Human Capital and the Wealth of Nations[†]

By Rodolfo E. Manuelli and Ananth Seshadri*

We reevaluate the role of human capital in determining the wealth of nations. We use standard human capital theory to estimate stocks of human capital and allow the quality of human capital to vary across countries. Our model can explain differences in schooling and earnings profiles and, consequently, estimates of Mincerian rates of return across countries. We find that effective human capital per worker varies substantially across countries. Cross-country differences in Total Factor Productivity (TFP) are significantly smaller than found in previous studies. Our model implies that output per worker is highly responsive to changes in TFP and demographic variables. (JEL E23, I25, J24, J31, O47)

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 the classic work of Solow (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 Rodríguez-Clare (1997); Hall and Jones (1999); Parente and Prescott (2000); and Bils and Klenow (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 in accounting for relative incomes across countries. The key difference between our work and previous analyses is in the measurement of human capital. The standard approach largely 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 Mincerian 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 setup is flexible enough that individuals can choose not only the length of the schooling period (which we identify as a measure of the *quantity* of human capital) but also 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 US economy and 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. To be precise, in this exercise, the change in TFP results in endogenous changes in all variables, including the stocks of human and physical capital. We find that, in the model economy, the differences in TFP needed to account for the differences in output per worker do not exceed 35 percent. Our result is mostly driven by the model's implications about the stocks of human and physical 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 other recent estimates. Since the model matches actual years of education reasonably well, we conclude that it is differences in the *quality* of human capital that account for our findings.

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) when viewed through the eyes of our model. Moreover, the model is able to reproduce the negative cross-country correlation between output per worker and the return to schooling.

This research is related to the recent analysis of the effect of human capital on cross-country income differences. It provides an alternative way of computing human capital to that advanced by Klenow and Rodríguez-Clare (1997) and Bils and Klenow (2000). The main difference lies in the use of a Mincer based estimate of human capital stocks (taking schooling as exogenous) versus a model based measure. The papers closest to ours are Erosa, Koreshkova, and Restuccia (2010) and Cordoba and Ripoll (2013). From the point of view of the computation of the stock of human capital and the impact of changes in productivity on the equilibrium level of the stock there are two key differences. First, these papers assume that post-schooling human capital is independent of economic forces and hence does not enter in the determination of a country's stock of human capital. Second, they do not allow for differences in early childhood human capital. These differences mitigate the effect of economic conditions on human capital accumulation and in turn, results in a much smaller sensitivity of stocks of human capital to changes in TFP. Despite

these differences, this body of work suggests that standard Mincerian techniques underestimate the importance of human capital for economic development.

I. The Model

In this section we describe the basic model, characterize its solution, and derive the implications for quantity and quality of schooling of differences in total factor productivity. The basic structure is, essentially, the Ben-Porath (1967) model augmented to incorporate an early childhood sector, and allowing for different market goods to be used in the production of human capital.

A. *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

$$(1) \qquad \max_{x_S, x_w, x_E, n, s} \int_6^R e^{-r(a-6)}[(1 - \tau)wh(a)(1 - n(a))$$

$$- (\aleph_{[n(a)=1]} \, p_s \, x_s(a) \, + \, (1 - \aleph_{[n(a)=1]})p_w \, x_w(a))] \, da \, - \, p_E \, x_E,$$

subject to

$$(2a) \qquad \dot{h}(a) \, = \, z_h[n(a)h(a)]^{\gamma_1} x_s(a)^{\gamma_2} - \delta_h h(a), \quad a \in [6, 6 + s),$$

$$(2b) \qquad \dot{h}(a) \, = \, z_h[n(a)h(a)]^{\gamma_1} x_w(a)^{\gamma_2} - \delta_h h(a), \quad a \in [6 + s, R]$$

and

$$(3) \qquad h(6) \, = \, h_E \, = \, h_B \, x_E^\upsilon$$

with h_B given.[1] In this formulation $\aleph_A$ is the indicator function of the set A.

Equations (2) and (3) describe the human capital accumulation technology. Here we distinguish the schooling period, which we identify with ages 6 to $6 + s$, and the on-the-job training period, corresponding to ages $6 + s$ to R, the retirement age. The differences across stages are associated with the nature of the market inputs, x_s or x_w, that are used to produce human capital. In this formulation, $n(a)$ is the fraction of the individual's human capital allocated to the production of human capital.[2] Investments in early childhood, which we denote by x_E (e.g., medical care, nutrition, and development of learning skills), determine the level of each individual's human

---

[1] 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.

[2] The technology for early childhood production assumes that early childhood education and schooling investments enter in a Cobb-Douglas fashion to determine human capital levels. There are few attempts to estimate this technology. Cunha, Heckman, and Schennach (2010) estimate this technology and find that these investments are complementary inputs. Complementarity will further increase the elasticity of the stock of human capital with respect to TFP.

capital at age six, $h(6)$, or h_E for short.[3] In what follows, we will allow for the possibility that the production of x_s and x_E use labor more intensively than other goods.[4] Let $\gamma \; (= \gamma_1 + \gamma_2)$ denote the returns to scale in the production of human capital.

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 except for the potentially different market inputs. 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—$x_j(a)$—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.[5] 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 Appendix A. 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. Assume that* $\gamma_2 - \upsilon(1 - \gamma_1) > 0$, *and that condition* Interiority (*in Appendix A*) *is satisfied. Then schooling,* s, *is strictly positive and*

(i) *Increasing in after tax wages.*

(ii) *Decreasing in the price of schooling goods,* p_s, *and on-the-job training goods,* p_w.

(iii) *Increasing in the measure of innate ability,* z_h, *and decreasing in the initial endowment of human capital,* h_B.

Moreover the level of human capital at the end of the schooling period is given by

$$(4) \qquad h(s + 6) = \left[\frac{z_h \gamma_1^{1-\gamma_2} \gamma_2^{\gamma_2}}{r + \delta_h} \left(\frac{(1 - \tau)w}{p_w} \right)^{\gamma_2} m(6 + s) \right]^{1/(1-\gamma)},$$

where

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

If condition Interiority *does not hold, individuals choose* $s = 0$ (*no schooling*).

---

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

[4] It is clear that parents' time is also important. However, given exogenous fertility, it seems best to ignore this dimension.

[5] 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.

PROOF:
See Appendix A.
There are several interesting features of the solution.

(i) *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. Equation (31a) in Appendix A shows that if $\gamma_2 - \upsilon(1 - \gamma_1) = 0$, changes in TFP—which get mapped in the steady state into changes in the wage rate w—have no impact on schooling. Thus, a model that ignores the role of market goods in producing human capital (i.e., $\gamma_2 = \upsilon = 0$) would have the counterfactual implication that the equilibrium level of schooling is the same in every country. Such a model would have to rely on differences in interest rates or the working horizon as the only source of *equilibrium* differences in schooling across countries.[6] Our formulation is flexible enough so that the impact of changes in the wage rate on the equilibrium level of 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 six 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 cross-country differences in the equilibrium years of schooling.

(ii) *Development and Schooling Quality.*—Holding both innate ability, z_h, and schooling, s, constant, equation (4) shows that the elasticity of human capital—which is a measure of quality per year of schooling—is $\gamma_2/(1 - \gamma)$, which is fairly large in our preferred parameterization.[7] 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

---

[6] Cross-country differences in the ability to learn, z_h and the endowment of human capital, h_B can also account for differences in s. Since we have no evidence of systematic differences across countries, we do not pursue this possibility.

[7] To be precise, we find that $\gamma_2 = 0.40$ and $\gamma = 0.886$. Thus the elasticity of the quality of human capital with respect to wages is 3.5.

"close" to constant returns, then the model will predict large cross-country differences in human capital even if TFP differences are small.[8]

(iii) *Early Human Capital.*—The model implies that, holding innate ability and schooling constant, the level of early childhood human capital, h_E, increases with the wage rate (TFP), and decreases with the price (p_E) of inputs necessary to produce it and the cost of on-the-job training (p_w). (See equation (29) in Appendix A.)

(iv) *Individual Characteristics, Schooling, and Human Capital.*—The equilibrium level of schooling depends on the ratio z_h/h_B. Specifically, the equilibrium level of schooling is monotone increasing in this ratio, which we interpret as a measure of innate ability (relative to the endowment at birth).

It is of interest to estimate the level of human capital per year of schooling corresponding to different countries indexed by their wage rate (or TFP). The results described in the previous proposition are inadequate for this analysis. The reason is simple: if we wish to compare the level of human capital of two individuals that have the *same schooling* but live in economies with *different levels of TFP*, then those individuals should be different in some dimension so that, even though when faced with different environments (different w), they choose the same level of schooling. In the context of the model, one way of resolving this is to adjust the level of innate ability for each individual as a function of wage rates and prices so that his choice of schooling is independent of country characteristics. The quality of human capital, also endogenously determined, does vary with the characteristics of the economic environment. The basic result is described in the following proposition.[9]

PROPOSITION 2: *There exists a function $G(s)$ such that the level of human capital at the end of the schooling period, $h(6 + s)$, is given by*

$$(5) \qquad h(6 + s) = G(s)\left(\frac{(1 - \tau)w}{p_E}\right)^{\frac{v}{1-v}}.$$

PROOF:
   See Appendix A.
   Equation (5) illustrates two important mechanisms underlying our results. First, "identical" individuals (in the sense of having exactly the same level of schooling, s) have different levels of human capital depending on the wage rate (TFP) of the country of origin. The higher the level of TFP, the higher the quality of human capital. Second, these differences in quality hinge on differences in early human capital. To be precise, if all individuals had exactly the same human capital at age

---

[8]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.
   [9]It is also possible to vary the initial endowment of human capital but we find that heterogeneity in ability to learn is a more interesting dimension to study.

six, i.e., if $\upsilon = 0$, then the average individual in each country would choose different levels of schooling but the quality of schooling—defined as the level of human capital per year of schooling—would be the same across countries.[10]

Equilibrium Age-Earnings Profiles.—Even though the model is explicit about market income and investments in human capital, it says very little about the timing of payments and who pays for what. Earnings at age a are (assuming that individuals pay for all the costs of on-the-job training),

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

Given the solution to the income maximization problem (see the details in Appendix A), measured income as a function of experience, defined as $p = a - s - 6$, and schooling, s, is

$$(6) \quad \hat{y}(s, p) = (1 - \tau)w\left[\frac{z_h\gamma_1^{1-\gamma_2}\gamma_2^{\gamma_2}}{r + \delta_h}\left(\frac{(1 - \tau)w}{p_w}\right)^{\gamma_2}\right]^{1/(1-\gamma)}$$

$$\times \left\{e^{-\delta_h p}h(6 + s) + \frac{r + \delta_h}{\gamma_1}\int_{6+s}^{p+6+s}e^{-\delta_h(p+6+s-t)}m(t)^{\frac{\gamma}{1-\gamma}}dt - \frac{\gamma}{\gamma_1}m(a)^{1/(1-\gamma)}\right\}.$$

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.[11]

B. *Equilibrium*

To close the model we study the steady state of this economy. 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 and its distribution across sectors. 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.

---

[10] The nature and technology of early childhood human capital is the subject of much recent research. See, for example, Cunha, Heckman, and Schennach (2010) and Schoellman (2013).

[11] Proposition 2 is one way of doing that. It essentially assumes that the changes in s are driven by heterogeneity in innate ability.

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

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

It is easy to check that in the steady state

(7)
$$N(a, t) = \phi(a)e^{\eta t},$$

where

(8)
$$\phi(a) = \eta \frac{e^{-\eta a}}{1 - e^{-\eta T}},$$

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

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

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

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.—Let F^c and F^s describe the production functions of the aggregate good and the schooling good, respectively. Optimization on the part of firms implies that

(9)
$$p_k(r + \delta_k) = z F_k^c(\kappa_c, 1),$$

where κ_c is the physical capital–human capital ratio in the production of the aggregate good and p_k is the relative (to consumption) price of capital. The wage rate per unit of human capital, w, is

(10)
$$w = z F_h^c(\kappa_c, 1).$$

The price of the schooling good (and also of the good used in producing early child-
hood human capital), $p_s = p_E$, and the capital–human capital ratio in this sector,
κ_s, satisfy

$$p_s z F_h^s(\kappa_s, 1) = w,$$

$$p_s z F_k^s(\kappa_s, 1) = p_k(r + \delta_k).$$

Then, output per worker (measured in terms of the aggregate good) is

$$y = z F^c(\kappa_c, 1) \overline{h}_c^e + z F^s(\kappa_s, 1) \overline{h}_s^e,$$

where

$$\overline{h}_c^e + \overline{h}_s^e = \overline{h}^e.$$

The allocation of human capital to the schooling sector can be computed from the
following expression:

$$\frac{\int_6^{6+s} x_s(a)\phi(a)da}{\int_{6+s}^R \phi(a)da} + \frac{\phi(6)x_E}{\int_{6+s}^R \phi(a)da} = z F^s(\kappa_s, 1) \overline{h}_s^e.$$

Here the left-hand side is the total demand for the schooling good on a per worker
basis. Finally, $\overline{h}_c^e = \overline{h}^e - \overline{h}_s^e$.

II. Calibration

Our calibration strategy involves choosing the parameters so that the steady state
implications of the model economy presented above are consistent with observa-
tions for the United States (circa 2010). Following a standard approach in macro we
assume that the production function for consumption goods is Cobb-Douglas and
given by

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

The technology to produce early childhood human capital and educational services,
(x_E and x_s) is assumed to be less capital-intensive, consistent with the available evi-
dence, and takes the form $z k^\beta (\overline{h}^e)^{1-\beta}$. In what follows, we set θ to 0.33 and β to 0.2.
Given this, and a capital–output ratio of 2.52, the interest rate is 5.5 percent. The
depreciation rate is set at $\delta_k = 0.075$.

Less information is available on the fraction of job training expenditures that are
not reflected in wages. We follow a long-standing tradition in equating earnings with
$wh(1 - n) - x$.[12]

---

[12] Measured earnings likely exclude some part of the training expenditures. Assuming that half of the goods
expenditures are not counted toward measured earnings hardly changes the quantitative findings.

Our theory implies that it is only the ratio $h_B^{1-\upsilon}/\left(z_h^{1-\upsilon}(1-\tau)w^{\gamma_2-\upsilon(1-\gamma_1)}\right)$ that matters for all the moments of interest. Since we assume that initial human capital h_B and ability to learn z_h are common across all countries, we can choose z, p_k (which determines w) for the US economy and h_B arbitrarily and calibrate z_h (common to all countries) to match a desired moment. We normalize to $h_B = 1$ and $z = 1$ (for the US economy). We set $B = 25$ and $R = \min\{64, T\}$ and $T = 78.8$. This leaves us with five parameters, z_h, γ_1, γ_2, δ_h, and υ. The moments we seek in order to pin down these parameters are:

(i)   Wage rate at age 55/wage rate at age 25 of 2 (authors' calculations using the American Community Survey 2011).

(ii) Years of schooling of 12.05 (Barro and Lee 2010).

(iii) Schooling expenditures as a fraction of GDP of 4.5 percent (UNESCO 2007).

(iv) Pre-primary expenditures as a fraction of GDP of 1.1 percent (Weiss and Brandon 2010).

(v)   Wage rate at 64/wage rate at 55 = 0.79 (French 2005).

A few comments are in order. We measure wage growth as the average wage of an individual aged 54 to 56 relative to the average wage between 24 and 26. Wage growth estimates from the NLSY between ages 25 and 55 are slightly lower than a factor of 2. Targeting a number at the high end of the estimates leads to a lower calibrated value of $\gamma_1 + \gamma_2$.[13] Public expenditures on schooling are around 4 percent of GDP while a measure that includes public and private exceeds 5 percent. We target a number in the middle. UNESCO data on pre-primary educational expenditures relative to GDP include only purchased inputs and does not incorporate parental time and expenses at home.[14] Recent estimates of this same number (see Weiss and Brandon 2010) that incorporate parental time peg the value at 1.1 percent of GDP. We use this as our target. Finally wage rates decline at the end of the life-cycle. French (2005) reports a decline in wage rates of more than 20 percent between ages 55 and 64. Our own estimates from the Health and Retirement Study (HRS) also imply a 21 percent decline in earnings between the very same age range. For details on the HRS earnings data, see Scholz et al. (2006). We choose a target of 0.79 for the wage ratio between ages 64 and 55. A higher rate of depreciation leads to a lower elasticity of stocks of human capital with respect to TFP.[15]

---

[13] Our estimate of the degree of returns to scale, $(\gamma = \gamma_1 + \gamma_2 = 0.886)$, lies in the range reported by Browning, Hansen, and Heckman (1999).

[14] http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=136&IF_Language=eng&BR_Topic=0.

[15] Since the Ben-Porath model is silent on the labor leisure choice, a more appropriate moment to target is the wage rate. The cross-country implications are similar if we were to use a zero depreciation rate. The elasticity of schooling with respect to GDP per worker is even higher when we reduce the rate of depreciation on human capital.

TABLE 1—CALIBRATED PARAMETERS

Parameter	z_h	γ_1	γ_2	ν	δ_h
Value	0.334	0.486	0.400	0.618	0.0270

The previous equations correspond to moments of the model when evaluated at the steady state. Calibration requires us to solve a system of five equations in five unknowns. The resulting parameter values are given in Table 1.

In Appendix B we present some additional evidence that can be used to evaluate our choice of parameter values.

III. Results

A. *Evidence*

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 Penn World Tables (PWT 8.0) and arrange them in deciles according to their output per worker, y which is measured as the average over the years 2003 and 2007.

Output per worker is calculated as real GDP at constant 2005 national prices (rgdpna) divided by the number of persons engaged (emp). We exclude countries with population less than 1,000,000. We also exclude top ten oil exporting countries measured as barrels per capita in 2009 from the CIA World Factbook.[16] Next, we combine them with observations on years of schooling (s) for the total population aged 25 or older taken from Barro and Lee (2010). The public expenditures on education relative to GDP (x_s) comes from UNESCO.[17] The life expectancy (T) comes from the life expectancy at birth from the World Data Bank.[18] The total fertility rate ($e^f/2$) also comes from the World Data Bank which is measured as the total fertility rate adjusted for the infant mortality rate taken from the CIA World Factbook.[19] All variables are averaged over the years 2003 to 2007. The relative price of capital (p_k) measured as the ratio of the "price level of capital formation" relative to the "price level of household consumption" from the PWT 8.0 and averaged over the years 2003–2007 after dividing by the price of capital for the United States. We break the countries into ten deciles after eliminating countries with any missing variables. This leaves us with 101 countries. The population values are displayed in the following table.

Table 2 illustrates the wide disparities in incomes across countries. The United States possesses an output per worker (normalized to one) that is about 53 times as high as the countries in the bottom decile. Years of schooling also vary systematically with the level of income from about three years at the bottom deciles to about

---

[16] http://world.bymap.org/OilExports.html.
[17] http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=136.
[18] http://data.worldbank.org/indicator/SP.DYN.LE00.IN.
[19] http://data.worldbank.org/indicator/SP.DYN.TFRT.IN.

TABLE 2—WORLD DISTRIBUTION

Decile	y (relative to United States)	s	x_s	T	$e^f/2$ (TFR/2)	p_k
90–100	0.872	10.36	5.1	80	0.82	1.00
80–90	0.743	9.77	5.7	79	0.93	0.97
70–80	0.508	9.79	4.6	76	0.72	1.14
60–70	0.348	8.79	4.7	74	0.95	1.23
50–60	0.251	8.45	3.9	70	0.94	1.35
40–50	0.187	6.29	4.6	71	1.20	1.10
30–40	0.125	7.64	4.7	66	1.34	1.47
20–30	0.077	5.18	3.8	62	1.79	1.44
20–10	0.037	3.61	4.6	54	2.22	1.34
0–10	0.019	2.75	3.6	53	2.39	1.22

11 at the top. Expenditures on primary and secondary schooling as a fraction of GDP do not systematically vary 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 vary much in the top half. Finally, the relative price of capital in the richest countries is about two-fifths of the level in the poorest countries.

B. *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 be precise, we take demographic variables (e^f (fertility rate), T (life expectancy), and R (retirement age)) for the "average" country in each decile and the price of capital p_k as given, and we choose the level of TFP for each decile so that the model's predictions for output per worker match that for the chosen decile. Table 3 presents the predictions of the model and the data.

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 over-predict educational attainment for the richer set of countries. In terms of a rough measure of quality such as schooling expenditures as a fraction of output, the model actually under-predicts investment at the high end of the world income distribution, and slightly over-predicts expenditures for the poor countries.

The striking results are the estimates of TFP. According to our model, TFP in the poorest countries (i.e., countries in the lowest decile of the world income distribution) is estimated to be only 63 percent 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 Rodríguez-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 percent. Thus, our estimate of TFP in the poorest countries is more than two times higher.

TABLE 3—OUTPUT AND SCHOOLING WITH p_k, f, T VARYING ACROSS COUNTRIES: DATA AND MODEL

Decile	y (relative to United States)	*TFP*	s		x_s	
			Data	Model	Data	Model
90–100	0.872	0.97	10.36	11.68	5.1	3.2
80–90	0.743	0.95	9.77	11.50	5.7	3.7
70–80	0.508	0.94	9.79	10.32	4.6	2.5
60–70	0.348	0.92	8.79	9.47	4.7	3.3
50–60	0.251	0.91	8.45	8.70	3.9	3.0
40–50	0.187	0.81	6.29	8.49	4.6	4.0
30–40	0.125	0.85	7.64	7.06	4.7	4.0
20–30	0.077	0.79	5.18	5.98	3.8	5.2
20–10	0.037	0.71	3.61	4.25	4.6	5.8
0–10	0.019	0.63	2.75	2.83	3.6	4.3

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 5.7. Thus, according to the model, changes in TFP have a large multiplier effect on output per worker.[20]

As indicated, countries differ both in terms of TFP and demographics and both of these sets of variables affect human capital investments. 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 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 an important role.

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 by about 53 percent. This is accompanied by a 26 percent increase in the level of schooling (from 2.83 to 3.59). 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.

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 percent and 90 percent of the richest countries) there is almost no change in output per worker (from 0.743 to 0.730). 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

---

[20] 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.

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 separate work, we studied the impact that changes in TFP have upon (endogenously chosen) fertility.[21]

IV. 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. *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.[22] The Mincerian framework assumes that the average human capital of a worker of age a, at time t in country j with s years of schooling is

$$(11) \qquad \hat{h}_j(a,\, t) \;=\; e^{\beta_0(X_{jt})} e^{\beta_1(s)+\beta_2(a-s)},$$

where X_{jt} represents variables that influence the quality of human capital, and the β_j are functions. In this specification, the log of earnings is given by

$$(12) \qquad \ln(w_{jt}\, \hat{h}_j(a,\, t)) \;=\; \ln(w_{jt}) \;+\; \beta_0(X_{jt}) \;+\; \beta_1(s) \;+\; \beta_3(a-s).$$

Klenow and Rodríguez-Clare (1997) and Bils and Klenow (2000) use versions of equation (12) to estimate $\hat{h}_j(a,\, t)$. It is not easy to summarize their results but the general finding is that for all the specifications that they tried, the term $\beta_0(X_{jt})$ did not contribute much to the stock of human capital.[23]

To understand why we find that quality differences are important it is useful to contrast equation (12) with equation (6), which is the model's implication for earnings. First, the model's version is nonlinear, and it does not factor out like equation (12). Second, the same factors that account for differences in schooling also affect both the

---

[21] In related research (Manuelli and Seshadri 2009) we study a version of the model with endogenous fertility and find that the basic results do not change much.

[22] There is a large literature examining the importance of school quality as measured by expenditures. Hanushek (2006) in his review of the literature finds that additional resources have not led to discernible improvements in student achievement. Case and Yogo (1999) find that the quality of schools in black South Africa has a large and significant effect on the rate of return to schooling for black men. Glewwe and Kremer (2006) find that school quality in developing countries is quite low. Teacher quality and availability is a common problem. In their review of the importance of schooling quality, they find that while earlier retrospective studies suggest that providing additional resources may have little impact on learning, more recent evidence from randomized evaluations paints a more uniformly positive picture.

In our model, additional improvements to the quality of schooling have only a temporary impact on the stock of human capital. The idea is that the individual realizes that he or she possesses more than the desired level of human capital and proceeds to cut back on human capital accumulation later in life. Hence the micro finding that additional resources do not matter much for wages later in life is consistent with our framework.

[23] Caselli (2005) extends their work and tries a variety of different specifications. He does not find that adding quality makes a significant difference.

return to experience and the quality of human capital. Last, the model implies that the shape of the age-earnings profile is not independent of prices and wage rates.

Are these differences quantitatively important? First, consider a simple version that ignores quality differences first.[24] Assume that $\beta_1(s) \approx 0.10 \times s$, which corresponds to a 10 percent rate of return on schooling. This is a commonly used value. Thus, if we take a country from the lowest decile with $s_P = 3$, and assume that the average worker in the United States has 12 years of schooling, the average human capital of the poor country relative to the United States is

$$\frac{\hat{h}_P}{\hat{h}_{US}} = e^{-0.1 \times 9} = 0.41.$$

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

$$\overline{h}_j = e^{\beta_{j0}} e^{\beta_1(s)}.$$

If, as before, we compare a country from the bottom decile of the output distribution with the United States, Table 4 implies that its relative average human capital is 0.07. It follows that our measure of quality, for this pair of countries, is simply

$$e^{(\beta_{P,0} - \beta_{US,0})} = \frac{\overline{h}_P}{\overline{h}_{US}} e^{\beta_1(s_{US} - s_P)} = 0.07 \times 2.46 = 0.17.$$

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 United States. In our model, this ratio is driven by differences in wages and demographics.[25] The magnitude of the differences in relative quality suggests that ignoring this dimension can induce significant biases in the estimates of human capital.

B. *The Importance of Early Childhood and On-the-Job Training* (*OJT*)

Our model implies that, even at age six, there are substantial differences between the human capital of the average child in rich and poor countries. In Table 4 we present the values of human capital at age six (h_E) and aggregate human capital per worker ($\overline{h}^e$) for each decile relative to the United States. Even though the

---

[24] Specifically, we assume that in equation (12), $\beta_0 = 0$, and since we are comparing similar age distributions we can ignore β_3.

[25] See Heckman, Lochner, and Todd (2008) and Hanushek and Woessmann (2011) on differences in quality.

The latter find a weak association between resources and test scores. As a conceptual matter, test scores reflect both acquired human capital h as well as innate ability z_h. Consider our model with heterogeneity in ability z_h. At a given point in time, average test scores reflect an expectation over some function $F(h(z_h), z_h)$. A linear regression of test scores on schooling resources can yield a zero coefficient especially if ability is controlled for (this corresponds to equation (3) in their paper). The mechanism is straightforward. Increases in resources lead only to temporary changes in stocks of human capital at that age. The individual realizes that he possesses more than the privately optimal level of human capital and hence cuts back on human capital accumulation later on in life. Changes in TFP can generate permanent increases in stocks of human capital (which presumably will be reflected in test scores) but these cross-country regressions typically control for the level of income in a given country as well.

TABLE 4—UNDERSTANDING HUMAN CAPITAL DIFFERENCES

	Relative to United States			Contribution (shares)		
Decile	y	h_E	$\bar{h}$	OJT	Schooling	Early childhood
90–100	0.872	0.93	0.90	0.43	0.48	0.08
80–90	0.743	0.88	0.80	0.43	0.49	0.09
70–80	0.508	0.80	0.64	0.45	0.45	0.10
60–70	0.348	0.73	0.50	0.44	0.44	0.12
50–60	0.251	0.67	0.41	0.45	0.42	0.13
40–50	0.187	0.56	0.31	0.43	0.42	0.15
30–40	0.125	0.53	0.24	0.44	0.38	0.18
20–30	0.077	0.45	0.16	0.40	0.37	0.23
20–10	0.037	0.35	0.08	0.32	0.33	0.35
0–10	0.019	0.28	0.05	0.32	0.24	0.44

differences in early childhood capital are small for the relatively rich countries, they are significant 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 percent of the human capital of a US 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 (4)). The last three columns of Table 4 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.

C. *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 seem to be close to 10 percent when linearity is imposed.[26] Thus, one challenge for the model economy is to reproduce the rate of return in a Mincer-style regression.

---

[26] 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.

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 same wage and interest rate, differences in s are driven by differences in (z_h, h_B). These two variables have very different effects 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 1 and 20 years of education. Given the nonlinearity of the earnings function, we need population weights of individuals in different categories of experience and schooling. We obtain these population weights from the American Community Survey (2011), with schooling ranging from 1 through 20 and experience going from 5 to 45.[27] We then proceed in two steps. First, we divide the population into groups based on schooling levels: {1–6, 7–8, 9–10, 11, 12, 13, 14, 15–16, 17–18, 19–20}. We calculate the ratio of earnings around age 50 to around age 30 and average years of schooling for each group. For each group, we solve for (z_h, h_B) so as to match the earnings ratio and the average years of schooling. Next, using the estimated z_h, we further adjust h_B so that it generates all of the schooling levels within each group.[28] Thus, there will be as many pairs of (z_h, h_B) as there are schooling levels 1–20. 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 0.1 while the group with 12 years of schooling and 30 years of experience has a mass of 0.05, 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 obtain a Mincer coefficient of 7.3 percent when restricting the education 1 to 17. This number is consistent with the Mincer coefficient attained using the actual ACS data of 9.0 percent.[29] The coefficients on experience and experience squared in the regression are 0.0976 and −0.0018 respectively.

As a second test, we computed for each representative country in our world distribution of output (ten countries in all), the effect on log earnings of an additional year of education. We took this to be the return on schooling in country (decile) i. The procedure is as follows: we used *exactly the same distribution* of pairs (z_h, h_B) that we estimated for the United States with the appropriate (for each country) population weights to compute for each "type" and country age-earnings profiles. We computed the Mincer coefficient for each decile by regressing earnings against education, experience, and experience squared using the artificial dataset. As in the case

---

[27] We focus on males born in the 50 US states aged 25 to 55. We further exclude samples with zero earnings and samples in the top and bottom 10 percent of the earnings distribution.

[28] Adjusting by z_h while keeping h_B fixed does not have much of an impact on the results.

[29] We restrict to schooling years 1 to 17 because there can be measurement errors in mapping the reported educational attainment to the years of schooling among post-secondary education.

of the United States, we restrict to schooling attainment of 1 to 17 and experience of 5 to 45. Finally, we regressed the estimated return on GDP per capita and obtained a coefficient of -0.074, which is statistically significant at the 5 percent level.[30] This is to be compared with a similar exercise using actual data run by Banerjee and Duflo (2005) using different datasets. Their estimate is -0.084.

To summarize, the cross-sectional relationship within a country 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 flatter than the rich country profile.[31] Since the return to formal education is, approximately, the derivative of the earnings-schooling profile, it is the smaller slope of the earnings-schooling profile as TFP increases (a cross-country effect) that accounts for the cross-country observations.

D. *The Importance of Early Childhood*

The baseline model we presented is the workhorse of modern labor economics. While it is one of the most commonly used frameworks in thinking about age wage profiles, it is by no means the dominant model of schooling choices. In this section we highlight the role played by early childhood human capital. To demonstrate the quantitative importance, we will redo the cross-country exercise by eliminating early childhood production and by assuming that physical capital is not used in the production of human capital. In particular, we explore the implications of assuming that all children (regardless of the country of origin) begin schooling with the same stock of human capital (that is, we set $\nu = 0$). In addition, we assume that the technology to produce early childhood and schooling does not use physical capital at all. By suppressing the use of physical capital in the production of educational services, and by eliminating early childhood human capital production we aim to demonstrate the importance of these features in understanding cross-country differences in schooling and the estimated TFP.

We re-calibrate the parameters $(z_h, \gamma_1, \gamma_2)$ so that the income ratio, schooling, and educational expenditure to GDP ratio are exactly the same as the estimated values. The corresponding new values are: $z_h = 0.221$, $\gamma_1 = 0.529$, $\gamma_2 = 0.359$.

As in the baseline case, we vary TFP to match output per worker when demographic variables and the relative price of capital are set according to the data. The results are in Table 5.

TFP in the bottom decile is now 60 percent of that in the top decile (which compares with 63 percent in the baseline model). Eliminating early childhood human capital accumulation has an important effect on schooling decisions. Recall that

---

[30] Standard error of 0.024.
[31] This is consistent with the findings in Lagakos et al. (2013).

TABLE 5—OUTPUT AND SCHOOLING WITHOUT p_k, f, T VARYING ACROSS COUNTRIES

Decile	y (relative to United States)	TFP	s Data	Model
90–100	0.872	0.97	10.36	11.28
80–90	0.743	0.94	9.77	10.80
70–80	0.508	0.93	9.79	8.79
60–70	0.348	0.91	8.79	7.37
50–60	0.251	0.89	8.45	6.07
40–50	0.187	0.80	6.29	5.03
30–40	0.125	0.83	7.64	3.21
20–30	0.077	0.77	5.18	1.27
20–10	0.037	0.69	3.61	0.00
0–10	0.019	0.60	2.75	0.00

a higher initial stock of human capital, all else equal, will reduce the amount of schooling. Indeed, eliminating the early childhood technology effectively increases the stock of initial human capital at age six for a child in a poor country. This reduces schooling. Table 5 demonstrates that the effect is quantitatively large. Children in the bottom decile do not go to school and the model displays too much sensitivity of schooling with respect to TFP.

This also has implications for Mincerian rates of return. We argued above that the baseline model can account for the Mincerian rates of return in the United States. By contrast, without the early childhood educational production function, the model implies too high a Mincerian rate of return. The Mincer coefficient generated by variation in ability alone is around 4 percent. To see why the model without the early childhood technology implies a lower Mincerian rate of return, consider two individuals in the baseline model, one with a higher ability level than the other. The individual with a higher ability level also acquires more early childhood human capital, which in part goes toward decreasing schooling differences across these two individuals. Hence, to match a given difference in schooling levels, the baseline model will require a larger difference in abilities across these two individuals (to counteract the fact the individual with the higher ability level will possess a greater stock of human capital before schooling begins) than the model without the early childhood technology. Hence the Mincerian return generated here is lower.

The bottom line is that the particular choice of the early childhood technology is important in generating the right macro elasticity of schooling with respect to TFP as well as the right micro elasticity of schooling with respect to ability. The elasticity of the stock of human capital at age six with respect to expenditures does not affect the estimated TFP. While a lower value of υ reduces the amount of investment in early childhood human capital in the United States, it does so in every other country as well. Consequently, the implied TFP levels do not change much.

E. *Implications for TFP Growth*

Does our analysis allow any room for TFP growth in the United States over say the last 50 years? Between 1950 and 2000, real GDP per worker in the United States increased by a factor of 2.6. At the same time, the price of capital (relative to consumption) decreased by a factor of 0.78. If we were to simply feed the relative price

of capital into the model and hold TFP unchanged, output per worker increases by a factor of 1.71 and the rise in schooling is about one-half the observed increase over the same time frame.

Hence there is a role for TFP changes. TFP increases by around 33 percent between 1950 and 2000 in order to account for the rise in GDP per worker in the time series. (When relative y is $1/2.6$ and relative p_k is $1/0.78$, the estimated TFP $= 0.75$ and schooling is 6.3.) Is this reasonable? It is not easy to compare this figure with available estimates. For instance, Jorgenson (1988, p. 23) reports that "The findings presented here allocate more than three-fourths of US economic growth during the period 1948–1979 to growth in capital and labor inputs and less than one-fourth to productivity growth." While our figure is lower than the TFP growth figures reported in Jorgensen, the central thesis of our paper is that traditional growth accounting underestimates the growth in stock of human capital by not accounting for quality changes in the stock of human capital over time.

V. Conclusion

The quantitative importance of human capital in understanding cross-country income differences has been and will continue to be a hotly debated issue. Most of the existing work employs a version of the Mincer earnings regressions to estimate stocks of human capital, that puts little or no weight on quality differences, and concludes that human capital differences are small. In this paper we show that a standard human capital framework, which Mincer draws upon to derive his pioneering earnings equation, generates large differences in the stocks of human capital driven by small differences in TFP. Our results suggest that human capital has a central role in determining the wealth of nations and that the quality of human capital varies systematically with the level of development. The model is successful in capturing the large variation in levels of schooling across countries and is also consistent with the cross-country evidence on Mincerian rates of return. 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. The typical individual in a poor country not only chooses to acquire fewer years of schooling but also acquires less human capital per year of schooling. From a policy perspective, our results indicate that even modest increases in productivity can result in large gains in output per worker in the long run. Our model suggests that there are large payoffs to understanding what explains productivity differences.

Appendix A

The first-order conditions of the income maximization problem are

(13a) $\qquad (1 - \tau)whn \ \leq \ q\gamma_1 z_h(nh)^{\gamma_1} x_j^{\gamma_2}, \quad$ with equality if $n < 1$,

(13b) $\qquad\qquad p_j x \ = \ q\gamma_2 z_h(nh)^{\gamma_1} x_j^{\gamma_2}, \ j \in \{s, w\}$

(13c) $\qquad\qquad \dot{q} \ = \ rq - [q\gamma_1 z_h(nh)^{\gamma_1} x_j^{\gamma_2} h^{-1} - \delta_h] - w(1 - n),$

(13d) $\qquad\qquad \dot{h} \ = \ z_h(nh)^{\gamma_1} x_j^{\gamma_2} - \delta_h h,$

where $a \in [6, R]$, and $q(a)$ is the costate variable that must satisfy $q(R) = 0$.

Define the function $V(h, a)$ by

$$(14) \quad V(h, a) = \max \int_a^R e^{-r(t-a)}[(1 - \tau)wh(t)(1 - n(t)) - p_w x_w(a)] \, da$$

subject to

$$(15) \quad \dot{h}(t) = z_h(n(t)h(t))^{\gamma_1} x_w(t)^{\gamma_2} - \delta_h h(t), \quad t \in [a, R].$$

A standard calculation shows that

$$(16) \quad V(h, a) = (1 - \tau)w \left\{ \frac{m(a)}{r + \delta_h} h + \frac{1 - \gamma}{\gamma_1} \left[\frac{z_h \gamma_1}{r + \delta_h} \left(\frac{\gamma_2}{\gamma_1} \frac{(1 - \tau)w}{p_w} \right)^{\gamma_2} \right]^{1/(1-\gamma)} \right.$$

$$\left. \times \int_a^R e^{-r(t-a)} m(t)^{1/(1-\gamma)} \, dt \right\},$$

where

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

It is easy to show that the optimal solution has the property that

$$(17) \quad n(a)h(a) = \left[\frac{z_h \gamma_1^{1-\gamma_2} \gamma_2^{\gamma_2}}{r + \delta_h} \left(\frac{(1 - \tau)w}{p_w} \right)^{\gamma_2} \right]^{1/(1-\gamma)} m(a)^{1/(1-\gamma)}.$$

It is useful to record the optimal values of goods investment and human capital along the solution. They are given by

$$(18) \quad x_w(a) = \left(\frac{\gamma_2}{\gamma_1} \frac{(1 - \tau)w}{p_w} \right) \left[\frac{z_h \gamma_1}{r + \delta_h} \left(\frac{\gamma_2}{\gamma_1} \frac{(1 - \tau)w}{p_w} \right)^{\gamma_2} \right]^{1/(1-\gamma)} m(a)^{1/(1-\gamma)},$$

and

$$(19) \quad h(a') = e^{-\delta_h(a'-a)} h(a) + \frac{r + \delta_h}{\gamma_1} \left[\frac{z_h \gamma_1}{r + \delta_h} \left(\frac{\gamma_2}{\gamma_1} \frac{(1 - \tau)w}{p_w} \right)^{\gamma_2} \right]^{1/(1-\gamma)}$$

$$\times \int_a^{a'} e^{-\delta_h(a'-t)} m(t)^{\frac{\gamma}{1-\gamma}} \, dt, \quad a' \geq a > 6 + s.$$

Thus, once the solution is properly described by this problem (i.e., once equation (11) is satisfied), the optimal decision rules are given by equations (17), (18), and (19).

Finally, let $q(a)$ denote the marginal value of human capital in the post-schooling period. Thus,

$$(20) \qquad q(a) = \frac{\partial V}{\partial h}(h, a) = (1 - \tau)w\frac{m(a)}{r + \delta_h}.$$

The next lemma describes the solution during the schooling period.

LEMMA 1: *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 $a \in [6, 6 + s)$,*

$$(21) \qquad x_s(a) = \left(h_E^{\gamma_1}\frac{q_E}{p_s}\gamma_2 z_h\right)^{\frac{1}{1-\gamma_2}} e^{\frac{r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}(a-6)}, \quad a \in [6, 6 + s),$$

and

$$(22) \quad h(a) = h_E e^{-\delta_h(a-6)}\left[1 + \frac{(1 - \gamma_1)}{\mu}\left(h_E^{-(1-\gamma)}\left(\frac{q_E}{p_s}\right)^{\gamma_2}\gamma_2^{\gamma_2}z_h\right)^{\frac{1}{1-\gamma_2}}\left(e^{\mu(a-6)} - 1\right)\right]^{\frac{1}{1-\gamma_1}},$$

$$a \in [6, 6 + s).$$

PROOF OF LEMMA 1:
From (13b) we obtain that

$$(23) \qquad p_s x_s(a) = (q(a)h(a)^{\gamma_1})^{\frac{1}{1-\gamma_2}}(\gamma_2 z_h)^{\frac{1}{1-\gamma_2}}.$$

Let $M(a) = q(a)h(a)^{\gamma_1}$. Then,

$$(24) \qquad \dot{M}(a) = M(a)\left[\frac{\dot{q}(a)}{q(a)} + \gamma_1\frac{\dot{h}(a)}{h(a)}\right].$$

However, it follows from (13c) and (13d) after substituting (23) that

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

Using this in (24),

$$(25) \qquad q(a)h(a)^{\gamma_1} = q_E h_E^{\gamma_1} e^{(r+\delta_h(1-\gamma_1))(a-6)}.$$

Using this result to determine expenditures, $x_s(a)$, and using this in equation (13d),

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

It can be verified, by direct differentiation, that (22) is a solution.

LEMMA 2: *Assume that* $\gamma_2 - \upsilon(1 - \gamma_1) > 0$. *The solution* $(h_E, h(6+s), s, q_E)$ *of the income maximization problem is given by the solution to*

$$(26) \qquad \frac{m(6+s)}{r+\delta_h} = \frac{q_E h_E^{\gamma_1}}{(1-\tau)w} \frac{e^{-(r+\delta_h(1-\gamma_1))s}}{\left[\frac{z_h \gamma_1^{1-\gamma_2}\gamma_2^{\gamma_2}}{r+\delta_h}\left(\frac{(1-\tau)w}{p_w}\right)^{\gamma_2} m(6+s)\right]^{\gamma_1/(1-\gamma)}},$$

$$(27) \qquad h(6+s) = \left[\frac{z_h \gamma_1^{1-\gamma_2}\gamma_2^{\gamma_2}}{r+\delta_h}\left(\frac{(1-\tau)w}{p_w}\right)^{\gamma_2} m(6+s)\right]^{1/(1-\gamma)},$$

$$(28) \quad h(6+s) = h_E e^{-\delta_h s}\left[1 + \frac{(1-\gamma_1)}{\mu}\left(h_E^{-(1-\gamma)}\left(\frac{q_E}{p_s}\right)^{\gamma_2}\gamma_2^{\gamma_2} z_h\right)^{\frac{1}{1-\gamma_2}}(e^{\mu(a-6)}-1)\right]^{\frac{1}{1-\gamma_1}},$$

$$(29) \qquad h_E = \upsilon^{\frac{\upsilon}{1-\upsilon}} h_B^{\frac{1}{1-\upsilon}} \left(\frac{q_E}{p_E}\right)^{\frac{\upsilon}{1-\upsilon}}.$$

PROOF OF LEMMA 2:

Since the switch from the schooling to the working period occurs when $n(6+s) = 1$, equation (17) implies equation (27). Next, equation (25) and equation (20) evaluated at $a = 6 + s$ implies that

$$\frac{m(6+s)}{r+\delta_h} = \frac{q_E h_E^{\gamma_1}}{(1-\tau)w} \frac{e^{-(r+\delta_h(1-\gamma_1))s}}{h(6+s)^{\gamma_1}}.$$

Substituting equation (27) into that expression one obtains equation (26). Equation (29) is simply equation (22) evaluated at $a = 6 + s$. At $a = 6$, the individual solves the optimal early human capital investment problem, given the shadow price of human capital q_E.

$$\max q_E h_B x_E^{\upsilon} - p_E x_E,$$

which implies that equation (29) holds.

Combining equations (26), (27), (28), and (29) it is possible to show that the previous system of equations can be reduced to one equation in s, given by

$$(30) \qquad F_0(s) = A_0 + A_1 F_1(s),$$

with

$$F_0(s) \equiv (r + \delta_h)^{-\frac{1-\gamma_1}{1-\gamma}\left[\frac{v\gamma_1+\gamma_2(1-v)}{1-v(1-\gamma_1)}\right]} m(6 + s)^{\frac{1-\gamma_1}{1-\gamma}\left[\frac{v\gamma_1+\gamma_2(1-v)}{1-v(1-\gamma_1)}\right]} e^{(1-\gamma_1)\frac{vr+\delta_h}{1-v(1-\gamma_1)}s},$$

$$F_1(s) = (r + \delta_h)^{-\left(\frac{1}{1-\gamma}\right)\frac{\gamma_2-v(1-\gamma_1)}{1-v(1-\gamma_1)}} m(6 + s)^{\left(\frac{1}{1-\gamma}\right)\frac{\gamma_2-v(1-\gamma_1)}{1-v(1-\gamma_1)}} e^{-\frac{\gamma_2-v(1-\gamma_1)}{1-v(1-\gamma_1)}\frac{r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}s}(e^{\mu s} - 1),$$

where

$$\mu = \frac{\gamma_2 r + \delta_h(1 - \gamma_1)}{(1 - \gamma_2)},$$

and

$$(31a) \quad A_0 = \left[\left(\frac{h_B v^v}{z_h}\right)^{\frac{1-v}{1-\gamma}} \frac{(p_w)^{\frac{\gamma_2(1-v)}{1-\gamma}}}{(p_E)^v((1 - \tau)w)^{\frac{\gamma_2-v(1-\gamma_1)}{1-\gamma}}(\gamma_2^{\gamma_2}\gamma_1^{(1-\gamma_2)})^{\frac{(1-v)}{1-\gamma}}}\right]^{\frac{1-\gamma_1}{1-v(1-\gamma_1)}},$$

$$(31b) \quad A_1 = \frac{1 - \gamma_1}{\gamma_1\mu}(p_w p_s)^{\frac{\gamma_2}{1-\gamma_2}}.$$

The impact of changes in each element are indicated with a $+$ or a $-$ just below the relevant term.

$$(32) \qquad A_0 = \hat{A}_0\underset{(-)}{((1 - \tau)w,} \underset{(+)}{p_w,} \underset{(-)}{p_E,} \underset{(-)}{z_h,} \underset{(+)}{h_B}).$$

$$(33) \qquad A_1 = \hat{A}_1\underset{(+)}{(p_w,} \underset{(+)}{p_s}).$$

CONDITION (Interiority): $F_0(0) > A_0$.

PROOF OF PROPOSITION 1:

Existence of a solution is standard for this problem. Uniqueness follows from the fact that the objective function is linear and, given $\gamma < 1$, the constraint set is strictly convex. The functions F_i are continuous and differentiable in the interior, and

$$F_0(0) > 0, \quad F_0(R - 6) = 0, \quad F_0'(s) < 0,$$

and

$$F_1(0) = F_1(R - 6) = 0, \quad F_1(s) > 0 \text{ for } s \in (0, R - 6).$$

Define $F(s) \equiv F_0(s) - (A_0 + A_1 F_1(s))$. It follows that $F(s)$ is continuous and, if the Interiority condition is satisfied then $F(s) > 0$ and $F(R - 6) < 0$ and, hence, that there is an s^* that satisfies $F(s^*) = 0$. Since $F_0(s)$ is downward sloping, it must be the case that $A_1 F_1'(s^*) > F_0'(s^*)$ and that the function $A_0 + A_1 F_1(s)$ intersects $F_0(s)$ from below. These properties and the signs of the partial derivatives in the equations (32) and (33) imply the comparative statics results stated in the proposition.

Given the solution s^*, equation (27) gives the level of human capital at the end of the schooling period, $h(6 + s)$. Lastly, equations (26) and (29) can be used to determine q_E and h_E.

PROOF OF PROPOSITION 2:

The proposition requires that we hold s constant across economies and, hence, the same four equations (26)–(29) can be used if they are viewed as giving the appropriate values of $(h_E, h(6 + s), q_E)$ and z_h. Standard substitutions imply that the solution for $h(6 + s)$ is given by

$$h(6 + s) = G(s) \left(\frac{(1 - \tau)w}{p_E} \right)^{\frac{v}{1-v}},$$

where

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

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

APPENDIX B: MORE ON CALIBRATION

Early Childhood.—As additional evidence we used data from "Expenditures on Children by Families, 2011" (USDA 2012). For the average family, the sum of expenditures on health care, child care and education, and one-half of miscellaneous (includes personal care items, entertainment, and reading materials) amounts to approximately $4,167 per year, while output per worker was 92,247. This calculation yields an estimate of $x_E/y = 0.260$ while the calibrated value using a different approach is 0.237.

Schooling.—If capital share is β, our technology implies that the ultimate share of capital in the production of schooling is $\beta\gamma_2$, which, in our parameterization, is equal to $0.20 \times 0.40 = 0.08$. We note that the goods input share is calibrated to match the share of *purchased inputs* relative to GDP. The true share is likely to be significantly higher since purchased inputs exclude the value of parental time and resources.

Cumulative Effects.—Finally, we note that the human capital production function implies a form dynamic complementarity. The higher the human capital acquired early on, the higher the return associated with resources allocated to the production of future human capital.

REFERENCES

Banerjee, Abhijit V., and Esther Duflo. 2005. "Growth Theory Through the Lens of Development Economics." In *Handbook of Economic Growth.* Vol. 1A, edited by Philippe Aghion and Steven Durlauf, 473–552. Amsterdam: Elsevier B.V.

Barro, Robert J., and Jong-Wha Lee. 2010. "A New Data Set of Educational Attainment in the World, 1950–2010." National Bureau of Economic Research Working Paper 15902.

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* Chicago: University of Chicago Press.

Belzil, Christian, and Jorgen Hansen. 2002. "Unobserved Ability and the Return to Schooling." *Econometrica* 70 (5): 2075–91.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75 (4): 352–65.

Bils, Mark, and Peter J. Klenow. 2000. "Does Schooling Cause Growth?" *American Economic Review* 90 (5): 1160–83.

Browning, Martin, Lars Peter Hansen, and James J. Heckman. 1999. "Micro Data and General Equilibrium Models." In *Handbook of Macroeconomics.* Vol. 1A, edited by John B. Taylor and Michael Woodford, 543–633. Amsterdam: Elsevier B.V.

Case, Anne, and Motohiro Yogo. 1999. "Does School Quality Matter? Returns to Education and the Characteristics of Schools in South Africa." National Bureau of Economic Research Working Paper 7399.

Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences." In *Handbook of Economic Growth.* Vol. 1A, edited by Philippe Aghion and Steven Durlauf, 679–741. Amsterdam: Elsevier B.V.

Cordoba, Juan Carlos, and Marla Ripoll. 2013. "What Explains Schooling Differences across Countries?" *Journal of Monetary Economics* 60 (2): 184–202.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3): 883–931.

Erosa, Andres, Tatyana Koreshkova, and Diego Restuccia. 2010. "How Important Is Human Capital? A Quantitative Theory Assessment of World Income Inequality." *Review of Economic Studies* 77 (4): 1421–49.

French, Eric. 2005. "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour." *Review of Economic Studies* 72 (2): 395–427.

Glewwe, Paul, and Michael Kremer. 2006. "Schools, Teachers, and Education Outcomes in Developing Countries." In *Handbook of the Economics of Education.* Vol. 2, edited by Eric Hanushek and Finis Welch, 945–1017. Amsterdam: Elsevier B.V.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *The Quarterly Journal of Economics* 114 (1): 83–116.

Hanushek, Eric A. 2006. "School Resources." In *Handbook of the Economics of Education.* Vol. 2, edited by Eric Hanushek and Finis Welch, 865–908. Amsterdam: Elsevier B.V.

Hanushek, Eric A., and Ludger Woessmann. 2011. "The Economics of International Differences in Educational Achievement." In *Handbook of the Economics of Education.* Vol. 3, edited by Eric A. Hanushek, Stephen J. Machin, and Ludger Woessmann, 89–200. Amsterdam: Elsevier B.V.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2008. "Earnings Functions and Rates of Return." *Journal of Human Capital* 2 (1): 1–31.

Hsieh, Chang-Tai, and Peter J. Klenow. 2007. "Relative Prices and Relative Prosperity." *American Economic Review* 97 (3): 562–85.

Jorgenson, Dale W. 1988. "Productivity and Postwar U.S. Economic Growth." *Journal of Economic Perspectives* 2 (4): 23–41.

Karoly, Lynn A., M. Rebecca Kilburn, and Jill S. Cannon. 2005. *Early Childhood Interventions: Proven Results, Future Promise.* Santa Monica: RAND Corporation.

Klenow, Peter J., and Andrés Rodríguez-Clare. 1997. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" In *National Bureau of Economic Research Macroeconomics Annual 1997.* Vol. 12, edited by Ben S. Bernanke and Julio J. Rotemberg, 73–114. Cambridge: MIT Press.

Lagakos, David, Benjamin Moll, Tommaso Porzio, and Nancy Qian. 2013. "Experience Matters: Human Capital and Development Accounting." Unpublished.

Lucas, Robert E. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (1): 3–42.

Manuelli, Rodolfo E., and Ananth Seshadri. 2009. "Explaining International Fertility Differences." *Quarterly Journal of Economics* 124 (2): 771–807.

Manuelli, Rodolfo E., and Ananth Seshadri. 2014. "Human Capital and the Wealth of Nations: Dataset." *American Economic Review.* http://dx.doi.org/10.1257/aer.104.9.2736.

Mincer, Jacob A. 1974. *Schooling, Experience, and Earnings.* New York: Columbia University Press.

Parente, Stephen L., and Edward C. Prescott. 2000. *Barriers to Riches.* Cambridge, MA: MIT Press.

Schoellman, Todd. 2013. "Refugees and Early Childhood Human Capital." Unpublished.

Scholz, John Karl, Ananth Seshadri, and Surachai Khitatrarun. 2006. "Are Americans Saving 'Optimally' for Retirement?" *Journal of Political Economy* 114 (4): 607–43.

Solow, Robert M. 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70 (1): 65–94.

United Nations Educational, Scientific and Cultural Organization (**UNESCO**)**, Institute for Statistics.** 2007. *Global Education Digest 2007: Comparing Education Statistics Across the World.* Montreal: UNESCO Institute for Statistics.

United States Census Bureau. 2011. *The American Community Survey 2011.* United States Department of Commerce. http://www2.census.gov/acs2011_1yr/pums/csv-pus.zip (accessed December 2, 2013).

United States Department of Agriculture (USDA). 2012. *Expenditures on Children by Families, 2011.* Center for Nutrition Policy and Promotion Miscellaneous Publication 1528-2011.

Weiss, Elaine, and Richard Brandon. 2010. *The Economic Value of The U.S. Early Childhood Sector.* Washington, DC: The PEW Center on the States.