

Department of Economics

Working Paper Series

**How Important is Human Capital? A Quantitative Theory Assessment
of World Income Inequality**

09-007 | Andres Erosa
Instituto Madrileño de Estudios Avanzados
Tatyana Koreshkova
Concordia University
Diego Restuccia
University of Toronto



UNIVERSITÉ

Concordia

UNIVERSITY

How Important is Human Capital? A Quantitative Theory Assessment of World Income Inequality*

Andrés Erosa

University of Toronto and IMDEA

Tatyana Koreshkova

Concordia University and CIREQ

Diego Restuccia

University of Toronto

March 2009

Abstract

We develop a quantitative theory of human capital investments in order to evaluate the magnitude of cross-country differences in total factor productivity (TFP) that explains the variation in per-capita incomes across countries. We build a heterogeneous-agent economy with cross-sectional variation in ability, schooling, and expenditures on schooling quality. By embedding our analysis in a growth model with tradable and non-tradable sectors, we model sectorial productivity differences across countries, as documented in Hsieh and Klenow (2007). The parameters governing human capital production and random ability and taste processes are restricted by a set of cross-sectional data moments such as variances and intergenerational correlations of earnings and schooling, as well as slope coefficient and R^2 in a Mincer regression. Our main finding is that human capital accumulation strongly amplifies TFP differences across countries: To explain a 20-fold difference in the output per worker the model requires a 5-fold difference in the TFP of the tradable sector, versus an 18-fold difference if human capital is fixed across countries. Moreover, we find that sectorial productivity differences play a prominent role in quantitative implications of the theory.

*The authors would like to thank seminar and workshop participants at Queens University, Arizona State University, ASU Development Workshop, HEC Montreal, University of Montreal, Federal Reserve Bank of Richmond, 2006 NBER Summer Institute Growth Workshop, McGill University, University of Texas-Austin, University of Southern California, Bank of Canada Productivity Workshop, Concordia University, Carleton University, University of Tokyo, Seoul National University, Washington University in St. Louis, Yale University, University of Pennsylvania, 2007 SAET Conference. The paper has previously circulated under the title “On the Aggregate and Distributional Implications of Productivity Differences Across Countries.” *Contact Information:* Department of Economics, Concordia University, 1455 de Maisonneuve Bl. W, Montreal, Quebec H3G 1M8, Canada. E-mail: andres.erosa@imdea.org; tkoreshk@alcor.concordia.ca; and diego.restuccia@utoronto.ca.

1 Introduction

One of the most important challenges faced by economists is to explain the large observed differences in per-capita income across countries. In this paper, we develop a quantitative theory of human capital investments in order to evaluate the magnitude of cross-country differences in the total factor productivity (TFP) that explains the variation in per-capita income across countries. Building a quantitative theory allows us to circumvent two major problems faced by growth accounting exercises. First, to date, there are no reliable cross-country measures of the quality of schooling across countries. If this quality is positively associated with the level of economic development, the residual in growth accounting exercises overstates the cross-country differences in TFP. A second problem arises due to the (unobserved) covariance of TFP with measures of physical and human capital, which renders output variance decomposition difficult.

Our approach consists of developing a theory of human capital investments – schooling time and expenditures on schooling quality – that can be used to quantitatively assess the sources of cross-country income differences. It is well known that the quantitative implications of human capital theory hinge crucially on the value of the elasticity of human capital with respect to the expenditure on goods (see Trostel (1993) and Erosa and Koreshkova (2007)). The intuition is simple: If schooling requires only time inputs, a change in the wage rate affects equally the benefits and the costs of human capital accumulation, leaving the optimal level of human capital unchanged. On the other hand, when schooling requires only input of goods, an increase in the wage rate raises benefits but not the costs of schooling, hence increasing the optimal human capital stock. Therefore, the relative importance of time versus goods input determines the responsiveness of human capital to differences in the wage rate or TFP.¹ Developing a quantitative theory, in turn, is a challenging task due to the lack

¹Bils and Klenow (2000) point out that the production of human capital is more intensive in time input than the production of output goods. They and Klenow and Rodriguez-Clare (1997) argue that, by using a one-sector growth model, Mankiw, Romer, and Weil (1992) overstate the importance of goods input in the production of human capital and, thus, obtain results that understate TFP differences across countries.

of conclusive micro evidence on the expenditure elasticity of human capital: Some key human capital determinants, such as individual ability and private expenditures on education (including those outside of formal schooling), are not observed. We address this problem by building a theory of heterogeneous agents – in terms of ability, schooling tastes, and parental resources – where the parameters governing human capital accumulation have important implications for schooling and earnings inequality and intergenerational mobility *within* a country. Because the model generates a set of cross-sectional statistics, such as variances and intergenerational correlations of earnings and schooling, as well as slope coefficient and R^2 in a Mincer regression, that can be compared to actual data, we can use the restrictions of the theory and U.S. household data to pin down the key parameters – elasticities of human capital with respect to time and goods inputs – driving the quantitative implications of the theory. We then use the theory to study income inequality across countries and, in particular, to quantitatively assess how variations in TFP are *amplified* through human capital accumulation into larger differences in output per worker across countries.

Our analysis builds on a multi-sector growth model that allows for sectorial productivity differences across countries, as documented in Hsieh and Klenow (2007) and Herrendorf and Akos (2007). The cross-country experiments assume that, relative to the calibrated model economy, a one-percent reduction in the productivity of the manufacturing (tradable) sector is associated with an ε -percent reduction in the productivity of the service (nontradable) sector. When ε is set to 1, sectorial productivities vary uniformly, and there are no differences in the relative prices across countries (apart from the wage rate). When $\varepsilon < 1$, poor countries exhibit a low productivity in all sectors, but their productivity in services relative to manufacturing is high. Hence, relative prices vary across countries, with services being cheap in poor economies. Our baseline experiments minimize the role of human capital in amplifying income differences across countries by assuming that human capital investments only require services. When poor countries have a comparative advantage in the production of human capital inputs ($\varepsilon < 1$), their low aggregate productivity is not too detrimental

to human capital accumulation. Nonetheless, our main finding is that human capital accumulation strongly amplifies TFP differences across countries: The elasticity of output per worker – at PPP prices – with respect to TFP in the tradable sector is 1.94. This implies that a 5-fold difference in TFP explains a 20-fold difference in the output per worker, observed between the top 10 percent and bottom 10 percent of countries in the world income distribution. In contrast, without human capital accumulation, an 18-fold difference in the TFP of the tradable sector is required to account for the income difference between rich and poor countries.

The mechanism generating the large income disparity in our model economy is that a low TFP in poor countries leads individuals to invest few resources in accumulating both physical and human capital relative to individuals in rich countries. Human capital is an important source of income differences across countries, not only because it directly contributes to cross-country output differences, but also because a higher human capital stock stimulates physical capital accumulation by raising the marginal product of capital.

At the theoretical level, it is interesting to answer the following question: How does a one-percent change in TFP in all sectors in the economy affect output per worker? The answer is provided by the one-sector version of the model economy ($\varepsilon = 1$), and it is startling: The TFP elasticity of output increases to 2.8. In a world where TFP varies uniformly across sectors, the TFP ratio needed to generate a PPP-income ratio of a factor of 20 would be only 2.9, which is about two thirds of what is implied by the model parameterization consistent with productivity estimates in the data ($\varepsilon \in [.3, .4]$). Nevertheless, from a development accounting view, the relevant amplification effect is the one estimated with the two-sector model as the evidence suggests that TFP does not vary uniformly across sectors. We emphasize that the amplification role of human capital is large even for implausibly low values of ε . When $\varepsilon = .1$ the TFP elasticity of PPP output is 1.53 in the model with human capital accumulation and .86 in the economy with exogenous human capital. This differential in TFP elasticities across model economies is not minor: To generate an income ratio of 20, the economy with

endogenous human capital requires a 7-fold difference in the TFP of the tradable sector while the economy with exogenous human capital requires a 33-fold difference. We conclude that it is important to model both human capital and sectorial productivity differences for assessing the cross-country variation in productivity.

The paper proceeds as follows. The next section describes in detail the economic environment. In section 3, we consider a simple version of the model in order to illustrate the main features of our theory driving human capital investments and to motivate our calibration strategy. Section 4 lays out the calibration strategy for the benchmark economy and shows that the model economy is consistent with several dimensions of heterogeneity in the data that were not targeted in the calibration. Furthermore, the predictions of the benchmark economy are tested using results from the micro literature on the enrollment effects of college tuition changes. In section 5, we evaluate the aggregate impact of TFP differences across countries, perform a sensitivity analysis, and examine the predictions of the theory for the variation in relative prices across countries. Finally, we compare our findings to related papers in the literature. Section 6 concludes.

2 Economic Environment

We consider an economy populated by overlapping generations of people who are altruistic toward their descendants. People are heterogeneous in skills and physical assets and face idiosyncratic (uninsurable) uncertainty about their labor earnings. Investments in human capital involve children's time and expenditures by parents that affect the quality of the human capital of their children. Parents cannot borrow to finance investment in human capital. Since the analysis in this paper focuses on steady states, time subscripts are omitted in the description of the model and use a prime to indicate the next period value of a given variable.

Demographic Structure There is a large number of dynasties (mass one). Individuals live for three periods, so that the model period is set to 20 years. An individual is referred to as a child in the first period of his life (real age 6-26 years), a young parent in the second period (real age 26-46 years), and an old parent in the third period of his life (real age 46-66 years).² A household is composed of 3 people: old parent, young parent, and a child.

Production Technologies We assume that production takes place in two sectors – manufacturing and services – with the following technologies:

$$Y_M = A_M K_M^\alpha H_M^{1-\alpha}, \quad (1)$$

$$Y_S = A_S K_S^\alpha H_S^{1-\alpha} \quad (2)$$

where Y_M and Y_S denote the output of the manufacturing and service sectors; K_i and H_i represent the services of physical and human capital used in sector $i \in \{M, S\}$. Parameter $\alpha \in (0, 1)$ is the elasticity of output with respect to physical capital and is assumed to be equal across sectors. Parameter A_i , $i \in \{M, S\}$, represents a sectorial TFP, which is allowed to vary across sectors.

Manufacturing output can be consumed (C_M), invested in physical capital (X), and invested in human capital (E_M). Services can be either consumed (C_S) or invested in human capital (E_S). Feasibility requires

$$C_M + X + E_M = Y_M,$$

$$C_S + E_S = Y_S.$$

²In an earlier version of the paper, we modeled a retirement period. Since it did not affect the quantitative implications of the theory, we decided to abstract from retirement in the current version of the paper.

Physical capital is accumulated according to

$$K' = (1 - \delta)K + X,$$

where investment goods X are produced in the manufacturing sector.

We model human capital investments that take place ‘early’ in the life of an individual and that include schooling as well as investments outside of formal schooling. Consistent with the view of Kendrick (1976), Becker and Tomes (1986), Haveman and Wolfe (1995), Neal and Johnson (1996), Mulligan (1997), Keane and Wolpin (1997, 2001), among many others, we think that households invest a lot of resources in their children outside of school (health, food, shelter, books, recreational activities and extracurricular educational activities). This view motivates our focus on a broad notion of human capital investments. The human capital of a child is produced with the inputs of schooling time ($s \in [0, 1]$) and expenditures in human capital quality ($e > 0$) according to the following production function:

$$h_c = A_H z (s^\eta e^{1-\eta})^\xi \quad \eta, \xi \in [0, 1], \quad (3)$$

A unit of schooling time (quantity of schooling) is produced with one unit of a child’s time and \bar{l} units of market human capital services. In other words, schooling requires own time and human capital purchased in the market.³ Educational expenditures in quality are assumed to be a composite of manufacturing goods and services

$$e = e_M^\phi e_S^{1-\phi}, \quad \phi \in [0, 1].$$

To model heterogeneity across individuals, we follow the micro literature in allowing individuals to differ in terms of their ability z and their taste for schooling θ . We assume that the shocks to z and θ are idiosyncratic to each dynasty and that they are observed

³Schooling s is a Leontief function of own time t and market human capital services h_s : $s = \min\{t, \frac{h_s}{\bar{l}}\}$.

at the beginning of the period, that is, before human capital investments take place. The ability z is transmitted across generations according to a discrete Markov transition matrix $Q(z, z')$, where $q_{i,j} = Pr(z' = z_i | z = z_j)$. The taste shock θ is *iid* across individuals and, possibly, correlated with the current realization of the ability shock z . The distribution of the taste shock is thus described by a discrete matrix $Q_\theta(z, \theta)$.

The parameter A_H in the human capital production function (3) is common across all individuals in the economy and is normalized to 1 in our baseline economy. Later, in some quantitative experiments, we consider cross-country variation in the efficiency of the human capital technology by allowing the parameter A_H to vary across economies (Klenow and Rodriguez-Clare (2001) allow for the possibility that countries differ in the productivity of the education sector).

Preferences A per-period utility function of the household is

$$\frac{1}{1-\sigma} \left((C_M)^\gamma (C_S)^{1-\gamma} \right)^{1-\sigma} + v(s, \theta),$$

where C_M represents consumption of manufacturing goods and C_S consumption of services. The term $v(s, \theta)$ represents utility of schooling, where θ is a taste shock that varies across individuals. Thus, consistent with the micro literature on schooling (see, for instance, Keane and Wolpin (2001) and Card (2001)), in our model heterogeneity in schooling decisions is driven by variation not only in parental wealth and labor market returns (ability) but also in schooling tastes.

Market Structure and Relative Prices We assume competitive markets for factor inputs and outputs. Firms in the manufacturing and service sectors maximize

$$\pi_M = Y_M - wH_M - (r + \delta)K_M, \tag{4}$$

$$\pi_S = P_S Y_S - wH_S - (r + \delta)K_S, \tag{5}$$

where we have set the price of manufacturing goods to 1 (numeraire). Profit maximization in the manufacturing and services sectors imply

$$p_S = \frac{A_M}{A_S}. \quad (6)$$

The division of a total expenditure E between e_M and e_S to produce the composite input e in the human capital technology is a static problem

$$\begin{aligned} e(E) &\equiv \max_{e_N, e_T} (e_M)^\phi (e_S)^{1-\phi} \\ \text{s.t. } p_S e_S + e_M &= E. \end{aligned}$$

The optimal choices are $e_M(E) = \phi E$, $e_S(E) = (1 - \phi) \frac{E}{p_S}$, $e(E) = \phi^\phi (1 - \phi)^{1-\phi} (p_S)^{1-\phi} E$.

Setting $\hat{\phi} \equiv \frac{1}{\phi^\phi (1-\phi)^{1-\phi}}$ and

$$p_e \equiv \hat{\phi} (p_S)^{1-\phi}, \quad (7)$$

expenditures in education satisfy $E = \frac{(p_S)^{1-\phi}}{\phi^\phi (1-\phi)^{1-\phi}} e = p_e e$. Note that the price of the composite education input e increases with the relative price of services.

Similarly, the optimal allocation of total expenditure C between consumption of manufacturing goods C_M and services C_S solve the static problem

$$\begin{aligned} c(C) &\equiv \max_{C_N, C_T} (C_M)^{\gamma_j} (C_S)^{1-\gamma_j} \\ \text{s.t. } p_S C_S + C_M &= C. \end{aligned}$$

Optimal behavior implies that total consumption expenditures $C = \frac{(p_S)^{1-\gamma_j}}{\gamma_j^{\gamma_j} (1-\gamma_j)^{1-\gamma_j}} c = p_c c$, where $p_c \equiv \frac{(p_S)^{1-\gamma_j}}{\gamma_j^{\gamma_j} (1-\gamma_j)^{1-\gamma_j}}$.

Public Education Since our calibration strategy is to use cross-sectional heterogeneity within a country to restrict the parameters governing human capital accumulation, we cannot abstract from the role of public education on education and labor market outcomes. We

model public education by assuming that education expenditures are subsidized at the rate p per unit of schooling time. These expenditures are financed with a proportional tax τ on households' income. Public and private expenditures are perfect substitutes in the production of human capital.

Decision Problem of the Household All decisions of the household are made by the young parent. The state of a young parent is given by a quadruple (q, h_p, z, θ) : resources of the old parent q , human capital of the young parent h_p , child's ability z , and child's taste for schooling θ . Households face uncertainty over the realization in future periods of ability, school taste and market luck, hence, they maximize the expected discounted lifetime utility of all generations in the dynasty. Young parents choose consumption c , assets a' , time spent in school by their children s (where $1 - s$ is the working time of the children), and resources spent on the quality of education of their children e . A parent who provides his child with s years of schooling and a quality of education e incurs expenditures $P_e e + (w\bar{l} - p)s$, where $w\bar{l}$ is the cost of market human capital services per year of education, and p denotes public education expenditures (or subsidies) per year of education.

The decision problem of a young household can be written using the dynamic programming language as follows:

$$V(q, h_p, z, \theta) = \max_{c, e, s, h', a} \left\{ U(c) + v(s, \theta) + \beta \sum_{z', \theta', \mu'} Q_z(z, z') Q_\theta(z', \theta') Q_\mu(\mu') V(q', \mu' h_c, z', \theta') \right\}, \quad (8)$$

subject to

$$\begin{aligned} P_c c + P_e e + (w\bar{l} - p)s + a &= (1 - \tau)w [\psi_2 h_p + \psi_1 h_c (1 - s)] + q, \\ h_c &= A_H z (s^\eta e^{1-\eta})^\xi, \\ q' &= (1 - \tau) [w\psi_3 h_p + ra] + a \\ a &\geq 0, \quad s \in [0, 1], \end{aligned}$$

where (ψ_1, ψ_2, ψ_3) are life-cycle productivity parameters. The first two terms in the objective function are the current period utility and the third term is the expected discounted future utility. The expectations of the next period's value function is taken over the market luck of the current child μ' and over the ability z' and school taste θ' of the child born in the next period. The first constraint is the household budget constraint, where the right hand side is given by the sum of the earnings of the young parent and the child upon finishing school $(1 - \tau)w[\psi_2 h_p + \psi_1 h_c(1 - s)]$ and the resources q (after tax earnings and gross asset return) brought to the current household by the old parent. The third constraint defines the parental wealth q' of the next household in the dynasty line.

We emphasize that when young parents make education decisions for their children, they know the ability z and the taste for schooling θ . However, they face uncertainty regarding the market luck of their children μ' , which is realized in the adult stage of the individual's life cycle. The human capital h_c of an individual at the end of period 1 evolves stochastically, according to a realization of a market luck shock μ' : $h_p = \mu' h_c$, where μ' is *iid* across individuals and time according to a density $Q_\mu(\mu)$ with a mean equal to 1. We assume that markets are imperfect in that households cannot perfectly insure against labor-market risk and the human-capital shocks affecting their descendants. Moreover, individuals cannot borrow.

There are many sources of heterogeneity in parental investment decisions for human capital: First, children differ in ability z and taste for schooling θ . Second, parents differ in their asset holdings and human capital. Parental resources play an important role because incomplete markets, together with a taste for schooling, imply that human capital investment decisions do not maximize expected lifetime income. Summing up, our baseline economy has many factors inducing heterogeneity in schooling and earnings.

The theory abstracts from on-the-job human capital accumulation. This assumption is motivated by tractability reasons as well as some empirical evidence across countries. Using the coefficients for returns to experience for each country reported in Bils and Klenow (2000),

we found that the earnings of a worker with 20 years of experience relative to a worker with 10 years of experience is not systematically related to the level of per-capita income across countries. In fact, we found a small negative correlation between returns to labor market experience (wage growth) and per-capita income across countries, which suggests that on-the-job investments in human capital are not likely to be an important source of income differences across countries.

3 Human Capital Investments in a Complete Markets Environment

This section provides some analytical results that shed light on how the parameters of the human capital technology determine the quantitative implications of the theory. To study a simplified version of the model economy, we assume complete markets and abstract from tastes for schooling. As a result, human capital investment decisions are independent of consumption decisions and maximize lifetime income. We show that the quantitative implications of the theory for income inequality – within and across countries – depend crucially on the expenditure elasticity of human capital. We also study how cross-country differences in sectorial productivities generate heterogeneity in relative prices and human capital investments.

3.1 Human capital investments across individuals and countries

Consider a world with a large number of countries. Each country is populated by measure 1 of dynasties and by a vector of prices (w, p_e) that varies across countries. Capital markets are assumed to be perfect so that in equilibrium individuals make efficient investments in human capital. The attention is confined to the steady-state analysis. The equilibrium interest rate is given by the individuals' rate of time preference ρ . Although the theory

makes no predictions for the distribution of income, consumption, and wealth, it does have important implications for the variation of schooling and earnings across individuals and countries.

3.1.1 The decision problem

We analyze how variation in wages and variation in ability (z) lead to heterogeneous human capital investments across countries (derive macro elasticities) and across individuals (derive micro elasticities). The goal is to isolate the effects of the parameters of the human capital technology on micro and macro elasticities in our model.

Consider the decision problem of an individual with ability z in a country with a wage rate w and a price of education goods p_e , where these prices are expressed in terms of the manufactured good. The human capital investment decision can be formulated as choosing schooling time (s) and expenditures (e) to maximize the present value of the lifetime earnings net of the education costs:

$$\max_{e,s,h} \{w(1-s)h\psi_1 + wh\Psi - p_e e - w\bar{l}s\} \quad (9)$$

$$s.t. \quad h = A_H z (s^\eta e^{1-\eta})^\xi, \quad (10)$$

where $\Psi = \sum_{i=2}^3 \beta^{i-1}\psi_i$ with ψ_i representing the life cycle productivity parameters described in the previous section, and $\beta = \frac{1}{1+r}$ provided $r = \rho$. The cost of schooling includes expenditures in human capital quality (e), time-purchases on the market (tuition costs) per unit of schooling time ($w\bar{l}$), and foregone earnings in the first period of life ($sw h\psi_0$).

Assuming an interior solution, the corresponding first-order conditions are:

$$-wh\psi_0 + wh_s [(1-s)\psi_1 + \Psi] = w\bar{l}, \quad (11)$$

$$wh_e [(1-s)\psi_1 + \Psi] = p_e, \quad (12)$$

where $h_s = \frac{h}{s}\eta\xi$ and $h_e = \frac{h}{e}(1-\eta)\xi$. These equations can be expressed as

$$A_H z (s^\eta e^{1-\eta})^\xi \left\{ -\psi_0 + \frac{\eta\xi}{s} [(1-s)\psi_1 + \Psi] \right\} = \bar{l} \quad (13)$$

$$e = \left\{ \frac{w}{p_e} A_H z (1-\eta)\xi [(1-s)\psi_1 + \Psi] s^{\eta\xi} \right\}^{\frac{1}{1-(1-\eta)\xi}}. \quad (14)$$

We are ready to analyze how individual decisions depend on the parameters of the maximization problem. In the absence of tuition costs ($\bar{l} = 0$), it is easy to solve for s from (13) and verify that the optimal quantity of schooling does not vary across individuals with different values of z . Intuitively, when there are no tuition costs of schooling ($\bar{l} = 0$) a change in z raises proportionally the benefits and costs of schooling and has no effect on the optimal choice of years of schooling. Moreover, when $\bar{l} = 0$, there is no variation in schooling across countries (w, p_e, A_H). We thus maintain $\bar{l} > 0$. Similarly, in the absence of education expenditures ($\eta = 1$), the quality of schooling does not vary across individuals and countries. On the contrary, when $0 < \eta < 1$, equations (13) and (14) imply that both quantity and quality of schooling vary across individuals (z) and countries (w, p_e, A_H).

Proposition 1: *The theory requires $\bar{l} > 0$ and $0 < \eta < 1$ in order to generate differences in the quantity and quality of schooling across individuals (z, η) and countries (w, p_e, A_H).*

3.1.2 Micro-elasticities

Having characterized individual decisions, we can examine the effects of the parameters of the human capital technology on variation in schooling and earnings across individuals and countries. To gain insights with simple algebra, it is convenient to set $\psi_1 = 0$. Combining (13) and (14), taking logs, and differentiating with respect to $\ln z$, gives an expression for the individual (ability) elasticity of schooling:

$$E_{sz} \equiv \frac{\partial \ln(s)}{\partial \ln z} = \frac{1}{1-\xi}. \quad (15)$$

The elasticity of expenditures with respect to ability is obtained by differentiating (14) with respect to $\ln z$:

$$E_{ez} \equiv \frac{\partial \ln(e)}{\partial \ln z} = \frac{1}{1 - (1 - \eta)\xi} \left(1 + \eta\xi \frac{\partial \ln(s)}{\partial \ln z} \right) = \frac{1}{1 - \xi}, \quad (16)$$

where the last equality used (15). The elasticity of human capital with respect to the ability, is obtained by log-differentiating (10) with respect to $\ln z$ and by using (15) and (16):

$$E_{hz} \equiv \frac{1}{1 - \xi}.$$

Taking stock of the above findings, we note that the elasticities of schooling and human capital with respect to ability are the same. The magnitude of this elasticity is determined by the returns to scale in the human capital accumulation technology (parameter ξ). For a given distribution of z in the population, the variations in both schooling and human capital increase with the returns-to-scale parameter ξ . Hence this parameter is important for the predictions of the theory on the cross-sectional inequality in schooling and earnings.

3.1.3 Macro-elasticities

Combining (13) and (14), taking logs, and differentiating with respect to $\ln w$, we obtain the cross-country (wage) elasticity of schooling:

$$E_{sw} \equiv \frac{\partial \ln(s)}{\partial \ln w} = \frac{(1 - \eta)\xi}{1 - \xi}. \quad (17)$$

Using (17), the variation in the schooling quality across countries (w) satisfies

$$E_{ew} \equiv \frac{\partial \ln(e)}{\partial \ln w} = \frac{1}{1 - (1 - \eta)\xi} \left(1 + \eta\xi \frac{(1 - \eta)\xi}{1 - \xi} \right). \quad (18)$$

Log-differentiating (10) with respect to $\ln w$, together with (17) and (18), gives

$$E_{hw} \equiv \frac{(1-\eta)\xi}{1-\xi}. \quad (19)$$

Since E_{hw} does not vary across individuals, the aggregation is trivial: If two countries differ in TFP by a ratio A_R , then their ratio of aggregate human capital is: $H_R = \int (W_R)^{E_{hA}} dG_z = (W_R)^{E_{hA}}$. We conclude $E_{Hw} = E_{hw} = \frac{(1-\eta)\xi}{1-\xi}$. Table 1 summarizes the mapping from the model parameters into the micro and macro elasticities.

Table 1: Elasticities for the Deterministic Model

Macro	
$E_{Hw} = E_{hw} = E_{sw}$	$\frac{(1-\eta)\xi}{1-\xi}$
E_{ew}	$\frac{1}{1-(1-\eta)\xi} \left(1 + \frac{\xi^2 \eta (1-\eta)}{1-\xi}\right)$
$E_{x,pe} = -E_{x,w}$	for $x \in \{s, e, h\}$
$E_{x,\frac{w}{pe}} = E_{x,w}$	
$E_{x,A_H} = E_{x,z}$	
Micro	
$E_{sz} = E_{hz} = E_{ez}$	$\frac{1}{1-\xi}$

We are now ready to explore the sensitivity of the wage elasticity of human capital to the parameters of the human capital technology. Since E_{hw} increases with the returns-to-scale parameter ξ and decreases with the time share parameter η , E_{hw} increases with the expenditure elasticity of human capital and is maximized when $\eta = 0$ and $\xi = 1$. As the time-share parameter η decreases from 1 to 0, E_{Hw} takes values in the interval $[0, \frac{\xi}{1-\xi}]$. For instance, when $\xi = .9$, E_{Hw} takes values between 0 and 9, depending on the time-share parameter. In other words, a wage ratio of 3 can generate differences in human capital per worker anywhere from a factor of 0 to 20 thousand.⁴

⁴While E_{hw} is determined both by η and ξ , note that the expenditure elasticity alone provides a lower bound to the amplification effect. This is because $\eta \geq 0$ implies $(1-\eta)\xi \leq \xi$, which together with (19) and $\xi < 1$ implies that $E_{hw} \in \left[\frac{(1-\eta)\xi}{1-(1-\eta)\xi}, \infty\right)$. On the other hand, the parameter ξ implies an upper bound for E_{hw} since E_{hw} varies from 0 to $\frac{\xi}{1-\xi}$ for all feasible values of η .

Proposition 2: *The amplification effect of human capital, given by (19), depends crucially on the expenditure elasticity of human capital $(1 - \eta) \xi$. In particular, if the expenditure share is zero ($\eta = 1$), then human capital does not amplify TFP differences across countries, no matter how close ξ is to 1.*

Countries might not only differ in wage rates but on the relative price of education-goods (p_e) and the efficiency parameter in the human capital technology (A_H). An inspection of the individual's optimization conditions (13) and (14) implies that the elasticities of the variable $x \in \{s, e, h\}$ with respect to p_e and A_H are

$$\begin{aligned} E_{x,p_e} &= -E_{x,w}, \\ E_{x,\frac{w}{p_e}} &= E_{x,w} \\ E_{x,A_H} &= E_{x,z}. \end{aligned}$$

Note that human capital investment decisions are determined by the ratio of the wage rate to the price of education goods. Moreover, an increase in the efficiency of the human capital accumulation technology is tantamount to a change in the distribution of z .

3.2 Amplification with cross-country variation in sectorial productivities

We have shown that cross-country differences in relative prices (w, p_e) generate a cross-country variation in human capital investments. The next step is to analyze how cross-country differences in sectorial productivities generate heterogeneity in prices across countries and, hence, in human capital investments. To make progress, assume that

$$A_S^j = (A_M^j)^\varepsilon, \text{ for all } j, \text{ with } A_S^{US} = A_M^{US} = 1 \text{ and } \varepsilon < 1. \quad (20)$$

This assumption implies that a 1 percent change in the TFP of the manufacturing sector is associated with an ε -percentage change in the TFP of the service. Note that the case $\varepsilon = 1$ corresponds to the standard one-sector growth model with no variation in relative prices across countries. Using (6), it follows that

$$P_s^j = (A_M^j)^{1-\varepsilon}$$

so that $\varepsilon < 1$ implies that services are cheaper, in terms of manufacturing goods, in poor countries than in rich countries. The findings in Hsieh and Klenow (2007) suggest that $\varepsilon = 1/3$ gives a reasonably good approximation of the cross-country data on relative prices. We also assume that there are no cross-country differences in the efficiency of the human capital technology: $A_H^j = 1$ for all j . This assumption will be relaxed in the next section.

The firms' first-order conditions in the service sector imply

$$w^j = P_S^j (1 - \alpha) A_S^j \left(\frac{K_S^j}{H_S^j} \right)^\alpha, \quad (21)$$

$$R^j = \rho + \delta = P_S^j \alpha A_S^j \left(\frac{K_S^j}{H_S^j} \right)^{\alpha-1}. \quad (22)$$

Solving for the capital-to-labor ratio in the service sector gives

$$\frac{K_S^j}{H_S^j} = \left\{ \frac{P_S^j \alpha A_S^j}{\rho + \delta} \right\}^{\frac{1}{1-\alpha}}. \quad (23)$$

Combine (6), (21), (23) to obtain

$$w^j = c_w (A_M^j)^{\frac{1}{1-\alpha}}, \quad (24)$$

where $c_w = (1 - \alpha) \left(\frac{\alpha}{R} \right)^{\frac{\alpha}{1-\alpha}}$ does not depend on j . The real wage rate, measured in terms of manufacturing goods, increases with the TFP of the manufacturing sector.

The price of the composite education input in terms of the TFP of the manufacturing sector is obtained using (6), (7), and (20):

$$P_e^j = \widehat{\phi} (A_M^j)^{(1-\varepsilon)(1-\phi)} \quad (25)$$

From $\varepsilon < 1$ and $\phi \in (0, 1)$, it follows that the price of the composite education input good is cheaper (in terms of manufacturing goods) in countries with low productivity in manufacturing (A_M^j).

We are now ready to focus our attention on the price ratio $\frac{w^j}{p_e^j}$, which drives the variation in human capital investments across countries. Using (24) and (25), gives the following expression:

$$\frac{w^j}{p_e^j} = c_{wp} (A_M^j)^{\frac{1}{1-\alpha} - (1-\varepsilon)(1-\phi)},$$

where c_{wp} is constant across countries. Hence, the A_M -elasticity of $\frac{w}{p_e}$ is

$$E_{(w/p_e)A_M} = \frac{1}{1-\alpha} - (1-\varepsilon)(1-\phi).$$

In a one-sector growth model ($\varepsilon = 1$), this elasticity is equal to $\frac{1}{1-\alpha}$. When $\varepsilon < 1$, the cross-country variation in the relative price $\frac{w^j}{p_e^j}$ is lower than in the one sector growth model ($\varepsilon = 1$). This is quite intuitive: When $\varepsilon < 1$, poor countries are very inefficient in producing manufacturing goods, but they are not so inefficient in producing services. Services are cheap in poor countries because these countries have a high TFP in this sector relative to manufacturing. Since services are an input in a composite education good, the real wage – expressed in terms of the composite education good – does not increase with per-capita income across countries as much as in the one-sector growth model. It is also intuitive that the elasticity $E_{w/p_e, A_M}$ decreases both with ε and ϕ . That is, the elasticity is lower the higher the comparative advantage of poor countries is in producing education goods, and the lower the share of manufacturing goods is in the composite input e .

The amplifier effect of TFP differences in the manufacturing sector on human capital differences across countries is given by

$$\begin{aligned} E_{hA_M} &= E_{h,w/pe} E_{w/pe,A_M} \\ &= \frac{(1-\eta)\xi}{1-\xi} \left(\frac{1}{1-\alpha} - (1-\varepsilon)(1-\phi) \right). \end{aligned}$$

The fact that human capital production requires some services ($\phi < 1$) makes human capital less sensitive to a reduction of TFP in the manufacturing sector.

In a similar manner, we obtain

$$\begin{aligned} E_{sA_M} &= E_{s,w/pe} E_{w/pe,A_M} \\ &= \frac{(1-\eta)\xi}{1-\xi} \left(\frac{1}{1-\alpha} - (1-\varepsilon)(1-\phi) \right), \end{aligned}$$

and

$$\begin{aligned} E_{eA_M} &= E_{e,w/pe} E_{w/pe,A_M} \\ &= \frac{1}{1-(1-\eta)\xi} \left(1 + \eta\xi \frac{(1-\eta)\xi}{1-\xi} \right) \cdot \left(\frac{1}{1-\alpha} - (1-\varepsilon)(1-\phi) \right). \end{aligned}$$

Proposition 3: *Assume that countries differ in their relative productivities across sectors (for all j : $A_S^j = (A_M^j)^\varepsilon$ with $\varepsilon < 1$). The amplification effect of human capital accumulation is driven by the expenditure elasticity of human capital as in the one-sector growth model. The quantitative response to a change in the TFP in the manufacturing sector diminishes with the extent of a comparative advantage in producing services (a reduction in ε) and with the importance of services in the composite education goods input (a reduction in ϕ).*

4 Calibration

As discussed in the previous section, the aggregate implications of TFP differences across countries in our model hinge on the parameters determining human capital accumulation. Our calibration strategy is to restrict these parameters using cross-sectional heterogeneity of schooling and earnings in the data for the United States.

4.1 Parameters and Targets

We calibrate our benchmark economy (B.E.) to data for the United States. We normalize the units in which output is measured so that $A_S = A_M = 1$. The calibration of the baseline economy does not require taking an explicit stand on the shares of manufactured goods in consumption expenditures (γ) and in human capital expenditures (ϕ). In particular, we calibrate a one-sector economy with no manufacturing sector ($\gamma = \phi = 0$). It is easy to show that for any fixed γ and $\phi \in (0, 1)$, the two-sector model economy delivers, after an appropriate normalization of the distribution of ability and of the distribution of taste for schooling, the same equilibrium statistics as the calibrated one-sector model economy. Hence, without loss of generality, we proceed by calibrating a one-sector model economy. The parameters γ and ϕ will have important consequences for the cross-country experiments in the next section of the paper, and their values will be determined later.⁵

The mapping between model parameters and targeted data moments is multidimensional, and we thus solve for parameter values jointly. The discussion of calibration is divided into two parts: first, we discuss parameters that relate to preferences, demographics, and the production of goods, and second – parameters that relate to human capital accumulation.

A summary of parameter values and data targets is provided in Table 2.

⁵Letting c_1 and e_1 denote expenditures in consumption and human capital in a one-sector model, an equivalent two-sector growth model can be constructed as follows: Define the quantity of consumption c_2 and human capital (composite) input e_2 so that $c_1 = p_c c_2$ and $e_1 = p_e e_2$, for $p_c = \hat{\gamma}$ and $p_e = \hat{\phi}$. Then, normalize the distribution of ability and the taste shock in the two-sector model as follows: $z_2 = z_1(p_e)^{(1-\eta)\varepsilon}$ and $\theta_2 = \theta_1/(p_c)^{1-\sigma}$. This insures that all the equilibrium statistics are identical across the one-sector and two-sector model economies.

Table 2: Parameters and Data Targets

Parameter		Value	Target	U.S.	B.E.
<i>Consumption Preferences</i>					
CRRA	σ	2	Empirical literature	—	—
Discount factor	$\beta^{1/20}$	0.9646	Interest rate (%)	5	5
<i>Goods/Services Technologies</i>					
Capital share	α	0.33	Capital income share	0.33	0.33
Annual depreciation	δ	0.0745	Investment to output ratio	0.2	0.2
<i>Human Capital Technology</i>					
Schooling cost	\bar{l}	2.65	Educ. inst. salaries (% of GDP)	5	5
H.C. RTS	ξ	1.00	Variance of fixed effects	0.67	0.67
H.C. time share	η	0.6	Correlation of schooling	0.46	0.48
<i>Tastes for Schooling</i>					
Low	θ_L	0.00075	Mean years of schooling	12.6	12.6
High	θ_H	0.01285	R^2 in Mincer regression	0.22	0.21
Ability-taste corr. control	b	1.09	Mincer return	0.1	0.1
Ability std	σ_z	0.23	Variance of schooling	8.5	8.3
Ability correlation	ρ_z	0.78	Correlation of earnings	0.5	0.49
Market luck std	σ_μ	0.375	Variance of earnings	0.36	0.38
Tax rate on income	τ	0.044	Public Education (% of GDP)	3.9	3.9

Preferences, Demographics, and Production of Goods We set the relative-risk-aversion parameter σ to 2. There is no direct empirical counterpart for this parameter in the empirical literature since our model period is 20 years, and there is an infinite intertemporal substitution of consumption within a period. However, we consider a value of σ that is in the range of values considered in quantitative studies with heterogeneous agents.⁶ The discount factor β is set to target an annual interest rate of 5 percent, which is roughly the return on capital in the U.S. economy as measured by the average return on non-financial corporate capital net of taxes in 1990-96 (Poterba (1997)). The capital-share parameter is set to 0.33, consistent with the capital-income share in the U.S. economy from the National Income and Products Accounts. The depreciation rate δ is selected to match an investment-to-output ratio of 20 percent as documented in the Economic Report of the President (2004).⁷

⁶See Keane and Wolpin (2001) and Restuccia and Urrutia (2004) for discussions of these estimates.

⁷A similar target is obtained using the average of the investment-to-output ratio in the PWT6.1 for the period 1990 to 1996 (Heston, Summers, and Aten (2002)).

Human Capital Accumulation Recall that the human capital technology is given by $h_c = z(s^\eta e^{1-\eta})^\xi$, where s denotes schooling time and e denotes educational expenditures. Thus, we need to specify two elasticity parameters, ξ and η . Ability follows an $AR(1)$ process (in logs):

$$\ln(z') = \rho_z \ln(z) + \epsilon_z, \quad \epsilon_z \sim N(0, \sigma_z^2).$$

In our computations, we approximate this stochastic process with a discrete first-order Markov chain that takes 7 possible values for ability z , using the procedure in Tauchen (1986) to compute transition probabilities. The market luck μ is *iid* according to $\ln(\mu) \sim N(0, \sigma_\mu^2)$, approximated over 5 values similarly to z .

On the cost side, human capital accumulation is affected by the schooling cost \bar{l} and the public education subsidy p . The latter is determined by the tax rate on income τ in equilibrium. On the preferences side, human capital investments depend on the tastes for schooling. The functional form $v(s, \theta)$ specified for the utility of schooling allows for a diminishing marginal utility from schooling and a bounded marginal utility from schooling at zero level of schooling:

$$v(s, \theta) = \theta [1 - \exp\{-s\}]$$

where $\theta \in \{\theta_L, \theta_H\}$. To allow for tastes for schooling to be correlated with ability, we let the probability of the high-taste shock to increase with ability: $Prob(\theta_H | \ln(z)) = \min\{0.5 + b \ln z, 1\}$.⁸ Note that $b > 0$ implies that taste for schooling and ability are positively correlated. Thus, three parameters need to be specified for the tastes of schooling: two values for schooling tastes $\{\theta_L, \theta_H\}$ and parameter b , governing the correlation of the abilities and schooling tastes.

To sum up, there are ten parameters determining human capital accumulation:

$$\{\xi, \eta, \rho_z, \sigma_z, \sigma_\mu, \bar{l}, \tau, \theta_L, \theta_H, b\}.$$

These parameters are calibrated so that in equilibrium the model economy matches the

⁸Ability is drawn first, then the schooling taste is determined.

following ten targets from the U.S. data:

1. Intergenerational correlation of log-earnings of 0.5 (Mulligan (1997); see also excellent surveys of the empirical literature on the intergenerational correlation of earnings by Stokey (1998) and Solon (1999)).
2. Variance of log of permanent earnings of 0.36 (Mulligan (1997, 1999)).
3. Average years of schooling of 12.63, computed from CPS data for 1990.
4. Variance of schooling of 8.47, computed from CPS data for 1990.
5. Public education expenditures on all levels of education as a fraction of GDP of 3.9 percent from the Statistical Abstract of the United States (1999). In computing this statistic in the data, we treat as public expenditures all state and federal expenditures. We view local public expenditures as private because they are closely tied to property values and, therefore, to the incomes of parents. (See Restuccia and Urrutia (2004) for a discussion.)
6. A variance of individual fixed effects accounts for $\frac{2}{3}$ of the variance of log-earnings (Zimmerman (1992)). In our model, fixed effects are due to heterogeneity in parental resources, abilities, and tastes for schooling. The rest of the variation in earnings is due to market luck. Thus, the variance of fixed effects relative to the variance of earnings (in logs) is given by $1 - \frac{\sigma_{\mu}^2}{\text{var}(\ln(h_c))}$.
7. A Mincer return to schooling of 10 percent. Heckman, Lochner, and Todd (2005) report a Mincer return between 10 to 13 percent during the period 1980 to 1990. Psacharopoulos (1994) and Banerjee and Duflo (2005) estimate a Mincer return of 10 percent for the United States for the period 1990-95. All of these studies use data on annual earnings. Since our theory is about lifetime inequality, we estimate Mincer returns using NLSY to proxy lifetime earnings with 6-year averages of the earnings of

males aged 30-45.⁹ We obtained Mincer returns in the range .09 to .11, depending on the age group considered (see Table 3). In our model, we measure returns to education by regressing individual log-wages, wh_p , on years of education, given by the model period times s :

$$\ln(wh_{p,i}) = b_0 + b_1 (20s_i) + u_i,$$

where b_1 gives the Mincer returns to schooling.

8. R^2 in the Mincer regression of .22. We find that the R^2 tends to increase with the age-group considered, taking values between 0.16 and 0.26 (Table 3). Because the average value of R^2 over the life cycle is about 0.22, we set this value as a calibration target.

Table 3: Mincer Regression Results, NLSY

Age Group	Constant	Mincer Return	R^2	Num. Obs.
30-35	5.68	0.08	0.16	1857
35-40	5.56	0.10	0.21	1307
40-45	5.38	0.12	0.26	427

9. Teacher and staff compensation share in GDP of 0.05. According to the U.S. Department of Education, Digest of Education Statistics, 2007, (public and private) institutional costs for all levels of education amounted to 7 percent of GDP in 1990-1995. Seventy two percent of these expenditures were on teacher and staff compensation (Education at a Glance, OECD, 2007). In the model, this expenditure corresponds to the $w\bar{l}s$ cost of schooling aggregated across households.
10. Intergenerational correlation of schooling of 0.46 as obtained by regressing children's years of schooling on parental education, where the latter is defined as the average years of schooling among mothers and fathers (see Hertz, Jayasundera, Piraino, Selcuk, Smith, and Verashchagina (2007)).

⁹Each 6-year age group includes all males who worked full time for at least 3 out of 6 years, with observed wages and hours.

The calibration solves a rather complicated multidimensional mapping. Nonetheless, it is useful to discuss how model parameters affect some specific targets. Given mean years of schooling, the cost of teachers \bar{l} , and the income tax rate τ to finance public education expenditures, are almost directly pinned down by the share of teacher and staff salaries on GDP and by the share of public education expenditures on GDP (that is, the distribution of schooling matters little). The variance of market luck σ_μ^2 is set to match the variance of earnings. The parameter ξ – determining returns to scale in the human capital technology – targets the variance of individual fixed effects. In our theory, the earnings of parents and children are correlated in part due to differences in parental resources (the poor invest less), and in part due to an exogenous correlation of parental and children’s abilities. Thus, the correlation of ability ρ_z targets the intergenerational correlation of earnings.

Given the parameters just discussed, the five remaining parameters – variance of ability (σ_z), mean and variance of schooling tastes (controlled by θ_H and θ_L), correlation of ability and schooling taste (controlled by b), and the time share (η) – jointly determine the mean and variance of schooling, the R^2 and schooling coefficient in a Mincer regression, and the intergenerational correlation of schooling. The mean value for schooling taste $0.5(\theta_L + \theta_H)$ can be targeted to mean years of schooling, as the utility of schooling increases the benefits of schooling time. However, this parameter has important consequences for other targets as well. To develop this point, note that the return to schooling is affected by tastes and ability, where the former determines the utility of schooling and the latter the labor market returns to schooling. By making utility of schooling more prominent in schooling decisions, an increase in the mean of schooling tastes reduces the explanatory power of schooling on earnings. On the other hand, an increase in the variance of ability raises the importance of labor market returns in schooling decisions, hence raising the R^2 and the schooling coefficient in a Mincer regression. Moreover, while the targets for Mincer returns and for the R^2 tend to move together in response to parameter changes, the relative magnitudes of these responses

depend on the parameter being changed.¹⁰ The variance of schooling increases with a rise in the heterogeneity in the returns to schooling, which can be attained with an increase in both the variance of schooling tastes and ability or with a decrease in the time share (η). Furthermore, these alternative ways of increasing the variance of schooling have different implications for a Mincer regression: While the explanatory power of schooling in a Mincer regression decreases with the time-share parameter, it increases with both the variance of schooling tastes and ability (the former by increasing the R^2 and the latter by increasing the slope coefficient). Moreover, it also increases with the parameter b – controlling the correlation of schooling and ability – by making ability more important than taste shocks as a source of schooling variance. Nonetheless, the parameter b has a distinctive effect: While the intergenerational correlation of schooling increases with the parameter b , by making schooling tastes correlated across generations in a dynasty, this target is unaffected by the variance of ability or decreases with the variance of schooling tastes. Altogether, in spite of the high interdependence of the targeted moments, the parameters have distinctive quantitative effects on those moments.

4.2 The Benchmark Economy

The benchmark economy matches all the calibration targets quite closely (see Table 2). Below, we show that the model is also consistent with several dimensions of heterogeneity in the data that were not targeted in the calibration: schooling distribution as well as some evidence on the relationship between schooling attainment of the children and resources/background of their parents. Finally, we use results in the micro literature on the enrollment effects of college tuition changes to test the benchmark economy. We conclude that the model represents a good quantitative theory of a *within*-country heterogeneity in schooling and earnings in the U.S. economy.

¹⁰For instance, changes in schooling tastes (mean and variance) tend to have a strong impact on the R^2 , and changes in the standard deviation of ability tend to have a strong impact on the schooling coefficient of the Mincer regression.

Schooling Distribution While the calibration only targeted the mean and the variance of schooling, the model economy accounts surprisingly well for the distribution of schooling.¹¹ Table 4 reports maximum attained school years by population percentiles obtained from CPS 1990 data and those generated by our model. The model slightly overpredicts educational attainments at the bottom of the distribution and underpredicts them at the top of the distribution.

Table 4: Schooling Distribution – Model vs. Data

Percentile	Distribution of Schooling Years						
	5	10	25	50	75	90	95
Data	8	9	12	12	14.1	16	18
Model	10	10.2	10.6	11.2	14.1	17.2	19.4

Data source: CPS 1990

Schooling and Parental Background Although the calibration targeted the intergenerational correlation of schooling, the benchmark economy is consistent with other statistics relating parental background to offspring’s schooling. According to the statistics reported in Keane and Wolpin (2001), the probability that a child attains schooling less than or equal to 12 years conditional on his highest-educated parent having less than or equal to 12 years of schooling is .71 in the data. This probability is .72 in the model. Similarly, the probability that a child attains more than 12 years of schooling conditional on his highest-educated parent having more than 12 years of schooling is .60 in the data. This probability is .67 in the model economy.

In reviewing the literature on children’s educational attainment, Haveman and Wolfe (1995) report that the elasticity of children’s educational attainments with respect to family economic resources varies in the range of .02 to .20. In many of the studies cited in their

¹¹We note, however, that time in school is a continuous variable in our model, making its comparison with the data non-trivial. In particular, the distribution of schooling in the data has clear spikes at levels of education where an educational degree is completed.

survey, family income is recorded only in a single year and hence measures permanent income with an error. Haveman and Wolfe (1995) argue that when income is measured over a long period of time, the estimated impact of income is far greater. Our model economy produces an elasticity of .16, which is well within the range of values in the empirical literature.

Expenditures on Education In a well-known study, Haveman and Wolfe (1995) estimated the annual investment in children in the US economy in the year 1992. Their calculations distinguish the investments made by public institutions from those made by parents, as well as between direct and indirect private costs. They report that direct non-institutional private costs of education of children aged 0-18 accounted for 8 percent of GDP. Private and public institutional costs for all levels of education add 7.5 percent of GDP (U.S. Census Bureau). The total direct cost of education in the U.S. is thus 15.5 percent of GDP. In our model, the calibration did not target the aggregate amount of expenditures in education. Computed as the sum of $(p_e e + w \bar{l}s)$ over all students, the total cost of education accounts for 14.3 percent of GDP, a value slightly below the estimate of Haveman and Wolfe (1995).

Experiment: Effects of Tuition on College Enrollment. There is a large literature on enrollment effects of college tuition changes to which the predictions of the model can be compared. This literature is surveyed by Leslie and Brinkman (1987) and discussed by Keane and Wolpin (2001). Typically, the college costs effects on enrollment are identified from time series and cross-state variations in tuition rates and grant levels. To compare results across studies, it has become standard to use the percentage change in the overall enrollment rate of 18-24 year olds in response to a tuition increase of \$100 per year, expressed in dollars for the academic year 1982-1983. In a survey covering 25 empirical studies, Leslie and Brinkman conclude that, for national studies including the full range of public and private institutions, estimates of the effects of a \$100 increase in 1982-83 dollars tend to tightly pack in the range of a 1.8 to 2.4 percentage decline in the enrollment rate of 18-24 year olds.

To evaluate the response of college enrollment to a price change in the model economy,

we simulate a one-period unanticipated increase in college tuition of \$1000 in 1982-1983 dollars. This experiment is done in partial equilibrium so that factor prices are kept fixed. We find that college attendance declines by 1.5% per 100 dollars increase in college tuition, which is close to the consensus estimates in the empirical literature review by Leslie and Brinkman and to the recent estimates in Keane and Wolpin (2001).¹² In a schooling model structurally estimated with NLSY data on white young males, the authors found a decline in the college enrollment rate of 1.2 percent per \$100 tuition increase in 1982-83 dollars. Using estimates from Kane (1994), Keane and Wolpin report that a \$1000 tuition increase in 1982-1983 dollars leads to declines in the enrollment rate of 34.0, 20.0, 12.3, and 3.0 percent, respectively, for white males whose parents are in the first through fourth income quartiles. In comparison, our model economy predicts declines of 23.6, 21.9, 18.0, and 9.8 percent for individuals with parents in the first through fourth income quartiles. The model is thus consistent with the evidence that tuition effects are much stronger among individuals born in families with a low parental income, although tuition effects decline with parental income more steeply in the data than in the model. Altogether, the model is consistent with the micro evidence on the enrollment effects of college tuition changes.

5 Quantitative Results

This section uses the theory developed to quantitatively assess the consequences of TFP differences across countries. We assume that countries are identical in terms of preferences and technologies but only differ in their level of TFP. We asked the following questions: What cross-country differences in TFP are required for the model economy in order to account for a 20-fold income ratio between rich and poor countries? Does human capital play an important role in amplifying income differences across countries?

¹²More precisely, we found that a \$1000 increase in tuition raises the enrollment rate by 15%. Following the literature, we divide by 10 to obtain the response to a change in tuition of 100 dollars. We obtained quite similar results when we simulated an increase in tuition of 500 dollars. In this case, the decrease in tuition was 7.3 percent which implies a decline of 1.45 percent in enrollment per \$100 increase in tuition.

In answering the above questions, it will prove important to consider seriously the finding of Hsieh and Klenow (2007) that TFP does not vary uniformly across sectors between rich and poor countries. The fact that poor countries are relatively more productive at producing services than manufacturing implies that services are relatively cheap in poor countries. If the production of human capital is intensive in services, then poor countries have a comparative advantage in the production of human capital input. As a result, the low aggregate productivity of poor countries might not be as detrimental to human capital accumulation as one would conclude by ignoring the sectorial productivity differences across countries.

5.1 The experiment

To assess the magnitude of the TFP differences needed to account for the observed disparity in per-capita income across countries, we first need to take a stand on the values of three key parameters $(\varepsilon, \phi, \gamma)$. The first parameter, ε , determines the elasticity of the TFP in the service sector to a change of the TFP in the manufacturing sector. The other two parameters, ϕ and γ , pin down the share of manufacturing goods in consumption and in educational expenditures. Intuitively, ε determines the importance of cross-country heterogeneity in relative prices while ϕ and γ affect how the variation in relative prices impact on investment decisions and output per worker across countries.

The quantitative experiments below assume that the data counterpart to the service and manufacturing sectors in the model economy are the nontradable and tradable sectors in the data analyzed by Hsieh and Klenow (2007). In a cross-country study, these authors find that a one-percent variation in the TFP of the tradable sector is associated with a .3-percent variation in the TFP of the nontradable sector in 1996, and that this elasticity was about .4 in 1985 (see Table 7 on page 581).¹³ We thus consider experiments with $\varepsilon = .3$ and .4. To evaluate the sensitivity of the results, we also consider a ‘low’ and a ‘high’ value for the

¹³Hsieh and Klenow report that the elasticity of TFP with respect to PPP-output is about one third lower in the nontradable sector than in the tradable sector in the year 1996 (see Table 7 on page 581). This ratio is .5 in 1980 and .4 in 1985.

parameter ε by setting $\varepsilon = .1$ and $\varepsilon = 1$.

The calibration of the benchmark economy normalized the units in which output is measured in the manufacturing sector and in the service sector so that $A_S = A_M = 1$. Because these two sectors have the same TFP in the benchmark economy, the parameters γ and ϕ – determining the share of manufacturing goods in consumption and in education expenditures – do not affect equilibrium statistics. However, these parameters determine the quantitative implications of the theory across countries when sectorial productivities decrease unevenly relative to the benchmark economy. This observation should not be surprising: Since poor countries are relatively more efficient in the production of services (nontradables) than in the production of manufactured goods (tradables), cross-country income differences decrease with the importance of services in aggregate expenditures.

The discussion above implies that it is important to take a stand on the values assigned to the parameters γ and ϕ . In doing so, two issues need to be confronted: First, in the data the share of services in aggregate consumption expenditures increases with per-capita income across countries, suggesting that parameter γ varies with the level of economic development. Second, there is no direct evidence on the share of services in education expenditures so that it is not obvious how to determine a value for ϕ in the model economy. We proceed by setting $\gamma = .27$ in the Benchmark Economy so that this economy is consistent with the share of services in consumption expenditures in the US.¹⁴ To match the variation in the share of services in consumption expenditures across countries, the experiments below assume that the parameter γ varies with TFP in the manufacturing sector with constant elasticity. For each of the values of ε considered, we calibrate the value of this elasticity so that the theory is consistent with the fact that the elasticity of the share of tradables in consumption expenditures with PPP output is -0.13 . The share of services in education

¹⁴Note that in the data the education services provided by private non-profit institutions and government are included in the final consumption of households at their cost (see the Handbook of the International Comparison Programme). To be consistent, we define total consumption in our model as the sum of household consumption c and expenditures in education e . Hence, in the baseline economy the parameter γ determining the share of tradable in aggregate consumption is set so that $\frac{\gamma c}{c+e} = 0.23$ (author's estimate using PWT data), which implies $\gamma > .23$.

expenditures e is assumed to be 100% (ϕ is set to 0). Note that by assuming that all education inputs are produced by the service sector, our results will minimize the role of human capital accumulation in accounting for income differences across countries. Section 5.5 analyzes the sensitivity of the quantitative findings in some key dimensions. In particular, we investigate how the quantitative findings change when a fraction of educational expenses are in the form of tradable goods (such as pencils, paper, books, and computers), and when we allow countries to differ in their efficiency at producing human capital.

5.2 Measurement

To measure GDP at PPP prices, a set of ‘international prices’ needs to be chosen in a manner consistent with the methodology in the PWT. The set of ‘international prices’ in the PWT is constructed by averaging prices among all countries, according to the procedures established by the International Comparison Program (ICP) of the United Nations. In order to calculate the average price for a product across countries, each with its own currency, the prices in the individual countries are converted into a common numeraire currency using PPP exchange rates. The average price for good i is defined as:

$$P_i = \sum_j \frac{p_i^j}{E^j} \frac{q_i^j}{\sum_j q_i^j} \text{ for } i = 1, \dots, n, \quad (26)$$

where p_i^j and q_i^j represent the price and quantity of product i in country j . Each national price is converted into a common numeraire currency by dividing by the country’s PPP exchange rate E^j , and then averaged across all countries. The resulting price P_i is a weighted arithmetic average of the converted national price using the quantity shares as weights. Thus, P_i is the total value of the world transactions for good i , expressed in terms of PPP exchange rates, divided by the total quantity of the good.

Note that the set of international prices and the PPP exchange rates in the PWT are jointly determined as the solution to a system of equations involving prices for all goods

and PPP exchange rates for all countries. Solving such a system of equations in our model economy is a very demanding task as it involves simulating a set of model economies. The simulated model economies should mimic the world distribution of countries in terms of their population sizes and income distribution. In this way, the distribution of quantities transacted across countries for various commodities can be aggregated as in the PWT data. To circumvent this difficult problem, researchers typically calibrate the baseline economy to the US and set international prices equal to the prices in the baseline economy. This approach is motivated by the fact that, because the PWT use aggregate quantities to aggregate country prices, rich country prices are weighted more than poor-country prices (see the discussion in Hsieh and Klenow). Below, we argue that this approach has some serious drawbacks when applied to a model of schooling, such as the one in this paper.

As discussed in the Handbook of the ICP, there is very little basis for comparing education prices across countries using tuition or fees because they usually do not cover full cost and are not market-prices due to heavy government subsidies. It is thus not possible to use (26) to determine an international schooling price that can be used to value schooling output across countries. As a result, the ICP uses an “indirect approach” which involves using data on the PPP prices of inputs to aggregate at international prices expenditures on education inputs.¹⁵ Since the salaries of teachers are a major schooling cost, the international salary for teachers is a crucial determinant of schooling expenditures at international prices. However, the U.S. wage badly approximates the real wage used in the PWT to value education costs across countries. While for most products, such as cars and airline tickets, the US and similarly rich economies account for most of the world transactions, this is not the case for schooling, for two reasons. First, the variation in average years of schooling across rich and poor countries is easily an order of magnitude smaller than the variation in the consumption of cars and airline tickets. Second, poor countries account for the bulk of the world population of school-aged individuals. Thus, (26) is likely to put a significant mass on the salaries of teachers in

¹⁵The price of education is then obtained as the ratio of education expenditures at domestic prices to expenditures at international prices.

poor countries.¹⁶

The above discussion stresses a novel point: The choice of an international wage rate to value teachers' services across countries in a schooling model is a delicate issue. Moreover, as we have verified in our computational experiments, the results for output inequality across countries critically depend on the choice made. To circumvent these difficulties, we compute GDP at international prices net of teacher output or, equivalently, we measure national income net of the salary of teachers. The advantage of this approach is that we avoid taking a stand on how to set the international real wage for teachers. Moreover, using data from the PWT on institutional expenditures in education, we have made a similar adjustment to the GDP data in the PWT by computing GDP at international prices net of institutional expenditures. We have verified that all statistics of interest (such as the dispersion in income per capita and the income elasticity of schooling) are not affected in a significant manner by this adjustment. To sum up, the experiments below measure GDP at international prices as follows:

$$GDP_j^* = Y_M + Y_s,$$

where the price of services is set as in the baseline economy ($p_s^* = p_s^{US} = 1$). Later on, in a sensitivity analysis, we evaluate how the results change when the US wage rate is used as an international price to value teachers' services in GDP. We show that this procedure has highly counterfactual implications.

¹⁶While it is obviously important to aggregate all international prices in the model economy in a manner consistent with the PWT, this issue is of a first-order importance when it comes to aggregating teachers' wages. Because the cross-country variation in real wages is very large, incorrect weights can lead to an international salary for teachers in the model economy that is grossly at odds with the PWT. To deal with this problem, one approach would be to calibrate the model economy to replicate the world population distribution across rich and poor economies and use (26) to jointly solve for the set of international prices and countries' PPP exchange rates. This is a daunting task. Moreover, there is no guarantee that our simulations can mimic the world distribution of years of schooling because our calibration only targets average years of schooling in the baseline economy.

5.3 Amplification Effect

Unlike the results in Section 3, the amplification effect in the benchmark economy cannot be characterized with an analytical expression. However, there is a simple way of measuring the amplification effect of TFP in the calibrated model economy. For each value of ε , we simulate the model economy for different values of A_M and run the following regression in the simulated data

$$\ln Y = a_1 + a_2 \ln A_M + u_i,$$

where Y denotes GDP. The values considered for A_M are 1, .5, .25, .125. We run the regression for GDP measured at domestic prices and PPP prices. The fact that the R^2 in all the regressions are close to 1 implies that the estimated regressions represent a good description of how A_M and Y covary in the simulated data. The coefficient a_2 can then safely be interpreted as the elasticity of output with respect to A_M .

Table 5: Amplification

ε	.1	.3	.4	1
Human Capital Model				
<i>TFP elasticity of GDP</i>				
PPP prices	1.53	1.94	2.08	2.8
Domestic prices	1.98	2.16	2.26	2.8
<i>A_M ratio for GDP, PPP, ratio of 20</i>	7.1	4.7	4.0	2.9
TFP elasticity of Physical Capital	1.97	2.15	2.23	2.8
TFP elasticity of Human Capital	0.46	.63	.70	1.24
Exogenous Human Capital Model				
<i>TFP elasticity of GDP</i>				
PPP prices	.856	1.046	1.12	1.49
Domestic prices	1.49	1.49	1.49	1.49
<i>A_M ratio for GDP, PPP, ratio of 20</i>	33.1	17.5	14.5	7.5
TFP elasticity of Physical Capital	1.49	1.49	1.49	1.49

Table 5 reports the main results in the paper. The elasticity of GDP – at PPP prices – with respect to A_M is 1.94 when ε is .3 and 2.08 when ε is .4. To assess what the estimated

elasticities imply for understanding the observed income differences across countries, we compute the TFP ratio in the manufacturing sector needed to generate a ratio of aggregate income at PPP prices of 20. This ratio is roughly the PPP-income ratio between the 10% richest countries to the 10% poorest countries in the world income distribution. The ratio of TFP in tradables to explain a PPP-income ratio of a factor of 20 is 4.7 when $\varepsilon = .3$ and 4.0 when $\varepsilon = .4$. These findings imply a substantial amplification of TFP differences across countries. The mechanism generating a large income disparity is that a low TFP leads individuals in poor countries to invest few resources in accumulating both physical and human capital relative to individuals in rich countries.

One way of assessing the amplification results in our paper is to compare them with the findings in Hsieh and Klenow (2007). These authors perform a development accounting exercise using a growth model with no human capital accumulation and find that a 1-percent change in the TFP of the tradable sector leads to an increase in output per worker of 1.04.¹⁷ To show that the much larger amplification in our theory is due to human capital accumulation, we calibrate a version of the model economy with no investments in human capital.¹⁸ When $\varepsilon = .3$ and human capital is exogenous, the A_M elasticity of output at PPP prices is 1.05, which is quite close to the 1.04 estimated by Hsieh and Klenow. The ratio of TFP in manufacturing to explain a PPP-income ratio of a factor of 20 is now 17.5, which is much higher than the 4.7 ratio in the economy with investments in human capital.

Human capital is an important source of amplification of income differences across countries for two reasons: First, human capital directly contributes to cross-country output differences because poor countries accumulate less human capital than rich countries. Second, a higher human capital stock stimulates more physical capital accumulation by raising the marginal product of capital. As a result, human capital accumulation amplifies the effects of TFP differences on physical capital: While the A_M elasticity of physical capital is 2.15 in

¹⁷They report an elasticity of TFP in the tradable sector with respect to income per capita in 1996 of .962, which implies a TFP elasticity of income of $1/.962 = 1.04$.

¹⁸Labor productivity is fixed at $h = z s^{n\varepsilon}$, where s is set at its average value in the benchmark economy.

the economy with human capital accumulation, it is only 1.49 in the model with no human capital investments (as documented in Table 5 for the economy with $\varepsilon = .3$). Note that the strength of these effects increases when sectorial productivity differences across countries are small (ε is high). The TFP elasticities of human and physical capital rise with ε as the higher the value of this parameter the lower the comparative advantage of poor countries at producing services (see Figure 1).

Given the role of human capital in amplifying income differences, it is interesting that poor countries in our model economy are characterized by a high ratio of human capital stocks to PPP-output. The assumption that the human capital production technology does not vary across countries implies that poor countries have a comparative advantage in accumulating human capital relative to physical capital. The relative inefficiency in producing physical capital explains why these countries exhibit a low ratio of physical capital to output at PPP prices, as emphasized by Hsieh and Klenow (2007). The quantitative importance of these effects can be assessed by comparing the A_M -elasticities of physical capital, output, and human capital (see Table 5). When $\varepsilon = .3$, we have that $E_{KA_M} = 2.15$, $E_{YA_M} = 1.94$, and $E_{HA_M} = .63$. The ranking of elasticities ($E_{KA_M} > E_{YA_M} > E_{HA_M}$) implies that the physical-capital-to-output ratio increases with TFP but that the human-capital-to-output ratio decreases with TFP.

The results in Table 5 indicate that it is important to model sectorial productivity differences for assessing the role of human capital in amplifying income differences. When $\varepsilon = 1$, the A_M elasticity of PPP-output is about 1.31 points higher in the model with endogenous human capital than in the model with no human capital investments. This elasticity-differential decreases to 1.046 when $\varepsilon = .40$ and to 0.89 when $\varepsilon = .3$. As the TFP elasticity of PPP output decreases with ε , so does the amplification provided by human capital (see Figure 1, bottom). When ε is set at a low value, poor countries exhibit a high relative TFP in the service sector, which, in turn, leads to a low price of services relative to rich countries. Since services are a key input in the production of human capital, cheap services in poor countries

operate as a force towards reducing income inequality. Nevertheless, the amplification role of human capital is large even for implausibly low values of ε . When $\varepsilon = .1$ the TFP elasticity of PPP output is 1.53 in the model with human capital accumulation and .86 in the economy with exogenous human capital. This differential in TFP elasticities across model economies is not minor: To generate an income ratio of 20 the economy with endogenous human capital requires an A_M ratio of 7.1 while the economy with exogenous human capital requires an A_M ratio of 33.1 (see Table 5).

At the theoretical level, it is interesting to answer the following question: How does a one-percent change in TFP in all sectors in the economy affect output per worker? The answer is provided by the one-sector version of the model economy ($\varepsilon = 1$), and it is startling: The amplification effect is now 2.8. In a world were TFP varies uniformly across sectors, the TFP ratio needed to generate a PPP-income ratio of a factor of 20 would be only 2.9, which is about two thirds of what is implied by the two-sector model with $\varepsilon \in [.3, .4]$. Nevertheless, from a development accounting view, the relevant amplification effect is the one estimated with the two-sector model as the evidence suggests that TFP does not vary uniformly across sectors. We conclude that it is important to model both human capital and sectorial productivity differences for assessing the cross-country variation in productivity.

5.4 Discussion on Relative Prices and Human Capital

We have shown that human capital is an important source of amplification and that the TFP elasticity of output depends critically on the parameter ε . Since ε determines the variation in relative prices across countries, it is important to examine the implications of the theory for relative prices and test them with the evidence from the Penn World Tables (Heston, Summers, and Aten (2002), hereafter, PWT). Moreover, to address the concern that our quantitative theory may be exaggerating the TFP elasticity of human capital investments, we examine evidence on the variation in human capital investments across countries.

5.4.1 Theory and Evidence on Relative Price Variation

We can construct a proxy for the price of education inputs across countries using data on the price of services from the PWT. This seems a reasonable proxy in the view that educational inputs are mostly provided by the service industry. Figure 2 (top) plots cross-country data on the price of services versus per-capita income from the PWT for the year 1996 as well as the simulated model data. Note that the elasticity of the relative price of services with respect to PPP output is .30 in the PWT data, which is quite close to the .29 value predicted by the economy with $\varepsilon = .4$ and to the .36 value obtained in the economy with $\varepsilon = .3$. The economies with $\varepsilon = 1$ and $\varepsilon = .1$ have counterfactual predictions for the variation in the price of services across countries: The former implies no variation in relative prices across countries while the latter grossly overpredicts the variation in the data (2, top). We conclude that the evidence supports values of ε within the range [.3, .4], with the best fit of the data obtained when ε is parameterized with values close to .4.

The PWT define PPP exchange rates for education as education expenditures in national currency divided by their real value in international dollars. The education PPP exchange rates are computed with data on expenditures by *educational institutions*, for there are no cross-country data on educational expenditures at the level of the household. Figure 2 (middle) plots the PWT data on the price of education, normalized by the PPP price of GDP, and the PPP output per capita. For comparison, we plot model-simulated data on the prices of two key schooling inputs normalized by the PPP price of GDP: the wage rate and the price of services (nontradables) (Figure 2, middle and bottom). By focusing on the prices of these two schooling inputs, we avoid taking a stand on how to aggregate and value at international prices expenditures in our model economy (see the discussion in the Measurement section).

Figure 2 (middle) documents that the price of institutional expenditures in education tends to increase with income across countries albeit the relationship is not very strong. While the income elasticity of the education price is .046, in our model economy the income

elasticity of the real wage rate – in terms of PPP output – takes values above .60 when $\varepsilon = .3$ and .4. The one-sector economy ($\varepsilon = 1$) exhibits the smallest elasticity, .53, which is still well above the value in the data. Figure 2 (bottom) shows that the income elasticity of the relative price of services – in terms of the PPP price of output – takes values of .25 and .21 when $\varepsilon = .3$ and .4. Hence, the findings suggest that in our simulations the price of education inputs rises too fast with the level of economic development relative to the data, suggesting that the quantitative results *understate* human capital differences across countries.

5.4.2 Theory and Evidence on Variation in Human Capital Investments

Next, we turn to the question: Are the cross-country differences in human capital investments implied by the theory plausible? For each value of ε , we obtain observations for average years of schooling and output per worker by simulating economies that vary in their relative levels of TFP. In Figure 3 (top), we plot cross-country data on schooling and income, taken from Cohen and Soto (2007) and Heston, Summers, and Aten (2002), together with the simulated data from the model economy. The figure reveals that the income semi-elasticity of schooling in the cross-country data is 2.56. All the simulated model economies generate a lower schooling-income semielasticity than in the data. The highest value of the semielasticity, 1.86, is generated by the economy with $\varepsilon = 1$. Given that our theory does not overstate schooling differences across countries, we then ask if the the cross-country differences in the *quality* of schooling across similar schooling levels are reasonable.

In an influential paper, Hendricks (2002) measures cross-country differences in schooling quality using data on relative earnings (adjusted by schooling levels) of immigrants in the United States. Table 6 provides summary statistics on the data analyzed by Hendricks. The population of U.S. immigrants is divided into 4 groups according to the income-percentile of the country of origin relative to the United States. The country groups considered are the 20th, 30th, 40th and 50th to 65th percentiles of the U.S. per-capita income. The average years

of schooling among immigrants in these country groups are, respectively, 12.5, 12.8, 12.4, 14.3, and the average earnings of these immigrants are 97%, 92%, 94% and 107% of the earnings of similarly educated workers in the United States (see Figure 4). While Mexico and Puerto Rico¹⁹ have a per-capita income of roughly 45% of the U.S. level, we did not include these two countries in the 40th income percentile group because immigrants born in Mexico and Puerto Rico have on average 7.4 years of schooling – a schooling level well below that of all other immigrants in Hendricks’ sample. Nevertheless, we examine the data for Mexico and Puerto Rico separately in a fifth country group.

Table 6: U.S. Immigrant Characteristics (computed from Hendricks (2002))

<i>GDP, PPP, percentile</i>	20 – 30%	30 – 40%	40 – 50%	50 – 65%	MEX-PRT
Number of countries	11	10	7	5	2
Relative home GDP, PPP	24.4	33.3	44.8	58.5	45.75
Years of schooling	12.5	12.8	12.4	14.3	7.45
Relative earnings	0.97	0.92	0.94	1.07	0.93

We simulate immigrants from five potential source countries differing with respect to their TFP in order to generate comparable statistics from the model economy. Immigrants are selected so that they have an average level of schooling consistent with the data reported in Hendricks. We then compute the ratio of earnings between immigrants and equally schooled workers in the Benchmark Economy and compare these results to Hendricks’ data.

To proceed, we need to take a stand on how immigrants are selected from the population in the source country. In our model economy, equally schooled individuals can be heterogeneous in many characteristics (taste for schooling, ability, parental human capital, and wealth) and, thus, in their earnings. As a result, selection into emigration from the distribution of these characteristics has important consequences for their average earnings. We now study in detail how different forms of selection by wealth affect the results. For each

¹⁹To simplify the terminology here, we slightly abuse the language and refer to Puerto Rico as a ‘country’ rather than an incorporated U.S. territory and call migrants from Puerto Rico in the U.S. ‘immigrants’

source country, we compute the wealth distribution for individuals within a given schooling bracket, and we entertain two possibilities on how immigrants are selected from these populations. First, as a benchmark, we assume that immigrants are randomly drawn from the entire wealth distribution. Second, we examine selection into emigration based on the household wealth and show that this type of selection successfully reconciles the relative earnings of immigrants predicted by the model with those obtained from the data.²⁰ Figure 4 presents results for economies with $\varepsilon = .3$ and $\varepsilon = .4$.

When immigrants are a random draw from the entire wealth distribution (100th percentile on Figure 4), our model tends to overpredict the earnings gap between workers of the same schooling level in rich and poor countries. While the earnings gap for immigrants born in countries below the 50th percentile are about .80 in the model, it is above .90 in the data. On the other hand, the model overpredicts the the relative earnings for immigrants born in Mexico and Puerto Rico, producing a ratio of 1.04 relative to a .93 in the data. Moreover, the model cannot account for the fact that immigrants from countries in the 50th percentile earn about 7% more than americans, an observation suggesting that schooling quality is higher in this group of countries than in the U.S.

We now show that the model can account well for these non-trivial patterns of the immigrant earnings data, provided that selection of immigrants by wealth is allowed to play a role. We assume that immigrants are a random draw from the bottom tail of the wealth distribution in a given schooling bracket. Then, we right-censor this distribution at different wealth percentiles and assume that immigrants are a random draw from the resulting wealth distribution. Notice that selection is more important the more truncated the wealth distribution is. The case of no selection corresponds to the assumption that immigrants are drawn from the bottom 100% of the wealth distribution (e.g. no truncation). Figure 4 graphs, for $\varepsilon = 0.3$ and $\varepsilon = 0.4$, how average earnings vary as immigrants are increasingly

²⁰Note that selection by wealth matters because wealth is correlated with ability and schooling expenditures. Alternatively, we could directly select immigrants in terms of their ability and schooling expenditures but this would not affect our main conclusion that selection can reconcile the predictions of the theory with the data.

drawn from more wealthy backgrounds. We find that immigrants' human capital tends to decline with parental wealth for the first four country groups. On the other hand, earnings increase with parental wealth in the fifth country group representing Mexico and Puerto Rico.

Why does the relationship between parental wealth and immigrants' human capital switch signs across countries? The key is that while immigrants from the first four country groups have on average more than 12 years of schooling, immigrants in the country group representing Mexico and Puerto Rico have on average only about 7 years of schooling. As a result, immigrants from the first four country groups exhibit high average years of schooling relative to their source-country population, and the opposite is true for the case of Mexico and Puerto Rico. When immigrants are positively selected from the schooling distribution, they tend to exhibit a relatively high taste for schooling. In this case, the rate of return to schooling is a relatively less important determinant of schooling decisions. Moreover, the importance of returns to education for schooling decisions diminishes with parental wealth: Wealthy individuals tend to care more about the utility of schooling as they have a low marginal utility of consumption. As a result, conditional on a level of schooling, individuals with wealthy backgrounds tend to be of a relatively low ability and to spend little on education. This accounts for the negative relationship between earnings and parental wealth among highly-schooled immigrants. Our findings point that the case of Mexico and Puerto Rico is quite different. Immigrants from these countries are relatively low-schooled and care little about the utility of schooling. The rate of return to schooling is the main driving force behind their schooling decisions. A more favourable parental background is associated with more human capital expenditures and, hence, higher human capital.

Figure 4 (top) indicates with a dot, for each country group, the amount of selection needed to account for the immigrants' earnings data for the economy with $\varepsilon = .3$. The earnings data for the first three country groups can be rationalized if immigrants are drawn from the bottom 20th, 30th, and 50th percent of the wealth distribution. Moreover, assuming that

immigrants are drawn from the bottom 40 percent of the wealth distribution accounts for the observed earnings ratio of 1.07 among immigrants born in countries in the 50th to 65th percentile group. The average earnings ratio observed among immigrants born in Mexico and Puerto Rico can be accounted for if immigrants from these countries were drawn from the bottom 60% of the wealth distribution. Figure 4 (bottom) illustrates similar findings for the economy with $\varepsilon = .4$. We conclude that the model can account well for Hendricks' data.

5.5 Sensitivity Analysis

We now assess the effects of TFP under different assumptions on the importance of traded goods in education expenditures and on the cross-country differences in the efficiency in the human capital production. Moreover, in light of the discussion in the Measurement Section, we evaluate the sensitivity of the results to alternative ways of measuring TFP.

The cross-country experiments presented above assumed that all education inputs are nontradable ($\phi = 0$) and that countries have the same efficiency in producing human capital ($A_H^j = 1$ for all j). To the extent that education requires some tradable inputs (such as books, paper, pencils, computers), it is interesting to evaluate the sensitivity of the results to allowing for some expenditures in tradable inputs. To this end, we consider an experiment in which the share of traded goods in education inputs is set to .23 for all countries. This value corresponds to the share of traded goods in consumption expenditures in the U.S. ($\gamma^{US} = .23$). Recall that the U.S. exhibits the lowest share of traded goods in consumption expenditures, as in the data this share decreases with the level of per-capita income across countries. We thus view the choice of $\phi = .23$ for all j as likely to be conservative.

To the extent that in the data poor countries tend to have low productivity in all sectors of the economy, it seems reasonable to consider the possibility that poor countries also have a low efficiency in producing human capital. To this end, we consider an experiment in which countries differ in their efficiency in producing human capital (the share of tradable goods in education expenditures is set to 0). Specifically, countries are assumed to vary in their

efficiency at producing human capital according to their TFP in the service sector ($A_H^j = A_S^j$ for all j).

We find that modeling traded inputs in human capital production and modeling cross-country differences in the efficiency of human capital production improves the predictions of the theory for the cross-country variation in schooling (see Figure 3, bottom). In fact, the latter formulation matches the schooling-income elasticity in the data quite closely. Amplification does not rise much in the first experiment, but it increases substantially in the second one. When countries differ in the efficiency of human capital production, the A_M ratio that generates an output ratio of 20 drops to 3.3, which should be compared to the 4.7 ratio in the Baseline Economy. Despite differences in the results across experiments, we emphasize two consistent findings. First, for all the specifications of the human capital technology considered, we find that human capital plays a crucial role in amplifying income differences across countries. Second, accounting for the income differences across countries always requires substantial differences in TFP.

Table 7: Sensitivity

	Baseline($\varepsilon = .3$)	$\phi = .23$	$A_H = A_S$
<i>TFP elasticity of</i>			
GDP, PPP	1.94	2.05	2.49
Physical Capital	2.23	2.3	3.1
Human Capital	.63	.76	1.22
<i>for GDP, PPP, ratio of 20</i>			
TFP ratio	4.7	4.3	3.3

In Section 5.2, we have argued that it is quite difficult to take a stand on how to value teacher services at international prices. These difficulties led us to value GDP net of teacher services in our computational experiments. We now evaluate how our results change if we use the prices of our baseline economy (calibrated to the US) to value teacher services when measuring GDP at international prices. Table 8 presents the results. When $\varepsilon = .3$, using

the U.S. wage rate to value teacher services across countries reduces amplification from 1.94 to 1.5. As a result, the TFP ratio in the Manufacturing sector needed to explain an output ratio of 20 increases from 4.7 to 7.4. In assessing the relevance of these findings, it is worth making two observations. First, TFP amplification is substantially above the one obtained in the exogenous human capital model as indicated by the fact that a TFP ratio of 17.5 is needed to account for an output ratio of 20. Thus, our main finding – that human capital is an important source of amplification of cross-country income differences – stands regardless of the method used to value teacher services. Second, when U.S. wages are used to value the contribution of teacher services to GDP, the model economy has implications that are grossly at odds with the data: It implies that the share of teacher salaries in GDP is about 65% of GDP in poor economies. However, among countries with a per-capita income of *1/20th* of the U.S. level, teacher salaries account for 6.5% of GDP at PPP prices. That is, the theory overstates this share by a factor of 10. We conclude that using U.S. wages to value teacher services severely biases upwards the GDP of poor countries at international prices and, hence, biases downward our estimates of TFP amplification.

Table 8: Robustness: using w^{US} to value teacher services

ε	.1	.3	.4	1
<i>TFP elasticity of GDP, PPP</i>				
Baseline	1.53	1.94	2.08	2.8
$w^* = w^{US}$	1.26	1.50	1.66	2.29
Exogenous HC	.856	1.046	1.12	1.49
<i>A_M ratio for GDP, PPP, ratio of 20</i>				
Baseline	7.1	4.7	4.0	2.9
$w^* = w^{US}$	10.8	7.4	6.1	3.7
Exogenous HC	33.1	17.5	14.5	7.5

5.6 Brief Literature Discussion

In a provocative paper, Mankiw, Romer, and Weil (1992) [hereafter MRW] argue that the Solow growth model augmented to include human capital can account for most of the variation in output per capita across countries. However, Klenow and Rodriguez-Clare (1997) and Bils and Klenow (2000) argue that MRW have overstated the importance of human capital in accounting for cross-country income differences by focusing on a one-sector model with no distinction between the production of goods and human capital. To address these concerns, our paper develops a theory of human capital investments based on a multi-sector heterogeneous agent model, calibrates it to micro data on schooling and earnings, and uses the theory to quantitatively assess the sources of cross-country income differences. In a similar pursuit, but using a different approach, Manuelli and Seshadri (2005) [MS hereafter] build a representative-agent model with life-cycle human capital accumulation. They abstract from sectorial heterogeneity in TFPs and calibrate the human capital technology to the age profile of wages in the data. This approach produces a TFP elasticity of output per worker of 6.6, which is substantially larger than the elasticity of 1.94 in our baseline calibration and 2.81 when we abstract from sectorial productivity differences.²¹ The discrepancy between these elasticities is not minor: While MS find that a factor of 20 differences in output per worker can be explained with a TFP difference of 60 percent, our results point to a TFP difference of 400 percent. Alternatively, an amplification effect with a TFP elasticity of output of 1.94 in our baseline calibration implies that an annual rate of TFP growth of .92 percent accounts for the post-war output growth in the United States (about 1.8 percent a year), whereas the amplification effect found by MS requires a much lower annual rate of technological progress (.27 percent).

In a closely related study that follows a methodology much different from ours, Hendricks

²¹Manuelli and Seshadri (2005) report a TFP elasticity of output per worker of 9 when both TFP and demographic factors are allowed to vary across countries. We estimate the elasticity to be 6.6 when demographic factors are kept constant to U.S. levels using the results in Table 4, page 24. In comparing results across papers, it should be noted that the TFP differences in MS apply to the whole economy as opposed to the Manufacturing (tradable) sector as in our paper.

(2002) concludes that important TFP differences are needed to account for the cross-country data. Hendricks' growth accounting exercise does not rely on a specific functional form for human capital accumulation, but instead uses data on relative (adjusted by schooling levels) earnings of immigrants in the United States to directly infer cross-country differences in schooling quality. He estimates that for the five poorest countries in his sample – with output per capita of 5.8 percent of the US level – a low factor accumulation reduces income per capita to 47% of the US level. In our paper, the reduction in output per capita accounted for by a low factor accumulation is 25% – almost twice as big as estimated by Hendricks. Nonetheless, we think that the conclusions in our paper are consistent with the findings of Hendricks. Like Hendricks, we find that substantial TFP differences are required to account for the income disparity in the data. We thus side with Hendricks in concluding that accounting for the observed cross-country income differences on the basis of human and physical capital alone would require implausibly large degrees of self-selection in unobserved skills among immigrants. Our findings suggest that TFP is more important than is apparent in Hendricks' careful analysis since TFP differences can account for most of the cross-country variation in average years of schooling and in schooling quality.²² In addition, our paper emphasizes that human capital accumulation also plays an important role in amplifying income differences across countries. We show this by comparing TFP amplification in the benchmark economy to that in a model with exogenous human capital. We find that a 20-fold difference in income between rich and poor countries is generated by a 5-fold TFP difference in the tradable sector in our model with human capital, while the model without human capital requires a stunning 18-fold TFP difference.

²²While the cross-country differences in schooling quantity and quality are taken as given by Hendricks, in our paper they are the result of TFP differences across countries.

6 Conclusions

We developed a quantitative theory of human capital with heterogeneous agents in order to assess the magnitude of the cross-country differences in TFP that are needed to explain the variation in cross-country output per worker. To this end, we calibrate to micro evidence a rich schooling model in which heterogeneity in schooling is driven by variations in labor-market returns, schooling tastes, and wealth. In our theory, the parameters governing human capital production, tastes for schooling, and a random ability process have important implications for a set of cross-sectional statistics: variances and intergenerational correlations of earnings and schooling, as well as slope coefficient and R^2 in a Mincer regression. These restrictions of the theory and U.S. household data are used to pin down the key parameters driving the quantitative implications of the theory.

To address the findings of Hsieh and Klenow (2007) that services tend to be cheaper in poor countries relative to rich countries, we embed our human capital theory in a multi-sector growth model that allows for sectorial productivity differences across countries. Our cross-country simulations allow for relative prices to vary across countries, with services being relatively cheap in poor economies. Our baseline experiments minimize the role of human capital in amplifying income differences across countries by assuming that human capital investments only require services. Nonetheless, our main finding is that human capital accumulation strongly amplifies TFP differences across countries: In our preferred parameterization, the elasticity of output per worker – at PPP prices – with respect to TFP in the tradable sector is 1.94. This implies that a 5-fold difference in TFP explains a 20-fold difference in the output per worker, observed between the 10 percent richest and 10 percent poorest countries in the world. In contrast, without human capital accumulation, an 18-fold difference in the TFP of the tradable sector is required to account for the income difference between rich and poor countries.

We leave for future work two important extensions of our analysis. First, we plan to

explore the distributional implications of cross-country differences in TFP, fiscal policies and support for public education. Second, we would like to examine the sensitivity of our results to modeling heterogeneity in the marginal returns to schooling. This heterogeneity has been emphasized by a recent micro literature as a potential explanation for why the estimates of returns to schooling using instrumental variables are bigger than the ones obtained using ordinary least squares (see Card (2001)). This extension could also be interesting because the impact of a public-policy reform is driven by the marginal returns of the individuals affected by the reform and not by the average return in the whole population.

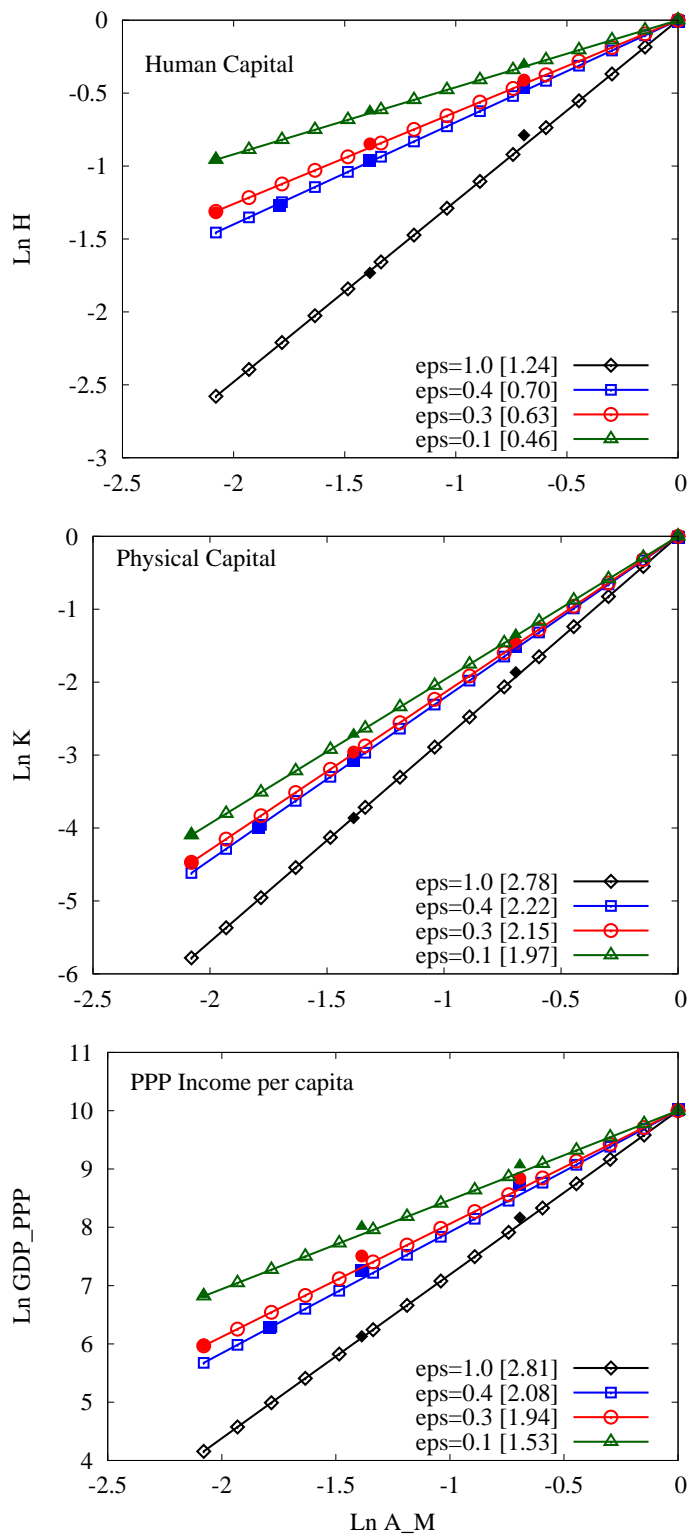
References

- BANERJEE, A., AND E. DUFLO (2005): “Growth Theory through the Lens of Development Economics,” in *Handbook of Economic Growth, Vol. 1A*, ed. by S. Durlauf, and P. Aghion, pp. 473–552. Elsevier Science Ltd. - North Holland.
- BECKER, G., AND N. TOMES (1986): “Human Capital and the Rise and Fall of Families,” *Journal of Labor Economics*, 4(3), S1–S39.
- BILS, M., AND P. KLENOW (2000): “Does Schooling Cause Growth?,” *American Economic Review*, 90(5), 1160–83.
- CARD, D. (2001): “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems,” *Econometrica*, 69(5), 1127–60.
- COHEN, D., AND M. SOTO (2007): “Growth and Human Capital: Good Data, Good Results,” *Journal of Economic Growth*, 12(1), 51–76.
- Economic Report of the President (2004):. United States Government Printing Office, Washington.
- EROSA, A., AND T. KORESHKOVA (2007): “Progressive Taxation in a Dynastic Model of Human Capital,” *Journal of Monetary Economics*, 54, 667–85.
- HAVEMAN, R., AND B. WOLFE (1995): “The Determinants of Children’s Attainments: A Review of Methods and Findings,” *Journal of Economic Literature*, 33, 1829–78.
- HECKMAN, J. J., L. J. LOCHNER, AND P. E. TODD (2005): “Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond,” Working Paper 11544, NBER.
- HENDRICKS, L. (2002): “How Important is Human Capital for Development? Evidence from Immigrants’ Earnings,” *American Economic Review*, 92(1), 198–219.
- HERRENDORF, B., AND V. AKOS (2007): “Which Sectors Make the Poor Countries So Unproductive,” mimeo.
- HERTZ, T., T. JAYASUNDERA, P. PIRAINO, S. SELCUK, N. SMITH, AND A. VERASHCHAGINA (2007): “The Inheritance of Education Inequality: International Comparisons and Fifty-Year Trends,” *The B.E. Journal of Economic Analysis and Policy*, 7(2), Article 10.
- HESTON, A., R. SUMMERS, AND B. ATEN (2002): *Penn World Table Version 6.1*. Center for International Comparisons at the University of Pennsylvania (CICUP).
- HSIEH, C., AND P. KLENOW (2007): “Relative Prices and Relative Prosperity,” *American Economic Review*, 97(3), 562–585.

- KANE, T. J. (1994): “College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education,” *Journal of Political Economy*, 102(5), 878–911.
- KEANE, M. P., AND K. I. WOLPIN (1997): “The Career Decisions of Young Men,” *Journal of Political Economy*, 105(3), 473–522.
- (2001): “The Effects of Parental Transfers and Borrowing Constraints on Educational Attainment,” *International Economic Review*, 42(4), 1051–1103.
- KENDRICK, J. W. (1976): *The Formation and Stocks of Total Capital*. Columbia University Press (for the National Bureau of Economic Research), New York.
- KLENOW, P., AND A. RODRIGUEZ-CLARE (1997): “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?,” in *NBER Macroeconomics Annual*, ed. by B. S. Bernanke, and J. J. Rotemberg. MIT Press, Cambridge, MA.
- LESLIE, L. L., AND P. T. BRINKMAN (1987): “Student Price Response in Higher Education: The Student Demand Studies,” *The Journal of Higher Education*, 58(2), 181–204.
- MANKIW, N. G., D. ROMER, AND D. N. WEIL (1992): “A Contribution to the Empirics of Economic Growth,” *Quarterly Journal of Economics*, 107(2), 402–37.
- MANUELLI, R., AND A. SESHADRI (2005): “Human Capital and the Wealth of Nations,” mimeo.
- MULLIGAN, C. B. (1997): *Parental Priorities and Economic Inequality*. University of Chicago Press, Chicago.
- (1999): “Galton versus the Human Capital Approach to Inheritance,” *Journal of Political Economy*, 107(6), S184–S224.
- NEAL, D. A., AND W. R. JOHNSON (1996): “The Role of Premarket Factors in Black-White Wage Differences,” *Journal of Political Economy*, 104(5), 869–895.
- POTERBA, J. (1997): “The Estate Tax and After-Tax Investment Returns,” Working Paper 6337, NBER.
- PSACHAROPOULOS, G. (1994): “Returns to Investment in Education: A Global Update,” *World Development*, 22, 1325–43.
- RESTUCCIA, D., AND C. URRUTIA (2004): “Intergenerational Persistence of Earnings: The Role of Early and College Education,” *American Economic Review*, 94(5), 1354–78.
- SOLON, G. (1999): “Intergenerational Mobility in the Labor Market,” in *Handbook of Labor Economics*, Vol. 3A, ed. by O. C. Ashenfelter, and D. Card, pp. 1761–1800. North-Holland, Amsterdam.
- Statistical Abstract of the United States (1999):. U.S. Government Printing Office, Washington, DC.

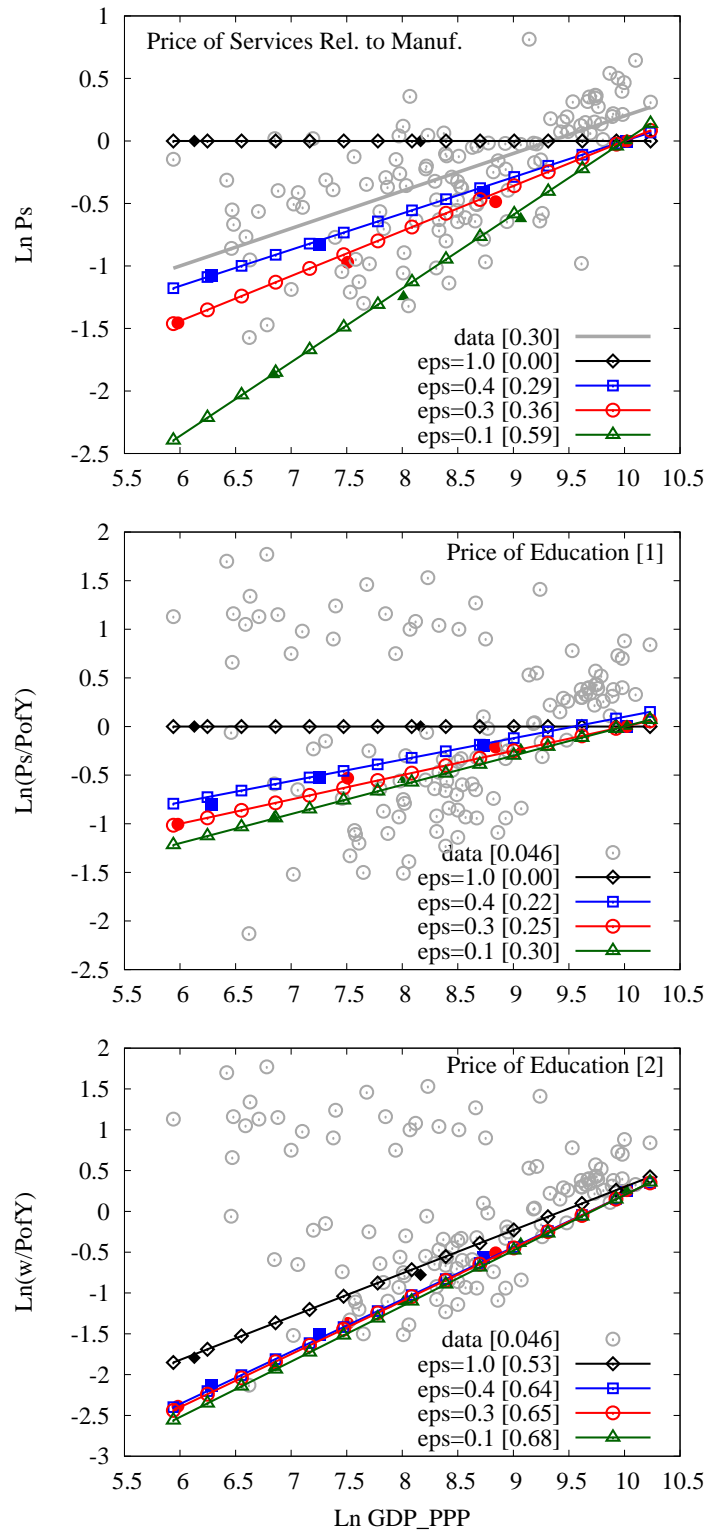
- STOKEY, N. L. (1998): "Shirtsleeves to Shirtsleeves: The Economics of Social Mobility," in *Frontiers of Research in Economic Theory: The Nancy L. Schwartz Memorial Lectures, 1983-1997.*, ed. by D. P. Jacobs, E. Kalai, and M. I. Kamien, pp. 210–41. Cambridge University Press, New York, NY.
- TAUCHEN, G. (1986): "Finite State Markov Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters*, 20(2), 177–81.
- TROSTEL, P. (1993): "The Effect of Taxation on Human Capital," *Journal of Political Economy*, 101(2), 327–50.
- ZIMMERMAN, D. J. (1992): "Regression Toward Mediocrity in Economic Stature," *American Economic Review*, 82(3), 409–429.

Figure 1: TFP Amplification



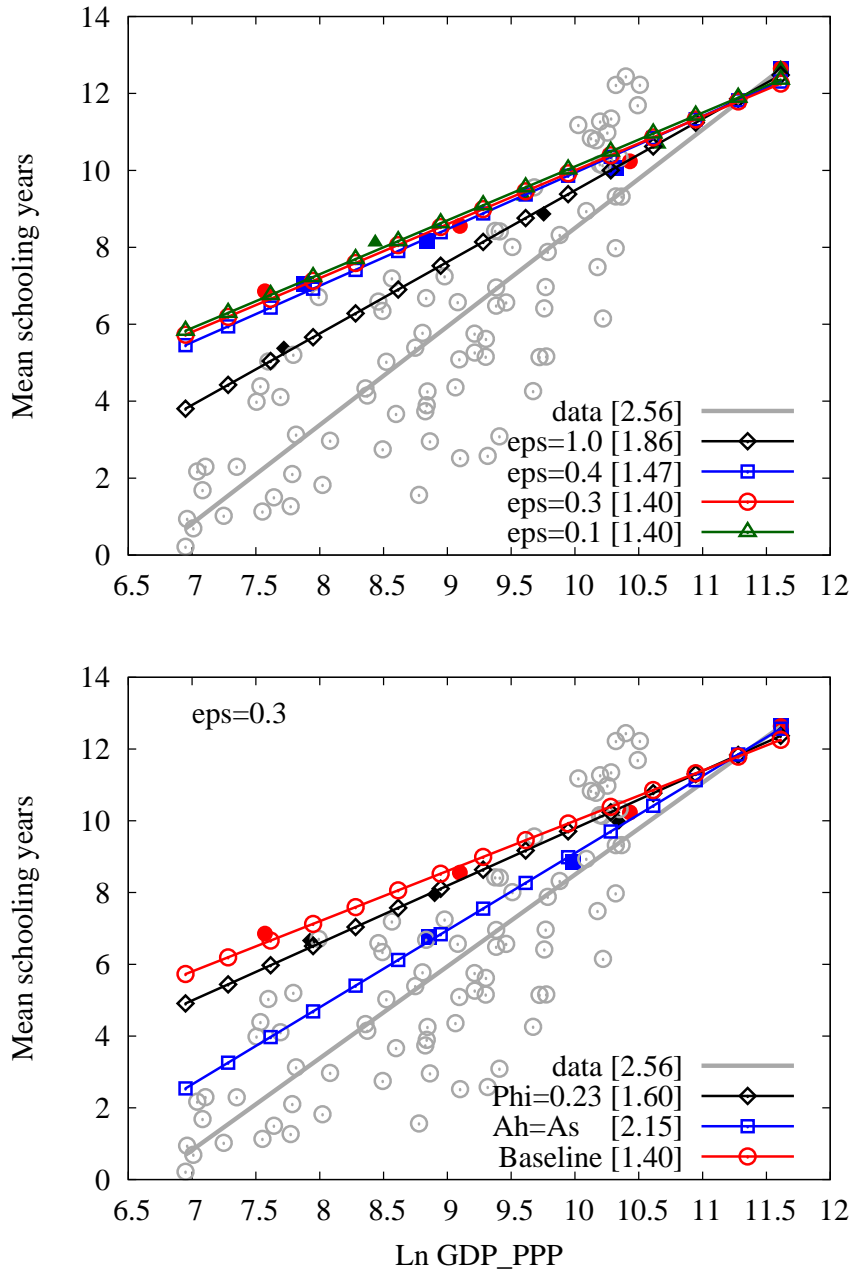
Model capital stocks are normalized by those in the benchmark economy. Model PPP income is normalized by the 1990 U.S. GDP at PPP prices. Solid points represent simulated economies; lines are OLS regressions with log-TFP in manufacturing sector as an explanatory variable. TFP elasticities (regression coefficients on TFP) are indicated in square brackets.

Figure 2: Prices and Income



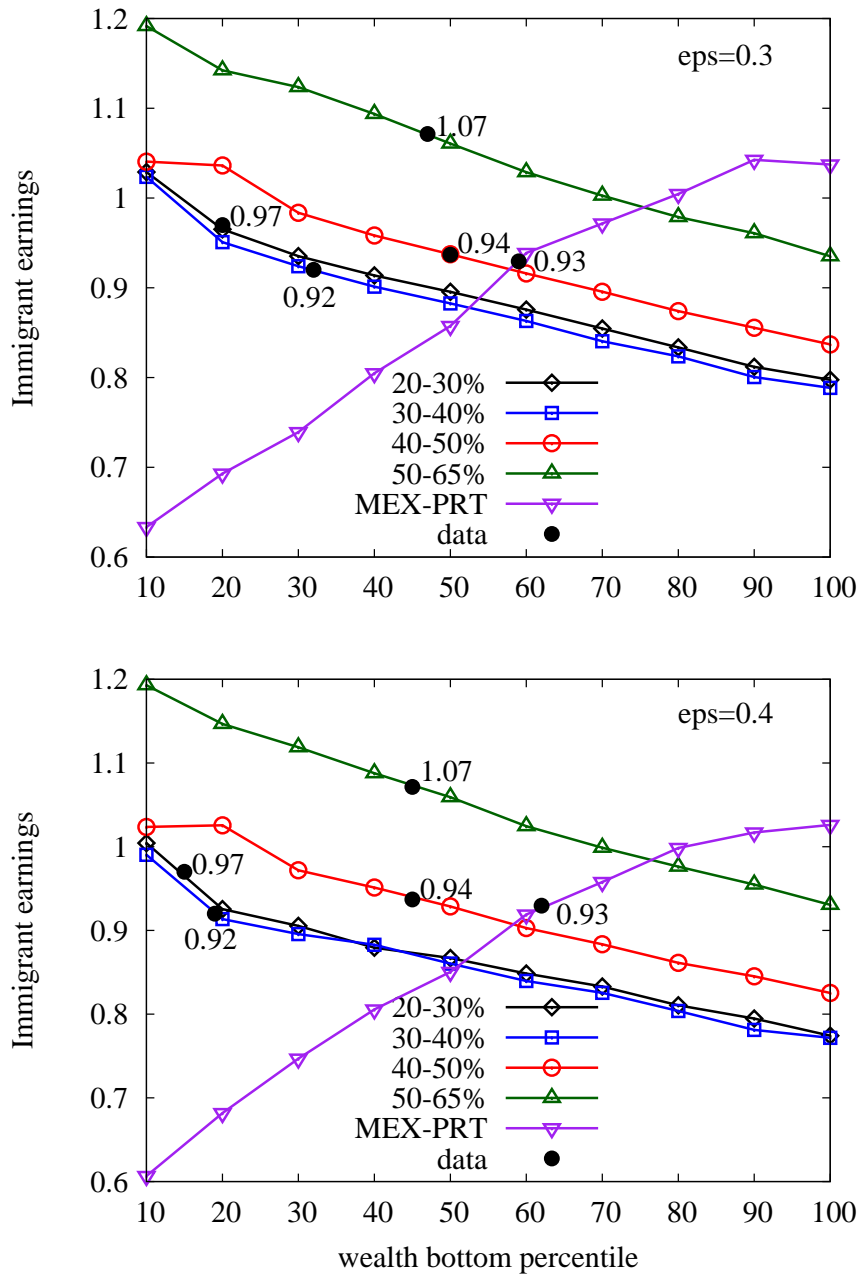
Model PPP income is normalized by the 1990 U.S. GDP at PPP prices. Solid points represent simulated economies; lines are OLS regressions with per-capita income at PPP prices as an explanatory variable. Income elasticities (regression coefficients on GDP, PPP) are indicated in square brackets. Data on prices of non-tradables and education are from Heston, Summers, and Aten (2002).

Figure 3: Schooling and Income



Model PPP income is normalized by the 1990 U.S. GDP at PPP prices. Solid points represent simulated economies; lines are OLS regressions with per-capita income at PPP prices as an explanatory variable. Income semi-elasticities (regression coefficients on GDP, PPP) are indicated in square brackets. Data is from Cohen and Soto (2007)

Figure 4: Relative Earnings of Immigrants



Data on relative earnings of immigrants across countries is from Hendricks (2002), adjusted by the level of schooling of the immigrant population. Each curve corresponds to an economy with a per-capita PPP income relative to that of the U.S. in the percentile indicated. Wealth percentiles are for populations with similar schooling.