zero for $|t| > 25$, except for a few wiggles. Hence, λ_ξ was chosen to set the integrand in (A6) equal to zero for values of t beyond that point. Since $g(t) = 0$ for $|t| > 1$, $\lambda_\xi = .04$. (The bandwidth used in estimating η was chosen in the same manner; $\lambda_\eta = .04$.)

Another tuning parameter is K_n , the number of terms in the power series expansion (A6). Experimentation in fitting known mixture distributions suggested that K_n has a limited effect on the ability to estimate the density, provided it is not too small; K_n of 10 or more usually provided a good fit. In estimating the PSID data, I set $K_n = 25$. Small changes the parameter did not substantially effect results. Similarly, since the data contain unbalanced panels, T had to be chosen. I set $T = 5$, leaving the bulk of individuals in the sample. Slight variations in T did not seem to matter much.

Less obvious parameter choices arise when estimating (A7). Specifically, it is unclear how to choose τ_{\min} and τ_{\max} . Let \mathcal{D} be the interval $[\tau_{\min}, \tau_{\max}]$. Experimentation, using mixed distributions, suggested that the choice of \mathcal{D} has a great impact on the ability of the deconvolution estimator to capture asymmetry in ξ . A plot of the LHS of (A7) against τ is smooth for τ near zero but chaotic for larger τ (see Figure A1). This phenomenon induced a tradeoff in selecting the width of \mathcal{D} . In the experimentation, an insufficiently wide \mathcal{D} captured too little information about the overall shape of (A7) to give accurate estimates of the known density. However, for wide \mathcal{D} , the chaotic pattern at large τ made it impossible for a regression to capture the shape of (A7). Using larger K_n did not help, due to the extreme nonlinearity of the function at large τ . In the experimentation, the best estimates of the known density were obtained when \mathcal{D} was chosen to be as large as possible, while still giving the appearance of a good fit in the regression (A7), in the region where the function is relatively smooth (*e.g.*, roughly $\tau \in [-9, 9]$ in figure A1). Hence, when estimating the PSID I chose \mathcal{D} by repeatedly estimating (A7), slightly widening \mathcal{D} each time, and stopping when the predicted values of (A7) appeared to become poor in the relatively smooth area. That is, I used the largest \mathcal{D} that visually gave a good fit; $\mathcal{D} = [-16, 16]$. (This interval is smaller than that over which the empirical CF of ξ was evaluated, as determined by the smoothing function g and the bandwidth λ_ξ .) Figure A1 shows the LHS of (A7) and the corresponding fitted values over this range.

The tails of the estimated distributions of η and ξ are wiggly and sometimes negative, and therefore not useful in estimating probabilities. To cope with this problem, I assigned area to the tails of the estimated distributions. Since the density η did not appear non-normal, I assigned area to its tails ($|\eta| > 3$) by using the same mean and variance of η as in the data. For the tails of ξ (*i.e.*, $\xi \notin [-.5, .65]$), I used the mixture of two normals with the first 5 moments equal to those found in the actual data. To compensate for the fact that the resulting densities did not integrate to one, I normalized them. Relatively little area is found in the tails, so this solution should cause minimal distortion. (The density of ε was not used in the simulations.)

Finally, estimating the densities can be computationally expensive. Therefore, in simulating the mobility probabilities, rather than estimating the corresponding density each time an error component needed to be drawn, I made look-up tables. I divided the domain over which I estimated the densities (*i.e.* $\eta \in [-3, 3]$ and $\xi \in [-.5, .65]$) into bins of width .012 and estimated the density at the midpoint of each bin. The value at the midpoint served as the value of the density over the entire bin. Similarly, I divided the region in which I assumed a mixture of normals into bins.

References

- Arellano, M. and S.R. Bond. 1991. Some tests of specification for panel data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies* 58: 277-297.
- Bai, Jushan. 2003. Testing parametric conditional distributions of dynamic models. *The Review of Economics and Statistics* 85: 531 - 548.
- Bane, Mary Jo and David T. Ellwood. 1985. Slipping into and out of Poverty: the Dynamics of Spells. *The Journal of Human Resources* 21: 2-23.
- Blundell, Richard and Stephen Bond. 1997. Initial conditions and moment restrictions in dynamic panel data models. UCL Discussion Papers in Economics Working Paper 97-07.
- Chamberlain, G. 1983. Panel Data. In *Handbook of Econometrics*, Griliches and M. Intriligator, (eds), North Holland: Amsterdam.
- Coe, Richard D., Greg J. Duncan, and Martha S. Hill. 1978. Dependency and Poverty in Short and Long Run. In *5000 American Families: Patterns and Economic Progress, Vol 6.*, Institute for Social Research: Ann Arbor Mich.
- Daly, Mary and Rob Valletta. 2003. Earnings Inequality and Earnings Mobility in the U.S. *FRBSF Economic Letter* 28.
- Datcher – Loury, Linda. 1986. Racial Differences in the Stability of High Earnings among Young Men. *Journal of Labor Economics* 4: 301-316.
- Duncan, Greg J. 1984. *Years of Poverty and Plenty*. Institute for Social Research: Ann Arbor, Mich.
- Freeman, Richard B. 1972. *Labor Economics*. Prentice-Hall, Inc: Englewood Cliffs, NJ.
- Geweke, John and Michael Keane. 2000. An empirical analysis of income dynamics among men in the PSID: 1968-1989. *Journal of Econometrics* 96: 293-356.
- Hill, Martha S. 1981. Some Dynamic Aspects of Poverty. In *5000 American Families: Patterns and Economic Progress, Vol 9*. Institute for Social Research : Ann Arbor Mich.
- Horowitz, Joel L. 1997. *Semiparametric Methods in Econometrics*. Springer-Verlag New York, Inc.: New York.
- Horowitz, Joel L. and Marianthi Markatou. 1996. Semiparametric estimation of regression models for panel data. *Review of Economics Studies* 63: 145-168.
- Hsiao, C. 1986. *Analysis of Panel Data*. Cambridge University Press: Cambridge.
- Levy, Frank. 1977. How big is the American Underclass? Working paper 0090-1. The Urban Institute: Washington.
- Lillard, L. A. and R.J Willis. 1978. Dynamic aspects of earnings mobility. *Econometrica* 46: 985-1012.
- McCall, John J. 1973. *Income Mobility, Racial Discrimination, and Economic Growth*. D.C. Heath and Company: Lexington, Massachusetts.
- Moffitt, Robert and Peter Gottschalk. 1998. Trends in the Variances of Permanent and Transitory Earnings in the U.S. and their Relation to Earnings Mobility. Working paper, Johns Hopkins University.
- Psacharopoulos, George. 1992. Returns to Education: A further International Update & Implications. In *The Economic Value of Education, International Library of Critical Writings in Economics, vol. 17*. Ashgate: Brookfield, VT.
- Rainwater, Lee. 1982. Persistent and Transitory Poverty: a New Look, working paper 70. Joint Studies for Urban Studies: Cambridge, MA.
- Robin, Jean-Marc and Stéphane Bonhomme. 2004. Modeling Individual Earnings Trajectories using Copulas with an Application to the Study of Earnings Inequality: France, 1990-2002. Working paper, Université de Paris 1-Panthéon-Sorbonne (EUREQua), CREST-INSEE.
- Silverman, B.W. 1985. *Density estimation for Statistics and Data Analysis*. Chapman and Hall: New York.
- White, Halbert and G. MacDonald. 1980. Some Large Sample Tests for Nonnormality of the Linear Regression Model. *Journal of the American Statistical Association* 76: 419-433.

Table 1—Means and standard deviations of the variables

VARIABLE	MEAN	ST. DEV.	DESCRIPTION
earn	8.711003	0.7673415	Log Annual Earnings
age*	0.4133124	0.112503	Age
race†	0.2729287	0.4454691	race = 1 if Black
educ†	11.70469	3.379222	Years of Education
married	0.8399606	0.3666463	Dummy Indicating Married
fathed1†	0.0391891	0.1940467	Dummy Indicating Father's Education Missing
mothed1†	0.0391481	0.1939492	Dummy Indicating Mother's Education Missing
fathed2†	0.270364	0.4441524	Dummy Indicating Father's Education H.S.
mothed2†	0.4542041	0.4979034	Dummy Indicating Mother's Education H.S.
fathed3†	0.0613074	0.2398958	Dummy Indicating Father's Education College
mothed3†	0.0526078	0.2232516	Dummy Indicating Mother's Education College
age2	0.1834838	0.0977576	age ²
age3	0.086756	0.0672389	age ³
edage	4.744865	1.811688	educ*age
edage2	2.067019	1.23901	educ*age ²
edage3	0.960918	0.8041386	educ*age ³

* measured as age ÷ 100.

† denotes time invariant variables.

Table 2 - Estimated parameters

VARIABLE	ESTIMATE	ST. ERR.
Learn (Lagged Earnings)	-.3047257	.2186512
married	.1198652	.0162113
age	-66.17857	11.4648
age2	143.6497	24.53757
age3	-105.1305	17.39596
edage	5.209332	1.01026
edage2	-10.34064	2.05162
edage3	6.781913	1.370012
race	-.3657833	.007294
educ	-.7637857	.0010205
fathed1	-.0481818	.0159195
mothed1	-.0378569	.0174562
fathed2	.0517658	.0080009
mothed2	-.0191424	.0064059
fathed3	-.0100347	.0140393
mothed3	.055004	.0140592
constant	20.56714	.013087
σ^2_{η} :	.5433	
Moments of ξ :	μ_2 : .154201677	μ_3 : -.271924211
	μ_4 : .906143108	μ_5 : -2.7302413
ρ :	.43	

Table 3 - Predicted percentage increase in annual earnings for corresponding increase in age/education

Education in years	Age changes from (top) to (bottom)	Predicted percentage increase in earnings
12	25 years old 35 years old	16%
12	35 years old 45 years old	5%
12	45 years old 55 years old	-20%
Age	Education changes from (top) to (bottom)	Predicted percentage increase in earnings
25	12 years 16 years	17%
35	12 years 16 years	51%
45	12 years 16 years	60%
55	12 years 16 years	58%

Table 4- Quantiles of earnings in the PSID sample vs. the percentage that lie below each quantile in the simulated earnings

Quantile	Quantile value	% below normal	% below nonparametric
5%	7.44	9%	4%
10%	7.87	19%	10%
25%	8.33	36%	27%
50%	8.74	54%	53%
75%	9.08	68%	74%
90%	9.36	78%	87%
95%	9.53	84%	93%
99%	10.03	93%	99%

Quantile value: the value of the given quantile in the PSID sample

% below normal: the percentage who lie below each quantile in the normal simulated data

% below nonparametric: the percentage who lie below each quantile in the nonparametric simulated data

Table 5 – Estimated transition probabilities: the probability of remaining in bottom quintile of sample for θ time periods conditional on being initially there

θ	BLACK		8		12		16	
	normal	nonpar.	normal	nonpar.	normal	nonpar.	normal	nonpar.
1	0.665	0.710	0.584	0.627	0.506	0.533		
3	0.447	0.530	0.344	0.425	0.251	0.320		
5	0.349	0.436	0.240	0.321	0.152	0.226		
8	0.273	0.351	0.158	0.236	0.084	0.151		
11	0.227	0.299	0.118	0.186	0.054	0.113		

θ	WHITE		8		12		16	
	normal	nonpar.	normal	nonpar.	normal	nonpar.	normal	nonpar.
1	0.573	0.618	0.491	0.515	0.421	0.451		
3	0.341	0.433	0.246	0.327	0.185	0.252		
5	0.243	0.337	0.157	0.244	0.103	0.177		
8	0.178	0.267	0.094	0.180	0.054	0.107		
11	0.140	0.222	0.066	0.145	0.033	0.066		

nonpar.: the probabilities obtained by using the nonparametrically estimated densities
normal: the probabilities obtained by assuming normality

Table 6 –Mean of estimated parameters vs. true value used in data generation process

Variable:	Learn	Married	Age	Age2	Age3	Edage	Edage2	Edage3
Actual:	-0.30473	0.119865	-66.1786	143.6497	-105.131	5.209332	-10.3406	6.78193
Estimated:	-0.31115	0.120298	-66.3196	143.841	-105.189	5.219202	-10.3461	6.773175
St.Dev:	0.091161	0.009035	9.573385	22.76997	17.66739	0.790875	1.878125	1.457724

Actual: true value used in data generation process
Estimated: mean of the 1000 estimated parameters
St. Dev: Standard deviation of the 1000 estimated parameters

Table 7 - True mobility probabilities vs. estimated probabilities: the probability of remaining in bottom quintile of sample for θ time periods conditional on being initially there

θ	BLACK			12			16		
	8	8	normal	true	nonpar.	normal	true	nonpar.	normal
1	0.820	0.831	0.701	0.763	0.762	0.642	0.678	0.683	0.571
2	0.743	0.755	0.562	0.660	0.647	0.482	0.543	0.529	0.393
3	0.691	0.703	0.479	0.590	0.570	0.383	0.457	0.427	0.289
4	0.655	0.666	0.421	0.537	0.514	0.319	0.391	0.351	0.219
5	0.628	0.636	0.378	0.494	0.470	0.274	0.341	0.295	0.175
6	0.605	0.612	0.343	0.459	0.433	0.239	0.302	0.250	0.143
7	0.586	0.592	0.316	0.433	0.402	0.212	0.270	0.212	0.118
8	0.569	0.576	0.293	0.408	0.377	0.190	0.246	0.183	0.100
9	0.554	0.562	0.274	0.387	0.355	0.173	0.223	0.161	0.085
10	0.542	0.551	0.258	0.370	0.335	0.157	0.206	0.141	0.074
11	0.531	0.540	0.246	0.354	0.319	0.144	0.190	0.125	0.064

θ	WHITE			12			16		
	8	8	normal	true	nonpar.	normal	true	nonpar.	normal
1	0.735	0.755	0.618	0.650	0.655	0.553	0.543	0.560	0.484
2	0.643	0.649	0.466	0.533	0.508	0.381	0.412	0.425	0.302
3	0.583	0.580	0.377	0.454	0.414	0.286	0.329	0.331	0.205
4	0.540	0.534	0.316	0.403	0.348	0.225	0.273	0.269	0.150
5	0.509	0.499	0.274	0.365	0.298	0.186	0.231	0.228	0.118
6	0.483	0.472	0.240	0.332	0.260	0.157	0.197	0.195	0.093
7	0.463	0.449	0.217	0.309	0.231	0.134	0.174	0.170	0.074
8	0.447	0.430	0.198	0.286	0.207	0.117	0.149	0.151	0.061
9	0.433	0.415	0.181	0.270	0.188	0.102	0.133	0.136	0.051
10	0.421	0.401	0.169	0.252	0.170	0.091	0.121	0.124	0.042
11	0.409	0.388	0.157	0.239	0.155	0.082	0.109	0.112	0.035

true: the “true” probability in the Monte Carlo experiment
nonpar.: the estimated probabilities obtained from using the nonparametric distributions
normal: the estimated probabilities obtained from assuming normality

Figure 1 – Normal probability plot of $\xi_{it} - \xi_{i,t-1}$

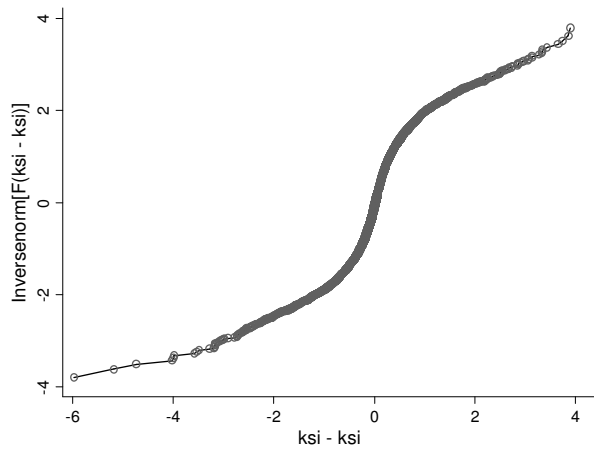


Figure 2 - Estimated empirical c.f. of η_i vs. square of its argument

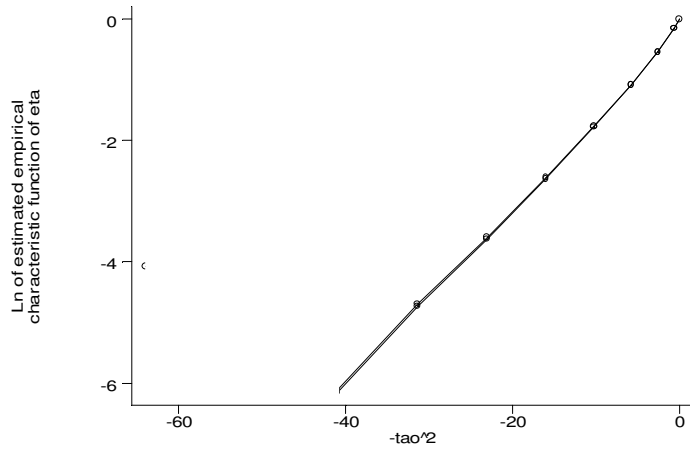


Figure 3 - Distribution of earnings in sample (solid line) vs. that predicted by normal (dotted line) and semiparametric (dashed line) models

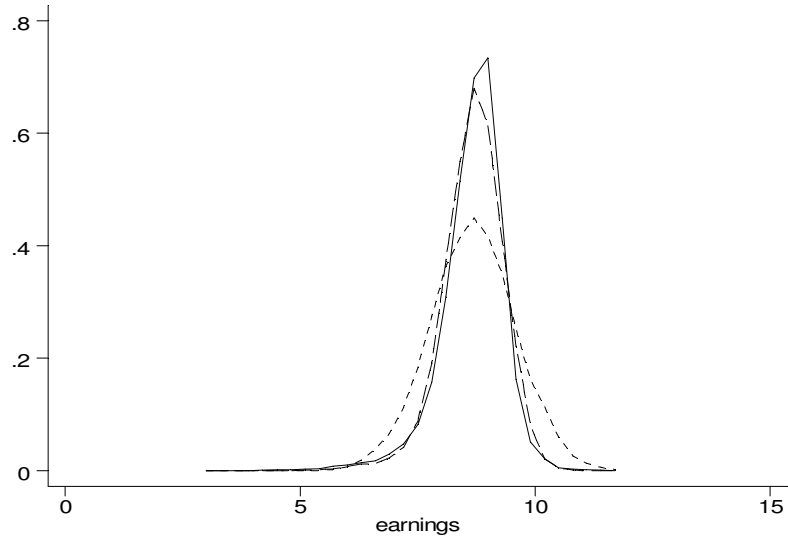
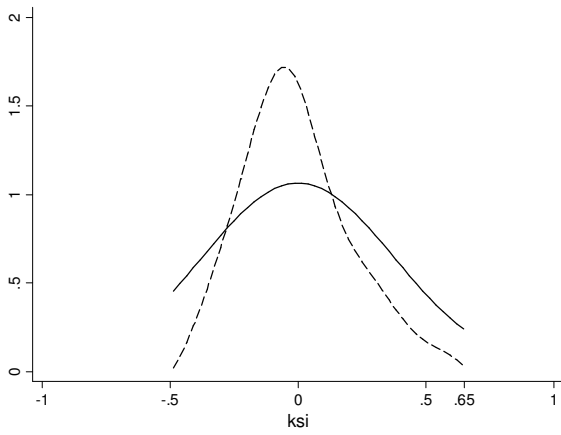


Figure 4 – Nonparametrically estimated density function of ξ overlaid with normal (dotted line)*



* The nonparametrically estimated density is negative for $\xi \notin [-.5, .65]$. Therefore, the figure was not plotted for these ξ . (As stated in the text, the tails of the density were estimated with a mixture of two normals when estimating the earnings mobility probabilities.)

Figure 5 – Estimated density (dotted line) of ξ vs. its true density (solid line)

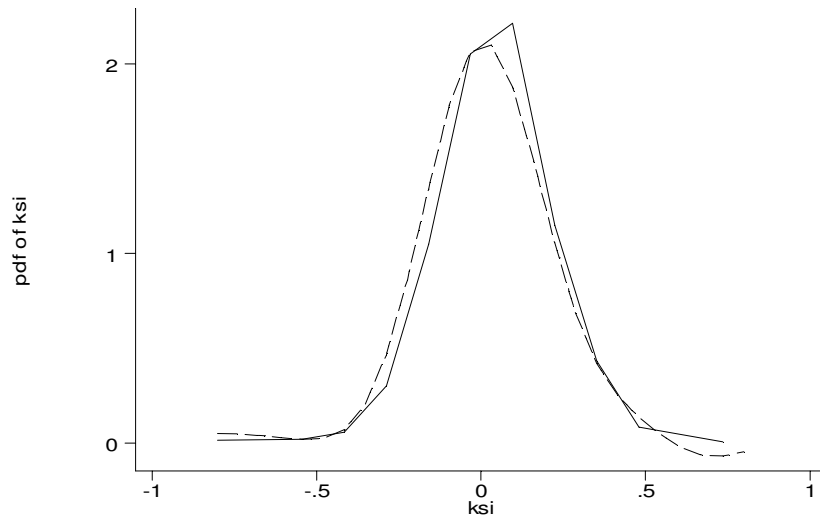
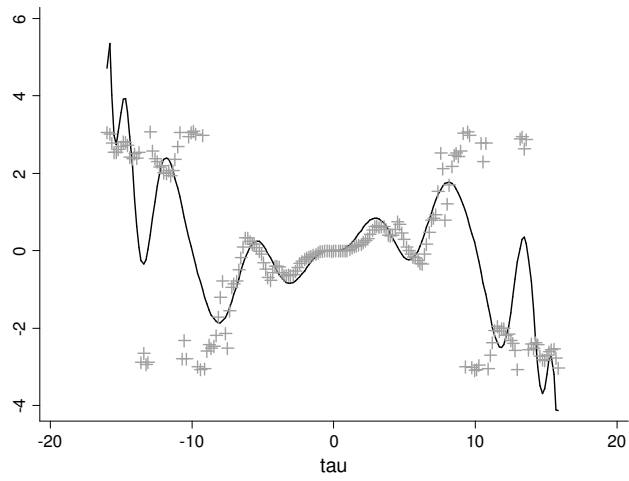


Figure A1 – Fitted regression of power series expansion in (A7)*



* Pluses show the actual values of the LHS of (A7), *i.e.*, $\text{Im}\{\log[\hat{q}(\tau, \dots, \tau) / |\hat{q}(\tau, \dots, \tau)|]\}$, for each τ . Solid line represents fitted values.