CHAPTER 1

The Basic Theory of Human Capital

1. General Issues

One of the most important ideas in labor economics is to think of the set of marketable skills of workers as a form of capital in which workers make a variety of investments. This perspective is important in understanding both investment incentives, and the structure of wages and earnings.

Loosely speaking, human capital corresponds to any stock of knowledge or characteristics the worker has (either innate or acquired) that contributes to his or her "productivity". This definition is broad, and this has both advantages and disadvantages. The advantages are clear: it enables us to think of not only the years of schooling, but also of a variety of other characteristics as part of human capital investments. These include school quality, training, attitudes towards work, etc. Using this type of reasoning, we can make some progress towards understanding some of the differences in earnings across workers that are not accounted by schooling differences alone.

The disadvantages are also related. At some level, we can push this notion of human capital too far, and think of every difference in remuneration that we observe in the labor market as due to human capital. For example, if I am paid less than another Ph.D., that must be because I have lower "skills" in some other dimension that's not being measured by my years of schooling—this is the famous (or infamous) *unobserved heterogeneity* issue. The presumption that all pay differences are related to skills (even if these skills are unobserved to the economists in the standard data sets) is not a bad place to start when we want to impose a conceptual structure on empirical wage distributions, but there are many notable exceptions, some of which will be discussed later. Here it is useful to mention three:

(1) Compensating differentials: a worker may be paid less in money, because he is receiving part of his compensation in terms of other (hard-to-observe) characteristics of the job, which may include lower effort requirements, more pleasant working conditions, better amenities etc.
(2) Labor market imperfections: two workers with the same human capital may be paid different wages because jobs differ in terms of their productivity and pay, and one of them ended up matching with the high productivity job, while the other has matched with the low productivity one.
(3) Taste-based discrimination: employers may pay a lower wage to a worker because of the worker's gender or race due to their prejudices.

In interpreting wage differences, and therefore in thinking of human capital investments and the incentives for investment, it is important to strike the right balance between assigning earning differences to unobserved heterogeneity, compensating wage differentials and labor market imperfections.

2. Uses of Human Capital

The standard approach in labor economics views human capital as a set of skills/characteristics that increase a worker's productivity. This is a useful starting place, and for most practical purposes quite sufficient. Nevertheless, it may be useful to distinguish between some complementary/alternative ways of thinking of human capital. Here is a possible classification:

(1) The Becker view: human capital is directly useful in the production process. More explicitly, human capital increases a worker's productivity in all tasks, though possibly differentially in different tasks, organizations, and situations. In this view, although the role of human capital in the production process may be quite complex, there is a sense in which we can think of it as represented (representable) by a unidimensional object, such as the stock

of knowledge or skills, h, and this stock is directly part of the production function.

(2) The Gardener view: according to this view, we should not think of human capital as unidimensional, since there are many many dimensions or types of skills. A simple version of this approach would emphasize mental vs. physical abilities as different skills. Let us dub this the Gardener view after the work by the social psychologist Howard Gardener, who contributed to the development of multiple-intelligences theory, in particular emphasizing how many geniuses/famous personalities were very "unskilled" in some other dimensions.

(3) The Schultz/Nelson-Phelps view: human capital is viewed mostly as the capacity to adapt. According to this approach, human capital is especially useful in dealing with "disequilibrium" situations, or more generally, with situations in which there is a changing environment, and workers have to adapt to this.

(4) The Bowles-Gintis view: "human capital" is the capacity to work in organizations, obey orders, in short, adapt to life in a hierarchical/capitalist society. According to this view, the main role of schools is to instill in individuals the "correct" ideology and approach towards life.

(5) The Spence view: observable measures of human capital are more a signal of ability than characteristics independently useful in the production process.

Despite their differences, the first three views are quite similar, in that "human capital" will be valued in the market because it increases firms' profits. This is straightforward in the Becker and Schultz views, but also similar in the Gardener view. In fact, in many applications, labor economists' view of human capital would be a mixture of these three approaches. Even the Bowles-Gintis view has very similar implications. Here, firms would pay higher wages to educated workers because these workers will be more useful to the firm as they will obey orders better and will be more reliable members of the firm's hierarchy. The Spence view is different from the others, however, in that observable measures of human capital may be rewarded because they are signals about some other characteristics of workers. We will discuss different implications of these views below.

3. Sources of Human Capital Differences

It is useful to think of the possible sources of human capital differences before discussing the incentives to invest in human capital:

(1) Innate ability: workers can have different amounts of skills/human capital because of innate differences. Research in biology/social biology has documented that there is some component of IQ which is genetic in origin (there is a heated debate about the exact importance of this component, and some economists have also taken part in this). The relevance of this observation for labor economics is twofold: (i) there is likely to be heterogeneity in human capital even when individuals have access to the same investment opportunities and the same economic constraints; (ii) in empirical applications, we have to find a way of dealing with this source of differences in human capital, especially when it's likely to be correlated with other variables of interest.

(2) Schooling: this has been the focus of much research, since it is the most easily observable component of human capital investments. It has to be borne in mind, however, that the R^2 of earnings regressions that control for schooling is relatively small, suggesting that schooling differences account for a relatively small fraction of the differences in earnings. Therefore, there is much more to human capital than schooling. Nevertheless, the analysis of schooling is likely to be very informative if we presume that the same forces that affect schooling investments are also likely to affect non-schooling investments. So we can infer from the patterns of schooling investments what may be happening to non-schooling investments, which are more difficult to observe.

(3) School quality and non-schooling investments: a pair of identical twins who grew up in the same environment until the age of 6, and then completed the same years of schooling may nevertheless have different amounts of human capital. This could be because they attended different schools with varying qualities, but it could also be the case even if they went to the same school. In this latter case, for one reason or another, they may have chosen to make different investments in other components of their human capital (one may have worked harder, or studied especially for some subjects, or because of a variety of choices/circumstances, one may have become more assertive, better at communicating, etc.). Many economists believe that these "unobserved" skills are very important in understanding the structure of wages (and the changes in the structure of wages). The problem is that we do not have good data on these components of human capital. Nevertheless, we will see different ways of inferring what's happening to these dimensions of human capital below.

(4) Training: this is the component of human capital that workers acquire after schooling, often associated with some set of skills useful for a particular industry, or useful with a particular set of technologies. At some level, training is very similar to schooling in that the worker, at least to some degree, controls how much to invest. But it is also much more complex, since it is difficult for a worker to make training investments by himself. The firm also needs to invest in the training of the workers, and often ends up bearing a large fraction of the costs of these training investments. The role of the firm is even greater once we take into account that training has a significant "matching" component in the sense that it is most useful for the worker to invest in a set of specific technologies that the firm will be using in the future. So training is often a joint investment by firms and workers, complicating the analysis.

(5) Pre-labor market influences: there is increasing recognition among economists that peer group effects to which individuals are exposed before they join the labor market may also affect their human capital significantly. At some level, the analysis of these pre-labor market influences may be "sociological". But it also has an element of investment. For example, an altruistic parent deciding where to live is also deciding whether her offspring will be exposed to good or less good pre-labor market influences. Therefore, some of the same issues that arise in thinking about the theory of schooling and training will apply in this context too.

4. Human Capital Investments and The Separation Theorem

Let us start with the partial equilibrium schooling decisions and establish a simple general result, sometimes referred to as a "separation theorem" for human capital investments. We set up the basic model in continuous time for simplicity.

Consider the schooling decision of a single individual facing exogenously given prices for human capital. Throughout, we assume that there are perfect capital markets. The separation theorem referred to in the title of this section will show that, with perfect capital markets, schooling decisions will maximize the net present discounted value of the individual. More specifically, consider an individual with an instantaneous utility function $u(c)$ that satisfies the standard neoclassical assumptions. In particular, it is strictly increasing and strictly concave. Suppose that the individual has a planning horizon of T (where $T = \infty$ is allowed), discounts the future at the rate $\rho > 0$ and faces a constant flow rate of death equal to $\nu \geq 0$. Standard arguments imply that the objective function of this individual at time $t = 0$ is

$$\max \int_0^T \exp(-(\rho + \nu) t) u(c(t)) dt. \tag{1.1}$$

Suppose that this individual is born with some human capital $h(0) \geq 0$. Suppose also that his human capital evolves over time according to the differential equation

$$\dot{h}(t) = G(t, h(t), s(t)), \tag{1.2}$$

where $s(t) \in [0,1]$ is the fraction of time that the individual spends for investments in schooling, and $G : \mathbb{R}_+^2 \times [0,1] \to \mathbb{R}_+$ determines how human capital evolves as a function of time, the individual's stock of human capital and schooling decisions. In addition, we can impose a further restriction on schooling decisions, for example,

$$s(t) \in \mathcal{S}(t), \tag{1.3}$$

where $\mathcal{S}(t) \subset [0,1]$ and may be useful to model constraints of the form $s(t) \in \{0,1\}$, which would correspond to the restriction that schooling must be full-time (or other such restrictions on human capital investments).

The individual is assumed to face an exogenous sequence of wage per unit of human capital given by $[w(t)]_{t=0}^T$, so that his labor earnings at time t are

$$W(t) = w(t)[1 - s(t)][h(t) + \omega(t)],$$

where $1 - s(t)$ is the fraction of time spent supplying labor to the market and $\omega(t)$ is non-human capital labor that the individual may be supplying to the market at time t. The sequence of non-human capital labor that the individual can supply to the market, $[\omega(t)]_{t=0}^T$, is exogenous. This formulation assumes that the only margin of choice is between market work and schooling (i.e., there is no leisure).

Finally, let us assume that the individual faces a constant (flow) interest rate equal to r on his savings. Using the equation for labor earnings, the lifetime budget constraint of the individual can be written as

$$\int_0^T \exp(-rt)\, c(t)\, dt \leq \int_0^T \exp(-rt)\, w(t)[1 - s(t)][h(t) + \omega(t)]\, dt. \tag{1.4}$$

The Separation Theorem, which is the subject of this section, can be stated as follows:

THEOREM 1.1. ***(Separation Theorem)*** *Suppose that the instantaneous utility function $u(\cdot)$ is strictly increasing. Then the sequence $\left[\hat{c}(t), \hat{s}(t), \hat{h}(t)\right]_{t=0}^T$ is a solution to the maximization of (1.1) subject to (1.2), (1.3) and (1.4) if and only if $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^T$ maximizes*

$$\int_0^T \exp(-rt)\, w(t)[1 - s(t)][h(t) + \omega(t)]\, dt \tag{1.5}$$

subject to (1.2) and (1.3), and $[\hat{c}(t)]_{t=0}^{T}$ *maximizes (1.1) subject to (1.4) given* $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$. *That is, human capital accumulation and supply decisions can be* separated *from consumption decisions.*

PROOF. To prove the "only if" part, suppose that $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ does not maximize (1.5), but there exists $\hat{c}(t)$ such that $\left[\hat{c}(t), \hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ is a solution to (1.1). Let the value of (1.5) generated by $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ be denoted Y. Since $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ does not maximize (1.5), there exists $[s(t), h(t)]_{t=0}^{T}$ reaching a value of (1.5), $Y' > Y$. Consider the sequence $[c(t), s(t), h(t)]_{t=0}^{T}$, where $c(t) = \hat{c}(t) + \varepsilon$. By the hypothesis that $\left[\hat{c}(t), \hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ is a solution to (1.1), the budget constraint (1.4) implies

$$\int_0^T \exp(-rt)\,\hat{c}(t)\,dt \leq Y.$$

Let $\varepsilon > 0$ and consider $c(t) = \hat{c}(t) + \varepsilon$ for all t. We have that

$$\begin{aligned} \int_0^T \exp(-rt)\,c(t)\,dt &= \int_0^T \exp(-rt)\,\hat{c}(t)\,dt + \frac{[1-\exp(-rT)]}{r}\varepsilon. \\ &\leq Y + \frac{[1-\exp(-rT)]}{r}\varepsilon. \end{aligned}$$

Since $Y' > Y$, for ε sufficiently small, $\int_0^T \exp(-rt)\,c(t)\,dt \leq Y'$ and thus $[c(t), s(t), h(t)]_{t=0}^{T}$ is feasible. Since $u(\cdot)$ is strictly increasing, $[c(t), s(t), h(t)]_{t=0}^{T}$ is strictly preferred to $\left[\hat{c}(t), \hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$, leading to a contradiction and proving the "only if" part.

The proof of the "if" part is similar. Suppose that $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ maximizes (1.5). Let the maximum value be denoted by Y. Consider the maximization of (1.1) subject to the constraint that $\int_0^T \exp(-rt)\,c(t)\,dt \leq Y$. Let $[\hat{c}(t)]_{t=0}^{T}$ be a solution. This implies that if $[c'(t)]_{t=0}^{T}$ is a sequence that is strictly preferred to $[\hat{c}(t)]_{t=0}^{T}$, then $\int_0^T \exp(-rt)\,c'(t)\,dt > Y$. This implies that $\left[\hat{c}(t), \hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ must be a solution to the original problem, because any other $[s(t), h(t)]_{t=0}^{T}$ leads to a value of (1.5) $Y' \leq Y$, and if $[c'(t)]_{t=0}^{T}$ is strictly preferred to $[\hat{c}(t)]_{t=0}^{T}$, then $\int_0^T \exp(-rt)\,c'(t)\,dt > Y \geq Y'$ for any Y' associated with any feasible $[s(t), h(t)]_{t=0}^{T}$. □

The intuition for this theorem is straightforward: in the presence of perfect capital markets, the best human capital accumulation decisions are those that maximize the lifetime budget set of the individual. It can be shown that this theorem does not hold when there are imperfect capital markets. Moreover, this theorem also fails to hold when leisure is an argument of the utility function of the individual. Nevertheless, it is a very useful benchmarkas a starting point of our analysis.

5. Schooling Investments and Returns to Education

We now turn to the simplest model of schooling decisions in partial equilibrium, which will illustrate the main tradeoffs in human capital investments. The model presented here is a version of Mincer's (1974) seminal contribution. This model also enables a simple mapping from the theory of human capital investments to the large empirical literature on returns to schooling.

Let us first assume that $T = \infty$, which will simplify the expressions. The flow rate of death, ν, is positive, so that individuals have finite expected lives. Suppose that (1.2) and (1.3) are such that the individual has to spend an interval S with $s(t) = 1$—i.e., in full-time schooling, and $s(t) = 0$ thereafter. At the end of the schooling interval, the individual will have a schooling level of

$$h(S) = \eta(S),$$

where $\eta(\cdot)$ is an increasing, continuously differentiable and concave function. For $t \in [S, \infty)$, human capital accumulates over time (as the individual works) according to the differential equation

$$\dot{h}(t) = g_h h(t), \tag{1.6}$$

for some $g_h \geq 0$. Suppose also that wages grow exponentially,

$$\dot{w}(t) = g_w w(t), \tag{1.7}$$

with boundary condition $w(0) > 0$.

Suppose that

$$g_w + g_h < r + \nu,$$

so that the net present discounted value of the individual is finite. Now using Theorem 1.1, the optimal schooling decision must be a solution to the following maximization problem

$$\max_{S} \int_{S}^{\infty} \exp\left(-\left(r+\nu\right)t\right) w\left(t\right) h\left(t\right) dt. \tag{1.8}$$

Now using (1.6) and (1.7), this is equivalent to:

$$\max_{S} \frac{\eta\left(S\right) w\left(0\right) \exp\left(-\left(r+\nu-g_w\right)S\right)}{r+\nu-g_h-g_w}. \tag{1.9}$$

Since $\eta\left(S\right)$ is concave, the objective function in (1.9) is strictly concave. Therefore, the unique solution to this problem is characterized by the first-order condition

$$\frac{\eta'\left(S^*\right)}{\eta\left(S^*\right)} = r+\nu-g_w. \tag{1.10}$$

Equation (1.10) shows that higher interest rates and higher values of ν (corresponding to shorter planning horizons) reduce human capital investments, while higher values of g_w increase the value of human capital and thus encourage further investments.

Integrating both sides of this equation with respect to S, we obtain

$$\ln \eta\left(S^*\right) = \text{constant} + \left(r+\nu-g_w\right)S^*. \tag{1.11}$$

Now note that the wage earnings of the worker of age $\tau \geq S^*$ in the labor market at time t will be given by

$$W\left(S,t\right) = \exp\left(g_w t\right) \exp\left(g_h\left(t-S\right)\right) \eta\left(S\right).$$

Taking logs and using equation (1.11) implies that the earnings of the worker will be given by

$$\ln W\left(S^*,t\right) = \text{ constant} + \left(r+\nu-g_w\right)S^* + g_w t + g_h\left(t-S^*\right),$$

where $t-S$ can be thought of as worker experience (time after schooling). If we make a cross-sectional comparison across workers, the time trend term $g_w t$, will also go into the constant, so that we obtain the canonical Mincer equation where, in the cross section, log wage earnings are proportional to schooling and experience.

Written differently, we have the following cross-sectional equation

$$\ln W_j = \text{ constant} + \gamma_s S_j + \gamma_e \text{experience}, \tag{1.12}$$

where j refers to individual j. Note however that we have not introduced any source of heterogeneity that can generate different levels of schooling across individuals. Nevertheless, equation (1.12) is important, since it is the typical empirical model for the relationship between wages and schooling estimated in labor economics.

The economic insight provided by this equation is quite important; it suggests that the functional form of the Mincerian wage equation is not just a mere coincidence, but has economic content: the opportunity cost of one more year of schooling is foregone earnings. This implies that the benefit has to be commensurate with these foregone earnings, thus should lead to a proportional increase in earnings in the future. In particular, this proportional increase should be at the rate $(r + \nu - g_w)$.

Empirical work using equations of the form (1.12) leads to estimates for γ in the range of 0.06 to 0.10. Equation (1.12) suggests that these returns to schooling are not unreasonable. For example, we can think of the annual interest rate r as approximately 0.10, ν as corresponding to 0.02 that gives an expected life of 50 years, and g_w corresponding to the rate of wage growth holding the human capital level of the individual constant, which should be approximately about 2%. Thus we should expect an estimate of γ around 0.10, which is consistent with the upper range of the empirical estimates.

6. A Simple Two-Period Model of Schooling Investments and Some Evidence

Let us now step back and illustrate these ideas using a two-period model and then use this model to look at some further evidence. In period 1 an individual (parent) works, consumes c, saves s, decides whether to send their offspring to school, $e = 0$ or 1, and then dies at the end of the period. Utility of household i is given as:

$$\ln c_i + \ln \hat{c}_i \tag{1.13}$$

where $\hat{c}$ is the consumption of the offspring. There is heterogeneity among children, so the cost of education, θ_i varies with i. In the second period skilled individuals (those with education) receive a wage w_s and an unskilled worker receives w_u.

First, consider the case in which there are no credit market problems, so parents can borrow on behalf of their children, and when they do so, they pay the same interest rate, r, as the rate they would obtain by saving. Then, the decision problem of the parent with income y_i is to maximize (1.13) with respect to e_i, c_i and $\hat{c}_i$, subject to the budget constraint:

$$c_i + \frac{\hat{c}_i}{1+r} \leq \frac{w_u}{1+r} + e_i \frac{w_s - w_u}{1+r} + y_i - e_i \theta_i$$

Note that e_i does not appear in the objective function, so the education decision will be made simply to maximize the budget set of the consumer. This is the essence of the Separation Theorem, Theorem 1.1 above. In particular, here parents will choose to educate their offspring only if

$$\theta_i \leq \frac{w_s - w_u}{1+r} \tag{1.14}$$

One important feature of this decision rule is that a greater skill premium as captured by $w_s - w_u$ will encourage schooling, while the higher interest rate, r, will discourage schooling (since schooling is a form of investment with upfront costs and delayed benefits).

In practice, this solution may be difficult to achieve for a variety of reasons. First, there is the usual list of informational/contractual problems, creating credit constraints or transaction costs that introduce a wedge between borrowing and lending rates (or even make borrowing impossible for some groups). Second, in many cases, it is the parents who make part of the investment decisions for their children, so the above solution involves parents borrowing to finance both the education expenses and also part of their own current consumption. These loans are then supposed to be paid back by their children. With the above setup, this arrangement works since parents are fully altruistic. However, if there are non-altruistic parents, this will create obvious problems.

Therefore, in many situations credit problems might be important. Now imagine the same setup, but also assume that parents cannot have negative savings, which is a simple and severe form of credit market problems. This modifies the constraint set as follows

$$\begin{aligned} c_i &\leq y_i - e_i\theta_i - s_i \\ s_i &\geq 0 \\ \hat{c}_i &\leq w_u + e_i\left(w_s - w_u\right) + (1+r)\,s \end{aligned}$$

First note that for a parent with $y_i - e_i\theta_i > w_s$, the constraint of nonnegative savings is not binding, so the same solution as before will apply. Therefore, credit constraints will only affect parents who needed to borrow to finance their children's education.

To characterize the solution to this problem, let us look at the utilities from investinging and not investing in education of a parent. Also to simplify the discussion let us focus on parents who would not choose positive savings, that is, those parents with $(1+r)\,y_i \leq w_u$. The utilities from investing and not investing in education are given, respectively, by $U(e = 1 \mid y_i, \theta_i) = \ln(y_i - \theta_i) + \ln w_s$, and $U(e = 0 \mid y_i, \theta_i) = \ln y_i + \ln w_u$. Comparison of these two expressions implies that parents with

$$\theta_i \leq y_i \frac{w_s - w_u}{w_s}$$

will invest in education. It is then straightforward to verify that:

(1) This condition is more restrictive than (1.14) above, since $(1+r)\,y_i \leq w_u < w_s$.

(2) As income increases, there will be more investment in education, which contrasts with the non-credit-constrained case.

One interesting implication of the setup with credit constraints is that the skill premium, $w_s - w_u$, still has a positive effect on human capital investments. However, in more general models with credit constraints, the conclusions may be more nuanced. For example, if $w_s - w_u$ increases because the unskilled wage, w_u, falls, this may reduce the income level of many of the households that are marginal for the education decision, thus discourage investment in education.

7. Evidence on Human Capital Investments and Credit Constraints

This finding, that income only matters for education investments in the presence of credit constraints, motivates investigations of whether there are significant differences in the educational attainment of children from different parental backgrounds as a test of the importance of credit constraints on education decisions. In addition, the empirical relationship between family income and education is interesting in its own right.

A typical regression would be along the lines of

$$\text{schooling=controls} + \alpha \cdot \log \text{parental income}$$

which leads to positive estimates of α, consistent with credit constraints. The problem is that there are at least two alternative explanations for why we may be estimating a positive α:

(1) Children's education may also be a consumption good, so rich parents will "consume" more of this good as well as other goods. If this is the case, the positive relationship between family income and education is not evidence in favor of credit constraints, since the "separation theorem" does not apply when the decision is not a pure investment (enters directly in the utility function). Nevertheless, the implications for labor economics are quite similar: richer parents will invest more in their children's education.

(2) The second issue is more problematic. The distribution of costs and benefits of education differ across families, and are likely to be correlated with income. That is, the parameter θ_i in terms of the model above will be

correlated with y_i, so a regression of schooling on income will, at least in part, capture the direct effect of different costs and benefits of education.

One line of attack to deal with this problem has been to include other characteristics that could proxy for the costs and benefits of education, or attitudes toward education. The interesting finding here is that when parents' education is also included in the regression, the role of income is substantially reduced.

Does this mean that credit market problems are not important for education? Does it mean that parents' income does not have a direct affect on education? Not necessarily. In particular, there are two reasons for why such an interpretation may not be warranted.

(1) First, parents' income may affect the quality rather than the quantity of education. This may be particularly important in the U.S. context where the choice of the neighborhood in which the family lives appears to have a major effect on the quality of schooling. This implies that in the United States high income parents may be "buying" more human capital for their children, not by sending them to school for longer, but by providing them with better schooling.

(2) Parental income is often measured with error, and has a significant transitory component, so parental education may be a much better proxy for permanent income than income observations in these data sets. Therefore, even when income matters for education, all its effect my load on the parental education variable.

Neither problem is easy to deal with, but there are possible avenues. First, we could look at the incomes of children rather than their schooling as the outcome variable. To the extent that income reflects skills (broadly defined), it will incorporate unobserved dimensions of human capital, including school quality. This takes us to the literature on intergenerational mobility. The typical regression here is

$$\text{log child income} = \text{controls} + \alpha \cdot \text{log parental income} \quad (1.15)$$

Regressions of this sort were first investigated by Becker and Tomes. They found relatively small coefficients, typically in the neighborhood of 0.3 (while others, for example Behrman and Taubman estimated coefficients as low as 0.2). This means that if your parents are twice as rich as my parents, you will typically have about 30 to 40 percent higher income than me. With this degree of intergenerational dependence, differences in initial conditions will soon disappear. In fact, your children will be typically about 10 percent (α^2 percent) richer than my children. So this finding implies that we are living in a relatively "egalitarian" society.

To see this more clearly, consider the following simple model:

$$\ln y_t = \mu + \alpha \ln y_{t-1} + \varepsilon_t$$

where y_t is the income of t-th generation, and ε_t is serially independent disturbance term with variance σ_ε^2. Then the long-term variance of log income is:

$$\sigma_y^2 = \frac{\sigma_\varepsilon^2}{1-\alpha^2} \tag{1.16}$$

Using the estimate of 0.3 for α, equation (1.16) implies that the long-term variance of log income will be approximately 10 percent higher than σ_ε^2, so the long-run income distribution will basically reflect transitory shocks to dynasties' incomes and skills, and not inherited differences.

Returning to the interpretation of α in equation (1.15), also note that a degree of persistence in the neighborhood of 0.3 is not very different from what we might expect to result simply from the inheritance of IQ between parents and children, or from the children's adoption of cultural values favoring education from their parents. As a result, these estimates suggest that there is a relatively small effect of parents income on children's human capital.

This work has been criticized, however, because there are certain simple biases, stacking the cards against finding large estimates of the coefficient α. First, measurement error will bias the coefficient α towards zero. Second, in typical panel data sets, we observe children at an early stage of their life cycles, where differences in earnings may be less than at later stages, again biasing α downward. Third, income mobility may be very nonlinear, with a lot of mobility among middle income families, but very little at the tails. Work by Solon and Zimmerman has dealt with the first two problems. They find that controlling for these issues increases the degree of persistence substantially to about 0.45 or even 0.55. The next figure shows Solon's baseline estimates.

Table 4—OLS and IV Estimates of ρ for Various Single-Year Income Measures in 1967

Income measure	OLS	IV	Sample si
Log earnings	0.386 (0.079)	0.526 (0.135)	322
Log wage	0.294 (0.052)	0.449 (0.095)	316
Log family income	0.483 (0.069)	0.530 (0.123)	313
Log (family income/poverty line)	0.476 (0.060)	0.563 (0.103)	313

Note: Standard-error estimates are in parentheses.

Figure 1.1

A paper by Cooper, Durlauf and Johnson, in turn, finds that there is much more persistence at the top and the bottom of income distribution than at the middle.

That the difference between 0.3 and 0.55 is in fact substantial can be seen by looking at the implications of using $\alpha = 0.55$ in (1.16). Now the long-run income distribution will be substantially more disperse than the transitory shocks. More specifically, we will have $\sigma_y^2 \approx 1.45 \cdot \sigma_\varepsilon^2$.

To deal with the second empirical issue, one needs a source of exogenous variation in incomes to implement an IV strategy. There are no perfect candidates, but some imperfect ones exist. One possibility, pursued in Acemoglu and Pischke (2001), is to exploit changes in the income distribution that have taken place over the past 30 years to get a source of exogenous variation in household income. The basic idea is that the rank of a family in the income distribution is a good proxy for parental human capital, and conditional on that rank, the income gap has widened over the past 20 years. Moreover, this has happened differentially across states. One can exploit this source of variation by estimating regression of the form

$$s_{iqjt} = \delta_q + \delta_j + \delta_t + \beta_q \ln y_{iqjt} + \varepsilon_{iqjt}, \tag{1.17}$$

where q denotes income quartile, j denotes region, and t denotes time. s_{iqjt} is education of individual i in income quartile q region j time t. With no effect of income on education, β_q's should be zero. With credit constraints, we might expect lower quartiles to have positive β's. Acemoglu and Pischke report versions of this equation using data aggregated to income quartile, region and time cells. The estimates of β are typically positive and significant, as shown in the next two tables.

However, the evidence does not indicate that the β's are higher for lower income quartiles, which suggests that there may be more to the relationship between income and education than simple credit constraints. Potential determinants of the relationship between income and education have already been discussed extensively in the literature, but we still do not have a satisfactory understanding of why parental income may affect children's educational outcomes (and to what extent it does so).

8. The Ben-Porath Model

The baseline Ben-Porath model enriches the models we have seen so far by allowing human capital investments and non-trivial labor supply decisions throughout the lifetime of the individual. It also acts as a bridge to models of investment in human capital on-the-job, which we will discuss below.

Let $s(t) \in [0,1]$ for all $t \geq 0$. Together with the Mincer equation (1.12) above, the Ben-Porath model is the basis of much of labor economics. Here it is sufficient to consider a simple version of this model where the human capital accumulation equation, (1.2), takes the form

$$\dot{h}(t) = \phi(s(t)h(t)) - \delta_h h(t), \tag{1.18}$$

Table 4
Fixed effects regressions for the probability of attending college within two years of high school controlling for income quartile region by income quartile cells, 1972–1992[a]

	Ever attending any college				Ever attending four-year college			
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log mean family income	0.218 (0.101)	0.107 (0.044)	0.102 (0.044)	0.146 (0.107)	0.212 (0.065)	0.148 (0.041)	0.142 (0.040)	0.093 (0.108)
Return to college	1.336 (0.491)	—	− 0.887 (0.616)	—	0.817 (0.314)	—	− 0.994 (0.556)	—
Region effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income quartile effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Income quartile × Region effects	No	No	No	Yes	No	No	Yes	Yes
Income quartile × Year effects	No	No	No	Yes	No	No	Yes	Yes
Region × Year effects	No	No	No	Yes	No	No	No	Yes

[a]Data are cell level means for 4 Census regions, 4 years, and 4 quartiles for the income of the student's family. Number of cells is 64. Dependent variable is the fraction of students enrolled in any college or in a four-year college within two years of high school graduation calculated from the NLS-72, HSB Senior and Sophomore cohorts, and the NELS. Students left high school in 1972, 1980, 1982, and 1992. Return to college is the relative wage of those with exactly 4 years of college to those with a high school degree (for workers with 1–5 years of experience) calculated from the Census for 1970, 1980, and 1990.

Figure 1.2

where $\delta_h > 0$ captures "depreciation of human capital," for example because new machines and techniques are being introduced, eroding the existing human capital of the worker. The individual starts with an initial value of human capital $h(0) > 0$. The function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is strictly increasing, continuously differentiable and strictly concave. Furthermore, we simplify the analysis by assuming that this function satisfies the Inada-type conditions,

$$\lim_{x \to 0} \phi'(x) = \infty \text{ and } \lim_{x \to h(0)} \phi'(x) = 0.$$

The latter condition makes sure that we do not have to impose additional constraints to ensure $s(t) \in [0,1]$..

Let us also suppose that there is no non-human capital component of labor, so that $\omega(t) = 0$ for all t, that $T = \infty$, and that there is a flow rate of death $\nu > 0$. Finally, we assume that the wage per unit of human capital is constant at w and the interest rate is constant and equal to r. We also normalize $w = 1$ without loss of any generality.

Again using Theorem 1.1, human capital investments can be determined as a solution to the following problem

$$\max \int_0^\infty \exp\left(-\left(r+\nu\right)\right)\left(1-s\left(t\right)\right)h\left(t\right)dt$$

subject to (1.18).

This problem can then be solved by setting up the current-value Hamiltonian, which in this case takes the form

$$\mathcal{H}\left(h,s,\mu\right) = \left(1-s\left(t\right)\right)h\left(t\right) + \mu\left(t\right)\left(\phi\left(s\left(t\right)h\left(t\right)\right) - \delta_h h\left(t\right)\right),$$

where we used $\mathcal{H}$ to denote the Hamiltonian to avoid confusion with human capital. The necessary conditions for an optimal solution to this problem are

$$\begin{aligned}
\mathcal{H}_s\left(h,s,\mu\right) &= -h\left(t\right) + \mu\left(t\right)h\left(t\right)\phi'\left(s\left(t\right)h\left(t\right)\right) = 0 \\
\mathcal{H}_h\left(h,s,\mu\right) &= \left(1-s\left(t\right)\right) + \mu\left(t\right)\left(s\left(t\right)\phi'\left(s\left(t\right)h\left(t\right)\right) - \delta_h\right) \\
&= \left(r+\nu\right)\mu\left(t\right) - \dot{\mu}\left(t\right) \\
\lim_{t\to\infty} \exp\left(-\left(r+\nu\right)t\right)\mu\left(t\right)h\left(t\right) &= 0.
\end{aligned}$$

To solve for the optimal path of human capital investments, let us adopt the following transformation of variables:

$$x\left(t\right) \equiv s\left(t\right)h\left(t\right).$$

Instead of $s(t)$ (or $\mu(t)$) and $h(t)$, we will study the dynamics of the optimal path in $x(t)$ and $h(t)$.

The first necessary condition then implies that

$$1 = \mu\left(t\right)\phi'\left(x\left(t\right)\right), \tag{1.19}$$

while the second necessary condition can be expressed as

$$\frac{\dot{\mu}(t)}{\mu(t)} = r + \nu + \delta_h - s(t)\phi'(x(t)) - \frac{1-s(t)}{\mu(t)}.$$

Substituting for $\mu(t)$ from (1.19), and simplifying, we obtain

$$\frac{\dot{\mu}(t)}{\mu(t)} = r + \nu + \delta_h - \phi'(x(t)). \tag{1.20}$$

The steady-state (stationary) solution of this optimal control problem involves $\dot{\mu}(t) = 0$ and $\dot{h}(t) = 0$, and thus implies that

$$x^* = \phi'^{-1}(r + \nu + \delta_h), \tag{1.21}$$

where $\phi'^{-1}(\cdot)$ is the inverse function of $\phi'(\cdot)$ (which exists and is strictly decreasing since $\phi(\cdot)$ is strictly concave). This equation shows that $x^* \equiv s^*h^*$ will be higher when the interest rate is low, when the life expectancy of the individual is high, and when the rate of depreciation of human capital is low.

To determine s^* and h^* separately, we set $\dot{h}(t) = 0$ in the human capital accumulation equation (1.18), which gives

$$\begin{aligned} h^* &= \frac{\phi(x^*)}{\delta_h} \\ &= \frac{\phi\left(\phi'^{-1}(r+\nu+\delta_h)\right)}{\delta_h}. \end{aligned} \tag{1.22}$$

Since $\phi'^{-1}(\cdot)$ is strictly decreasing and $\phi(\cdot)$ is strictly increasing, this equation implies that the steady-state solution for the human capital stock is uniquely determined and is decreasing in r, ν and δ_h.

More interesting than the stationary (steady-state) solution to the optimization problem is the time path of human capital investments in this model. To derive this, differentiate (1.19) with respect to time to obtain

$$\frac{\dot{\mu}(t)}{\mu(t)} = \varepsilon_{\phi'}(x)\frac{\dot{x}(t)}{x(t)},$$

where

$$\varepsilon_{\phi'}(x) = -\frac{x\phi''(x)}{\phi'(x)} > 0$$

is the elasticity of the function $\phi'(\cdot)$ and is positive since $\phi'(\cdot)$ is strictly decreasing (thus $\phi''(\cdot) < 0$). Combining this equation with (1.20), we obtain

$$\frac{\dot{x}(t)}{x(t)} = \frac{1}{\varepsilon_{\phi'}(x(t))}\left(r + \nu + \delta_h - \phi'(x(t))\right). \tag{1.23}$$

Figure 1.4 plots (1.18) and (1.23) in the h-x space. The upward-sloping curve corresponds to the locus for $\dot{h}(t) = 0$, while (1.23) can only be zero at x^*, thus the locus for $\dot{x}(t) = 0$ corresponds to the horizontal line in the figure. The arrows of motion are also plotted in this phase diagram and make it clear that the steady-state solution (h^*, x^*) is globally saddle-path stable, with the stable arm coinciding with the horizontal line for $\dot{x}(t) = 0$. Starting with $h(0) \in (0, h^*)$, $s(0)$ jumps to the level necessary to ensure $s(0)h(0) = x^*$. From then on, $h(t)$ increases and $s(t)$ decreases so as to keep $s(t)h(t) = x^*$. Therefore, the pattern of human capital investments implied by the Ben-Porath model is one of high investment at the beginning of an individual's life followed by lower investments later on.

In our simplified version of the Ben-Porath model this all happens smoothly. In the original Ben-Porath model, which involves the use of other inputs in the production of human capital and finite horizons, the constraint for $s(t) \leq 1$ typically binds early on in the life of the individual, and the interval during which $s(t) = 1$ can be interpreted as full-time schooling. After full-time schooling, the individual starts working (i.e., $s(t) < 1$). But even on-the-job, the individual continues to accumulate human capital (i.e., $s(t) > 0$), which can be interpreted as spending time in training programs or allocating some of his time on the job to learning rather than production. Moreover, because the horizon is finite, if the Inada conditions were relaxed, the individual could prefer to stop investing in human capital at some point. As a result, the time path of human capital generated by the standard Ben-Porath model may be hump-shaped, with a possibly declining portion at the end. Instead, the path of human capital (and the earning potential of the individual) in the current model is always increasing as shown in Figure 1.5.

The importance of the Ben-Porath model is twofold. First, it emphasizes that schooling is not the only way in which individuals can invest in human capital and there is a continuity between schooling investments and other investments in human capital. Second, it suggests that in societies where schooling investments are high we may also expect higher levels of on-the-job investments in human capital. Thus there may be systematic mismeasurement of the amount or the quality human capital across societies.

This model also provides us with a useful way of thinking of the lifecycle of the individual, which starts with higher investments in schooling, and then there is a period of "full-time" work (where $s(t)$ is high), but this is still accompanied by investment in human capital and thus increasing earnings. The increase in earnings takes place at a slower rate as the individual ages. There is also some evidence that earnings may start falling at the very end of workers' careers, though this does not happen in the simplified version of the model presented here (how would you modify it to make sure that earnings may fall in equilibrium?).

The available evidence is consistent with the broad patterns suggested by the model. Nevertheless, this evidence comes from cross-sectional age-experience profiles, so it has to be interpreted with some caution (in particular, the decline at the very end of an individual's life cycle that is found in some studies may be due to "selection," as the higher-ability workers retire earlier).

Perhaps more worrisome for this interpretation is the fact that the increase in earnings may reflect not the accumulation of human capital due to investment, but either:

(1) simple age effects; individuals become more productive as they get older. Or
(2) simple experience effects: individuals become more productive as they get more experienced—this is independent of whether they choose to invest or not.

It is difficult to distinguish between the Ben-Porath model and the second explanation. But there is some evidence that could be useful to distinguish between age effects vs. experience effects (automatic or due to investment).

Josh Angrist's paper on Vietnam veterans basically shows that workers who served in the Vietnam War lost the experience premium associated with the years they served in the war. This is shown in the next figure.

Presuming that serving in the war has no productivity effects, this evidence suggests that much of the age-earnings profiles are due to experience not simply due to age. Nevertheless, this evidence is consistent both with direct experience effects on worker productivity, and also a Ben Porath type explanation where workers are purposefully investing in their human capital while working, and experience is proxying for these investments.

9. Selection and Wages—The One-Factor Model

Issues of selection bias arise often in the analysis of education, migration, labor supply, and sectoral choice decisions. This section illustrates the basic issues of selection using a single-index model, where each individual possesses a one-dimensional skill. Richer models, such as the famous Roy model of selection, incorporate multi-dimensional skills. While models with multi-dimensional skills make a range of additional predictions, the major implications of selection for interpreting wage differences across different groups can be derived using the single-index model.

Suppose that individuals are distinguished by an unobserved type, z, which is assumed to be distributed uniformly between 0 and 1. Individuals decide whether to obtain education, which costs c. The wage of an individual of type z when he has no education is

$$w_0(z) = z$$

and when he obtains education, it is

$$w_1(z) = \alpha_0 + \alpha_1 z, \tag{1.24}$$

where $\alpha_0 > 0$ and $\alpha_1 > 1$. α_0 is the main effect of education on earnings, which applies irrespective of ability, whereas α_1 interacts with ability. The assumption that $\alpha_1 > 1$ implies that education is complementary to ability, and will ensure that high-ability individuals are "positively selected" into education.

Individuals make their schooling choices to maximize income. It is straightforward to see that all individuals of type $z \geq z^*$ will obtain education, where

$$z^* \equiv \frac{c - \alpha_0}{\alpha_1 - 1},$$

which, to make the analysis interesting, we assume lies between 0 and 1. Figure 1.7 gives the wage distribution in this economy.

Now let us look at mean wages by education group. By standard arguments, these are

$$\bar{w}_0 = \frac{c - \alpha_0}{2(\alpha_1 - 1)}$$

$$\bar{w}_1 = \alpha_0 + \alpha_1 \frac{\alpha_1 - 1 + c - \alpha_0}{2(\alpha_1 - 1)}$$

It is clear that $\bar{w}_1 - \bar{w}_0 > \alpha_0$, so the wage gap between educated and uneducated groups is greater than the main effect of education in equation (1.24)—since $\alpha_1 - 1 > 0$. This reflects two components. First, the return to education is not α_0, but it is $\alpha_0 + \alpha_1 \cdot z$ for individual z. Therefore, for a group of mean ability $\bar{z}$, the return to education is

$$w_1(\bar{z}) - w_0(\bar{z}) = \alpha_0 + (\alpha_1 - 1)\bar{z},$$

which we can simply think of as the return to education evaluated at the mean ability of the group.

But there is one more component in $\bar{w}_1 - \bar{w}_0$, which results from the fact that the average ability of the two groups is not the same, and the earning differences resulting from this ability gap are being counted as part of the returns to education. In fact, since $\alpha_1 - 1 > 0$, high-ability individuals are selected into education increasing the wage differential. To see this, rewrite the observed wage differential as follows

$$\bar{w}_1 - \bar{w}_0 = \alpha_0 + (\alpha_1 - 1)\left[\frac{c - \alpha_0}{2(\alpha_1 - 1)}\right] + \frac{\alpha_1}{2}$$

Here, the first two terms give the return to education evaluated at the mean ability of the uneducated group. This would be the answer to the counter-factual question of how much the earnings of the uneducated group would increase if they were to obtain education. The third term is the additional effect that results from the fact that the two groups *do not* have the same *ability* level. It is therefore the selection effect. Alternatively, we could have written

$$\bar{w}_1 - \bar{w}_0 = \alpha_0 + (\alpha_1 - 1)\left[\frac{\alpha_1 - 1 + c - \alpha_0}{2(\alpha_1 - 1)}\right] + \frac{1}{2},$$

where now the first two terms give the return to education evaluated at the mean ability of the educated group, which is greater than the return to education evaluated at the mean ability level of the uneducated group. So the selection effect is somewhat smaller, but still positive.

This example illustrates how looking at observed averages, without taking selection into account, may give misleading results, and also provides a simple example of how to think of decisions in the presence of this type of heterogeneity.

It is also interesting to note that if $\alpha_1 < 1$, we would have negative selection into education, and observed returns to education would be less than the true returns. The case of $\alpha_1 < 1$ appears less plausible, but may arise if high ability individuals do not need to obtain education to perform certain tasks.

Table 5
Fixed effects regressions for the probability of attending college within two years of high school effects by income quartile region by income quartile cells, 1972–1992[a]

	Ever attending any college				Ever attending four-year college			
Independent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log mean family income Quartile 1	0.018 (0.143)	0.154 (0.056)	0.139 (0.064)	– 0.039 (0.187)	0.010 (0.085)	0.108 (0.052)	0.064 (0.053)	– 0.016 (0.190)
Log mean family income Quartile 2	0.229 (0.258)	0.189 (0.113)	0.167 (0.117)	0.201 (0.334)	0.151 (0.153)	0.128 (0.105)	0.087 (0.101)	– 0.205 (0.339)
Log mean family income Quartile 3	0.617 (0.273)	0.161 (0.116)	0.148 (0.129)	0.328 (0.283)	0.428 (0.162)	0.174 (0.107)	0.150 (0.112)	– 0.039 (0.287)
Log mean family income Quartile 4	0.405 (0.152)	0.012 (0.071)	– 0.005 (0.072)	0.231 (0.132)	0.392 (0.092)	0.212 (0.066)	0.183 (0.063)	0.147 (0.134)
Return to college Quartile 1	0.691 (1.052)	—	– 1.049 (0.759)	—	– 0.053 (0.623)	—	– 1.577 (0.659)	—
Return to college Quartile 2	1.144 (0.938)	—	– 1.032 (0.726)	—	0.599 (0.556)	—	– 1.121 (0.630)	—
Return to college Quartile 3	0.481 (1.050)	—	– 0.963 (0.722)	—	0.171 (0.622)	—	– 1.115 (0.627)	—
Return to college Quartile 4	1.367 (0.952)	—	– 0.438 (0.723)	—	1.304 (0.564)	—	– 0.226 (0.627)	—
Region effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Income quartile effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Income quartile × Region effects	No	No	No	Yes	No	Yes	Yes	Yes
Income quartile × Year effects	No	No	No	Yes	No	Yes	Yes	Yes
Region × Year effects	No	No	No	Yes	No	No	No	Yes

[a]Data are cell level means for 4 Census regions, 4 years, and 4 quartiles for the income of the student's family. Number of cells is 64. Dependent variable is the fraction of students enrolled in any college or in a four-year college within two years of high school graduation calculated from the NLS-72, HSB Senior and Sophomore cohorts, and the NELS. Students left high school in 1972, 1980, 1982, and 1992. Return to college is the relative wage of those with exactly 4 years of college to those with a high school degree (for workers with 1–5 years of experience) calculated from the Census for 1970, 1980, and 1990.

Figure 1.3

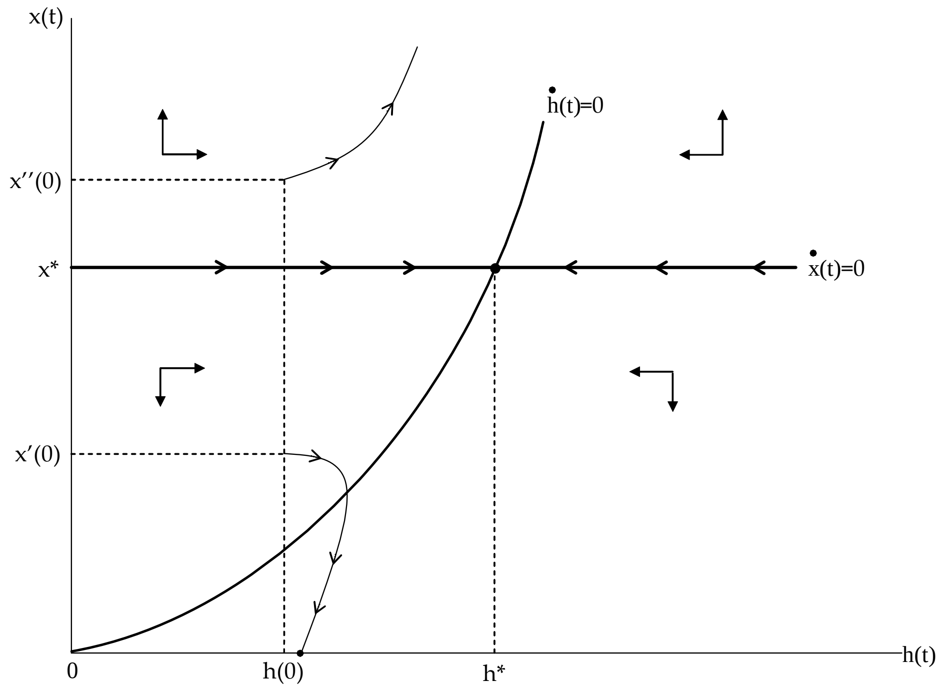


Figure 1.4. Steady state and equilibrium dynamics in the simplified Ben Porath model.

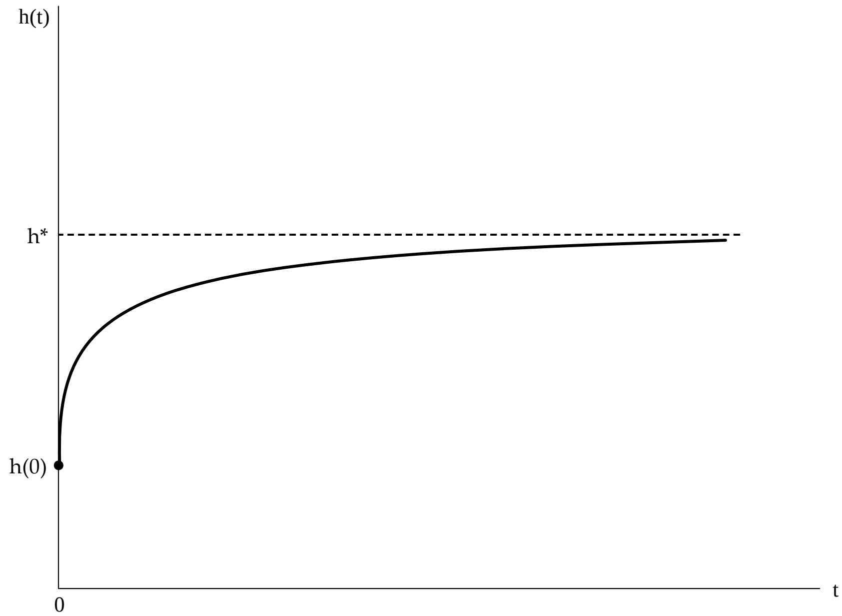
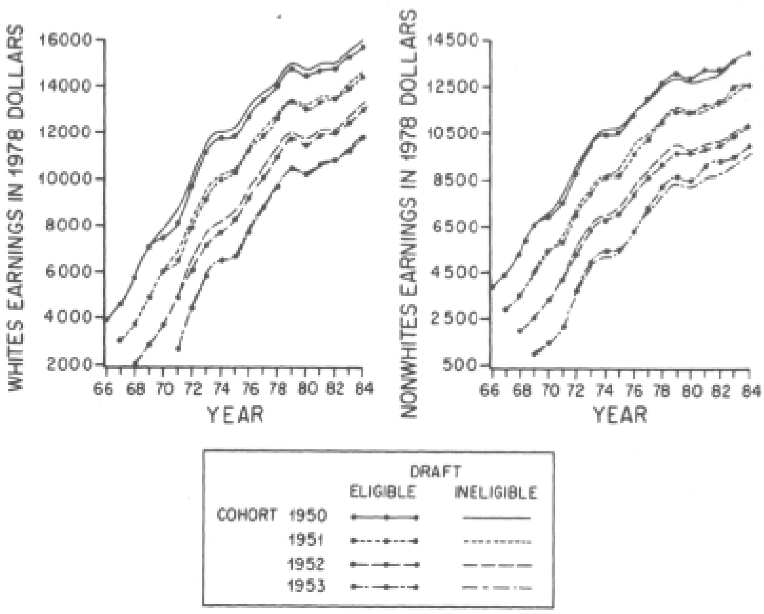


Figure 1.5. Time path of human capital investments in the simplified Ben Porath model.



Notes: The figure plots mean W-2 compensation in 1981–84 against probabilities of veteran status by cohort and groups of five consecutive lottery numbers for white men born 1950–53. Plotted points consist of the average residuals (over four years of earnings) from regressions on period and cohort effects. The slope of the least squares regression line drawn through the points is −2,384 with a standard error of 778, and is an estimate of α in the equation

$$\bar{y}_{ctj} = \beta_c + \delta_t + \hat{p}_{cj}\alpha + \bar{u}_{ctj}.$$

FIGURE 3. EARNINGS AND THE PROBABILITY OF VETERAN STATUS BY LOTTERY NUMBER

FIGURE 1.6

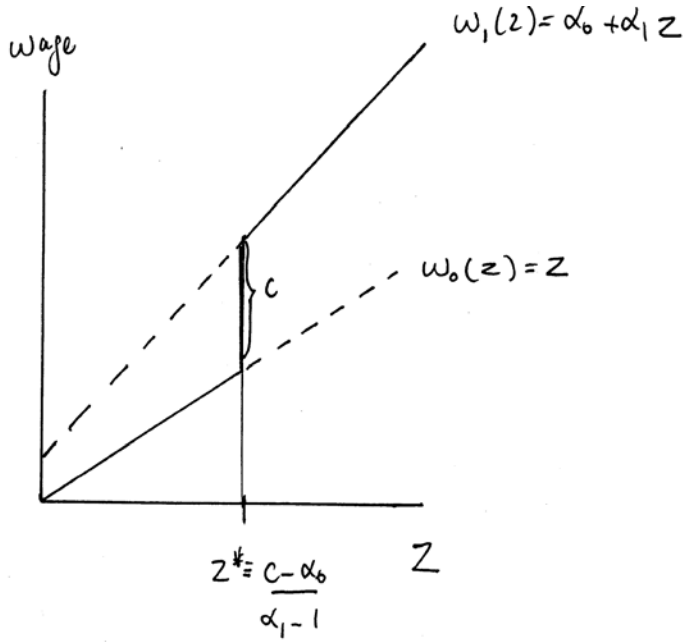


Figure 1.7. Selection in the One-Factor Model.