

Human Capital and the Wealth of Nations

Rodolfo E. Manuelli and Ananth Seshadri*

Department of Economics

University of Wisconsin-Madison

March, 2005

Abstract

No question has perhaps attracted as much attention in the economics literature as “Why are some countries richer than others?” In this paper, we revisit the development problem. Contrary to recent work, we find that only relatively small (of at most 27%) differences in Total Factor Productivity (TFP) are required to explain the large differences in output per worker. We estimate the (long-run) elasticity of output with respect to TFP to be around 8, and we find a substantial impact associated with (exogenous) changes in life expectancy and fertility.

The key difference between our model and recent work in this area is that we use theory to estimate the stocks of human capital, and we allow the *quality* of human capital to vary across countries. We find that the cross-country differences in average human capital stocks are large.

*We thank Larry Jones, John Kennan, Pete Klenow, Ellen Mcgrattan and Ed Prescott for comments. We thank the NSF for financial support.

1 Introduction

No question has perhaps attracted as much attention in the economics literature as “Why are some countries richer than others?” Much of the current work traces back to Solow’s classic work (1956). Solow’s seminal paper suggested that differences in the rates at which capital is accumulated could account for differences in output per capita. More recently, following the work of Lucas (1988), human capital disparities were given a central role in the analysis of growth and development. However, the best recent work on the topic reaches the opposite conclusion. Klenow and Rodriguez-Clare (1997), Hall and Jones (1999) and Parente and Prescott (2000) argue that most of the cross country differences in output per worker are not driven by differences in human capital (or physical capital); rather they are due to differences in a residual, total factor productivity (TFP).

In this paper we revisit the development problem. In line with the earlier view, we find that factor accumulation is more important than TFP to explain relative incomes. The key difference between our work and previous analyses is in the measurement of human capital. The standard approach — inspired by the work of Mincer (1974) — takes estimates of the rate of return to schooling as building blocks to *directly* measure a country’s stock of human capital. Implicitly, this method assumes that the marginal contribution to output of one additional year of schooling is equal to the rate of return. One problem with this procedure is that it is not well suited to handle cross-country differences in the *quality* of human capital. Following the pioneering work of Becker (1964) and Ben-Porath (1967), we model human capital acquisition as part of a standard utility maximization problem. Our set up is flexible enough so that individuals can choose the length of the schooling period —which we identify as a measure of the *quantity* of human capital— and the amount of human capital per year of schooling and post-schooling training, which we view as a measure of *quality*. We use evidence on schooling and age-earnings profile to determine the parameters of the human capital production function. We then compute stocks of human capital as

the output of this technology, evaluated at the (individually) optimal choice of inputs given the equilibrium prices. Thus, we use theory — disciplined by observations — to indirectly estimate the stocks of human capital in each country.

We calibrate the model to match some moments of the U.S. economy and, following the standard development accounting approach, we compute the levels of TFP that are required to explain the observed cross-country differences in output per worker. According to the model, only relatively modest (of at most 27%) differences in TFP across countries suffice to explain the (large) observed differences in output per worker. Thus, TFP does not explain a large share —in the conventional way that this is estimated— of the differences in output per worker. Our result is mostly driven by our estimates of the average stocks of human capital. We find that cross-country differences in average human capital per worker are much larger than suggested by recent estimates. Since the model matches actual years of education quite well, we conclude that it is differences in the *quality* of human capital account for our findings..

We go beyond the development accounting exercise and compute the impact on a country’s output per worker of changes in any of the exogenous variables. We consider two exercises. First, we estimate the impact on (long run) output of an exogenous increase in TFP. We find that the resulting elasticity is fairly large: a 1% increase in (relative) TFP results in a 8% (long run) increase in (relative) output per worker. This is mostly due to the response of human capital to the change in productivity. The second exercise is designed to evaluate the contribution of demographic characteristics to underdevelopment. In the model, countries differ in terms of life expectancy, retirement age and fertility. We conduct the following counterfactual experiment: imagine ‘endowing’ the average country in the lowest decile of the world income distribution with the demographic characteristics of the average country in the top decile. Then let individuals adjust their choices of physical and human capital. We find that this demographic change doubles the level of output in the poor country.¹

¹For an excellent review of the connection between demographics and growth see Galor (2005).

Even though we do not use estimates of a Mincer style regression to construct stocks of human capital, we show that the model generates estimated rates of return to schooling that are in the range of those observed in the data. Since international quality differentials in human capital play such an important role we use the model to predict the path of earnings of an immigrant to the U.S. as a function of country of origin. The model is fairly successful at reproducing the time path of income for immigrants given their level of schooling.

The baseline economy relies on differences in TFP and demographics to account for the variability in output per capita. This is an extreme view. It is well documented (see, for example, Chari, Kehoe and McGrattan (1997) and Hsieh and Klenow (2003)), that there are significant cross country differences in the price of capital. When we allow the price of capital to vary in the same way as in the data, our model predicts that to account for differences in output per worker *no differences* in TFP are needed.

Broadly speaking, our approach emphasizes a new dimension (quality of education) that helps explain differences across countries. In this sense it is related to a variety of papers. Chari, Kehoe and McGrattan (1997) rely on organization capital, Parente, Rogerson and Wright (2000) emphasize the role of home production, and Jeong and Townsend (2004), rely on market imperfections.

In section 2 we present the theoretical model. In section 3 we describe the calibration, and in section 4 we present the results. In section 5, we discuss the results and in section 6, we use the model to compute the implications for the return on schooling and for the relative income of immigrants. Section 7 presents some concluding comments.

2 The Model

In this section we describe the basic model. We present an economic environment with imperfect altruism and we show that, under some conditions, the solution to the utility

maximization problem is identical to the solution of an income maximization problem. We then compute the aggregate variables in this economy using the exogenously specified demographic structure.

2.1 The Individual Household Problem

The representative household is formed at age I (age of independence). At age B , e^f children are born. The period of ‘early childhood’ (defined by the assumption that children are not productive during this period) corresponds to the (parent) age B to $B + 6$. The children remain with the household (and as such make no decisions of their own) until they become independent at (parent) age $B + I$. The parent retires at age R , and dies at age T .

Each parent chooses his own consumption over his lifetime, $c(a)$, and each of his children consumption during the years that they are part of his household, $c_k(a)$ for $a \in [B, B + I)$ to maximize his utility (which depends on his own consumption) and the discounted value of his children utility. The discount factor reflects the possibility of imperfect altruism. In addition to consumption, the parent chooses the amount of market goods to be used in the production of new human capital, $x(a)$, and the fraction of the time allocated to the formation of human capital, $n(a)$ (and, consequently, what fraction of the available time to allocate to working in the market, $1 - n(a)$) for him and each of his children while they are still attached to his household. The parent also decides to make investments in early childhood, which we denote by x_E (e.g. medical care, nutrition and development of learning skills), that determine the level of each child’s human capital at age 6, $h_k(B + 6)$, or h_B for short. Finally, the parent chooses how much to bequeath to each children at the time they leave the household, b_k . We assume that each parent has unrestricted access to capital markets, but that he cannot commit his children to honor his debts. Thus, we restrict b_k to be non-negative.

The maximization problem solved by a representative household is

$$V^P(h, b) = \max \int_I^T e^{-\rho(a-I)} u(c(a)) da + e^{-\alpha_0 + \alpha_1 f} e^{-\rho B} V^k(h_k(B+I), b_k) \quad (1)$$

subject to

$$\begin{aligned} & \int_I^T e^{-r(a-I)} c(a) da + e^f \int_B^{B+I} e^{-r(a-I)} c_k(a) da + \int_I^R e^{-r(a-I)} x(a) da + \\ & e^f \int_B^{B+I} e^{-r(a-I)} x_k(a) da + e^f e^{-rB} b_k + e^f e^{-r(B+6)} x_E \\ \leq & \int_I^R e^{-r(a-I)} w h(a) (1 - n(a)) da + e^f \int_B^{B+I} e^{-r(a-I)} w h_k(a) (1 - n_k(a)) da + b, \end{aligned} \quad (2)$$

and

$$\dot{h}(a) = z_h [n(a)h(a)]^{\gamma_1} x(a)^{\gamma_2} - \delta_h h(a), \quad a \in [I, R] \quad (3)$$

$$\dot{h}_k(a) = z_h [n_k(a)h_k(a)]^{\gamma_1} x_k(a)^{\gamma_2} - \delta_h h_k(a), \quad a \in [B+6, B+I], \quad (4)$$

$$h_k(B+6) = h_B x_E^v, \quad (5)$$

$$h(I) \quad \text{given}, \quad 0 < \gamma_i < 1, \quad \gamma = \gamma_1 + \gamma_2 < 1,$$

where $V^k(h_k(B+I), b_k)$ is the utility of a child at the time he becomes independent. The term $e^{-\alpha_0 + \alpha_1 f}$ captures the degree of altruism. If $\alpha_0 = 0$, and $\alpha_1 = 1$, this is a standard infinitively-lived agent model. Positive values of α_0 , and values of α_1 less than 1 capture the degree of imperfect altruism. The budget constraint is standard, and its present value formulation captures the assumption of unrestricted borrowing and lending within one's lifetime.

The technology to produce human capital of each child at the beginning of the potential school years, $h_k(B+6)$ or h_E is given by (5). Our formulation captures the idea that nutrition and health care are important determinants of early levels of human capital, and those inputs are, basically, market goods.² Equations (3) and (4) correspond to the standard human capital accumulation model initially developed

²It is clear that parents time is also important. However, given that we assume that the number of children is exogenous, it is not necessary to explicitly model the time cost of raising children.

by Ben-Porath (1967). There are two important features of our formulation. First, we assume that the technology for human capital accumulation is the same during the schooling and the training periods. We resisted the temptation to use a more complicated parameterization so as to force the model to use the same factors to account for the length of the schooling period and the shape of the age-earnings profile. Second, we assume that the market inputs used in the production of human capital — $x(a)$ and $x_k(a)$ — are privately purchased. In the case of the post-schooling period, this is not controversial. However, this is less so for the schooling period. Here, we take the ‘purely private’ approach as a first pass. In an extension of the basic model we explore the role of public education³, even though all that is needed for our assumption is that, at the margin, individuals pay for the last unit of market goods allocated to the formation of human capital during the schooling period.

In this economy, the critical discount factor captures how parents value the utility of their children at the time they (the children) leave the home. This factor is given by $e^{-[\alpha_0+(1-\alpha_1)f]}e^{-\rho B}$. In a steady state in which the intergenerational bequest constraint is not binding, it must be the case that the discount factor equal the interest rate. This requires

$$e^{-rB} = e^{-[\alpha_0+(1-\alpha_1)f]}e^{-\rho B},$$

or,

$$r = \rho + [\alpha_0 + (1 - \alpha_1)f]/B. \tag{6}$$

Proposition 1 *Assume that $r = \rho + [\alpha_0 + (1 - \alpha_1)f]/B$, then the solution to the optimal human capital accumulation corresponding to the maximization of (1) subject to (2)-(5) is identical to the solution of the following income maximization problem*

$$\max \int_6^R e^{-r(a-6)} [wh(a)(1 - n(a)) - x(a)] da - x_E \tag{7}$$

³An alternative explanation is that Tiebout like arguments effectively imply that public expenditures on education play the same role as private expenditures. The truth is probably somewhere in between.

subject to

$$\dot{h}(a) = z_h[n(a)h(a)]^{\gamma_1}x(a)^{\gamma_2} - \delta_h h(a), \quad a \in [6, R), \quad (8)$$

and

$$h(6) = h_E = h_B x_E^v \quad (9)$$

with h_B given. In this notation a indicates an individual's age.

Proof. : See the Appendix ■

An intuitive (and heuristic) argument that shows the correspondence between the utility maximization and the income maximization problem is as follows: Suppose that parents (who make human capital accumulation decisions for their children until age I) do not choose to maximize the present value of income of their children (only part of which they keep). In this case, and since $b_k > 0$, the parent could increase the utility of each child by adopting the income maximizing human capital policy and adjusting the transfer to finance this change. It follows that the cost to the parent is the same and the child is made better off. Since the parent appropriates the income generated by child labor, one might wonder if it is not in the best interest of the parent to take the child out of school early and send him to work. However, this cannot be optimal as the parent can choose the optimal —from the point of view of the child— human capital policy and change the bequest as necessary. Since the parent's income is unchanged and the child is better off, this results in an increase in the utility of the parent.

As our informal discussion suggests, the key ingredient is that the intergenerational no borrowing constraint is not binding. Since this option is, effectively, another technology that the parent can use to transfer wealth to his children, standard arguments show that there will be no distortions. In related work we show that, when the non-negative bequest is binding, this is no longer true. In that case, which requires $r > \rho + [\alpha_0 + (1 - \alpha_1)f]/B$, the equilibrium human capital choices no longer maximize the present value of income. (See Manuelli and Seshadri (2004).)

In the unconstrained case, it is possible to fully characterize the solution to the income maximization problem. The main features of the solution are summarized in

Proposition 2 *There exists a unique solution to the income maximization problem.*

The number of years of schooling, s , satisfies

1.

$$F(s) = \frac{h_B^{1-\gamma}}{z_h^{1-v} w \gamma_2^{-v(1-\gamma_1)}} \left(\frac{v}{r + \delta_h} \right)^{(1-\gamma)v} \left(\frac{\gamma_2^{\gamma_2} \gamma_1^{(1-\gamma_2)}}{r + \delta_h} \right)^{-(1-v)}, \quad (10)$$

where

$$F(s) \equiv m(6+s)^{1-v(2-\gamma)} e^{(1-\gamma)(\delta_h+rv)s} \left[1 - \frac{r + \delta_h}{\gamma_1} \frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r + \delta_h(1-\gamma_1)} \frac{1 - e^{-\frac{\gamma_2 r + \delta_h(1-\gamma_1)}{(1-\gamma_2)} s}}{m(s+6)} \right]^{\frac{(1-\gamma)(1-v(1-\gamma_1))}{1-\gamma_1}},$$

and

$$m(a) = 1 - e^{-(r+\delta_h)(R-a)},$$

provided that

$$m(6)^{1-v(2-\gamma)} > \frac{h_B^{1-\gamma}}{z_h^{1-v} w \gamma_2^{-v(1-\gamma_1)}} \left(\frac{v}{r + \delta_h} \right)^{(1-\gamma)v} \left(\frac{\gamma_2^{\gamma_2} \gamma_1^{(1-\gamma_2)}}{r + \delta_h} \right)^{-(1-v)}.$$

Otherwise the privately optimal level of schooling is 0.

2. The level of human capital at the age at which the individual finishes his formal schooling is given by

$$h(s+6) = \left[\frac{\gamma_2^{\gamma_2} \gamma_1^{\gamma_1} z_h w \gamma_2}{(r + \delta_h)^\gamma} \right]^{\frac{1}{1-\gamma}} \frac{\gamma_1}{r + \delta_h} m(6+s)^{\frac{1}{1-\gamma}} \quad (11)$$

Proof. : See the Appendix ■

There are several interesting features of the solution.

1. **The Technology to Produce Human Capital and the Impact of Macroeconomic Conditions.** The proposition illustrates the role played by economic

forces in inducing a feedback from aggregate variables to the equilibrium choice of schooling. To be precise, had we assumed that market goods do not appear in the production of human capital (i.e. $\gamma_2 = v = 0$), the model implies that changes in wage rates have *no impact* on schooling decisions. (See equation 10) Thus, the standard formulation that assumes that market goods are not used in the production of human capital has to rely on differences in interest rates or the working horizon as the only source of *equilibrium* differences in schooling across countries.⁴ Our formulation is flexible enough so that the impact of wages on equilibrium schooling is ambiguous. The reason is simple: Pre-schooling investments in human capital and schooling are substitutes; hence, depending on the productivity of market goods in the production of early childhood human capital relative to schooling human capital, increases in wages may increase or decrease schooling. To be precise, if v is sufficiently high (and $\gamma_2 - v(1 - \gamma_1) < 0$), increases in market wages make parents more willing to invest in early childhood human capital. Thus at age 6 the increase in human capital (relative to a low v economy) is sufficiently large that investments in schooling are less profitable. In this case, the equilibrium level of s decreases. Even though theoretically possible, this requires extreme values of v . In our parameterization $\gamma_2 - v(1 - \gamma_1) > 0$, and we obtain the more ‘normal’ response: high wage (and high TFP) economies are also economies with high levels of schooling. This is an important source of differences in the equilibrium years of schooling that individuals in different countries decide to acquire.

2. Development and Early Childhood Human Capital. Early childhood human capital accumulation captures an important difference between an average child entering school at a rich country vis a vis a poor country. This is

⁴It is clear from the formulation that cross-country differences in z_h —ability to learn— and h_B —the endowment of human capital— can also account for differences in s . Since we have no evidence of systematic differences across countries, we do not pursue this possibility in this paper.

particularly important as an additional unit of investment early on in life has an important effect throughout the life cycle of an individual. One could view early childhood as influencing either the stock of human capital at age 6—as we do here—or the z_h parameter in the production function of human capital. The two formulations have very different effects. An increase in h_B —the amount of human capital that an individual is born with—decreases the desired amount of schooling. The reason is simple: since the individual starts with more human capital, he can leave school earlier to generate income. Note that this happens even though one can view h_B as a factor that increases the productivity of time allocated to the production of human capital. On the other hand, increases in z_h increase the equilibrium years of schooling. (For details see Manuelli and Seshadri, 2004). In our formulation, the stock of human capital at age 6, h_E is endogenously determined. Since goods inputs are the only input into the production of childhood human capital, children in richer countries will start off with a higher level of human capital before entry into school when compared to children from poorer countries, other things equal. (See equation (42) in the Appendix.)

3. **Development and Schooling Quality.** In the context of this model there is a natural way to distinguish between quantity and quality of schooling. We specify that if two individuals choose the same value of s their levels of schooling are identical. However, the quality of schooling is measured by the differences in human capital, as given by $h(6 + s)$. To illustrate the implications of the model for the impact of development on quality, consider two countries with level of real wages given by $w' > w$, and no differences in interest rates. Now suppose that two individuals residing in these two countries choose the same level of schooling. This, of course, requires that these individuals differ along some other dimension. For the sake of simplicity, we assume that their initial levels of human capital, h_B , are adjusted to that the (endogenous) value of s is

the same. In this case, (11) implies that the individual in the country with the higher wage rate also has more capital. The elasticity of $h(s+6)$ with respect to w is $\gamma_2/(1-\gamma)$, which is fairly large in our preferred parameterization.⁵ This result illustrates one of the major implications of the approach that we take in measuring human capital in this paper: differences in years of schooling are not perfect (or even good in some cases) measures of differences in the stock of human capital. Cross-country differences in the quality of schooling can be large, and depend on the level of development. If $R-s$ does not vary much across countries, and since, in the steady state, wages are proportional to TFP, it follows from (11) that

$$h(6+s) \propto z^{\frac{\gamma_2}{(1-\theta)(1-\gamma)}}, \quad (12)$$

where θ stands for capital's share of income, and z is TFP. Thus, the elasticity of initial human capital with respect to TFP is $\gamma_2/(1-\theta)(1-\gamma)$. It is clear that if the human capital production technology is 'close' to constant returns, then the model will predict large cross country differences in human capital even if TFP differences are small.

4. **The Allocation of Time and Goods Over the Life Cycle.** It can be shown (see equation (38) in the Appendix) that, initially, an individual allocates all his time to producing human capital. This is the period that we identify as schooling. In the post-schooling period, and as the individual ages, he allocates less and less time to accumulating human capital. In terms of the value of goods allocated to the formation of human capital, the model implies increasing costs as a function of years spent in school, and a decreasing function of age in the

⁵To be precise, we find that $\gamma_2 = 0.33$, and $\gamma = 0.93$. Thus the elasticity of the quality of human capital with respect to wages is 4.71. Had we chosen to adjust z_h so that years of schooling are the same, the qualitative result is the same. In this case, the elasticity is $v/(1-v)$, which in our preferred parameterization is approximately 1.2

post-schooling period.⁶ (See equations (39) and (40) in the Appendix.)

2.1.1 Equilibrium Age-Earnings Profiles

Even though the model is very explicit about market income and investments in human capital, it says very little about the timing of payments and who pays for what. In particular, during the post-schooling period it is necessary to determine who pays for the time and good costs associated with training. In order to define measured income at age a , $y(a)$ we assume that a fraction π of post-schooling expenses in market goods are paid for by employers, and subtracted from measured wages. Thus,

$$y(a) = wh(a)(1 - n(a)) - \pi x(a).$$

Given the solution to the income maximization problem (see equation (41) in the Appendix), measured income is

$$\begin{aligned} y(a) = & \left[\frac{\gamma_2^{\gamma_2} \gamma_1^{\gamma_1} z_h w^{\gamma_2}}{(r + \delta_h)^\gamma} \right]^{\frac{1}{1-\gamma}} w \{ \gamma_1 e^{-\delta_h(a-6-s)} \frac{m(6+s)^{\frac{1}{1-\gamma}}}{r + \delta_h} \\ & - (\gamma_1 + \pi\gamma_2) \frac{m(a)^{\frac{1}{1-\gamma}}}{r + \delta_h} + \frac{e^{-\delta_h(a-R)}}{\delta_h} \int_{e^{\delta_h(6+s-R)}}^{e^{\delta_h(a-R)}} [(1 - x^{\frac{r+\delta_h}{\delta_h}})^{\frac{\gamma}{1-\gamma}} dx] \}. \end{aligned} \quad (13)$$

Let $p = a - s - 6$ be the level of experience. In this case (13) is given by

$$\begin{aligned} \hat{y}(s, p) = & \left[\frac{\gamma_2^{\gamma_2} \gamma_1^{\gamma_1} z_h w^{\gamma_2}}{(r + \delta_h)^\gamma} \right]^{\frac{1}{1-\gamma}} w \{ \gamma_1 e^{-\delta_h p} \frac{m(6+s)^{\frac{1}{1-\gamma}}}{r + \delta_h} \\ & - (\gamma_1 + \pi\gamma_2) \frac{m(p+6+s)^{\frac{1}{1-\gamma}}}{r + \delta_h} + \frac{e^{-\delta_h(p+6+s-R)}}{\delta_h} \int_{e^{\delta_h(6+s-R)}}^{e^{\delta_h(p+6+s-R)}} [(1 - x^{\frac{r+\delta_h}{\delta_h}})^{\frac{\gamma}{1-\gamma}} dx] \}. \end{aligned}$$

The function $\hat{y}(s, p)$ summarizes the implications of the model for the age-earnings profile of an individual. In some sense, one could view this expression as the model's analog of a Mincer-style relationship. However, it is necessary to exercise some caution in order to make this comparison. These are two important reasons why our set-up differs from the Mincerian framework.

⁶If we interpret the post schooling period as being mostly on the job training, the model implies that young workers receive more training than older workers.

First, unlike in Mincer’s theory, schooling is endogenous and varying s in $\hat{y}(s, p)$ can give rise to biased estimates.⁷ Thus, in the context of this model it is necessary to be explicit about the factors that **induce** different individuals to choose different levels of s . For individuals *within a country*, we consider only variations in the parameters (z_h, h_B) as potential sources of heterogeneity. When comparing returns *across countries* both the relevant wage rate and the working horizon (w, R) are allowed to vary in accordance with the model and the data.

Second, $\ln(\hat{y}(s, p))$ is a highly nonlinear function of s . This is the case regardless of whether differences in schooling are due to differences in ability (z_h) or due to differences in the initial stock of human capital (h_B) .

2.2 Equilibrium

Given the individual decision on human capital accumulation and investment as a function of age, all we need is to compute the age structure of the population to determine aggregate human capital. Since the capital-human capital ratio is pinned down by the condition that the marginal product of capital equal the cost of capital, this suffices to determine output per worker.

Demographics Since we consider only steady states, we need to derive the stationary age distribution of this economy. Let $N(a, t)$ be the number of people of age a at time t . Thus, our assumptions imply

$$N(a, t) = e^f N(B, t - a)$$

and

$$N(T, t) = 0.$$

⁷For a discussion of the problems associated with viewing schooling as an exogenous variable see Heckman, Lochner and Todd (2003), and Card (2000).

It is easy to check that in the steady state

$$N(a, t) = \phi(a)e^{\eta t}, \quad (14)$$

where

$$\phi(a) = \eta \frac{e^{-\eta a}}{1 - e^{-\eta T}}, \quad (15)$$

and $\eta = f/B$ is the growth rate of population.

Aggregation It turns out that to compute the equilibrium of the model we only need to determine the per capita aggregate amount of human capital effectively supplied to the market. Let $\bar{h}(r, w)$ be the average (per person) level of human capital as a function of r and w . Thus, $\bar{h}(r, w)$ is given by

$$\bar{h}(r, w) = \int_{6+s}^R h(a)(1 - n(a))\phi(a)da.$$

Equilibrium From (6) it follows that if the bequest constraint is not binding, the interest rate is given by

$$r = \rho + \frac{\alpha_0}{B} + (1 - \alpha_1)\eta. \quad (16)$$

Optimization on the part of firms implies that

$$p_k(r + \delta_k) = zF_k(\kappa, 1), \quad (17)$$

where κ is the physical capital - human capital ratio. The wage rate per unit of human capital, w , is,

$$w = zF_h(\kappa, 1). \quad (18)$$

Let \bar{h} be the average level of human capital when r is given by (16), and w is given by (18). Then, aggregate output and consumption per person, \bar{c} , satisfy

$$\bar{c} = [zF(\kappa, 1) - (\delta_k + \eta)\kappa p_k]\bar{h}. \quad (19)$$

For this to be an equilibrium, we need to verify that, at the candidate solution, $b > 0$. Let $c(a)$ be the individual level of consumption computed from the utility

maximization problem. Then,

$$\bar{c} = \int_0^R c(a)\phi(a)da. \tag{20}$$

It can be shown that the function $c(a)$ depends on one level of consumption, say $c(I)$, and its value can be pinned down by requiring that (19) and (20) hold. This completes the description of the steady state.

The level of output is given by

$$y = zF(\kappa, 1)\bar{h}$$

3 Calibration

We use standard functional forms for the utility function and the final goods production function. The utility function is assumed to be of the CRRA variety

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma},$$

while the production function is assumed to be Cobb-Douglas

$$F(k, h) = zk^\theta h^{1-\theta}.$$

Our calibration strategy involves choosing the parameters so that the steady state implications of the model economy presented above is consistent with observations for the United States (circa 2000). Thus, we calibrate the model to account for contemporaneous observations in the U.S. We then vary the exogenous demographic variables and choose the level of TFP for other countries so that the model's predictions for output per worker match that for the chosen country. Consequently, while output per worker for other countries are chosen so as to match output per worker by construction, the model makes predictions on years of schooling, age earnings profiles and the amount of goods inputs used in the production of human capital.

There are some parameters that are standard in the macro literature. Thus, following Cooley and Prescott (1995), the discount factor is set at $\rho = 0.04$ and the

depreciation rate is set at $\delta_k = .06$. Less information is available on the fraction of job training expenditures that are not reflected in wages. There are many reasons why earnings ought not to be equated with $wh(1-n) - x$. First, some part of the training is off the job and directly paid for by the individual. Second, firms typically obtain a tax break on the expenditures incurred on training. Consequently, the government (and indirectly, the individual through higher taxes) pays for the training and this component is not reflected in wages. Third, some of the training may be firm specific, in which case the employer is likely to bear the cost of the training, since the employer benefits more than the individual does through the incidence of such training. Finally, there is probably some smoothing of wage receipts in the data and consequently, the individual's marginal productivity profile is likely to be steeper than the individual's wage profile. For all these reasons, we set $\pi = 0.5$. We experiment later with $\pi = 0$ and $\pi = 1$. We also assume that the same fraction π is not measured in GDP.

Finally α_1 determines the degree of curvature in the altruism function of the individual. Note that this also determines the real interest rate in other countries. Also, since $f = 0$ in the United States, the value of α_1 is irrelevant to the calibration of the model economy to US data. However, the choice of α_1 does affect the outcomes for other countries where $f > 0$. We proceed by setting $\alpha_1 = 1$ as this implies that real interest rates are constant across countries and in turn implies that (measured) capital output ratios do not vary across countries.⁸ Finally, we assume that $B = 25$.

Our theory implies that it is only the ratio $h_B^{1-\gamma} / (z_h^{1-v} w^{\gamma_2 - v(1-\gamma_1)})$ that matters for all the moments of interest. Consequently, we can choose z , p_k (which determine w) and h_B arbitrarily and calibrate z_h to match a desired moment. The calibrated value of z_h is common to all countries. Thus, the model does not assume any cross-country differences in an individual's 'ability to learn.' This leaves us with 7 parameters,

⁸In future work we plan on studying the impact of demographic changes when $\alpha_1 < 1$. Since in this model we need not be explicit about the consumption side we do not report values for the coefficient of risk aversion, σ and the parameters governing the cost function for children, ν_0 and ν_1 .

$\theta, \alpha_0, \delta_h, z_h, \gamma_1, \gamma_2$ and ν . The moments we seek in order to pin down these parameters are:

1. Capital's share of income of 0.33. Source: NIPA
2. Capital output ratio of 2.52. Source: NIPA
3. Earnings at age R/Earnings at age 55 of 0.8. Source: SSA
4. Earnings at age 50/Earnings at age 25 of 2.17. Source: SSA
5. Years of schooling of 12.08. Source: Barro and Lee
6. Schooling expenditures per pupil (primary and secondary) relative to GDP per capita of 0.214. Source: OECD
7. Pre-primary expenditures per pupil relative to GDP per capita of 0.14. Source: OECD

Thus, we use the properties of the age-earnings profile to identify the parameters of the production function of human capital. This, of course, follows a standard tradition in labor economics. The previous equations correspond to moments of the model when evaluated at the steady state. This, calibration requires us to solve a system of 7 equations in 7 unknowns. The resulting parameter values are

Parameter	θ	α_0	δ_h	z_h	γ_1	γ_2	ν
Value	0.315	0.75	0.018	0.361	0.63	0.3	0.55

Of some interest are our estimates of α_0 and γ_i . Since the first one is positive, it implies that agents are imperfectly altruistic. Our estimate of γ_2 is fairly large, and indicates that, in order for the model to be consistent with both average schooling in the U.S. as well as the pattern of the age-earnings in the data, market goods have to enter in the production function of human capital.⁹

⁹The coefficient of capital in the production function, θ , does not coincide with capital share due to the unmeasured component of post-schooling training. In our calculations, unmeasured post-schooling training is approximately 4.7% of measured output.

4 Results

Before turning to the results, we first describe the data so as to get a feel for the observations of interest. We start with the countries in the PWT 6.1 and put them in deciles according to their output per worker, y . Next, we combine them with observations on years of schooling (s), expenditures per pupil relative to output per worker (x_s), life expectancy (T), total fertility rate (f), and the relative price of capital (p_k) for each of these deciles. The population values are displayed in the following table.

Table 1: World Distribution						
Decile	y (relative to US)	s	x_s	T	f (TFR/2)	p_k
90-100	0.921	10.93	3.8	78	.85	1.02
80-90	0.852	9.94	4.0	76	.9	1.11
70-80	0.756	9.72	4.3	73	1	1.06
60-70	0.660	8.70	3.8	71	1.2	1.04
50-60	0.537	8.12	3.1	69	1.35	1.52
40-50	0.437	7.54	2.9	64	1.6	1.77
30-40	0.354	5.88	3.1	57	2.05	1.56
20-30	0.244	5.18	2.7	54	2.5	1.93
10-20	0.146	4.64	2.5	51	2.7	2.11
0-10	0.052	2.45	2.8	46	3.1	2.78

Table 1 illustrates the wide disparities in incomes across countries. The United States possesses an output per worker that is about 20 times as high as the countries in the bottom decile. Further notice that years of schooling also varies systematically with the level of income —from about 2 years at the bottom deciles to about 11 at the top. The quality of education as proxied by the expenditures on primary and secondary schooling as a fraction of GDP also seems to increase with the level of development. This measure should be viewed with a little caution as it includes only

public inputs and not private inputs (including the time and resources that parents invest in their kids). Next, notice that demographic variables also vary systematically with the level of development - higher income countries enjoy greater life expectancies and lower fertility rates. More important, while demographics vary substantially at the lower half of the income distribution, they do not move much in the top half. Finally, the relative price of capital in the richest countries is about a third of the level in the poorest countries.

Development Accounting We now examine the ability of the model to *simultaneously* match the cross country variation in output per capita and years of schooling. To isolate the role of human capital, we ignore cross-country differences in the price of capital. Thus, we set $p_k = 1$ in every country (we relax this later). To be clear, we change R (retirement age) and e^f (fertility rate) across countries and choose the level of TFP in a particular country so as to match output per worker. We then see if the predictions for years of schooling are in accordance with the data. Note that changes in f do not affect the individual allocations but affect aggregates. Also, differences in T (life expectancy) have no effect on individual human capital accumulation, but affect aggregates through the distribution of the population. Note also that since $\alpha_1 = 1$ and we hold fixed the relative price of capital, there is no variation whatsoever in the physical capital to output ratio. Consequently, the only source of variation in GDP per worker (apart from TFP) is human capital per worker

Table 2: Output and Schooling - Data and Model						
(Constant p_k)						
Decile	y (relative to US)	TFP	s		x_s	
			Data	Model	Data	Model
90-100	0.921	0.99	10.93	11.64	3.8	3.7
80-90	0.852	0.98	9.94	10.92	4.0	3.5
70-80	0.756	0.97	9.72	10.40	4.3	3.4
60-70	0.660	0.95	8.70	9.64	3.8	3.9
50-60	0.537	0.93	8.12	8.90	3.1	4.9
40-50	0.437	0.90	7.54	6.79	2.9	4.7
30-40	0.354	0.88	5.88	5.69	3.1	3.9
20-30	0.244	0.85	5.18	4.29	2.7	3.2
10-20	0.146	0.82	4.64	3.01	2.5	2.7
0-10	0.052	0.73	2.45	2.19	2.8	2.0

Table 2 presents the predictions of the model and the data. The striking results are the estimates of TFP. In our model, TFP in the poorest countries (i.e. countries in the lowest decile of the world income distribution) is estimated to be only 73% of the level of TFP in the United States. This is in stark contrast to the results of Parente and Prescott (2000), Hall and Jones (1999) and Klenow and Rodriguez-Clare (1997) who find that large differences in TFP are necessary to account for the observed differences in output per worker. By way of comparison, the corresponding number in their studies is around 20%. Thus, their estimate of TFP in the poorest countries is between two to three times lower than ours.

What is driving our results? Since we take schooling to be endogenous, it is possible that our model requires smaller differences in TFP because it implies large differences in schooling or, alternatively, that it allocates too many resources to schooling and, hence, it exaggerates the differences across countries. We now argue that this is not the case. The implications for average schooling across deciles of the world income distribution match the available data rather closely. In terms of a rough measure of quality such as expenditures per pupil, the model actually underpredicts investment at the two ends of the world income distribution.¹⁰ Thus, this cannot explain our findings.¹¹ We used the model to compute the elasticity of output with respect to TFP when all factors are allowed to vary (this is the very long run), and the economy has adjusted to the new steady state. We estimate this elasticity to be around 8. This estimate suggest that, in the long run, there are large payoffs in terms of output per worker of small changes in TFP.

A second source of differences across countries is demographics. At the individual

¹⁰The model overpredicts x_s for countries in the middle of the distribution.

¹¹As mentioned before, the model makes predictions on the total amount of goods used in the production of schooling including the value of goods and time parents allocate to educating their children outside of formal schooling. The data includes only the expenditures classified as (public) school expenditures. Moreover, it is not clear to what extent capital costs are included in this measure.

level earlier retirement (lower R) induces less demand for human capital, as it can only be used for fewer periods. Since poor countries have lower effective values of R , this results in lower levels of human capital. At the aggregate level, differences in fertility result in differences in the fraction of the population that is at different stages of their working life. Since poorer countries tend to have a larger fraction of the working age population concentrated in the younger segments, and since human capital increases with age (except near the end of working life), aggregation results in smaller levels of human capital for poorer countries. Thus, as we will argue later, differences in demographics play a significant role.

Differences in the Price of Capital So far we have assumed that there are no distortions in the price of capital. Following Chari Kehoe and McGrattan (1997) we now allow p_k to vary according to the values in Table 1. Table 3 presents the results.

Table 3: Output and Schooling - Data and Model				
(Varying p_k)				
Decile	y	p_k	TFP	TFP
	(relative to US)		baseline	p_k varies
90-100	0.921	1.02	0.99	1.00
80-90	0.852	1.11	0.98	1.01
70-80	0.756	1.06	0.97	0.99
60-70	0.660	1.04	0.95	0.96
50-60	0.537	1.52	0.93	1.05
40-50	0.437	1.77	0.90	1.07
30-40	0.354	1.56	0.88	1.01
20-30	0.244	1.93	0.85	1.05
10-20	0.146	2.11	0.82	1.04
0-10	0.052	2.78	0.73	1.01

When the price of capital varies according to the data, no differences in the level of productivity are needed to account for the world income distribution. Thus, differences in the price of capital and endogenous accumulation of inputs (mostly human capital) can account for all of the observed differences in output per worker.

Changing Demographics Before we discuss the estimated impact on human capital, it is of interest to explore the impact of demographics. To gain insight into this question, imagine holding TFP fixed at the baseline level (where the relative price of capital is also held fixed) and imagine changing all the demographic variables to the US level. The results of such an experiment are presented in Table 4.

Table 4: Output and Schooling - Data and Model					
Changing Demographics					
Decile	y		s		
	baseline	demog	Data	baseline	demog
90-100	0.921	0.913	10.93	11.64	11.70
80-90	0.852	0.851	9.94	10.92	11.21
70-80	0.756	0.756	9.72	9.40	10.2
60-70	0.660	0.664	8.70	8.64	9.33
50-60	0.537	0.572	8.12	7.30	8.56
40-50	0.437	0.483	7.54	6.49	7.92
30-40	0.354	0.402	5.88	5.49	7.12
20-30	0.244	0.331	5.18	4.29	5.97
10-20	0.146	0.251	4.64	3.01	4.79
0-10	0.052	0.123	2.45	2.19	4.04

For example, if countries in the lowest decile were to have the same demographic profile as the United States, their output per worker would increase more than 100% (from 5.2% to 12.3% of the U.S. level). This is accompanied by the doubling in the level of schooling. Thus, in this experiment, demographic change drives both

schooling and output. Thus, the model is consistent with the view that changes in fertility can have large effects on output. It is important to emphasize that our quantitative estimates reflect long run changes. The reason is that they assume that the level of human capital has fully adjusted to its new steady state level. Given the generational structure, this adjustment can take a long time.

As expected, even though demographic change will substantially help poor countries, it will not have much of an impact among the richest countries. For example, for countries in the second decile (with initial income between 80% and 90% of the richest countries) there is no change in output per worker.

Even though we find large effects associated with demographic change our results should be viewed with caution since we assume that demographic change is orthogonal to changes in TFP, while in a model of endogenous fertility it is likely that macro conditions will affect fertility decisions (and longevity). The important observation is that changes in fertility induced by aggregate changes can have large effects on income through their impact on human capital accumulation decisions.

5 The Role of Human Capital: Discussion

In this section we describe some of the implications of the model. We emphasize those aspects that provide us insights on how cross-country differences in TFP can account for differences in schooling and the quality of human capital.

Quality vs. Quantity of Human Capital As indicated before, a key element of our model is that the quality of human capital varies systematically with the level of development. Equation (12) displays the elasticity of the stock of human capital at the time an individual completes his formal schooling with respect to TFP. It is given by $\gamma_2 / (1 - \theta)(1 - \gamma)$.¹² Given our estimated parameters ($\gamma_1 = 0.6, \gamma_2 = 0.33$

¹²This formula assumes that the interest rate and the term $R - s$ does not vary much across countries, which is not a bad approximation in our data.

and $\theta = 0.315$), this elasticity is 6.88. This, in turn, implies that the differences in the stock of human capital at the end of schooling are substantial. For example, the difference in TFP between a country in the bottom decile and the top decile is approximately 37% ($1/0.73$, see Table 2). Thus, the model implies that the stock of human capital of the typical individual at the time he leaves school in a country from the top decile is approximately 8.7 times the level of human capital of an individual from a poor country.

This difference is due both to differences in the level of schooling (quantity) and differences in the quality of human capital. In order to get a sense of the relative importance of these two factors, let's specify that

$$h_i(s + 6) = h_i^q e^{\phi_i s},$$

where h_i^q is an index of the quality of human capital and ϕ_i is the rate of return on schooling in country i . Let US and P index the U.S. and a poor (bottom decile) country, respectively. It follows that

$$\frac{h_{US}^q}{h_P^q} = \left(\frac{z_{US}}{z_P} \right)^{\frac{\gamma}{(1-\theta)(1-\gamma)}} e^{\phi_P s_P - \phi_R s_R}.$$

Given our parameter estimates and assuming that the rate of return to schooling is approximately 10% in both countries (i.e. $\phi_{US} = \phi_P = 0.10$), and that the difference $s_{US} - s_P$ is 10 years, it follows that our estimate of the ratio of the qualities is over 3.2. Thus, in this simple calculation, over 50% of the differences in the stocks of human capital (at the time the average individual leaves school) is due to differences in quality.

The Importance of Early childhood, and On-the-Job Training Our model implies that, even at age 6, there are large differences between the human capital of the average child in rich and poor countries. In Table 5 we present the values of human capital at age 6 (h_E) for each decile relative to the U.S.

Table 5: Human Capital			
(relative to U.S.)			
Decile	y	h_E	\bar{h}
	(relative to US)	(relative to US)	(relative to US)
90-100	0.921	0.96	0.95
80-90	0.852	0.91	0.88
70-80	0.756	0.88	0.79
60-70	0.660	0.86	0.71
50-60	0.537	0.79	0.60
40-50	0.437	0.72	0.50
30-40	0.354	0.65	0.43
20-30	0.244	0.60	0.32
10-20	0.146	0.53	0.20
0-10	0.052	0.47	0.08

Even though the differences in early childhood capital are small for the relatively rich countries (output per worker at least 75% of the U.S.), the differences are large when comparing rich and poor countries. Our estimates suggest that a six year old from a country in the bottom decile has less than 50% of the human capital of a U.S. child.

A Comparison with the Mincerian Approach At this point it is useful to compare the differences between our analysis based on an explicit optimizing approach (where schooling and the earnings profile are endogenous) with an approach that takes the results of a Mincer regression as estimates of a production function. The Mincerian framework implies that the average human capital of a worker in country i with s_i years of schooling is

$$\hat{h}_i = Ce^{\phi_i s_i}.$$

The standard approach uses an estimate of $\phi_i = \phi \approx 0.10$, which corresponds to a 10% return. Thus, if we take a country from the lowest decile with $s_P = 2$, and assuming that the average worker in the U.S. has 12 years of schooling, we estimate that the average human capital of the poor country (relative to the U.S.) is

$$\frac{\hat{h}_P}{\hat{h}_{US}} = e^{-.1 \times 10} = 0.37.$$

Our approach, in a reduced form sense, allows for the Mincerian intercepts to vary across countries. Thus in our specification, we can view average human capital in country i as

$$\bar{h}_i = C_i e^{\phi_i s_i}.$$

If, as before, we compare a country from the bottom decile of the output distribution with the U.S., Table 5 implies that its relative average human capital is 0.08. It follows that our measure of quality, for this pair of countries, is simply

$$\frac{C_P}{C_{US}} = \frac{\bar{h}_P}{\bar{h}_{US}} e^{\phi(s_{US} - s_P)} = 0.08 \times 2.71 = 0.22.$$

Thus, our numerical estimate is that the quality of human capital in a country in the lowest decile is approximately one fifth of that of the U.S. In our model, this ratio is driven by differences in wages and demographics. The magnitude of the differences in relative quality points to the quantitative importance of ignoring differences in quality.¹³

The Role of Curvature in the Production Function of Human Capital

As the previous elasticity estimates make clear, our results depend on the degree of returns to scale in the production of human capital as given by γ_1 and γ_2 . Our

¹³In a recent paper, Caselli (2003) explicitly models, in a reduced form sense, differences in C_i across countries. He then uses some empirical results to estimate how much of the differences in country characteristics can explain differences in quality and concludes that these cannot be important factors. Our results differ from his in that we use an explicit model to compute quality differentials.

estimate is $\gamma = \gamma_1 + \gamma_2 = 0.93$. Prior research found a range of values from about 0.5 to 1. The key element that allows us to pin down a value of γ is that we require that our parameterization explain both the number of years of schooling and the evidence on age-earnings profile. Most previous work including Ben-Porath, Haley and others estimate the model taking schooling as given. Consequently, estimations of the Ben Porath model have focused on the period of specialization (post schooling).¹⁴ The (apparently) high value of γ is needed to match the age earnings profile *and the schooling level* that we see in the data. Lower values of γ (with z_h adjusted so as to match years of schooling) imply too steep an age earnings profile. One might be led to believe that if we were to reduce γ sufficiently, at some point the steepness of the profile will be approximately what will prevail with the high γ that we calibrate it to. (After all, when $\gamma = 0$, the age earnings profile is flat.) While true in principle, it turns out that with a low γ , say 0.5, the model cannot match the implications for years of schooling even if ability (or wages for that matter) were infinitely large!

The Fraction of the Time Allocated to Producing Human Capital If we ignore the contribution of individuals of age less than 6, the fraction of the available time allocated to the production of human capital is

$$\bar{n} = \frac{\int_6^R n(a)\phi(a)da}{\int_6^T \phi(a)da}.$$

For our baseline parameterization $\bar{n} = 0.29$. This compares with a value of 0.24 used by Parente and Prescott (2002) for their preferred parameterization. Moreover, it shows that the degree of curvature in the production function of human capital ($\gamma = \gamma_1 + \gamma_2$) can be large and still keep the model from predicting an unusually large allocation of time to human capital creation. In fact, with a lower value of γ ,

¹⁴As exception is important work by Heckman, Lochner and Taber who use different technologies for producing schooling and training on the job - in particular they assume that while schooling is goods intensive in the sense that there are (exogenous) tuition costs, goods do not matter for the production of training.

the fraction of time allocated to human capital accumulation \bar{n} is a lot higher.

6 Some Implications of the Model

The main difference between our set-up and other approaches lies in the specification of the production function of human capital. It seems natural then to ‘test’ the model by confronting some of its implications with the data. There are two dimensions that seem worth exploring. First, since our estimates of the stock of human capital are very different from those obtained using estimates of a Mincer-style regression, it is not clear whether data generated by our model can match the estimated return to schooling. Second, since our model relies on cross country differences in the quality of human capital it has sharp implications about the incomes of immigrants. To test how reasonable the model is, we compare the predictions of the model with the evidence on the behavior of earnings of immigrants.

6.1 Mincer Regressions

Even though the interpretation and the precise point estimate of the schooling coefficient in a Mincer regression are controversial, most estimates—at least when linearity is imposed—seem to be close to 10%.¹⁵ Thus, one challenge for the model economy is to reproduce the rate of return in a Mincer-style regression.

However, since the model predicts that all (homogeneous) individuals choose exactly the same level of schooling, it is necessary to introduce some source of microeconomic heterogeneity. As we mentioned before, to induce differences among individuals within a country the model has two natural candidates: differences in z_h

¹⁵The assumption that the relationship between log earnings and schooling is linear is also controversial. Heckman, Lochner and Todd document significant non-linearities. More recently, Belzil and Hansen (2002) find that, when the return is allowed to be a sequence of spline functions, the function is convex.

(ability to learn), and differences in h_B (initial human capital). From the results in Proposition 2 it follows that the equilibrium years of schooling depend on the ratio $h_B^{1-\gamma} / (z_h^{1-\nu} w^{\gamma_2 - \nu(1-\gamma_1)})$. Since in a given country all individuals face the same wage, differences in s are driven by differences in (z_h, h_B) . These two variables have very different impacts on lifetime earnings. Heterogeneity in z_h results in lifelong differences in earnings (lack of convergence across individuals), while differences in h_B get smaller with age.

For our computations we varied z_h (and h_B) so as to generate lifetime earnings for individuals who choose to acquire between 0 and 20 years of education. Given the non-linearity of the earnings function, we need population weights of individuals in different categories of experience and schooling. We obtain these population weights from the NLSY, with schooling ranging from 0 through 20 and experience going from 5 to 45. We then proceed in two steps: If the only source of heterogeneity is in ability, we adjust z_h from its baseline value in order to obtain the ability levels that lead to the different schooling levels. Thus, there will be as many ability levels as there are schooling levels. We also have their predicted age earnings profiles. Next, we draw observations from the experience-schooling categories depending on their population weights. For instance, if the group with 12 years of schooling and 10 years of experience has a mass of .1 while the group with 12 years of schooling and 30 years of experience has a mass of .05, we then draw twice as many observations from the first category relative to the second. We then run a standard Mincer regression with schooling, experience and the square of experience as independent variables and the logarithm of earnings on the left. We repeat these steps and recover the Mincerian return when the only source of variation is in initial human capital.¹⁶

The Mincer coefficient generated by variation in ability alone is around 13% while that obtained from variation in h_B alone is close to 0. In order to obtain a point estimate of the return, we need to know the joint distribution of z_h and h_B . However,

¹⁶We follow the same procedure when we adjust h_B .

given the rather tight bounds that we obtain, we conclude that the model is consistent with the ‘stylized fact’ that the Mincerian return for the United States is around 8%.

As a second test, we computed for each representative country in our world distribution of output (10 countries in all) the effect on log earnings of an additional year of education, and we took this to be the return on schooling in country (decile) i . We then regressed this return on the log of GDP per capita and obtained a coefficient of -0.10 (when z_h is the only source of heterogeneity), and -0.04 (when h_B varies). This is to be compared with a similar exercise —with actual data— run by Banerjee and Duflo (2004) using different data sets. Their estimate is -0.08. Thus, depending on the mix between z_h and h_B the model can account for the cross-country evidence on Mincerian returns.

Thus, to summarize, the cross-section (within a country) relationship implied by the model between returns to schooling and years of schooling is positive, while the cross-country estimate is negative. Even though this looks like a contradiction, that is not the case. The key observation is that along a given earnings-schooling profile (for a given country) only individual characteristics are changing, while the profiles of different countries reflect differences in demographics and wage rates. It is possible to show that demographic differences and differences in wage rates imply that the earnings-schooling profile of a poor country lies below that of a rich country. It turns out, that the poor country profile is also steeper than the rich country profile. Since the return to formal education is, approximately, the derivative of the earnings-schooling profile, it is the increased steepness of the earnings-schooling profile as TFP decreases (a cross-country effect) that dominates the convexity of the profile as schooling increases (for a given level of TFP) that is the dominant effect that accounts for the cross-country observations.

6.2 Immigrant Evidence

A key prediction of the model is that the quality of human capital varies (inversely) with the level of development. Thus, it implies that if we compare two individuals with the same level of schooling acquired in different countries, their effective amount of human capital will also be different. A simple test of the model would be to bring an individual from a poor country to a rich country and observe his income relative to a native with the same schooling.

One imperfect piece of evidence related to this thought experiment is provided by the experience of immigrants in relatively rich countries. Our casual reading of this literature suggests the following stylized facts (related to immigrants in the U.S.).

Fact 1 Immigrants earn initially lower income than comparable natives with the same level of schooling. This wage differential has been increasing over time.

Fact 2 The growth rate of earnings of immigrants is higher than the growth rate of earnings of similar—in terms of measurable characteristics—natives.

Fact 3 The level of earnings of recent immigrants, holding schooling constant, is positively related to the level of per capita output in their country of origin.

The model presented in this paper is too simple to account for the complex and changing patterns of migration. Nevertheless, in order to get a rough idea of the quantitative predictions of the model, we study earnings in the U.S. (predicted by the model) of an immigrant who has approximately the years of schooling of the average immigrant in the U.S., and originates from different countries as measured by the level of development. Once this individual has migrated, he chooses investment in human capital optimally given the prices that he faces and the new working horizon (we assume that the migrant retires at the same time as the natives). We analyze earnings of a 25 year old migrant who chooses not to go back to school.¹⁷ Our theory

¹⁷Had we chosen to study a younger immigrant, the model would have predicted that some of

implies that an immigrant from a poor country will earn less than a native and that he will choose to invest more in human capital since he starts with a lower stock of human capital than the comparable native. Thus, our theory also predicts some catch-up. It is clear that, qualitatively, the model is consistent with the facts.

We now discuss the ability of the model to match the data from a quantitative point of view.

Fact 1 Borjas (1994) estimates that, for recent arrivals, the percentage wage differential between immigrant and native men increased from -16.6% in 1970 to -31.7% in 1990 (see Borjas (1994), Table 3). In order to estimate the implications of the model, we need time series estimates of the schooling levels for natives and immigrants, as well as the ‘identity’ of their country of origin, so that we can estimate the change in ‘quality’ of the human capital of the average immigrant. Borjas (1992) estimates that, in 1970, the average immigrant had .2 years of education less than a native (who had 11.3 years at the time), while in 1990 we estimate that the average immigrant had 12.5 years of schooling (natives had 13).¹⁸ He also reports that the level of GNP per capita in the country of origin of the typical recent immigrant in 1970 was slightly above 50% of the U.S., while in the 1980s (we do not have data for 1990) it had decreased to approximately 39% of U.S. GNP per capita. Using those values the model predicts that initial —defined as the average over the first five years after immigration— earnings of the average immigrant are 15% lower than those of the natives in 1970, and 23% in 1990. The model is consistent with the view that the ‘quality’ of the average immigrant has decreased, and this is one reason why recent immigrants

them —depending on the country of origin— would enroll in school after migrating to the U.S. This is consistent with the findings of Betts and Lofstrom (2000), but we do not pursue this line here.

¹⁸This estimate implies that the gap between immigrants and natives that was estimated to be large in 1980 by Borjas, has narrowed in the 1990s. For evidence on this see Betts and Lofstrom (2000).

earn less than natives (see Borjas (1994)).¹⁹

Fact 2 Borjas (1994) reports the evidence on the growth rate of earnings of immigrants relative to natives. The precise amount of catch-up is controversial (see Borjas (1994) for a discussion), but it is in the range of 6-15% for the first decade after immigration to 10-25% for the first two decades after immigration. We analyzed the 10 and 20 year average growth rate of earnings (relative to natives with the same years of schooling) for two individuals: one that comes from a country in the middle of the world income distribution and the other that comes from a country in the lowest decile. As before, we considered both individuals that differ in terms of their h_B , as well as immigrants who differ (from their fellow country men) in terms of z_h . The results (the first number corresponds to h_B , while the second gives the predictions for z_h) are presented in Table 6. Our estimates fall within the range reported in the literature and capture the actual amount of catch-up

Table 6. Growth Rate of Relative Income		
GNP (origin)	10-year	20-year
Middle Income	5-14%	7-20%
Low Income	8-19%	11-27%

Fact 3 We used the model to estimate the level of initial income (first five years) for an individual with 12 years of schooling as a function of output per worker in his country of origin. We computed earnings in the U.S. of an immigrant from a

¹⁹The model underpredicts the drop in income. However, this is due to our choice of concentrating on selecting immigrants in terms of z_h . If immigrants were selected only in terms of differences in h_B , the model predicts differences of -37% and -51%. Thus if the proportion of ‘ z_h immigrants’ was 92% in 1970 and 68% in 1990, the model would perfectly predict the observed differences in income. The increase in the proportion of immigrants who gain less from immigrating is consistent with the change in U.S. immigration policy that reduced the number of ‘economic’ migrants in favor of individuals with family ties.

middle income country (50% of U.S. output per worker) and a poor country (7% of U.S. output per worker). Given the non-linearity of the model, the computed elasticity is sensitive to the choice of country (as well as source of variation). We find estimates that range from 0.01 to 0.04 (when only z_h is varied) to 0.05 to 0.17 (when only h_B is varied). Borjas (1994) indicates that the elasticity of earnings with respect to GDP is around 0.04, while Borjas (2000, Table 1.6 column 4) estimates the elasticity around 0.05. Thus, the model’s predictions are roughly in agreement with the evidence.

Jasso, Rosenzweig and Smith (2002) report earnings of immigrants in their country of origin and in the U.S. They find that migration is associated with significant increases in earnings. The reported differences correspond to average ‘before’ and ‘after’ wages. Large differences are evidence of substantial TFP differences. However, Jasso and Rosenzweig also report that 24% of the sample earned less in the U.S. than in their country of origin. The small sample size (230 observations), the imputation criteria²⁰, and the lack of data that would allow us to use our model to predict earnings²¹, makes us cautious about using these estimates.²²

In research that is quite close in spirit to our work, Hendricks (2002), uses earnings of immigrants to the U.S. to estimate the human capital in the country of origin. He concludes that human capital differences cannot account for a large share of the cross-country variation in output per worker. In Hendricks’ framework —and ignoring self-selection— if the ‘efficiency’ of human capital in country j is close to one, then the Mincer estimate of human capital is approximately correct, and the

²⁰Jasso and Rosenzweig imputed full time earnings on the basis of their info. They do not report what fraction of the sample had its earnings imputed.

²¹To be precise, we need data on the time elapsed between last job in the country and first job in the U.S. so that we can use our model to adjust for changes in human capital.

²²In addition, Hendricks (2002) reports that in an earlier paper, Jasso, Rosenzweig and Smith (1998), they find no effects on Mexican immigrants who, we suspect, comprise a significant proportion of the sample.

approach pioneered by Klenow and Rodriguez-Clare (1997), Hall and Jones (1999) yields the correct estimate of human capital for each country. Hendricks' estimates of the efficiency parameter are quite high, with only 7 out of 67 countries in the sample displaying efficiency levels below 80% of the U.S., and the 'low income' sample having an average efficiency level of 90% of the U.S.²³

Hendricks' results are not inconsistent with the model in this paper. In order to use our model to 'predict' the efficiency parameter we need to estimate the human capital of a migrant from a given country with a given number of years of education. For example, the average years of schooling of immigrants from the only two countries in the second lowest decile in Hendricks' sample²⁴ is over 15, while the average years of schooling in the decile of origin is 4.64. In order to 'produce' an individual with 15 years of schooling from the second lowest decile, we followed the same strategy and picked his z_h to match the level of schooling of immigrants. In this case, our estimates of efficiency are around 80-90%, and broadly consistent with Hendricks'.²⁵

Overall, the model does a reasonable job at tracking the earnings dynamics of immigrants, even though it was not designed for that purpose. In particular, the evidence on immigrant income lends support to the view that some of the differences in output per worker are driven by differences in the quality of human capital.²⁶

²³The efficiency levels in the one type of human capital model can be estimated from the information in Table 1 (p. 204) in Hendricks (2002). They are given by the ratio w_c^h/w_c^N , which can be inferred from the Table. Hendricks sample is not comparable to ours. There is no data for any country in the lowest decile, and data for only two countries for the second poorest decile. We thank Lutz Hendricks for his help.

²⁴These countries are Ghana and Kenya. Ethiopia, which is in the poorest 10%, is listed in the Appendix, but no data are presented in Table 1.

²⁵Hendricks finds that for many countries, including some relatively poor countries, the efficiency parameter exceeds 100%. Our model, by construction, has to find efficiency levels below 100%. Of course, this ignores selection.

²⁶In our discussion we completely ignored the impact of differences in languages and learning about the host country environment. These are important considerations and a search model or a set-up along the lines of Jovanovic (1979) models can also account for steeper age-earnings profiles and

7 Conclusion

Our results show that human capital has a central role in determining the wealth of nations. In particular, we show that an extended neoclassical model that incorporates a human capital sector is capable of generating large differences in the stocks of human capital with these differences arising out of small differences in TFP. The novelty is that the model implies that the quality of human capital varies systematically with the level of development. The model is quite successful in capturing the large variation in levels of schooling across countries and is also consistent with the cross-country evidence on rates of return, as well the behavior of earnings of immigrants. The model also implies that a large fraction of the cross-country differences in output are due to differences in the quality of human capital. To be precise, the typical individual in a poor country not only chooses to acquire fewer years of schooling, he also acquires less human capital per year of schooling.

The conventional wisdom is that cross-country differences in human capital are small and that consequently differences in TFP are large. Hence policies that achieve small changes in TFP cannot have large effects on output per capita. Moreover, using the Mincer approach that takes schooling as exogenous, those models effectively give up on trying to understand the impact of TFP on human capital accumulation. We find that, the elasticity of output per worker with respect to TFP is approximately 8. Thus, a mere 10% permanent increase in relative TFP, is predicted to increase output per worker in the long run by 80%. The model suggests that there are huge payoffs to understanding what explains productivity differences. Thus, in our model, productivity differences play a central role in explaining development.

We also find a significant role for policies that induce demographic change. We estimate that if a country in the lowest decile of the world income distribution was endowed with the demographic characteristics of the representative country in the

lower initial wages. These generalizations are beyond the scope of this paper. For a nice exposition of other theories of the earnings distribution, see Neal and Rosen (1999).

top decile, output per worker would double.

Naturally, the consideration of capital market imperfections such as binding inter-generational loan markets (which will result in the steady state of the open economy version of the model presented above) will only increase the role played by demographics and further reduce the importance of TFP. In ongoing work, we are studying the impact of a variety of human capital policies in the presence of distortions.

8 Appendix

Proof of Proposition 1. : We show that the first order conditions corresponding to both problems coincide. Since the problems are convex, this suffices to establish the result. Consider first the first order conditions of the income maximization problem given the stock of human capital at age 6, $h(6) = h_E$. Let $q(a)$ be the costate variable. A solution satisfies

$$whn \leq q\gamma_1 z_h (nh)^{\gamma_1} x^{\gamma_2}, \quad \text{with equality if } n < 1, \quad (21a)$$

$$x = q\gamma_2 z_h (nh)^{\gamma_1} x^{\gamma_2}, \quad (21b)$$

$$\dot{q} = rq - [q\gamma_1 z_h (nh)^{\gamma_1} x^{\gamma_2} h^{-1} - \delta_h] - w(1 - n), \quad (21c)$$

$$\dot{h} = z_h (nh)^{\gamma_1} x^{\gamma_2} - \delta_h h, \quad (21d)$$

where $a \in [6, R]$. The transversality condition is $q(R) = 0$.

Let Φ be the Lagrange multiplier associated with the budget constraint (2). Then, the relevant (for the decision to accumulate human capital) problem solved by a parent is

$$\begin{aligned} \max \Phi \{ & \int_I^R e^{-r(a-I)} [wh(a)(1 - n(a)) - x(a)] da \\ & + e^f \int_B^{B+I} e^{-r(a-I)} [wh_k(a)(1 - n_k(a)) - x_k(a)] da \\ & - e^f e^{-rB} b_k - e^f e^{-r(B+6)} x_E \} + e^{-\alpha_0 + \alpha_1 f} e^{-\rho B} V^k(h_k(B+I), b_k), \end{aligned}$$

where, in this notation, a stands for the parent's age. It follows that the first order conditions corresponding to the choice of $[h(a), n(a), x(a), q_p(a)]$ are identical to those corresponding to the income maximization problem (21), including the transversality condition $q_p(R) = 0$ for $a \in [I, R]$. It follows that $q_p(a) = q(a)$. Simple algebra shows that the first order conditions corresponding to the optimal choices of $[h_k(a), n_k(a), x_k(a), q_k(a)]$ also satisfy (21) for $a \in [6, I]$. However, the appropriate transversality condition for this problem is

$$q_k(B+I) = e^{-[\alpha_0 + (1-\alpha_1)f]} e^{-(\rho-r)B} \frac{1}{\Phi} \frac{\partial V^k(h_k(B+I), b_k)}{\partial h_k(B+I)}.$$

However, given (6), and the envelope condition

$$\frac{\partial V^k(h_k(B+I), b_k)}{\partial h_k(B+I)} = \Phi_k q_p(I),$$

evaluated at the steady state $\Phi = \Phi_k$, it follows that

$$q_k(B+I) = q_p(I).$$

Thus, the program solved by the parent (for $a \in [I, R]$) is just the continuation of the problem he solves for his children for $a \in [6, I]$. It is clear that if (6) does not hold, then there is a ‘wedge’ between how the child values his human capital after he becomes independent, $q_p(I)$, and the valuation that his parent puts on the same unit of human capital, $q_k(B+I)$. ■

For simplicity, we prove a series of lemmas that simplify the proof of Proposition 2. It is convenient to define several functions that we will use repeatedly.

Let

$$C_h(z_h, w, r) = \left[\frac{\gamma_2^{\gamma_2} \gamma_1^{\gamma_1} z_h w^{\gamma_2}}{(r + \delta_h)^\gamma} \right]^{\frac{1}{1-\gamma}},$$

and

$$m(a) = 1 - e^{-(r+\delta_h)(R-a)}.$$

The following lemma provides a characterization of the solution in the post schooling period.

Lemma 3 *Assume that the solution to the income maximization problem stated in Proposition 1 is such that $n(a) = 1$ for $a \leq 6 + s$ for some s . Then, given $h(6 + s)$ the solution satisfies, for $a \in [6 + s, R]$,*

$$x(a) = \left(\frac{\gamma_2 w}{r + \delta_h} \right) C_h(z_h, w, r) [1 - e^{-(r+\delta_h)(R-a)}]^{\frac{1}{1-\gamma}}, \quad a \in [6 + s, R), \quad (22)$$

$$h(a) = e^{-\delta_h(a-6-s)} \left\{ h(6+s) + \frac{C_h(z_h, w, r)}{\delta_h} e^{-\delta_h(6+s-R)} \int_{e^{\delta_h(6+s-R)}}^{e^{\delta_h(a-R)}} \left(1 - x \frac{r+\delta_h}{\delta_h} \right)^{\frac{\gamma}{1-\gamma}} dx \right\}, \quad a \in [6 + s, R), \quad (23)$$

and

$$q(a) = \frac{w}{r + \delta_h} [1 - e^{-(r+\delta_h)(R-a)}], \quad a \in [6 + s, R]. \quad (24)$$

Proof of Lemma 3. : Given that the equations (21) hold (with the first equation at equality), standard algebra (see Ben-Porath, 1967 and Haley, 1976) shows that (24) holds. Using this result in (21b) it follows that

$$x(a) = \left[\frac{\gamma_2^{\gamma_2} \gamma_1^{\gamma_1} z_h w^{\gamma_2}}{(r + \delta_h)^\gamma} \right]^{\frac{1}{1-\gamma}} \left(\frac{\gamma_2 w}{r + \delta_h} \right) [1 - e^{-(r+\delta_h)(R-a)}]^{\frac{1}{1-\gamma}},$$

which is (22). Next substituting (22) and (24) into (21d) one obtains a non-linear non-homogeneous first order ordinary differential equation. Straightforward, but tedious, algebra shows that (23) is a solution to this equation. ■

The next lemma describes the solution during the schooling period.

Lemma 4 *Assume that the solution to the income maximization problem stated in Proposition 1 is such that $n(a) = 1$ for $a \leq 6 + s$ for some s . Then, given $h(6) = h_E$ and $q(6) = q_E$, the solution satisfies, for $a \in [6, 6 + s)$,*

$$x(a) = (h_E^{\gamma_1} q_E \gamma_2 z_h)^{\frac{1}{1-\gamma_2}} e^{\frac{r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}(a-6)}, \quad a \in [6, 6 + s) \quad (25)$$

and

$$h(a) = h_E e^{-\delta_h(a-6)} \left[1 + \left(h_E^{-(1-\gamma)} q_E^{\gamma_2} \gamma_2^{\gamma_2} z_h \right)^{\frac{1}{1-\gamma_2}} \frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r + \delta_h(1-\gamma_1)} \right. \\ \left. \left(e^{\frac{\gamma_2 r + \delta_h(1-\gamma_1)}{(1-\gamma_2)}(a-6)} - 1 \right) \right]^{\frac{1}{1-\gamma_1}}, \quad a \in [6, 6 + s) \quad (26)$$

Proof of Lemma 4. : From (21b) we obtain that

$$x(a) = (q(a)h(a)^{\gamma_1})^{\frac{1}{1-\gamma_2}} (\gamma_2 z_h)^{\frac{1}{1-\gamma_2}}. \quad (27)$$

Since we are in the region in which the solution is assumed to be at a corner, (21a) implies

$$h(a) \leq \left(\frac{\gamma_1}{w} \right)^{\frac{1-\gamma_2}{1-\gamma}} (\gamma_2^{\gamma_2} z_h)^{\frac{1}{1-\gamma}} q(a)^{\frac{1}{1-\gamma}} \quad (28)$$

In order to better characterize the solution we now show that the shadow value of the total product of human capital in the production of human capital grows at a constant rate. More precisely, we show that For $a \in [6, 6 + s)$, $q(a)h^{\gamma_1}(a) = q_E h_E^{\gamma_1} e^{[r+\delta_h(1-\gamma_1)](a-6)}$. To see this, let $M(a) = q(a)h^{\gamma_1}(a)$. Then,

$$\dot{M}(a) = M(a) \left[\frac{\dot{q}(a)}{q(a)} + \gamma_1 \frac{\dot{h}(a)}{h(a)} \right].$$

However, it follows from (21c) and (21d) after substituting (27) that

$$\begin{aligned} \frac{\dot{h}(a)}{h(a)} &= z_h h(a)^{\gamma_1-1} x(a)^{\gamma_2} - \delta_h, & a \in [6, 6 + s) \\ \frac{\dot{q}(a)}{q(a)} &= r + \delta_h - \gamma_2 z_h h(a)^{\gamma_1-1} x(a)^{\gamma_2}, & a \in [6, 6 + s). \end{aligned}$$

Thus,

$$\frac{\dot{q}(a)}{q(a)} + \gamma_1 \frac{\dot{h}(a)}{h(a)} = r + \delta_h(1 - \gamma_1).$$

The function $M(a)$ satisfies the first order ordinary differential equation

$$\dot{M}(a) = M(a)[r + \delta_h(1 - \gamma_1)]$$

whose solution is

$$M(a) = M(6) e^{[r+\delta_h(1-\gamma_1)](a-6)}$$

which establishes the desired result.

Using this result the level of expenditures during the schooling period is given by

$$x(a) = (h_E^{\gamma_1} q_E \gamma_2 z_h)^{\frac{1}{1-\gamma_2}} e^{\frac{r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}(a-6)}, \quad a \in [6, 6 + s).$$

Substituting this expression in the law of motion for $h(a)$ (equation (21d)), the equilibrium level of human capital satisfies the following first order non-linear, non-homogeneous, ordinary differential equation

$$\dot{h}(a) = (h_E^{\gamma_1 \gamma_2} q_E^{\gamma_2} \gamma_2^{\gamma_2} z_h)^{\frac{1}{1-\gamma_2}} e^{\frac{\gamma_2[r+\delta_h(1-\gamma_1)]}{(1-\gamma_2)}(a-6)} h^{\gamma_1}(a) - \delta_h h(a).$$

It can be verified, by direct differentiation, that (26) is a solution. ■

The next lemma describes the joint determination, given the age 6 level of human capital h_E , of the length of the schooling period, s , and the age 6 shadow price of human capital, q_E .

Lemma 5 *Given h_E , the optimal shadow price of human capital at age 6, q_E , and the length of the schooling period, s , are given by the solution to the following two equations*

$$q_E = \left[\frac{\gamma_1^{\gamma_1(1-\gamma_2)} \gamma_2^{\gamma_1\gamma_2} z_h^{\gamma_1} w^{(1-\gamma_1)(1-\gamma_2)}}{(r + \delta_h)^{(1-\gamma_2)}} \right]^{\frac{1}{1-\gamma}} h_E^{-\gamma_1} e^{-(r+\delta_h(1-\gamma_1))s} m(s+6)^{\frac{1-\gamma_2}{1-\gamma}}, \quad (29)$$

and

$$\begin{aligned} q_E^{\frac{\gamma_2}{1-\gamma_2}} h_E^{\frac{\gamma_1\gamma_2}{1-\gamma_2}} e^{-\delta_h(1-\gamma_1)s} & \left(\frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r + \delta_h(1-\gamma_1)} \right) (\gamma_2^{\gamma_2} z_h)^{\frac{1}{1-\gamma_2}} \\ & [e^{\frac{\gamma_2 r + \delta_h(1-\gamma_1)}{(1-\gamma_2)}s} - 1] + h_E^{1-\gamma_1} e^{-\delta_h(1-\gamma_1)s} \\ & = \left(\frac{\gamma_1^{(1-\gamma_2)} \gamma_2^{\gamma_2}}{(r + \delta_h)} \right)^{\frac{1-\gamma_1}{1-\gamma}} (z_h w^{\gamma_2})^{\frac{1-\gamma_1}{1-\gamma}} [m(s+6)]^{\frac{1-\gamma_1}{1-\gamma}}. \end{aligned} \quad (30)$$

Proof of Lemma 5. To prove this result, it is convenient to summarize some of the properties of the optimal path of human capital. For given values of (q_E, h_E, s) the optimal level of human capital satisfies

$$\begin{aligned} h(a) & = h_E e^{-\delta_h(a-6)} \left[1 + \left(h_E^{-(1-\gamma)} q_E^{\gamma_2} \gamma_2^{\gamma_2} z_h \right)^{\frac{1}{1-\gamma_2}} \frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r + \delta_h(1-\gamma_1)} \right. \\ & \left. \left(e^{\frac{\gamma_2 r + \delta_h(1-\gamma_1)}{(1-\gamma_2)}(a-6)} - 1 \right) \right]^{\frac{1}{1-\gamma_1}}, \quad a \in [6, 6+s] \end{aligned} \quad (31)$$

$$\begin{aligned} h(a) & = e^{-\delta_h(a-s-6)} \left\{ h(6+s) + \frac{C_h(z_h, w, r)}{\delta_h} e^{-\delta_h(6+s-R)} \right. \\ & \left. \int_{e^{\delta_h(6+s-R)}}^{e^{\delta_h(a-R)}} \left(1 - x^{\frac{r+\delta_h}{\delta_h}} \right)^{\frac{\gamma}{1-\gamma}} dx \right\}, \quad a \in [6+s, R]. \end{aligned} \quad (32)$$

Moreover, during at age $6+s$, (28) must hold at equality. Thus,

$$h(6+s) = \left(\frac{\gamma_1}{w} \right)^{\frac{1-\gamma_2}{1-\gamma}} (\gamma_2^{\gamma_2} z_h)^{\frac{1}{1-\gamma}} q(6+s)^{\frac{1}{1-\gamma}}.$$

Using the result in Lemma 4 in the previous equation, it follows that

$$q(6+s) = \frac{(h_E^{\gamma_1} q_E)^{\frac{1-\gamma}{1-\gamma_2}} e^{\frac{1-\gamma}{1-\gamma_2}(r+\delta_h(1-\gamma_1))(6+s)}}{\left(\frac{\gamma_1}{w}\right)^{\gamma_1} (\gamma_2^{\gamma_2} z_h)^{\frac{\gamma_1}{1-\gamma_2}}}. \quad (33)$$

Since

$$q(6+s) = \frac{w}{r+\delta_h} [1 - e^{-(r+\delta_h)(R-s-6)}],$$

it follows that

$$q_E = \left[\frac{\gamma_1^{\gamma_1(1-\gamma_2)} \gamma_2^{\gamma_1\gamma_2} z_h^{\gamma_1} w^{(1-\gamma_1)(1-\gamma_2)}}{(r+\delta_h)^{(1-\gamma_2)}} \right]^{\frac{1}{1-\gamma}} h_E^{-\gamma_1} e^{-(r+\delta_h(1-\gamma_1))s} m(s+6)^{\frac{1-\gamma_2}{1-\gamma}},$$

which is (29). Next, using (31) evaluated at $a = 6 + s$, and (28) at equality (and substituting out $q(6+s)$) using either one of the previous expressions we obtain (30).

■

We now discuss the optimal choice of h_E . Since q_E is the shadow price of an additional unit of human capital at age 6, the household chooses x_E to solve

$$\max q_E h_B x_E^v - x_E.$$

The solution is

$$h_E = v^{\frac{v}{1-v}} h_B^{\frac{1}{1-v}} q_E^{\frac{v}{1-v}}. \quad (34)$$

Proof of Proposition 2. Uniqueness of a solution to the income maximization problem follows from the fact that the objective function is linear and, given $\gamma < 1$, the constraint set is strictly convex. Even though existence can be established more generally, in what follows we construct the solution. To this end, we first describe the determination of years of schooling. Combining (29) and (30) it follows that

$$h_E = e^{\delta_h s} m(s+6)^{\frac{1}{1-\gamma}} (z_h w^{\gamma_2})^{\frac{1}{1-\gamma}} \left(\frac{\gamma_2^{\gamma_2} \gamma_1^{(1-\gamma_1)}}{r+\delta_h} \right)^{\frac{1}{1-\gamma}} \left[1 - \frac{r+\delta_h}{\gamma_1} \frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r + \delta_h(1-\gamma_1)} \frac{1 - e^{-\frac{\gamma_2 r + \delta_h(1-\gamma_1)}{(1-\gamma_2)} s}}{m(s+6)} \right]^{\frac{1}{1-\gamma_1}}. \quad (35)$$

Next, using (29) in (34), h_E must satisfy

$$h_E = h_B^{\frac{1}{1-v(1-\gamma_1)}} v^{\frac{v}{1-v(1-\gamma_1)}} \left(\frac{\gamma_1^{\gamma_1(1-\gamma_2)} \gamma_2^{\gamma_1\gamma_2}}{(r + \delta_h)^{1-\gamma_2}} \right)^{\frac{v}{(1-\gamma)(1-v(1-\gamma_1))}} \quad (36)$$

$$(z_h^{\gamma_1} w^{(1-\gamma_1)(1-\gamma_2)})^{\frac{v}{(1-\gamma)(1-v(1-\gamma_1))}} e^{-\frac{v(r+\delta_h(1-\gamma_1))}{1-v(1-\gamma_1)}s} m(s+6)^{\frac{v(1-\gamma_2)}{(1-\gamma)(1-v(1-\gamma_1))}}.$$

Finally, (35) and (36) imply that the number of years of schooling, s , satisfies

$$m(s+6)^{1-v(2-\gamma)} e^{(1-\gamma)(\delta_h+rv)s} \quad (37)$$

$$\left[1 - \frac{r + \delta_h}{\gamma_1} \frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r + \delta_h(1-\gamma_1)} \frac{1 - e^{-\frac{\gamma_2 r + \delta_h(1-\gamma_1)}{(1-\gamma_2)}s}}{m(s+6)} \right]^{\frac{(1-\gamma)(1-v(1-\gamma_1))}{1-\gamma_1}}$$

$$= \frac{h_B^{1-\gamma}}{z_h^{1-v} w^{\gamma_2-v(1-\gamma_1)}} \left(\frac{v}{r + \delta_h} \right)^{(1-\gamma)v} \left(\frac{\gamma_2^{\gamma_2} \gamma_1^{(1-\gamma_2)}}{r + \delta_h} \right)^{-(1-v)}.$$

As in the statement of the proposition, let the left hand side of (37) be labeled $F(s)$.

Then, an interior solution requires that $F(0) > 0$, or,

$$m(6)^{1-v(2-\gamma)} > \frac{h_B^{1-\gamma}}{z_h^{1-v} w^{\gamma_2-v(1-\gamma_1)}} \left(\frac{v}{r + \delta_h} \right)^{(1-\gamma)v} \left(\frac{\gamma_2^{\gamma_2} \gamma_1^{(1-\gamma_2)}}{r + \delta_h} \right)^{-(1-v)}.$$

Inspection of the function $F(s)$ shows that there exists a unique value of s , say \bar{s} , such that $F(s) > 0$, for $s < \bar{s}$, and $F(s) \leq 0$, for $s \geq \bar{s}$. It is clear that $\bar{s} < R - 6$. Hence, the function $F(s)$ must intersect the right hand side of (37) from above. The point of intersection is the unique value of s that solves the problem. ■

It is convenient to collect a full description of the solution as a function of aggregate variables and the level of schooling, s .

Solution to the Income Maximization Problem It follows from (21a), and the equilibrium values of the other endogenous variables, the time allocated to human capital formation is 1 for $a \in [6, 6 + s)$, and

$$n(a) = \frac{m(a)^{\frac{1}{1-\gamma}}}{e^{-\delta_h(a-s-6)} m(6+s)^{\frac{1}{1-\gamma}} + \frac{(r+\delta_h)e^{-\delta_h(a-R)}}{\gamma_1 \delta_h} \int_{e^{\delta_h(6+s-R)}}^{e^{\delta_h(a-R)}} (1-x)^{\frac{r+\delta_h}{\delta_h}} \frac{\gamma}{1-\gamma} dx}, \quad (38)$$

for $a \in [6 + s, R]$.

The amount of market goods allocated to the production of human capital is given by

$$x(a) = \left(\frac{\gamma_2 w}{r + \delta_h} \right) C_h(z_h, w, r) m(6 + s)^{\frac{1}{1-\gamma}} e^{\frac{r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}(a-s-6)}, \quad a \in [6, 6 + s), \quad (39)$$

$$x(a) = \left(\frac{\gamma_2 w}{r + \delta_h} \right) C_h(z_h, w, r) m(a)^{\frac{1}{1-\gamma}}, \quad a \in [6 + s, R). \quad (40)$$

The level of human capital of an individual of age a in the post-schooling period (i.e. $a \geq 6 + s$) is given by

$$h(a) = C_h(z_h, w, r) \left\{ e^{-\delta_h(a-s-6)} \frac{\gamma_1}{r + \delta_h} m(6 + s)^{\frac{1}{1-\gamma}} + \frac{e^{-\delta_h(a-R)}}{\delta_h} \int_{e^{\delta_h(6+s-R)}}^{e^{\delta_h(a-R)}} \left(1 - x^{\frac{r+\delta_h}{\delta_h}} \right)^{\frac{\gamma}{1-\gamma}} dx \right\}, \quad a \in [6 + s, R). \quad (41)$$

The stock of human capital at age 6, h_E , is

$$h_E = v^v h_B \left[\frac{\gamma_1^{\gamma_1(1-\gamma_2)} \gamma_2^{\gamma_1 \gamma_2} z_h^{\gamma_1} w^{(1-\gamma_1)(1-\gamma_2)}}{(r + \delta_h)^{(1-\gamma_2)}} \right]^{\frac{v}{1-\gamma}} e^{-v(r+\delta_h(1-\gamma_1))s} m(6 + s)^{\frac{v(1-\gamma_2)}{1-\gamma}} \quad (42)$$

References

- [1] Banerjee, A. V., and E. Duflo, 2004, "Growth Theory Through the Lens of Development Economics," MIT working paper.
- [2] Barro, Robert J. and Lee, Jong-Wha, 1996, "International Measures of Schooling Years and Schooling Quality," *American Economic Review*, May (Papers and Proceedings), 86(2), pp. 218-23.
- [3] Becker, Gary S. *Human Capital : A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: National Bureau of Economic Research, 1964.
- [4] Belzil, Christian & Hansen, Jörgen, 2002, "Unobserved Ability and the Return to Schooling," *Econometrica*, vol. 70(5), pp. 2075-2091.
- [5] Betts, Julian and Magnus Lofstrom, 2000, "The Educational Attainment of Immigrants: Trends and Implications," in G. J. Borjas (ed) **Issues in the Economics of Immigration**, University of Chicago Press for the NBER, Chicago.
- [6] Bils, M. and P. Klenow. 2000. "Does Schooling Cause Growth?" *American Economic Review* 90(5):1160-83
- [7] Ben Porath, Y., 1967, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy* 75, pt. 1, 352-365.
- [8] Borjas, George J., 1994, "The Economics of Immigration," *Journal of Economic Literature*, Vol. XXXII, pp: 1667-1717.
- [9] Borjas, George J., 2000, "The Economic Progress of Immigrants," in G. J. Borjas (ed) **Issues in the Economics of Immigration**, University of Chicago Press for the NBER, Chicago.

- [10] Card, D., 2000, "The Causal Effects of Education on Earnings," in D. Card and O. Ashenfleter (eds.) **Handbook of Labor Economics**, North-Holland.
- [11] Caselli, F., 2003, "The Missing Input: Accounting for Cross-Country Income Differences," forthcoming in Aghion P. and S. Durlauf (eds) **Handbook of Economic Growth**, North Holland publishers
- [12] Chari, V.V., P.J. Kehoe and E.R. McGrattan, 1997, "The Poverty of Nations: A Quantitative Investigation," Research department Staff Report 204, Federal Reserve Bank of Minneapolis.
- [13] Denison, Edward F., 1967, *Why Growth Rates Differ: Postwar Experience in Nine Western Countries* (Washington, DC: The Brookings Institution).
- [14] Galor, Oded, 2005, "From Stagnation to Growth: Unified Growth Theory," forthcoming in P. Aghion and S. Durlauf (eds) **Handbook of Economic Growth**, North Holland.
- [15] Haley, W., 1976, "Estimation of the Earnings Profile from Optimal Human Capital Accumulation," *Econometrica*, 44, 1223-38.
- [16] Hall, Robert E., and Jones, C. I., 1999, "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, February 1999, vol. 114(1), pages 83-116.
- [17] Hendricks, Lutz, 2002, "How Important is Human Capital for Development? Evidence from Immigrant Earnings," *American Economic Review*, Vol. 92, No. 1, pp: 198-219.
- [18] Heckman, J., L. Lochner, and C. 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.

- [19] Heckman, J., L. Lochner, and P. Todd, 2003, "Fifty Years of Mincer Earnings Regressions," NBER Working Paper 9732.
- [20] Hsieh, Chang-Tai and P. Klenow, 2003, "Relative Prices and Relative Prosperity," NBER Working Paper 9701.
- [21] Jasso, G. and M. Rosenzweig, 2002, "The Earnings of U.S. Immigrants: World Skill Prices, Skill Transferability and Selectivity," working paper.
- [22] Jeong, H. and R. Townsend, 2004, "Discovering Sources of TFP Growth: Occupation Choice, Capital Heterogeneity, and Financial Deepening," working paper.
- [23] Jovanovic, B., 1979, "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87, 972-990
- [24] Kendrick, John W., 1961. **Productivity Trends in the United States**, Princeton NJ, Princeton University Press.
- [25] Klenow, Peter and Rodríguez-Clare, Andrés, 1997, "The Neoclassical Revival in Growth Economics: Has it Gone Too Far?" in Ben Bernanke and Julio Rotemberg, eds., **Macroeconomics Annual 1997**. Cambridge, MA: MIT Press, 1997, pp. 73-102.
- [26] Lucas, Robert, 1988, "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 3-42
- [27] Mankiw, N. Gregory; Romer, David and Weil, David N. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics*, May 1992, 107(2), pp. 407-37.
- [28] Manuelli, R. and A. Seshadri, 2004, "Schooling and Human Capital: Theory and Implications," manuscript.

- [29] Mincer, J., 1974, **Schooling, Experience and Earnings**, Columbia University Press, New York, 1974.
- [30] Neal, D. and S. Rosen (1999), "Theories of the Distribution of Earnings," in **Handbook of Income Distribution**, ed. A. Atkinson and F. Bourguignon, North Holland Publishers.
- [31] Parente, S.L. and E.C. Prescott. (2000). **Barriers to Riches**, Cambridge, MA: MIT Press.
- [32] OECD, 1996, **Education at a Glance: OECD Indicators**.
- [33] Prescott, Edward C., 1998, "Needed: A Theory of Total Factor Productivity," *International Economic Review*, 39(3), pp. 525-51.
- [34] Solow, Robert M., 1956, "A contribution to the theory of economic growth", *Quarterly Journal of Economics*, Vol.70 No.1, pp: 65 - 94.