Human Capital and the Wealth of Nations

Rodolfo E. Manuelli and Ananth Seshadri*
Department of Economics
University of Wisconsin-Madison

April 2007

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 and reevaluate the role of human capital. The key difference between our paper and recent work in this area is that we use theory to estimate the stocks of human capital, and that we allow the *quality* of human capital to vary across countries. When quality differences are allowed, we find that effective human capital per worker varies substantially across countries.

As a result of this finding, we estimate that cross-country differences in Total Factor Productivity (TFP) are significantly smaller than those reported in previous studies. Moreover, our model implies that output per worker is highly responsive to differences in TFP and in demographic variables.

---

*We thank Gary Becker, Bart Hobijn, Chad Jones, Larry Jones, John Kennan, Pete Klenow, Bob Lucas, Kevin Murphy and Richard Rogerson for comments. We are grateful to the NSF and Wisconsin Alumni Research Foundation for generous 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), Parente and Prescott (2000) and Bils and Klenow (2000a) 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 income 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. We restrict our analysis to steady states. According to the model, 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 and by the cross-country differences in demographic structure. 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 that 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 (holding demographic variables constant). We find that the resulting elasticity is fairly large: a 1% increase in (relative) TFP results in a 6.5% (long run) increase in (relative) output per worker. The second exercise is designed to evaluate the contribution of demographic characteristics to underdevelopment.[1] In the model, countries differ in terms of life expectancy, retirement age and fertility. We conduct the following counterfactual experiment: we 'endow' each country with the demographic characteristics of the U.S. Then, we let individuals adjust their choices of physical and human capital. We find that this demographic change doubles the

---

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

level of output in the poor country.

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. Thus, the Mincerian equation can be viewed as an equilibrium relationship between two endogenous variables (schooling and earnings) in the Ben-Porath model.

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 relative 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.

This research is related to the recent analysis of the effect of human capital on cross-country income differences. As such, it provides an alternative way of computing human capital to that advanced by Klenow and Rodriguez-Clare (1997) and Bils and Klenow (2000a). The main difference is in the use of Mincer based estimate of human capital stocks (taking schooling as exogenous) vs a Ben-Porath based measure. The papers closest to ours is Erosa, Koreshkova and Restuccia (2007). From the point of view of the computation of the stock of human capital —and the effect of changes in TFP— the key difference is that they assume that post-schooling human capital is independent of economic forces. This lowers one channel through which wages affect human capital accumulation and this, in turn, results in a lowers estimated elasticity of TFP —2.8 versus 6.5.

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. Section 7 presents some concluding comments.

2 The Model

In this section we describe the basic model, characterize its solution, and compute the implications for output per worker using the exogenously specified demographic structure.

2.1 The Individual's Problem

The representative individual maximizes the present discounted value of net income. We assume that each agent lives for T periods and retires at age $R \leq T$. The maximization problem is

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

subject to

$$\dot{h}(a) = z_h[n(a)h(a)]^{\gamma_1}x(a)^{\gamma_2} - \delta_h h(a), \qquad a \in [6, R), \tag{2}$$

and

$$h(6) = h_E = h_B x_E^{\upsilon} \tag{3}$$

with h_B given.[2] Equations (2) and (3) correspond to the standard human capital accumulation model initially developed by Ben-Porath (1967). This formulation allows for both market goods, $x(a)$, and a fraction $n(a)$ of the individual's human capital, to be inputs in the production of human capital.[3] Investments in early childhood[4], which we denote by x_E (e.g. medical care, nutrition and development of learning

---

[2] The assumption of linear utility is without loss of generality. It can be shown that the solution to the income maximization problem is also the solution to a utility maximization problem when the number of children is given, parents have a bequest motive, and bequests are unconstrained. For details, see Manuelli and Seshadri (2005).

[3] In particular, we assume that there are no external effects and attempt to see how far this takes us.

[4] For a review of the evidence on the effects of early interventions on future outcomes, see Karoly, Kilburn and Cannon (2005).

skills), determine the level of each individual's human capital at age 6, $h(6)$, or h_E for short.[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.[6]

There are two important features of our formulation. First, we assume that the human capital accumulation technology 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 —$x(a)$— 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.[7] In fact, for our argument to go through, it suffices that, at the margin, individuals pay for the last unit of market goods allocated to the formation of human capital.

The full solution to the income maximization problem, which to our knowledge is novel, is presented in the Appendix. The solution to the problem is such that $n(a) = 1$, for $a \leq 6 + s$. Thus, we identify s as *years of schooling*. The following proposition characterizes s and the level of human capital at the end of the schooling period, $h(6 + s)$.

Proposition 1 *There exists a unique solution to the income maximization problem. The number of years of schooling, s, satisfies*

---

[5] It should be made clear that market goods ($x(a)$ and x_E) are produced using the same technology as the final goods production function. Hence, the production function for human capital is more labor intensive than the final goods technology.

[6] It is clear that parents' time is also important. However, given exogenous fertility, it seems best to ignore this dimension. For a full discussion see Manuelli and Seshadri (2005).

[7] An 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.

1.

$$F(s) = \frac{h_B^{1-\gamma}}{z_h^{1-\upsilon} w^{\gamma_2 - \upsilon(1-\gamma_1)}}, \tag{4}$$

where

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

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

and

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

provided that

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

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

$$h(s+6) = \left[\frac{\gamma_2^{\gamma_2}\gamma_1^{\gamma_1} z_h w^{\gamma_2}}{(r+\delta_h)^{\gamma}}\right]^{\frac{1}{1-\gamma}} \frac{\gamma_1}{r+\delta_h} m(6+s)^{\frac{1}{1-\gamma}} \tag{5}$$

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. $\gamma_2 = \upsilon = 0$), the model implies that changes in wage rates have *no impact* on schooling decisions. (See equation

(4)). 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.[8] 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 υ is sufficiently high (and $\gamma_2 - \upsilon(1-\gamma_1) < 0$), 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 υ 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 υ. In our parameterization $\gamma_2 - \upsilon(1-\gamma_1) > 0$, 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 choose to acquire.

2. **Development and Schooling Quality.** The total impact of changes in wages (or TFP) on the stock of human capital at the end of schooling is given by totally differentiating (5).

$$\frac{dh(s+6)}{dw} = \frac{\partial h(s+6)}{\partial s}\frac{ds}{dw} + \frac{\partial h(s+6)}{\partial w}.$$

The first term on the right hand side can be interpreted as the effect of changes in the wage rate on the *quantity* of human capital (years of schooling), while

---

[8] It 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.

the second term captures the impact on the level of human capital per year of schooling, a measure of *quality*. Direct calculations (see equation (5)) show that the elasticity of quality with respect to the wage rate is $\gamma_2/(1-\gamma)$, which is fairly large in our preferred parameterization.[9] 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 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.[10]

3. **Individual Characteristics, Schooling and Human Capital.** Individuals with higher ability —as measured by z_h— choose longer schooling periods. At the other end, high levels of initial human capital, h_B, result in lower schooling. The model implies that even at age 6, there are differences in human capital among identical children that live in different countries due to differences in the cost (measured in wage units) of market goods needed to build early childhood capital. (See equation (32) 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

---

[9] To 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.

[10] It can be shown that the elasticity of quality with respect to TFP is $\gamma_2/[(1-\theta)(1-\gamma)]$, where θ is capital share.

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 (31) in the Appendix), measured income as a function of experience, defined as $p = a - s - 6$, and schooling, s, is

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

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 theoretical version of the equilibrium relationship between schooling, experience and earnings. Since schooling is endogenous, the relationship cannot be viewed as a (nonlinear) regression, even if one were willing to tack on an error term. In order to interpret the model's predictions about education and earnings it is necessary to be explicit about the factors that **induce** different individuals to choose different levels of s.

2.2 Equilibrium

Given the interest rate, standard profit maximization pins down the equilibrium capital-human capital ratio. To determine output per worker, it is necessary to compute 'average' human capital in this economy. Given that we are dealing with finite lifetimes —and full depreciation of human capital at death— there is no aggregate version of the law of motion of human capital since, as the previous derivations show, the amount of human capital supplied to the market depends on an individual's age. Thus, to compute average 'effective' human capital we need to determine the age structure of the population.

Demographics We assume that each individual has e^f children at age B. Since we consider only steady states, we need to derive the stationary age distribution of this economy associated with this fertility rate. Our assumptions imply

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

and

$$N(t',t) = 0, \qquad t' > T.$$

It is easy to check that in the steady state

$$N(a,t) = \phi(a)e^{\eta t}, \tag{6}$$

where

$$\phi(a) = \eta \frac{e^{-\eta a}}{1 - e^{-\eta T}}, \tag{7}$$

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

Aggregation To compute output per worker it suffices to estimate the per capita aggregate amount of human capital effectively supplied to the market, and the physical capital-human capital ratio. The average amount of human capital per worker allocated to market production, $\bar{h}^e$, is given by

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

This formulation shows that, even if R —the retirement age— is constant, changes in the fertility rate and life expectancy, η and T respectively, can have an impact on the average stock of human capital.

Equilibrium Optimization on the part of firms implies that

$$p_k(r + \delta_k) = zF_k(\kappa, 1), \tag{8}$$

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

$$w = zF_h(\kappa, 1). \tag{9}$$

Then, output per worker is

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

3 Calibration

We use standard functional forms. The production function is assumed to be Cobb-Douglas

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

Our calibration strategy involves choosing the parameters so that the steady state implications of the model economy presented above are consistent with observations for the United States (circa 2000). We then vary the exogenous demographic variables in accordance with the data, and we 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 TFP for other countries is chosen so as to match output per worker by construction, the predictions of the model about years of schooling and the amount of goods inputs used in the production of human capital can be compared with the evidence.

Following Cooley and Prescott (1995), 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 several 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 $\pi = 0.5$.[11] We also assume that the same fraction π is not measured in GDP.

Our theory implies that it is only the ratio $h_B^{1-\gamma}/(z_h^{1-\upsilon} w^{\gamma_2-\upsilon(1-\gamma_1)})$ that matters for all the moments of interest. Since we assume that h_B (initial human capital) and z_h (ability to learn) are common across all coutnries, we can choose z, p_k (which determines w) and h_B arbitrarily and calibrate z_h to match a desired moment. We set $B = 25$ and $R = \min\{64, T\}$. This leaves us with 7 parameters, $\theta, r, \delta_h, z_h, \gamma_1, \gamma_2$ and υ. 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 as a fraction of GDP of 3.77. Source: OECD, Education at a Glance.
7. Pre-primary expenditures per pupil relative to GDP per capita of 0.14. Source: OECD, Education at a Glance.

---

[11] If we were to take the view that $\pi = 1$, our estimate of the returns to scale, $\gamma = \gamma_1 + \gamma_2$ increases to 0.96 thereby further increasing the elasticity of output with respect to TFP. In a sense, choosing $\pi = 0.5$ understates our case.

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	r	δ_h	z_h	γ_1	γ_2	ν
Value	0.315	0.07	0.018	0.361	0.63	0.3	0.55

Our estimate of the degree of returns to scale ,γ ($\gamma = \gamma_1 + \gamma_2$), lies in the range reported by Browning, Hansen and Heckman (1999). More recent estimates by Heckman, Lochner and Taber (1998) and Kuruscu (2006) are around 0.93. Some of the earlier estimates used substantially higher real interest rates (20% and higher). Such a high real interest explains their finding of a significantly smaller estimate of returns to scale. Furthermore, Haley (1976) arbitrarily restricts a key parameter to take on a value of 0.5 (which is effectively $h_B^{1-\gamma}/z_h$ in the context of our model) in his work and also discusses at length the issues of identification. This explains his lower estimate of γ

Finally, even though we did not seek to match the entire profile of earnings by age, our calibrated parameters come very close to mimicking the actual data as can be seen from Figure 1. In that figure we also show the impact of changing γ (keeping the ratio of γ_1 to γ_2 unchanged) on the predictions of the model for the ratio of earnings between ages 25 and 50. Low values of γ (i.e. $\gamma = 0.50$) imply too steep an age earnings profile, while values closer to the constant returns to scale model (i. e. $\gamma = 0.99$) result in a decreasing age earnings profile.[12] The intuition for these results parallels the findings in growth models: In the case of (nearly) constant returns to scale the optimal path requires slow adjustment and, given the finite lifetime, it implies a decreasing age-earnings profile. In this case, earnings at age 50 are, counterfactually, below earnings at age 25. If the production function for human

---

[12] If one fixes $\gamma_2 = 0.33$ and changes γ_1 the results are similar. If $\gamma_1 = .55$, the earnings ratio is 4.03, while if it is set to .65 the earnings ratio decreases to 0.72.

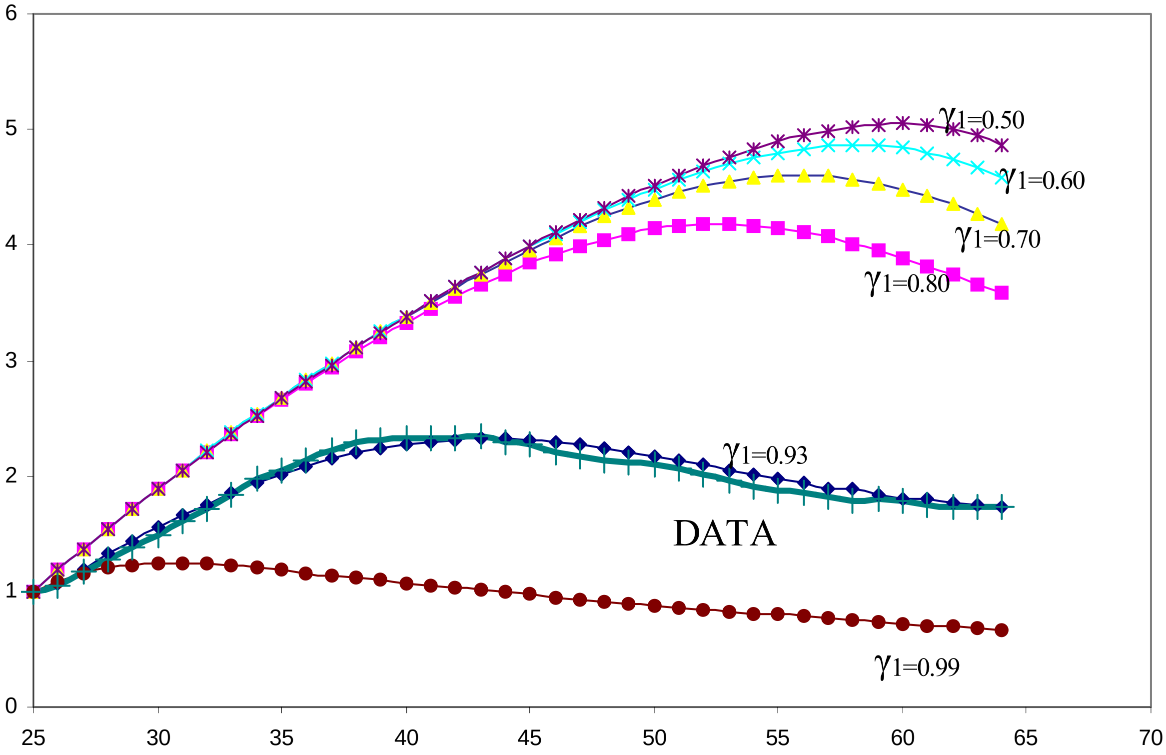

Figure 1: Effect of changing γ on the shape of the age earnings profile

capital displays returns to scale that are lower than the calibrated values, the model predicts a high ratio of earnings at age 50 relative to earnings at age 25.

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 on education relative to GDP (x_s), life expectancy (T), total fertility rate ($e^f/2$), and the relative price of capital (p_k) for each of these deciles. The population values are displayed in the following table.

Table 2: World Distribution

Decile	y (relative to US)	s	x_s	T	$e^f/2$ (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 2 illustrates the wide disparities in incomes across countries. The United States possesses an output per worker (normalized to one) that is about 20 times as high as the countries in the bottom decile. Years of schooling also vary 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 some of the market inputs that are used in the educational process, and it excludes expenses paid by parents (including the time and resources that parents invest in their kids). 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, years of schooling, and measures of spending in education. 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) and T (life expectancy) across countries (according to the data) and choose the level of TFP in a particular country so as to match output per worker. Table 3 presents the predictions of the model and the data.

Table 3: 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

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 25%. Thus, our estimate of TFP in the poorest countries is more than two times higher.

The model does fairly well matching the two variables that it predicts: schooling and expenditures in formal education. The results are in Table 3 in the columns labeled s and x_s. The predictions for schooling are close to the data although they tend to underpredict educational attainment for the poorer set of countries. In terms of a rough measure of quality such as schooling expenditures as a fraction of output, the model actually underpredicts investment at the two ends of the world income distribution.[13] Thus, this cannot explain our findings.[14]

We used the model to compute the elasticity of output with respect to TFP when all endogenous variables are allowed to reach their new steady state (this is the very long run). We estimate this elasticity to be around 6.5. Thus, according to the model, changes in TFP have a large multiplier effect on output per worker.[15]

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

---

[13] The model overpredicts x_s for countries in the middle of the distribution.

[14] As 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 school expenditures. Moreover, it is not clear to what extent capital costs are included in this measure.

[15] The elasticity that can be inferred from Table 3 is much higher, around 9.4. The reason is that those values reflect changes in TFP and demographic variables.

in fertility and life expectancy 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 argue next, differences in demographics play a significant role.

Changing Demographics 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. 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.

From a methodological point of view, ignoring the age structure of the population would have resulted in a much smaller estimated impact of TFP on output per worker. To be precise, had we assumed that every country has the same demographic structure as the U.S., we would have found that the level of TFP required to match (relative) output per worker in the lowest decile is 65% of the U.S. (instead of 73%).

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. In ongoing work, we study the impact that changes in TFP have upon (endogenously chosen) fertility.[16]

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)

---

[16] In related research (Manuelli and Seshadri (2007)) we study a version of the model with endogenous fertility and find that the basic results do not change much.

we now allow p_k to vary according to the values in Table 2. Table 5 presents the results.

Table 5: 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.

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.

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 Minncerian framework assumes 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 6 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 suggests that ignoring this dimension can induce significant biases in the estimates of human capital.[17]

---

[17] In a recent paper, Caselli (2003) explicitly models, in a reduced form sense, differences in C_i

The Importance of Early childhood and On-the-Job Training (OJT) Our model implies that, even at age 6, there are substantial differences between the human capital of the average child in rich and poor countries. In Table 6 we present the values of human capital at age 6 (h_E) and aggregate human capital per worker ($\bar{h}^e$) for each decile relative to the U.S. 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.

The differences in stocks of human capital produced by our model is a result of investments undertaken over the three phases - early childhood, schooling and job training. It is only natural to further investigate the importance of each of these channels in contributing to human capital differences. One possible way to arrive at the contribution of each of the three phases is

$$1 = \underbrace{\frac{\bar{h}^e - h(6+s)}{\bar{h}^e}}_{\text{OJT}} + \underbrace{\frac{h(6+s) - h_E}{\bar{h}^e}}_{\text{Schooling}} + \underbrace{\frac{h_E}{\bar{h}^e}}_{\text{Early Childhood}} .$$

Recall that $h(6+s)$ is the stock of human capital that an individual possesses at the age at which he leaves school (see equation (5)). The last 3 columns of Table 6 present the results. Notice that while on the job training and schooling are the dominant contributors in the top deciles, early childhood contributes a lot more to the bottom deciles. This transpires mainly because in poorer nations, children constitute a significant part of the work force. Since a large fraction of the working population is young, this large mass contributes a lot more to human capital per worker differences than in a richer country where the population distribution is close to uniform.

---

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.

Table 6: Understanding Human Capital Differences

	Relative to U.S.			Contribution (Shares)		
Decile	y	h_E	$\bar{h}$	OJT	Schooling	Early Childhood
90-100	0.921	0.96	0.95	0.40	0.50	0.10
80-90	0.852	0.91	0.88	0.40	0.50	0.10
70-80	0.756	0.88	0.79	0.38	0.51	0.11
60-70	0.660	0.86	0.71	0.41	0.48	0.12
50-60	0.537	0.79	0.60	0.39	0.48	0.13
40-50	0.437	0.72	0.50	0.43	0.43	0.14
30-40	0.354	0.65	0.43	0.42	0.44	0.15
20-30	0.244	0.60	0.32	0.37	0.45	0.19
10-20	0.146	0.53	0.20	0.34	0.43	0.23
0-10	0.052	0.47	0.08	0.32	0.40	0.28

6 Implications for Mincer Regressions

Even though the interpretation and the precise point estimate of the schooling coefficient in a Mincer regression is controversial, most estimates —at least when linearity is imposed— seem to be close to 10%.[18] Thus, one challenge for the model economy is to reproduce the rate of return in a Mincer-style regression.

Since the model predicts that all (homogeneous) individuals choose exactly the same level of schooling, it is necessary to introduce some source of heterogeneity to match the observed differences in schooling. The two natural candidates are differences in z_h (ability to learn), and differences in h_B (initial human capital). From the results in Proposition 1 it follows that the equilibrium years of schooling depend on the ratio $h_B^{1-\gamma}/\left(z_h^{1-\upsilon}w^{\gamma_2-\upsilon(1-\gamma_1)}\right)$. Since in a given country all individuals face the

---

[18] The assumption that the relationship between log earnings and schooling is linear is also controversial. Belzil and Hansen (2002) find that, when the return is allowed to be a sequence of spline functions, the relationship is convex.

same wage and interest rate, 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.

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, 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.

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.

7 Conclusion

In this paper, we show that an extended neoclassical convex model that incorporates a human capital sector is capable of generating large differences in the stocks of human capital driven by small differences in TFP. Our results show that human capital has a central role in determining the wealth of nations. The novelty is that the framework that we use 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.[19] We find that, the elasticity of output per worker with respect to TFP is slightly over 9. 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.[20]

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 top decile, output per worker would double.

Naturally, the consideration of capital market imperfections such as binding intergenerational 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

---

[19] Very few papers attempt to explain schooling differences across countries - a notable exception is Bils and Klenow (2000b).

[20] The model is also consistent with the evidence on earnings of immigrants in the U.S. Some preliminary results are in Manuelli and Seshadri (2007b)

impact of a variety of human capital policies in the presence of distortions, as well as the role of endogenously chosen fertility.

8 Appendix

The first order conditions of the income maximization problem given the stock of human capital at age 6, $h(6) = h_E$ are,

$$\begin{aligned} whn &\leq q\gamma_1 z_h (nh)^{\gamma_1} x^{\gamma_2}, \qquad \textit{with equality if } n < 1, && \text{(10a)} \\ x &= q\gamma_2 z_h (nh)^{\gamma_1} x^{\gamma_2}, && \text{(10b)} \\ \dot{q} &= rq - [q\gamma_1 z_h (nh)^{\gamma_1} x^{\gamma_2} h^{-1} - \delta_h] - w(1-n), && \text{(10c)} \\ \dot{h} &= z_h (nh)^{\gamma_1} x^{\gamma_2} - \delta_h h, && \text{(10d)} \end{aligned}$$

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

For simplicity, we prove a series of lemmas that simplify the proof of Proposition 1. 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 2 *Assume that the solution to the income maximization problem is such that $n(a) = 1$ for $a \leq 6 + s$ for some s. Then, given $h(6 + s)$ the solution satisfies, for $a \in [6 + s, R)$,*

$$x(a) = \left(\frac{\gamma_2 w}{r + \delta_h} \right) C_h(z_h, w, r) \left[1 - e^{-(r+\delta_h)(R-a)} \right]^{\frac{1}{1-\gamma}}, \qquad a \in [6 + s, R), \tag{11}$$

$$\begin{aligned} h(a) &= e^{-\delta_h(a-6-s)} \{ h(6+s) + \frac{C_h(z_h, w, r)}{\delta_h} e^{-\delta_h(6+s-R)} \\ &\quad \int_{e^{\delta_h(6+s-R)}}^{e^{\delta_h(a-R)}} (1 - x^{\frac{r+\delta_h}{\delta_h}})^{\frac{\gamma}{1-\gamma}} dx \}, \qquad a \in [6+s, R), \end{aligned} \tag{12}$$

and

$$q(a) = \frac{w}{r+\delta_h}[1 - e^{-(r+\delta_h)(R-a)}], \qquad a \in [6+s, R). \tag{13}$$

Proof of Lemma 2. : Given that the equations (10) hold (with the first equation at equality), standard algebra (see Ben-Porath, 1967 and Haley, 1976) shows that (13) holds. Using this result in (10b) it follows that

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

which is (11). Next substituting (11) and (13) into (10d) one obtains a non-linear non-homogeneous first order ordinary differential equation. Straightforward, but tedious, algebra shows that (12) is a solution to this equation. ■

The next lemma describes the solution during the schooling period.

Lemma 3 *Assume that the solution to the income maximization problem 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 $a \in [6, 6+s)$,*

$$x(a) = (h_E^{\gamma_1} q_E \gamma_2 z_h)^{\frac{1}{1-\gamma_2}}\, e^{\frac{r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}(a-6)}, \qquad a \in [6, 6+s) \tag{14}$$

and

$$\begin{aligned} h(a) \;=\; & h_E e^{-\delta_h(a-6)}[1 + \left(h_E^{-(1-\gamma)} q_E^{\gamma_2} \gamma_2^{\gamma_2} z_h\right)^{\frac{1}{1-\gamma_2}} \frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r + \delta_h(1-\gamma_1)} \\ & (e^{\frac{\gamma_2 r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}(a-6)} - 1)]^{\frac{1}{1-\gamma_1}}, \qquad a \in [6, 6+s) \end{aligned} \tag{15}$$

Proof of Lemma 3. : From (10b) we obtain that

$$x(a) = (q(a)h(a)^{\gamma_1})^{\frac{1}{1-\gamma_2}} (\gamma_2 z_h)^{\frac{1}{1-\gamma_2}}. \tag{16}$$

Since we are in the region in which the solution is assumed to be at a corner, (10a) implies

$$h(a) \leq \left(\frac{\gamma_1}{w}\right)^{\frac{1-\gamma_2}{1-\gamma}} (\gamma_2^{\gamma_2} z_h)^{\frac{1}{1-\gamma}} q(a)^{\frac{1}{1-\gamma}} \tag{17}$$

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 $a \in [6, 6+s)$, $q(a)h(a)^{\gamma_1} = q_E h_E^{\gamma_1} e^{[r+\delta_h(1-\gamma_1)](a-6)}$. To see this, let $M(a) = q(a)h(a)^{\gamma_1}$. Then,

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

However, it follows from (10c) and (10d) after substituting (16) that

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

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

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

Substituting this expression in the law of motion for $h(a)$ (equation (10d), the equilibrium level of human capital satisfies the following first order non-linear, non-homogeneous, ordinary differential equation

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

It can be verified, by direct differentiation, that (15) 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 4 *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*

$$\begin{aligned} q_E &= \left[\frac{\gamma_1^{\gamma_1(1-\gamma_2)}\gamma_2^{\gamma_1\gamma_2}z_h^{\gamma_1}w^{(1-\gamma_1)(1-\gamma_2)}}{(r+\delta_h)^{(1-\gamma_2)}}\right]^{\frac{1}{1-\gamma}} h_E^{-\gamma_1} \\ &\quad e^{-(r+\delta_h(1-\gamma_1))s}m(s+6)^{\frac{1-\gamma_2}{1-\gamma}}, \end{aligned} \tag{18}$$

and

$$\begin{aligned} & q_E^{\frac{\gamma_2}{1-\gamma_2}} h_E^{\frac{\gamma_1\gamma_2}{1-\gamma_2}} e^{-\delta_h(1-\gamma_1)s}\left(\frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r+\delta_h(1-\gamma_1)}\right)(\gamma_2^{\gamma_2}z_h)^{\frac{1}{1-\gamma_2}} \\ & [e^{\frac{\gamma_2 r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}s}-1] + h_E^{1-\gamma_1}e^{-\delta_h(1-\gamma_1)s} \\ = & \left(\frac{\gamma_1^{(1-\gamma_2)}\gamma_2^{\gamma_2}}{(r+\delta_h)}\right)^{\frac{1-\gamma_1}{1-\gamma}} (z_h w^{\gamma_2})^{\frac{1-\gamma_1}{1-\gamma}}\,[m(s+6)]^{\frac{1-\gamma_1}{1-\gamma}}. \end{aligned} \tag{19}$$

Proof of Lemma 4. 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

$$\begin{aligned} h(a) &= h_E e^{-\delta_h(a-6)}[1+\left(h_E^{-(1-\gamma)}q_E^{\gamma_2}\gamma_2^{\gamma_2}z_h\right)^{\frac{1}{1-\gamma_2}}\frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r+\delta_h(1-\gamma_1} \\ &\quad (e^{\frac{\gamma_2 r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}(a-6)}-1)]^{\frac{1}{1-\gamma_1}},\ a\in[6,6+s) \end{aligned} \tag{20}$$

$$\begin{aligned} h(a) &= e^{-\delta_h(a-s-6)}\{h(6+s)+\frac{C_h(z_h,w,r)}{\delta_h}e^{-\delta_h(6+s-R)} \\ &\quad \int_{e^{\delta_h(6+s-R)}}^{e^{\delta_h(a-R)}}(1-x^{\frac{r+\delta_h}{\delta_h}})^{\frac{\gamma}{1-\gamma}}dx\}, \qquad a\in[6+s,R). \end{aligned} \tag{21}$$

Moreover, at age $6+s$, (17) 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 3 in the previous equation, it follows that

$$q(6+s) = \frac{\left(h_E^{\gamma_1} q_E\right)^{\frac{1-\gamma}{1-\gamma_2}} e^{\frac{1-\gamma}{1-\gamma_2}(r+\delta_h(1-\gamma_1))(6+s)}}{\left(\frac{\gamma_1}{w}\right)^{\gamma_1} \left(\gamma_2^{\gamma_2} z_h\right)^{\frac{\gamma_1}{1-\gamma_2}}}. \tag{22}$$

Since

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

it follows that

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

which is (18). Next, using (20) evaluated at $a = 6+s$, and (17) at equality (and substituting out $q(6+s)$) using either one of the previous expressions we obtain (19). ■

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

$$\max q_E h_B x_E^{\upsilon} - x_E.$$

The solution is

$$h_E = \upsilon^{\frac{\upsilon}{1-\upsilon}} h_B^{\frac{1}{1-\upsilon}} q_E^{\frac{\upsilon}{1-\upsilon}}. \tag{23}$$

Proof of Proposition 1. 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 (18) and (19) it follows that

$$\begin{aligned} h_E \quad = \quad & e^{\delta_h s} m(s+6)^{\frac{1}{1-\gamma}} \left(z_h w^{\gamma_2}\right)^{\frac{1}{1-\gamma}} \left(\frac{\gamma_2^{\gamma_2} \gamma_1^{(1-\gamma_1)}}{r+\delta_h}\right)^{\frac{1}{1-\gamma}} \\ & \left[1 - \frac{r+\delta_h}{\gamma_1} \frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r + \delta_h(1-\gamma_1)} \frac{1 - e^{-\frac{\gamma_2 r + \delta_h(1-\gamma_1)}{(1-\gamma_2)} s}}{m(s+6)}\right]^{\frac{1}{1-\gamma_1}}. \end{aligned} \tag{24}$$

Next, using (18) in (23), h_E must satisfy

$$\begin{aligned} h_E &= h_B^{\frac{1}{1-\upsilon(1-\gamma_1)}} \upsilon^{\frac{\upsilon}{1-\upsilon(1-\gamma_1)}} \left(\frac{\gamma_1^{\gamma_1(1-\gamma_2)} \gamma_2^{\gamma_1\gamma_2}}{(r+\delta_h)^{1-\gamma_2}} \right)^{\frac{\upsilon}{(1-\gamma)(1-\upsilon(1-\gamma_1))}} \qquad (25) \\ & \quad \left(z_h^{\gamma_1} w^{(1-\gamma_1)(1-\gamma_2)} \right)^{\frac{\upsilon}{(1-\gamma)(1-\upsilon(1-\gamma_1))}} e^{-\frac{\upsilon(r+\delta_h(1-\gamma_1))}{1-\upsilon(1-\gamma_1)} s} m(s+6)^{\frac{\upsilon(1-\gamma_2)}{(1-\gamma)(1-\upsilon(1-\gamma_1))}}. \end{aligned}$$

Finally, (24) and (25) imply that the number of years of schooling, s, satisfies

$$\begin{aligned} & m(s+6)^{1-\upsilon(2-\gamma)} e^{(1-\gamma)(\delta_h + r\upsilon)s} \qquad (26) \\ & \left[1 - \frac{r+\delta_h}{\gamma_1} \frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r + \delta_h(1-\gamma_1)} \frac{1 - e^{-\frac{\gamma_2 r + \delta_h(1-\gamma_1)}{(1-\gamma_2)} s}}{m(s+6)} \right]^{\frac{(1-\gamma)(1-\upsilon(1-\gamma_1))}{1-\gamma_1}} \\ & \left(\frac{\upsilon}{r+\delta_h} \right)^{-(1-\gamma)\upsilon} \left(\frac{\gamma_2^{\gamma_2} \gamma_1^{(1-\gamma_2)}}{r+\delta_h} \right) \qquad (27) \\ = & \ \frac{h_B^{1-\gamma}}{z_h^{1-\upsilon} w^{\gamma_2 - \upsilon(1-\gamma_1)}}. \end{aligned}$$

As in the statement of the proposition, let the left hand side of (26) be labeled $F(s)$. Then, an interior solution requires that $F(0) > 0$, or,

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

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 (26) 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 (10a), 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

$$n(a) = \frac{m(a)^{\frac{1}{1-\gamma}}}{e^{-\delta_h(a-s-6)} m(6+s)^{\frac{1}{1-\gamma}} + \frac{(r+\delta_h)e^{-\delta_h(a-R)}}{\gamma_1 \delta_h} \int_{e^{\delta_h(6+s-R)}}^{e^{\delta_h(a-R)}} (1 - x^{\frac{r+\delta_h}{\delta_h}})^{\frac{\gamma}{1-\gamma}} dx}, \qquad (28)$$

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

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

$$x(a) = \left(\frac{\gamma_2 w}{r+\delta_h}\right) C_h(z_h, w, r) m(6+s)^{\frac{1}{1-\gamma}} e^{\frac{r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}(a-s-6)}, \; a \in [6, 6+s), \tag{29}$$

$$x(a) = \left(\frac{\gamma_2 w}{r+\delta_h}\right) C_h(z_h, w, r) m(a)^{\frac{1}{1-\gamma}}, \qquad a \in [6+s, R). \tag{30}$$

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

$$\begin{aligned} h(a) \;&=\; C_h(z_h, w, r)\{e^{-\delta_h(a-s-6)} \frac{\gamma_1}{r+\delta_h} m(6+s)^{\frac{1}{1-\gamma}} + \frac{e^{-\delta_h(a-R)}}{\delta_h} \\ &\quad \int_{e^{\delta_h(6+s-R)}}^{e^{\delta_h(a-R)}} (1 - x^{\frac{r+\delta_h}{\delta_h}})^{\frac{\gamma}{1-\gamma}} dx\}, \qquad a \in [6+s, R). \end{aligned} \tag{31}$$

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

$$\begin{aligned} h_E \;&=\; \upsilon^{\upsilon} h_B \left[\frac{\gamma_1^{\gamma_1(1-\gamma_2)} \gamma_2^{\gamma_1 \gamma_2} z_h^{\gamma_1} w^{(1-\gamma_1)(1-\gamma_2)}}{(r+\delta_h)^{(1-\gamma_2)}}\right]^{\frac{\upsilon}{1-\gamma}} \\ &\quad e^{-\upsilon(r+\delta_h(1-\gamma_1))s} m(6+s)^{\frac{\upsilon(1-\gamma_2)}{1-\gamma}} \end{aligned} \tag{32}$$

References

[1] Banerjee, A. V., and E. Duflo, 2004, "Growth Theory Through the Lens of Development Economics," MIT working paper.

[2] Barro, R. J., and Lee, J, 1996, "International Measures of Schooling Years and Schooling Quality," *American Economic Review*, May (Papers and Proceedings), 86(2), pp. 218-23.

[3] Becker, G. S., 1964, *Human Capital : A Theoretical and Empirical Analysis, with Special Reference to Education.* New York: National Bureau of Economic Research.

[4] Belzil, C., and Hansen, J., 2002, "Unobserved Ability and the Return to Schooling," *Econometrica*, vol. 70(5), pp. 2075-2091.

[5] Ben Porath, Y., 1967, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy* 75, pt. 1, 352-365.

[6] Bils, M., and Klenow, P., 2000a, "Does Schooling Cause Growth?" *American Economic Review* 90(5):1160-83

[7] Bils, M., and Klenow, P., 2000b, "Explaining Differences in Schooling Across Countries," mimeo, University of Rochester.

[8] Browning, M., L. P. Hansen and J. J. Heckman, 1999, "Micro Data and General Equilibrium Models," in J. B. Taylor and M. Woodford (eds), **Handbook of Macroeconomics**, Vol. 1, North Holland, pp: 544-633.

[9] Caselli, F, 2006, "The Missing Input: Accounting for Cross-Country Income Differences," in Aghion P. and S. Durlauf (eds) **Handbook of Economic Growth**, North Holland publishers.

[10] Chari, V.V:, Kehoe, P.J., and McGrattan, E.R., 1997, "The Poverty of Nations: A Quantitative Investigation," Research Department Staff Report 204, Federal Reserve Bank of Minneapolis.

[11] Cooley, T. (ed.), 1995, **Frontiers of Business Cycle Research**, Princeton University Press.

[12] Cooley, T. and E.C. Prescott, 1995, "Economic Growth and Business Cycles," in T. Cooley (ed.) **Frontiers of Business Cycle Research**, Princeton University Press.

[13] Erosa, A., T. Koreshkova and D. Restuccia, 2007, "How Important is Human Capital? A Quantitative Theory Assessment of World Income Inequality," mimeo.

[14] Galor, O., 2006, "From Stagnation to Growth: Unified Growth Theory," 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, R. 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] 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.

[18] Hsieh, C. and P. Klenow, 2003, "Relative Prices and Relative Prosperity," NBER Working Paper 9701.

[19] Karoly, L., M. Kilburn, and J. S. Cannon, 2005, **Early Childhood Interventions: Proven Results, Future Promise**, Rand Corporation.

[20] Klenow, P. and Rodríguez-Clare, A., 1997, "The Neoclassical Revival in Growth Economics: Has it Gone Too Far?" in Ben Bernanke and Julio Rotenberg, eds., **Macroeconomics Annual 1997**. Cambridge, MA: MIT Press, 1997, pp. 73-102.

[21] Kuruşçu, B., 2006, "Training and Lifetime Income," *American Economic Review*, forthcoming.

[22] Lucas, R., 1988, "On the Mechanics of Economic Development," J*ournal of Monetary Economics*, 22, 3-42

[23] Manuelli, R. and A. Seshadri, 2005, "Schooling and Human Capital: Theory and Implications," manuscript.

[24] Manuelli, R. and A. Seshadri, 2007a, "Explaining International Fertility Differences," manuscript.

[25] Manuelli, R. and A. Seshadri, 2007b, "Human Capital and Inmigration," manuscript.

[26] Mincer, J., 1974, **Schooling, Experience and Earnings**, Columbia University Press, New York, 1974.

[27] Parente, S.L. and E.C. Prescott., 2000. **Barriers to Riches**, Cambridge, MA: MIT Press.

[28] Parente, S., Rogerson, R.D., and R. Wright, R., 2000 "Homework in Development: Household Production and the Wealth of Nations," *Journal of Political Economy*, pp: 680-687

[29] OECD, 1996, **Education at a Glance: OECD Indicators**.

[30] Solow, Robert M., 1956, "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, Vol.70 No.1, pp: 65 - 94.