

# Heterogeneity in the Production of Human Capital

Solomon W. Polachek<sup>\*</sup>  
Department of Economics  
State University of New York at Binghamton and IZA  
[polachek@binghamton.edu](mailto:polachek@binghamton.edu)

Tirthatanmoy Das  
Department of Economics  
University of Central Florida and IZA  
[Tirthatanmoy.Das@ucf.edu](mailto:Tirthatanmoy.Das@ucf.edu)

Rewat Thamma-Apiroam  
Faculty of Economics  
Kasetsart University, Thailand  
[fecoret@hotmail.com](mailto:fecoret@hotmail.com)

October 5, 2014

**Key words: Life-Cycle Model; Ability to Learn; Ability to Earn; Heterogeneity**  
**JEL Classifications: J24; J29; J31; J39**

## Abstract

We derive a tractable nonlinear earnings function which we estimate separately individual-by-individual using the NLSY79 data. These estimates yield five important parameters for each individual: three ability measures (two representing the ability to learn and one the ability to earn), a rate of skill depreciation, and a time discount rate. In addition, we obtain a population wide estimate of the rental rate of human capital. To illustrate heterogeneity in the production of human capital, we plot the distribution of these parameters along with NLSY79 reported AFQT scores. By utilizing these parameters, we are able to verify a number of heretofore untested theorems relating to human capital investments. In addition, we are able to show how these human capital production function parameters relate to cognitive ability, personality traits, and family background. Further, accounting for this individual specific heterogeneity dramatically reduces estimates of population-wide persistence of (unit-root) permanent and (mean-reverting) transitory shocks.

---

\* The initial version of this paper was written while Polachek was a Visiting Scholar at the NBER in Cambridge, Massachusetts. It was presented at the 2011 IZA Cognitive and Non-Cognitive Skills Workshop in Bonn, Germany. The current version now focuses on the heterogeneity of individual human capital production function parameters. We thank Vikesh Amin, Christian Belzil, Hwan-Sik Choi, Armin Falk, Alfonso Flores-Lagunes, Richard Freeman, Claudia Goldin, James Heckman, Larry Katz, Subal Kumbhakar, Dennis Pixton, Arnab Roy, Anton Schick, Xiangin Xu, Francis Yammarino, and Bong Joon Yoon for valuable advice and discussion.

## 1. Introduction

Parameters of life-cycle models are used in various branches of economics. For example, they are employed to calibrate dynamic general equilibrium models (King and Rebelo, 1999), to interpret skill formation (Cunha, Heckman, Lochner, and Masterov, 2005), but there are numerous other applications.<sup>1</sup> Typically, such parameters are estimated population-wide. However, as Browning, Hansen and Heckman (1999) show, heterogeneity of these human capital parameters is crucial because often representative agent models are severely limited and can yield erroneous results. Because of this insight, macroeconomics has started to use models where heterogeneity is present (Heathcote, et al., 2009). Currently, studies that get at parameter heterogeneity of human capital models do so in a limited way. None, to our knowledge, examine how *each* parameter varies individual-by-individual.

In this paper, we estimate individual-specific human capital parameters. We use these parameters to answer a number of questions: First, knowing individual-specific human capital parameters enables us to test heretofore untested aspects of the life-cycle human capital model. An example is whether a greater “ability to learn” (in Heckman et al.’s, 1998, terminology) is associated with more years of school and whether a greater “ability to earn” is associated with less years of school. Another is whether a higher rate of time preference is associated with fewer years of school. Second, knowing individual-specific human capital parameters enables us to assess how ability as well as time preference and skill depreciation are related to personality. For example, do individuals with a high internal locus of control have a greater ability to learn? Is emotional depression associated with time preference? Third, knowing individual-specific human capital parameters enables us to examine how family background, including getting educational stimuli as a child, is related to one’s abilities to learn and earn. Fourth, knowing individual-specific parameters enables us to assess how accounting for heterogeneity affects estimated responses to permanent (unit-root) and transitory (mean-reverting) income shocks. In short, utilizing these parameters has implications for a wide range of issues, from the nature versus nurture debate to macroeconomic policy.

With the advent of speedier computers, better optimization routines, and longer panels than in the past, one can retrieve individual-specific parameters of the human capital life-cycle model by estimating appropriate earnings functions individual-by-individual. In this paper, we

---

<sup>1</sup> See Cunha and Heckman (2007), Guvenen (2011) and Polachek (2008) for additional applications.

obtain individual specific measures for five such parameters. As will be explained, these include each person's rate of time preference, each person's skill depreciation rate, and (again using Heckman et al.'s 1998 nomenclature) three ability parameters – two of which measure a person's "ability to learn" and one a person's "ability to earn".

To estimate these parameters person-by-person, we adopt a number of innovations. First, we devise a tractable formulation of an earnings function by modifying Haley's (1976) nonlinear specification. This enables us to estimate and identify the five basic parameters mentioned above for each person. We prove identification based on Newey and McFadden's (1986) conditional mean criterion. Second, we adopt a maximization routine more prone to converge to a global optimum than traditional hill-climbing techniques. Third, we examine the plausibility of our estimates by testing whether they are consistent with individual choices based on life-cycle human capital theory. Fourth, we plot the distribution of each parameter across the population and compare these distributions by race. Fifth, we show how family background as well as skill-based tests and personality are related to one's abilities to earn and learn. Finally, sixth, we assess how taking account of such heterogeneity affects responses to permanent and transitory shocks estimated in earnings dynamics models.

Our estimates yield a number of new findings. For example, on the micro level, we find that blacks have higher rates of time discount and skill depreciation than whites. Individuals with both higher time discount rates and greater rates of the skill depreciation have fewer years of school. Individuals with a high internal locus of control, and individuals who demonstrate high levels of self-esteem, exhibit greater ability as well as lower skill depreciation and time discount rates. Individuals inclined towards mental depression have a higher time discount. At the same time, family background, such as higher parental education, is associated with a greater ability to learn, lower skill depreciation, and a smaller rate of time discount. Educational stimuli, such as growing up in a household that subscribed to newspapers and magazines, are associated with a higher ability. Conversely, growing up poor is associated with lower levels of ability. On the macro level, we find that accounting for heterogeneity reduces estimates of population-wide reactions to permanent and transitory shocks by over 50%.

Of course a number of assumptions underlie our approach. First, we assume individuals plan their human capital investment strategy based on expectations that they seek to work each year of their working life. This is why we concentrate on males who generally have continuous

work histories. Second, we assume individuals use their time and existing human capital to create new human capital, but we do not include other inputs such as books and computers as well as parental, teacher, and school quality which also can be used to create additional human capital.<sup>2</sup> Third, we assume labor markets reward individuals based on human capital, and neither incomplete information nor incentive pay governs worker earnings. Fourth, we assume human capital is homogeneous in that remuneration per unit of human capital is constant both across the population and over the life-cycle of each individual.<sup>3</sup> Fifth, we assume all human capital production function parameters remain constant throughout each person's life. In the context of our model, this means we assume that ability does not change over one's lifetime, though modifications can be made to parameterize changes in measured ability as environmental factors such as job, industry, or location change (Borghans et al., 2008). Finally, we rely on individuals with a significant work history. Obviously, those with a more complete work-history constitute a select sample of the population. Such selectivity biases could come about when making inferences about racial differences in ability, for example, if black workers are relatively more able than white workers compared to black and white non-workers. However, we find that the ability advantage of workers to non-workers is similar for both blacks and whites, so that this bias is at worst very small.

The remainder of the paper is organized as follows: Section 2 derives the life-cycle human capital model that forms the basis of our estimation. Section 3 describes our estimation, including issues regarding identification. Section 4 explains the data. Section 5 gives our results. Finally, section 6 concludes.

## **2. Using the Life-Cycle Human Capital Model to Estimate Ability**

The human capital life-cycle investment model yields a complex nonlinear earnings function.<sup>4</sup> From this nonlinear function we identify three ability parameters based on the production

---

<sup>2</sup> Heckman (2008) describes how to modify the underlying human capital production function to include these factors as well as family background and personality.

<sup>3</sup> We cannot test this assumption because age variation in the NLSY79 is limited. However, we find cohort effects to be negligible. Further, we test whether rental rates per unit of human capital differ across individuals based on occupation and other characteristics, but by and large, we find these rental rates vary little based on these characteristics.

<sup>4</sup> Mincer's log-linear specification gets around these nonlinearities by assuming time-equivalent human capital declines linearly with age. There are conceptual issues regarding estimation as well as interpretation of the schooling and experience coefficients in the Mincer earnings function. See Heckman, Lochner and Todd (2006).

function of human capital. Using Heckman, Lochner and Taber's (1998) terminology, two of these parameters depict "ability to learn" because they measure the ease which an individual can create new human capital from old human capital.<sup>5</sup> The third parameter depicts an individual's initial "ability to earn" because it represents earnings power devoid of human capital investments.<sup>6</sup> In addition, for each individual, we estimate a skill depreciation parameter as well as a rate of time preference.

The derivation of the nonlinear earnings function containing these parameters entails the typical lifetime maximization paradigm. In the model, one's earnings are directly proportional to human capital stock. Each year one's human capital stock is augmented by the amount of new human capital one creates through schooling or on-the-job training, and one's human capital stock is diminished by the amount human capital depreciates. Creating new human capital entails using time and existing human capital to produce new human capital, given one's ability. The greater one's ability, the more human capital one can produce, and the more rapidly one can increase earnings power from year-to-year (Ben-Porath, 1967).

Whereas not everyone accepts the human capital framework as the basis for modeling earnings, the approach is surprisingly robust compared to other models in explaining earnings patterns.<sup>7</sup> For example, screening models explain why education enhances earnings, occupational segregation models explain why women earn less, efficiency wage models explain certain wage premiums, and productivity enhancing contract models explain upward sloping (though not necessarily concave) earnings profiles; but none of these theories simultaneously explain all these phenomena, whereas the human capital model does. But more important, these other models do not allow one to identify ability, skill depreciation, or time discount rates from an empirical specification. For this reason, we adopt the human capital model.

## **2.1 The Ben-Porath Model**

---

<sup>5</sup> The first of these two parameters is an individual's human capital production function output elasticity, and the second is the individual's human capital production function total factor productivity parameter.

<sup>6</sup> From these we also provide estimates of the value of human capital stock measured at the time one graduates from school and enters the labor market.

<sup>7</sup> See Rubinstein and Weiss (2006) for a survey of these approaches as well as a discussion of the importance of heterogeneity.

The Ben-Porath (1967) model assumes individuals invest to maximize expected lifetime earnings.<sup>8</sup> Investment is governed by a production function in which one combines own time and ability along with past human capital investments to create new human capital. At each time period, the marginal cost of each unit of investment is essentially the foregone earnings associated with the time needed to produce an additional unit of human capital.<sup>9</sup> The marginal gain is the present value of each additional unit of human capital. Ben-Porath's innovation was to realize that the finite life constraint implies a monotonically declining marginal gain over the life-cycle (at least for individuals that work continuously throughout their lives).<sup>10</sup> The equilibrium yields a human capital stock that rises over the life-cycle at a diminishing rate. This results in the commonly observed concave earnings profile.

The solution to Ben-Porath's earnings function is nonlinear. At the time of this breakthrough in 1967, few computers were fast enough to easily estimate its parameters. However, given the advent of faster computers and longer panels containing individual data, we feel now is a good time to examine the implications of the life-cycle model. Our innovation is to do so person-by-person. Obtaining person-specific life-cycle parameters gets at heterogeneity. As already mentioned, this heterogeneity is a relatively important issue in micro-based econometric research. In addition, it has implications for calibrating macroeconomic models, and it has implications regarding the earnings-dynamics literature. We adopt a generalized Haley (1976) specification.

## 2.2 Generalizing the Haley Model

The human capital model assumes an individual's potential earnings  $Y_t^*$  (what a person *could* earn) in time period  $t$  are directly related to human capital stock  $E_t$ . As such,

$$Y_t^* = RE_t \tag{1}$$

---

<sup>8</sup> Incorporating labor supply enables one to maximize utility potentially enabling one to identify specific taste parameters, but doing so requires a number of additional assumptions to identify key earnings function parameters.

<sup>9</sup> In more complicated models this cost also includes expenses for goods such as tuition, books, computers, and other material inputs to create human capital. As does Mincer and others, we assume the goods components are offset by earnings during the investment process.

<sup>10</sup> See Polachek (1975) for the case of discontinuous labor force participation.

where for simplicity  $R$  is assumed to be the constant rental rate per unit of human capital.<sup>11</sup> Human capital stock is accumulated over one's lifetime by judicious investments in oneself via schooling and on-the-job training (as well as health, job search and other earnings augmenting types of human capital).<sup>12</sup> The rate of change in human capital stock,  $\dot{E}_t$  is the amount of human capital produced ( $Q_t$ ) minus depreciation, so that

$$\dot{E}_t = Q_t - \delta E_t \quad (2)$$

where  $\delta$  is the constant rate of human capital stock depreciation. For simplicity, we assume individuals create human capital using a Cobb-Douglas production function such that

$$Q_t = \beta K_t^b \quad (3)$$

where  $K_t$  is the fraction of human capital stock reinvested in time period  $t$  and  $b \in (0,1)$  and  $\beta$  are production function parameters.<sup>13</sup> The parameter  $b$  reflects the rate at which current (invested) human capital stock is transformed to new human capital. It indicates how one acquires new knowledge from old, and as such denotes how quickly one learns. We designate  $b$  to depict the “scale” at which one learns. The “technology” parameter  $\beta$  represents total factor productivity. Both  $\beta$  and  $b$  reflect an individual's ability to learn.

The individual's objective is to maximize discounted disposable earnings,  $Y_t$ , over the working life-cycle.<sup>14</sup> This goal is achieved by choosing the amount of human capital,  $K_t$ , to reinvest each year,  $t$ , in order to maximize the present value of lifetime earnings

$$\max_{K_t} J = \int_0^N e^{-rt} Y_t dt \quad (4)$$

---

<sup>11</sup> Heckman et al. (1998) assume that the rental rates can vary by schooling level. Polachek (1981) assumes the rental rate can vary by type of human capital. Polachek and Horvath (1977) assume the rental rate can vary by geographic location. Earnings dynamics models (Meghir and Pistaferri, 2011) assume rental rate shocks can affect the investment process. To identify each parameter we maintain the assumption of a constant rental rate. Later in the paper, we test whether rental rates vary significantly by population strata.

<sup>12</sup> Specific training is also included because, according to Kuratani (1973), in equilibrium workers receive remuneration for the exact same portion of specific training they pay for, which they finance by taking lower wages during the training period.

<sup>13</sup> As already mentioned, we assume no additional inputs other than one's own human capital. Less simplified production functions could entail individuals employing “goods” inputs such as teachers, books, and study time. For example, Ben-Porath (1967) assumes  $q_t = \beta K_t^{b_1} D_t^{b_2}$  where  $D_t$  equals other inputs. Later empirical analysis precludes taking account of these other factors of production because no data are available for these other inputs. Thus we adopt the above more simplified human capital production function used by Haley (1976).

<sup>14</sup> As already mentioned, we abstract from labor supply.

where  $J$  is the total discounted disposable earnings over the working life-cycle,  $r$  is the personal time discount rate, and  $N$  is the age at which one retires (assumed known with certainty).<sup>15</sup>

Disposable earnings are

$$Y_t = R[E_t - K_t]. \quad (5)$$

Maximization of (4) subject to equations (2) and (3) can be done by maximizing the Hamiltonian

$$H(K_t, E_t, \lambda_t, t) = e^{-rt} R[E_t - K_t] + \lambda_t [\beta K_t^b - \delta E_t] \quad (6)$$

with constraints  $E_t - K_t \geq 0$ , which means one cannot invest using more human capital than one currently has; and the transversality condition  $\lambda_N = 0$ , which indicates a zero (labor market) gain from human capital investing in one's final year at work. The solution involves three phases: (1) specialization in human capital investment, the time periods when  $K_t = E_t$ , which we denote as "school" since in these time periods one invests full-time; (2) working, which defines the time periods when one both works and invests; and (3) retirement, which denotes the time periods when one ceases investing completely. We are concerned with Phase 2 since this depicts the only time periods we can observe earnings.

This maximization yields a nonlinear (in the parameters) earnings function<sup>16</sup>

$$\begin{aligned} Y_t = & W^{\frac{1}{(1-b)}} \left[ \left( \frac{1}{\delta} + \left( E^{1-b} - \frac{1}{\delta} \right) e^{\delta(b-1)t^*} \right)^{\frac{1}{(1-b)}} - \left( \frac{1}{\delta} \left[ \frac{b}{r+\delta} \right]^{\frac{b}{(1-b)}} \right) e^{\delta(t^*-t)} \right] \\ & + \left\{ \frac{1}{\delta} \left[ \frac{b}{r+\delta} \right]^{\frac{b}{(1-b)}} \left( 1 - \frac{b\delta}{r+\delta} \right) \right\} + \left\{ \left[ \frac{b}{r+\delta} \right]^{\frac{1}{(1-b)}} \frac{1}{(1-b)} e^{(r+\delta)(t-N)} \right\} \\ & - \left\{ 0.5 \left[ \frac{b}{r+\delta} \right]^{\frac{1}{(1-b)}} \left( \frac{1}{(1-b)} \right) \left( \frac{b}{(1-b)} \right) e^{2(r+\delta)(t-N)} \right\} + \epsilon_t \end{aligned} \quad (7)$$

where  $W = \beta R^{1-b}$ ;  $E = \frac{E_0}{\beta^{\frac{1}{(1-b)}}}$ ;  $t^*$  is the age at which the individual graduates from school;  $N$  is the anticipated retirement age which we take as 65, a reasonable assumption for this cohort; and  $E_0$  is the human capital stock when training begins. In reality parents begin training their children at (or prior to) birth, but for our purposes we consider period 0 to begin when the child starts

<sup>15</sup> We define  $t=0$  to be the time when one begins full-time schooling because we have no data on individual investments prior to school.

<sup>16</sup> Appendix A contains the derivation. Note this specification differs from Haley's because in our derivation we assume a two-term Taylor expansion for the third term in Haley's earnings function. Our specification enables us to identify skill depreciation, which Haley's specification could not do. Importantly, as is shown in Appendix A, this identification does not arise from approximating Haley's third term, nor does it introduce an endogeneity bias by introducing a non-zero truncation error.



formal schooling because this is the point we know children begin learning full-time. Given likely measurement error in  $Y_t$  and the influence of other unobservable factors, we add a time varying error term  $\varepsilon_t$  for each individual.

We fit (7) separately for each individual in the NLSY79 to estimate  $b, W, E, \delta$ , and  $r$ . The dependent variable is the individual's weekly earnings (adjusted to 1982-84 dollars). The independent variable is current age,  $t$ ; and  $t^*$  denotes the age when one completes school. As will be explained later, we use data only on those individuals who completed school, and thus we do not consider individuals with school intermittent trajectories. In principle,  $t^*$  can be solved in terms of the ability, time-discount, and depreciation parameters  $E_0, b, \beta, r$ , and  $\delta$ . Inserting this solution into (7) would complicate the specification as well as omit valuable information we already have for  $t^*$ . Further, including  $t^*$  does not result in an endogeneity bias when estimating (7) person-by-person because  $t^*$  is a constant, and hence uncorrelated with the error term  $\varepsilon_t$ . Instead, later in the paper, we show how school level varies across individuals based on the above earnings function parameters we estimate. Our dataset has 1928 individuals. Thus, we run 1928 regressions, to obtain parameter estimates for each individual.

One point about  $E_0$  is noteworthy before we describe how we estimate (7). In the formal model (see Appendix A),  $E_0$  corresponds to human capital stock when one *begins* specialization, that is when one begins school. One can also derive estimates for potential earnings when one just begins work. We do so by defining  $E_S$  as the amount of human capital when one just completes school.  $E_S$  is computed by augmenting  $E_0$  by the human capital produced during each year of school. Multiplying  $E_S$  by the rental rate per unit of human capital yields potential earnings. Of course, at this stage of the life-cycle, potential earnings exceed actual earnings because individuals are still heavily investing in human capital, though not full-time (on-the-job training). Later in life, the gap between potential earnings and actual earnings should diminish as the proportion of available time spent investing declines. Later in the paper, when presenting our empirical estimates, we verify the validity of these predictions.

### 3. Estimation

Fitting earnings function (7) gives rise to four major considerations.<sup>17</sup> First, proving structural identifiability of each parameter is complicated, given the intricate nonlinearity of (7). Second, choosing an efficient optimization routine is important, given the large number of nonlinear regressions and the necessity to achieve global rather than local convergence. Third, we need a method to retrieve individual-specific parameters estimates  $\beta$ ,  $E_0$ , and the population estimate  $R$  from composite term estimates  $E$  and  $W$ . Fourth, specifying standard errors for each parameter estimate is nontrivial given only 15-24 observations per individual. In the following five subsections, we handle each of these four issues as well as test whether  $R$  varies across the population.

#### 3.1 Structural Identification

We prove identification in Appendix B by applying the nonlinear least-squares conditional mean identification criteria (Newey and McFadden, 1986). This requires there to exist a unique set of parameters  $\phi^*$  that minimizes the mean square error,  $E[y - f(t, \phi)]^2$ , given the first moment,  $E[\varepsilon] = 0$ , and the condition  $E[\varepsilon|t] = 0$ , where  $f(t, \phi)$  is the conditional mean of  $y$ ,  $E[y|t] = f(t, \phi)$ . To show this we consider two parameter vectors,  $\phi$  and  $\phi^*$ , each of which minimizes the mean square error. We utilize a proof by contradiction to demonstrate that  $\phi = \phi^*$  thereby implying a unique  $\phi$ .

#### 3.2 Global Convergence and Parameter Restrictions

For estimation we employ the Genetic Algorithm (GA), a recently available parallel processing optimization routine originally developed by Holland (1975).<sup>18</sup> The GA technique optimizes numeric strings using genetic reproduction, crossover, and mutation concepts (Goldberg, 1989). It is more prone to converge at a global optimum compared to Newton-Raphson hill-climbing algorithms which rely on a point-to-point gradient-based search (Dorsey and Mayer, 1995).

---

<sup>17</sup> We use nonlinear least squares, but as an alternative, one can use the shrinkage estimator which improves the standard error at the expense of introducing bias. Whereas the bias is easy to express in linear models, this is not the case in complex nonlinear models. For nonlinear models, additional research remains to analyze standard measures of asymptotic distributional quadratic bias and risk measures (Ahmed and Nicol, 2012).

<sup>18</sup> We use a version of GA written by Czarnitzki and Doherr (2009).

We implement the algorithm to estimate  $\hat{b}_i$ ,  $\hat{W}_i$ ,  $\hat{E}_i$ ,  $\hat{\delta}_i$ , and  $\hat{r}_i$  in (7) for each individual  $i$ .<sup>19</sup> As in the Ben-Porath life-cycle model we restrict the parameter space to positive real numbers and given the Cobb-Douglas production function for human capital, we restrict  $b$  to be less than 1. Further, we restrict the coefficients  $r$  and  $\delta$  to be less than 0.2 and 0.1 respectively, given that  $r$  and  $\delta$  depict rates of time preference and depreciation. Finally, we choose multiple seeds (search procedures) to insure convergence to a global minimum. As will be reported, we achieve convergence for 1868 of the 1928 individuals.<sup>20</sup>

### 3.3 Identification of Individual-Specific $\beta_i$ , $E_{0i}$ , and Population-Wide $R$

To identify  $\beta_i$ ,  $E_{0i}$ , and the population-wide  $R$ , we adopt the following approach: First, we specify  $\beta_i$  to equal  $\beta e_i$  where  $\beta$  is the population average and  $e_i$  is the individual deviation. Second, we rewrite (8) as

$$W_i = R^{(1-b_i)} \beta e_i. \quad (8)$$

Taking the logarithm, yields

$$\ln W_i = (1-b_i) \ln R + \ln \beta + \ln e_i. \quad (9)$$

Assuming  $E[\ln e_i] = 0$ , one can fit (9) using each individual's values for  $\hat{W}_i$  and  $\hat{b}_i$  obtained from estimating (7) to identify the population value of  $R$  (the coefficient of  $(1-b_i)$ ), the average  $\beta$  (the constant term), and individual-specific values of  $\beta_i$  obtained by taking the anti-log<sub>e</sub> of the

---

<sup>19</sup> We assume iid residuals based on the results of a Durbin-Watson serial correlation test for the nonlinear least squares model (White, 1992). Only six cases exhibit positive autocorrelation. The absence of significant autocorrelation for any one individual (5% significance level) allows us to abstract from possible persistent macroeconomic shocks often considered in the earnings dynamics literature. Were there any persistent wage shocks, one would need to modify (7) to accommodate the serial correlation. Pooling the residuals over all individuals yields a Durbin-Watson statistic of 1.199 based on Bhargava et al. (1982). On the other hand, and even more importantly, the Durbin-Watson statistic based on rerunning (7) for all individuals pooled (i.e., for the whole sample combined rather than individual-by-individual) yields a Durbin-Watson statistic of only 0.401, suggesting far more serial correlation when one does not adjust for heterogeneity. This result (i.e., the difference in Durbin-Watson values between the individual-specific and the pooled regressions) motivates our analysis on earnings dynamics, later in the paper.

<sup>20</sup> Those for whom we do not achieve convergence have a higher standard deviation in weekly earnings (\$263 versus \$225), lower schooling, (12.3 versus 12.9 years), and lower AFQT scores (34 versus 41). The more erratic earnings in these observations probably cause us not to achieve convergence.

sum of the latter two terms in (9).<sup>21</sup> Utilizing  $b_i$  and  $\beta_i$  values along with the coefficient

$$\hat{E}_i = \frac{E_{0i}}{\beta_i^{1/b_i}}$$

obtained from estimating (7) yields individual-specific  $E_{0i}$ .

Heckman, Lochner and Taber (1998) adopt an alternative identification strategy to determine  $R$ . Their approach exploits the fact that all observed earnings changes (adjusted for hours) between two time periods must be attributed to rental rates changes when in “flat periods” human capital stock ( $E_t$ ) remains constant. Typically, flat spots occur late in life, usually around the mid-fifties, an age greater than any current respondent in the NLSY. As will be shown in Section 5.2, Bowlus and Robinson (2012), who apply the flat spot identification approach with CPS data, obtain similar results to ours.

### 3.4 Variation in $R$ across the Population: A Test of Human Capital Homogeneity

Equation (9) can be modified to test for human capital homogeneity. Homogeneity implies each basic human capital unit rents for a common price,  $R$ , determined in the market. Under homogeneity,  $R$  is the same across all occupations, all education levels, and all other characteristics. In short, earnings differ across individuals in the amount of human capital acquired, not because remuneration for each unit differs. On the other hand, heterogeneity implies rental rates per unit of human capital can vary if the market rewards each type of human capital differently. We test for homogeneity by determining whether rental rates differ across segments of the population. Human capital is homogeneous if rental rates are constant. Human capital is heterogeneous if rental rates differ. Obviously, discrimination, incomplete information, and “non-market” effects will weaken the test. Also, economy-wide aggregate demand shocks (for example, as manifested by the unemployment rate) can influence  $R$ , since supply and demand fluctuations can affect spot market prices for all human capital equally.<sup>22</sup>

---

<sup>21</sup> We identify  $R$  because we estimate  $W_i$  and  $b_i$  for each individual, whereas others simply estimate population averages (albeit for a different specification). As a result, we can estimate (9).

<sup>22</sup> We have not examined detailed fields of study such as engineering or medical specialties, which could yield different types of human capital.

Let  $X_i^*$  depict a vector of individual, regional, and job characteristics. Define  $X_i = (1, X_i^*)$ , and  $A$  to be the corresponding vector of coefficients. Rewrite (9) as

$$\begin{aligned} \ln \widehat{W}_i &= (1 - \widehat{b}_i)X_i A + \ln \beta + \ln(e_i) \\ &= (1 - \widehat{b}_i)A_0 + (1 - \widehat{b}_i)X_i^* A_1 + \ln \beta + \ln(e_i). \end{aligned} \quad (9')$$

Here  $\ln R = A_0 + X_i^* A_1$ . A statistically *insignificant*  $A_1$  is consistent with homogeneity.

### 3.5 Precision of the Estimates: Bootstrapped Standard Errors

We construct paired bootstrapped standard errors to get at the precision of our estimates (MacKinnon, 2006).<sup>23</sup> We run 200 regressions for each individual with randomly drawn (with replacement) bootstrapped samples of size equal to the number of observations.<sup>24</sup> From these, we construct the bootstrapped standard errors for each parameter.<sup>25</sup>

## 4. The Data

We utilize the 2012 NLSY79, which contains up to 24 years of data for each respondent through 2010.<sup>26</sup> We do not apply sampling weights since we are estimating (7) for each individual separately.<sup>27</sup> To estimate (7) we use data on age ( $t$ ), the age at which one leaves school ( $t^*$ , defined as schooling plus five), and weekly earnings (annual earnings divided by number of weeks worked) deflated using the 1982-84 urban CPI as base.<sup>28</sup> Because our earnings function specification is designed for those who work continuously, we concentrate only on males given that females are more likely to have discontinuous labor force participation, making the

---

<sup>23</sup> The absence of autocorrelation allows us to compute bootstrap standard errors with i.i.d. residuals. In the presence of significant serial correlation, simple bootstrapping may not work well (Li and Maddala, 1996). In such a case, one could apply alternative bootstrapping methods, such as moving block bootstrapping, to get correct standard errors.

<sup>24</sup> Chernick (2007, p. 89) reports results adopting a similar procedure for a two-parameter nonlinear least-squares regression, but with 20 bootstrap replications on 51 observations.

<sup>25</sup> In considering examples with a large number of parameters ( $N$ ) and a small sample ( $n$ ), Fan et al. (2007) show that for cases of  $t$  simultaneous tests, a necessary and sufficient condition for achieving asymptotic-

level accuracy is  $\ln N = o(n^{\frac{1}{2}})$ . In our case, given 5 parameters,  $\ln(5)=1.6 < 3.9-4.9$ , based on a sample of 15-24. Our sample ( $n$ ) spans 40-60% of an individual's 40-year work life, thereby including a large portion of the maximum number of observations possible.

<sup>26</sup> A description of the data is available at: <http://www.bls.gov/nls/>.

<sup>27</sup> We use the sample weights when we aggregate the results to get inferences about particular segments of the population.

<sup>28</sup> CPI data were obtained from Table 24 (Historical Consumer Price Index for All Urban Consumers), p. 73 (<http://www.bls.gov/cpi/cpid1408.pdf>).

measurement of experience (t) more difficult and resulting in a far more complex nonlinear earnings equation (Polachek, 1975). Further, we use data only on individuals that have completed school because those working while in school (or those working with the intention of going back to school) earn less than commensurately schooled individuals who completed their education (Lazear, 1977). This approach avoids measurement errors associate with intermittent school trajectories.

A main virtue of the NLSY79 is the information on ability, indicators of personality, as well as family background, all of which are independent of our estimated human capital production function parameters. Of these we concentrate on the 1980 AFQT<sup>29</sup> and its particular cognitive skills components (general science, arithmetic reasoning, numerical operations, and math knowledge), craftsmanship skills (mechanical skills, electronics knowledge, coding speed<sup>30</sup>, and automobile repair knowledge), indicators of personality (Rotter locus of control score, Rosenberg self-esteem score, and the CES-D depression scale), family background (mother's and father's schooling, father's occupation, living in an urban area, lived with parents at age 14, household poverty at age 14, household had magazines at age 14), and outcome measures (years of schooling completed and indicators of mental health status). We compare these ability, personality, background and outcome measures respondent-by-respondent to the Ben-Porath parameters we estimated using (7) and (9).

## 5. Estimation Results

We use non-linear least-squares to evaluate (7) for each person with 15-24 years of data.<sup>31</sup> We employ an algorithm (denoted as GA) used in genetic research (Czarnitzki and Doherr, 2007) which is less susceptible to getting stuck at local optima than traditional gradient based optimization techniques. We estimate five crucial parameters: an ability parameter ( $\hat{b}_i$ ), the discount rate ( $\hat{r}_i$ ), the human capital depreciation rate ( $\hat{\delta}_i$ ), and the composite parameters

$\hat{E}_i = \frac{E_{0i}}{\beta_i^{1/\hat{b}_i}}$  and  $\hat{W}_i = (R^{1-\hat{b}_i})\beta_i$ . Table 1 contains average estimates for the entire sample as

<sup>29</sup> The 1980 AFQT score differs slightly from the 1989 and 2006 scores because of the way each component is weighted.

<sup>30</sup> Some interpret this as an indication of motivation, a personality trait. See Segal (2012).

<sup>31</sup> Most individuals have 19 or more years of data.

well as for blacks and whites separately.<sup>32</sup> The mean  $b$  and  $W$  values for whites exceed those of blacks, whereas the mean black time discount rate ( $r$ ) and  $E$  exceed white values. Also, we find the mean black human capital depreciation rate ( $\delta$ ) exceeds the mean white human capital depreciation rate. We shall discuss the implications of these parameter values shortly, but first we address statistical precision which, as already mentioned, we compute via bootstrap techniques.<sup>33</sup> For this, we run 200 nonlinear regression replications utilizing a randomly drawn sample (with replacement) equal to the number of observations available for each person.<sup>34</sup> Mean values of the coefficient standard errors averaged across all individuals are given in row (2) of each panel. Median values which deemphasize outliers are given in row (3). On average, most observations contain coefficients that are statistically significant with the exception of  $E$  and  $r$ , for which the parameter distributions are more right-skewed.

Three factors can affect the accuracy of these estimates. First, estimation equation (7) is based on a Taylor approximation of (A-13). We argue (Appendix A) the approximation error is small. However, we did not show the error has no influence on the actual estimates. Second, the sample size for each individual is limited to between 15 and 24 observations. Based on Fan et al. (2007) we claim sufficient asymptotic accuracy, but we do not test for this. Finally, third, one might question the accuracy of the nonlinear optimization routine used to estimate the parameters.

Simulation techniques constitute one approach to evaluate the extent of these potential biases.<sup>35</sup> To implement this simulation, we randomly pick 100 individuals. For each individual, we draw 100 samples of size 15-24. These samples are based on prediction from equation (A-13) derived from coefficient estimates of equation (7). Appendix C contains more detail. We compute the bias, variance, and root-mean-square-error for each of the five parameters for each individual. These are plotted in Figures C.1 to C.3 in Appendix C. With the exception of a small number of outliers, these values center around zero. Finally, we compute the overall mean and variance of these coefficient biases over all 10,000 observations (Table 2) and test whether these biases are

---

<sup>32</sup> These estimates do not include 42 individuals for whom 1980 sampling weights were not provided in the NLSY data. Also, we do not present estimates for Hispanics. However, the Hispanic parameters are generally between whites and blacks in magnitude.

<sup>33</sup> As for goodness of fit, a pseudo  $R^2$  measure for the entire population is 0.81, computed as  $R^2 = 1 - \frac{\sum_i \sum_t (y_{it} - \hat{y}_{it})^2}{\sum_i \sum_t (y_{it} - \bar{y})^2}$ .

<sup>34</sup> These computations took 281 hours using an i7-vpro chip parallel processor computer running seven STATA programs in tandem, each utilizing the GA algorithm.

<sup>35</sup> We thank an anonymous referee and the editor of this journal for suggesting such a simulation technique. Indeed Eisenhauer, Heckman, and Mosso (2013) adopt this approach to evaluate alternative tuning parameters when estimating simulated method of moments models.

different from zero. The t-tests for each of the five standardized biases are given in Appendix Table C.2. Four are insignificantly different from zero at the 5% level.

Based on the identification strategy we described earlier, we get values for  $b_i$ ,  $\beta_i$ ,  $E_{0_i}$ ,  $E_{s_i}$ ,  $\delta_i$ , and  $r_i$ , as well as a population-wide value of  $R$ . Mean values of the individual-varying parameters are given in Table 3.<sup>36</sup>

### 5.1 Consistency with Prior Population-Wide Estimates

Mean values of our parameters compare favorably to past studies that estimate aggregate Ben-Porath based models, though understandably there are differences due to alternative methodologies and data. For example, we obtain an  $r$  of 0.041 compared to Haley's (1976) 0.055. We obtain a mean  $b$  of 0.35 compared to Haley's 0.58, Heckman's (1975) 0.67, Heckman's (1976) 0.51-0.54, Heckman et al.'s (1998) 0.80, Song and Jones's (2006) 0.5, and Liu's (2009) 0.52, and we obtain a  $\delta$  of 0.027 compared to Johnson and Hebein's (1974) 0.022 and Heckman's (1976) 0.04-0.07. Of course, our results are based on weighted averages of individual values whereas the other studies examine one function for the population as a whole. Further, each uses slightly different human capital production functions, and some incorporate life-cycle labor supply.<sup>37</sup>

Our results are also consistent with computations of Mincer's "time-equivalent" post-school investment as well as with Ben-Porath's declining time-equivalent investment. Figure 1 plots potential and actual earnings for individuals who began work immediately following school.<sup>38</sup> Actual earnings come from the data, and as such are observed for each person. Potential earnings are computed by multiplying predicted human capital stock ( $E_s$ ) by the population-wide market rental rate per unit of human capital stock ( $R$ ), both of which are parameter estimates. Innate to the model, potential earnings exceed actual earnings; and one can see this to be the case by comparing the two distributions. The mode for actual weekly earnings is approximately \$100 per week (in 1982-84 dollars) and the modal value for potential earnings is about \$250. The ratio

---

<sup>36</sup> To conserve space, summary statistics for each of the ability, personality, family background and outcome measures contained in the NLSY79 are available upon request.

<sup>37</sup> See Browning, Hansen, and Heckman (1999) for a survey describing the results from a number of such studies.

<sup>38</sup> These exclude those with very low schooling levels and those who took a year or more to find their first job.



implies a “time-equivalent” investment for new entrants to be about 0.60 which compares favorably to the 0.7 range based on Mincer’s original earnings function regressions.<sup>39</sup> Re-computing these two distributions for older workers (Figure 2) shows a definite narrowing of the distance between potential and actual earnings. In short, according to life-cycle theory, older workers reinvest less of their existing human capital as they age, and this is what our estimates show.

## 5.2 Homogeneity of Human Capital

Table 4 presents results from estimating (9) and two versions of (9’). In each,  $A_0$  is 2.7 implying a rental rate in 1982-84 dollars per week of about \$15.  $LnR$  does *not* vary significantly based on cohort or occupation. In column (2), it varies by 0.012 ln points per year of school (but only at the 10% significance level), by as much as 0.12 ln points by race, by 0.06 ln points in urban areas and by -0.03 ln points per week of unemployment spell. Of these, only the unemployment rate and race remain statistically significant after adjusting for personality and AFQT in column (3). As such, we find that race explains only about 4% of the human capital rental rate, but this possibly includes discrimination. The unemployment coefficient remains about the same value (-0.03). This negative coefficient is consistent with an economy-wide response to aggregate demand shocks, rather than negating homogeneous human capital. The positive AFQT and Pearlin coefficients (in column (3)) might be consistent with better labor market search and/or matching.

Interestingly, despite their different identification strategy, Bowlus and Robinson (2012) find a “close correspondence [in rental rates] ... for such diverse educational groups as high school dropouts and college graduates” and that dropouts tend “to suffer larger price [rental rate] declines in recessions” (page 3498). However, our approach is more general in that we can relate rental rates to individual characteristics. As such, we show (column 3) that the differences in rental rates attributable to schooling (in column 2) may be due to ability and personality, rather than school, per se.

## 5.3 Heterogeneity of the Ben-Porath Parameters

---

<sup>39</sup> One obtains 0.56 and 0.81 respectively when one solves for  $k_0$  (the equivalent of our  $E_0$ ) using Mincer’s (1974) Gompertz specification G(2a) and G(2b), p. 92.

One way to examine population heterogeneity is to plot kernel density distributions of our estimated parameters. Also, it is instructive to compare our estimates to the distribution of AFQT reported in the NLSY79, especially given we have three ability measures ( $b$  and  $\beta$  representing abilities to learn, and  $E_0$  representing ability to earn). For ease of comparison, we scale each of our parameters because each has a different measurement range and metric. For example,  $b$  ranges from 0.01 to 0.70 with a mean of .35 and  $\beta$  varies from 0.13 to 1.59 with a mean of 0.64.

Thus we scale each parameter by  $\frac{x_i - L_i}{H_i - L_i}$  where  $L_i$  is the lowest value and  $H_i$  is the highest. This

POMP (percent of maximum possible) score yields a scaling between zero and a hundred, where  $i$  indexes each parameter. Figures 3 plots the kernel density functions for AFQT,  $b_i$ ,  $\beta_i$ ,  $E_{0_i}$ ,  $E_{S_i}$ , and  $\delta_i$  for blacks and whites.<sup>40</sup> Ability parameters  $b$  and  $\beta$  are relatively bell-shaped. Parameters  $E_0$  and  $r$  are skewed to the right, and the skill depreciation parameter  $\delta$  appears to have a double peak. Generally blacks (solid line) fare worse than whites (dashed line). Kolmogorov-Smirnov tests for the difference in these ability distributions are given in Table 5 (rows 1-7). The race differences for each distribution (with the exception of  $r$ ) are significantly different statistically, but the distance measure is largest for the AFQT.<sup>41</sup>

For comparison purposes we also utilize the Kolmogorv-Smirnov test to determine the similarity of black and white distributions for each other NLSY79 ability, personality, and background variable used in the remainder of the paper. With the exception of the Rosenberg Self-Esteem Measure, and the Pearlin Mastery Score, all are statistically different. Moreover, the distance measures of all the ability test scores (AFQT and the ASVAB tests) are larger than our ability measures ( $b$ ,  $\beta$ , and  $E_0$ ) and the personality indicators (Rotter Locus of Control, Rosenberg Self-Esteem Score, and the CES-D 20 depression index). Race differences in the distance measure for poverty is between the AFQT and ASVAB test scores, our ability measures, and the personality indicators. In short, the AFQT and ASVAB test scores appear to accentuate race differences compared to our and other measures.

#### 5.4 Consistency with Human Capital Theory

---

<sup>40</sup> The AFQT test scores are given in percentiles ranging from 1 to 99. We compute raw scores based on appropriately summing the scores for each component part. We then scale these as indicated above. These rescaled scores are contained in Figure 3. To conserve space, we do not plot racial differences in the distributions of other ability, personality, or family background variables.

<sup>41</sup> Note the Kolmogorov-Smirnov test is independent of units. As such, one can compare the distance measure across variables. The greater the distance, the larger the relative race difference.

Our estimated ability, skill depreciation, time preference measures, as well as the AFQT scores are related to schooling in a predictable way. These patterns are presented in Table 6 which contains a specific regression. The dependent variable is  $\log_e$  years of school completed. The independent variables are  $b$ ,  $\beta$ ,  $E_0$ ,  $\delta$ , and  $r$  (also in logs so that the effects of each can be compared in percent terms), cohort, race, and background variables. The ability to learn parameters  $b$  and  $\beta$  have the largest positive coefficients. Human capital theory predicts a positive correlation between this type ability and schooling level because a higher ability to learn raises the amount of human capital one can produce per unit of time. Holding rental rates per unit of human capital constant, producing more human capital per time unit lowers the opportunity cost of going to school, thereby increasing the amount of school purchased. On the other hand, schooling levels do not rise with the ability to earn parameter,  $E_0$ . This is expected because an individual's higher initial human capital substitutes for schooling, and as Ben-Porath (1967) predicts, leads one to stop school earlier.<sup>42</sup> Schooling decreases with skill depreciation. This result also is as expected since a higher depreciation rate lowers the value of what is learned because more is "forgotten". Finally, schooling levels decrease with the estimated time discount rate ( $r$ ). This latter result, too, is noteworthy because higher time discount rates should imply *fewer* years of schooling. Individuals with high discount rates are more reluctant to put off the gratification of current market earnings, given that they discount the future heavily.

### **5.5 How the Ben-Porath Parameters Vary by Cognitive Ability, Personality, and Family Background**

One criticism of the Mincer (1958) framework is it does not explain why people choose a particular level of education. In his model people are indifferent between various levels of school because all levels yield the same lifetime earnings. Individuals are homogeneous in every respect except levels of schooling. No explanation is given why people vary in the number of years of schooling they undertake.

The Ben-Porath model argues that ability, depreciation and time preference affect human capital investments, and hence levels of school attained. As we illustrated above, by examining each individual separately, we obtain individual specific ability, depreciation, and time preference

---

<sup>42</sup>  $E_s$  is positively related to schooling since by construction it incorporates what is learned in school.

parameters. These are related to years of schooling in a predictable way, as was shown in Table 6. But this begs the question why, in the first place, these ability and time preference parameters differ from person to person. In this section we examine this issue. We ask what personality and background factors are associated with the various Ben-Porath parameters (skill depreciation, time preference and ability to learn and earn) we estimated. Getting at this question gives some indication of the relative roles of nature versus nurture, and may serve as the underlying reasons why these parameters, which are related to years of school, differ across the population. In short, innate ability (e.g., quantitative versus verbal test scores), personality traits, and family background might determine the Ben-Porath parameters that predict years of school. At the same time, our estimated Ben-Porath parameters can be used to test hypotheses psychologists raise about ability, intelligence and personality. In the next three subsections, we examine each of these issues.

### **5.5.1 Ben-Porath Parameters and ASVAB Test Scores**

Psychologists define intelligence as a “general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience” (Gottfredson, 1997a, p. 13). General intelligence is often referred to simply as “g”. According to De Young (2011, p. 3) “the most widely used distinction between abilities, at the level of the hierarchy immediately below *g*, is ... *fluid* and *crystallized* intelligence (Horn & Cattell, 1966).” Fluid intelligence is supposed to be innate and crystallized is supposed to be learned, in essence knowledge based. However, according to De Young, recent factor analysis by Johnson and Bouchard (2005a and 2005b) finds crystallized intelligence to be mostly verbal and fluid intelligence to be mostly nonverbal so that “most past findings regarding fluid and crystallized intelligence ... can be translated cleanly into a verbal-nonverbal framework” (p. 3).

We find some evidence that this is the case with our ability estimates. Table 7 presents a correlation table between our ability estimates and ASVAB test scores. The top panel presents correlations for cognitive skills and the lower panel for what we call craftsmanship skills. Rows 1-3 of the top panel give these correlations for math, row 4 for science and rows 5-6 for verbal scores. Almost without exception, math abilities are more highly correlated with  $b$ ,  $\beta$  and  $E_0$  than are the verbal test scores. As such, our measures more represent innate fluid ability rather than crystallized or learned ability. This makes sense because  $E_0$  depicts one’s ability to earn just

*before* one begins school, and  $b$  and  $\beta$  reflect one's ability to create new human capital from old, again independent of human capital stock. The correlation between ability and craftsmanship skills are the lowest which also makes sense since these skills are more learned. Also, we find the more able depreciate less quickly (forget less) and have lower time discount rates.

### **5.5.2 Ben-Porath Parameters and Personality**

McAdams and Pal (2006, p. 212) define personality as “encompassing dispositional traits, characteristic adaptations, and integrative life stories, complexly and differentially situated in culture.” Whether personality and ability are related is still debated in the psychology literature. Eysenck (1994) argues that personality and intelligence are unrelated. Chamorro-Premuzic and Furnham (2005) claim that both are related but categorically distinct. DeYoung's (2011, p. 6) review of current research “rules out the possibility that intelligence is unrelated to personality.”

Table 8 presents correlation coefficients for four psychological indicators. The Rotter score (varying from 0 to 16), which denotes a person's external locus of control, defines the extent one views life chances to be determined by external factors. The Mastery scale defines the extent individuals can control the environment where they operate. Scores vary from 7 (low mastery) to 28 (high mastery). The Rosenberg Self-Esteem Scale (Rosenberg, 1965) is a measure of self-worth based on how a respondent answers ten questions regarding self-acceptance. The CES-D scale measures symptoms associated with depression. The empirical results indicate a negative correlation between external locus of control and ability to learn, a positive correlation between mastery and ability to learn, a positive correlation between self-esteem and ability to learn, and a negative relation between depression and ability to learn. Self-esteem is positively related to one's ability to earn, but locus of control, mastery, and depression are not. Those with high external locus of control, those with low mastery scores, and those exhibiting depression have high discount rates and high skill depreciation rates.

### **5.5.3 Ben-Porath Parameters and Family Background**

Family background is also correlated with ability. Table 9 indicates a positive correlation between parental education, father's success (being in a professional or managerial occupation) and one's ability to learn. Similarly these parental background variables exhibit an inverse correlation with

time-discount and skill depreciation. Having had magazines in the home at age 14 are also positively correlated with the ability to learn and negatively correlated with skill depreciation and the time discount rate. The correlation with poverty during childhood (in year 1978) is the opposite. To the extent learning goes on in the home, these results are consistent with parental investments in children’s human capital, as the correlations mimic the results obtained from AFQT and ASVAB ability measures, exhibited earlier in Table 8.

We observe no relation between living in an urban area and our ability measures. This is an expected result to the extent being in an urban area has no relationship to human capital investments and no relationship to ability.

## 5.6 Earnings Dynamics Implications

Examining the autoregressive structure of residuals has implications regarding earnings dynamics. To see this, define two sets of residuals. The first ( $\varepsilon_{it}^I$ ) is obtained from estimating (7) individual-by-individual. Here  $\varepsilon_{it}^I = Y_{it} - f_i(t, t^*; \hat{b}_i, \hat{W}_i, \hat{E}_i, \hat{\delta}_i, \hat{r}_i)$ . The second ( $\varepsilon_{it}^S$ ) is obtained by re-estimating (7) for all individuals pooled (i.e., for the whole sample combined rather than individual-by-individual). Define this latter residual to be  $\varepsilon_{it}^S = Y_{it} - f(t, t^*; \hat{b}, \hat{W}, \hat{E}, \hat{\delta}, \hat{r})$  where  $f(\cdot)$  is un-subscripted because it now depicts a population-wide estimate. The mean first-order autoregressive coefficient,  $\rho_i^I$ , based on  $\varepsilon_{it}^I$  is 0.26. However,  $\rho_i^S$  based on  $\varepsilon_{it}^S$  is 0.80. The difference is statistically significant at  $\alpha < .01$ . This stark difference highlights how parameters estimating individual reactions to earnings shocks can change when adjusting for individual-specific heterogeneity.

A growing earnings dynamics literature decomposes earnings in terms of permanent and transitory components. The most minimal specification depicts earnings simply as the sum of these two components so that  $Y_{it} = a_i + v_{it}$ , but typically one augments this specification in a number of ways. These modifications include introducing explanatory variables to account for earnings levels, allowing the relative importance of the error components to vary with calendar time, introducing persistence in the transitory shock via ARMA processes, permitting the permanent component to evolve over time, and adjusting for heterogeneity in various ways. One question arising in this literature is how allowing for heterogeneity affects the estimation of these shocks. Clearly as Krueger et al. (2010) indicate “what may pass as a permanent shock may sometimes be heterogeneity in disguise,” and responses to transitory shocks can vary with

measured and unmeasured individual characteristics. Indeed a number of recent studies concentrate on heterogeneity by allowing ARMA processes to vary across individuals (e.g., Browning and Ejrnæs, 2013). This question is important because earnings dynamics parameters have been applied to a number of policy related economics issues. These include consumption and savings (Browning, Hansen, and Heckman, 1999), schooling choices (Cunha, Heckman, and Navarro, 2005), poverty (Lillard and Willis, 1978), and more. Our results can be used to help address the question of how heterogeneity affects the earnings dynamics error structure parameters.

A number of studies present decile ranges of key parameters illustrating that heterogeneity affects the speed individuals respond to shocks (e.g., Browning, Ejrnæs, and Alvarez, 2010; and Browning and Ejrnæs, 2013). However, it is also possible that heterogeneity manifests itself in the way individuals accumulate human capital over their lives. Varying human capital accumulation trajectories, in turn, can alter the way one responds to shocks. In addition, this type heterogeneity might also alter what one actually perceives to be a shock in the first place. As such, some of what past literature perceives as a permanent shock might reflect person-specific differences in human capital acquisition. Similarly, person-specific differences in human capital accumulation might be perceived as adjustments to transitory shocks. To get at these possibilities, we compare the parameters obtained from a relatively simple error structure under the two regimes described above. In each case we assume an error structure similar to ones used in a number of published studies surveyed in Doris et al (2013) and Hryshko (2008). Specifically, we assume

$$\begin{aligned}\epsilon_{it} &= \Gamma_t + \mu_{it} \\ \mu_{it} &= a_i + v_{it} \\ v_{it} &= \rho v_{it-1} + \theta e_{it-1} + e_{it}\end{aligned}$$

where, of course, each parameter has an I or S superscript, and where  $\Gamma_t$  depicts year-specific effects,  $a_i$  individual-effects,  $\rho$  the autoregressive component, and  $\theta$  the moving average component of the remaining individual-year specific shock after eliminating year and individual heterogeneity. Clearly the difference in these parameters (e.g.,  $\rho^I$  and  $\rho^S$ ,  $\theta^I$  and  $\theta^S$ , etc.) indicate the effect of heterogeneity on estimated earnings dynamics. In Table 10 we present estimates of these parameters using the residuals computed based on the two-regimes outlined above. Column (1) presents results for the residuals adjusted for heterogeneity, whereas column

(2) provides the non-heterogeneity adjusted parameters.<sup>43</sup> Like most studies, the variance in the transitory shock ( $\sigma_\epsilon^2$ ) trounces the variance in the permanent shock ( $\sigma_a^2$ ). In virtually all studies  $\rho$  is positive, and at 0.92, our estimate of  $\rho$  is in the same ballpark as the 0.956 obtained by Dickens (2000), the 0.847 obtained by Moffit and Gottschalk (1995), and the 0.964 obtained by Sologon and O’Donoghue (2010), all also using an ARMA(1,1) process. On the other hand, our estimate of  $\rho$  for the heterogeneity adjusted model, is 0.44, about half the magnitude obtained in the non-heterogeneity adjusted case. Also, as in most studies,  $\theta$  is negative. For the non-heterogeneity case  $\theta = -0.347$ , which again is comparable to those studies above. For the heterogeneity adjusted case  $\theta = -0.074$ , about  $\frac{1}{6}$  the size of the non-heterogeneity adjusted case. In short, adjusting for heterogeneity dramatically decreases estimates measuring the persistence of transitory earnings shocks. This means previous measures of such responses may indeed be contaminated as “what may pass as a permanent shock may sometimes be heterogeneity in disguise,” and estimated responses to transitory shocks can vary with measured and unmeasured individual characteristics, as Kreuger et al. (2010) speculated.

## 5.7 Selectivity

One might argue that black workers (working 15 or more years) are relatively more able than white workers (working 15 or more years) because only the relatively “better” blacks compared to whites are able to sustain such a long work history. One can assess this bias by utilizing the NLSY79 reported AFQT scores for non-workers of each race. If the “worker” compared to “non-worker” AFQT advantage is greater for blacks than whites, then our Ben-Porath measures overstate black compared to white ability, and as a result understate the racial ability gap. In contrast, if the relative AFQT advantage is greater for whites, then the opposite is true, and as such, we then overstate black-white ability differences. A t-test rejects the hypotheses that the overall difference is unequal. In short, those working 15 or more years tend to be more able than those working less than 15 years (or not at all); but the difference between black “workers” and “non-workers” is not statistically different than for white “workers” and “non-workers”.<sup>44</sup> This result is consistent with small, if any, selectivity biases when considering racial

---

<sup>43</sup> Our empirical analysis of the covariance structure is based on residuals for years 1979-1994. After 1994, data were collected every two years instead of every year.

<sup>44</sup> The actual test comprises an insignificant  $\alpha_3$  white-worker interaction coefficient in the following regression:  $AFQT_i = 21.06 (0.9) + 26.39(1.1)white_i - 1.24(1.5)worker_i + 2.55(1.8)white_i * worker_i + \epsilon_i$ , where the standard errors are given in parentheses.



differences in our estimated Ben-Porath ability parameters obtained by concentrating on blacks and whites working at least 15 years of their lifetime.

## **6. Summary and Conclusion**

As early as the 1990s, Browning, Hansen and Heckman (1999) used parameters of the life-cycle model to illustrate inherent biases in the representative agent framework because the approach fails to account for individual heterogeneity. Recent studies acknowledge this insight about heterogeneity, but none estimate a complete set of Ben-Porath parameters individual-by-individual. To our knowledge, this paper represents the first to do so.

These parameters are important because they can be used to help calibrate macroeconomic models, because they have implications with regard to the distribution of income, because they can be used to test theorems regarding schooling decisions, and because they are related to underlying psychological personality and family background variables. Further, to the extent that personality traits are innate, knowing these parameters and how they relate to personality can shed light on aspects of nature versus nurture.

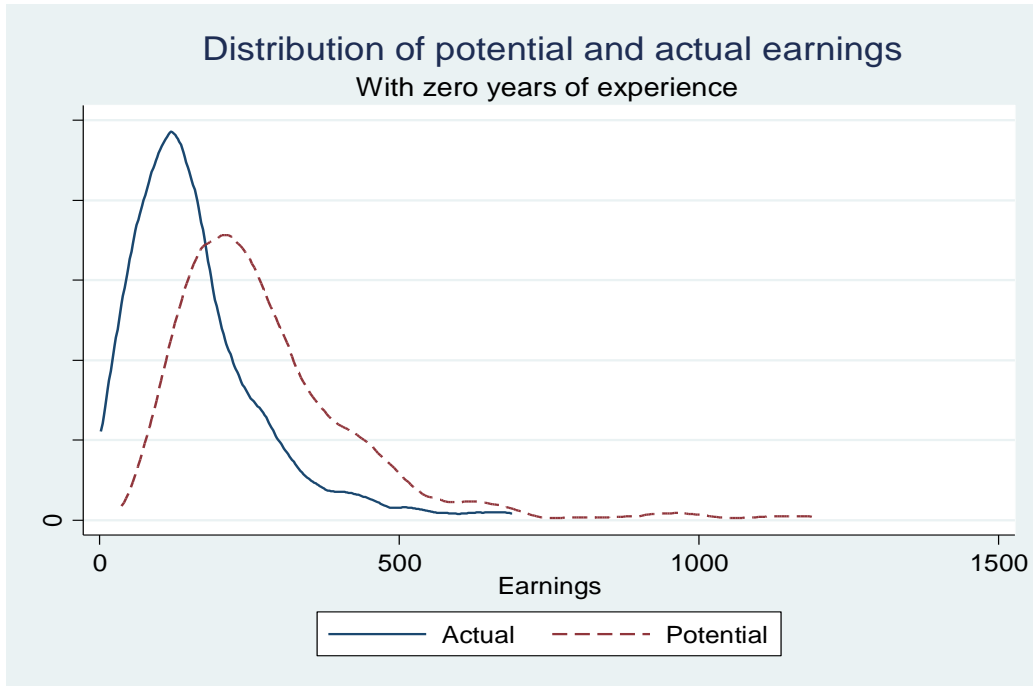
To obtain these parameters, we adopt four methodological innovations. First, we devise a tractable, albeit complex, nonlinear formulation of the Ben-Porath model that enabled us to identify five basic human capital parameters. Second, we prove these parameters are identified based on Newey and McFadden's (1986) nonlinear least-squares conditional mean identification criteria. Third, we estimate these parameters, individual-by-individual, using a novel algorithm. Fourth, we are able to estimate a plausible population-wide human capital rental rate.

A number of new findings emerge from our analysis. First, we find blacks have slightly higher skill depreciation and time discount rates, which could account for one reason blacks obtain less years of school than whites. Second, we confirm important relationships based on the life-cycle human capital model. We find one's ability to learn to be positively correlated with years of school, but not so with one's ability to earn. Further, both a higher discount rate and a greater degree of skill depreciation are associated with fewer years of school. Third, we confirm that ability is related to a number of personality traits and family background variables. For example, a high internal locus of control and a high mastery (belief one controls events) are related to one's ability to learn, but unrelated to one's ability to earn. Also, individuals who score

high on the AFQT exhibit a high ability to learn, and only a marginally higher ability to earn. Finally, our results are relevant to the macroeconomic earnings-dynamics literature. For example, we find the autoregressive and moving average parameters on transitory shocks are more than halved when using our heterogeneity-adjusted residuals.

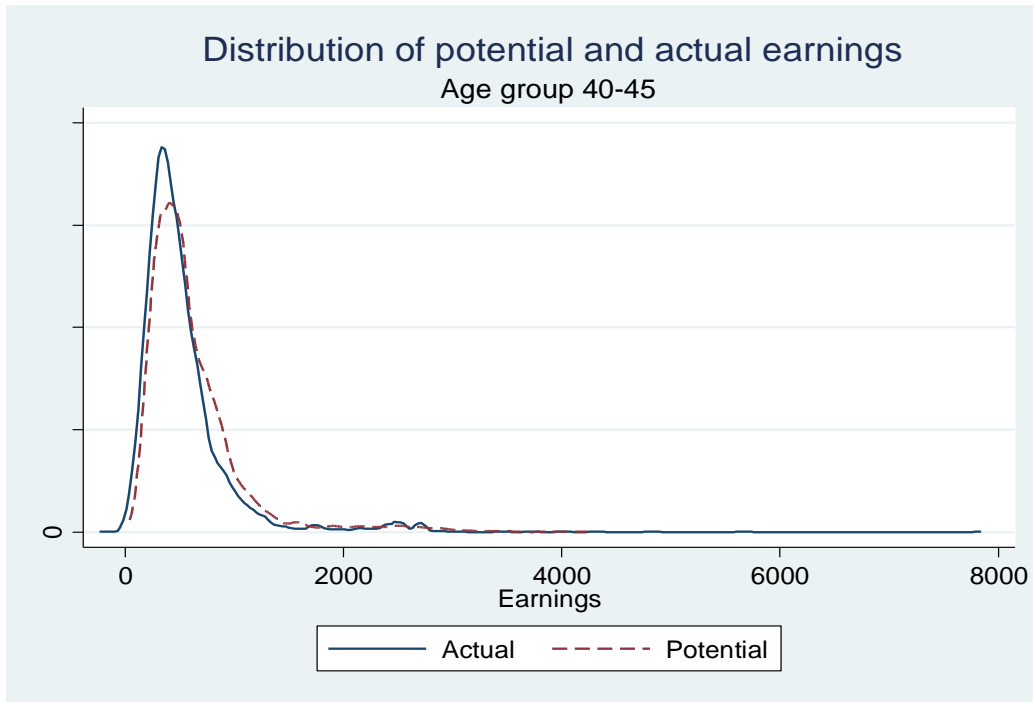
Our results are promising enough to warrant pursuing the approach further. For example, zeroing in on various types of ability might enable one to gain insights into occupational choice decisions including answering questions relating to gender differences in scientific professions. Also, accounting for this heterogeneity can alter how one assesses macroeconomic policy, particularly how individuals respond to economic shocks. Further, linking ability, time discount, and skill depreciation parameters to understand innate own and parental personality characteristics could get at important questions regarding nature versus nurture. In any case, the whole approach gets at heterogeneity in a new way that can be valuable to analyze other aspects of human behavior.

Figure 1\*



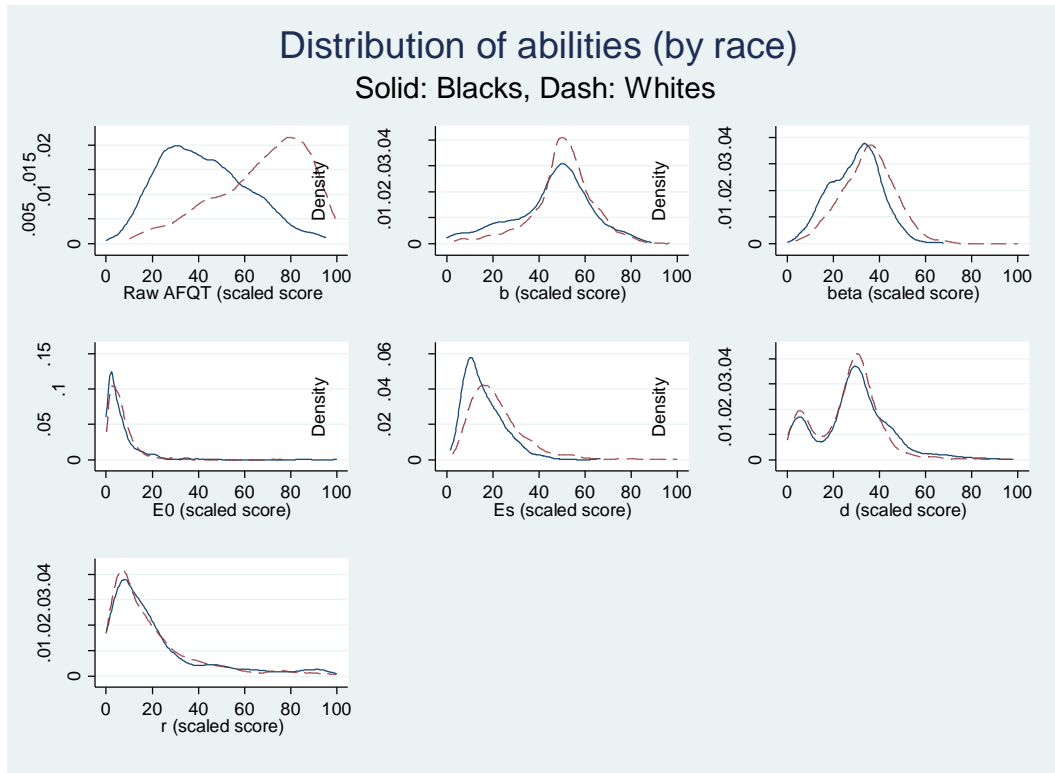
\*Weekly earnings in 1982-4 dollars. Potential earnings computed by multiplying predicted human capital at zero experience ( $E_0$ ) by the population rental rate per unit of human capital ( $R$ ). Actual earnings are from the NLSY79.

Figure 2\*



\*Weekly earnings in 1982-4 dollars. Potential earnings computed by multiplying predicted human capital for 40-45 year olds ( $E_{40-45}$ ) by the population rental rate per unit of human capital ( $R$ ). Actual earnings are from the NLSY79.

Figure 3\*



\* Estimated parameters are scaled from 0-100 to be compatible with AFQT scores. See text for details.

**Table 1: Earnings Function Parameter Estimates\***

	$b$	$W$	$E$	$\delta$	$r$
All (N=1826)	0.351	3.808	5.392	0.027	0.041
<i>Bootstrapped SE(mean)</i>	0.054	0.778	2.98	0.008	0.022
<i>Bootstrapped SE(median)</i>	0.049	0.685	2.39	0.007	0.016
Proportion of observations with sig(5%)	0.943	0.95	0.430	0.792	0.445
Blacks (N=596)	0.323	3.748	6.253	0.029	0.043
<i>Bootstrapped SE(mean)</i>	0.062	0.866	3.470	0.009	0.024
<i>Bootstrapped SE(median)</i>	0.059	0.765	2.58	0.0079	0.0182
Proportion of observations with sig(5%)	0.899	0.930	0.367	0.783	0.398
Whites (N=1230)	0.355	3.817	5.265	0.026	0.041
<i>Bootstrapped SE(mean)</i>	0.049	0.735	2.73	0.0075	0.02
<i>Bootstrapped SE(median)</i>	0.046	0.657	2.33	0.0066	0.0143
Proportion of observations with sig(5%)	0.965	0.96	0.461	0.797	0.469

\*Weighted average coefficients of equation (7) estimated for NLSY79 individuals. Row 1 of each panel gives average parameter estimates of (7) over the entire sample. Rows (2) and (3) give mean and median bootstrapped standard errors. Each observation is weighted by the NLSY79 weights when computing the averages and medians. Average  $R^2$  values for All, Blacks, and Whites are 0.480, 0.397, and 0.492.

**Table 2: Simulated Biases<sup>1</sup>**

Coefficient	Mean Bias	Variance
b	-0.005	0.00520
W	-0.046	0.86421
$\delta$	-0.002	0.000144
r	0.004	0.00083
E	0.897	13.3536

<sup>1</sup> Biases based on 100 simulated regressions of sample size 15-24 for each of 100 individuals randomly chosen from the 1826 individuals of Table 1. See Appendix C for details on how these biases are computed.

**Table 3: Mean Values of Individual-Specific Parameters\***

Parameter	Observations	Mean	SD
<i>b</i>	1826	0.35	0.10
$\beta$	1826	0.64	0.17
$E_0$	1826	2.75	2.79
$E_s$	1826	18.11	10.53
$\delta$	1826	0.03	0.01
r	1826	0.04	0.04
Average weekly earnings (1982-84 dollars)	34872	426.1	348.1
t (age)	34872	32.0	8.48
t* (School leaving age)	1826	18.12	2.9

\* Computed from our estimates of (7) and (9). Parameter definitions are given in text.

**Table 4: Determination of Rental Rates per Unit of Human Capital\***

Dependent variable: log(W)	(1)	(2)	(3)
Independent Variables			
(1-b)	2.719*** (0.0704)	2.694*** (0.193)	2.671*** (0.231)
(1-b)*school		0.0116* (0.00624)	-0.00647 (0.00716)
(1-b)*unemployment		-0.0309*** (0.00376)	-0.0292*** (0.00392)
(1-b)*white		0.116*** (0.0329)	0.0699* (0.0371)
(1-b)*urban		0.0601** (0.0260)	0.0394 (0.0270)
(1-b)*manufacturing		0.0216 (0.0293)	0.0320 (0.0304)
(1-b)*father professional or managerial		0.0504* (0.0293)	0.0326 (0.0304)

		(0.0282)	(0.0292)
(1-b)*cohort-age14in1979		-0.000367	0.0304
		(0.0716)	(0.0753)
(1-b)*cohort-age15in1979		0.0599	0.0709
		(0.0673)	(0.0707)
(1-b)*cohort-age16in1979		0.0473	0.0412
		(0.0674)	(0.0707)
(1-b)*cohort-age17in 1979		0.0699	0.0595
		(0.0670)	(0.0701)
(1-b)*cohort-age18in1979		0.0339	0.0237
		(0.0658)	(0.0690)
(1-b)*cohort-age19in1979		0.0192	0.00982
		(0.0664)	(0.0695)
(1-b)*cohort-age20in1979		-0.00939	-0.0315
		(0.0670)	(0.0701)
(1-b)*cohort-age21in1979		0.0490	0.0368
		(0.0668)	(0.0704)
(1-b)*professional		0.161	0.168
		(0.150)	(0.150)
(1-b)*service		-0.0111	0.0134
		(0.152)	(0.152)
(1-b)*sales		0.181	0.191
		(0.151)	(0.151)
(1-b)*construction		0.102	0.131
		(0.150)	(0.151)
(1-b)*farm		-0.0566	0.0149
		(0.190)	(0.192)
(1-b)*production worker		0.114	0.130
		(0.150)	(0.150)
(1-b)*Locus of control			-0.00582
			(0.00516)
(1-b)*Self esteem			0.00424
			(0.00328)
(1-b)*AFQT, 1980			0.00226***
			(0.000568)
(1-b)*Pearlin mastery score			0.00908**
			(0.00436)
(1-b)*Depression index			0.00159
			(0.00162)
Constant	-0.501***	-0.739***	-0.780***
	(0.0462)	(0.0516)	(0.0534)
Observations	1,826	1,557	1,464
R <sup>2</sup>	0.450	0.516	0.527

\* Estimation of (9) and (9'). The dependent variable is  $\ln \hat{w}_i^*$ . The (1-b) coefficient depicts  $\ln R$ . The other coefficients represent percent deviations associated with the NLSY79 independent variables. See text for an explanation. Standard errors in parentheses, \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . The cohort coefficients are relative to age 22 in 1979.

**Table 5: Kolmogorov-Smirnov Test Comparing the Black and White Distributions of the Indicated Variable\***

<b>Variable</b>	<b>Distance</b>	<b>P-value</b>
<i>b</i>	0.140	0.000
$\beta$	0.199	0.000
$E_0$	0.119	0.000
$E_s$	0.252	0.000
$\delta$	0.097	0.001
<i>r</i>	0.037	0.643
AFQT, 80	0.487	0.000
Locus of control (Rotter)	0.078	0.014
Self-esteem (Rosenberg)	0.022	0.992
Pearlin mastery	0.066	0.063
CES-D 20	0.174	0.000
ASVAB, general science	0.520	0.000
ASVAB, arithmetic	0.458	0.000
ASVAB, word knowledge	0.481	0.000
ASVAB, paragraph comprehension	0.395	0.000
ASVAB, numeric ability	0.331	0.000
ASVAB, coding speed	0.354	0.000
ASVAB, auto shop knowledge	0.597	0.000
ASVAB, math knowledge	0.349	0.000
ASVAB, mechanical knowledge	0.519	0.000
ASVAB, electronics	0.505	0.000
Agreeableness (principal component)	0.223	0.000
Extraversion (principal component)	0.151	0.004
Openness (principal component)	0.533	0.000
Conscientiousness (principal component)	0.132	0.000
Neuroticism (principal component)	0.161	0.000
Mother's years of schooling	0.274	0.000
Father's years of schooling	0.239	0.000
Fathers occupation (if professional/managerial)	0.189	0.000
Urban	0.065	0.064
Poverty	0.303	0.000
Household subscribes magazine (at age 14)	0.316	0.000

\* Computed from our estimates of (7), (9) and data contained in the NLSY79.

**Table 6: Schooling Level as a Function of the Ben-Porath Parameters\***

Variables	Coefficient.	t-value
ln(b)	0.0528	5.81
ln( $\beta$ )	0.0716	4.69
ln( $E_0$ )	0.0058	1.26
ln( $\delta$ )	-0.0342	-6.49
ln(r)	-0.0322	-7.98
Constant	2.3349	68.52
Observations	1701	
Adj. R <sup>2</sup>	0.2356	

\* Dependent Variable: Ln(Completed Years of School) from NLSY79. Independent variables are individual-specific coefficient estimates of (7) and (9). Also adjusted for cohort, race, household poverty in 1978, and whether household subscribed to magazines at age 14,

**Table 7: Correlation: Ben-Porath Parameters and Cognitive and Craftsmanship Skills\***

	<i>b</i>	$\beta$	$E_0$	$\delta$	<i>r</i>
<b>Cognitive Measures</b>					
Arithmetic reasoning	0.222	0.196	0.049	-0.141	-0.043
Sig (p-value)	0.000	0.000	0.039	0.000	0.071
Numeric operations	0.210	0.259	0.051	-0.113	-0.043
Sig (p-value)	0.000	0.000	0.031	0.000	0.071
Math knowledge	0.233	0.208	0.024	-0.133	-0.103
Sig (p-value)	0.000	0.000	0.317	0.000	0.000
General science	0.190	0.175	0.052	-0.129	-0.027
Sig (p-value)	0.000	0.000	0.029	0.000	0.254
Word knowledge	0.186	0.185	0.018	-0.142	-0.052
Sig (p-value)	0.000	0.000	0.450	0.000	0.029
Paragraph comprehension	0.158	0.163	0.033	-0.148	-0.034
Sig (p-value)	0.000	0.000	0.161	0.000	0.153
<b>Craftsmanship Skills</b>					
Mechanical	0.149	0.167	0.043	-0.132	-0.005
Sig (p-value)	0.000	0.000	0.069	0.000	0.838
Electronics	0.137	0.198	0.080	-0.116	0.020
Sig (p-value)	0.000	0.000	0.001	0.000	0.403
Coding speed	0.184	0.199	0.034	-0.125	-0.061
Sig (p-value)	0.000	0.000	0.152	0.000	0.010
Auto shop	0.062	0.187	0.072	-0.094	0.034
Sig (p-value)	0.009	0.000	0.003	0.000	0.151

\* Computed from our estimates of (7), (9) and data contained in the NLSY79. N=1784.



**Table 8: Correlation: Ben-Porath Parameters and Personality\***

	$b$	$\beta$	$E_0$	$\delta$	$r$
Locus of control (Rotter)	-0.120	-0.080	0.000	0.120	0.050
Sig (p-value)	0.000	0.000	0.900	0.000	0.040
Self-esteem (Rosenberg)	0.070	0.120	0.080	-0.080	0.060
Sig (p-value)	0.000	0.000	0.000	0.000	0.010
Mastery score (Pearlin)	0.130	0.120	-0.010	-0.110	-0.110
Sig (p-value)	0.000	0.000	0.820	0.000	0.000
CES-D Depression Scale	-0.101	-0.066	0.024	0.106	0.049
Sig (p-value)	0.000	0.005	0.303	0.000	0.040

\* Computed from our estimates of (7), (9) and data contained in the NLSY79.  $N_{\text{Rotter}}=1807$ ,  $N_{\text{Rosenberg}}=1808$ ,  $N_{\text{Pearlin}}=1790$ , and  $N_{\text{CES-D}}=1790$ .

**Table 9: Correlation: Ben-Porath Parameters and Family Background\***

	$b$	$\beta$	$E_0$	$\delta$	$r$
Mothers schooling	0.162	0.067	-0.039	-0.114	-0.076
Sig (p-value)	0.000	0.006	0.104	0.000	0.002
Father's schooling	0.151	0.127	-0.039	-0.097	-0.100
Sig (p-value)	0.000	0.000	0.115	0.000	0.000
Professional/Managerial Father	0.138	0.088	-0.007	-0.085	-0.063
Sig (p-value)	0.000	0.000	0.764	0.000	0.007
HH poverty, 1978	-0.081	-0.149	-0.049	0.034	-0.037
Sig (p-value)	0.001	0.000	0.042	0.161	0.131
HH had magazine, age 14	0.153	0.099	0.011	-0.134	-0.062
Sig (p-value)	0.000	0.000	0.654	0.000	0.008
Urban	0.022	0.048	0.015	0.016	-0.006
Sig (p-value)	0.340	0.040	0.520	0.500	0.780

\* Computed from our estimates of (7), (9) and data contained in the NLSY79.  $N_{\text{Mother's schooling}}=1721$ ,  $N_{\text{Father's schooling}}=1596$ ,  $N_{\text{Prof/managerial father}}=1826$ ,  $N_{\text{HH pvery}}=1715$ ,  $N_{\text{HH magazine}}=1812$ ,  $N_{\text{Urban}}=1820$ .

**Table 10: Covariance Structure\***

Parameter	Individual-Specific Model	Pooled Model
$a_i$	-1.78 (0.977)	-6.06 (3.311)
$\sigma^2_{a_i}$	0.00	0.01
$\rho$	0.44 (0.007)	0.92 (0.003)
$\theta$	-0.07	-0.35
$\sigma^2_e$	8687.75	7805.66

\*Parameters are based on Earnings Dynamics Model in Section 5.7. We use SAS version 9.2 to compute the ARMA parameters. It reports the estimate of  $\rho$  (the auto regressive parameter). It does not report  $\theta$  (the moving average parameter) directly. Instead it reports  $\gamma$  (a re-parameterization involving  $\rho$  and  $\theta$ ). The values are ( $\gamma$ -pooled=0.8030 (se: 0.004482); and  $\gamma$ -indiv= 0.3753 (se: 0.006783). Based on Fuller (1976), page 68), we calculate  $\theta$  from the following equation of  $\rho$  and  $\gamma$ :  $\theta = \frac{\beta - \sqrt{\beta^2 - 4\alpha^2}}{2\alpha}$  where  $\alpha = \gamma - \rho$  and  $\beta = 1 + \rho^2 - 2\gamma\rho$ .  $\sigma_e^2$  is computed as:  $\sigma_e^2 = \frac{\sigma_{i,t}^2 - \rho^2 \sigma_{i,t-1}^2}{1 + \theta^2}$

## Appendix A: Derivation of the Human Capital Earnings Function

Assume the individual's objective is to maximize discounted disposable earnings  $Y_t$  over the working life-cycle.<sup>45</sup> This objective is achieved by choosing the amount of human capital  $K_t$  to reinvest each year in order to maximize the present value of lifetime earnings

$$\underset{K_t}{Max} \quad J = \int_0^N e^{-rt} Y_t dt \quad (A1)$$

where  $J$  is the total discounted disposable earnings over the working life-cycle,  $r$  is the personal discount rate, and  $N$  is the number of years one works, assumed known with certainty. Disposable earnings are

$$Y_t = R[E_t - K_t] \quad (A2)$$

where  $R$  is the rental rate per unit of human capital,<sup>46</sup>  $E_t$  denotes human capital stock in time period  $t$  and  $K_t$  the amount of human capital stock reinvested in time period  $t$  to create new human capital. We assume the individual begins with an innate stock of human capital  $E_0$  which can be augmented by investing all or part of this. The period-to-period change in human capital is denoted by

$$\dot{E}_t = Q_t - \delta E_t = \beta K_t^b - \delta E_t \quad (A3)$$

where we assume  $\delta$  is a constant rate of stock depreciation of existing human capital stock, and where we assume individuals create human capital using a Cobb-Douglas production  $Q_t = \beta K_t^b$ .

Maximization of (A1) subject to equations (A2) and (A3) entails maximizing the Hamiltonian

$$H(K_t, E_t, \lambda_t, t) = e^{-rt} R[E_t - K_t] + \lambda_t [\beta K_t^b - \delta E_t] \quad (A4)$$

with constraints  $E_t - K_t \geq 0$ . and making use of the transversality condition  $\lambda_N = 0$ .

The function  $\lambda_t$  is the marginal contribution to the total discounted disposable earnings if there is one more unit of human capital investment in period  $t$ . Assuming that no corner solutions are binding, the necessary conditions are as follows.

$$\frac{\partial H}{\partial K_t} = 0 \quad (A5.1)$$

$$\frac{\partial H}{\partial E_t} = -\dot{\lambda}_t \quad (A5.2)$$

$$\frac{\partial H}{\partial \lambda_t} = \dot{E}_t \quad (A5.3)$$

$$\lambda_N = 0 \quad (A5.4)$$

From equation (A5.2), we obtain  $\dot{\lambda}_t = -(\text{Re}^{-rt} - \delta \lambda_t)$ . Solving this differential equation and using the transversality condition (A5.4) we obtain

<sup>45</sup> As noted in the text, we abstract from labor supply considerations.

<sup>46</sup> In the empirical work we test whether  $R$  varies by race, occupation, industry and other variables.

$$\lambda_t = \frac{R}{r + \delta} e^{-rt} \left[ 1 - e^{-(r+\delta)(N-t)} \right] \quad (\text{A6})$$

From (A6),  $\dot{\lambda} < 0$ , indicating a diminishing value of human capital over time.

From (A5.1)

$$\frac{\partial H}{\partial K_t} = 0 = -R e^{-rt} + \lambda_t \beta b K_t^{b-1}$$

implying that

$$K_t = \left[ \frac{R e^{-rt}}{\lambda_t \beta} \right]^{1/b-1} = \left[ \frac{b \beta e^{rt} \lambda_t}{R} \right]^{1/1-b} \quad (\text{A7})$$

Substituting (A6) into (A7) yields

$$K_t = \left( \frac{b \beta}{r + \delta} \right)^{1/1-b} \left( 1 - e^{-(r+\delta)(N-t)} \right)^{1/1-b}, \text{ for } t \in [t^*, N]. \quad (\text{A8})$$

Of course,  $K_t = E_t$  during school since one devotes full-time to investing while in school.

To obtain human capital stock ( $E_t$ ), we combine (A8) with (A3) and (A5.3) which yields a differential equation whose closed form solution entails an infinite hypergeometric series

$$E_t = B e^{\delta(t^*-t)} + \left( \frac{\beta}{r + \delta} \right)^{1/(1-b)} b^{b/(1-b)} \sum_{j=0}^{\infty} \binom{b/(1-b)}{j} \frac{e^{j(r+\delta)(t-N)}}{j + (\delta/(r+\delta))} (-1)^j, \quad t \in [t^*, N], \quad (\text{A9})$$

where

$$B = \left[ \frac{\beta}{\delta} + \left( E_0^{1-b} - \frac{\beta}{\delta} \right) e^{\delta(b-1)t^*} \right]^{1/(1-b)} - \frac{\beta^{1/(1-b)}}{r + \delta} \left( \frac{b}{r + \delta} \right)^{b/(1-b)} \sum_{j=0}^{\infty} \binom{b/(1-b)}{j} \frac{e^{j(r+\delta)(t^*-N)}}{j + (\delta/(r+\delta))} (-1)^j. \quad (\text{A10})$$

Haley shows that the infinite hypergeometric series converges to a particular value from the second term. In Haley's derivation the convergence criterion is set for 6 decimal places. A simpler form can be obtained by setting the convergence at 4 decimal places. We use this slightly less stringent convergence criterion to construct the earnings function.

At  $j = 0$ , the infinite sum of the hyper-geometric series becomes

$$\sum_{j=0}^{\infty} \binom{b/(1-b)}{j} \frac{e^{j(r+\delta)(t^*-N)}}{j + \frac{\delta}{(r+\delta)}} (-1)^j = \frac{(r + \delta)}{\delta}$$

Thus,

$$B = \left[ \frac{\beta}{\delta} + (E_0^{1-b} - \frac{\beta}{\delta}) e^{\delta(b-1)t^*} \right]^{1/(1-b)} - \frac{\beta^{1/(1-b)}}{(r+\delta)} \left( \frac{b}{r+\delta} \right)^{b/(1-b)} \frac{(r+\delta)}{\delta}$$

or,

$$B = \left[ \frac{\beta}{\delta} + (E_0^{1-b} - \frac{\beta}{\delta}) e^{\delta(b-1)t^*} \right]^{1/(1-b)} - \frac{\beta^{1/(1-b)}}{\delta} \left( \frac{b}{r+\delta} \right)^{b/(1-b)} \quad \text{A(11)}$$

As such, the stock of human capital at time  $t$  can be expressed as

$$E_t = B e^{\delta(t^*-t)} + b^{b/(1-b)} \left( \frac{\beta}{r+\delta} \right)^{\frac{1}{1-b}} \left( \frac{r+\delta}{\delta} \right)$$

or

$$E_t = B e^{\delta(t^*-t)} + \frac{\beta^{1/(1-b)}}{\delta} \left( \frac{b}{r+\delta} \right)^{\frac{b}{1-b}}$$

or

$$E_t = \left\{ \left[ \frac{\beta}{\delta} + (E_0^{1-b} - \frac{\beta}{\delta}) e^{\delta(b-1)t^*} \right]^{1/(1-b)} - \frac{\beta^{1/(1-b)}}{\delta} \left( \frac{b}{r+\delta} \right)^{b/(1-b)} \right\} e^{\delta(t^*-t)} + \frac{\beta^{1/(1-b)}}{\delta} \left( \frac{b}{r+\delta} \right)^{b/(1-b)}$$

or

$$E_t = \beta^{1/(1-b)} \left[ \frac{1}{\delta} + \left( \frac{E_0^{1-b}}{\beta} - \frac{1}{\delta} \right) e^{\delta(b-1)t^*} \right]^{1/(1-b)} e^{\delta(t^*-t)} + \frac{\beta^{1/(1-b)}}{\delta} \left( \frac{b}{r+\delta} \right)^{b/(1-b)} (1 - e^{\delta(t^*-t)}) \quad \text{A(12)}$$

Observed earnings can be expressed as following

$$Y_t = R[E_t - K_t]$$

where,  $R$  is the rental rate of human capital. Thus,

$$Y_t = A_0 e^{\delta(t^*-t)} + A_1 [1 - e^{\delta(t^*-t)}] - A_2 [1 - e^{(r+\delta)(t-N)}]^{1/(1-b)} \quad \text{A(13)}$$

where

$$A_0 = R \beta^{1/(1-b)} \left[ \frac{1}{\delta} + \left( \frac{E_0^{1-b}}{\beta} - \frac{1}{\delta} \right) e^{\delta(b-1)t^*} \right]^{1/(1-b)}$$

$$A_1 = R \beta^{1/(1-b)} \left[ \frac{b}{r+\delta} \right]^{b/(1-b)} \frac{1}{\delta}$$

$$A_2 = R \beta^{1/(1-b)} \left[ \frac{b}{r+\delta} \right]^{1/(1-b)}$$

Letting  $x = e^{(r+\delta)(t-N)}$  in the third term of (A13), one can rewrite  $[1 - e^{(r+\delta)(t-N)}]^{1/(1-b)}$  in terms of  $x$  as  $f(x) = [1 - x]^{\frac{1}{1-b}}$ . Expanding  $f(x)$  with a second degree Taylor's series approximation around  $x = 0$  yields

$$f(x) = f(0) + f'(0)x + \frac{1}{2!}f''(0)x^2$$

or

$$f(x) = 1 - \frac{1}{(1-b)}e^{(r+\delta)(t-N)} + \frac{1}{2!} \frac{1}{(1-b)} \frac{b}{(1-b)} e^{2(r+\delta)(t-N)} \quad (A14)$$

Note that  $x = 0$  is a good approximation point because  $e^{(r+\delta)(t-N)}$  assumes a value close to zero for any reasonable  $(r + \delta)$  during one's work life  $t < N$ . To support this claim, we simulate  $f(x)$  with various plausible values of  $(r + \delta)$  ranging from (0.04 – 0.08) and  $b$  (ranging from 0.2 to 0.6) and plot them against age in Figure A1, Figure A2, Figure A3. All three figure show the approximated function closely matches the actual function.

As an additional test to show a statistically insignificant truncation error, we compute values for  $[1 - e^{(r+\delta)(t-N)}]^{1/(1-b)}$  over the life-cycle  $15 < t < 55$  (the  $t$ -range of our data). We denote these 39 values as  $z_t$ , where  $t = 16, \dots, 54$ . Next we approximate  $z_t$  by  $f(x_t)$  as in (A14). This yields 39  $z_t$  and  $f(x_t)$  pairs on which we run the regression  $z_t = \alpha f(x_t) + \varepsilon_t$ . We repeat this 5040 times for  $z_t$  and  $f(x_t)$  values computed for  $0.01 < r < 0.15$ ,  $0.01 < \delta < 0.1$ , and  $0.2 < b < 0.7$  in order to test the hypothesis that  $\alpha = 1$ . We find only 1.8% of the regressions reject the hypothesis at the 1% significance level.

Substituting (A14) in (A13) yields

$$\begin{aligned} Y &= (A_0 - A_1)e^{\delta(t^*-t)} + A_1 - A_2 \left[ 1 - \alpha e^{(r+\delta)(t-n)} + \frac{1}{2} \alpha (\alpha - 1) e^{2(r+\delta)(t-N)} \right] \\ &= (A_0 - A_1)e^{\delta(t^*-t)} + (A_1 - A_2) + A_2 \alpha e^{(r+\delta)(t-n)} - \frac{1}{2} A_2 \alpha (\alpha - 1) e^{2(r+\delta)(t-N)}. \end{aligned}$$

Finally, given that  $\alpha = \frac{1}{1-b}$  and  $W = R^{(1-b)}\beta$ , we obtain the earnings function (7):

$$\begin{aligned} Y &= W^{\frac{1}{(1-b)}} \left\{ \left( \frac{1}{\delta} + \left( E^{1-b} - \frac{1}{\delta} \right) e^{\delta(b-1)t^*} \right)^{\frac{1}{(1-b)}} - \left( \frac{1}{\delta} \left[ \frac{b}{r+\delta} \right]^{\frac{b}{(1-b)}} \right) e^{\delta(t^*-t)} \right\} \\ &\quad + \left\{ \frac{1}{\delta} \left[ \frac{b}{r+\delta} \right]^{\frac{b}{(1-b)}} \left( 1 - \frac{b\delta}{r+\delta} \right) \right\} + \left\{ \left[ \frac{b}{r+\delta} \right]^{\frac{1}{(1-b)}} \frac{1}{(1-b)} e^{(r+\delta)(t-N)} \right\} \\ &\quad - \left\{ 0.5 \left[ \frac{b}{r+\delta} \right]^{\frac{1}{(1-b)}} \left( \frac{1}{(1-b)} \right) \left( \frac{b}{(1-b)} \right) e^{2(r+\delta)(t-N)} \right\} \quad (7) \end{aligned}$$

We fit equation (7) to estimate  $b$ ,  $W$ ,  $E$ ,  $\delta$ , and  $r$ .

Figure A.1

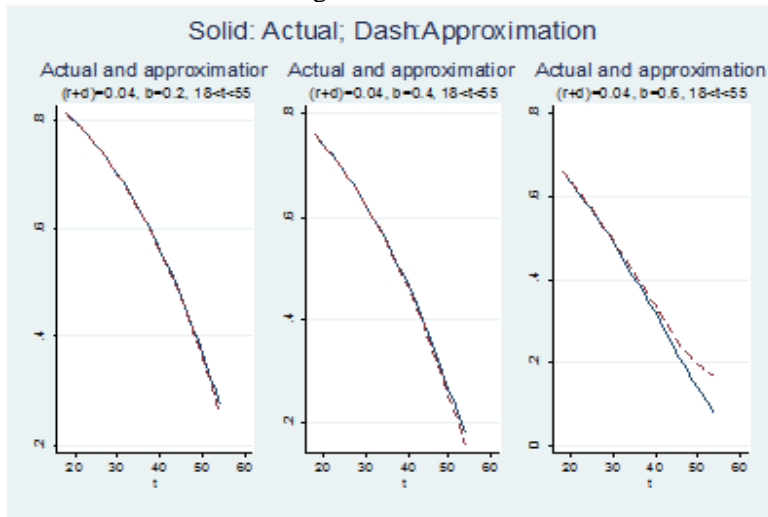


Figure A.2

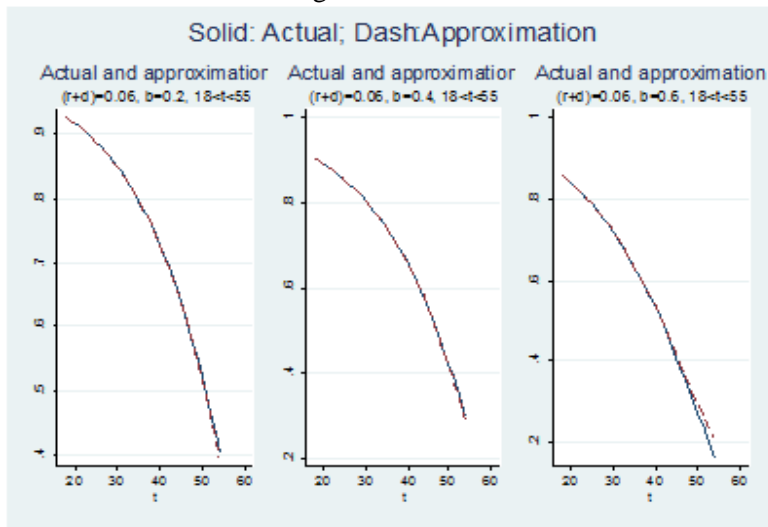
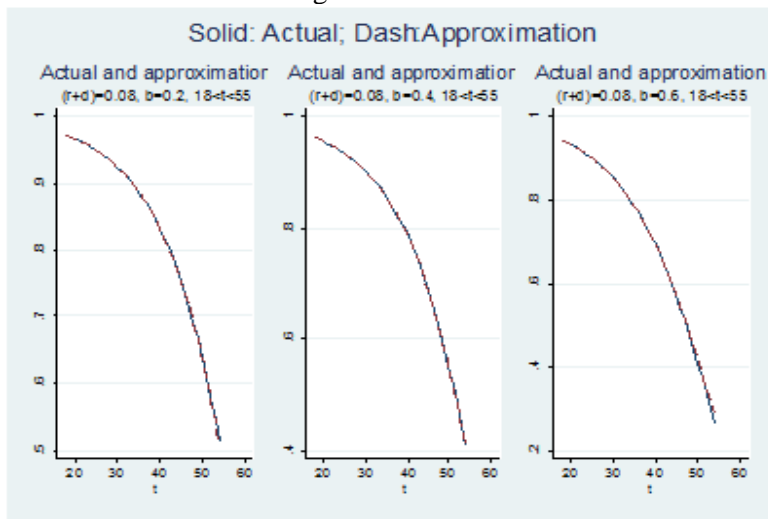


Figure A.3



## Appendix B: Structural identification

Least-squares estimation seeks a set of parameters that minimizes the mean square error  $E[y - f(t, \phi)]^2$  where  $f(t, \phi)$  is the regression function which represents the conditional mean of  $y$  ( $E[y|t] = f(t, \phi)$ ) under the assumption of  $E[\varepsilon] = 0$  and  $E[\varepsilon|t] = 0$ . Identification implies a unique  $\phi$  minimizing  $E[y - f(t, \phi)]^2$  that satisfies both moment conditions.

To show this, we employ a proof by contradiction. Let  $\phi$  be the conditional mean that minimizes the mean square error. Next, suppose there is another parameter  $\phi^*$  such that  $f(t, \phi^*)$  minimizes the mean square error. Identification requires that there exist a unique set of parameters  $\phi^*$  such that  $f(t, \phi) = f(t, \phi^*)$ .

Proof:

To show this, re-parameterize (7) as follows:

$$y = F + Me^{-\delta t} + Ke^{\theta_1 t} - Le^{\theta_2 t} + \varepsilon \quad (7')$$

where

$$F = W^{1/(1-b)} \left[ \left\{ \left( \frac{1}{\delta} + (E^{(1-b)} - \frac{1}{\delta}) e^{\delta(b-1)t^*} \right)^{1/(1-b)} \right\} + \left\{ \frac{1}{\delta} \left[ \frac{b}{(r+\delta)} \right]^{b/(1-b)} \left( 1 - \frac{b\delta}{(r+\delta)} \right) \right\} \right]$$

$$M = -W^{1/(1-b)} \left[ \left( \frac{1}{\delta} \left[ \frac{b}{(r+\delta)} \right]^{b/(1-b)} \right) e^{\delta t^*} \right]$$

$$K = W^{1/(1-b)} \left[ \left[ \frac{b}{(r+\delta)} \right]^{1/(1-b)} \frac{1}{1-b} e^{-(r+\delta)N} \right]$$

$$L = W^{1/(1-b)} \left[ 0.5 \left[ \frac{b}{(r+\delta)} \right]^{1/(1-b)} \frac{1}{1-b} \frac{b}{1-b} e^{-2(r+\delta)N} \right]$$

$$\theta_1 = (r + \delta)$$

$$\theta_2 = 2(r + \delta)$$

Suppose (7') is not identified. This means that there exists at least another  $\phi \neq \phi^*$  such that  $F^* + M^* e^{-\delta^* t} + K^* e^{\theta_1^* t} - L^* e^{\theta_2^* t} = F + M e^{-\delta t} + K e^{\theta_1 t} - L e^{\theta_2 t}$ , or  $F^* - F + M^* e^{-\delta^* t} - M e^{-\delta t} + K^* e^{\theta_1^* t} - K e^{\theta_1 t} - L^* e^{\theta_2^* t} + L e^{\theta_2 t} = 0$ .

Suppose  $\delta^* \neq \delta$ ;  $\theta_1^* \neq \theta_1$ ;  $\theta_2^* \neq \theta_2$ , then the only possible parameter values that satisfy the above equation are  $F^* = F = M^* = M = K^* = K = L^* = L = 0$  which means the regression function does not exist. This is a contradiction.

On the other hand, if  $\delta^* = \delta$ ,  $\theta_1^* = \theta_1$ , and  $\theta_2^* = \theta_2$ , then the only combination that satisfies the above equation is  $F^* = F, M^* = M, K^* = K, L^* = L$  which means that the parameter vectors are identical i.e.  $\theta_1^* = \theta_1$ ;  $\theta_2^* = \theta_2$ . Thus, (7') enables one to uniquely retrieve  $\hat{F}, \hat{M}, \hat{K}, \hat{L}, \hat{\theta}_1, \hat{\theta}_2$ , and  $\hat{\delta}$ .

Next, from this, we show a unique solution for the five parameters of interest ( $b, W, E, \delta$ , and  $r$ ) obtained when solving the seven equations underlying the parameter estimates ( $\hat{F}, \hat{M}, \hat{K}, \hat{L}, \hat{\theta}_1, \hat{\theta}_2$ , and  $\hat{\delta}$ ) obtained from (7'). To do this we write out the seven equations:

$$\hat{F} = W^{1/(1-b)} \left[ \left\{ \left( \frac{1}{\delta} + (E^{(1-b)} - \frac{1}{\delta}) e^{\delta(b-1)t^*} \right)^{1/(1-b)} \right\} + \left\{ \frac{1}{\delta} \left[ \frac{b}{(r+\delta)} \right]^{b/(1-b)} \left( 1 - \frac{b\delta}{(r+\delta)} \right) \right\} \right] \quad (i)$$



$$\widehat{M} = -W^{1/(1-b)} \left[ \left( \frac{1}{\delta} \left[ \frac{b}{(r+\delta)} \right]^{b/(1-b)} \right) e^{\delta t^*} \right] \quad (\text{ii})$$

$$\widehat{K} = W^{1/(1-b)} \left[ \left[ \frac{b}{(r+\delta)} \right]^{\frac{1}{1-b}} \frac{1}{1-b} e^{-(r+\delta)N} \right] \quad (\text{iii})$$

$$\widehat{L} = W^{1/(1-b)} \left[ 0.5 \left[ \frac{b}{(r+\delta)} \right]^{\frac{1}{1-b}} \frac{1}{1-b} \frac{b}{1-b} e^{-2(r+\delta)N} \right] \quad (\text{iv})$$

$$\widehat{\theta 1} = (r + \delta) \quad (\text{v})$$

$$\widehat{\theta 2} = 2(r + \delta) \quad (\text{vi})$$

$$\widehat{\delta} = \delta \quad (\text{vii})$$

From the above,  $\delta$  is uniquely identified from (vii). Given (vii) and the linear restrictions (v) and (vi), we obtain a unique value for  $r$ .

Next, divide (ii) by (iii) to obtain:  $\frac{\widehat{M}}{\widehat{K}} = -\frac{1}{\delta} e^{\delta t^*} (r + \delta) e^{(r+\delta)N} \frac{(1-b)}{b}$ .

Let  $\gamma_1 = -\frac{1}{\delta} e^{\delta t^*} (r + \delta) e^{(r+\delta)N}$  to yield  $\frac{\widehat{M}}{\widehat{K}} = \gamma_1 \frac{(1-b)}{b}$ . (viii)

Divide (iii) by (iv) to yield:  $\frac{\widehat{K}}{\widehat{L}} = 2(r + \delta) e^{(r+\delta)N} \frac{(1-b)}{b}$ .

Let  $\gamma_2 = 2(r + \delta) e^{(r+\delta)N}$  to obtain  $\frac{\widehat{K}}{\widehat{L}} = \gamma_2 \frac{(1-b)}{b}$ . (ix)

Combining (viii), and (ix) yields  $\frac{\widehat{K}^2}{\widehat{L}\widehat{M}} = \frac{\gamma_2}{\gamma_1}$ , or  $\widehat{K}^2 = \frac{\gamma_2}{\gamma_1} \widehat{L}\widehat{M}$ , or  $\widehat{K} = \sqrt{\frac{\gamma_2}{\gamma_1} \widehat{L}\widehat{M}}$

Note that  $\frac{\gamma_2}{\gamma_1}$  is expressed as a function of school leaving age  $t^*$ , Retirement age  $N$ , and  $\delta, r$ . Any solution of  $b, W, E, \delta, r$  that solves for  $M$  and  $L$  must also solve for  $K$ . Thus, one can ignore (iii).

From equation (i), (ii), and (iv), we solve for  $b, W, E$ . Divide (ii) by (iv) to obtain

$$\frac{\widehat{M}}{\widehat{L}} = -\frac{2}{\delta} e^{\delta t^*} (r + \delta) e^{(r+\delta)N} \frac{(1-b)^2}{b^2} \quad \text{or} \quad \frac{(1-b)}{b} = \pm \sqrt{-\frac{\delta}{2e^{\delta t^*} (r+\delta) e^{(r+\delta)N}} \frac{\widehat{M}}{\widehat{L}}} \quad (\text{x})$$

The solution thus depends on the value of  $\widehat{M}$ , and  $\widehat{L}$ . For any  $\widehat{M} \leq 0$ , and  $\widehat{L} \geq 0$  (x) will yield real solutions. Imposing the non-negativity restriction on the parameters (i.e.  $W, b, E, r, \delta > 0$ ) ensures that  $\widehat{M} \leq 0$ , and  $\widehat{L} \geq 0$  conditions are met. Further the restriction  $0 < b < 1$  ensures that the negative solution of  $\frac{(1-b)}{b}$  in (x) is also not possible. Hence,  $b$  is uniquely identified. Once  $b$  is uniquely identified, one can uniquely solve for  $W$  either from (ii) or from (iv). Finally, given the values of  $W, b, r, \delta$ , one can solve for  $E$  from (i). Thus, the parameters  $W, b, E, r, \delta$  are structurally identified.

## Appendix C: Simulated Parameter Estimates

We select a sample of 100 individuals at random from the 1826 individuals in our data. For each of these individuals, we have a set of parameters ( $b$ ,  $W$ ,  $E$ ,  $\delta$ , and  $r$ ), which were computed from equation (7). From these, we compute  $A_0$ ,  $A_1$ , and  $A_2$  of equation (A13). For each of the 100 individuals, we use equation (A13) along with schooling and work experience to generate 100 sets of earnings data (using A13) for samples of 15 to 24 observations (ten of size 15, ten of size 16, and so on, up to ten of size 24). We then augment earnings by random draws from a normal distribution with zero mean and a standard deviation based on the estimated error from equation (7) adjusted for the degrees of freedom. Using equation (A13) to simulate earnings enables us to avoid possible truncation biases arising from the Taylor Series approximation described in Appendix B. These varying sample sizes enable us to mimic the sample sizes we previously used from the NLSY. This procedure yields 10,000 individual panels of earnings data.

For each of these simulated datasets (of size 15 to 24) we re-estimate (7) to obtain the parameter estimates for  $b_{is}^e$ ,  $W_{is}^e$ ,  $\delta_{is}^e$ ,  $r_{is}^e$ , and  $E_{is}^e$  (where  $i = 1 \dots 100$  depicts each individual and  $s = 1 \dots 100$  depicts each panel). For each individual, we calculate the average of these estimated parameters ( $b_i^e$ ,  $W_i^e$ ,  $\delta_i^e$ ,  $r_i^e$ , and  $E_i^e$ ), their variances, and their root mean squared errors. Then we construct the bias by subtracting our original parameter estimates ( $b_i$ ,  $W_i$ ,  $\delta_i$ ,  $r_i$ , and  $E_i$ ) from the average of these estimated parameters to obtain  $B_i^b = (b_i^e - b_i)$ ,  $B_i^W = (W_i^e - W_i)$ ,  $B_i^\delta = (\delta_i^e - \delta_i)$ ,  $B_i^r = (r_i^e - r_i)$ , and  $B_i^E = (E_i^e - E_i)$ .

Figures C.1, C.2, and C.3 plot kernel densities of the 100 sets of these biases and their variances and root mean-square errors. Table C.1 presents the overall mean and the overall variance. Given that the coefficients for each  $i$  are obtained from simulated data, each set of biases comes from distributions with different standard deviations based on the original estimate of (7). For hypothesis testing, we standardized,  $B_i^b$ ,  $B_i^W$ ,  $B_i^\delta$ ,  $B_i^r$ , and  $B_i^E$ , such that  $\tilde{B}_i^b = \frac{B_i^b}{\sigma_{B_i^b}/\sqrt{n}}$ ,  $\tilde{B}_i^W = \frac{B_i^W}{\sigma_{B_i^W}/\sqrt{n}}$ ,  $\tilde{B}_i^\delta = \frac{B_i^\delta}{\sigma_{B_i^\delta}/\sqrt{n}}$ ,  $\tilde{B}_i^r = \frac{B_i^r}{\sigma_{B_i^r}/\sqrt{n}}$ ,  $\tilde{B}_i^E = \frac{B_i^E}{\sigma_{B_i^E}/\sqrt{n}}$  where  $\sigma_{B_i^b}$ ,  $\sigma_{B_i^W}$ ,  $\sigma_{B_i^\delta}$ ,  $\sigma_{B_i^r}$ , and  $\sigma_{B_i^E}$  are the standard deviations and  $n$  is the number of replications (in our case 100). These standardized biases are distributed according to a standardized normal distribution allowing us to test if they statistically differ from zero. We find four biases ( $\tilde{B}_i^b$ ,  $\tilde{B}_i^W$ ,  $\tilde{B}_i^\delta$ , and  $\tilde{B}_i^r$ ) are insignificantly different from zero at the 5% level. Table C.2 contains the 95% confidence intervals for these.

Figure C.1

Kernel Density Plot of Biases for Each Parameter

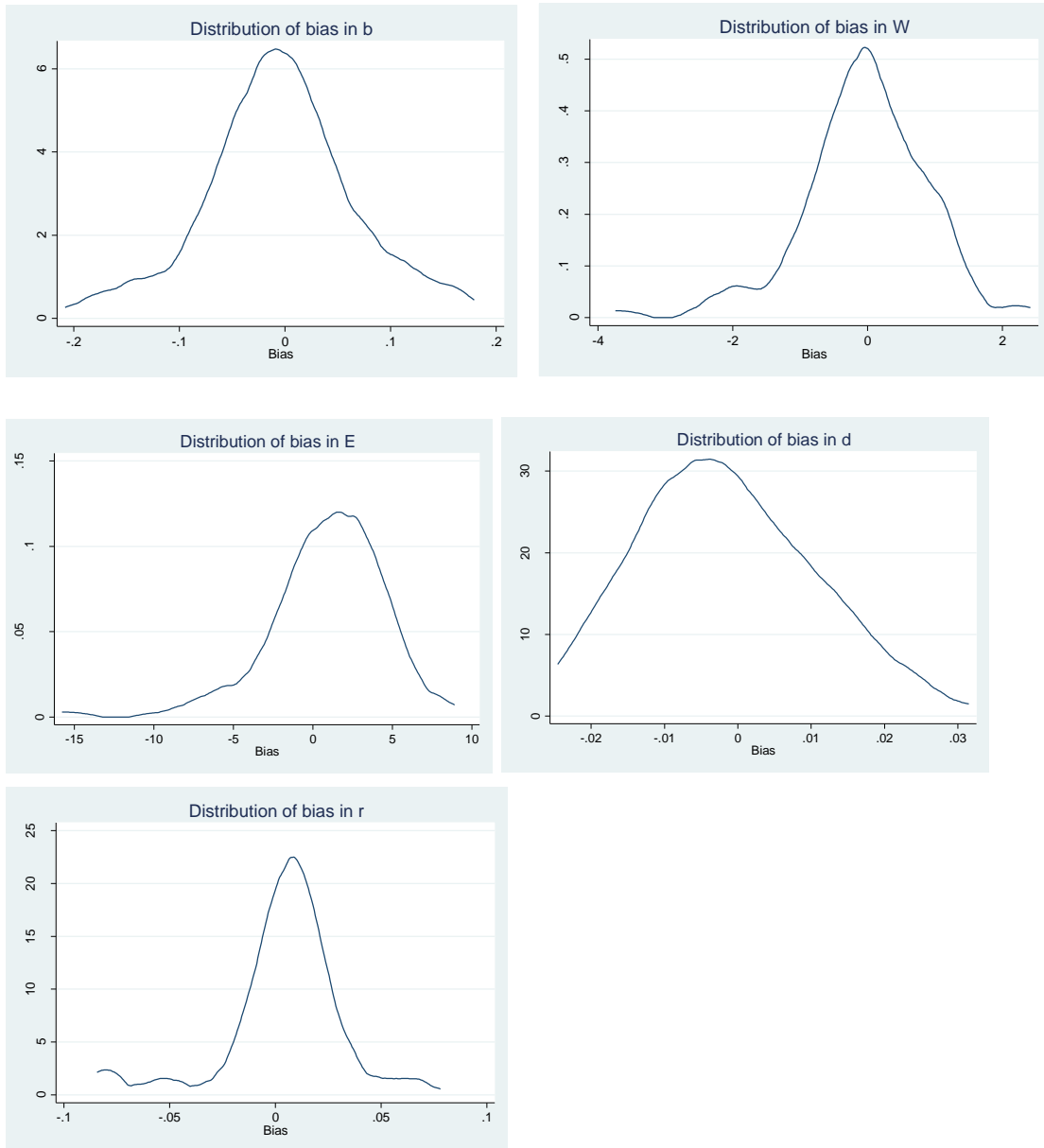


Figure C.2

Kernel Density Plot of Parameter Variances

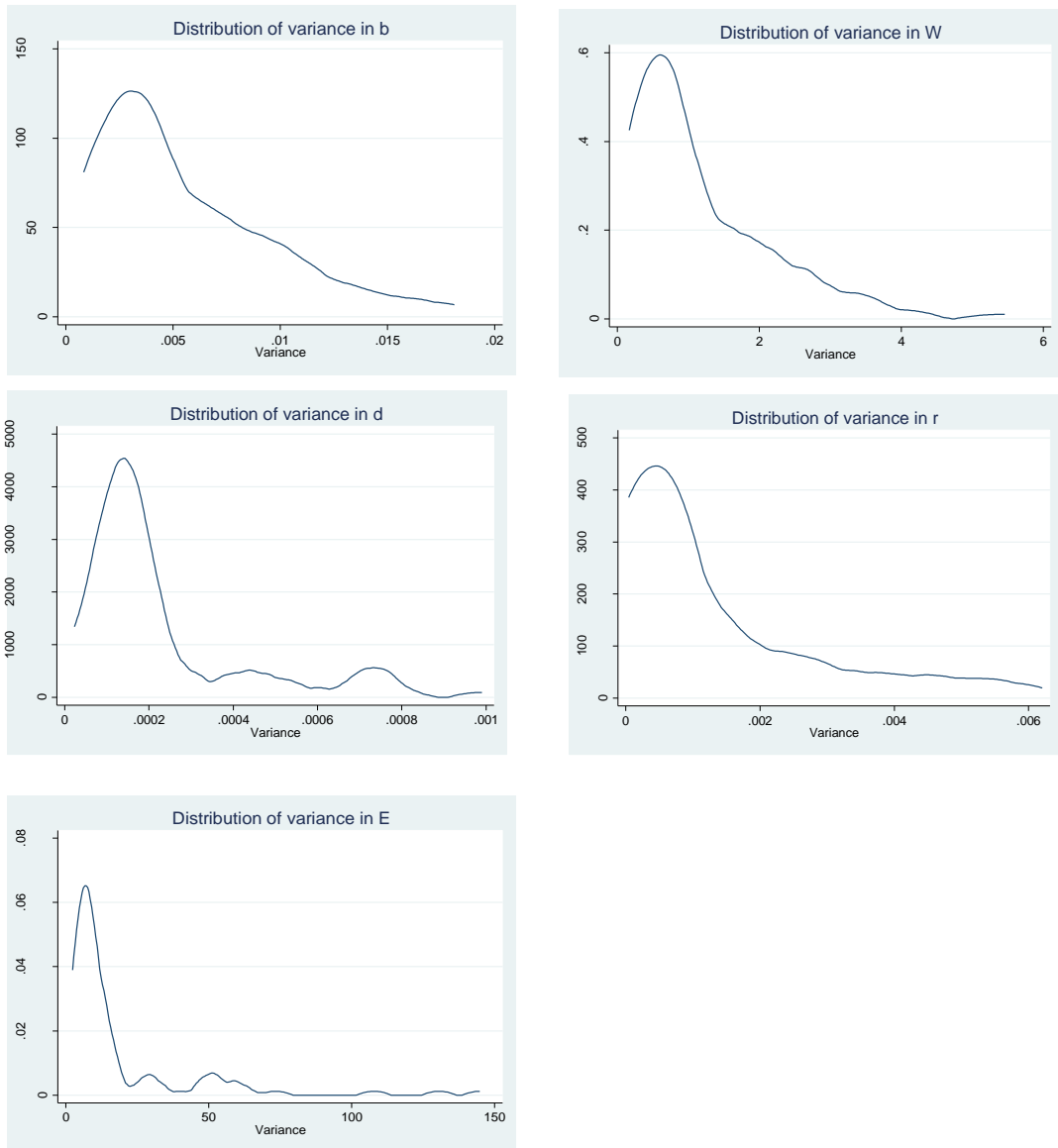


Figure C.3

Kernel Density Plot of Parameter Root Mean Square Errors

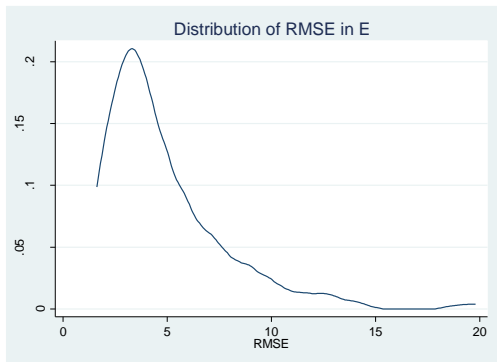
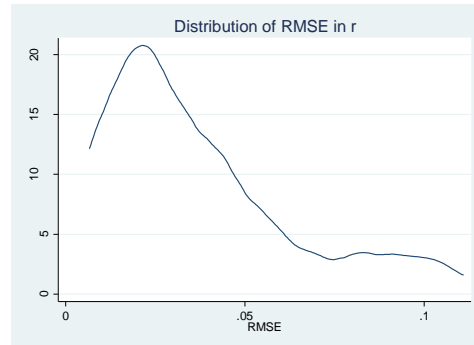
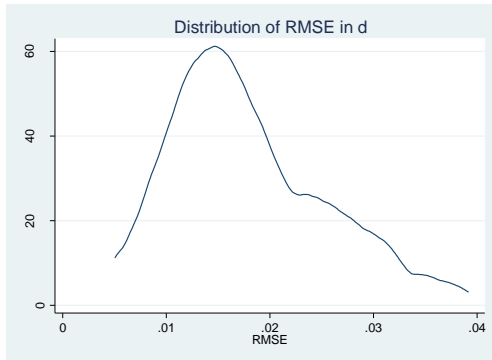
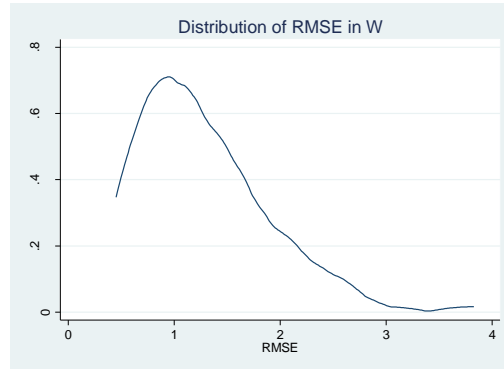
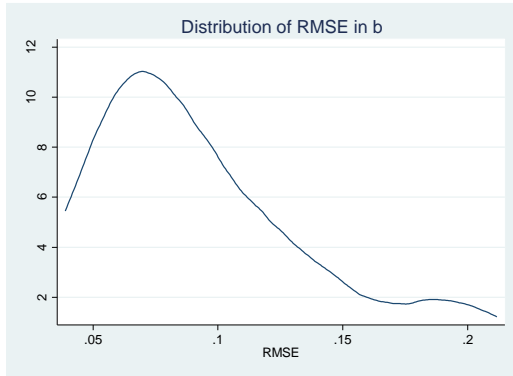


Table C.1

Simulated Biases

<b>Coefficient</b>	<b>Mean Bias</b>	<b>Variance</b>
<b>b</b>	<b>-0.005</b>	<b>0.00520</b>
<b>W</b>	<b>-0.046</b>	<b>0.86421</b>
$\delta$	<b>-0.002</b>	<b>0.000144</b>
<b>r</b>	<b>0.004</b>	<b>0.00083</b>
<b>E</b>	<b>0.897</b>	<b>13.3536</b>

Table C.2

Standardized Confidence Intervals of Biases

<b>Standardized Coefficient</b>	<b>95% Confidence Interval</b>		
<b>b</b>	<b>-3.483</b>	<b>to</b>	<b>0.941</b>
<b>W</b>	<b>-2.392</b>	<b>to</b>	<b>1.746</b>
$\delta$	<b>-3.398</b>	<b>to</b>	<b>0.034</b>
<b>r</b>	<b>-1.142</b>	<b>to</b>	<b>3.003</b>
<b>E</b>	<b>0.826</b>	<b>to</b>	<b>4.289</b>

## References

- Ahmed, S. Ejaz and Christopher Nicol. 2012. "An Application of Shrinkage Estimation to the Nonlinear Regression Model." *Computational Statistics and Data Analysis* 56:3309-3321.
- Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life-Cycle of Earnings." *Journal of Political Economy*, 75 (4): 352-365.
- Bhargava, A. and L. Franzini and W. Narendranathan. 1982. "Serial Correlation and the Fixed Effects Model." *Review of Economic Studies*, 49(4): 533-549.
- Borghans, Lex, Angela Lee Duckworth, James J. Heckman and Bas ter Weel. 2008. "The Economics and Psychology of Personality Traits." *Journal of Human Resources*, 43(4): 972-1059.
- Bowlus, Audra and Chris Robinson. 2012. "Human Capital Prices, Productivity, and Growth." *American Economic Review* 102(7): 3483-3515.
- Browning, Martin and Mette Ejrnæs. 2013. "Heterogeneity in the Dynamics of Labor Earnings," *Annual Review of Economics* 5: 219-245.
- Browning, Martin, Mette Ejrnæs and Javier Alvarez. 2010. "Modelling Income Processes with Lots of Heterogeneity." *Review of Economic Studies*, 77(4): 1353-1381.
- Browning, Martin, Lars Hansen and James Heckman. 1999. "Micro Data and General Equilibrium Models." in *Handbook of Macroeconomics*, Volume 1A (John Taylor and Michael Woodford, eds.) Amsterdam: Elsevier Press pp543-633.
- Chamorro-Premuzic, T. and A. Furnham. 2005. "The Relationship between Personality Traits, Subjectively-Assessed and Fluid Intelligence." *Personality and Individual Differences*, 38:1517-1528.
- Chernick, Michael. 2007. *Bootstrap Methods: A Guide for Practitioners and Researchers* (2<sup>nd</sup> ed.), Hoboken, NJ: John Wiley & Sons.
- Cunha, Flavio and James Heckman. 2007. "Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling." *Labour Economics*, 14(6): 870-93.
- Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2005. "Interpreting the Evidence on Life Cycle Skill Formation." NBER Working Paper No. 11331
- Cunha, Flavio, James J. Heckman, and Salvador Navarro. 2005. "Separating Uncertainty from Heterogeneity in Life Cycle Earnings." *Oxford Economic Papers* 57(2): 191-261.
- Czarnitzki, Dirk and Thorsten Doherr. 2009. "Genetic Algorithms for Econometric Optimization." Working Paper.
- DeYoung, Colin G. 2011. "Intelligence and Personality," in Sternberg, R. J., & Kaufman, S. B., Eds. (2011). *The Cambridge Handbook of Intelligence*. New York: Cambridge University Press. pp. 711-737.

- Doris, Aedín, Donal O’Neill, and Olive Sweetman. 2013. “Identification of the Covariance Structure of Earnings Using the GMM Estimator,” *Journal of Economic Inequality* 11(3): 343–372.
- Dorsey, Robert and Walter Mayer. 1995. “Genetic Algorithms for Estimation Problems with Multiple Optima, Nondifferentiability, and Other Irregular Features,” *Journal of Business & Economic Statistics*, 13(1): 53-66.
- Eisenhauer, Philipp, James Heckman, and Stefano Mosso. 2013. “Estimation of Dynamic Discrete Models by Maximum Likelihood and the Simulated Method of Moments,” Working Paper, <https://heckman.uchicago.edu/MLvsSMM>.
- Eysenck, H. J. 1994. “Personality and Intelligence: Psychometric and Experimental Approaches. In R. J. Sternberg and P. Ruzgis (Eds.), *Personality and Intelligence*. New York: Cambridge University Press. pp. 3-31.
- Fan, Jianqing, Peter Hall, and Qiwei Yao. 2007. “To How Many Simultaneous Hypothesis Tests Can Normal, Student’s t or Bootstrap Calibration Be Applied?” *Journal of the American Statistical Association* 102:1282-1288.
- Fuller, W. A. 1976. *Introduction to Statistical Time Series*. New York: John Wiley & Sons.
- Goldberg, David. 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, Massachusetts: Addison-Wesley.
- Gottfredson, L. S. 1997. “Mainstream Science on Intelligence: An Editorial with 52 Signatories, History, and Bibliography.” *Intelligence*, 24: 13-23.
- Guvenen, Faith. 2011. “Macroeconomics with Heterogeneity: A Practical Guide,” NBER Working Paper #17622.
- Haley, William J. 1976. “Estimation of the Earnings Profile from Optimal Human Capital Accumulation.” *Econometrica*, 44 (6): 1223-1238.
- Heathcote, Jonathan, Kjetil Storesletten, and Giovanni Violante. 2009. “Quantitative Macroeconomics with Heterogeneous Households,” NBER Working Paper No. 14768.
- Heckman, James. 1975. "Estimates of Aging Capital Production Function Embedded in a Life-Cycle Model of Labor Supply." in Nester Terleckyi, ed. *Household Production and Consumption*. New York: Columbia University Press for the National Bureau of Economic Research. pp. 227-58.
- Heckman, James. 1976. “A Life-Cycle model of Earnings, Learning and consumption,” *Journal of Political Economy*, 84(4): S11-S44.
- Heckman, James. 2008. “School, Skill, and Synapses.” *IZA Discussion Paper No 3515*.



- Heckman, James, Lance Lochner and Christopher Taber. 1998. "Explaining rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics*, 1: 1-58.
- Heckman, James, Lance Lochner, and Petra Todd. 2006. "Earnings Functions, Rates of Return And Treatment Effects: The Mincer Equation And Beyond," in Eric A. Hanushek and Finis Welch, eds. *Handbook of the Economics of Education*, 1: 307-458.
- Heckman, James, Lance Lochner, and Petra, Todd. 2008. "Earnings Functions and Rates of Return," Discussion Paper No. 3310, IZA.
- Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems*. Ann Arbor, Michigan: University of Michigan Press.
- Horn, John L. and Raymond B. Cattell. 1966. "Refinement and Test of the Theory of Fluid and Crystallized General Intelligences. *Journal of Educational Psychology*. 57: 253-270.
- Hryshko, Dmytro. 2008. "RIP to HIP: The Data Reject Heterogeneous Labor Income Profiles." University of Alberta Working Paper.
- Johnson, Thomas and Frederick Hebein. 1974. "Investments in Human Capital and Growth in Personal Income 1956-1966." *The American Economic Review*, 64(4), 604 - 15.
- Johnson, W. and T.J. Bouchard. 2005a. "The Structure of Human Intelligence: It's Verbal, Perceptual, and Image Rotation (VPR), not Fluid Crystallized." *Intelligence*, 33: 393-416.
- Johnson, W. and T. J. Bouchard. 2005b. "Constructive Replication of the Visual-Perceptual-Image Rotation Model in Thurstone's (1941) Battery of 60 Tests of Mental Ability." *Intelligence*, 33: 417-430.
- King, Robert and Sergio Rebelo. 1999. "Resuscitating Real Business Cycles." *Handbook of Macroeconomics*, Volume 1B (John Taylor and Michael Woodford, eds.) Amsterdam: Elsevier Press pp. 927-1007.
- Krueger, Dirk, Fabrizio Perri, Luigi Pistaferri, Giovanni L. Violante." 2010. "Cross-Sectional Facts for Macroeconomists," *Review of Economic Dynamics* 13: 1-14.
- Kuratanui, Masatoshi. 1973. *A Theory of Training, Earnings and Employment in Japan*. Ph.D. Dissertation. Columbia University.
- Lazear, Edward. 1977. "Schooling as a Wage Depressant." *Journal of Human Resources*, 12(2): 164-76.
- Li, Hongyi and G.S. Maddala. 1996. "Bootstrapping Time Series Models." *Econometric Reviews*, 15(2): 115-158.
- Lillard, Lee A. and Robert J. Willis. 1978. "Dynamic Aspects of Earning Mobility," *Econometrica*, 46(5): 985-1012.

- Liu, Hujun. 2009. "Life Cycle Human Capital Formation, Search Intensity and Wage Dynamics." University of Western Ontario Working Paper.
- MacKinnon, James. 2006. "Bootstrap Methods in Econometrics," *The Economic Record*, 82:S2-S18.
- McAdams, D. P. and J.L. Pals. 2006. "A New Big Five: Fundamental Principles for an Integrative Science of Personality." *American Psychologist*, 61, 204-217.
- Meghir, Costas and Luigi Pistaferri 2011. "Earnings, consumption and life cycle choices," in *Handbook of Labor Economics*. O. Ashenfelter and D. Card (eds), 4B: 774-854.
- Mincer, Jacob. 1958. "Investment in Human Capital and Personal Income Distribution." *Journal of Political Economy*, 66(4): 281-302.
- Mincer, Jacob. 1974. *Schooling, Experience and Earnings*. New York: Columbia University Press for the NBER.
- Moffitt, Robert and Peter Gottschalk. 2002. "Trends in the Transitory Variance of Earnings in the U.S." *The Economic Journal*, 112 (478): C68-C73.
- Newey, Whitney K. and Daniel McFadden. 1986. "Large sample estimation and hypothesis testing," *Handbook of Econometrics*, R. F. Engle & D. McFadden (ed.), edition 1, volume 4, chapter 36, pages 2111-2245 Elsevier.
- Pearlin, L., M. Lieberman, E. Menaghan and J. Mullan. 1981. "The Stress Process." *Journal of Health and Social Behavior*, 22: 337-356.
- Polachek, Solomon. 1975 "Differences in Expected Post-School Investment as a Determinant of Market Wage Differentials." *International Economic Review*, 16:451-70.
- Polachek, Solomon. 1981. "Occupational Self Selection: A Human Capital Approach to Sex Differences in Occupational Structure." *Review of Economics and Statistics*, 63(1) 60-69.
- Polachek, Solomon. 2008. "Earnings Over the Lifecycle: The Mincer Earnings Function and Its Applications," *Foundations and Trends in Microeconomics*, 4(3): 165-272.
- Polachek, Solomon and Francis Horvath. 1977. "A Life-Cycle Approach to Migration: Analysis of the Perspicacious Peregrinator." *Research in Labor Economics*, 1:103-49.
- Rosenberg, M. 1965. *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.
- Rubinstein, Yona and Yoram Weiss. 2006. "Post Schooling Wage Growth: Investment, Search and Learning," in Eric A. Hanushek and Finis Welch, eds. *Handbook of the Economics of Education*, Elsevier. 1: 1-67.
- Segal, Carmit. 2012. "Working When No One Is Watching: Motivation, Test Scores, and Economic Success." *Management Science*, 58(8): 1438-1457.

Sologon, D., and C. O'Donoghue. 2010. "Earnings Dynamics And Inequality Among Men In Luxembourg, 1988–2004: Evidence From Administrative Data," IZA Discussion Paper 3751.

Song, Xueda and John Jones. 2006. "The Effects of Technological Change on Life-Cycle Human Capital Investment." Working Paper.

White, Kenneth J. 1992. "The Durbin-Watson Test for Autocorrelation in Nonlinear Models." *The Review of Economics and Statistics*, 74(2): 370-373.