What Explains Schooling Differences Across Countries?*

Juan Carlos Córdoba†

and

Marla Ripoll‡

May, 2011

Abstract

This paper provides a theory that explains the cross-country distribution of average years of schooling, as well as the so called human capital premium puzzle. In our theory, credit frictions as well as differences in access to public education, fertility and mortality turn out to be the key reasons why schooling differs across countries. Differences in growth rates and in wages are second order.
Keywords: human capital, per capita income differences, life expectancy, public education spending, life cycle model
JEL classification: I22, J24, O11

*An early version of this paper circulated under the title "The Role of Education in Development".

†Department of Economics, Iowa State University, 279 Heady Hall, Ames IA 50010, e-mail: `cordoba@iastate.edu`.

‡Department of Economics, University of Pittsburgh, 4933 W.W. Posvar Hall, Pittsburgh PA 15260, e-mail: `ripoll@pitt.edu`.

1 Introduction

School enrollment data from UNESCO indicates that as of 2005, a child in Niger was expected to attend school for just 3.7 years, while a child in Norway would attend for 17.4 years. Although school enrollment rates are higher than a decade ago everywhere in the world, a salient feature of the data is that large educational gaps continue to exist between rich and poor countries. What explains this large schooling differences?

Schooling models in the tradition of Becker (1964), Ben-Porath (1967), Mincer (1974) and Rosen (1976), have been used to quantitatively assess the causes of schooling differences. A prominent example is the work of Bils and Klenow (2000a), BK henceforth. They argue that differences in the rates of economic growth are instrumental in explaining differences in schooling attainments. Figure 1 plots a measure of expected years of schooling based on enrollments for a large sample of countries in 2005 (horizontal axis) against schooling predicted by a version of BK's model (vertical axis).[1] Two observations emerge from this exercise. First, the model has a low explanatory power as illustrated by a low R^2 of 0.27. Figure 1 shows that the model overpredicts schooling in fast growing countries such as South Korea, Hong Kong, China, and Macao, while it underpredicts schooling in most high-income countries whose growth rates are modest. The reason for the low explanatory power is that in the data schooling is highly correlated with per capita income (correlation is about 0.8), but long-run growth is mostly uncorrelated with income levels (correlation is about 0.1). Second, the exercise assumes a high interest rate, between 9.5 to 10.3%, in order to match average schooling in the sample, an issue known in the literature as the "human capital premium puzzle" (Elias, 2003; Palacios-Huerta, 2006; Kaboski, 2007). Given that riskless rates are typically much lower, between 1 to 3%, this suggests the existence of non-trivial credit frictions in the accumulation of human capital.

This paper introduces credit frictions in an otherwise standard Ben-Porath model and shows that they significantly improve its ability to account for the cross-country distribution of average years of schooling, and the gap between the returns to schooling and the returns to riskless assets. Credit frictions bring about explanatory variables for schooling that are irrelevant in frictionless models. Specifically, optimal schooling in standard frictionless models is obtained from a simple income maximization problem in which family characteristics such as family income over the entire life cycle, family size, or parental bequests do not play any significant role. In contrast, these and other variables, such as the supply of public education, acquire central importance in the presence of credit frictions. Since these variables are disperse and correlated with per capita income, we find that models with credit constraints can better explain the dispersion of schooling across countries. Furthermore, our model predicts a wedge between asset returns and human capital returns similar to that found in the data.

Credit frictions can take different forms. In this paper we consider two alternative formulations: borrowing constraints for students and non-negative bequest constraints. Both alternatives turn out to produce very similar results in terms of steady state schooling, educational expenditures,

[1] The results are robust to different definitions of schooling and different versions of the model. Appendix A provides details of the exercise.

and human capital accumulation. In the benchmark model individuals are unable to borrow while in school. This assumption is consistent with evidence that indicates students face borrowing constraints. For instance, Jacoby (1994) uses Peruvian data to document that borrowing constraints affect primary school attendance and completion, and DeGregorio (1996) provides evidence that borrowing constraints affect human capital accumulation in both OECD and developing countries. Historically, the presence of borrowing constraints has been cited as an important element in understanding the evolution of schooling in the US. For instance, Goldin and Katz (2008) explain the high school movement in the US using a model in which educational returns, opportunity costs and capital constraints affect private human capital investments. They cite the presence of capital-market imperfections as key in explaining the emergence of support for public secondary schooling in the early twentieth century (p. 208). Finally, more recent evidence for the US indicates that more youth are borrowing constrained today than were in the early 1980s (Belley and Lochner 2007; Lochner and Monge-Naranjo 2008, Lovenheim 2008).[2]

We also consider, as an alternative to the benchmark, a model with a non-negative bequest constraint. This alternative assumption can be justified both theoretically and empirically. Non-negative bequests effectively imply that parents cannot legally impose debt obligations on their children. As discussed in Rangazas (2000), one may think that such a constraint may be circumvented by an informal agreement in which the child could agree to pay a portion of the consumption and education expenses to the parent when the child becomes a worker. Such an agreement seems quite unlikely for primary and secondary education expenditures, which are the relevant ones for most countries in the world. Even at the college level, two-sided altruism could be consistent with such an informal agreement, but as discussed in Chakrabarti, Lord and Rangazas (1993), uncertainty about the degree of the child's altruism toward the parent would keep education investment levels in children inefficiently low. On the empirical side, Gale and Scholz (1994) provide evidence that gifts from children to parents occur in a very small percentage of families in the US.

Our benchmark model is a life-cycle economy populated by individuals who have access to a public education system, face a stochastic life span, and are altruistic toward their children. Individuals receive an optimal transfer, or bequest, from their parents upon birth and make their own lifetime consumption and human capital decisions, including years of schooling and educational investments. In our model, the parental bequest at birth is the only state variable and it collects all the effects of the family on their children. We model demographic variables (fertility and mortality) as well as public education in detail, as those variables turn out to be important in the presence of credit frictions. Our modelling of the public education system seeks to capture two salient features of the data. First, public education is the predominant form of education. It accounts on average for 83.7% and 78.1% of primary and secondary enrollment respectively around the world. Second, richer countries invest significantly more resources in education per pupil than poor countries. In the model, the government finances schooling for a given number of years. Individuals can use their own funds to complement the given public funds (the intensive margin) and/or to finance

[2] Carneiro and Heckman (2003) document evidence against short-term borrowing constraints for US college students from the NLSY79, but they highlight the importance of family resources during the entire life-time of the student (p. 5).

additional years of schooling (the extensive margin).

We carefully calibrate the model to match educational data from OECD countries, assess its performance beyond the identification targets, and employ it for several purposes. First, we quantitatively assess the relative importance of differences in wages, mortality, fertility, and public education policies in explaining cross-country schooling differences. Second, we construct human capital stocks for a set of 74 countries and compare the results to existing alternatives. Finally, we study the implications of the model for the sources of cross-country income differences.

The analysis yields five main findings. First, we find that, depending on parameters, our model can explain between 83% to 94% of the standard deviation of schooling attainments across all the countries in the sample. A simple R^2 between the schooling data and the schooling predicted by the model is 0.71. These figures are remarkable given the difficulty frictionless models face in accounting for this dispersion, and given that only information about OECD countries is used to identify the key parameters. Our four next results identify the main sources of schooling differences according to the model.

We find that fertility rate differences are the most important determinant of schooling differences across countries. In a counterfactual exercise in which fertility rates around the world are equalized to US levels, the dispersion of schooling falls by 55%. This is because, in the benchmark, students rely heavily on parental resources to finance their consumption and other expenditures during schooling years. The presence of a large number of siblings dilutes the parental transfers per child causing a reduction in schooling years. This mechanism is further amplified over time as individuals with less schooling and earnings are able to transfer less resources to their descendants. The models thus displays a clear quantity-quality trade off: children in countries with high fertility rates stay less years in school and invest less in education. Such trade off does not occur when access to financial market is unrestricted.

The second most important determinant of schooling differences are mortality rates. The standard deviation of schooling falls by 31% when all mortality rates are equated to US levels. We also find that the bulk of this reduction is due to changes in adult mortality rather than child mortality. These results are consistent with empirical estimates but difficult to replicate by alternative models. There are two reasons why our model is more successful in this regard. First, we do not assume an unrealistically high interest rate and therefore future earnings are discounted less heavily in our model than in alternative models. As a result, changes in life expectancy have a larger effect on the present value of labor earnings which is a key determinant of schooling decisions. Second, parental transfers also increase when mortality decreases as children with longer life span weight more in their parents' utility.

The third main finding is that wage differentials play only a minor role in explaining schooling differences. According to the model, equating steady-state wages to the US level reduces the dispersion of schooling by only 3%. A small income effect on schooling is consistent with the empirical evidence summarized by Haveman and Wolfe (1995). In our model, a small income effect is the result of two opposing forces. On the one hand, individuals expecting higher wages optimally reduce schooling years as a way to increase consumption during the period in which credit is restricted. On the other hand, richer parents leave larger transfers to their children, and therefore increase schooling of their descendants.

Finally, we assess the role of public education. We find that public education policies could significantly affect the dispersion of schooling attainments. For instance, equating the duration of the public education subsidy in all countries to the number of years offered in the US would reduce the dispersion of schooling by 38%. Most of the action in this counterfactual comes from African countries, which are the poorest in the sample. In these countries years of schooling increase between 2 and 4 years, which is significant, but still students would drop out of public school early on. Thus, sufficient availability of public education can help to reduce schooling dispersion, but only to a certain degree.

Our paper is related to a handful of others in the human capital formation literature, but many of them either do not try to explain differences in schooling and/or returns to schooling across countries, or they share limitations similar to those discussed above. They include Becker and Tomes (1986), Mankiw, Romer and Weil (1992), Klenow and Rodríguez-Clare (1997), Glomm and Ravikumar (1998) and (2001), Hall and Jones (1999), de la Croix and Licandro (1999), Rangazas (2000), Bils and Klenow (2000b), Boucekkine, de la Croix and Licandro (2002), Ferreira and Pessoa (2005), Schoellman (2007), Hendricks (2010) and Restuccia and Vandenbroucke (2010). Our model is similar in spirit to BK's model, but it incorporates a Ben-Porath production function, credit market frictions, and public education. Manuelli and Seshadri (2007) study an extended Ben-Porath model that incorporates two production functions, one for early childhood investments and another for the remaining of life. They show that if early investments are less intensive in goods, then schooling is a positive function of wages. Their model can produce a large dispersion of schooling based on the cross-country dispersion of wages. However, this mechanism requires a large elasticity of schooling to income, of around 0.7, which is substantially higher (even ten times higher) than existing estimates for the US (see Hauser and Daymont 1977, Datcher 1982, Behrman and Taubman 1986 and 1989, Becker and Tomes 1986, Hill and Duncan 1987, Haveman and Wolfe 1995, and Erosa, Koreshkova and Restuccia 2010).[3] Our Ben-Porath model assumes a unique human capital production function and predicts a small elasticity of schooling to income, which is consistent with the empirical evidence. In Section 4.4.3 we document further difficulties of a frictionless Ben-Porath model in accounting for school dispersion, returns to schooling, and expenditures. Our work is complementary with Erosa, Koreshkova and Restuccia (2010). They analyze the role of human capital formation in explaining cross-country income differences in a model with credit frictions. Although not their focus, a limitation of their analysis is that their calibrated model does not do well in explaining the dispersion of schooling across countries (see their Figure 3, p. 1444). Their model abstracts from fertility, mortality and certain public educational variables, such as the duration of schooling subsidies, which we find are key determinants of the variation of schooling across countries.

The remainder of the paper is organized as follows. Section 2 lays down the model and discusses

[3] For example, Erosa, Koreshkova and Restuccia (2010) calibrate their model using cross-sectional US data and find an elasticity of schooling to income of 0.16. Appendix B shows the calculation of the elasticity in Manuelli and Seshadri (2007).

the main features of its solution. The calibration of the model is presented in Section 3. Results, including robustness checks and comparison with the literature are discussed in Section 4. Section 5 presents some development accounting results. Section 6 shows that the results of the benchmark model also arise in a model with binding bequest constraints. Finally, Section 7 concludes.

2 The benchmark model

Consider an economy populated by altruistic individuals who live to a maximum of T years, survive with probability $\pi(a)$ to age a, where $0 \leq a \leq T$, go to school from age 0 to age s, work from age s until retirement at age R, and have f children at age F. [4] Individuals receive a bequest b from their parents at birth, subsidies for education from the government between ages $\underline{s}$ to $\overline{s}$, earn wages during working years, save and pay taxes. Although for simplicity we call them "bequests," these are actually *inter vivos* transfers. Finally, in the model agents take prices as given, specifically the after-tax wage rate per unit of human capital, w, the risk-free interest rate, r, and age-contingent consumption prices, $q(a)$. We assume there are annuity markets and prices are actuarially fair.

The distinguishing feature of the model is that individuals face borrowing constraints during schooling years. As a result, parental transfers are the only resource available to support consumption and educational expenditures during schooling years. Moreover, optimal schooling decisions cannot be obtained from a simple income maximization problem, as is the case in standard frictionless models, but need to be derived from the full utility maximization problem.

2.1 Human capital

Human capital is accumulated through schooling and experience. Human capital of an individual with s years of schooling and no experience is given by:

$$h(s) = \left(\int_0^s (i(a))^\beta \, da \right)^{\gamma/\beta} = \left(\int_0^s \left(\frac{e(a)}{p_E} \right)^\beta da \right)^{\gamma/\beta}, \tag{1}$$

where $\beta \in (0, 1]$ and $\gamma \in (0, 1]$.[5] Term $i(a) = e(a)/p_E$ represents investments in education services at age a, $e(a)$ are educational expenditures in units of consumption goods, and p_E is the relative price of education in terms of consumption goods (the numeraire). Expenditures $e\,(a)$ are composed of public subsidies, $e_p(a)$, and private funds, $e_s(a)$. Equation (1) is a version of the Ben-Porath (1967) technology for human capital accumulation (see Section 4.4.2 for details).

Parameter β governs the degree of substitution of educational investments while γ determines the degree of returns to scale in the production of $h(s)$. To better understand the role of γ and β, it is useful to consider the simple case $e(a) = e$. In this case, equation (1) becomes $h\,(s) = (e/p_E)^\gamma \, s^{\gamma/\beta}$

[4] In this paper we treat f as an exogenous variable. Challenging new issues emerge when endogenizing fertility in models with parental altruism (see Jones, Tertilt and Schoondbroodt, 2008). We discuss some of the issues and propose new solutions in Cordoba and Ripoll (2011).

[5] The restriction on β is required so that $\partial h(s)/\partial s > 0$.

so that γ is the elasticity of expenditures and γ/β is the elasticity of years of schooling. Consider now the returns to schooling, $r_s(s)$, implied by (1). They are defined as the derivative of log-earnings with respect to schooling, $d\ln\left(wh\left(s\right)\right)/ds$. Using (1), one finds that:

$$r_s\left(s\right)=\frac{\gamma}{\beta}h(s)^{-\frac{\beta}{\gamma}}\left(e\left(s\right)/p_E\right)^{\beta} \tag{2}$$

so that returns to schooling diminish with the amount of human capital $\left(\partial r_s\left(s\right)/\partial h(s)<0\right)$, and increase with the amount of expenditures at age s $\left(\partial r_s\left(s\right)/\partial e(s)>0\right)$. For the case $e(a)=e$, $r_s\left(s\right)$ takes the simple form $r_s(s)=(\gamma/\beta)/s$. This case highlights the role of s and γ/β as the key determinants of returns to schooling.

Finally, we assume human capital is further enhanced by experience at work. Human capital at age a, where $R\geq a\geq s$, is given by $h(a)=h(s)e^{\nu(a-s)}$ where $a-s$ is experience and ν are returns to experience[6].

2.2 Individual's problem

An individual with initial assets b chooses years of schooling s, assets at age s, $\omega\left(s\right)$, bequests b' for each of his/her f children, and a lifetime path of consumption and private expenditures in education $\{c(a),e_s(a)\}_{a=0}^{T}$ that solves the following problem:

$$V\left(b\right)=\max_{\substack{\{c(a),e_s(a)\}_{a>0}^{T}\\ s,\omega(s),b'}}\int_0^T e^{-\rho a}u\left(c\left(a\right)\right)\pi\left(a\right)da+e^{-\rho F}\phi(f)V\left(b'\right)\pi\left(F\right)$$

subject to:

$$\int_0^s\left(c\left(a\right)+e_s\left(a\right)\right)q\left(a\right)da+q\left(s\right)\omega\left(s\right)\leq b; \tag{3}$$

$$\int_s^T c\left(a\right)q\left(a\right)da+q\left(F\right)fb'\leq\int_s^R wh(s)e^{\nu(a-s)}q(a)da+q(s)\omega\left(s\right); \tag{4}$$

$$h\left(s\right)=\left(\int_0^s\left(\left(e_p(a)+e_s\left(a\right)\right)/p_E\right)^{\beta}da\right)^{\gamma/\beta};$$

$$e_s\left(a\right)\geq 0;\qquad b'\geq 0;\qquad \omega\left(s\right)\geq 0;\qquad 0\leq s\leq F;$$

and

$$e_p(a)=\begin{cases} e_p \text{ if } \underline{s}\leq a\leq\overline{s}\\ 0 \text{ otherwise}\end{cases};$$

where ρ is the rate of time preference, $u(\cdot)$ is a momentary utility function, $\pi\left(a\right)$ is the probability of surviving up to age a, and $\phi(f)$ is an altruism function that weights the utility of children in

[6] Given our focus on schooling decisions, we simplify the human capital formation after schooling years as in Bils and Klenow (2000a).

the utility of the parents. The momentary utility and altruism functions are assumed to have the forms:

$$u(c) \equiv \frac{c^{1-\sigma}-1}{1-\sigma} \text{ and } \phi(f) \equiv \phi f^{\psi} \text{ with } 0<\psi<1. \tag{5}$$

The first restriction of the problem, equation (3), is the present-value budget constraint during schooling years. During this period, individual's only resources are parental bequests which can be used to consume, invest in education and save. Students cannot borrow since $\omega(s)$ is restricted to be non-negative.[7] The second restriction, equation (4), is the budget constraint during working years. During this period individuals use savings and earned labor income to pay for consumption and to leave non-negative bequests b' to their children. Notice that individuals are not credit constrained when they become workers and parents since $s \leq F$.[8] Parents can then borrow for the purpose of providing optimal bequests to their descendants, and there are no unintended bequests. The last restriction describes the public education policy. The government provides subsidies for education between the ages of $\underline{s}$ to $\leq \overline{s}$ in the amount e_p. Finally, define human wealth as $W(s) \equiv \int_s^R wh(s)e^{\nu(a-s)}q(a)da$.

2.3 Prices

Age-contingent prices, $q(a)$, are assumed to be actuarially fair: $q(a) = e^{-ra}\pi(a)$, where r is the after-tax riskless interest rate. This assumption presumes the existence of well functioning annuity markets where individuals can diversify mortality risk.

An additional assumption is required for the borrowing constraint to bind in steady state. Unless some restriction is imposed, parents may leave bequests large enough so that their children would like to save rather than borrow. To prevent this possibility, the following assumption bounds the degree of altruism.

Assumption 1 $\frac{\phi(f)}{f} < e^{(\rho-r)F}$.

To gain some intuition about this assumption, consider the case $\phi(1) = f = 1$ which describes a simple dynastic economy with perfect altruism. In this case, Assumption 1 simplifies to $r < \rho$ which is a standard way to induce a borrowing constraint to bind. Alternatively, if $r = \rho$, then imperfect altruism in the form of $\phi(f) < f$ would be required to satisfy the assumption above.

2.4 Optimal choices

We now describe the most relevant properties of the optimality conditions (details are in Appendix C). We focus on steady state solutions of the problem. A feature of the solution is that consumption would jump at age s if the borrowing constraint is binding. Denote $c^S(s)$ and $c^W(s)$ the optimal

[7] We do not allow for informal arrangements in which children borrow from their parents and agree to pay back when they become adults. As argued by Rangazas (2000), such an agreement would be farfetched for primary and secondary school students.

[8] Restriction $s \leq F$ is not binding for any country in the calibration below.

consumption at age s of an individual as a student and as a worker respectively. Absent borrowing constraints, $c^S(s)$ would be equal to $c^W(s)$.

2.4.1 Bequests and consumption

Given that children are unable to borrow and that some minimal pre-school is required in order to accumulate positive human capital, parents would always find optimal to leave positive bequests (otherwise, children's consumption would be zero). The optimal amount of bequests satisfies the condition:

$$u'(c(F)) = \frac{\phi(f)}{f} u'(c(0)), \tag{6}$$

which equates marginal cost of bequeating to its marginal benefit (for the parent).

Optimal saving during schooling years and during working/ retirement years results in the following pair of conditions:

$$\begin{cases} u'(c(0)) = e^{(r-\rho)a} u'(c(a)) \text{ for } 0 \leq a \leq s \\ u'(c(s)) = e^{(r-\rho)(a-s)} u'(c(a)) \text{ for } T \geq a \geq s \end{cases}.$$

In words, individuals fully smooth consumption within each subinterval of their life (as student or as worker/ retired) but not across sub-intervals. Using the previous equation, (6) can be written as:

$$\frac{u'(c^S(s))}{u'(c^W(s))} = G \equiv \frac{f}{\phi(f)} e^{-(r-\rho)F} > 1. \tag{7}$$

The last inequality follows from Assumption 1 and implies that the borrowing constraint is binding.[9] Thus, even altruistic parents do not leave enough bequests to prevent their children's consumption to jump upon entering the labor force. The reason is that the relative low interest rate (relative to the rate of time preference) makes it optimal to consume early in life, while bequests are a way to postpone consumption (via children's consumption). G is the key margin in the model and measures the tightness of the credit constraint. Notice that G only depends on r, ρ, f and F but not on wages, educational subsidies, or any other level variable.

It is instructive to derive the shadow price of "credit" in this environment. Parents implicitly act as the financial institution by providing bequests to their children, while collecting benefits from the consumption stream of all their descendants. Parents can transfer resources to themselves from age 0 to age F at the price $q(F) = e^{-rF}\pi(F)$, but transfer resources to their children using the implicitly higher price $Ge^{-rF}\pi(F)$. Let r_b be the shadow interest rate associated to this transfer defined by $e^{r_b F} = Ge^{rF}$ or $r_b = r + \ln(G)/F > r$. Using equation (7) and (5), one finds that:

$$r_b = \rho + \frac{(1-\psi)\ln f - \ln \phi}{F}. \tag{8}$$

[9] Without a borrowing constraint individuals would pick $\omega(s)$ so that $u'(c^S(s)) = u'(c^W(s))$.

This result shows that the shadow price of credit increases with the number of children, and that the strength of the association depends crucially on the parameter $\psi \in (0,1)$. Therefore, the model predicts that credit is more expensive in high fertility countries. This is the key mechanism explaining the quantitative finding below that fertility differences are the main determinant of schooling differences across countries. Notice also that if parents are perfectly altruistic towards each child, meaning $\psi = \phi = 1$, then $r_b = \rho$ and the shadow price of credit is independent of the number of children. However, even in that case the borrowing constraint still binds if $r < \rho$.

Let E^* be the present value of total educational expenditures. The steady state optimal bequests are given by:

$$b = \frac{G^{-\frac{1}{\sigma}} W(s) + E^* \Omega(s)}{\Omega(s) + q(F) f G^{-\frac{1}{\sigma}}}, \qquad \text{where } \Omega(s) \equiv \frac{\int_s^T e^{-(\rho - r)a/\sigma} q(a)\, da}{\int_0^s e^{-(\rho - r)a/\sigma} q(a)\, da}.$$

2.4.2 Optimal education spending

The optimality condition for total education spending at age a, $e^*(a)$, can be written as:

$$\underbrace{q(a)}_{\text{marginal cost}} \geq \underbrace{\frac{1}{G} \int_s^R w \frac{\partial h(s)}{\partial e^*(a)} e^{v(t-s)} q(t) dt}_{\text{marginal benefit of } e(a)} \text{ with equality if } e_s^*(a) > 0. \tag{9}$$

The left hand side is the cost of investing one unit of consumption at age a, while the right hand side is the present value of the associated additional labor income flow adjusted by the factor $1/G$ (< 1). The presence of this last factor means that binding borrowing constraints reduce educational investments because the associated gains are less valuable. The equation also implies that countries with higher mortality rates undertake lower educational investment because individuals in those countries discount future earnings more heavily ($q(t)/q(a)$ is lower). Similarly, early investments in education (pre-school) would be particularly low in countries with high infant mortality.

The optimal educational investments, $e^*(a)$, have the form:

$$e^*(a) = \max\{\widehat{e}^*(a), e_p(a)\} \text{ for } a \in [0, s].$$

In this formulation, $\widehat{e}^*(a)$ is the amount that individuals would optimally like to spend in education at age a while, $e_p(a)$ is the public subsidy for education. Figure 2 illustrates functions $e^*(a)$ and $e_p(a)$, where $e_p(a)$ is the horizontal line e_p between ages $\underline{s}$ and $\overline{s}$ and zero otherwise. Private funds are needed to finance education between ages 0 to age $\underline{s}$ (*pre-school* for short) and after age $\overline{s}$ (*college* for short).[10] It may also be optimal to complement the public subsidy, for example if the subsidy is small, between the ages of $\underline{s}$ to $\overline{s}$. Finally, upward sloping curves correspond to different scenarios for $\widehat{e}^*(a)$. Since $q(a)$ decreases with age then $\widehat{e}^*(a)$ increases with age.

[10] In the calibration, interval $[\underline{s}, \overline{s}]$ would be different for each country to reflect the availability of public education in different countries.

Case 1 in Figure 2 illustrates a situation in which there is only private spending in education during pre-school since optimal schooling, s_1, is lower than $\overline{s}$. Case 2 illustrates a case in which private spending includes pre-school and some college since optimal schooling, s_2, is larger than $\overline{s}$. In this case, private spending includes pre-school and some college. Finally in Case 3, optimal schooling is $s_3 > \overline{s}$ but now there is also some private spending in the interval $[\underline{s}, \overline{s}]$.

2.4.3 Optimal schooling choice

The optimality condition for the choice of schooling years, s, is:

$$\underbrace{e_s(s) + \sigma \frac{\Delta u(s)}{u'(c^S(s))}}_{\text{marginal cost of } s} = \underbrace{\frac{1}{q(s)} \frac{1}{G} \frac{\partial}{\partial s} \left[\int_s^R w(a)\, h(s)\, e^{v(a-s)} q(a) da \right]}_{\text{net marginal benefit of } s}, \tag{10}$$

where $\Delta u(s) \equiv u\left(c^W(s)\right) - u\left(c^S(s)\right) > 0$. The left-hand side of this equation is the marginal cost of additional schooling which is given by additional schooling expenditures, $e_s(s)$, plus a cost associated to the consumption jump at age s, $\sigma \times \Delta u(s) / u'\left(c^S(s)\right)$. The right-hand side of the equation is the marginal benefit of schooling given by the present value of additional labor income associated to additional schooling.

Notice that a binding borrowing constraint reduces years of schooling because it increases the schooling marginal cost, due to the consumption jump, and reduces its marginal benefit (by the factor $1/G$). A feature of the optimal schooling decision is that only survival probabilities after age s are relevant for the calculations of s. In contrast, optimal educational spending, $e(a)$, is a function of survival probabilities at early ages as well.

An alternative way of writing the optimal schooling choice is:

$$r_s(s) = \nu + \frac{q(s)}{\int_s^R e^{\nu(a-s)} q(a) da} + \frac{G e_s(s)\, q(s)}{W(s)} + \frac{\sigma}{1-\sigma} \left(G^{(1-\sigma)/\sigma} - 1 \right) e^{-\rho s} \pi(s) \frac{c^S(s)}{W(s)}, \tag{11}$$

which provides a link between schooling choices and returns to schooling. Equation (11) can be used together with (2) to solve for s. Notice that the last term of equation (11) reflects the curvature of the utility function.

Regarding the relationship between schooling and wages, it is well-known that in the frictionless Ben-Porath model steady-state schooling is independent of wages w. The following proposition states that the same result holds in the credit constrained model under two polar cases: public schooling only or private schooling only.

Proposition 1. Optimal schooling, s^*, is independent of w in the following two cases: *(i)* $e^* = e_p$ for all a (a pure public system); *(ii)* $e_p = 0$ for all a (a pure private system).

Proof. See Appendix C.

According to the proposition, credit constraints alone do not imply that countries with higher wages will have higher schooling. In our calibrated model below we find a modest positive relationship between schooling and wages due to mixed nature of the educational systems.

3 Calibration

We use a calibrated version of the model to assess its quantitative implications for a cross-section of countries. For the calibration and quantitative exercises below we use the most recent data we could assemble, typically 2005, for a set of 74 countries. Sample size is determined by data availability. We assume that some parameters are common across countries while other are country specific.

3.1 Parameters common across countries

Tables 1 and 2 show the parameters assumed to be common across countries. Parameters in Table 1 are set exogenously from micro evidence, while those in Table 2 are calibrated. We set σ to 1.5, a standard value in the macro literature (Cooley and Prescott, 1995). Returns to experience, ν, is set to 2% implying that wages are multiplied by a factor of 2.23 for 40 years of experience. This is consistent with estimates by Bils and Klenow (2000a) and Murphy and Welch (1990) who find this factor to be 2.5 and 2.2 respectively.

Table 1. Parameters common across countries

Parameter	Concept	Value	Source / Criteria
σ	relative risk aversion	1.5	Cooley and Prescott (1995)
ν	returns to experience	2%	Bils and Klenow (2000a), Murphy and Welch (1990)
$\underline{s}$	starting schooling age	6	UNESCO
F	parenthood age	25	Satisfies restriction $s \leq F$ in sample
R	retirement age	65	Binding level in richer countries
ϕ	level in $\phi(f) = \phi f^{\psi}$	1	Perfect altruism when $f = 1$
ψ	degree of altruism	0.4	Birchenall and Soares (2009)
r	riskless interest rate	3%	Mehra (2003)
α	capital share	0.33	Gollin (2002)

Starting school age, $\underline{s}$, is set to 6 years, a value that represents quite well most countries in the data (UNESCO). The age of parenthood, F, is set to 25, an age at which the average student in all countries has finished school, so the restriction $s \leq F$ is satisfied. There is data on the childbearing age (age at first birth) available for some countries from the Demographic and Health Survey and the OECD. Although there are differences across countries in this dimension, our quantitative results change little for more realistic values of F in poorer countries. Retirement age R is set to be

65, a value that binds mostly for rich countries in the sample and allows for a more realistic working life span in these countries. Introducing $R = 65$ allows us to address the concern that the positive effects of longer life expectancy in schooling may be overstated for rich countries, where individuals do retire and their working life span is not proportionally as large as their life expectancy relative to poorer countries.

Regarding the altruism function $\phi(f) = \phi \cdot (f)^{\psi}$, ϕ is set to 1 so that parents care about their children as much as they care about themselves when $f = 1$. As we discuss below, $f = 1$ closely characterizes many rich countries, including the US. This implies that for rich countries the model approximately behaves as a standard dynastic model. The parameter ψ is a key parameter in the models as it determines the tightness of the credit constraint (see equation 8). It has been estimated by Birchenall and Soares (2009) using micro evidence on the value of children's life. They obtain three different values for this parameter (0.39, 0.47 and 0.58) depending on the data set used. As a standard value for this parameter has not been established yet, we choose the intermediate value of $\psi = 0.47$ for the benchmark and perform robustness checks. In particular, in the robustness section we re-calibrate the model for the lower-end value in Birchenall and Soares, $\psi = 0.39$, and also for the higher-end value, $\psi = 0.58$.

Next, to choose the appropriate interest rate, notice that r is a riskless rate at which parents can save while r_b is the (shadow) rate of borrowing. The riskless interest rate is set to 3%, which is the historic value for the US between 1802 and 1998 according to Mehra (2003). Our results are robust to set r to 0.6% which is the average rate for the post war period reported by Mehra. A value of 3% is also standard in the labor and health literatures (e.g., Birchenall and Soares, 2009). Finally, a parameter that is needed below to compute wages is α, the share of capital income in total income. We set this share to a standard value of 0.33 (Gollin, 2002).

The remaining parameters γ, β, ρ are calibrated jointly to match three targets for the average of OECD countries in the sample. Results are shown in Table 2. The targets are: expected years of schooling, returns to schooling, and private education expenditures as a percentage of GDP (for 2003). The motivation for our identification strategy is the following. Remember that for constant educational expenditures e equation (32) becomes $h(s) = (e/p_E)^{\gamma} s^{\gamma/\beta}$ and returns to schooling are $r_s(s) = (\gamma/\beta)/s$. This representation makes clear that parameter γ controls optimal expenditures in education while parameter γ/β controls returns to schooling. We choose targets for OECD countries rather than for the US, which is the standard practice in the literature, because US education statistics tend to be somehow atypical among rich countries. This is specially true for private education expenditures as a percentage of GDP, which according to the World Bank are around 2.1% for the US in 2003, while they are only 0.65% for the average of OECD countries.

Table 2. Calibrated parameters

Parameter	Concept	Value	Target
ρ	rate of time preference	4.69%	Average schooling OECD: 16.14 years
γ	elasticity of $h(s)$ to expenditures	0.3	Private education spending % GDP OECD: 0.65%
γ/β	elasticity of $h(s)$ to schooling time	1.5	Returns to schooling OECD: 8.28%

Our measure of expected years of schooling is *school life expectancy* (SLE), as reported by UNESCO. SLE is defined as "the total number of years of schooling which a child of a certain age can expect to receive in the future, assuming that the probability of his or her being enrolled in school at any particular age is equal to the current enrollment ratio for that age."[11] In particular, for a child of age a in year t, SLE is given by

$$SLE_a^t = \sum_{i=a}^{n} \frac{\text{enrollment}_i^t}{\text{population}_i^t} \times 100$$

where n is a theoretical upper age-limit for schooling.[12] We choose SLE as our measure of years of schooling because it corresponds more closely to our theoretical construction of steady state schooling.[13] Average SLE in OECD countries is 16.14 years, so we calibrate the model to predict that the average individual in OECD countries is 22.14 years old at school completion.

Regarding returns to schooling, we follow BK's methodology and their intermediate parameter values. Specifically, returns to schooling are computed as $0.18 \times (SLE)^{0.28}$. The corresponding returns to schooling for the average SLE for OECD countries is 8.28%. These returns incorporate, in principle, a premium for human capital risk which our deterministic model abstracts from. Following Palacios-Huerta (2006), we assume that this premium is small, of about 1.1%, so we match returns to schooling of around 7.18% for OECD countries. These three targets result in the following calibrated values: $\rho = 4.69\%$, $\gamma = 0.3$ and $\gamma/\beta = 1.5$, which imply $\beta = 0.2$.

3.2 Country-specific parameters

Countries differ in: schooling-related variables e_p, $\bar{s}$ and grade repetition probabilities; demographic variables $\pi(a)$ and f; and prices p_E and w. We now consider these categories in turn.

3.2.1 Schooling

e_p is computed using the variable public education expenditures per pupil (all school levels) as a percentage of GDP per capita available from UNESCO. To compute $\bar{s}$, the average age at which the government stops providing education subsidies in each country, we combine UNESCO data on SLE, the duration of primary and secondary, and the percentage of total expenditures financed by

[11] From UNESCO's site http://www.uis.unesco.org/i_pages/indspec/tecspe_sle.htm

[12] Grade repetition creates a wedge between years spent in school and effective years of education for the purpose of accumulating of human capital. We adjust for grade repetition as discussed below in section 3.2.1.

[13] Bils and Klenow contruct a variable similar to SLE as as their measure of schooling. An alternative is average years of schooling from Barro and Lee (2000) which is systematically lower than SLE because older generations typically have lower schooling levels. Although levels are different, both variables display similar cross-country dispersion.

the government at different schooling levels. In particular, $\bar{s}$ is computed for each country as

$$\begin{aligned}\bar{s} &= 6 + \text{duration prim\&sec} \times \frac{\text{public expenditures prim\&sec}}{\text{total expenditures prim\&sec}} \\ &\quad + (SLE - \text{duration prim\&sec}) \times \frac{\text{public expenditures terciary}}{\text{total expenditures terciary}}\end{aligned}$$

which weights the years of duration of primary and secondary, as well as the duration of terciary ($SLE-$ duration prim&sec) by the respective percentage of public spending in total expenditures. Data on the latter variable is not available for all countries in the sample. For those countries with missing data, we proceed in either of the following two ways. For some countries there is data on public education expenditures as a fraction of the total, but not disaggregated by levels. In these case we computed $\bar{s}$ as $SLE \times$ (public expenditures/total expenditures). Second, for countries with no available data on the public share of expenditures, we use the duration of compulsory schooling from UNESCO as a measure of $\bar{s}$. Notice that by using SLE in measuring $\bar{s}$ we want to capture the number of years a "representative child" in each country receives public education subsidies, which corresponds to the definition of $\bar{s}$ in the model.

Figure 3 illustrates our computed years of duration of the public education subsidy ($\bar{s} - 6$) versus SLE in the data. Notice that the vertical distance between the 45-degree line and $\bar{s} - 6$ in each country corresponds to the number of schooling years fully financed with private resources (excluding preschool). The graph illustrates a strong positive correlation between $\bar{s}$ and SLE, which in our sample amounts to 80%. It also indicates that in general, the representative child in each country is enrolled in school for more years than those provided publicly. Notice also how for a few of the very poor countries, enrollments seem to be low relative to reported public provision.

Finally, we introduce an adjustment to equation (1) to take into account grade repetition which varies widely across countries. In particular, we rewrite (1) as:

$$h(s) = \left(\int_0^s \left(\frac{d \cdot e(t)}{p_E} \right)^\beta dt \right)^{\gamma/\beta}$$

where d represents the probability of passing a grade. In other words, s still captures the number of years students are enrolled in school, but if a student repeats a grade, expenditures invested in education contribute proportionally less to the formation of human capital. We construct d for each country by using a weighted average of school repeaters in primary and secondary from UNESCO.

3.2.2 Demographics

Demographics variables in the model include mortality and fertility. In modeling the survival probability, we differentiate between mortality in early childhood, schooling years, and adulthood.

We assume that the survival probability to age a is given by

$$\pi(a) = \begin{cases} e^{-p_c a} & \text{for } a \leq a_c \\ \pi(a_c)e^{-p_s(a-a_c)} & \text{for } a_c \leq a \leq a_s \\ \pi(a_s)\frac{e^{-p(a-a_s)}-\xi}{1-\xi} & \text{for } a_s < a \leq T \end{cases},$$

where p_c is the hazard (mortality) rate during early childhood years $a \leq a_c$, and p_s is that during schooling years $a_c \leq a \leq a_s$. The survival probability during adulthood follows Boucekkine, de la Croix and Licandro (2002). Under this specification, the maximum age T is such that $e^{-p(T-a_s)} = \xi$ or

$$T = -\frac{\log(\xi)}{p} + a_s. \tag{12}$$

Our formulation of $\pi(a)$ is a compromise between computational convenience and realism. We choose $a_c = 5$ and $a_s = 25$ for all countries. This interval represents well the potential ages for students. In order to calibrate p_c, p_s, p and ξ, we use the 2006 life tables from the World Health Organization for each country in the sample. Specifically, we use the survival probability at age 5 to compute p_c from $\pi(5) = e^{-p_c 5}$; that at age 25 to compute p_s from $\pi(25) = e^{-p_c a_c - p_s(25-a_c)}$; and those at ages 55 and 85 to jointly solve for p and ξ from

$$\pi(55) = \pi(25)\frac{e^{-p(55-a_s)} - \xi}{1-\xi} \text{ and } \pi(85) = \pi(25)\frac{e^{-p(85-a_s)} - \xi}{1-\xi}.$$

Notice that each country will have a different T as implied by (12). Figure 4 illustrates our calibrated survival probability functions for a few countries in our sample and compares them with the data.

Finally, we measure fertility f from World Development Indicators in 2005 as the number of births per woman divided by two. We divide by two as in the model there is a single parent to f children.

3.2.3 Prices

We allow the price of education p_E to differ across countries. Although there is no available measure for p_E, we use as a proxy the relative price of government spending from the Penn World Tables (PWT, v.6.2., 2004). Since most education around the world is public, and the largest share of education costs is represented by wages, we think this is a reasonable proxy. Recall that PWT relative prices are PPP adjusted, and so are comparable across countries.

Lastly, we compute after-tax wages per unit of human capital using a standard aggregate formulation:

$$w = \frac{(1-\tau)(1-\alpha)y_t^{data}}{h_t}, \tag{13}$$

where y_t^{data} is output per worker at time t obtained from the PWT, $1-\alpha$ is the share of labor in aggregate income, τ is a proportional income tax computed as government spending as a percentage of GDP in 2005 from the World Development Indicators, h_t is human capital per worker and t is a baseline year.

An issue that emerges here is how to compute h_t. Our model has implications for the steady state value of this variable but actual economies seem to be far from steady state as suggested by two observations: first, younger cohorts have significant more schooling than older ones; and second, the age composition of the population, which is needed to compute average human capital, is changing substantially in many countries. On the other hand, computing h_t out of steady state using (1) is currently unfeasible given the lack of historical series on educational expenditures per pupil across countries.[14]

With these limitations in mind, we proceed by computing h_t in a way that is roughly consistent with our approach. First, we define $h_t(s_t) \equiv h(s)\left(\frac{s_t}{s}\right)^{\gamma/\beta}$, which is a version of $h(s)$ adjusted by the fact s_t may differ from its steady state level s. The adjustment is motivated by the fact that $h(s) \approx i^{\gamma} s^{\gamma/\beta}$ when i is the same for all ages in the interval $[\underline{s}, s]$. The definition above implies that if $s_t = s$ then $h_t(s) = h(s)$. We choose $t = 2000$ and use average years of schooling, s_t, among adult population from Barro and Lee (2000).

We then define average human capital at time t as:

$$h_t = \Theta_t h_t(s_t) \tag{14}$$

where Θ_t is an adjustment for the average experience of the working force at time t. To compute Θ_t we use data on population by five-year age groups in each country from the World Population Prospects. Using the middle point for each age interval, together with the measure of average years of schooling from Barro and Lee (2000), we construct a weighted average of exponential functions $\exp(\nu \times (\text{age - schooling}))$, where the weights are given by population shares of the corresponding interval. In doing this computation we take into account that we calibrated the retirement age to be $R = 65$. Finally, plugging h_t into (13), w is obtained. In practice, term $h_t(s)$ is computed as part of the solution of the model while term Θ_t is directly computed from the data. Appendix C describes the solution algorithm.

4 Results

We use the calibrated model to study a cross-section of 74 countries in 2005. We now describe the main quantitative predictions of the model, as well as a number of counterfactual exercises.

4.1 Model's fit

Table 3 displays a number of measures to evaluate the model's performance. Recall that the model was calibrated to only match average years of schooling, returns to schooling and private education spending as a fraction of GDP for OECD countries. However, as Table 3 shows, the model does

[14] Most of the data on expenditures start in the 1990's.

a good job replicating average years of schooling in the whole sample, as well as private spending in education (for a subsample of 55 countries for which there is available data on spending). The model estimates an average return to schooling of 8.3% for the whole sample, while returns to schooling computed with BK's method using schooling data are 11.2%. Of this difference, 1.1% was embedded in the calibration as a premium for human capital risk, so the model underestimates average returns to schooling by about 1.8 percentage points.

Table 3. Model's performance

	Data	Model
Means		
Years of schooling	12.96	13.60
Returns to schooling	11.2%	8.3%
Private education spending % GDP	1.2%	1.2%
Standard deviations		
Years of schooling	3.35	2.78
Returns to schooling	2.1%	1.3%
Private education spending % GDP	1.25%	0.98%
Correlation between model and data		
Years of schooling	84.7%	
Returns to schooling	86.3%	
Private education spending % GDP	35.0%	

Although the model was not calibrated to match any standard deviation, Table 3 shows that it remarkably explains a substantial percent of the dispersion of years of schooling, returns to schooling and private education spending. Specifically, the model explains 83%, 61% and 79% of the dispersion of these variables respectively. Figures 5, 6 and 7 portray the model's performance in predicting schooling, returns to schooling and private education expenditures as a percentage of GDP respectively.

Finally, Table 3 reports the correlation between the model and the data for schooling, returns to schooling and private education expenditures. While the correlation is quite high for schooling and returns to schooling (84.7% and 86.3% respectively for the whole sample), it is lower at 35% for private education expenditures (subsample of 55 countries). However, the latter is mainly due to a few outliers in the sample in which reported private expenditures are quite high. If countries with expenditures larger than the average plus two standard deviations are dropped (Guyana with 5.9%, Iran 5.2%, and Iceland 3.7%) then the correlation between the model and the data goes up to about 51.4%, a much better performance. All in all, our streamlined model performs quite well in explaining key features of schooling around the world.

4.2 Human capital differences

Our model predicts human capital stocks that are different from standard estimates, such as those of BK or Hall and Jones, as these estimates abstract from investments in education beyond student's time. These estimates roughly define human capital as $\widehat{h}(s) \approx s^{\beta/\gamma}$.[15] Figure 8 displays the term $q \equiv h(s)/\widehat{h}(s)$ which is a measure of the "quality" of schooling. According to the model, standard measures of human capital should be adjusted downwards by an average of 60% for countries with per capita income below 50% of the US. In contrast, the adjustment is upwards for most richer countries but only as little as 10%. In other words, our model implies that the dispersion of human capital is larger than standard measures, mostly due to the low investments in education in poor countries.

Our quality adjustment to human capital series is more conservative that those implied by Manuelli and Seshadri (2007). They estimate that the quality of human capital in a country in the lowest decile is approximately one fifth of that of the U.S. Our equivalent measure is about two fifths. More importantly, while their estimate is driven mainly by cross-country differences in wages (and steady state demographics), in our model public education subsidies per pupil as well as fertility and mortality play the key role. Specifically, to the extent that in our model parental transfers serve as a substitute for credit markets during schooling years, the size and the number of children are an important determinant of private education spending.

Empirical evidence for the US finds a weak connection between education expenditures and schooling quality (Hanushek and Woessmann, 2007). However, as these same authors discuss, the question remains of "whether or not there is some minimum required level of resources even if impacts are not seen at higher levels of resources. This almost certainly is the case. It is consistent with the few "resource findings" ... about the availability of textbooks, the importance of basic facilities, the impact of having teachers actually show up for class, and similar minimal aspects of a school." (p. 67). We think our cross-country comparison of quality in Figure 8 exactly captures that: human capital in the poorest countries is only 40% of what estimates based only on schooling years would indicate, simply because of lack of minimal resources.

4.3 Counterfactuals

We now assess the relative quantitative importance of the exogenous parameters in explaining schooling differences across countries in our model. For this purpose, we equalize country specific parameters (p_c, p_s, p, f, e_p, $\overline{s}$, p_E and w) to their corresponding US value, one at a time. Table 4 presents the effects of these experiments on the standard deviation of schooling across countries, average cross country schooling, the variance of (log) parental transfers and average parental transfers.[16]

Regarding the dispersion of schooling across countries, we find that the strongest quantitative effect comes from equating fertility rates f in all countries to the US level. The model predicts

[15] This formula provides similar estimates as those of Bils and Klenow (2000a).

[16] As is standard in the literature, we use standard deviations to measure dispersion of time variables and variance in logs to measure dispersion of monetary variables.

that steady state schooling increases in around 3 years on average for a reduction of fertility in one child, a strong quantity-quality trade off. Moreover, since fertility rates are very different across countries, equating f to US levels reduces the standard deviation of schooling by around 56%.

Figure 9 portrays schooling profiles for the benchmark and for the counterfactual that equalizes fertility to US levels. Schooling substantially increases for the very poor countries and slightly decreases for countries with fertility rates below the US level, mostly European countries. Figure 10 and Table 4 illustrate the main mechanism at work. As families have less children, each child receives a larger parental transfer allowing them to finance consumption and educational investment, and to remain in school for a longer period. In fact, the variance of (log) parental transfers in the world falls by around 61% in this scenario. The key role of fertility in schooling decisions can be traced back to equation (8): credit constraints are tighter in high fertility countries.

Among demographic parameters, fertility is followed in quantitative importance by adult mortality. We find that an additional year of life expectancy increases schooling in around 0.11 years on average. This is consistent with the empirical estimates of BK (2000a) who find this effect to be between 0.125 and 0.25 (BK, footnote 27) although their model can only produce a factor of 0.03 to 0.04 (BK, p. 1176). Since life expectancy varies widely across countries, equating adult mortality p in every country to US levels reduces the standard deviation of schooling by 22.5%. In addition, equating all mortality rates (child p_c, student p_s and adult p) reduces the dispersion of schooling by 31%.

Key for understanding the role of mortality and life expectancy in schooling is the rate of discount. A high interest rate means that individuals discount future earnings heavily and therefore gains in life expectancy have only minor effects in present value calculations that are crucial for schooling decisions. As mentioned before, such high rate of return is required to produce realistic returns to schooling in frictionless models. Instead, in the presence of credit constraints large returns to schooling are compatible with a realistic low rate of return implying that future earnings are not discounted as heavily, and that gains in life expectancy have larger effects on schooling. A further important channel is altruism. A reduction in mortality increases the weight of children in their parents' utility and therefore increases bequests. In a frictionless model bequests play no role in schooling decisions. However, in the presence of borrowing constraints a larger bequest translates into more schooling years.

Table 4. Schooling counterfactuals (% change)

Parameter	$stdev(s)$	$mean(s)$	$var(ln(b))$	$mean(b)$
p_c	-3.7	0.4	-6.5	0.5
p_s	-3.5	0.3	-3.0	0.1
p	-22.5	2.4	-16.5	0.7
p_c, p_s, p	-30.8	3.2	-24.9	1.5
f	-56.2	3.5	-60.9	-6.4
e_p	22.7	-7.5	-17.6	2.4
$\overline{s}$	-35.3	1.8	-11.7	-1.7
p_E	2.0	-0.6	18.5	-13.0
w	-2.7	0.7	-53.7	67.2

Consider now the effect of wages on schooling. Our model predicts that schooling is mostly unresponsive to steady-state changes in wages: equating wages to US levels, which would be a drastic change for many poor countries, only reduces schooling dispersion by around 3%. A small response of schooling to income levels is consistent with a large empirical literature documenting a limited relationship between income and schooling (Hauser and Daymont 1977, Datcher 1982, Behrman and Taubman 1986 and 1989, Becker and Tomes 1986, Hill and Duncan 1987, and Haveman and Wolfe 1995). The small response of schooling to wages in our model is the result of two opposite forces that mostly cancel each other out: a substitution and an income effect. On the one hand, higher wages tend to reduce schooling as a way to increase consumption during schooling years, and therefore improve consumption smoothing when credit constraints are binding. On the other hand, higher wages tend to increase schooling because wealthier parents leave larger bequests. In fact, as reported in Table 4, average bequests increase by 67.2% mainly in poorer countries, which also reduces the dispersion of bequests.

Next, consider public education variables, $\overline{s}$ and e_p, the "extensive" and "intensive" margins of public education respectively. According to the model, one more year of public education availability, $\overline{s}$, translates into 0.22 more years of schooling on average. As seen in Figure 3, this extensive margin varies substantially across countries. As a result, the model predicts that equating $\overline{s}$ across countries to US levels, decreases the dispersion of schooling by 35.3%. The mechanism here is that individuals can take advantage of the subsidy only by staying longer in school. However, in many countries individuals drop out of public education due to other factors such as a particularly low life expectancy. We conclude that the potential effect of expanding the years of coverage of public education is large.

A perhaps surprising result is the effect of changes in public education subsidies e_p on schooling. We find that additional subsidies *decrease* rather than increase schooling. The reason, however, is simple. Since e_p is a subsidy only for education purposes, then its effect is to increase human capital and future labor earnings. Absent credit constraints individuals would borrow and increase consumption during all periods, particularly during schooling years. Binding credit constraints make this option unfeasible and therefore consumption would experience a larger jump at age s.

To improve consumption smoothing households adjust by reducing schooling years and private investments in education allowing them to consume more during the fewer schooling years but also more during working years thanks to their enhanced earning potential. As a consequence, the effect of equating e_p to US levels is to increase the cross-country dispersion of schooling by 22.7%.

4.4 Robustness analysis

In this section we provide a robustness check for the key parameter ψ and compare our calibrated human capital production function with others from related papers. In addition, we calibrate the frictionless version of our model and compare its quantitative fit with that of the credit constrained model.

4.4.1 Altruism

We first perform robustness checks on ψ, a key parameter of the model. According to Equation (8), ψ controls the tightness of the credit constraint. If $\psi = 1$, for example, the tightness of the constraint is independent of the number of children and therefore the same across countries. Our baseline calibration sets ψ to 0.47. For robustness purposes, we first use the value of ψ in the lower end of Birchenall and Soares' (2009) estimation: $\psi = 0.39$. We re-calibrate parameters γ, β, ρ for the same targets indicated in Table 2 and use the same values reported in Table 1 for all other parameters. The new values for the parameters are $\gamma = 0.305$, $\beta = 0.203$, $\rho = 4.78\%$. The low-ψ model predicts an average years of (expected) schooling in the sample of 13.21 years versus 12.96 in the data; average returns to schooling of 8.5% versus 11.2% in the data; and private education expenditures as a fraction of GDP of 1.07% versus 1.12% in the data. In terms of dispersion, the low-ψ model explains 94% of schooling's standard deviation, 73% of the standard deviation of returns to schooling, and 73% of the standard deviation of private education expenditures. The larger schooling dispersion in the low-ψ model is due to the tighter credit constraints. Specifically, for richer countries, whose fertility is close to one, changing ψ does not alter much the constraints. But for poorer countries with higher fertility rates, constraints become tighter, and schooling levels become lower with a lower ψ. Finally, the correlation between the low-ψ model and the data is 85.6% for schooling, 86.8% for returns to schooling, and 30.5% for private education expenditures as a fraction of GDP. As before, taking out the outliers countries, the latter increases to 48.3%. In sum, the low-ψ model is overall quite comparable to our baseline calibration.

It is also interesting to compare the counterfactuals under the low-ψ model. Table 5 reports those counterfactuals for which largest changes were obtained in Table 4, as a way to check for robustness. Overall, results are similar: the largest effects on schooling dispersion come from equalizing fertility and the duration of the public schooling subsidy. The quantitative effect is about the same for fertility (schooling dispersion drops 58.7%) and for $\overline{s}$ (dispersion drops 33.4% in the low-ψ model). Finally, the effects of equating mortality to US levels (child, student and adult) on schooling dispersion are slightly weakened (dispersion drops 22.9% in the low-ψ model, and 31% in the baseline). This reflects in part the fact that with a lower ψ, children are discounted more heavily in the parent's utility.

Table 5. Counterfactuals in the low-ψ model (% change)

Parameter	$stdev(s)$	$mean(s)$	$var(ln(b))$	$mean(b)$
p	-16.2	1.8	-15.0	0.4
p_c, p_s, p	-22.9	2.6	-23.6	1.2
f	-58.7	4.2	-64.7	-7.3
$\overline{s}$	-33.4	2.0	-13.7	-2.1
w	-1.2	0.4	-49.9	66.2

Next, we use the value of ψ on the higher end of Birchenall and Soares's (2009) estimation, $\psi = 0.58$, and re-calibrate parameters γ, β, ρ. We find that the re-calibrated values are $\gamma = 0.297$, $\beta = 0.198$, $\rho = 4.5\%$. The high-ψ model predicts an average years of (expected) schooling in the sample of 14.1 years (versus 12.96 in the data); average returns to schooling of 8.0% (11.2% in the data); and private education expenditures as a fraction of GDP of 1.52% (1.2% in the data). In terms of dispersion, relative to the baseline calibration, the high-ψ model predicts less dispersion of schooling (69% of that in the data), and of returns to schooling (46% of that in the data). In addition, the high-ψ model slightly overpredicts the dispersion of private education expenditures (112% of the data). In contrast to the low-ψ model, here schooling is not as disperse as in the baseline calibration because credit constraints are not as tight for poorer countries.

The correlation between the high-ψ model and the data is 79.8% for schooling, 84.7% for returns to schooling, and 32.1% for private education expenditures as a fraction of GDP (taking out the outliers countries, the latter increases to 47.4%). In sum, the high-ψ model is quite comparable with the baseline model in terms of the correlations with the data it predicts, but it explains relatively less dispersion of schooling and returns to schooling.

Table 6 reports counterfactuals as those in Table 5, but for the high-ψ model. Equalizing fertility continues to be the most important factor in reducing schooling dispersion in the high-ψ model (dispersion drops by 47%). Equalization of the duration of the public schooling subsidy also continues to be important (dispersion drops by 34.6%). The effects of equating mortality to US levels (child, student and adult) on schooling dispersion are increased (dispersion drops 41.6% in the high-ψ model, and 31% in the baseline). This is the case because with a higher ψ, there is lower discounting of children's utility in the parent's utility.

Table 6. Counterfactuals in the high-ψ model (% change)

Parameter	$stdev(s)$	$mean(s)$	$var(ln(b))$	$mean(b)$
p	-32.8	3.0	-18.0	1.6
p_c, p_s, p	-41.6	3.8	-25.7	2.4
f	-47.0	2.7	-55.5	-4.3
$\overline{s}$	-34.6	1.2	-9.0	-1.5
w	-6.2	1.4	-58.2	67.2

To summarize, the robustness checks for ψ indicate that, although there are some quantitative differences in the results across the three values we compare ($\psi = 0.39$, 0.47 and 0.58), the credit constrained model can account for a substantial fraction of the world schooling's dispersion. The strong prediction that emerges across all calibrations is that equating fertility and the duration of the public education subsidy to the US levels all around the world changes the dispersion of schooling in quantitatively important ways. The effect of mortality differences on schooling dispersion is relatively more sensitive to the alternative calibrations of ψ, but even at a minimum, equalization of mortality to US levels drops schooling dispersion by around 23%, a non-negligible amount.

4.4.2 Human capital production function

It is interesting to compare our human capital production function to existing estimates. Ben-Porath (1967) postulates the following law of motion for human capital $h(a)$ at age a:

$$\dot{h}(a) = z_h \left[n(a)h(a)\right]^{\gamma_1} i(a)^{\gamma_2} - \delta_h h(a), \tag{15}$$

where $0 \leq n(a) \leq 1$ is the fraction of time spent at school at age a, $i(a)$ are input goods, and δ_h is a depreciation rate. Let s be the age at which individuals stop full time schooling so that $n(a) = 1$ for $a \leq s$. The following proposition establishes a connection between Ben-Porath's formulation in (15) and our human capital production function in (1).

Proposition 2. The human capital stock $h(a)$ in Ben-Porath's model at age $a \leq s$ is given by:

$$h(a) = \left(\int_0^a e^{-\delta_h(1-\gamma_1)(a-t)} z_h \left(i(t)\right)^{\gamma_2} dt\right)^{\frac{1}{1-\gamma_1}}$$

which is identical to our human capital production function in (1) when $\beta = \gamma_2$, $\gamma/\beta = 1/(1-\gamma_1)$ and $\delta_h = 0$.[17]

Proof. Denote $M(a) = h(a)^{1-\gamma_1} = \int_0^a e^{-\delta_h(1-\gamma_1)(a-t)} i\left(t\right)^{\gamma_2} dt$ and notice that $h(a) = M(a)^{\frac{1}{1-\gamma_1}}$. Therefore,

$$\begin{aligned} \dot{h}\left(a\right) &= \frac{1}{1-\gamma_1} M(a)^{\frac{\gamma_1}{1-\gamma_1}} \left[z_h i\left(a\right)^{\gamma_2} - \delta_h \left(1-\gamma_1\right) M(a)\right] \\ &= z h\left(a\right)^{\gamma_1} i\left(a\right)^{\gamma_2} - \delta_h h(a) \end{aligned}$$

with $z = z_h/(1-\gamma_1)$, which corresponds to (15) when $n(a) = 1$ for $a \leq s$.

[17] We set $\delta_h = 0$ in the benchmark for simplicity. Our results are robust to different values δ_h.

Using our benchmark parameters in Table 2, the corresponding values for γ_1 and γ_2 are 0.33 and 0.2 respectively. On the other hand, estimates of the Ben-Porath function typically assume

$$\dot{h}(a) = \widetilde{z}_h \left[n(a)h(a)\right]^{\theta} - \delta_h h(a), \tag{16}$$

which abstracts from the input goods, $i(a)$. Now, if the true production function is (15), the optimal choice of $n(a)$ and $i(a)$ would result in the ratio of prices equal to the ratio of marginal products, or a relative price equal to $(\gamma_2/\gamma_1)\frac{n(a)h(a)}{i(a)}$. This means that $i(a)$ is proportional to $n(a)h(a)$. In that case, (15) could be written as (16) with $\theta = \gamma_1 + \gamma_2$.

Our implied estimate for θ is 0.53. Brownig et al. (1999, Table 2.3) summarize different estimates of θ found in the literature.[18] The four comparable estimates they report are 0.812 (Heckman), 0.52 (US Bureau of the Census), 0.578 (Haley), and the range $0.56 - 0.89$ (Brown). Our parameters of the human capital production function are clearly consistent with existing estimates.

4.4.3 The frictionless case

As a final robustness check, we calibrate the frictionless version of our model and compare its performance with the benchmark. The frictionless model is a simplified version of Manuelli and Seshadri (2007).[19] We find that the frictionless model can only account for a small fraction of the dispersion of schooling and the dispersion of other educational variables.

The optimal educational choices in the frictionless case can be obtained from the following net income maximization problem:

$$\max_{s,e_s(a),h(s)} \int_s^R wh\left(s\right)e^{\nu(a-s)}q(a)da - \int_0^s e_s\left(a\right)q\left(a\right)da$$

subject to

$$h\left(s\right) = z\left(\int_0^s \left(\left(e_p(a) + e_s\left(a\right)\right)/p_E\right)^{\beta} da\right)^{\gamma/\beta};$$

$$e_p(a) = \begin{cases} 0 \text{ if } a \leq \underline{s} \\ e_p(a) \text{ otherwise} \end{cases};$$

$$e_s\left(a\right) \geq 0.$$

Notice that individual's choices of schooling and education expenditures do not depend directly on the level of fertility f, and are also independent of bequests. In other words, the frictionless

[18] The corresponding parameter on their table is α and the relevant cases are those in which their β is restricted to be equal to α because the formulation above assumes the same exponent for $n(a)$ and $h(a)$.

[19] The main differences between our frictionless version and MS's are the formulation of the human capital production function for ages below six and for after schooling years. We consider a single production function describing human capital accumulation from age 0 to age s, while they assume one technlogy before age six and another technolog after age six. In addition, we assume that human capital formation in the job only comes from experience while they allow for further investments in time and goods.

model eliminates the connection between life-cycle family income and the choices of schooling and education expenditures. In addition, notice that parameters associated with utility such as σ and ρ do not play a role in the problem above. For instance, the equation that determines optimal schooling in the frictionless version of the model is:

$$r_s(s) = \nu + \frac{q(s)}{\int_s^R e^{\nu(a-s)} q(a) da},$$

which is similar to equation (11), except that now $G = 1$. This difference is key because with $G = 1$ there is no *direct* role for ρ, σ and fertility in determining schooling. There is, however, an *indirect* role of fertility in schooling. Average fertility affects the demographic structure of the population which in turn affects the determination of wages, still assumed to follow equations (13) and (14). This indirect demographic channel is present in our model as well as in Manuelli and Seshadri (2007). We find this indirect effect is weak and cannot explain a significant dispersion of schooling. Finally, notice that mortality enters directly in the equation above through $q(a)$, as is also the case in the model with frictions.

To provide a complete assessment, we now report the results obtained under four alternative calibrations of the frictionless model. Unless stated otherwise, we use the same parameters values reported in Table 1. For reference, recall that our benchmark model accounts for 83% of the dispersion of years of schooling, 61% of the dispersion of returns to schooling, and 79% of the dispersion of private education spending.

Our first calibration uses the same parameters used by Manuelli and Seshadri (2007) for the interest rate, of 7%, as well as for the human capital production function beyond pre-school. They use the formulation (15) with the following parameters values: $\gamma_1 = 0.63$ and $\gamma_2 = 0.3$. Notice that their implied value of θ corresponding to formulation (16) is 0.93, a value that is outside the range of estimates reported by Brownig et al. (1999, Table 2.3). As discussed in the previous section, the implied parameter values for our formulation would be: $\beta = \gamma_2 = 0.3$, and $\gamma/\beta = 1/(1-\gamma_1) = 2.7$ so that $\gamma = 0.81$. Recall from Table 2 that our benchmark calibration implies much more conservative numbers: γ is 0.3, while $\gamma/\beta = 1.5$. Under this parametrization the frictionless version of our model can only explain 34% of the schooling dispersion in the world, and 25% of the dispersion of returns to schooling, a poor performance compared to our model with credit frictions. It is interesting to notice that under these parameters the frictionless model predicts an average of 29 years of schooling for OECD countries. This high number is expected because in Manuelli and Seshadri "schooling" takes places over the whole life cycle. The key point though is that a frictionless Ben-Porath model with a unique production function cannot explain the schooling dispersion, even under Manuelli and Seshadri's parameters.

The second calibration identifies γ and β by matching two of the targets in Table 2: average years of schooling and average private education expenditures as a fraction of GDP in OECD countries. We obtain $\gamma = 0.251$ and $\beta = 0.183$. We find that this version of the model can only explain 28% of the schooling dispersion, and 15% of the dispersion of returns to schooling. By construction the model matches the average level of schooling in OECD countries, but it overestimates schooling in all other countries. In fact, while the world average schooling in the sample is 12.96 years, it is 15.61 in the model. In addition, the model underpredicts returns to schooling by a large amount: they are 11.2% in the data, but 5.97% in the model.

The third calibration uses an interest rate of 7%, the one used by Manuelli and Seshadri, instead of the 3% rate used in the benchmark. The calibration of γ and β in this case results in $\gamma = 0.428$ and $\beta = 0.23$. This calibrated model predicts a mean of schooling of 16.1 years, much higher than the data, but the mean of returns to schooling is 10.3%, somewhat closer to the data. In other words, in trying to better match returns to schooling, this calibration is still unable to correctly predict years of schooling. More importantly, under this calibration the model still suffers from the same issues mentioned above: it explains only 28% of the schooling dispersion and 23% of the dispersion of returns to schooling.

Finally, in a fourth calibration we choose r, γ and β in order to match all three targets in Table 2. This calibration should give the best chance to the model to simultaneously match schooling and returns to schooling. This fourth calibration yields $r = 5.5\%$, $\gamma = 0.36$ and $\beta = 0.222$. With these parameters the model matches average years of schooling, average private education expenditures as a fraction of GDP and average returns to schooling in OECD countries. But even this model overpredicts average years of schooling in the whole sample (15.87 years) and it only explains 28% of its dispersion.

These calibration exercises highlight the fact that the frictionless Ben-Porath model has problems in simultaneously matching levels of schooling and returns to schooling around the world, but more importantly, it explains only a small fraction of the schooling dispersion. The frictionless model fails to capture lower levels of schooling in poorer countries because it is missing the mechanism that family characteristics such as life-cycle parent's resources, family size, and parental transfers to children are important in determining schooling attainment.

5 Development accounting

In this section we study the implications of the model for cross-country income differences. This step requires to specify some additional aggregates and their determination. Assume that output is produced by a representative competitive firm operating the Cobb-Douglas technology

$$y_t = k_t^{\alpha}(Ah_t)^{1-\alpha}, \tag{17}$$

were h is the average human capital of the economy. In steady state, h is given by

$$h = \int_0^T h(a)n(a)da = h(s)\int_s^R e^{\nu(a-s)}n(a)da,$$

where $n(a)$ is the density of population of age a which is determined by demographic factors $\pi(a)$ and f. The firm hires labor and capital in competitive markets at pre-tax rates $\widetilde{w}$ per unit of human capital, and $\widetilde{r}$ per unit of capital. Profit maximization ensures the following conditions:

$$\widetilde{w} = (1-\alpha)\frac{y}{h}, \text{ and } \widetilde{r} = \alpha\frac{y}{k}. \tag{18}$$

Assume that individuals deposit their savings in mutual funds (MFs). MFs own the capital stock of the economy, and rent it to firms at the rate $\widetilde{r}$. MFs operate a constant returns to scale technology that transform p_I units of output into 1 unit of capital. Thus, p_I is the price of capital in terms of consumption goods, the numeraire. MFs are competitive and pay proportional taxes τ on earned income. Furthermore, assume that the following arbitrage condition between riskless bonds returns and physical capital returns holds:

$$r = (1-\tau)\frac{\widetilde{r}}{p_I} - \delta.$$

We continue to assume that the riskless rate r is exogenous and common across countries, but the price of investment goods p_I is different. We measure p_I from the Penn World Tables (PWT, v.6.2., 2004). Choosing $\delta = 10\%$ per year, the arbitrage condition above in combination with equation (18) imply a capital-output ratio k/y in each country. Finally, we compute A from equation (17) given information about y_t, k_t and h_t for a particular a period t. It is natural to pick $t = 2000$ given that we constructed h_t using equation (14).

Finally, steady state income y is computed as $y = (k/y)^{\frac{\alpha}{1-\alpha}} Ah$, which is another way to write (17). We use this expression to perform counterfactuals and evaluate the relative quantitative importance of the exogenous parameters in explaining the dispersion of per capita income. Table 7 summarizes these counterfactuals (for the purpose of comparison, here we also report the changes in standard deviation and mean schooling from Table 4).

Table 7. Per capita income counterfactuals (% change)

Parameter	$var(\ln(y))$	$mean(y)$	$stdev(s)$	$mean(s)$
p_c	3.5	-0.3	-3.7	0.4
p_s	-1.2	0.0	-3.5	0.3
p	-11.4	0.9	-22.5	2.4
f	-50.2	2.6	-56.2	3.5
e_p	-30.1	9.1	22.7	-7.5
$\overline{s}$	-7.5	-1.1	-35.3	1.8
p_E	22.7	-12.7	2.0	-0.6
A	-64.5	66.7	-2.7	0.7
p_I	-36.7	31.5	-3.1	0.7

There are a number of novel results on the table. First, among the mortality parameters, and similar to what we found for schooling, per capita income is most affected by adult mortality. We find that if mortality rates in all countries were equalized to the US level, then the variance of (log) per capita income would be reduced by 12.5%. Second, again similar to what we found for schooling, fertility rates have a large effect on per capita income. Specifically, equalizing fertility in all countries to the US level reduces the variance of per capita income by 52.5%. This result is impressive, and it is second only to changes in TFP levels A. Table 7 suggests that at least part of the effects of fertility on per capita income work through schooling years. Additional effects come through private spending in education. Lower fertility rates induce higher parental transfers per child, which are translated into higher private education spending and more years of schooling. Recall that the elasticity of spending in human capital is governed by parameter γ.

Third, opposite to what we found for schooling, equating public education spending per pupil to the US level has a much larger impact on per capita income than equating the maximum years of public subsidy $\overline{s}$. Specifically, equating e_p in each country to US levels decreases the variance of per capita income by 34.7%, while equating $\overline{s}$ only does by 8.8%. Again, the amplification effect on e_p works directly through parameter γ as a higher human capital "quality," which in turns results in higher per capita income.

Next, when TFPs A are equated to the US level the variance of per capita income is reduced by 60%, while in sharp contrast, the standard deviation of schooling only falls just by 3%. The first result is consistent with the findings of Manuelli and Seshadri (2007) and Erosa, Koreshkova and Restuccia (2010). TFP has a direct effect on per capita income, but in all these models it has an indirect effect through education spending, which impacts human capital via parameter γ. Finally, changes in the price of investment goods p_I have an important impact on per capita income dispersion, but a small one on schooling. We conclude that TFP levels and fertility rates are the two main determinants of per capita income dispersion. Quantitatively speaking, fertility rates have large effects both on schooling and per capita income dispersion.

6 A model with non-negative bequest constraints

In this section we analyze an alternative model in which the parent has full access to financial markets, makes optimal consumption and schooling choices on behalf of the children, but cannot leave a negative bequest to the children. Children live with the parent during schooling years, become independent upon finishing school and receive a non-negative bequest from the parent. Here we show that the steady state predictions of this model are almost identical to our benchmark model with credit constraints. As a result, parents underinvest in education and a human capital premium arises. Our benchmark model with credit constraints has the advantage of being much simpler, as individuals have a state vector with only one variable (bequest), while under this alternative model the state vector has three variables (parent's human capital, schooling, and initial bequest).

Individual's problem Consider an economy in which children live with the parent until age s, when they become workers. At that time they receive a non-negative bequest b from the parent.[20] Individuals have f children at age F, where $F \geq s'$. An individual with initial human capital h, schooling s and assets b chooses a lifetime path of consumption and private expenditures in education $\{c(a), e_s(a)\}_{a>0}^{T}$, schooling years for each child s', and bequests b' for each of the f children that solve the following problem:

$$
\begin{aligned}
V(h,s,b) \quad = \quad & \max_{\{c(a),e_s(a)\}_{a>0}^{T}, s', b'} \left\{ \int_s^T e^{-\rho(a-s)} u\left(c^W(a)\right) \frac{\pi(a)}{\pi(s)} da + \right. \\
& \left. \phi(f) e^{-\rho(F-s)} \left[\int_0^{s'} e^{-\rho a} u\left(c^S(a)\right) \pi(a)\, da + e^{-\rho s'} V\left(h', s', b'\right) \pi\left(s'\right) \right] \pi(F)/\pi(s) \right\}
\end{aligned}
$$

subject to:

$$
\int_s^T c^W(a)\, q(a)\, da + \int_0^{s'} f\left(c^S(a) + e_s(a)\right) q(F+a)\, da + q\left(F+s'\right) f b' \leq \int_s^R w h(s) e^{\nu(a-s)} q(a) da + q(s) b;
$$

$$
h' = \left(\int_0^{s'} \left(e_p(a) + e_s(a)/p_E\right)^{\beta} da \right)^{\gamma/\beta};
$$

$$
e_s(a) \geq 0; \qquad b' \geq 0; \qquad 0 \leq s' \leq F;
$$

where $c^W(a)$ denotes the consumption of the parent during working years, while $c^S(a)$ is the consumption of the child while living with the parent. The momentary utility and altruism functions are the same as before, and age-contingent prices $q(a)$ continue to be actuarially fair. Notice how the model above is less parsimonious than the one with credit constraints, as the state vector has three variables, rather than one.

An additional bound on the degree of altruism is required in order for the non-negative bequest constraint to bind. The following assumption, a slightly modified version of Assumption 1 in the paper, is enough for the bequest constraint to bind.

Assumption 1A $\frac{\phi(f)}{f} < e^{(\rho-r)F} \frac{\pi(F+s)}{\pi(F)\pi(s)}$.

To interpret this assumption, consider the special case of a dynastic model with constant mortality risk, $\phi(1) = f = 1$ and $\pi(F+s) = \pi(F)\pi(s)$. In this case, Assumption 1A reduces to $r < \rho$. Alternatively, if $r = \rho$ and the mortality risk is constant, then imperfect altruism in the form of $\phi(f) < f$ would be required to satisfy the assumption above. Either way, the bequest constraint is binding under Assumption 1A because it makes parents relatively impatient, more willing to consume early in life. This relative high consumption of the family early in life is costly and parents

[20]Similar results are obtained if the parental transfer occurs later in life since workers have unrestricted access to credit markets.

would like their children to help pay for it by leaving negative bequests, which forces the bequest constraint to bind.

We now show that a binding bequest constraint means that consumption jumps at age s when individuals leave their family and become workers. Such jump resembles the consumption jump of the model with credit constraints. To see this, notice that optimality conditions for the consumption of the parent $c^W(a)$ and that of the child $c^S(a)$ yield the following intratemporal allocation of consumption in the household:

$$u'\left(c^W\left(F+a\right)\right)\frac{\pi(a+F)}{\pi(F)\pi(a)}=\frac{\phi(f)}{f}u'\left(c^S\left(a\right)\right)\text{ for }a\le s',$$

which represents the equalization of the marginal utility of the parent at age $F+a$ conditional on surviving to that age (left-hand side), with that of each child at age a for $a \le s'$ (right-hand side). Combining this equation with the intertemporal consumption allocation and assuming a steady state situation results in:

$$\frac{u'(c^S(s))}{u'(c^W(s))}=G\equiv\frac{f}{\phi(f)}e^{-(r-\rho)F}\frac{\pi(s+F)}{\pi(F)\pi(s)}>1, \tag{19}$$

where the last inequality follows from Assumption 1A. Equation (19) is almost identical to (7) except for the right-most term, which involves a ratio of surviving probabilities. The reason consumption jumps at age s is because children free themselves from family debt when they leave their home thanks to the non-negative bequest constraint.

In addition to the almost exact resemblance of the consumption allocation, we also find that the schooling and private education expenditures choices are almost identical in the models with non-negative bequest constraints and credit constraints. Specifically, optimal schooling is given by:

$$e_s(s)+\sigma\frac{\Delta u(s)}{u'\left(c^S\left(s\right)\right)}=\frac{1}{G}\frac{1}{q(s)}\frac{\partial}{\partial s}\int_s^R we^{\nu(a-s)}\frac{\partial h'}{\partial s'}q(a)da$$

which exactly coincides with equation (10), except that the G in the equation above corresponds to the one defined in (19). Thus, as long as $\pi(F+s)\approx\pi(F)\pi\left(s\right)$, the schooling choices in the two models will be almost identical. Finally, optimal education spending satisfies (interior solution):

$$\frac{q\left(F+a\right)q(s)}{q(F+s)}=\frac{1}{G}\int_s^R we^{\nu(t-s)}\frac{\partial h}{\partial e\left(a\right)}q(t)dt$$

which is almost the same as equation (9) in the text. In fact, if the probability of dying was constant, they would exactly coincide.

In terms of quantitative implications, we computed the ratio $\pi(F+s)/[\pi(F)\pi\left(s\right)]$ for all countries in our sample and found it to be close to one for most countries, except for a few of the very poor (where it was at most 1.25). As a result, the quantitative predictions of the model with non-negative bequests and the one with credit constraints are almost the same.

7 Concluding comments

Understanding why educational outcomes vary so much across countries is challenging and important. For example, a major question in economics is how much of the cross-country income differences are due to differences in human capital. Since human capital is unobservable, credible estimates must come from models that perform well along observable dimensions, such as schooling years and returns to schooling. The fact that matching schooling data is challenging serves as a powerful way to sort out competing models. To the extent of our knowledge, our model significantly outperforms existing alternatives along these observable dimensions.

We carefully model key aspects of schooling decisions for a typical agent in an environment with credit frictions. The model provides new insights regarding the sources of differences in schooling as well as per capita income. Of major importance are demographic factors such as fertility and mortality. Our model suggests that controlling for demographics, income effects on schooling must be weak, a prediction that is consistent with a variety of empirical studies.

Our theory could explain the experience of Sub-Saharan African countries during the last 40 years. Per capita income in these countries mostly stagnated but schooling outcomes have improved. Although this is a topic we leave for future research, our model suggests that lower fertility, larger life expectancy and increased access to public education could explain these improved schooling outcomes. The natural next step in this research agenda is to study what explains the differences in demographics in the presence of long-term credit frictions. We leave this topic for future research.

A Figure 1

To construct Figure 1 we use Bils and Klenow's (2000) formula (11) (pp. 1164) with $\zeta = 0$ (no income effects):

$$(1+\mu)\, w(s)h(s) = \int_s^T \left[f'(s) - g'(t-s)\right] e^{-r(t-s)} w(t)h(t)dt.$$

where s is schooling, μ is the ratio of school tuition to the opportunity cost of student time, w is the wage, h is human capital, T is life expectancy, function $f(s)$ captures returns to schooling, while $g(s)$ captures returns to experience, and r is the real interest rate. We exclude income effects, as BK find them to be small. In BK's model $h(t) = h(s)e^{\gamma(t-s)}$, $w(t) = w(s)e^{g_A(t-s)}$ and $f'(s) = \theta/s^{\psi}$. Assume $g(t-s) = \gamma(t-s)$. In this case, the expression above can be written as:

$$1+\mu = \left[f'(s) - \gamma\right] \int_s^T e^{-r(t-s)} e^{g_A(t-s)} e^{\gamma(t-s)} dt = \frac{f'(s)-\gamma}{\eta} \left[1 - e^{-\eta(T-s)}\right]$$

where $\eta \equiv r - g_A - \gamma$. We use this equation to solve for s given the other parameters. From BK we use $\theta = 0.18$, $\psi = 0.28$, $\mu = 0.5$, and set r so that average schooling in the model equals average schooling in the data ($r = 7.73\%$). We set $\gamma = 0.01$, but results are not sensitive to this choice.

For T we use life expectancy at age 5 from the World Health Organization. For g_A we compute actual annual growth rates for the countries in the sample using data from the Penn World Tables for the period 1960-2000, and follow BK's procedure of setting g_A as an average between the actual rates and the average growth rate across countries. Notice that g_A captures expected growth, which in this case is measured by the long run per capita growth rate. In Figure 1 we measure s in the data using schooling life expectancy in 2005. Schooling life expectancy is the sum of current enrollment rates for all grades (primary to terciary). It captures the total number of years a child can currently expect to be enrolled at school. This variable corresponds the one constructed and used by BK as a measure of schooling. School life expectancy is available from UNESCO.

We perform two robustness checks. First, we cut the sample in Figure 1 to include the same countries we have in our own sample. The message is the same as in Figure 1: the model does not provide a good fit of the data ($R^2 = 0.17$). Second, we measure schooling as schooling life expectancy in 1999 (the earliest year available from UNESCO), and we use data for the period 1999-2003 (the latest years available from Penn World Tables) to construct the growth rates g_A. This captures BK's spirit of interacting schooling with future growth. Again, results are as in Figure 1 ($R^2 = 0.04$).

B Schooling in Manuelli and Seshadri's model

Equation (4) in Manuelli and Seshadri's paper reads (using our s which is $s+6$ in their notation):

$$\begin{aligned} Aw^{\upsilon(1-\gamma_1)-\gamma_2} &= B\left(1-e^{-(r+\delta)(R-s)}\right)e^{a_1(s-6)}\left[1-a_2\frac{1-e^{a_3(s-6)}}{1-e^{-(r+\delta)(R-s)}}\right]^{a_4} \\ &= B\left(1-e^{-(r+\delta)(R-s)}\right)e^{a_1(s-6)}\left[\frac{1-e^{-(r+\delta)(R-s)}-a_2\left(1-e^{a_3(s-6)}\right)}{1-e^{-(r+\delta)(R-s)}}\right]^{a_4} \\ &= B\left(1-e^{-(r+\delta)(R-s)}\right)^{1-a_4}e^{a_1(s-6)}\left[1-a_2-e^{-(r+\delta)(R-s)}+a_2e^{a_3(s-6)}\right]^{a_4} \end{aligned}$$

where $a_1 = (1-\gamma)(\delta_h + rv)$, $a_2 = \frac{r+\delta_h}{\gamma_1}\frac{(1-\gamma_1)(1-\gamma_2)}{\gamma_2 r+\delta_h(1-\gamma_1)} = -\frac{r+\delta_h}{\gamma_1}\frac{(1-\gamma_1)}{a_3}$; $a_3 = -\frac{\gamma_2 r+\delta_h(1-\gamma_1)}{(1-\gamma_2)}$; $a_4 = \frac{(1-\gamma)(1-\upsilon(1-\gamma_1))}{1-\gamma_1}$. According to this equation, the elasticity of schooling to w is given by:

$$\frac{ds/s}{dw/w} = \frac{\left(\upsilon(1-\gamma_1)-\gamma_2\right)/s}{-\frac{(1-a_4)(r+\delta)e^{-(r+\delta)(R-s)}}{1-e^{-(r+\delta)(R-s)}} + a_1 + a_4\frac{a_2a_3e^{a_3(s-6)}-(r+\delta)e^{-(r+\delta)(R-s)}}{1-a_2-e^{-(r+\delta)(R-s)}+a_2e^{a_3(s-6)}}}.$$

Using the parameter values provided in their paper, we find $(ds/s)/(dw/w) = 0.707$.

C Solution of benchmark model

To solve the individual's problem consider the associated Lagrangian:

$$\begin{gathered} \mathcal{L} = \int_0^s e^{-\rho a}u\left(c\left(a\right)\right)\pi\left(a\right)da + \int_s^T e^{-\rho a}u\left(c\left(a\right)\right)\pi\left(a\right)da + e^{-\rho F}\phi(f)V\left(b'\right)\pi\left(F\right) \\ +\lambda_1\left[b - q\left(s\right)\omega\left(s\right) - \int_0^s\left(c\left(a\right)+e_s\left(a\right)\right)q\left(a\right)da\right] \\ +\lambda_2\left[q\left(s\right)\omega\left(s\right) + \int_s^R w\theta\left(a-s\right)h\left(s\right)q(a)da - \int_s^T c\left(a\right)q\left(a\right)da - q\left(F\right)b'\right] \\ +\lambda_3\left[z\left(\int_0^s\left(\left(e_p(a)+e_s\left(a\right)\right)/p_E\right)^\beta da\right)^{\gamma/\beta} - h\left(s\right)\right] + \lambda_4 e_s(a) + \lambda_5\left[\omega\left(s\right)-\underline{\omega}\right]. \end{gathered}$$

The first order necessary conditions with respect to $c(a)$, $e_s(a)$, s and b' are, respectively:

$$\begin{cases} e^{-\rho a}u'(c(a))\pi(a) = \lambda_1 q(a) & \text{for } a \leq s \\ e^{-\rho a}u'(c(a))\pi(a) = \lambda_2 q(a) & \text{for } a > s \end{cases} \tag{20}$$

$$-\lambda_1 q(a) + \lambda_3 \frac{\partial h(s)}{\partial e(a)} \frac{\partial e(a)}{\partial e_s(a)} + \lambda_4 = 0 \tag{21}$$

$$\begin{aligned} \lambda_3 \tfrac{\partial h(s)}{\partial s} - \Delta u(s) \cdot e^{-\rho s}\pi(s) - \lambda_1 \left[\tfrac{\partial q(s)}{\partial s}\omega(s) + \left(c^S(s) + e_s(s)\right) q(s) \right] \\ +\lambda_2 \left[\tfrac{\partial q(s)}{\partial s}\omega(s) + \left(c^W(s) - wh(s)\theta(0)\right) q(s) + \textstyle\int_s^R w \tfrac{\partial \theta(a-s)}{\partial s} h(s) q(a) da \right] = 0 \end{aligned} \tag{22}$$

and

$$e^{-\rho F}\phi(f)V'(b')\pi(F) = \lambda_2 q(F) f \tag{23}$$

where

$$\frac{\partial h(s)}{\partial e(a)} = \gamma h(s)^{1-\frac{\beta}{\gamma}} e(a)^{\beta-1} p_E^{-\beta}, \tag{24}$$

$$\frac{\partial h(s)}{\partial s} = \frac{\gamma}{\beta} h(s)^{1-\frac{\beta}{\gamma}} \left(e(s)/p_E\right)^{\beta},$$

and

$$\Delta u(s) \equiv u\left(c^W(s)\right) - u\left(c^S(s)\right) \tag{25}$$

Further, the following envelope condition holds:

$$V'(b) = \lambda_1. \tag{26}$$

The equations above imply that:

$$\frac{\lambda_1}{\lambda_2} = \frac{u'\left(c^S(s)\right)}{u'\left(c^W(s)\right)} \text{ and} \tag{27}$$

$$\frac{\lambda_3}{\lambda_2} = \int_s^R w e^{\nu(a-s)} q(a) da = W(s)/h(s) \tag{28}$$

where we have used actuarially fair prices $q(a) = e^{-ra}\pi(a)$. Moreover, combining (23) and (26) yields

$$e^{-\rho F}\phi(f)\lambda_1^{child}\pi(F) = \lambda_2 q(F) f,$$

and using (20) this equation can be written as:

$$\frac{u'\left(c^{child}(0)\right)}{u'(c(F))} = \frac{f}{\phi(f)} \tag{29}$$

which dictates the relationship between the consumption of the child and the parent.

Notice that combining (25) and (27) yields:

$$\frac{\Delta u(s)}{u'(c^S(s))} = \frac{\Delta u(s) c^S(s)}{u'(c^S(s)) c^S(s)} = \left(\left(\frac{c^W(s)}{c^S(s)} \right)^{1-\sigma} - 1 \right) c^S(s) = \left(G^{(1-\sigma)/\sigma} - 1 \right) c^S(s) / (1-\sigma).$$

Education spending The optimal solution for education spending $e^*(a)$ has the form:

$$e^*(a) = \max\left\{\widehat{e}^*(a), e_p(a)\right\}$$

where $\widehat{e}^{*}(a)$ is the optimal solution for total education spending $e(a)$ when private spending is positive $e_s(a) > 0$.

Using (24) and (21), we can write:

$$\widehat{e}^{*}(a) = e^{*}(0)\,(q(a))^{-\frac{1}{1-\beta}} \tag{30}$$

where

$$e^{*}(0) = \left(\gamma h(s)^{-\frac{\beta}{\gamma}}\, p_E^{-\beta} W(s) \frac{1}{G}\right)^{\frac{1}{1-\beta}}. \tag{31}$$

For the purpose of solving the model (in a computer), the solution for $e^{*}(a)$ can be written in terms of age $\widehat{a}$ defined as the age at which $e^{*}(\widehat{a}) = e_p$. Defining $s_p \equiv \min\{s, \overline{s}, \max[\underline{s}, \widehat{a}]\}$, one finds that:

$$e^{*}(a) = \begin{cases} \widehat{e}^{*}(a) \text{ for } a \leq \min(s, \underline{s}) \\ e_p(a) \text{ for } \min(s, \underline{s}) \leq a \leq s_p \\ \widehat{e}^{*}(a) \text{ for } s_p \leq s \end{cases}$$

where

$$s_p = \min\{s, \overline{s}, \max[\underline{s}, \widehat{a}]\}$$

and $\widehat{a}$ denotes the final age at which spending in education is only public, i.e.,

$$e^{*}(\widehat{a}) = e_p(a).$$

These expressions allow to write $h(s)$ as:

$$h(s) = \left(\frac{e^{*}(0)}{p_E}\right)^{\gamma}\left[\left(\int_0^{\min(s,\underline{s})} q(a)^{-\frac{\beta}{1-\beta}} da + \int_{s_p}^{s} q(a)^{-\frac{\beta}{1-\beta}} da\right) + \left(\frac{e_p}{e(0)}\right)^{\beta}(s_p - \min(s, \underline{s}))\right]^{\gamma/\beta}. \tag{32}$$

An an example, for an individual who only goes to public school $s_p = s$ and her/his human capital is:

$$h(s) = \left(\frac{e^{*}(0)}{p_E}\right)^{\gamma}\left[\left(\int_0^{\underline{s}} q(a)^{-\frac{\beta}{1-\beta}} da\right) + \left(\frac{e_p}{e(0)}\right)^{\beta}(s - \underline{s})\right]^{\gamma/\beta}. \tag{33}$$

Alternatively, for an individual who goes to public school for primary and secondary and private college, $s_p = \overline{s}$ and her/his human capital is:

$$h(s) = \left(\frac{e^{*}(0)}{p_E}\right)^{\gamma}\left[\left(\int_0^{\underline{s}} q(a)^{-\frac{\beta}{1-\beta}} da + \int_{\overline{s}}^{s} q(a)^{-\frac{\beta}{1-\beta}} da\right) + \left(\frac{e_p}{e(0)}\right)^{\beta}(\overline{s} - \underline{s})\right]^{\gamma/\beta}. \tag{34}$$

Schooling Using equations (25), (27), and (28), together with the constraint $\underline{\omega} = 0$, we can write the optimality condition of the schooling choice (22) as:

$$\underbrace{\frac{1}{q(s)}\frac{\partial}{\partial s}\left[\int_s^R w(a)\,h(s)\,e^{v(a-s)} q(a) da\right] u'\left(c^W(s)\right)}_{\text{net marginal benefit of } s} = \underbrace{u'\left(c^S(s)\right) e_s(s) + \sigma \Delta u(s)}_{\text{marginal cost of } s}$$

so that the optimal schooling choice equates the marginal benefit to its marginal costs. An alternative way of writing the optimal schooling choice is

$$r_s(s) = g + \nu + \left[w(s)q(s) + \sigma \frac{1}{h(s)} \frac{\Delta u(s)}{\lambda_2} e^{-\rho s} \pi(s) + \frac{\lambda_1}{\lambda_2} \frac{e_s(s)\, q(s)}{h(s)} \right] \frac{\lambda_2}{\lambda_3}, \tag{35}$$

which provides a link between schooling choices and returns to schooling.

Consumption and bequests Using the definitions of $u(c)$, $\pi(a)$ and $q(a)$ into (20):

$$c(a) = \left[\lambda(a) e^{(\rho - r)a} \right]^{-1/\sigma}$$

where $\lambda(a) = \lambda_1$ if $a \leq s$ and $\lambda(a) = \lambda_2$ if $a > s$. Substituting this equation into (3) and (4) respectively and solving for λ_1 and λ_2 produces:

$$\lambda_1^{-\frac{1}{\sigma}} = \frac{b - \int_0^s e_s(a)\, q(a)\, da}{\int_0^s e^{-(\rho-r)a/\sigma} q(a)\, da} = \frac{b - E^*}{\int_0^s e^{-(\rho-r)a/\sigma} q(a)\, da} \tag{36}$$

and

$$\lambda_2^{-\frac{1}{\sigma}} = \frac{W(s) - q(F)\, f b'}{\int_s^T e^{-(\rho-r)a/\sigma} q(a)\, da} \tag{37}$$

where (28) has being used, and E^* is the present value of optimal private expenditures in education as given by:

$$E^* \equiv e^*(0) \left[\int_0^{\underline{s}} q(a)^{-\frac{\beta}{1-\beta}} da + \int_{s_p}^{s} q(a)^{-\frac{\beta}{1-\beta}} da \right] - e_p \int_{s_p}^{\min(s, \bar{s})} q(a) da. \tag{38}$$

Since we consider only steady state situations, let $b = b'$ in the two previous equations. Dividing one by the other, we derive the following optimal level of transfers:

$$b = \frac{W(s) G^{-\frac{1}{\sigma}} + E^* \Omega(s)}{\Omega(s) + q(F)\, f G^{-\frac{1}{\sigma}}}, \tag{39}$$

where

$$\Omega(s) = \frac{\int_s^T e^{-(\rho-r)a/\sigma} q(a)\, da}{\int_0^s e^{-(\rho-r)a/\sigma} q(a)\, da}.$$

Once b is obtained, one can go backwards and solve for $\lambda(a)$, $c(a)$ and Δu. In particular,

$$c^S(s) = \left[\lambda_1 e^{(\rho-r)s} \right]^{-1/\sigma}, \tag{40}$$

$$c^W(s) = \left[\lambda_2 e^{(\rho-r)s} \right]^{-1/\sigma}. \tag{41}$$

Solution algorithm We solve the model as follows. For some initial values of $e(0)$ and s, we first use equation (32) to compute $h(s)$. Second, we compute h_t and w using equations (14) and (13). Next, we solve for variables E^*, r_s, λ_3/λ_2, λ_3/λ_1, b, λ_1, λ_2, $c^S(s)$, $c^W(s)$, and Δu using equations

(2), (25), (28), (36), (37), (38), (39), (40), (41), together with:

$$\frac{\lambda_1}{\lambda_2} = \frac{u'(c(0))\,q(F)}{e^{-\rho F}u'(c(F))\,\pi(F)} = \frac{f}{\phi(f)}e^{-(r-\rho)F}$$

which is obtained using (20) and (29) in steady state and with actuarially fair prices. We iterate on the system of equations above by updating $e(0)$ and s using equations (31) and (35).

Proof Proposition 1: pure private education Equation (11) determines optimal schooling. We show that $r_s(s)$ is independent of w, and that $e_s(s)$ and $\Delta u(s)/u'\left(c^S(s)\right)$ are proportional to $W(s)$, so that w cancels out of the equation. In absence of public education, $e_s(a) = e^*(a) = \widehat{e}^*(a)$ and (32) becomes:

$$h(s) = \left[\left(\frac{e^*(0)}{p_E}\right)^{\beta}\int_0^s q(a)^{-\frac{\beta}{1-\beta}}da\right]^{\gamma/\beta}.$$

Using (31) and (28), and solving for $h(s)$ results in:

$$h(s) = \left[\frac{\gamma}{p_E}\frac{1}{G}\left(\int_0^s q(a)^{-\frac{\beta}{1-\beta}}da\right)^{(1-\beta)/\beta}\right]^{\gamma} W(s)^{\gamma}. \tag{42}$$

Plugging this result into (31):

$$e^*(0) = \frac{\gamma}{G}\frac{W(s)}{\int_0^s q(a)^{-\frac{\beta}{1-\beta}}da}. \tag{43}$$

Plugging this result into (30):

$$\widehat{e}^*(a) = \frac{\gamma}{G}\frac{W(s)}{\int_0^s \left(q(t)/q(a)\right)^{-\frac{\beta}{1-\beta}}dt} \tag{44}$$

Therefore, $e^*(a)$ is proportional to $W(s)$. Using (43) and (42) into(2):

$$r_s(s) = \frac{\gamma}{\beta}\frac{1}{\int_0^s q(a)^{-\frac{\beta}{1-\beta}}da} \tag{45}$$

Substituting (43) into (38) results in: $E^* \equiv \frac{\gamma}{G}W(s)$. Substituting this result into (39) produces:

$$b = \frac{1 + G^{\frac{1-\sigma}{\sigma}}\gamma\Omega(s)}{G^{\frac{1}{\sigma}}\Omega(s) + q(F)f}W(s), \tag{46}$$

Substituting this result in (36) and the result into (40)

$$c^S(s) = \left(\frac{1 + G^{\frac{1-\sigma}{\sigma}}\gamma\Omega(s)}{G^{\frac{1}{\sigma}}\Omega(s) + q(F)f} - \frac{\gamma}{G}\right)\frac{e^{-(\rho-r)s/\sigma}}{\int_0^s e^{-(\rho-r)a/\sigma}q(a)\,da}W(s) \tag{47}$$

Using the results above, (11) can be written as:

$$\frac{\gamma}{\beta}\frac{1}{\int_0^s q(a)^{-\frac{\beta}{1-\beta}}da} = \nu + \frac{1}{\int_s^R e^{\nu(a-s)}q(a)/q(s)da} + \frac{\gamma q(s)}{\int_0^s (q(t)/q(s))^{-\frac{\beta}{1-\beta}}dt}$$
$$+\frac{\sigma}{1-\sigma}\frac{\left(G^{(1-\sigma)/\sigma}-1\right)e^{-\rho s}\pi(s)}{\int_0^s e^{-(\rho-r)(a-s)/\sigma}q(a)\,da}\left(\frac{1+G^{\frac{1-\sigma}{\sigma}}\gamma\Omega(s)}{G^{\frac{1}{\sigma}}\Omega(s)+q(F)f}-\frac{\gamma}{G}\right).$$

Without credit frictions the last terms disappears. The credit constraint increases the right-hand side of the equation and therefore reduces schooling. Notice that wages do not appear in the equation.

Proof Proposition 1: pure public education In absence of private expenditures in education, (32) becomes $h(s) = (e_p/p_E)^{\gamma} s^{\gamma/\beta}$ and $r_s(s) = (\gamma/\beta)/s$. Moreover, $E^* \equiv 0$. Substituting this result into (39) produces:

$$b = \frac{G^{-\frac{1}{\sigma}}}{\Omega(s)+q(F)fG^{-\frac{1}{\sigma}}}W(s).$$

Substituting this result in (36), and the result into (40) yields:

$$c^S(s) = \frac{1}{G^{\frac{1}{\sigma}}\Omega(s)+q(F)f}\frac{e^{-(\rho-r)s/\sigma}}{\int_0^s e^{-(\rho-r)a/\sigma}q(a)\,da}W(s).$$

Equation (11) then becomes:

$$\frac{\gamma/\beta}{s} = \nu + \frac{q(s)}{\int_s^R e^{\nu(a-s)}q(a)da} + \frac{\sigma}{1-\sigma}\left(G^{(1-\sigma)/\sigma}-1\right)e^{-\rho s}\pi(s)\frac{1}{G^{\frac{1}{\sigma}}\Omega(s)+q(F)f}\frac{e^{-(\rho-r)s/\sigma}}{\int_0^s e^{-(\rho-r)a/\sigma}q(a)\,da}$$

Similar to the pure private education case, wages do not appear in the equation and credit frictions reduce schooling.

References

[1] Barro, R., Lee, J.W., 2000. International Data on Educational Attainment. Updates and Implications. NBER Working Paper 7911.

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

[3] Becker, G., Tomes, N., 1986. Human Capital and the Rise and Fall of Families. Journal of Labor Economics 4 (3), S1-S39.

[4] Belley, P., Lochner, L., 2007. The Changing Role of Family Income and Ability in Determining Educational Achievement. NBER Working Paper 13527.

[5] Behrman, J., Taubman, P., 1986. Birth Order, Schooling and Earnings. Journal of Labor Economics 4 (4 Part 2), S121-S145.

[6] Behrman, J., Taubman, P., 1989. Is Schooling 'Mostly in Genes'? Nature-Nurture Decomposition with Data on Relatives. Journal of Political Economy 97 (6), 1425-1446.

[7] Ben-Porath, Y., 1967. The Production Function of Human Capital and the Life Cycle of Earnings. Journal of Political Economy 75 (4, Part 1), 352-365.

[8] Bils, M., Klenow, P., 2000a. Does Schooling Cause Growth?. American Economic Review 90(5), 1160-1183.

[9] Bils, M., Klenow, P., 2000b. Explaining Differences in Schooling Across Countries?. Unpublished manuscript.

[10] Birchenall, J., Soares, R., 2009. Altruism, fertility, and the value of children: Health policy evaluation and intergenerational welfare. Journal of Public Economics 93 (1-2), 280-295.

[11] Boucekkine, R., de la Croix, D., Licandro, O., 2002. Vintage Human Capital, Demographic Trends, and Endogenous Growth. Journal of Economic Theory 104 (2), 340-375.

[12] Browning, M., Hansen, L. P., Heckman, J. J., 1999. Micro data and general equilibrium models. In J. B. Taylor and M. Woodford (ed.), Handbook of Macroeconomics, edition 1, volume 1, chapter 8, pages 543-633, Elsevier.

[13] Carneiro, P., Heckman, J., 2003. Human Capital Policy. NBER Working Paper 9495.

[14] Chakrabarti, S., Lord, W., Rangazas, P., 1993. Uncertain Altruism and Investment in Children. American Economic Review 83 (4), 994-1002.

[15] Cooley, T., Prescott, E., eds., 1995. Frontiers of Business Cycles Research. Princeton University Press.

[16] Cordoba, J. C., Ripoll, M., 2011. A Contribution to the Economic Theory of Fertility. Unpublished manuscript.

[17] Datcher, L., 1982. Effect of Community and Family Background on Achievement. Review of Economics and Statistics 64 (1), 32-41.

[18] De Gregorio, J., 1996. Borrowing Constraints, Human Capital Accumulation, and Growth. Journal of Monetary Economics 37, 49-71.

[19] de la Croix, D., Licandro, O., 1999. Life Expectancy and Endogenous Growth. Economic Letters 65 (2), 255-263.

[20] Elias, J., 2003. The Education Premium. Unpublished manuscript.

[21] Erosa, A., Koreshkova, T., Restuccia, D., 2010. How Important is Human Capital? A Quantitative Theory Assessment of World Income Inequality. Review of Economic Studies 77 (4), 1421-1449.

[22] Ferreirra, P. C., Pessoa, S., 2003. The Costs of Education, Longevity and the Poverty of Nations. Unpublished manuscript.

[23] Gale, W., Scholz, J. K., 1994. Intergenerational Transfers and the Accumulation of Wealth. Journal of Economic Perspectives 8 (4), 145-160.

[24] Glomm, G., Ravikumar, B., 1998. Flat-Rate Taxes, Government Spending on Education, and Growth. Review of Economic Dynamics 1 (1), 306-325.

[25] Glomm, G., Ravikumar, B., 2001. Human Capital Accumulation and Endogenous Public Expenditures. Canadian Journal of Economics 34 (3), 807-826.

[26] Goldin, C., Katz, L., 2008. The Race between Education and Technology. Belknap Press of Harvard University Press.

[27] Gollin, D., 2002. Getting Income Shares Right. Journal of Political Economy 110 (2), 458-474.

[28] Hall, R.E., Jones, C., 1999. Why Do Some Countries Produce So Much More Output Per Worker Than Others?. Quarterly Journal of Economics 114 (1), 83-116.

[29] Hanushek, E., Woessmann, L., 2007. The Role of School Improvement in Economic Development. Unpublished manuscript.

[30] Hauser, R. M., Daymont, T. N., 1977. Schooling, Ability, and Earnings: Cross-Sectional Findings 8 to 14 Years After High School Graduation. Sociology of Education 50 (July), 182-206.

[31] Haveman, R., Wolfe, B., 1995. The Determinants of Children's Attainments: A Review of Methods and Findings. Journal of Economic Literature 33(4), 1829-1878.

[32] Hendricks, L., 2010. Cross-country Variation in Educational Attainment: Structural Change or Within Industry Skill Upgrading?. Journal of Economic Growth, forthcoming.

[33] Hill, M., Duncan, G. J., 1987. Parental Family Income and the Socioeconomic Attainment of Children. Social Science Research 16, 39-73.

[34] Jacoby, H., 1994. Borrowing Constraints and Progress Through School: Evidence from Peru. Review of Economics and Statistics 76 (1), 151-160.

[35] Jones, L. E., Tertilt, A., Schoonbroodt, A. 2008. Fertility Theories: Can They Explain the Negative Fertility-Income Relationship? NBER working paper 14266.

[36] Kaboski, J., 2007. Explaining Schooling Returns and Output Levels Across Countries. Unpublished manuscript.

[37] Klenow, P. J., Rodríguez-Clare, A., 1997. The Neoclassical Revival in Growth Economics: Has It Gone Too Far?. In Bernanke, B., Rotemberg, J., eds., NBER Macroeconomics annual 1997. Cambridge, MA: MIT Press, 73-103.

[38] Lochner, L., Monge-Naranjo, A., 2008. The Nature of Credit Constraints and Human Capital. NBER Working Paper 13912.

[39] Lovenheim, M., 2008. Housing Wealth, Liquidity Constraints and College Enrollment. Unpublished manuscript.

[40] Mankiw, N. G., Romer, D., Weil, D. N., 1992. A Contribution to the Empirics of Economic Growth. Quarterly Journal of Economics 107 (2), 407-437.

[41] Manuelli, R., Seshadri, A., 2007. Human Capital and the Wealth of Nations. Unpublished manuscript.

[42] Mehra, R., 2003. The Equity Premium: Why is it a Puzzle?. Financial Analysts Journal, January/ February, 54-69.

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

[44] Murphy, K., Welch, F., 1990. Empirical Age-Earnings Profiles. Journal of Labor Economics 8 (2), 202-229.

[45] Palacios-Huerta, I. 2006. The Human Capital Premium Puzzle. Unpublished manuscript.

[46] Rangazas, P., 2000. Schooling and Economic Growth: A King-Rebelo Experiment with Human Capital. Journal of Monetary Economics 46 (2), 397-416.

[47] Restuccia, D., Vandenbroucke, G., 2010. Explaining Educational Attaintment Across Countries and Over Time. Unpublished manuscript.

[48] Rosen, S. A., 1976. Theory of Lifetime Earnings. Journal of Political Economy 84 (4), S45-S57.

[49] Schoellman, T., 2007. The Causes and Consequences of Cross-Country Differences in School Attainment. Unpublished manuscript.

Figure 1. Years of schooling - 2005
Data versus Bils and Klenow model

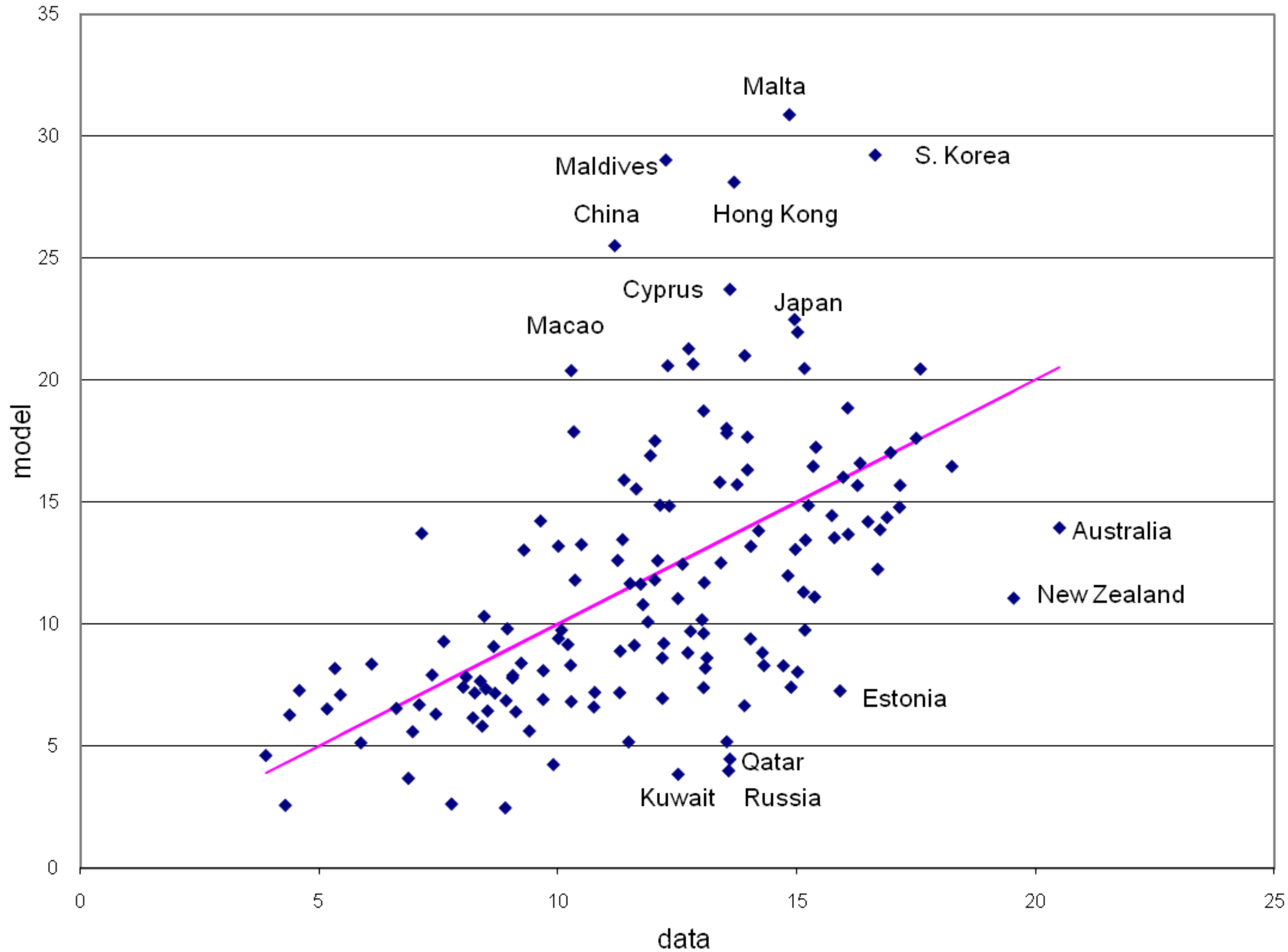
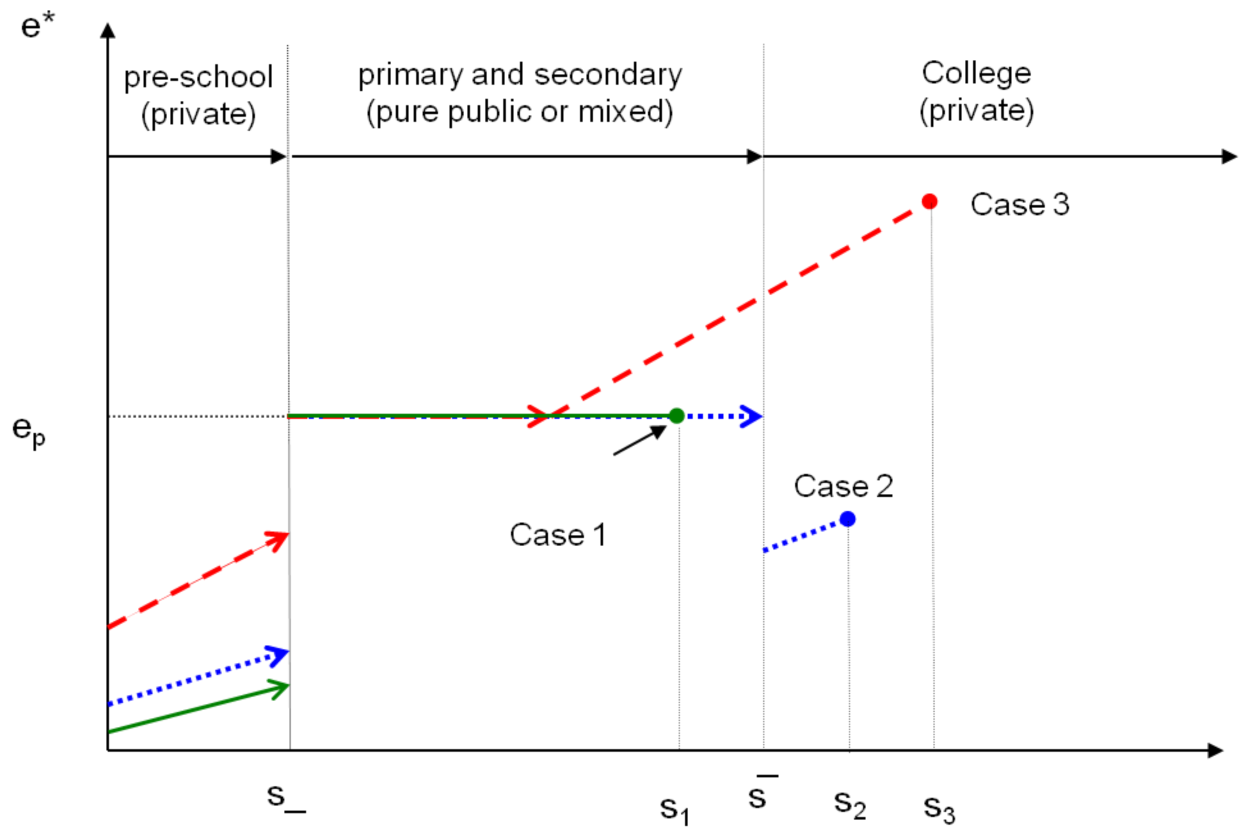


Figure 2. Individual expenditures in education: e*(a)



Case 1: Some public school **Case 2 :** Full public school + some private
Case 3: Full private and public school + some more private

Figure 3. Years of duration of public education subsidy versus school life expectancy in the data - 2005

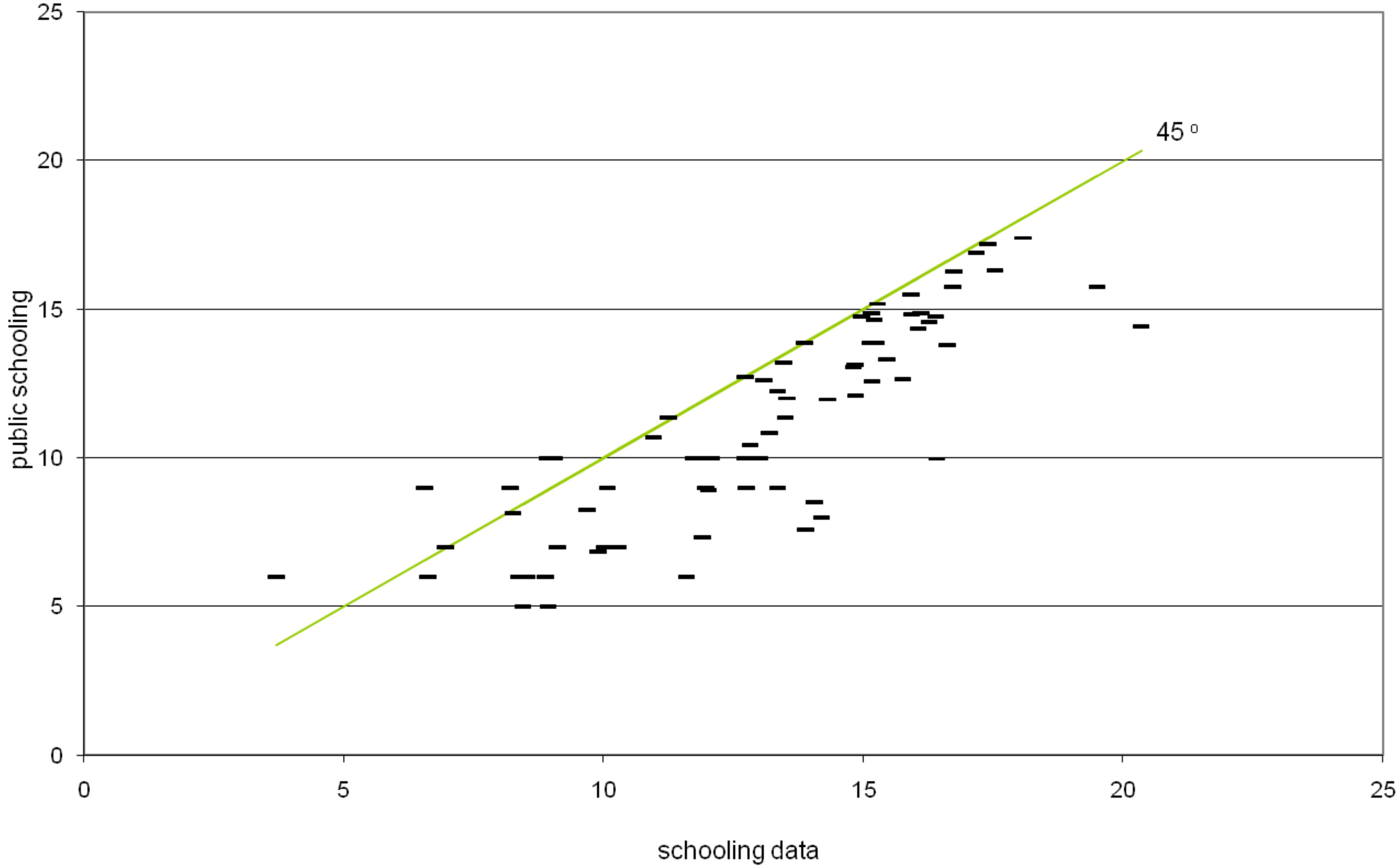


Figure 4. Survival probabilities at different ages
Precited (dashed) and Data (solid)

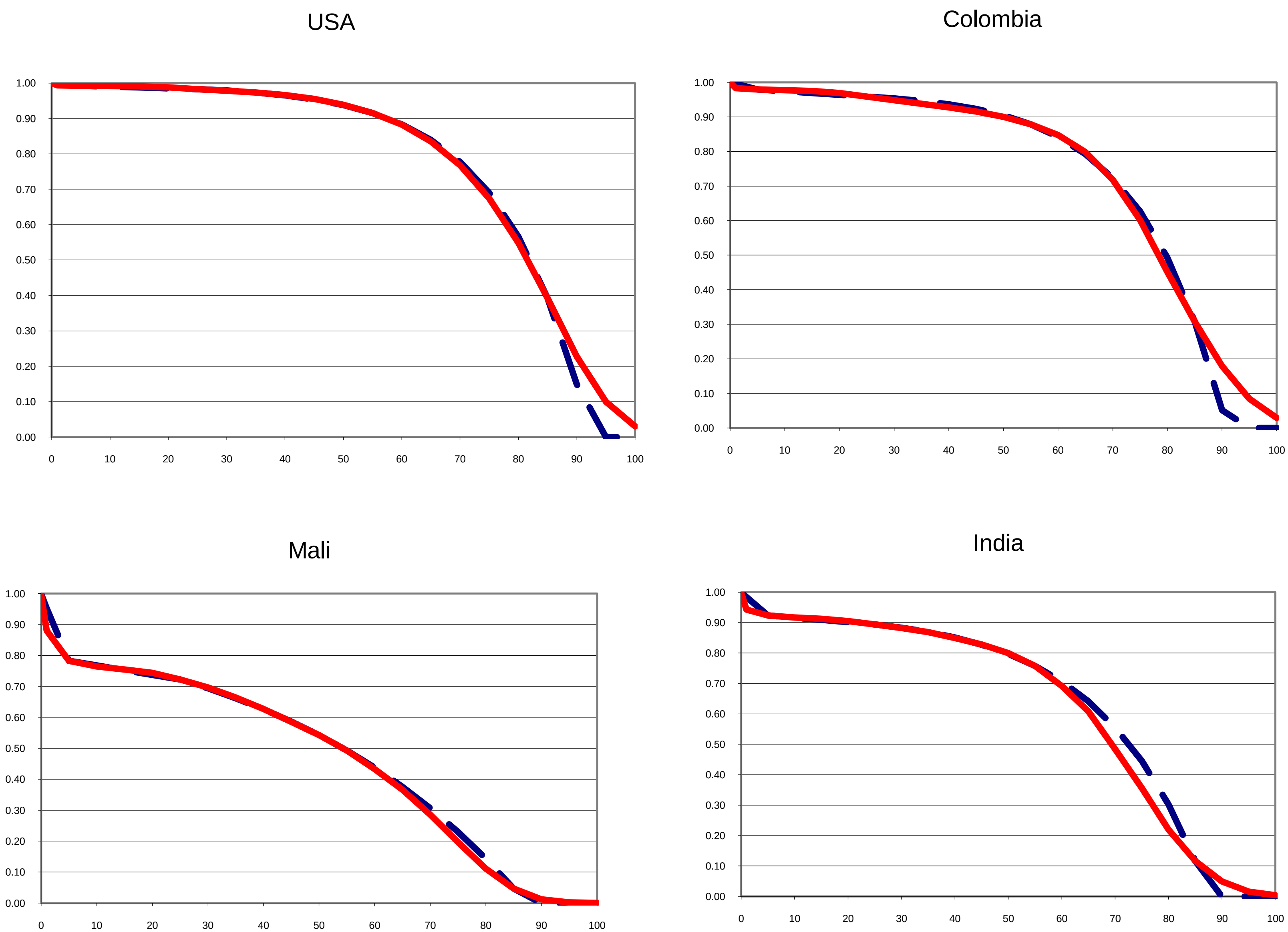


Figure 5. School life expectancy in the model and the data - 2005

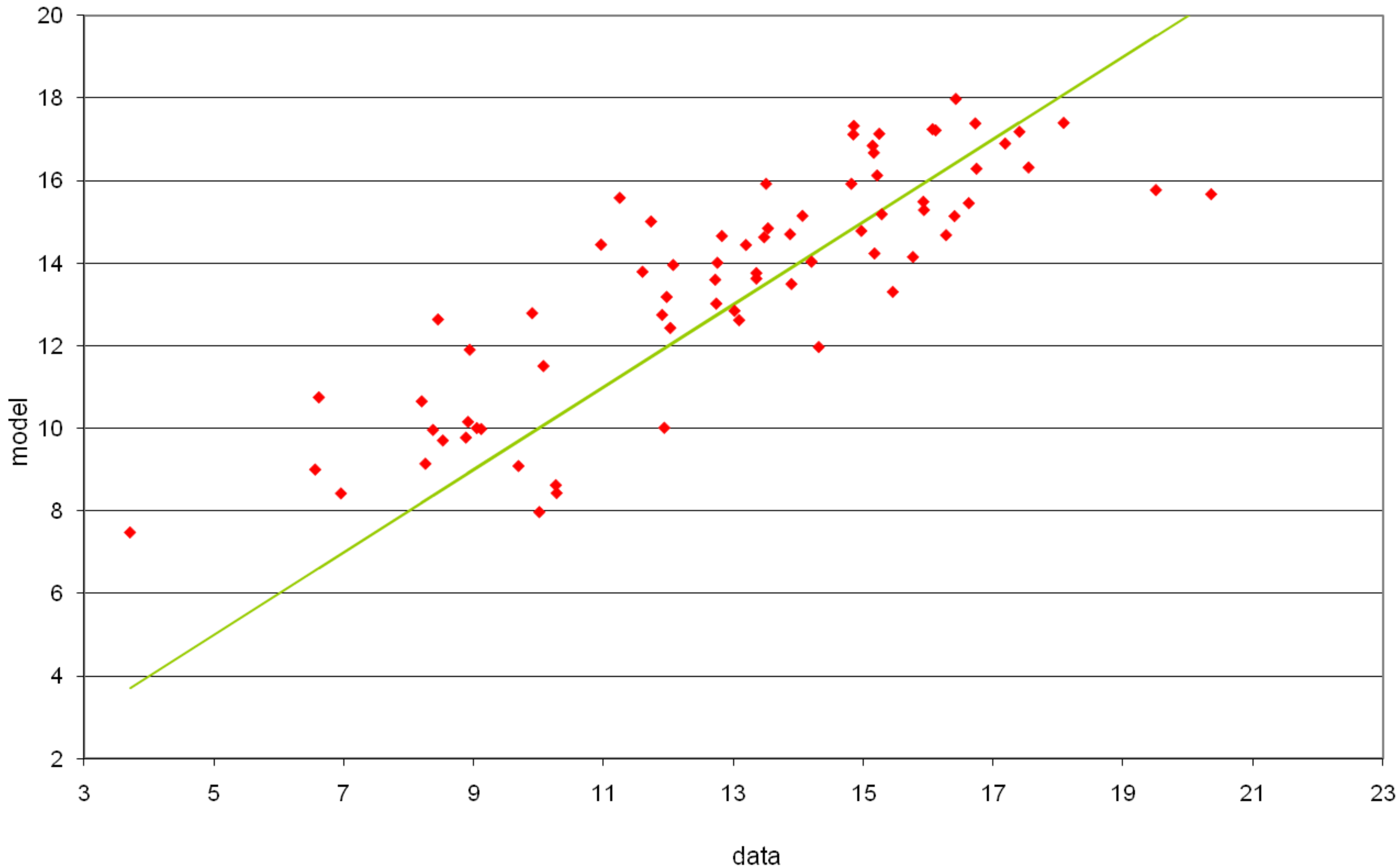
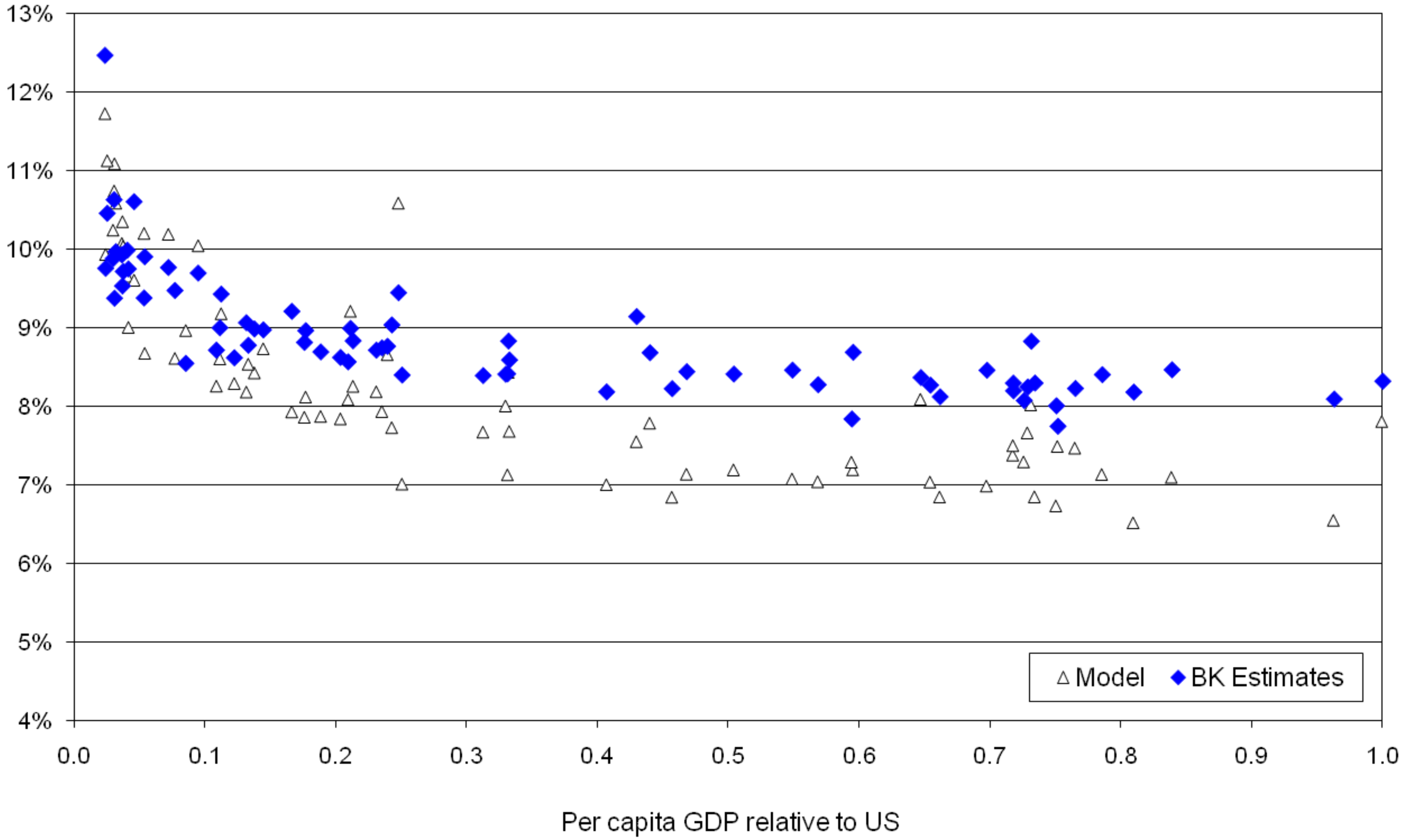


Figure 6. Returns to schooling
Model versus BK Estimates



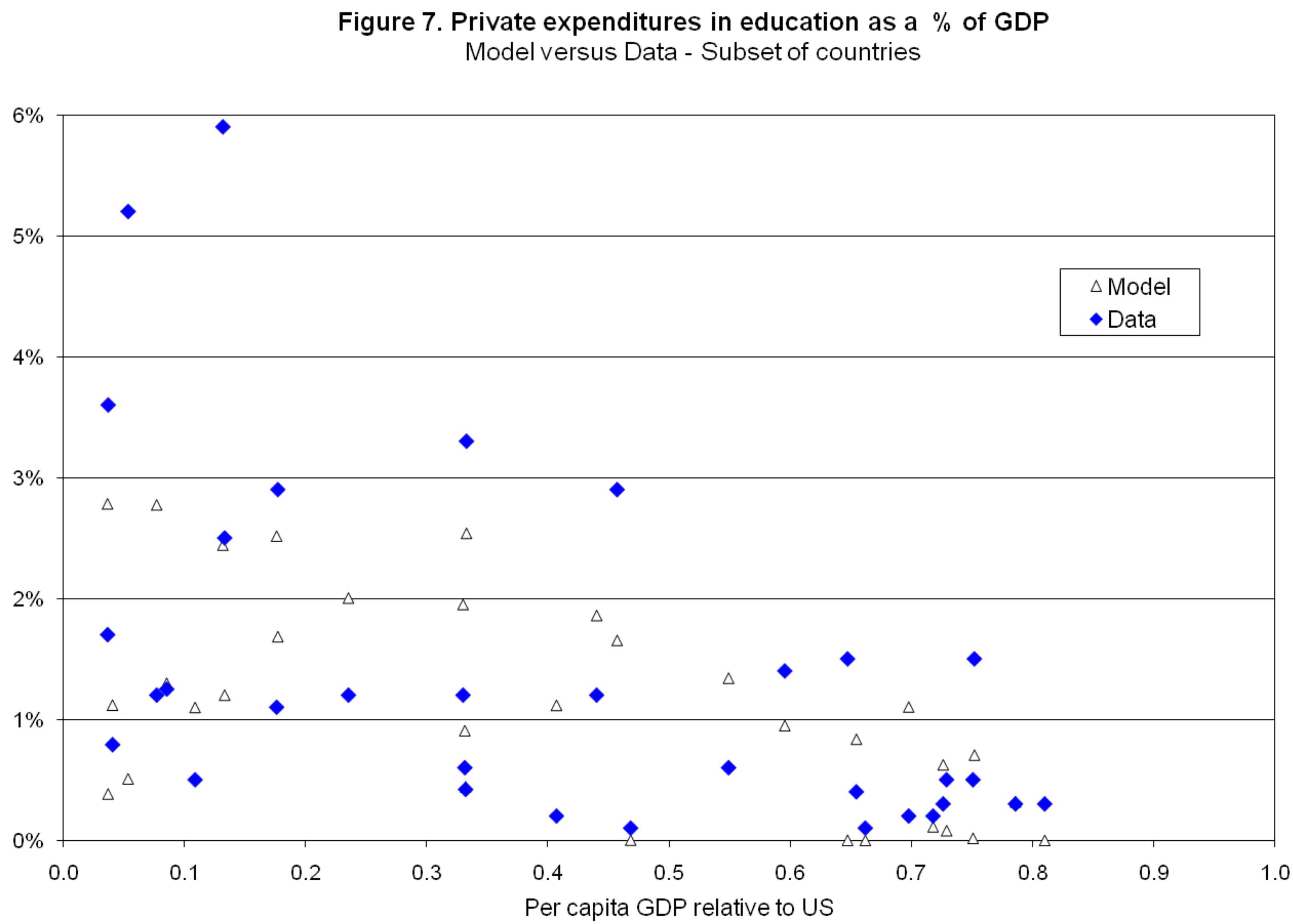


Figure 7. Private expenditures in education as a % of GDP
Model versus Data - Subset of countries
6%
5%
4%
3%
2%
1%
0%
Model
Data
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Per capita GDP relative to US

Figure 8. Quality of human capital

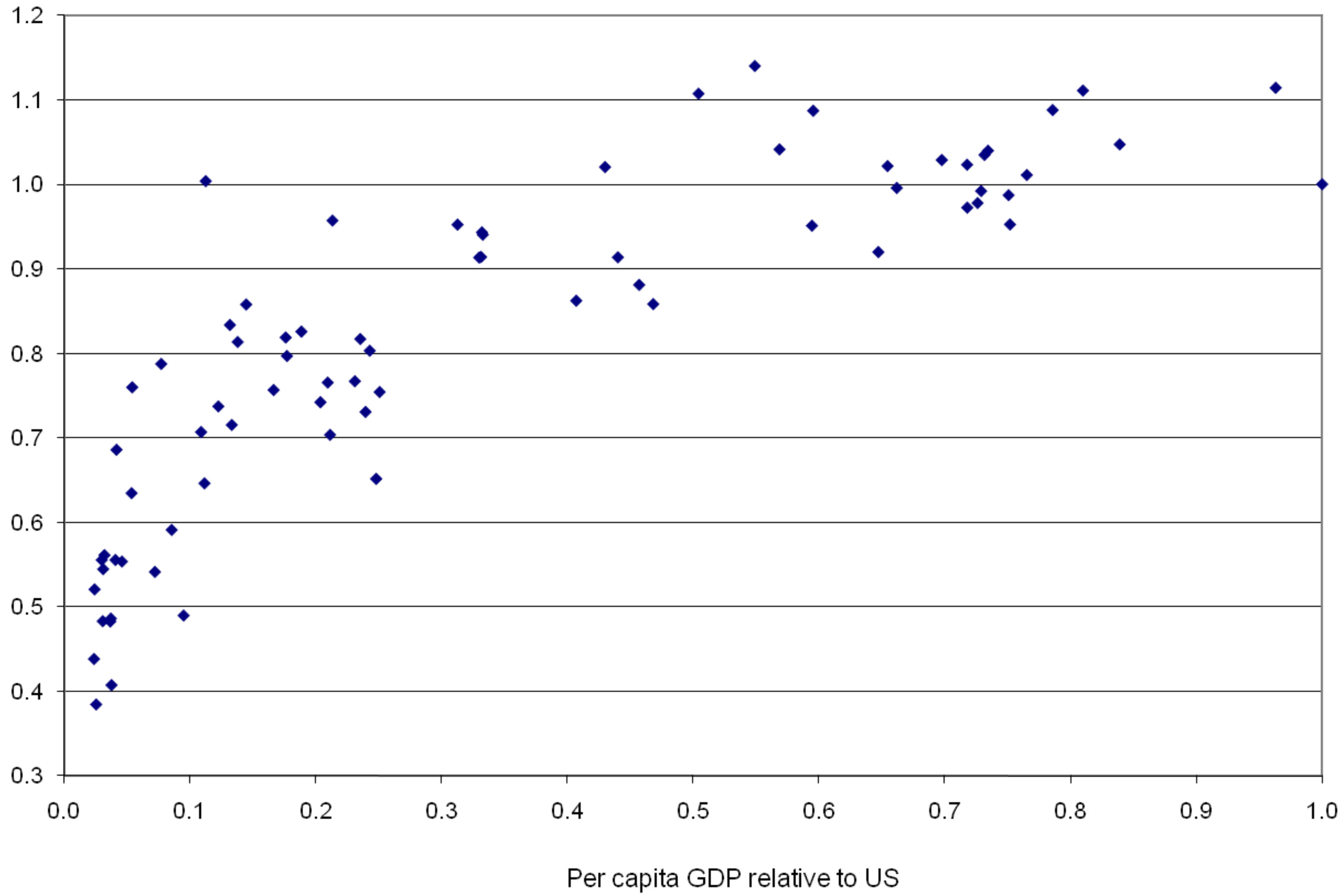


Figure 9. Schooling: benchmark and counterfactual

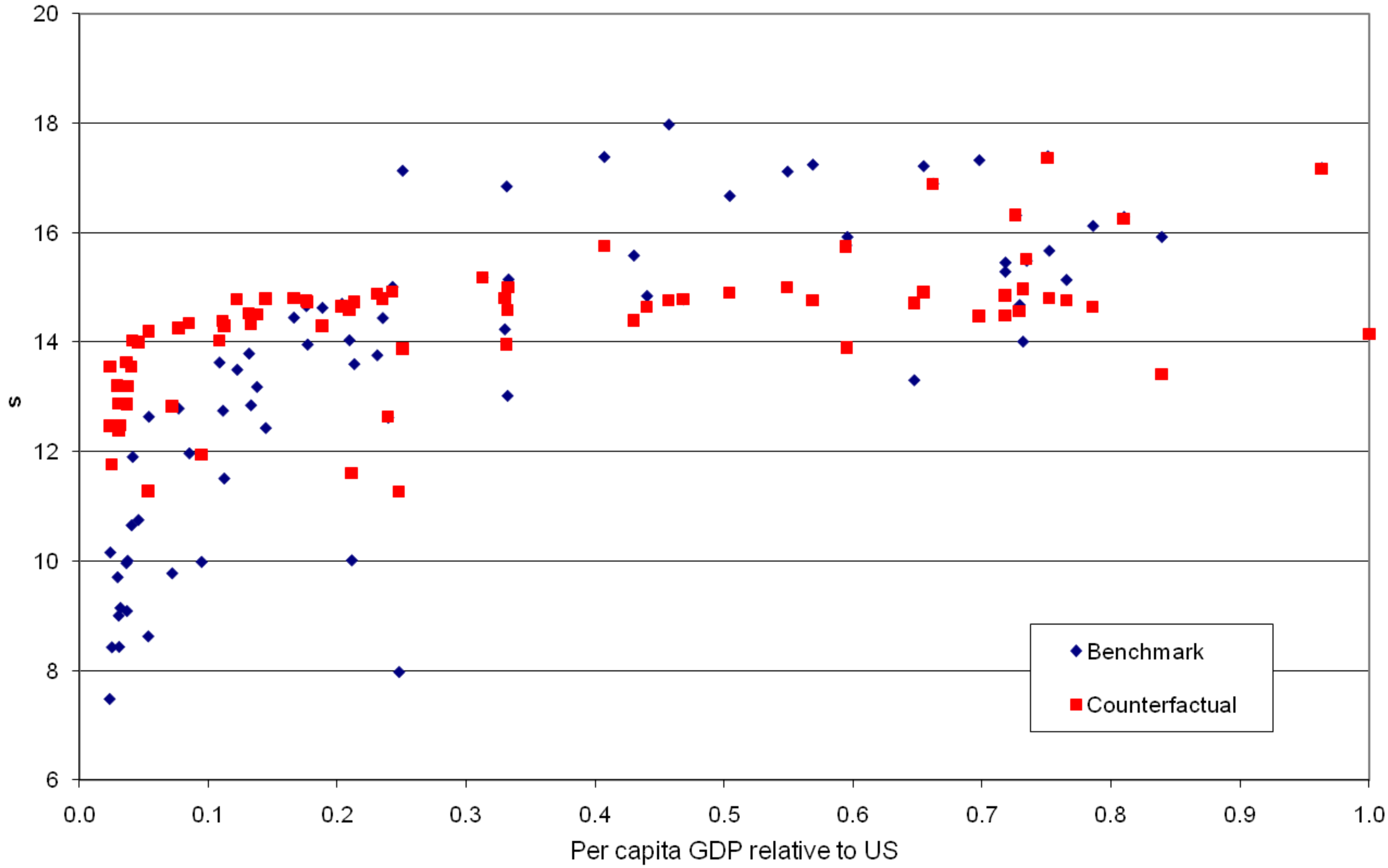


Figure 10. Parental transfers: benchmark and counterfactual