Heterogeneity in the Production of Human Capital

Solomon W. Polachek
Tirthatanmoy Das
Rewat Thamma-Apiroam

April 2013

Heterogeneity in the Production of Human Capital

Solomon W. Polachek
State University of New York at Binghamton
and IZA

Tirthatanmoy Das
Temple University
and IZA

Rewat Thamma-Apiroam
Kasetsart University

ABSTRACT

Heterogeneity in the Production of Human Capital[*]

We derive a tractable nonlinear earnings function which we estimate separately for each individual in 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 based on the life-cycle human capital model. In addition, we are able to show how these human capital production function parameters relate to cognitive ability, personality traits, and family background. Among our results, we find: Black-white differences in ability are smaller than those exhibited in standardized tests. Blacks have higher time discount and skill depreciation rates than whites. Individuals with higher time discount rates and greater rates of skill depreciation have fewer years of school. Individuals with both a high internal locus of control and self-esteem exhibit greater ability, lower skill depreciation, and smaller time discount rates. Individuals inclined towards depression have higher time discount rates. Agreeable, open, conscientious and extrovert individuals have a greater ability to learn but not necessarily a greater ability to earn. Neurotic individuals have a lower ability to learn. Higher parental education is associated with a greater ability to learn, lower skill depreciation, and a smaller time discount rate. Educational stimuli, such as growing up in a household that subscribed to magazines, are associated with higher ability. Conversely, growing up poor is associated with lower ability.

Corresponding author:

Solomon W. Polachek
Department of Economics
State University of New York at Binghamton (Binghamton University)
Binghamton, NY 13902-6000
USA
E-mail: polachek@binghamton.edu

1. Introduction

Much economics research analyzes the behavior of a typical economic agent, but ignores the possibility that agents differ from each other. This representative agent assumption pervades most areas of economics. For example, in microeconomics, theoreticians often assume standard utility and production functions but ignore the possibility that these functions can vary across individuals and firms. As an illustration, Jaimovich and Rebelo (2009) state "our model economy is populated by *identical* (emphasis ours) agents who maximize their lifetime utility." As examples in macroeconomics, Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2003) use a dynamic stochastic generalized equilibrium model to estimate the effect of a change in monetary policy but downplay the fact that various agents can react differently. In subfields, for example public finance, Parry and Small (2009) derive "tractable formulas for the welfare effects of fare adjustments in passenger peak and off-peak rail and bus transit" based on a representative agent framework, but recognize "that this can only approximate the aggregate behavior of a diverse population" (p. 273).

Often there are major discrepancies between the microeconomic evidence and the parameters upon which such representative agent models are based. Browning, Hansen and Heckman (1999) illustrate a number of important inconsistencies. As an example of this discordance, they show that typical micro-based labor supply estimates do not yield results compatible with often used steady-state growth models (p. 551). Browning, Hansen, and Heckman (1999) also outline the pitfalls of basing one's analysis on a representative agent, rather than taking account of the heterogeneity between agents. Here, they describe the drawback of using male labor supply parameters to calibrate steady-state growth, when female parameters (such as the inter-temporal elasticity of substitution) differ so greatly. As another illustration, they show how failing to take account of heterogeneity in rates of time preference can severely bias estimates of key macroeconomic indicators. In short, it is crucial to empirically document the heterogeneity of key behavioral parameters in order to understand many aspects of the economy.

Of late, acknowledging population-wide heterogeneity is becoming more widespread. For example, with regard to the debate whether prices are flexible or sticky, Boivin, Giannoni and Mihov (2009) allude to the fact that "empirical studies based on aggregate data . . . found stickiness" whereas "evidence based on the behavior of disaggregated prices suggests that prices are much more volatile than conventionally assumed in studies based on aggregate data" (pp. 350-351). Similarly in economic theory, models are beginning to address heterogeneity in a variety of ways. For example, Chiappori and Ekeland (2009) analyze "when, to what extent and under what conditions" one can recover underlying individual preferences and decision processes from a group's aggregate behavior. Another example of the importance of heterogeneity is the now widely available agent-based modeling software (Bonabeau, 2002).

Parameters of life-cycle models are widely used (as the workhorse) in various branches of economics. For example, in macroeconomics they are used to calibrate dynamic general equilibrium models of consumption and savings (Browning, Ejrnæs and Alvarez, 2010). In public finance they are used to explain aspects of earnings distribution over time (Moffitt and Gottschalk, 2002). In this vein, Huggett, Ventura and Yaron (2010) show how ability differences explain the widening (within age cohort) earnings distribution, but that differences in initial human capital indicate a narrowing of earnings distribution. In short, getting individual-specific parameters of the life cycle human capital accumulation model is important in many areas of economics. And as Browning, Hansen and Heckman (1999) show, estimating heterogeneity of these human capital parameters is crucial because often the representative agent model is severely limited.

Currently studies that get at parameter heterogeneity based on the human capital model do so in a limited way. They typically examine parameters for broad groups, such as old and young, male and female, or married and single. Some disaggregate to the industry level, some by level of education, and some by race. None, to our knowledge, examine how each of these parameters varies individual-by-individual. Surely, disaggregation on this fine a level will better shed light on the degree of heterogeneity.

Knowing more about individual heterogeneity of such key human capital parameters has implications regarding the labor market, particularly schooling decisions and earnings distribution, as well as social issues such as criminal behavior and teenage pregnancy. Also, this knowledge could yield better predictions of macroeconomic performance.

Estimating individual-specific human capital parameters is valuable for other reasons, as well. First, knowing individual-specific human capital parameters enables one to test heretofore untested aspects of the life-cycle human capital model. One 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 one 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 one to examine how family background, including getting educational stimuli as a child, is related to one's abilities to learn and earn. Exploring these relationships potentially have implications for the nature versus nurture debate, and hence are important for policymakers.

The point of this paper is to estimate person-specific parameters of the life-cycle human capital model in order to document the heterogeneity of these parameters across the population. In doing this, we analyze the plausibility of our estimates based on economic theory, but we go further and examine how these individual-specific parameters vary by race, by schooling level, by cognitive and other skills, by personality traits, and by family background.

Estimating heterogeneity has a long history. At least since 1950, econometricians considered the case when coefficients vary across individual observations (Rubin, 1950). Early empirical work consisted of random coefficients models. The problem, however, is

that these type models impose strict distributional restrictions (usually normality) on the heterogeneous coefficients, and anyway end up estimating the mean response, as in Hildreth and Houck (1968). This is also true of extensions of the random coefficients model, for example, the correlated random coefficients model.

Estimation with panel data often uses fixed-effects methods to adjust for heterogeneity. Of these, most analyses assume person specific intercepts. However, person specific intercepts imply individuals differ only with respect to the level of the outcome variable, but do not allow other factors to affect each individual differently. Clearly individual earnings profile slopes can vary just as do earnings profile intercepts. Models that adopt individual specific slopes generally do not consider heterogeneity for more than one exogenous variable. But even so, they still estimate single common population-wide parameters for each of the other independent variables (Polachek and Kim, 1994; Pesaran, 2006). At best, one gets at heterogeneity of one or two parameters, but not of each parameter. Mixed effect models, also used to estimate individual-specific parameters from panel data, likewise rely on distributional assumptions. Finally non-parametric approaches get at heterogeneity essentially by grouping individuals according to related (neighboring) measured characteristics within optimal "band widths" (Racine and Li, 2004). Rather than individual-specific parameters, they obtain "group-specific parameters."

With the advent of speedier computers, better optimization routines, and sufficiently longer panels, one can retrieve parameters of the human capital life-cycle model by estimating appropriate earnings functions for specific individuals. From these estimates, one can aggregate the data to obtain parameter averages for selected groups, such as all those employees of a particular race. In this paper, we make use of long enough panels to 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 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 in our analysis. First, we devise a tractable formulation of an earnings function based on 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 Pohjampalo's (1978) criteria for nonlinear time-continuous equations. Second, we adopt a maximization routine utilized in genetic biological research for nonlinear estimation. This permits us to search the parameter space more efficiently than traditional Newton-Raphson techniques. Third, we are able to examine the plausibility of our estimates by testing whether they are consistent with individual choices based on life-cycle human capital theory. Fourth, we are able to examine the distribution of each parameter across the population and compare such distributions by race. Fifth, we are able to show how family background as well as skill-based tests and personality are related to one's abilities to earn and learn.

Our estimates yield a number of new findings. For example, among our results, we find black-white differences in ability to earn and learn are smaller than those exhibited in standardized tests. 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 higher time discount rates. Conscientious, agreeable, open, and extrovert individuals have a greater ability to learn, but not necessarily a greater ability to earn. Neurotic individuals have a lower ability to learn. 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.

In addition, our results confirm findings obtained by psychologists. In this regard, psychologists break down personality into five basic factors: agreeableness, extroversion, openness, conscientiousness, and neuroticism. Of these they contend openness to be most correlated with intellectual ability and neuroticism the least correlated. Our results confirm this. Of the five personality traits, we also find the highest correlation between the ability to learn and openness, and we find the smallest (indeed negative) correlation between the ability to learn and neuroticism.

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 means individuals do not leave the labor force for family reasons and they do not alter their intended human capital investment plans based on unexpected spells of unemployment. This is why we concentrate on males who generally have continuous work histories. Second, we assume a relatively simple human capital production function. 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 inputs which can also be used to create additional human capital.[1] In our model high ability people can create a given amount of human capital with smaller time inputs. Third, we assume labor markets reward individuals based on human capital, and that 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 over the life-cycle of each individual.[2] As such, we claim most earnings variation is attributable to differences in the amount of human capital acquired over the life-cycle. 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

---

[1] Heckman (2008) describes how to modify the underlying human capital production function to include these factors as well as family background and personality.

[2] 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.

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 strong work-history constitute a select sample of the population. However, in our case, this selectivity does not preclude inferences obtained when relating our estimates to independently obtained individual characteristics because precisely the same individuals are used in obtaining information on both. On the other hand, selectivity biases could come about when making inferences about racial differences in ability if white workers are different in ability than black workers. These racial differences could come about, for example, if black workers are relatively more able than white workers compared to black and white non-workers. However, as will be shown, 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

Most empirical analyses adopt single equation log-linear Mincer earnings functions to parameterize earnings. The beauty of estimating simple Mincer earnings functions is computational ease.[3] Underlying the Mincer model is a Ben-Porath (1967) life-cycle earnings generating process that yields a complex nonlinear earnings function.[4] From this nonlinear function one is able to identify three cognitive ability parameters based on the production function of human capital. Using Heckman, Lochner and Taber's (1998) terminology, two of these parameters depict "ability to learn" because they measure the

---

[3] There are also some conceptual issues regarding estimation of the Mincer earnings function as well as interpretation of the schooling and experience coefficients. See Heckman, Lochner and Todd, 2006,

[4] Mincer's loge-linear specification gets around these nonlinearities by assuming time-equivalent human capital declines linearly with age. In reality, the time path of human capital acquisition is more complicated. Taking this investment time path into account yields a complex nonlinear earnings function.

ease which an individual can create new human capital from old human capital.[5] The third parameter depicts an individual's initial "ability to earn" because it represents earnings power devoid of human capital investments.[6] In addition, also for each individual, one can estimate a skill depreciation parameter as well as a rate of time preference.

The derivation of the earnings function containing these parameters entails the typical lifetime maximization paradigm. Based on this optimization process, one can derive optimal human capital investment, optimal human capital stock, and optimal earnings over a person's lifetime. 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 and 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). The result is a nonlinear earnings function with three parameters reflecting the different kinds of ability.

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. 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

---

[5] 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.

[6] 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.

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

The Ben-Porath (1967) model assumes individuals invest in themselves to maximize expected lifetime earnings.[7] 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, one equates the marginal cost and marginal gains of creating new human capital. 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.[8] 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).[9] 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 closed-form 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, a decade thereafter, Haley (1976) was able to estimate a version, but he simplified the estimation because not all parameters were readily identifiable. Given these computational difficulties, most scholars relied on Mincer's linear-in-the-parameters specification, which has become known as the Mincer earnings function. One problem is that Mincer's simplification does not allow one to identify the ability

---

[7] 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.

[8] 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.

[9] See Polachek (1975) for the case of discontinuous labor force participation.

parameters inherent in the Ben-Porath human capital production function. Another, is most specifications do not yield estimates of skill depreciation.[10]

Given the advent of faster computers and longer panels containing individual data, we feel now is a good time to reexamine Haley's approach. Further, as mentioned above, given sufficiently long panels for particular individuals, the approach enables one to compute ability, skill depreciation and time preference parameters person-by-person. Obtaining person-specific measures addresses unobserved heterogeneity. This is a relatively important issue in micro-based econometric research, and it has implications for calibrating macroeconomic models.

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,

$$Y_t = R E_t \qquad (1)$$

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

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

---

[10] T. Johnson (1970) as well as Mincer (1974) provide a specification whereby depreciation can be estimated, but few employ this strategy.

[11] 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. In order to preserve degrees of freedom we maintain the assumption of a constant rental rate. Instead, later in the paper, we test whether rental rates vary significantly by population strata.

[12] 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.

where δ 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

$$Q_t = \beta K_t^b \tag{3}$$

where K_t is the fraction of human capital stock reinvested in time period t and parameters $b \in [0,1]$ and β are production function parameters. [13] The parameter b reflects the rate at which current (invested) human capital stock is transformed to new human capital. It reflects how one acquires new knowledge from old, and as such reflects how quickly one learns. We denote b to depict the "scale" at which one learns. The "technology" parameter β represents total factor productivity. Both β 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. [14] 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

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

where J is the total discounted disposable earnings over the working life-cycle, r is the personal time discount rate and N is the number of years after which one retires (assumed known with certainly). [15] Disposable earnings are

$$Y_t = R[E_t - K_t]. \tag{5}$$

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

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

---

[13] As already mentioned, we assume no other 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).

[14] As already mentioned, we abstract from labor supply.

[15] 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.

with constraints $E_t - K_t \geq 0$, which means one cannot invest using more human capital than one currently has (i.e., no borrowing); 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. In school one plows back all one's earnings potential into more human capital investment and hence has no net earnings. Likewise during retirement one does not work so there are no earnings then either.

This maximization yields a nonlinear (in the parameters) earnings function[16]

$$Y_t = A_0 e^{\delta(t^*-t)} + A_1 [1 - e^{\delta(t^*-t)}] - A_2 [1 - e^{(r+\delta)(t-N)}]^{1/(1-b)} \tag{7}$$

where

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

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

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

and where t^* is the age at which one graduates from school (i.e., the age when Phase 1 ends); 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 one's training begins. In reality, parents initiate training their children at (or prior to) birth, but for our purposes we consider period 0 to be begin when a child starts formal schooling because this is the point we know children begin learning full-time.

---

[16] Appendix A contains the derivation. Note this differs a bit from the Haley specification because in our derivation we assume a one-term Taylor expansion whereas Haley uses a two-term Taylor expansion. Our approach yields a slightly simpler earnings function. This simplification will enable us to identify skill depreciation which Haley's specification could not do.

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 upon completing school. E_S is computed by augmenting E_0 by the amount of human capital produced in 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

Equation (7) contains six parameters: $R,\ \beta,\ b,\ r,\ \delta,$ and E_0. The parameters $r, \delta,$ and b all have no dimension. The parameters r and δ are percentages. The parameter b is the output elasticity in the human capital production function (3). It reflects returns to scale of human capital. It also can be construed as an ability parameter since it measures the productivity of old human capital in creating new human capital. These parameters are technically observable. The parameters β and E_0 are nominated in terms of units of human capital stock, and R is dimensioned as dollars per unit of human capital. Combining β and E_0 yields $E_0/\beta^{1/1-b}$ which is dimensionless. Thus we also treat

$E_0/\beta^{1/1-b}$ as a single parameter. For each individual i, we write $R\beta^{\frac{1}{1-b}}$ as

$$\hat{w}_i = [R^{(1-b_i)}\beta_i]^{\frac{1}{1-b_i}} = w_i^{*\frac{1}{1-b_i}}$$

(8)

14

where $w_i^* = R^{(1-b_i)}\beta_i$, which is the parameter we estimate. As a result of these identification restrictions we end up estimating five parameters: $\hat{b}_i = b_i$, $\hat{E}_i = \dfrac{E_{0_i}}{\beta_i^{1/1-b_i}}$,

$\hat{w}_i^* = R^{(1-b_i)}\beta_i$, $\hat{r}_i = r_i$ and $\hat{\delta}_i = \delta_i$ for each individual using nonlinear least squares for those individuals with at least fifteen years of data.

Fitting the earnings function (7) gives rise to four major considerations. First, proving the structural identifiability of each parameter is complicated, given the intricate nonlinearity of (7). Second, choosing an efficient optimization routine is important, both given the large number of nonlinear regressions we need to estimate and the necessity to achieve global rather than local convergence. Third, we need a method to retrieve individual-specific parameters estimates β_i, E_{0i}, and the population estimate R from composite term estimates E_i and w_i^* . Fourth, specifying standard errors for each parameter estimate is nontrivial given only 15-24 observations per individual. We handle each of these four issues as well as test whether R varies across the population in the following five subsections.

3.1 Structural Identification

H. Pohjanpalo (1978) provides criteria to prove structural identifiability of a nonlinear continuous time equation. His approach is based on a Taylor series expansion of the estimating equation. Because these Taylor series coefficients are unique, he shows identifiabilty is determined by the number of solutions of the set of nonlinear equations obtained when taking the estimating equation's time derivatives. A unique solution implies global identifiability; a finite number of solutions implies local identifiability; and an infinite number of solutions implies unidentifiability. To apply Pohjanpalo's approach, we take five derivatives, given that (7) has five unknown individual-specific parameters.

To implement the approach we approximate the third term of (7), i.e., by a second degree Taylor series expansion. This yields

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

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

$$\left. \left\{0.5\left[\frac{b}{r+\delta}\right]^{\frac{1}{(1-b)}}\left(\frac{1}{(1-b)}\right)\left(\frac{b}{(1-b)}\right)e^{2(r+\delta)(t-N)}\right\}\right] + \epsilon_t. \tag{7'}$$

Given likely measurement error in Y_t and the influence of other unobservable factors, we add a time varying error term ε_t for each individual. We fit (7') to estimate b, W, E, δ, and r. The dependent variable is the individual's weekly earnings data (adjusted to 1982-84 dollars). The independent variables are current age and age upon completion of school. Our dataset has 1928 individuals. Thus we run 1928 regressions to obtain parameter estimates for each individual.

To prove identification, we re-parameterize (7') to yield the following function

$$Y = Me^{-\delta t} + F(K, L, \delta, r) + Ke^{\theta t} - Le^{2\theta t} \tag{7''}$$

where $\theta = (r + \delta)$; F is a function of parameters M, K, L, δ, and r;

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

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

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

Identification of M, K, L, δ, and r implies the identification of $b, W, E, \delta,$ and r,

where $W = R^{(1-b)}\beta$ and $E = \dfrac{E_0}{\beta^{1/(1-b)}}$. As per Pohjampalo (1978), we take a power series

expansion of the first five derivatives of (7'') to obtain the following

$$\frac{dY}{dt}\Big|_{t=0} = -\delta M + \theta K - 2\theta L = z1$$

$$\frac{d^2Y}{dt^2}\Big|_{t=0} = \delta^2 M + \theta^2 K - 4\theta^2 L = z2$$

$$\frac{d^3Y}{dt^3}\Big|_{t=0} = -\delta^3 M + \theta^3 K - 8\theta^3 L = z3$$

$$\frac{d^4Y}{dt^4}\Big|_{t=0} = \delta^4 M + \theta^4 K - 16\theta^4 L = z4$$

$$\frac{d^5Y}{dt^5}\Big|_{t=0} = -\delta^5 M + \theta^5 K - 32\theta^5 L = z5$$

The solution of these five equations yields seven solutions for M, K, L, δ, and r; and δ.
This derivation proves that the parameters M, K, L, δ, and θ, and hence $b, W, E, \delta,$ and r
are locally identified. (See Appendix C.)

3.2 Global Convergence and Parameter Restrictions

For estimation we employ the Genetic Algorithm (GA), a recently available parallel
processing optimization routine originally used by biologists in genetic research.[18] This
algorithm reaches an optimum more efficiently than traditional Newton-Raphson hill-
climbing techniques. Also, it is less prone to converging at local rather than global optima.

We implement the algorithm to estimate $b, W, E, \delta,$ and r in (7') using nonlinear
least-squares.[19] As in the Ben-Porath life-cycle model we restrict the parameter space to

---

[18] We use a version of GA written by Czarnitzki and Doherr (2009).
[19] We assume iid residuals based on the results of a Durbin-Watson serial correlation test for the nonlinear
least squares model (White, 1992). Among the 1868 individuals in our data, only 6 cases exhibit positive
autocorrelation. The absence of significant autocorrelation (5% significance level) also allows us to abstract
from possible persistent macroeconomic shocks often considered in the earnings dynamics literature. Were
there any persistent wage shock, one would need to modify the regression technique to accommodate the
serial correlation.

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 δ to be less than 0.2 and 0.1 respectively, given that r and δ 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.

3.3 Identification of Individual-Specific β_i, E_{0_i}, and Population-Wide R

To identify β_i and E_{0_i} we adopt the following approach: First, we specify β_i to equal βe_i where β is the population average and e_i is the individual deviation. Second, we rewrite (8) as

$$w_i^* = R^{(1-b_i)} \beta e_i. \tag{9}$$

Taking the logarithm, yields

$$\ln w_i^* = (1-b_i) \ln R + \ln \beta + \ln(e_i). \tag{10}$$

Estimating (10) using each individual's values for $\widehat{w_i^*}$ and $\hat{b}_i$ obtained from the parameterization we employ to estimate (7') gives a population value of R (the coefficient of $(1-b_i)$, the average β (the constant term), and individual-specific values of β_i obtained by taking the anti-log$_e$ of the sum of the latter two terms in (10). Utilizing b_i and β_i values along with the coefficient $\hat{E}_i = \dfrac{E_{0_i}}{\beta_i^{1/1-b_i}}$ obtained from estimating (7') yields individual-specific E_{0_i}.

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

Equation (10) can be modified to test for human capital homogeneity. Ben-Porath assumes homogeneous human capital. As such, each basic unit rents for a price (R, in our notation) determined in the market. Given homogeneity, R is the same across all

occupations, education levels and all other characteristics because earnings differ across the population in the amount of human capital each person has acquired and not the remuneration of each unit, since each unit is the same. On the other hand, rental rates per unit of human capital can vary should human capital be heterogeneous. Here, heterogeneity implies the market rewards each type of human capital differently. Thus a test for human capital homogeneity is to test whether rental rates differ across the population. Human capital is homogeneous if rental rates are constant. Human capital is heterogeneous if rental rates differ. Obviously, discrimination and other type "non-market" effects will weaken the test.

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. Now, rewriting (10) as

$$\ln \widehat{w_i^*} = (1 - \hat{b}_i)X_i A + ln\beta + \ln(e_i)$$
$$= (1 - \hat{b}_i)A_0 + (1 - \hat{b}_i)X_i^* A_1 + ln\beta + \ln(e_i) \qquad (10')$$

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

3.5 Precision of the Estimates: Bootstrapped Standard Errors

The data comprise 15-24 observation for each individual. To get at the precision of our estimates, we construct bootstrapped standard errors, given that asymptotic theory does not apply to so few observations.[20] Thus, we run 300 regressions for each individual with randomly drawn (with replacement) bootstrapped samples of size equal to the number of observations. From these, we construct the bootstrapped standard errors.

4. The Data

Nowadays there are a number of panel micro-data sets containing information on schooling, work experience, and earnings over the life-cycle. However, as far as we know,

---

[20] 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 (See Li and Maddala (1996)). In such a cases, one may need to apply alternative bootstrapping methods such as moving block bootstrapping to get correct standard errors.

only the National Longitudinal Survey of Youth 1979 also contains family background, extensive cognitive and non-cognitive ability measures, as well as other information from which we construct the "Big Five" personality measures. (Appendix D indicates how we create the Big Five indices.) For this reason we utilize the NLSY79 in order to compare our estimated individual-specific earnings function parameters to the independently obtained test-score measured ability, personality and family background variables contained in the NLSY79.

As is well known, the NLSY79 is a nationally representative sample of young men and women aged 14 to 22 years old when first surveyed in 1979. The surveys have been conducted annually until 1994, and then performed every other year. We utilize the 2012 NLSY79, which contains up to 24 years of data for each respondent through 2010.[21] We do not apply sampling weights since we are estimating (7') for each individual separately rather than trying to use each individual's data to build a nationwide mean.[22] To estimate (7') we use data on age (t), years of school completed, and weekly earnings (i.e., annual earnings divided by number of weeks worked) deflated by the 1982-84 urban CPI index. From these we compute the age when schooling was completed ($t*$). Because our earnings function specification is designed for those who work continuously, we concentrate only on the males given that females are more likely to have discontinuous labor force participation, making the measurement of experience (t) more difficult and resulting in a far more complex nonlinear earnings equation (Polachek, 1975). In addition, current human capital acquisition is affected by future intermittent participation. Not being able to predict when and how long a woman will drop out precludes estimating female earnings functions, at least for the purposes of this paper. 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).

---

[21] The data and further explanations can be explored from the website http://www.bls.gov/nls/
[22] We use the sample weights when we aggregate the results to get inferences about particular segments of the population.

As was already mentioned, for the purposes of this study, the main virtue of the NLSY79 data is the information on cognitive and non-cognitive ability, indicators of personality (including sufficient detail to create indices of the "Big Five" personality measures), as well as family background, all of which were obtained independent of our estimated human capital production function parameters. Of these we concentrate on the 1980 AFQT[23] and its particular cognitive skills components (general science, arithmetic reasoning, numerical operations, and math knowledge), craftsmanship skills (mechanical skills, electronics knowledge, coding speed, and automobile repair knowledge), indicators of the "Big Five" personality traits (agreeableness, extraversion, openness, conscientiousness, and neuroticism), 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 individual-specific human capital production function parameters, to skill depreciation, and to time preference parameters we estimated using (7') and (10).

5. Estimation Results

We use non-linear least-squares to evaluate (7') for each person with 15 or more years of data.[24] 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. They are an ability parameter ($\hat{b}_i$), the discount rate ($\hat{r}_i$), the human capital depreciation rate (δ), and

the composite parameters $\hat{E}_i = \dfrac{E_{0_i}}{\beta_i^{1/1-b_i}}$ and $\hat{w}_i^* = (R^{1-b_i})\beta_i$. Table 1 contains average estimates for the entire sample as well as for blacks and whites separately.[25] The mean b

---

[23] The 1980 AFQT score differs slightly from the 1989 and 2006 scores because of the way each component is weighted.

[24] Most individuals have 19 or more years of data.

[25] We do not present estimates for Hispanics. However, they are generally between whites and blacks in magnitude.

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 (δ) 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.[26] For this, we run 300 nonlinear regression replications utilizing a randomly drawn sample (with replacement) equal to the number of observations available for each person.[27] 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.

Based on the identification strategy we described earlier, we get values for b_i, β_i, E_{0_i}, E_{s_i}, δ_i, and r_i, as well as a population-wide value of R. Mean values across all individuals are given in Table 2 along with mean values for each of the ability, personality, family background and outcome measures contained in the NLSY79 that were mentioned above. Obviously, sample sizes vary since not all variables are reported for each respondent.

5.1 Consistency with Prior Population-Wide Estimates

Interestingly, mean values of our parameters compare favorably to past studies that estimate aggregate earnings functions, though understandably there are some differences due to alternative methodologies and data. For example, we obtain an r of 0.041 compared to Haley's 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 δ of 0.027 compared

---

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

[27] These computations took 281 hours using an i7-vpro chip parallel processor computer running seven STATA programs in tandem, each utilizing the GA algorithm.

to Johnson and Hebein's (1974) 0.022 and Heckman's (1976) 0.04-0.07. However, we obtain a weekly rental rate per unit of human capital (R) of $14.85 which differs from Liu's $4.70. 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.[28]

Our results are also consistent with computations of Mincer's "time-equivalent" post-school investment as well as with Ben-Porath's prediction that time-equivalent investment decreases with age. Figure 1 plots potential and actual earnings for individuals who began work immediately following school.[29] 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. Theory predicts 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 $100 per week (in 1982-84 dollars) and the modal value for potential earnings is about $250. The ratio 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.[30] Re-computing these two distributions for older workers (Figure 2) shows a definite narrowing of the distance between potential and actual earnings, as predicted by theory. 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

---

[28] See Browning, Hansen, and Heckman (1996) for a survey describing the results from a number of these studies.

[29] These exclude those with very low schooling levels and those who took a year or more to find their first job.

[30] 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.

Table 3 presents results from estimating (10) and (10'). In both, 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. It varies by 0.0125 ln points per year of school, by as much as 0.11 ln points by race, and by -0.03 ln points per week of unemployment spell. As such, we find that race explains only as much as 11% of the human capital rental rate, but this possibly includes discrimination. Similarly, each extra year of school only increases the rental rate by 1¼ percent, which is far smaller than typical schooling rate of return estimates. Further, each extra week of unemployment decreases the human capital rental rate by only 3%, but given the typical worker's two week unemployment spell, this only amounts to 6%. As such, these results are consistent with a relatively homogeneous human capital.

Another way to look at this is to compare the variation in predicted (weekly) rental rates per unit of human capital based on (10') to the variation in overall (weekly) earnings. If the relative rental rate variation is small compared to earnings variation, one can conclude that most earning variation comes about because people invest differently, not because people are remunerated differently for the same human capital stock. Based on Table 3, we find the coefficient of variation (standard deviation divided by mean) for individual-specific estimates of the rental rate to be at most 12%. The coefficient of variation for weekly earnings is 61%. As such, our results are consistent with population-wide variations in the amount of human capital, and smaller variations in type. Again, human capital appears to be relatively homogeneous.

5.3 Heterogeneity of the Ben-Porath Parameters across the Population

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 β 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. For example, b ranges from 0.01 to 0.70 with a mean of .35 and β

varies from 0.13 to 1.59 with a mean of 0.64. Thus we scale each parameter by $\dfrac{x_i - L_i}{H_i - L_i}$

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 b_i, β_i, E_{0_i}, E_{S_i}, δ_i and r_i, as well as AFQT.[31] Ability parameters b and β are relatively bell-shaped. Parameters E_0 and r are skewed to the right, and the skill depreciation parameter δ appears to have a double peak, as does AFQT.

5.4 The Distribution of the Ben-Porath Earnings Function Parameters by Race

Next, in Figure 4, we plot black and white differences in these kernel densities for AFQT, $b_i, \beta_i, E_{0_i}, E_{S_i}$ and δ_i using these same individuals.[32] Generally blacks (solid line) fare worse than whites (dashed line) since for each measure the black distribution is further to the left than the white distribution. However, noteworthy is the large difference for AFQT compared the smaller estimated ability differences we measure. Kolmogorov-Smirnov tests for the difference in these ability distributions are given in Table 4 (rows 1-3 and 7). The race differences for each distribution are significantly different statistically, but the distance measure is largest for the AFQT.

For comparison purposes we also utilize the Kolmogorv-Smirnov test to determine the similarity of black and white distributions for each other NLY79 ability, personality, and background variable used in the remainder of the paper. With the exception of r, the Rosenberg Self-esteem Measure and the Pearlin Mastery Score, all are statistically different. However, interestingly, the distance measures of all the ability test scores (AFQT and the ASVAB tests) are larger than our ability measures (b, β, and E_0) and the personality indicators (Rotter Locus of Control, Rosenberg Self-Esteem Score,

---

[31] 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 what are contained in Figure 3.
[32] We do not plot r because Table 3 (discussed below) indicates no statistically significant differences in the distribution of r by race. To conserve space, we also do no plot racial differences in the distributions of the other ability, personality or family background variables.

25

the Pearlin Mastery Measure, the CES-D 20 depression index, as well as the Big Five personality components: agreeableness, extraversion, openness, conscientiousness, and neuroticism). Race differences in the distance measure for poverty is between the AFQT and ASVAB test scores, our ability measures, and the psychology-based personality indicators. In short, the AFQT and ASVAB test scores appear to accentuate race differences compared to our and other measures.

5.5 Consistency with Human Capital Theory

Our estimated ability, skill depreciation, time preference measures as well as the AFQT scores are related to schooling in a predictable way. Table 5 contains averages of these variables by years of school. The table is divided into two panels, the left seven columns for blacks and the right seven columns for whites. Columns 1-4 contain the estimated ability coefficients, column 5 gives the estimated skill depreciation rate, column 6 the time discount rate, and column 7 presents the AFQT score. A few patterns are noteworthy. First, even within schooling groups, b, β, E_0, and E_S values are higher for whites than blacks, but only marginally so; nevertheless there remain large differences in AFQT. Second, the ability to learn parameters b and β rise significantly with years of school, as do AFQT scores. Of course, a positive correlation between this type ability and schooling level is predicted by human capital theory 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 unit of time lowers the opportunity costs of going to school, thereby increasing the amount of school purchased. On the other hand, the ability to earn parameter, E_0, does not rise with schooling level, either for blacks or whites. 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. [33] Skill depreciation decreases with schooling. This result also is as expected since a higher depreciation rate lowers the value of what is learned because more is "forgotten". Finally, the estimated time discount rate (r) decreases with the level of school. This latter result, too, is noteworthy because higher time discount rates should

---

[33] E_s is positively related to schooling since by construction it incorporates what is learned in school.

imply *fewer* years of schooling, since individuals with high discount rates are more reluctant to put off the gratification of current market earnings, given that they discount the future heavily.

These patterns are also presented in Table 6 which contains a specific regression. The dependent variable is $\log_e$ years of school completed. The independent variables are b, β, E_0, δ, and r (also in logs so that the effects of each can be compared in percent terms), cohort, race, and background variables. Ability to learn (b and β) have the largest positive effect, depreciation (δ) and the rate of time preference (r) are negatively associated with years of school, and although not negative, the ability to earn (E_0) exhibits a statistically insignificant relation.

5.6 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 parameters, which in turn were related to years of schooling in a predictable way, as was shown in Tables 5 and 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

27

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.

5.6.1 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 crystalized is supposed to be learned, in essence knowledge based. However, according to De Young, recent factor analysis by Johnson and Bouchard (2005a and 2005b) find 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 crystalized or learned ability. This makes sense because E_0 depicts one's ability to earn just *before* one begins school,

and b and β 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.

5.6.2 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." As already discussed, the basic taxonomy used to describe personality is the Big Five. 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." Evidence shows "people scoring high in Intellect will, on average, score higher in Openness." The Extraversion and Neuroticism personality categories exhibit the least cognitive ability traits (De Young, 2011). Our results confirm these assertions.

Table 8 presents correlation coefficients for each of the Big Five personality traits. Interestingly, of the five personality categories, b and β are most correlated with Openness (0.22 and 0.21). This is consistent with intelligence being key to our measures of one's ability to learn, just as DeYoung suggests. Similarly, the correlations between b and β and the Extraversion and Neuroticism personality categories are lowest (0.09) and (-.10), meaning these categories are least related to one's ability to learn. Conscientiousness is unrelated to b and β, but has a significant correlation (0.06) with one's ability to earn (E_0). "Open" and "agreeable" people tend to have the lowest skill depreciation and time preference coefficients. Neurotics have the highest.

In addition, for comparison, we present correlation coefficients for four other psychological indicators. The Rotter score (varying from 0 to 13) denotes a person's

external locus of control, defines the extent one views life chances are determined by external factors.   A related measure is the Pearlin Mastery Scale (Pearlin, et. al, 1981). 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 (which is synonymous with a positive correlation between internal locus of control and ability), 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.

5.6.3 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 as a child (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 Table7.

Interestingly, 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.

5.7 Selectivity

Equation (7'), from which we obtain Ben-Porath parameters, is estimated only for those individuals with sufficiently long work histories. Clearly this sample is a select group because it does not include those respondents with shorter (or no) work histories. The "workers" we choose may be (and most likely are) different in ability (and probably other characteristics) than the "non-workers" we do not include. As such, making inferences about the whole population using our select sample may yield biased results because our ability estimates are not averaged over the entire population. However, our purpose is not to obtain population-wide estimates. Instead, our point is to use the difference in NLY79 reported measures and our estimated Ben-Porath ability parameters for the *same* workers to identify whether the same degree of heterogeneity (particularly difference in magnitude) is prevalent in both. In essence, we utilize the *difference* between the two measures (our Ben-Porath ability estimates and the NLSY79 AFQT and other attribute values) holding observable and unobservable individual characteristics constant.

On the other hand, 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. Table 10 presents differences in AFQT scores between those working 15 or more years (our sample) and the remainder of the population (those not meeting the work requirement) broken down by level of education. In six educational

categories the AFQT advantage for "workers" is higher for whites than blacks, while in six the ability advantage favors blacks. 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 no different statistically than for white "workers" and "non-workers".[34] This result is consistent with small, if any, selectivity biases when considering racial 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) emphasized biases inherent in the representative agent framework because the approach fails to account for individual heterogeneity. They illustrate these biases using parameters based on the life-cycle model. Recent studies acknowledge this insight about hetrogeneity, 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 three 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 estimate these

---

[34] The actual test comprises an insignificant α_3 white-worker interaction coefficient in the following regression: $AFQT_i = \alpha_0 + \alpha_1\ white_i + \alpha_2\ worker_i + \alpha_3\ white_i * worker_i + \epsilon_i$.

parameters, individual-by-individual, using a new genetic algorithm, a maximization routine first introduced to analyze the genome. Third, we prove these parameters are identified by using a criterion based on power series expansions first developed in bioscience research.

A number of new findings emerge from our analysis. First, we find that black-white differences in Ben-Porath ability parameters appear smaller than those obtained from AFQT test scores, and its components. On the other hand, 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, a higher discount rate and a greater degree of skill depreciation are associated with fewer years of school. Third, we confirm a number of findings in the psychology literature. For example, of the Big Five personality traits, we find Openness to be most related to ability and Neuroticism and Extroversion least related. Fourth, we find a number of skill, personality, and family background variables to be related to ability. For example, Conscientiousness is related to an individual's ability to earn, but *not* to an individual's ability to learn. On the other hand, Agreeableness, Extroversion and especially Openness is positively related with one's ability to learn, but *not* to one's ability to earn. Similarly, a high internal locus of control, a high mastery (belief one controls events) are related to a high ability to learn, but unrelated to one's ability to earn. Individuals scoring high on the AFQT also exhibit a high ability to learn and only a marginally higher ability to earn. Finally, individuals displaying high degrees of Neuroticism exhibit a high discount rate and a low ability to learn.

In short, Ben-Porath human capital production function parameters are heterogeneous across the population. They are related to innate verbal versus mathematical abilities, underlying psychological characteristics, and family background variables both in ways one would expect, and in ways not previously studied.

Of course, employing an earnings equation based on a life-cycle model to estimate person-specific abilities and skill-depreciation and time preference parameters are not free of bias. For example, both discrimination in the availability of high quality schooling, as well as discrimination in the labor market itself can cause racial differences in estimating ability parameters using earnings data. However, if so, racial discrimination in the labor market manifests itself in a similar way cultural biases might inculcate psychology-based models.

Technical simplifications could also mar interpretation of our results. Underlying our approach are the typical assumptions incorporated in life-cycle models. Obviously, our results may be suspect if earnings are determined by other frameworks such as incentive contracts or deferred compensation schemes. In addition, for computational simplicity, we utilize a relatively simple human capital production function. We envision more complicated versions incorporating non-cognitive skills directly into the human capital production function. These latter models would yield more complex earnings functions than the ones we already use. On the other hand, we feel strongly that our results are not simply verifying the well-known fact that high ability people simply earn more. Our ability measures encompass more than earnings level; they primarily arise from the curvature of the earnings profile.

Our results are promising enough to warrant pursuing the approach further. For example, zeroing in more on the various types of ability might enable one to gain insights into occupational choice decisions including answering questions relating to gender differences in one's inclination to go into scientific professions. In any case, the whole approach gets at heterogeneity across the population in a new way. Also, the data we generated on ability, skill depreciation, and time discount rates can be valuable to analyze other aspects of human behavior.

Figure 1*
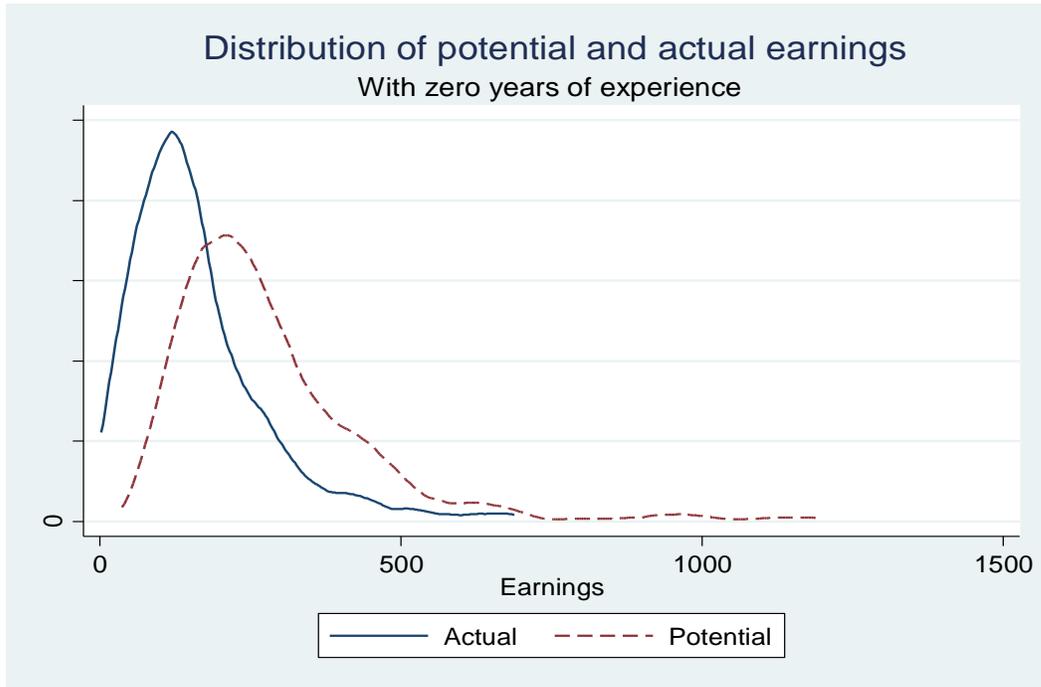
Distribution of potential and actual earnings
With zero years of experience

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

Figure 2*
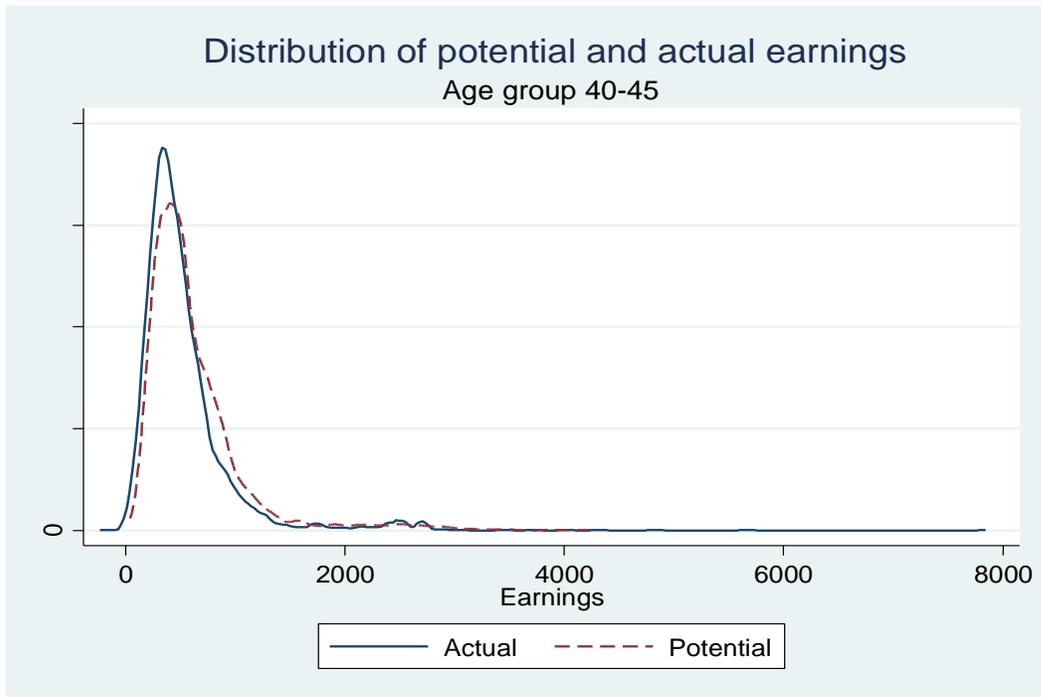
Distribution of potential and actual earnings
Age group 40-45

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

Figure 3
Distribution of Parameter Values*
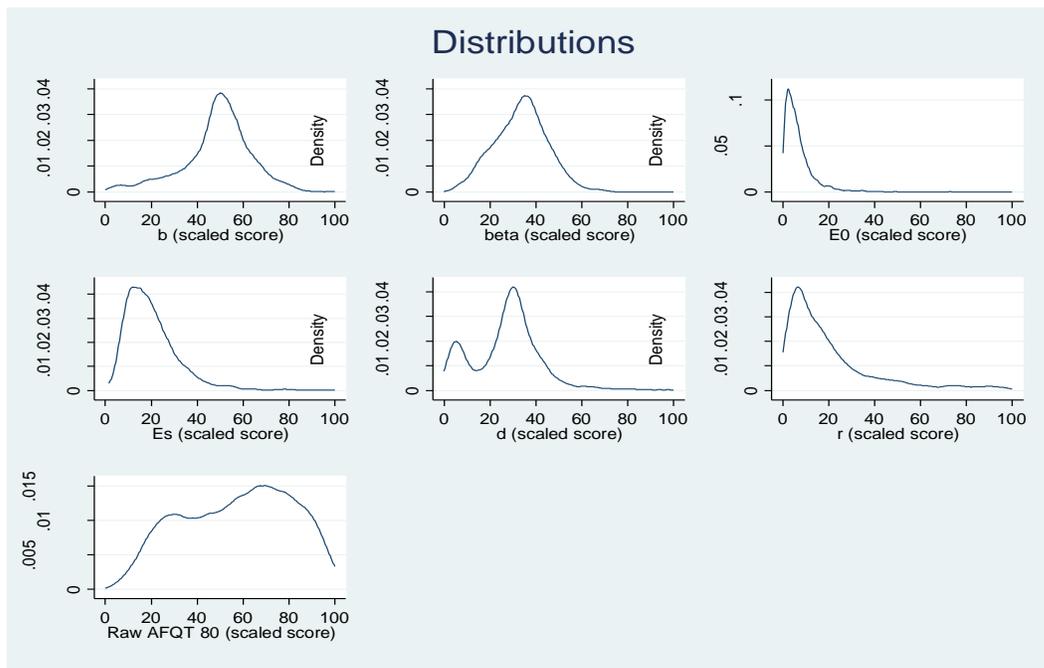
Distributions

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

Figure 4
Distribution of Parameter Values*
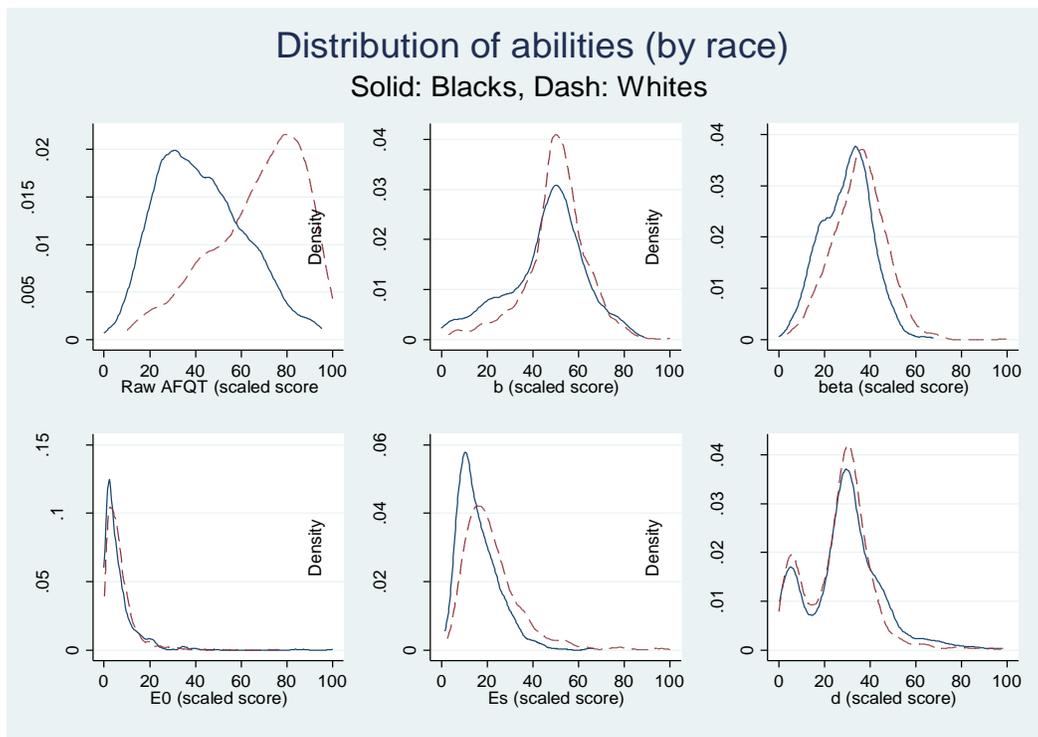
Distribution of abilities (by race)
Solid: Blacks, Dash: Whites

* 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	δ	r
All	0.351	3.808	5.392	0.027	0.041
Bootstrapped SE(mean)	0.054	0.778	2.98	0.008	0.022
Bootstrapped SE(median)	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	0.323	3.748	6.253	0.029	0.043
Bootstrapped SE(mean)	0.062	0.866	3.470	0.009	0.024
Bootstrapped SE(median)	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	0.355	3.817	5.265	0.026	0.041
Bootstrapped SE(mean)	0.049	0.735	2.73	0.0075	0.02
Bootstrapped SE(median)	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: Mean Values of Individual-Specific Parameters and Ability, Personality, and Family Background (Descriptive Statistics (weighted))*

Variables	Observation	Means	SD
b	1826	0.35	0.10
β	1826	0.64	0.17
E_0	1826	2.75	2.79
Es	1826	18.11	10.53
δ	1826	0.03	0.01
R	1826	0.04	0.04
AFQT, 80	1782	48.22	28.91
Locus of control (Rotter)	1807	8.49	2.36
Self-esteem (Rosenberg)	1808	22.65	3.98
Pearlin mastery	1790	22.24	2.99
ASVAB, general science	1784	16.41	4.87
ASVAB, arithmetic	1784	18.48	7.27
ASVAB, word knowledge	1784	25.43	7.72
ASVAB, paragraph comprehension	1784	10.43	3.48
ASVAB, numeric ability	1784	32.84	10.84
ASVAB, coding speed	1784	41.69	15.13
ASVAB, auto shop knowledge	1784	17.04	5.35
ASVAB, math knowledge	1784	13.37	6.26
ASVAB, mechanical knowledge	1784	15.89	5.28
ASVAB, electronics	1784	12.79	4.19
Agreeableness (principal component)**	897	60.24	25.60
Extraversion (principal component)**	718	55.47	17.94
Openness (principal component)**	996	63.02	20.29
Conscientiousness (principal component)**	1572	21.76	18.38
Neuroticism (principal component)**	1759	11.15	13.77
Mother's years of schooling	1721	11.73	2.33
Father's years of schooling	1596	11.86	3.25
Fathers occupation (if professional/managerial)	1826	0.24	0.43
If urban	1820	0.75	0.43
Poverty	1715	0.09	0.28
Household subscribes magazine (at age 14)	1812	0.69	0.46

* Computed from our estimates of (7'), (10) and data contained in the NLSY79. Variable definitions are given in text.

** POMP scaled values of principal component given for Big Five personality variables.

Table 3: Determination of Rental Rates per Unit of Human Capital*

RHS VARIABLES	Equation	
	(10)	(10')
(1-b)	2.698***	2.722***
	(0.0637)	(0.170)
(1-b)*school		0.0125**
		(0.00556)
(1-b)*unemployment		-0.0320***
		(0.00335)
(1-b)*white		0.114***
		(0.0306)
(1-b)*Hispanic		0.0636
		(0.0474)
(1-b)*urban		0.0586**
		(0.0237)
(1-b)*manufacturing		0.0289
		(0.0267)
(1-b)*father professional or managerial		0.0436*
		(0.0258)
(1-b)*cohort-age14in1979		-0.000897
		(0.0653)
(1-b)*cohort-age15in1979		0.0515
		(0.0616)
(1-b)*cohort-age16in1979		0.0412
		(0.0617)
(1-b)*cohort-age17in 1979		0.0489
		(0.0612)
(1-b)*cohort-age18in1979		0.0308
		(0.0602)
(1-b)*cohort-age19in1979		0.0206
		(0.0607)
(1-b)*cohort-age20in1979		-0.0111
		(0.0613)
(1-b)*cohort-age21in1979		0.0487
		(0.0611)
(1-b)*professional		0.135
		(0.130)
(1-b)*service		-0.0309
		(0.132)
(1-b)*sales		0.154
		(0.131)
(1-b)*construction		0.0764
		(0.130)
(1-b)*farm		-0.0962
		(0.164)
(1-b)*production worker		0.0816
		(0.130)
Constant	-0.491***	-0.740***
	(0.0419)	(0.0468)
Observations	2,262	1,915
R-squared	0.442	0.511

* Estimation of (10) and (10'). The dependent variable is $ln\widehat{w}_i^*$. 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 4: Kolmogorov-Smirnov Test Comparing the Black and White Distributions of the Indicated Variable*

Variable	Distance	P-value
b	0.140	0.000
β	0.199	0.000
E_0	0.119	0.000
Es	0.252	0.000
δ	0.097	0.001
R	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'), (10) and data contained in the NLSY79.

Table 5: Ben-Porath Parameters by Race, Level of Schooling*

School				Blacks							Whites			
	b	β	E_0	Es	δ	r	AFQT 1980	b	β	E_0	Es	δ	r	AFQT 1980
<12	0.31	0.52	3.15	9.60	0.033	0.053	9.6	0.32	0.61	2.91	11.72	0.030	0.049	22.1
12	0.32	0.55	2.60	11.78	0.029	0.046	15.6	0.34	0.65	2.84	15.84	0.028	0.045	44.4
13-15	0.32	0.57	2.88	14.78	0.028	0.037	24.3	0.36	0.64	2.68	18.51	0.025	0.040	59.0
>=16	0.37	0.68	2.54	23.59	0.027	0.029	44.2	0.40	0.67	2.56	28.62	0.022	0.029	77.2
All	0.32	0.57	2.73	13.67	0.029	0.043	20.4	0.36	0.65	2.76	18.76	0.026	0.041	52.4

* Computed mean values of parameters for equations (7') and (10) weighted using NLSY79 weights.

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

Variables	Coefficient.	t-value
ln(b)	0.0528	5.81
ln(β)	0.0716	4.69
ln(E0)	0.0058	1.26
ln(δ)	-0.0342	-6.49
ln(r)	-0.0322	-7.98
Constant	2.3349	68.52
Observations		1701
Adj. R2		0.2356

* Dependent Variable: Ln(Completed Years of School) from NLSY79. Independent variables are individual-specific coefficient estimates of (7') and (10). 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*

	b	β	E_0	δ	r
Cognitive Measures					
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
Obs	1784	1784	1784	1784	1784
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
Obs	1784	1784	1784	1784	1784
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
Obs	1784	1784	1784	1784	1784
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
Obs	1784	1784	1784	1784	1784
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
Obs	1784	1784	1784	1784	1784
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
Obs	1784	1784	1784	1784	1784
Craftsmanship Skills					
Mechanical	0.149	0.167	0.043	-0.132	-0.005
Sig (p-value)	0.000	0.000	0.069	0.000	0.838
Obs	1784	1784	1784	1784	1784
Electronics	0.137	0.198	0.080	-0.116	0.020
Sig (p-value)	0.000	0.000	0.001	0.000	0.403
Obs	1784	1784	1784	1784	1784
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
Obs	1784	1784	1784	1784	1784
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
Obs	1784	1784	1784	1784	1784

* Computed from our estimates of (7'), (10) and data contained in the NLSY79.

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

	b	β	E_0	δ	r
Big Five Personality Traits					
Agreeableness	0.137	0.138	0.058	-0.105	0.019
Sig(p-value)	0.000	0.000	0.081	0.002	0.575
Obs	897	897	897	897	897
Extraversion	0.090	0.075	0.025	-0.046	-0.009
Sig(p-value)	0.016	0.046	0.501	0.218	0.809
Obs	718	718	718	718	718
Openness	0.221	0.213	0.043	-0.159	-0.064
Sig(p-value)	0.000	0.000	0.172	0.000	0.042
Obs	996	996	996	996	996
Conscientiousness	0.023	0.031	0.057	-0.033	0.049
Sig(p-value)	0.367	0.216	0.023	0.190	0.053
Obs	1572	1572	1572	1572	1572
Neuroticism	-0.100	-0.075	0.027	0.075	0.064
Sig(p-value)	0.000	0.002	0.260	0.002	0.007
Obs	1759	1759	1759	1759	1759
Other Personality Meaures					
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
Obs	1807	1807	1807	1807	1807
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
Obs	1808	1808	1808	1808	1808
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
Obs	1790	1790	1790	1790	1790
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
Obs	1790	1790	1790	1790	1790

* Computed from our estimates of (7'), (10) and data contained in the NLSY79.

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

	b	β	E_0	δ	r
Mothers schooling	0.162	0.067	-0.039	-0.114	-0.076
	0.000	0.006	0.104	0.000	0.002
	1721	1721	1721	1721	1721
Father's schooling	0.151	0.127	-0.039	-0.097	-0.100
	0.000	0.000	0.115	0.000	0.000
	1596	1596	1596	1596	1596
Professional/Managerial Father	0.138	0.088	-0.007	-0.085	-0.063
	0.000	0.000	0.764	0.000	0.007
	1826	1826	1826	1826	1826
HH poverty, 1978	-0.081	-0.149	-0.049	0.034	-0.037
	0.001	0.000	0.042	0.161	0.131
	1715	1715	1715	1715	1715
HH had magazine, age 14	0.153	0.099	0.011	-0.134	-0.062
	0.000	0.000	0.654	0.000	0.008
	1812	1812	1812	1812	1812
Urban	0.022	0.048	0.015-	0.016	-.006-
	0.340	0.040	0.520	0.500	0.780
	1820	1820	1820	1820	1820

* Computed from our estimates of (7'), (10) and data contained in the NLSY79.

Table 10: AFQT Differences Between Workers and Non-Workers*

Years of school	Diff in AFQT, Black (workers - non-workers)	Diff in AFQT, Whites (workers - non-workers)
9	-6.60	-5.38
10	4.32	-7.05
11	1.44	-1.42
12	0.52	2.28
13	-1.07	4.50
14	-6.76	-1.85
15	-11.79	4.76
16	7.85	7.07
17	-8.95	1.27
18	7.59	0.70
19	11.79	4.46
20	30.65	9.56
All	-2.20	-4.18

Data computed from NLSY79.

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.[35] 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

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

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 certainly. Disposable earnings are

$$Y_t = R[E_t - K_t] \tag{A2}$$

where R is the rental rate per unit of human capital,[36] 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

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

where we assume δ 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 $Q_t = \beta K^b$.

Maximization of (A1) subject to equations (A2), (A3) and (A4) entails maximizing the Hamiltonian

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

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

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

$$\frac{\partial H}{\partial K_t} = 0 \tag{A5.1}$$

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

---

[35] As noted in the text, we abstract from labor supply considerations.
[36] In the empirical work we test whether R varies by race, occupation, industry and other variables.

$$\frac{\partial H}{\partial \lambda_t} = \dot{E}_t \tag{A5.3}$$

$$\lambda_N = 0 \tag{A5.4}$$

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

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

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

From (A5.1)

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

implying that

$$K_t = \left[\frac{\mathrm{Re}^{-rt}}{\lambda_t \beta}\right]^{1/b-1} = \left[\frac{b\beta e^{rt}\lambda_t}{R}\right]^{1/1-b} \tag{A7}$$

Substituting (A6) into (A7) yields

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

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 hypoergeometric series

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

where

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

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 points. A simpler form can be obtained by setting the convergence at 4 decimal points

level. 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

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

Thus,

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

or,

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

Thus the stock of human capital at time t can be expressed as

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

or

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

or

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

or

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

$$\text{(A12)}$$

Observed earnings can be expressed as following

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

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

$$Y_t = A_0 e^{\delta(t^*-t)} + A_1[1 - e^{\delta(t^*-t)}] - A_2[1 - e^{(r+\delta)(t-N)}]^{1/(1-b)} \qquad \text{(A13)}$$

where

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

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

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

Appendix B: Derivation of Estimation Equation (7')

The following provides a formal discussion of the approximation of the third term of equation (7) and derives estimation equation (7').

Let $x = e^{(r+\delta)(t-N)}$. Writing $[1 - e^{(r+\delta)(t-N)}]^{1/(1-b)}$ in terms of x yields $f(x) = [1 - x]^{\frac{1}{1-b}}$. Approximating $f(x)$ with a second degree Taylor's series approximation around $x = 0$ yields

$$f(x) = f(0) + f'(0)x + f''(0)x^2$$
or
$$f(x) = 1 - \frac{1}{(1-b)}e^{(r+\delta)(t-N)} + \frac{1}{2!}\frac{1}{(1-b)}\frac{b}{(1-b)}e^{2(r+\delta)(t-N)}$$

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 1, Figure 2, Figure 3. All three figure show the approximated function closely matches the actual function. Next we substitute $[1 - e^{(r+\delta)(t-N)}]^{1/(1-b)}$ by the approximated $f(x)$ to construct our approximated earnings function. This yields the following

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

$$= (A_0 - A_1)e^{\delta(t^*-t)} + (A_1 - A_2) + A_2\alpha\,e^{(r+\delta)(t-n)} - \frac{1}{2}A_2\alpha(\alpha-1)e^{2(r+\delta)(t-N)}.$$

Given $= \frac{1}{1-b}$ $W = R^{(1-b)}\beta$, we obtain

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

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

$$\left. - \left\{0.5[\frac{b}{r+\delta}]^{\frac{1}{(1-b)}}\left(\frac{1}{(1-b)}\right)\left(\frac{b}{(1-b)}\right)e^{2(r+\delta)(t-N)}\right\}\right] \qquad (7')$$

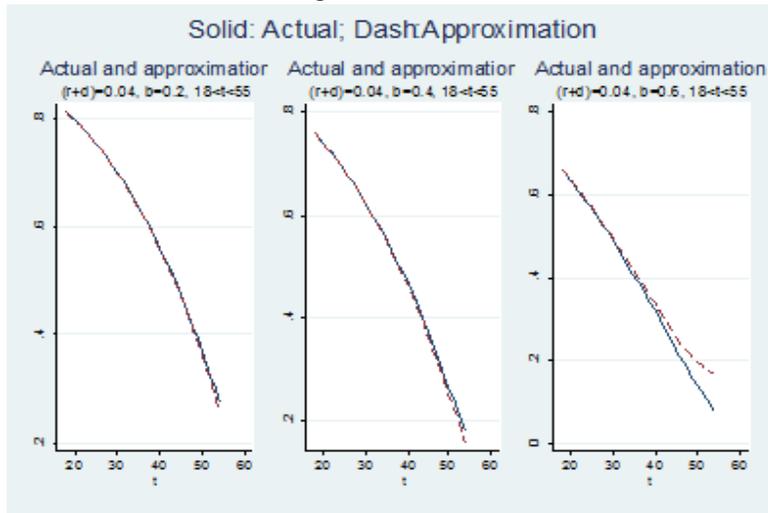We fit equation (7') to estimate b, W, E, δ, and r.

50

Figure A.1

Solid: Actual; Dash:Approximation

Figure A.2

Solid: Actual; Dash:Approximation
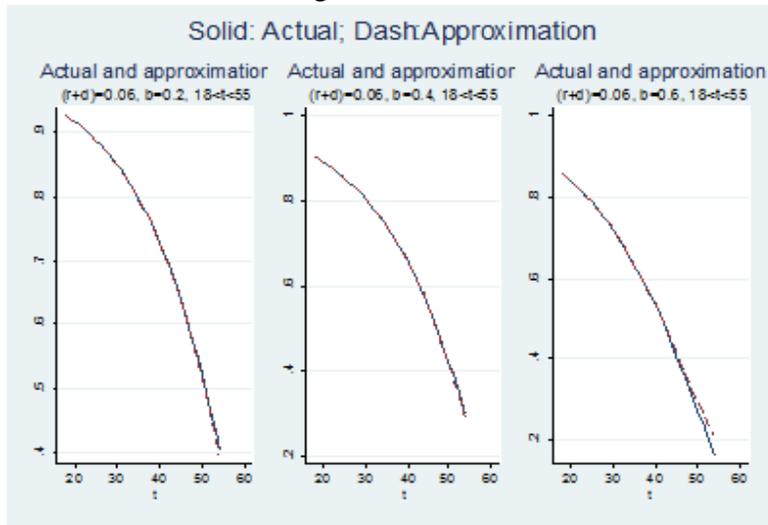
Figure A.3

Solid: Actual; Dash:Approximation

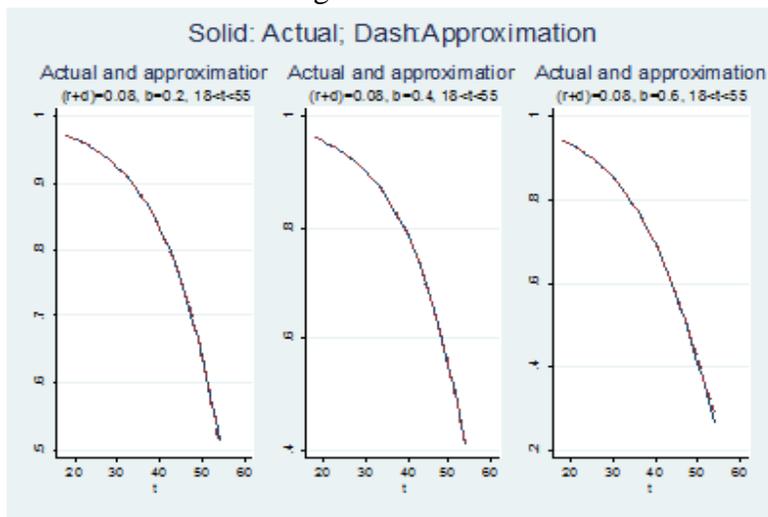## Appendix C: Structural Identification

As per Pohjampalo (1978) the power series expansion with the first five derivatives yields the following system of equations:

$$\frac{dY}{dt}(t=0) = -\delta M + \theta K - 2\theta L = z1 \qquad (C1)$$

$$\frac{d^2Y}{dt^2}(t=0) = \delta^2 M + \theta^2 K - 4\theta^2 L = z2 \qquad (C2)$$

$$\frac{d^3Y}{dt^3}(t=0) = -\delta^3 M + \theta^3 K - 8\theta^3 L = z3 \qquad (C3)$$

$$\frac{d^4Y}{dt^4}(t=0) = \delta^4 M + \theta^4 K - 16\theta^4 L = z4 \qquad (C4)$$

$$\frac{d^5Y}{dt^5}(t=0) = -\delta^5 M + \theta^5 K - 32\theta^5 L = z5 \qquad (C5)$$

Multiplying (C1) by 2θ and subtracting (C5) yields

$$-\delta M(2\theta + \delta) = 2\theta z1 - z2 - \theta^2 K \qquad (C6)$$

Multiplying (C2) by 2θ and subtracting (C3)

$$\delta^2 M(2\theta + \delta) = 2\theta z2 - z3 - \theta^3 K \qquad (C7)$$

Multiplying (C3) by 2θ and subtracting (C4) yields

$$-\delta^3 M(2\theta + \delta) = 2\theta z3 - z4 - \theta^4 K \qquad (C8)$$

Multiplying (C4) by 2θ and subtracting (C5) yields

$$\delta^4 M(2\theta + \delta) = 2\theta z4 - z5 - \theta^5 K \qquad (C9)$$

Dividing (C7) by (C6) yields

$$-\delta = \frac{2\theta z2 - z3 - \theta^3 K}{2\theta z1 - z2 - \theta^2 K} \qquad (C10)$$

Dividing (C8) by (C7) yields

$$-\delta = \frac{2\theta z3 - z4 - \theta^4 K}{2\theta z2 - z3 - \theta^3 K} \qquad (C11)$$

Dividing (C9) by (C8) yields

$$-\delta = \frac{2\theta z4 - z5 - \theta^5 K}{2\theta z3 - z4 - \theta^4 K} \qquad (C12)$$

Combining (C10) and (C11)

$$\frac{2\theta z2 - z3 - \theta^3 K}{2\theta z1 - z2 - \theta^2 K} = \frac{2\theta z3 - z4 - \theta^4 K}{2\theta z2 - z3 - \theta^3 K} \tag{C13}$$

Combining (C11) and (C12)

$$\frac{2\theta z3 - z4 - \theta^4 K}{2\theta z2 - z3 - \theta^3 K} = \frac{2\theta z4 - z5 - \theta^5 K}{2\theta z3 - z4 - \theta^4 K} \tag{C14}$$

Solving (C14) for K in terms of θ and z_i and substituting in (C13) yields the following equation

$$\frac{2\theta^2 z3 - 3\theta z4 + z5}{2\theta^2 z2 - 3\theta z3 + z4} = \tag{C15}$$

$$\frac{\theta(2\theta^2 z2 - 3\theta z3 + z4)^2}{4\theta^5 z1z2 - 2\theta^4(5z1z3 + z2^2) + \theta^3(8z1z4 + 5z2z3) + \theta^2(-2z1z5 - 8z2z4 + 4z3^2) + \theta(3z2z5 - 2z3z4) - z3z5 + z4^2}$$

Equation (C15) is a 7th degree polynomial in θ implying at most seven solutions if (C15) is well behaved. Solving for θ using MATHEMATICA yields seven solutions for θ. This means that we also obtain seven solutions for M, K, L, and δ. As per Pohjampalo (1978), this number of solutions means that M, K, L, θ, δ, and hence b, W, E, r, δ are locally identified.

Appendix D: Deriving Indices of the Big Five Personality Traits from the NLSY79

NLSY79 does not provide direct measures of the Big Five psychological traits. However, it contains a wide array of information closely resembling the components of the Big Five. Based on John, Neuman, and Soto's (2008) comprehensive classification of the Big Five, we select a set of variables from the NLSY79 which closely match their classification. Rather than do an exact inter-rater reliability test, we instead consulted with psychologists to assess the validity of our NLSY Big Five indicators. These are given in Appendix Table D-1. The five double columns give the Big Five personality traits. The first of the double columns give the John, Neuman, and Soto (JNS) descriptors. The second gives the NLSY79 variables we use. From these we construct one indicator for each of the five categories based on principal component analysis. For example, to compute the principal component of Extraversion, we utilize job characteristics data on dealing with people and sociability.

Appendix Table D-1
Variables Encompassing the Big Five Personality Traits

Openness		Conscientiousness		Extraversion		Agreeableness		Neuroticism	
JNS* Descriptors	NLSY79 Variables	JNS* Descriptors	NLSY79 Variables	JNS* Descriptors	NLSY79 Variables	JNS* Descriptors	NLSY79 Variables	JNS* Descriptors	NLSY79 Variables
Imaginative	Autonomy (Job characteristics file)	Responsible, dependable	Responsible for HH work such as cooking, laundry, dishwashing, grocery etc. (Time Use).	Assertive, outspoken	Dealing with people (Job characteristics); Dealing with people (Time Use)	Pleasant; Friendly	Friendship (Job characteristics)	Tense; Anxious; Nervous	Fearful (CES-Depression); Restless sleep (CES-Depression)
Curious	Mathematical knowledge (ASVAB); Word knowledge (ASVAB)	Organized	Inverse of: Could not get going (CES-Depression)	Sociable	Sociability at age 6; sociability at when adult (1985)	Generous	Donation, money, property >$25	Despondent; Unstable; Emotional	Inverse of: Hopeful (CES-Depression); Felt failure (CES-Depression); Unable to shake blues (CES-Depression); Inverse of: Felt good as other.
Intelligence	Arithmetic (ASVAB); General science (ASVAB); Numeric operation (ASVAB)	Slipshod	Easily bothered by things (CES-Depression)					Self-pity, Self-punishing	Inverse of: Happy (CES-Depression); Sad(CES-Depression); Inverse of: Enjoyed life; lonely (CES-Depression)

* Descriptors obtained from John, Naumann and Soto, 2008..

References

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life-Cycle of Earnings." *Journal of Political Economy,* 75 (4): 352-365.

Boivin, Jean, Marc Giannoni and Ilian Mihov. 2009. "Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data." *American Economic Review*, 99(1): 350-84.

Bonabeau, Eric. 2002. "Agent-Based Modeling: Methods and Techniques for Simulating Human Systems." *Proceeding of the National Academy of Sciences,* 99 (Suppl 3): 7280-7287.

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.

Browning, Martin, Metter 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.

Chiappori P.A. and I. Ekeland. 2009. "The Microeconomics of Efficient Group Behavior: Identification." *Econometrica,*77: 763—799.

Christiano L., M. Eichenbaum and C. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy,* 113(1):1-45.

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.

Eysenck, H. J. 1994. "Personality and intelligence: Psychometric and Experimental Approaches. In R. J. Sternberg and P. Ruzgis (Eds.), *Personality and Iintelligence.* New York: Cambridge University Press. pp. 3-31.

Goldberg, L. R. 1992. "The Development of Markers for the Big Five Factor Structure." *Psychological Assessment* **4** (1): 26–42.

Gottfredson, L. S. 1997. "Mainstream Science on Intelligence: An Editorial with 52 Signatories, History, and Bibliography." *Intelligence,* 24: 13-23.

Haley, William J. 1976 "Estimation of the Earnings Profile from Optimal Human Capital Accumulation." *Econometrica*, 44 (6): 1223-1238.

Heckman, James. 1975. "Estimates of Aging Capital Production Function Embedded in a Life-Cycle Model of Labor Supply." in Nester Terleckyji, 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.

Hildreth C. and J. Houck. 1968. "Some Estimators for a Linear Model with Random Coefficients." *Journal of the American Statistical Association,* 63: 584-595.

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.

Huggett, Mark, Ventura Gustavo, and Yaron Amir. 2010. "Sources of Lifetime Inequality." Working Paper.

Jaimovich Nir and Sergio Rebelo. 2009. "Can News about the Future Drive the Business Cycle?" *American Economic Review,* 99:1097-1118.

John, Oliver, Laura Naumann and Christopher Soto, 2008. "Paradigm Shift to the Integrative Big Five Trait Taxonomy: History, Measurement, and Conceptual Issues." in O. P. John, R. W. Robins, and L. A. Pervin (Eds.), *Handbook of Personality: Theory and Research*. New York, NY: Guilford Press. pp. 114-158.

Johnson, Thomas 1970 "Returns from Investment Human Capital," *American Economic Review,* 60(4): 546 560.

Johnson, Thomas. 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.

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.

Liu, Huju. 2009. "Life Cycle Human Capital Formation, Search Intensity and Wage Dynamics." University of Western Ontario Working Paper.

McAdams, D. P. and J.L. Pals. 2006. "A new Big Five: Fundamental Principles for an Integrative Science of Personality. *American Psychologist, 6:,* 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.

Parry, Ian and Kenneth Small. 2009. Should Urban Transit Subsidies Be Reduced? *American Economic Review,* 99: 700-724.

Pearlin, L., M. Lieberman, E. Menaghan and J. Mullan. 1981. "The Stress Process." *Journal of Health and Social Behavior,* 22: 337-356.

Pesaran, H. 2006. "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure." *Econometrica,* 74: 967-1012.

Pohjanpalo, H. 1978. "System Identifiability Based on the Power Series Expansion of the Solution." *Mathematical Biosciences*, 41:21–33.

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, Moon-Kak Kim. 1994. "Panel Estimates of the Gender Wage Gap: Individual-Specific Intercept and Individual Specific Slope Models." *Journal of Econometrics,* 61: 23-42.

Polachek, Solomon and Francis Horvath. 1977. "A Life-Cycle Approach to Migration: Analysis of the Perspicacious Peregrinator." *Research in Labor Economics,* 1:103-49.

Racine J and Q. Li. 2004. "Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data." *Journal of Econometrics,* 119: 99-130.

Rosenberg, M. 1965. Society and the Adolescent Self-Image." Princeton, NJ: Princeton University Press.

Rubin H. 1950. "Note on Random Coefficients." in T. Koopmans *Statistical Inference in Dynamic Economic Models* (NY and Londaon: John Wiley & Sons, Inc.). pp. 419-421.

Smets, F. and R. Wouters. 2003. "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association,* 1:1123-75.

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.