Schooling and Development: The Role of Credit Limits, Public Education, Fertility and Mortality

Juan Carlos Córdoba*

and

Marla Ripoll†

September, 2009

Abstract

This paper provides a theory of schooling that outperforms existing alternatives in explaining the cross-country distribution of average schooling, as well as the so called human capital premium puzzle. In our theory credit limits, 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, income differences, life expectancy, public education spending, life cycle model
JEL Classification: I22, J24, O11.

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

†Department of Economics, University of Pittsburgh, 4907 W.W. Posvar Hall, Pittsburgh PA 15260, e-mail: `ripoll@pitt.edu`. An early version of this paper circulated with the title "The Role of Education in Development."

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 Australia would attend for 20.4 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 (2000), 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 key issues emerge when analyzing Figure 1. First, the model substantially overpredicts schooling in fast growing countries (South Korea, Hong Kong, China, and Macao), while it underpredicts schooling in most high-income countries whose growth rates are modest (R^2 is 0.27). The reason for this lack of predictability 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).

A second issue with BK's model is that it requires implausibly high rates of riskless returns, between 9.5 to 10.3%, in order to match average schooling in their sample, an issue known in the literature as the "human capital premium puzzle" (Palacios-Huerta 2003, Elias 2003, Kaboski 2007). A potential problem in using a high rate of discount is that it may incorrectly downplay key determinants of schooling. For example, life span may be an important determinant of schooling since a typical individual spends around a third of his life in school, including pre-school, and the correlation between life expectancy and schooling in the data is 0.8. However, differences in life expectancy play almost no role in BK's model merely due to the fact that individuals discount the future heavily.

This paper provides a quantitative theory of schooling that outperforms existing alternatives

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

along the quantity (years of schooling) and price (returns to schooling and assets) dimensions. We consider a life-cycle economy populated by altruistic individuals who face borrowing constraints during their schooling years, have a stochastic life span and realistic fertility rates, and have access to a public education system. Our modeling of the public 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 additional years of schooling (the extensive margin).

The key feature of the model that explains its better fit is the borrowing constraint for students. A binding borrowing constraint brings about new explanatory variables for schooling choices that are irrelevant in frictionless models. Specifically, optimal schooling in frictionless models is obtained from a simple income maximization problem in which variables such as parental bequests, fertility, and public education play only a minimal role, if any. In our model, in contrast, these three variables acquire central importance. Furthermore, borrowing constraints also create a wedge between asset returns and human capital returns providing an explanation for the human capital premium puzzle. More importantly, our model does not require unrealistically high interest rates and, as a result, it is better suited to assess the role of mortality and life expectancy in explaining schooling differences. Such role turns out to be first order.

We calibrate the model 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 four main findings. We find that fertility rates are the single most important determinant of schooling differences across countries. In a counterfactual exercise in which fertility rates are equalized across the board to US levels, the dispersion of schooling falls by an impressive 53%. This is because, given the limited access to credit, individuals in the model rely heavily on parental transfers to finance their consumption and other expenditures during schooling years. But

a larger 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 would transfer less resources to their descendants. The model thus displays a clear quantity-quality trade off: children in countries with high fertility rates can afford less years of schooling. 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 36% 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 hard to replicate by frictionless models which typically find a small effect of mortality on schooling (Bils and Klenow, 2000; Manuelli and Seshadri, 2007; Hazan, 2009). 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 into 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 wages to the US level reduces the dispersion of schooling by only 4%. 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 opposite forces. On the one hand, individuals expecting higher wages during their working years optimally reduce schooling years as a way to increase consumption during schooling years given that 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. We find that equating years of education offered by the public system in all countries to that offered in the US would reduce the dispersion of schooling by 36.8%. 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 early out of public school. 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 (for example, 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; Boucekkine, de la Croix and Licandro, 2002; Hendricks, 2008; Ferreira and Pessoa, 2005; Schoellman, 2007), 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. As mentioned above, our model is similar in spirit to BK's model, but it incorporates credit market frictions and a role for public education. In their model, countries differ in schooling levels due to differences in growth rates: a higher growth rate of wages increases the incentives to accumulate schooling and human capital. The relative flat pattern of schooling across countries predicted by their model (Figure 1) results from the fact that growth rates are mostly uncorrelated with income levels. In our model growth rates are identical across countries and schooling differences are explained by differences in demographic factors (fertility and mortality) and in access to public education.

Manuelli and Seshadri (2007) also study a frictionless economy in which income levels, not income growth, explain schooling differences. As BK's, their model also requires a high riskless return (of 8%), and strong income effects on schooling which may be counterfactual, as argued by BK. Erosa, Koreshkova and Restuccia (2007) also proposed a model with strong income effects. In contrast, in our model, income effects on schooling are close to zero.

The remainder of the paper is organized as follows. Section 2 lays down the model and characterizes the balanced growth path. The calibration of the model is presented in Section 3, and results are discussed in Section 4. Section 5 summarizes development account results and Section 6 concludes.

2 The 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. Individuals receive a transfer 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. Moreover, individuals are credit constrained during

schooling years. For simplicity we call *bequests* the transfers at birth from parents to children.

Finally, 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, $q(a)$. There are annuity markets and prices are actuarially fair.

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 $\gamma \in (0, 1]$.[2] 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, $e_s(a)$.

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 $h(s) = (e/p_E)^\gamma s^{\gamma/\beta}$ so that γ is the elasticity of expenditures and γ/β is the elasticity of years of schooling. Notice that $h(s)$ satisfies Inada conditions so that optimal investments in education are positive for any $a \in [0, s]$. Thus, for example, it is suboptimal to leave school and return at a later time.

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, $d \ln (wh(s)) / ds$. Using (1), one finds that:

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

Therefore, returns to schooling diminish with the amount of human capital ($\partial r_s(s) / \partial h(s) < 0$) and increase with the amount of expenditures at age s ($\partial r_s(s) / \partial e(s) > 0$). For the case $e(a) = e$, $r_s(s)$ 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.

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

Finally, 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.

2.2 Individual's problem

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

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

subject to:

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

$$\int_s^T c(a)\, q(a)\, da + q(F)\, f b' \leq \int_s^R w h(s) e^{\nu(a-s)} q(a) da + q(s)\omega(s)\,; \tag{4}$$

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

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

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, $\pi(a)$ is the probability of surviving up to age a, and $\phi(f)$ is an altruism function that weights the utility of children in the utility of the parents. The momentary utility and altruism functions are assumed to have the forms:

$$u(c) \equiv \left(c^{1-\sigma} - 1\right)/(1-\sigma) \text{ and } \phi(f) \equiv \phi f^{\psi}.$$

The first restriction of the problem, equation (3), is the budget constraint during schooling years in present value. During this period, individual's only resources are parental bequests which can be used to consume, invest in education and save. Individuals cannot borrow since $\omega(s)$ is restricted

to be non-negative. 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 bequest to their offsprings, b'. Notice that individuals are not credit constrained when they become workers and parents since $s \leq F$.[3] Parents can thus borrow for the purpose of providing optimal bequests to their descendants. Moreover, 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, it is convenient to 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 perfectly 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 want 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, suppose $\phi(f) = 1$ and $f = 1$ 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.

2.4 Optimal allocations

We now describe the most relevant properties of the optimal allocation and leave the details for Appendix B. 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 constraints is binding. Denote $c^S(s)$ and $c^W(s)$ the optimal 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 $c^W(s)$.

[3] This last restriction is not binding for any country in the calibration below.

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:

$$u'\left(c\left(F\right)\right) = \frac{\phi(f)}{f} u'\left(c\left(0\right)\right), \tag{5}$$

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:

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

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, (5) can be written as:

$$\frac{u'(c^S(s))}{u'(c^W(s))} = G \equiv \frac{f}{\phi(f)} e^{-(r-\rho)F} > 1. \tag{6}$$

The last inequality follows from Assumption 1 and implies that the borrowing constraint is binding.[4] Thus, even altruistic parents do not leave enough bequests so that their children can fully smooth their consumption. The reason is that the relative low interest rate (relative to the rate of time preference) makes optimal to consume early in life, while bequests are a way to postpone consumption (via children's consumption). Notice that the ratio of marginal utilities at time s is completely determined by equation (6) and it only depends on r, ρ, f and F but not on wages, educational subsidies, or any other level variable.

Let E^* be the present value of total educational expenditures. Appendix B shows that, in steady state, bequests are given by:

$$b = \frac{G^{-\frac{1}{\sigma}} W(s) + E^* \Omega}{\Omega + q\left(F\right) f G^{-\frac{1}{\sigma}}} \text{ where } \Omega \equiv \frac{\int_s^T e^{-(\rho-r)a/\sigma} q\left(a\right) da}{\int_0^s e^{-(\rho-r)a/\sigma} q\left(a\right) da}. \tag{7}$$

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

2.4.2 Optimal education spending

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

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

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.

Appendix B shows that, for $a \in [0, s]$, $e^*(a)$ has the form:

$$e^*(a) = \max\{\widehat{e}^*(a), e_p(a)\}. \tag{8}$$

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).[5] 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.

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 $[\underline{s}, \overline{s}]$.

[5] In the calibration below the interval $[\underline{s}, \overline{s}]$ would be different for each country. Since many governments typically finance primary and secondary, $\underline{s}$ to $\overline{s}$ roughly corresponds to 6 and 18 years respectively.

2.4.3 Optimal schooling choice

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

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

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:

$$r_s(s) = \nu + \frac{wh(s)q(s) + \sigma \frac{\Delta u(s)}{u'(c^S(s))} e^{-\rho s} \pi(s) + G e_s(s) q(s)}{W(s)}, \tag{9}$$

which provides a link between schooling choices and returns to schooling, and as such is an important equation in calibrating the model. Equation (9) can be used together with (2) to solve for s.

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. We assume that some parameters are common across countries while other are country specific.

3.1 Parameters common across countries

Table 1 shows the parameters assumed to be common across countries and their calibrated values. We set σ to a standard value of 1.5. 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 (2000) and Murphy and Welch (1990) who find this factor to be 2.5 and 2.2 respectively.

Table 1. Parameters Common Across Countries

Parameter	Value	Description
σ	1.5	CRRA - Cooley and Prescott (1995)
ν	2%	returns to experience - Bils and Klenow (2000)
$\underline{s}$	6	starting schooling age
F	25	parenthood age
R	65	retirement age
ϕ	1	perfect altruism when $f = 1$
ψ	0.4	altruism - Birchenall and Soares (2009)
r	0.6%	US riskless rate - Mehra (2003)
ρ	4.0%	Average years of schooling (OECD)
γ	0.265	Private education spending % GDP (OECD)
β	0.18	Returns to schooling
α	0.33	capital share - Gollin (2001)

Starting school age, $\underline{s}$, is set to 6 years, a value that represents quite well most countries in the data. The age of parenthood, F, is set to 25, an age at which the average student in all countries has finished school.[6] Retirement age R is set to be 65, a value that is relevant mostly for rich countries in the sample and allows for a more realistic working life span.

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 see below, $f = 1$ closely characterizes many rich countries including the U.S. Moreover, ψ is set to 0.4, a middle-range value according to Birchenall and Soares (2009) who calibrate this parameter using micro evidence on

[6] The results of the paper are not sensitive to different plausible values of F.

the value of children's life.

The riskless interest rate is set to the U.S. historical experience. According to Mehra (2003), the riskless returns was on average 0.6% between 1947 and 2000. This is a big departure from Bils and Klenow (2000) who use a value of r much larger, in the order of 9.3 to 10.5% (pp. 1176).

Parameters γ, β, ρ are calibrated jointly to match three targets for OECD countries in the sample. The targets are: expected years of schooling, returns to schooling, and private education expenditures as a percentage of GDP (for 2003). We choose targets for OECD countries rather than only to the US, as is standard practice, 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. The calibrated value of γ is particularly sensitive to this target but is not critical for the main results.

Our measure of expected years of schooling is *school life expectancy* (SLE), as reported by UNESCO, plus 6 years (to include pre-school). 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 enrolment ratio for that age."[7] 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.[8] We choose SLE as our measure of years of schooling because it corresponds more closely to our theoretical construction of steady state schooling.[9] Average SLE in OECD countries is 16.14, and therefore we calibrate the model to predict 22.14 of total schooling for those countries.

Regarding returns to schooling, we use the methodology and intermediate parameter values used by BK. Specifically, returns to schooling are computed as $0.18 \times (SLE)^{0.28}$. The corresponding returns to schooling given the average SLE for OECD countries is 8.28%. These returns incorporate,

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

[8] 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.

[9] 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.

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 1.5%. Therefore, we match returns to schooling of 6.78% for OECD countries. These three targets result in the following calibrated values: $\rho = 4\%$, $\gamma = 0.265$ and $\beta = 0.18$.

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).

3.2 Country-specific parameters

Countries differ in: schooling-related variables e_p, $\overline{s}$ and grade repetition probabilities; demographic variables $\pi(a)$ and f; and prices p_E and w. We now consider these groups 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 $\overline{s}$, the 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 the government at different schooling levels. In particular, $\overline{s}$ is computed for each country as

$$\begin{aligned} \overline{s} &= \text{duration prim\&sec} \times \frac{\text{pub exp prim\&sec}}{\text{total exp prim\&sec}} \\ &\quad + (SLE - \text{duration prim\&sec}) \times \frac{\text{pub exp terciary}}{\text{total exp terciary}} \end{aligned}$$

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 $\overline{s}$ as $SLE \times$ (pub exp/total exp). Second, for countries with no available data on the public share of expenditures at all, we use the duration of compulsory schooling, from UNESCO, as a measure of $\overline{s}$. Notice that by using SLE in measuring $\overline{s}$ we want to capture the number of years a "representative child" in each country receives public education subsidies, which

corresponds to the definition of $\overline{s}$ in the model.

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

$$T = -\frac{\log(\xi)}{p} + a_s. \tag{10}$$

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 (10).

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.

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

$$w = \frac{(1-\tau)(1-\alpha)y_t^{data}}{h_t}, \tag{11}$$

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.

A key issue 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.[10]

With these limitations in mind, we decide to compute h_t in a way that is roughly consistent

[10] Most of the data on expenditures start in the 90's.

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 educational investments are roughly constant. The definition above implies that if $s_t = s$ then $h_t(s) = h(s)$. We choose $t = 2000$ because this is the last year for which Barro and Lee (2000) provide information about average years of schooling, s_t, among adult population.

We then define average human capital at time t as:

$$h_t = \Theta_t h_t(s_t) \tag{12}$$

where Θ_t is an adjustment for the average experience of the working force at time t.[11] Plugging h_t into (11), 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 B describes the solution algorithm.

4 Results

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

4.1 Model's fitness

Figure 3 portrays the schooling predicted by the model and calibrated values of $\overline{s}$ for each country against schooling in the data. Recall that we measure schooling in the data as SLE plus 6 years. If the predicted s perfectly matched the data, all observations would lie along the 45-degree line. The calibrated model predicts an average across countries of 19.57 years of schooling (6 of which are pre-school), slightly higher than the data (18.96). Moreover, the model explains 86% of the standard deviation of schooling, which is 3.35 years in the data, and 2.87 in the model. Since we calibrated the model to target only average schooling among OECD countries, these results are remarkable and show that the model is quite successful in explaining the average and dispersion of

[11] 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 $R = 65$.

schooling across countries.

Figure 3 also portrays our measure of $\overline{s}$ versus SLE in the data. As such, the vertical distance between the 45-degree line and $\overline{s}$ in each country corresponds to the number of schooling years fully financed with private resources, in addition to preschool. The graph clearly illustrates a strong positive correlation between $\overline{s}$ and SLE, which in our sample amounts to 80%.

Figure 4 shows returns to schooling predicted by the model and estimates based on BK. Our model predicts returns to schooling than are on average one percentage point below those implied by BK (7.9% versus 8.9%). This lower average is expected given that, by construction, our calibration matches one and a half percentage points less than the average returns to schooling in OECD countries, as explained above. Figure 4 also shows that the model does well in predicting the dispersion of returns to schooling. The standard deviation in the model is 1.3% versus 0.8% in the estimates.

Figure 5 portrays private education expenditures as a percentage of GDP in the data and in the model for a subsample of 55 countries for which there is available data. Even though we only targeted the mean of OECD countries, the model predicts almost exactly the mean and the standard deviation of this variable (1.2% and 1.1%) for the subsample of 55 countries.

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 $\widehat{h}(s) \approx s^{\beta/\gamma}$.[12] Figure 6 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 as much as a factor of 0.4 for almost all 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% (a factor of 1.1). 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

[12] This formula provides similar estimates as those of BK.

lowest decile is approximately one fifth of that of the U.S. Our equivalent measure is about two fifths. More importantly, while Manuelli and Seshadri's estimate is driven mainly by cross-country differences in wages and mortality, 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.

As documented by Hanushek and Woessmann (2007), although in the data there is no systematic relationship between resources spent and quality of education measured by student achievement, 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." (pp. 67). We think our cross-country comparison of quality in Figure 6 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 Extent of credit tightness

To develop some further intuition on the mechanisms at work in our model, we now comment on two other predictions. Figure 7 portrays $c^W(s)/c^S(s)$, the ratio of consumption of a worker relative to that of a student at age s. This ratio captures the extent to which individuals are credit constrained. In the absence of borrowing constraints, this ratio would be equal to one, as consumption would be perfectly smoothed across periods. As shown in the figure, poorer countries are substantially more credit constrained than richer countries.

To understand why this is the case, notice that the jump of consumption at age s is determined by equation (6). According to this equation, $u'\left(c^S(s)\right)/u'\left(c^W(s)\right)$ is a decreasing function of r and an increasing function of f. Since r is assumed to be identical across countries, the extent of the credit constraints is determined by f, the number of children. Individuals in countries with higher fertility rates face tighter credit constraints, or are less able to smooth consumption. The reason is that parental transfers act as a substitute for credit markets during schooling years but those transfer are more diluted the larger the number of children. This insight is confirmed by Figure

8 which shows how parental transfers, b, received by each child are much larger in rich countries where fertility rates are systematically lower.

4.4 Counterfactuals

We now assess the relative quantitative importance of the exogenous parameters in explaining schooling differences across countries according to our theory. 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 2 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.[13]

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 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 53%.

Figure 9 portrays schooling profiles for the counterfactual of equal fertility and for the benchmark. 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 2 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 remain in school for a longer period. In fact, the variance of (log) parental transfers in the world falls by around 62% in this scenario.

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 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, pp. 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 around 25%.

Key for understanding the role of mortality and life expectancy in schooling is the rate of

[13] 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.

discount. A high interest rate, as the one used by BK, 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. In a frictionless model altruism plays no role in schooling decisions. However, in the presence of borrowing constraints a reduction in mortality increases the weight that children receive in their parents' utility and therefore increases bequests. Larger bequests allow descendants to stay longer in school.

Table 2. Counterfactual exercises (% change relative to benchmark)

Parameter	$stdev(s)$	$mean(s)$	$var(ln(b))$	$mean(b)$
p_c	-5.1	0.6	-7.4	0.5
p_s	-3.9	0.4	-3.3	0.2
p	-25.3	2.8	-20.7	0.9
f	-53.1	3.5	-61.7	-7.2
e_p	17.7	-6.8	-15.4	0.5
$\overline{s}$	-36.8	2.1	-14.5	-1.9
p_E	1.9	-0.5	14.3	-11.5
w	-4.1	1.0	-51.8	70.3

Statistics are computed for the sample of 74 countries.

Consider now the effect of wages on schooling. Our model predicts that schooling is mostly unresponsive to permanent changes in wages: equating wages to US levels, which would be a drastic change for many poor countries, only reduces schooling dispersion by around 4%! This small response is consistent with the empirical findings summarized by Haveman and Wolfe (1995) who document mild income effects of schooling, and it is in clear contrast with the findings of Manuelli and Seshadri (2007) and Erosa, Koreskova and Restuccia (2007) who stress strong income effects in explaining schooling differences.

The small response of schooling to wages 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 3, average bequests increase by 70% 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 36.8%. The mechanism here is that individuals can only 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 increases 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 17.7%.

Finally, we assess the effects of relaxing borrowing constraint by allowing individuals to borrow each year up to 1% of their benchmark life-time earnings. The results of this exercise are not directly comparable to others in Table 2 since parameters are not being equalized across countries. We find that by relaxing the borrowing constraint, the standard deviation of schooling decreases

by 63.6% , but its mean *decreases* as well by 31.4%. In other words, schooling years become more similar across countries but at a *lower* level than the benchmark everywhere.

Parental transfers are the key to understand why more generous borrowing limits decrease steady state schooling. Individuals who can borrow more to consume and invest in education during schooling years would also have less resources available later in their life, particularly at the time of making transfers to their descendants. In fact, in the limit, if borrowing constraints were relaxed incrementally parents eventually would like to leave negative bequests to their children. In turn, children receiving smaller transfers from their parents would optimally choose less schooling. In fact, the steady state average value of transfers drops by 70% in the previous experiment. In other words, as long as credit is relatively inexpensive, namely as long as Assumption 1 is satisfied, individuals would prefer to spend more in the present than in the future. Bequests fall because they are a mechanism for parents to consume in the future through their children. A similar result is obtained by Soares (2008) in a different context.

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

$$y_t = k_t^{\alpha}(Ah_t)^{1-\alpha}, \tag{13}$$

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:

$$\widetilde{w} = (1-\alpha)\frac{y}{h}, \text{ and } \widetilde{r} = \alpha\frac{y}{k}. \tag{14}$$

Assume also 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 (14) imply a capital-output ratio k/y in each country. Finally, we compute A from equation (13) 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 (12).

Finally, steady state income y is computed as $y = (k/y)^{\frac{\alpha}{1-\alpha}} Ah$, which is another way to write (13). 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 3 summarizes these counterfactuals (for the purpose of comparison, here we also report the changes in standard deviation and mean schooling from Table 2).

Table 3. Income under counterfactuals (% change)

Parameter	$var(\ln(y))$	$mean(y)$	$stdev(s)$	$mean(s)$
p_c	3.2	-0.3	-5.1	0.6
p_s	-1.3	0.1	-3.9	0.4
p	-13.1	1.0	-25.3	2.8
f	-50.3	2.7	-53.1	3.5
e_p	-27.7	7.9	17.7	-6.8
$\overline{s}$	-8.7	-1.1	-36.8	2.1
p_E	18.5	-11.2	1.9	-0.5
A	-63.8	70.6	-4.1	1.0
p_I	-34.4	30.5	-3.5	0.8

Statistics are computed for the sample of 74 countries.

There are a number of novel results in 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 13.1%. 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 50.3%. This result is impressive, and it is second only to changes in TFP levels A. Table 3 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 27.7%, while equating $\overline{s}$ only does by 8.7%. 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 TFP levels A are equated to US levels the variance of per capita income is reduced by 63.8%, while in sharp contrast, the standard deviation of schooling only falls just by 4.1%. The first result is in line with the growth literature, while the second is consistent with BK, who also found that income effects are small in explaining schooling choices. Finally, changes in the price of investment goods p_I have an important impact on per capita income dispersion, but a small one in schooling. We conclude that TFP levels and fertility rates are the two main determinants of per capita income dispersion. Quantitatively speaking, the surprising result is that fertility rates have large effects both on schooling and per capita income dispersion.

6 Concluding comments

Understanding why educational outcomes varies so much across countries remains a challenge. Existing models that rely on differences in the rates of economic growth or on strong income effects have counterfactual implications. This paper carefully models key aspects of schooling decisions for a typical agent in a credit constraint environment. Such a model improves upon existing models along different dimensions, namely prices and quantities, and provides new insights regarding the causes of schooling as well as income differences across countries.

Of major importance are demographic factors such as fertility and mortality. Our model suggests that holding controlling for demographics, income effects on schooling must be weak, a prediction that is consistent with the evidence summarized by Haveman and Wolfe (1995). Moreover, changes in fertility and mortality must have important effect on schooling. This channel could explain the experience of Sub-Saharan African countries during the last 40 years. Per capita income in these countries have 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 this improved schooling outcomes.

References

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

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

[3] **Ben-Porath, Yoram**. 1967. "The Production Function of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy*, 75(4, Part 1): 352-65.

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

[5] **Boucekkine, Raouf, David de la Croix and Omar Licandro**. 2002. "Vintage Human Capital, Demographic Trends, and Endogenous Growth." *Journal of Economic Theory*, 104(2): 340-375.

[6] **de la Croix, David, and Omar Licandro**. 1999. "Life Expectancy and Endogenous Growth." *Economic Letters*, 65 (2): 255-263

[7] **Elias, Julio J**. 2003. "The Education Premium." Unpublished.

[8] **Erosa, Andres, Tatiana Koreshkova, and Diego Restuccia**. 2007. "On the Aggregate and Distributional Implications of Productivity Differences across Countries." Unpublished.

[9] **Ferreirra, Pedro Cavalcanti and Samuel Pessoa**. 2003. "The Costs of Education, Longevity and the Poverty of Nations." Unpublished.

[10] **Glomm, Gerhard and B. Ravikumar**. 1998. "Flat-Rate Taxes, Government Spending on Education, and Growth." *Review of Economic Dynamics,* 1(1): 306-325.

[11] **Glomm, Gerhard and B. Ravikumar**. 2001. "Human Capital Accumulation and Endogenous Public Expenditures." *Canadian Journal of Economics*, 34(3): 807-826.

[12] **Gollin, Douglas**. 2002. "Getting Income Shares Right." *Journal of Political Economy*, 110(2): 458-474.

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

[14] **Hanushek, Eric, and Ludger Woessmann**. 2007. "The Role of School Improvement in Economic Development." Unpublished.

[15] **Haveman, Robert, and Barbara Wolfe**. 1995. "The Determinants of Children's Attainments: A Review of Methods and Findings." *Journal of Economic Literature*, 33(4): 1829-1878.

[16] **Hazan, Moshe**. 2009. "Longevity and Lifetime Labor Supply: Evidence and Implications" forthcoming *Econometrica*.

[17] **Hendricks, Lutz**. 2008. "Why Does Education Differ Across Countries?" Unpublished.

[18] **Kaboski, Joshep**. 2007. "Explaining Schooling Returns and Output Levels Across Countries." Unpublished.

[19] **Klenow, Peter J. and Andrés Rodríguez-Clare**. 1997. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" in Ben S. Bernanke and Julio J. Rotemberg, eds., *NBER Macroeconomics annual 1997*. Cambridge, MA: MIT Press, 73-103.

[20] **Mankiw, N. Gregory, David Romer and David N. Weil**. 1992. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics*, 107(2): 407-437.

[21] **Manuelli, Rodolfo, and Ananth Seshadri**. 2007. "Human Capital and the Wealth of Nations." Unpublished.

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

[23] **Palacios-Huerta, Ignacio**. 2006. "The Human Capital Premium Puzzle." Unpublished.

[24] **Rosen, Sherwin A**. 1976. "Theory of Lifetime Earnings." *Journal of Political Economy*, 84(4): S45-S57.

[25] **Schoellman, Todd**. 2007. "The Causes and Consequences of Cross-Country Differences in School Attainment." Unpublished.

[26] **Soares, Jorge**. 2008. "Borrowing Constraints, Parental Altruism and Welfare." Unpublished.

7 Appendix 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 exercise is reported in Figure A1. 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. The exercise is reported in Figure A2. Again, results are as in Figure 1 ($R^2 = 0.04$).

8 Appendix B: Full model solution

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

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

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

$$\begin{cases} e^{-\rho a}u'(c(a))\pi(a) = \lambda_1 q(a) \text{ for } a \leq s \\ e^{-\rho a}u'(c(a))\pi(a) = \lambda_2 q(a) \text{ for } a > s \end{cases} \tag{15}$$

$$-\lambda_1 q(a) + \lambda_3 \frac{\partial h(s)}{\partial e(a)} \frac{\partial e(a)}{\partial e_s(a)} + \lambda_4 = 0 \tag{16}$$

$$\begin{gathered} \lambda_3 \frac{\partial h(s)}{\partial s} - \Delta u(s) \cdot e^{-\rho s}\pi(s) - \lambda_1 \left[\frac{\partial q(s)}{\partial s}\omega(s) + \left(c^S(s) + e_s(s)\right) q(s)\right] \\ +\lambda_2 \left[\frac{\partial q(s)}{\partial s}\omega(s) + \left(c^W(s) - w(s)h(s)\theta(0)\right) q(s) + \int_s^R w(a)\frac{\partial \theta(a-s)}{\partial s} h(s) q(a) da\right] = 0 \end{gathered} \tag{17}$$

and

$$e^{-\rho F}\phi(f)V'(b')\pi(F) = \lambda_2 q(F) f \tag{18}$$

where

$$\frac{\partial h(s)}{\partial e(a)} = \gamma h(s)^{1-\frac{\beta}{\gamma}} e(a)^{\beta-1} p_E^{-\beta}, \tag{19}$$

$$\frac{\partial h(s)}{\partial s} = \frac{\gamma}{\beta} h(s)^{1-\frac{\beta}{\gamma}} \left(e(s)/p_E\right)^{\beta}, \tag{20}$$

and

$$\Delta u(s) \equiv u\left(c^W(s)\right) - u\left(c^S(s)\right) \tag{21}$$

Further, the following envelope condition holds:

$$V'(b) = \lambda_1. \tag{22}$$

The equations above imply that:

$$\frac{\lambda_1}{\lambda_2} = \frac{u'\left(c^S(s)\right)}{u'\left(c^W(s)\right)} \text{ and} \tag{23}$$

$$\frac{\lambda_3}{\lambda_2} = \int_s^R w(a)\theta(a-s) q(a) da = w(0)\int_s^R e^{(g-r)a}\theta(a-s)\pi(a) da, \tag{24}$$

where we have used actuarially fair prices $q(a) = e^{-ra}\pi(a)$. Moreover, combining (18) and (22) yields

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

and using (15) this equation can be written as:

$$\frac{u'\left(c^{child}(0)\right)}{u'\left(c(F)\right)} = \frac{f}{\phi(f)} \tag{25}$$

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

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 (19) and (16), we can write:

$$\widehat{e}^*(a) = e^*(0)\left(q(a)\right)^{-\frac{1}{1-\beta}}$$

where

$$e^*(0) = \left(\gamma h(s)^{1-\frac{\beta}{\gamma}} p_E^{-\beta} \frac{\lambda_3}{\lambda_1}\right)^{\frac{1}{1-\beta}}. \tag{26}$$

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:

$$e^*(a) = \begin{cases} \widehat{e}^*(a) \text{ for } a \leq \min(s, \underline{s}) \\ e_p(a) \text{ for } \min(s, \underline{s}) \leq a \leq s_p \\ \widehat{e}^*(a) \text{ for } s_p \leq s \end{cases} \tag{27}$$

where

$$s_p = \min\{s, \overline{s}, \max[\underline{s}, \widehat{a}]\} \tag{28}$$

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

$$e^*(\widehat{a}) = e_p(a). \tag{29}$$

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

$$h(s) = \left[\left(\frac{e^*(0)}{p_E}\right)^{\beta} \left(\int_0^{\min(s,\underline{s})} q(a)^{-\frac{\beta}{1-\beta}} da + \int_{s_p}^{s} q(a)^{-\frac{\beta}{1-\beta}} da\right) + \left(\frac{e_p}{p_E}\right)^{\beta} (s_p - \min(s, \underline{s}))\right]^{\gamma/\beta}. \tag{30}$$

8.1 Optimal schooling choice

Using equations (21), (23), and (24), together with the constraint $\underline{\omega} = 0$, we can write the optimality condition of the schooling choice (17) 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

$$r_s(s) = g + \nu + \left[w(s)q(s) + \sigma \frac{1}{h(s)} \frac{\Delta u(s)}{\lambda_2} e^{-\rho s} \pi(s) + \frac{\lambda_1}{\lambda_2} \frac{e_s(s) q(s)}{h(s)}\right] \frac{\lambda_2}{\lambda_3}, \tag{31}$$

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

8.2 Optimal consumption and transfers

Using the definitions of $u(c)$, $\pi(a)$ and $q(a)$ into (15):

$$c(a) = \left[\lambda(a) e^{(\rho - r)a}\right]^{-1/\sigma} \tag{32}$$

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:

$$\lambda_1^{-\frac{1}{\sigma}} = \frac{b - \int_0^s e_s(a) q(a) da}{\int_0^s e^{-(\rho-r)a/\sigma} q(a) da} = \frac{b - E^*}{\int_0^s e^{-(\rho-r)a/\sigma} q(a) da} \tag{33}$$

and

$$\lambda_2^{-\frac{1}{\sigma}} = \frac{h\left(s\right)\frac{\lambda_3}{\lambda_2} - q\left(F\right)fb'}{\int_s^T e^{-(\rho-r)a/\sigma} q\left(a\right)da} \tag{34}$$

where (24) has being used, and E^* is the present value of optimal private expenditures in education as given by:

$$E^* \equiv e^*\left(0\right)\left[\int_0^{\underline{s}} q(a)^{-\frac{\beta}{1-\beta}}da + \int_{s_p}^{s} q(a)^{-\frac{\beta}{1-\beta}}da\right] - e_p\int_{s_p}^{\min(s,\overline{s})} q(a)da. \tag{35}$$

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:

$$b = \frac{h(s)\left(\frac{\lambda_1}{\lambda_2}\right)^{-\frac{1}{\sigma}}\frac{\lambda_3}{\lambda_2} + e^*\Omega}{\Omega + q\left(F\right)f\left(\frac{\lambda_1}{\lambda_2}\right)^{-\frac{1}{\sigma}}}, \tag{36}$$

where

$$\Omega = \frac{\int_s^T e^{-(\rho-r)a/\sigma} q\left(a\right)da}{\int_0^s e^{-(\rho-r)a/\sigma} q\left(a\right)da}. \tag{37}$$

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

$$c^S\left(s\right) = \left[\lambda_1 e^{(\rho-r)a}\right]^{-1/\sigma}, \tag{38}$$

$$c^W\left(s\right) = \left[\lambda_2 e^{(\rho-r)a}\right]^{-1/\sigma}. \tag{39}$$

8.3 Solution algorithm

We solve the model as follows. For some initial values of $e(0)$ and s, we first use equation (30) to compute $h(s)$. Second, we compute h_t and w using equations (12) and (11). 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), (21), (24), (33), (34), (35), (36), (38), (39), together with:

$$\frac{\lambda_1}{\lambda_2} = \frac{u'\left(c\left(0\right)\right)q(F)}{e^{-\rho F}u'\left(c\left(F\right)\right)\pi\left(F\right)} = \frac{f}{\phi(f)}e^{-(r-\rho)F} \tag{40}$$

which is obtained using (15) and (25) in steady state and with actuarially fair prices. We iterate on the system of equations above by updating $e(0)$ and s using equations (26) and (31).

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

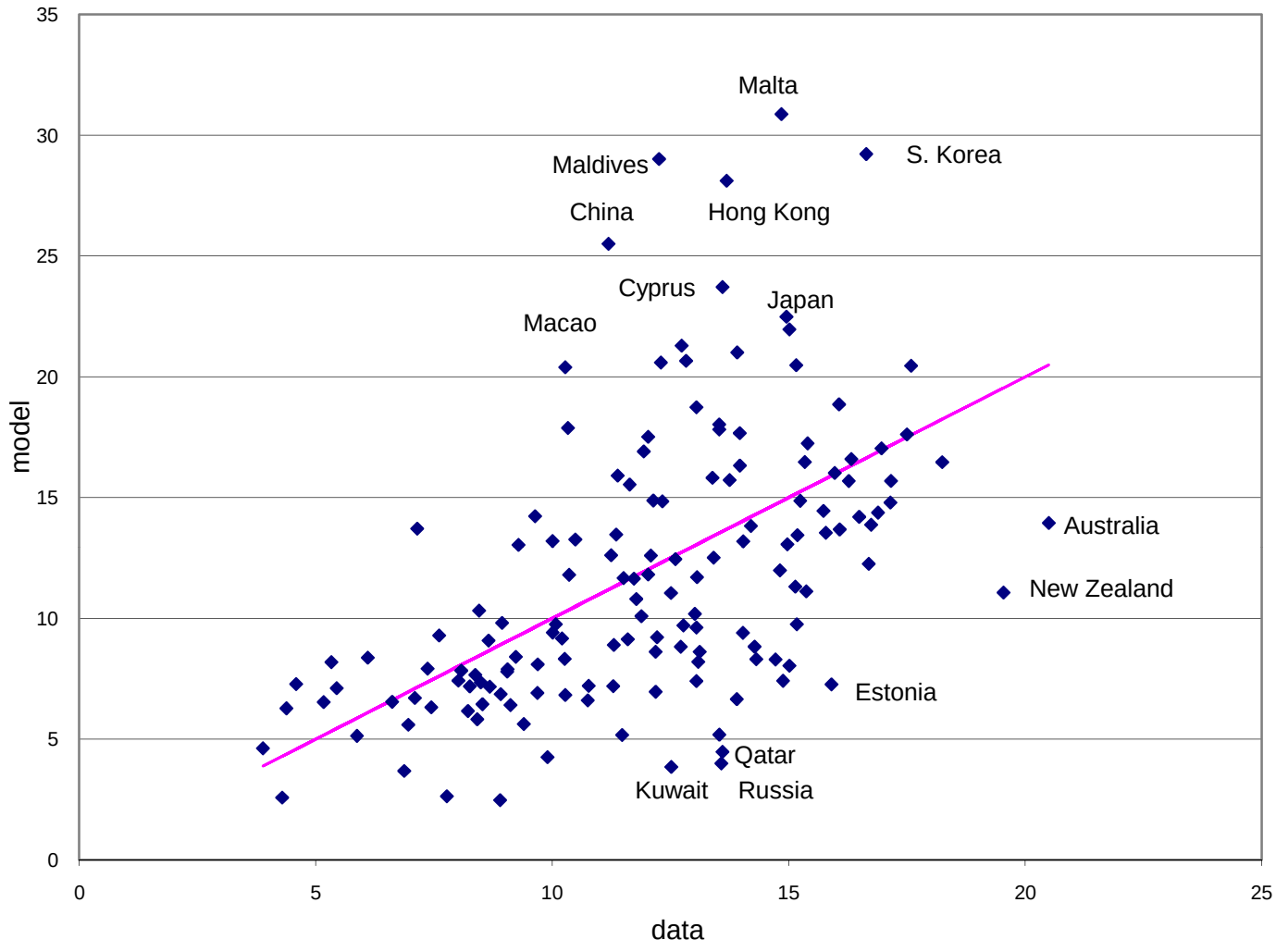
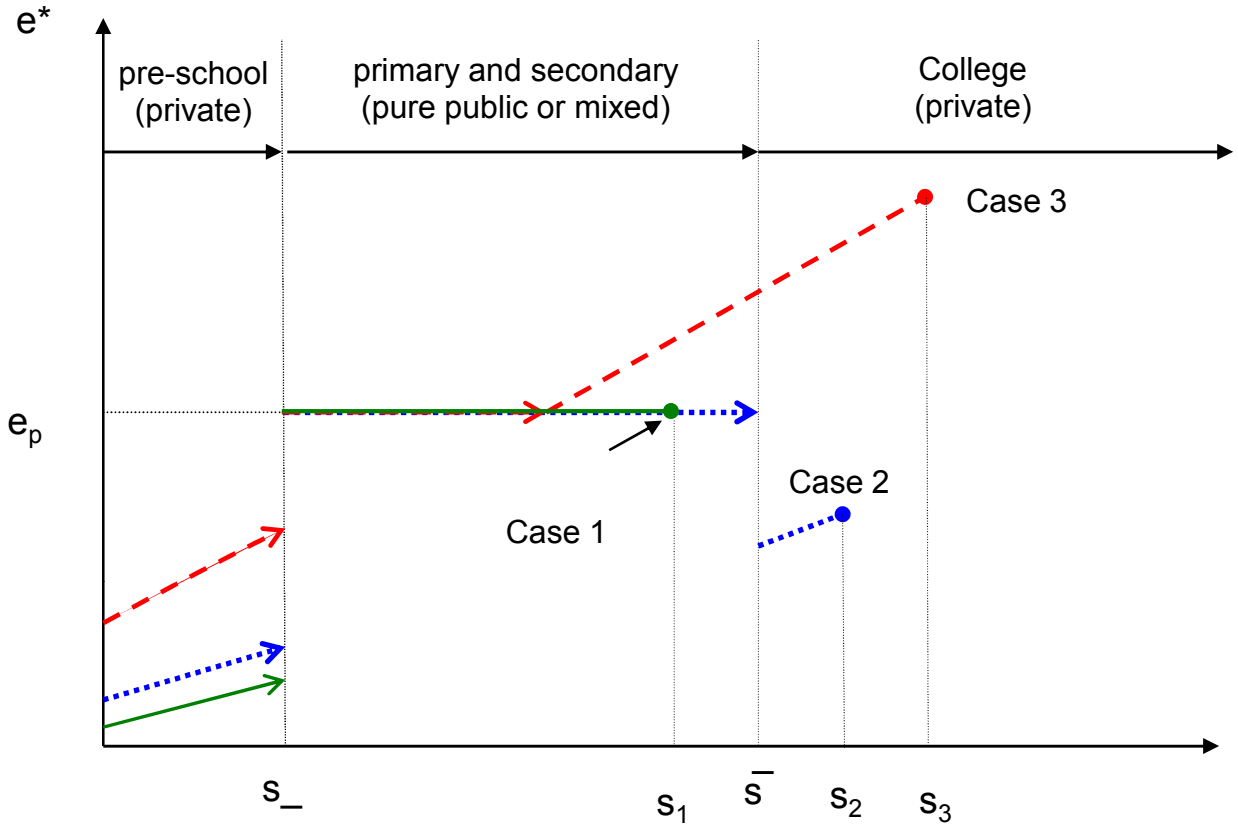


Figure 2. Individual expenditures in education: e*(a)



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

Figure 3. School life expectancy and maximum public schooling versus school life expectancy in the data

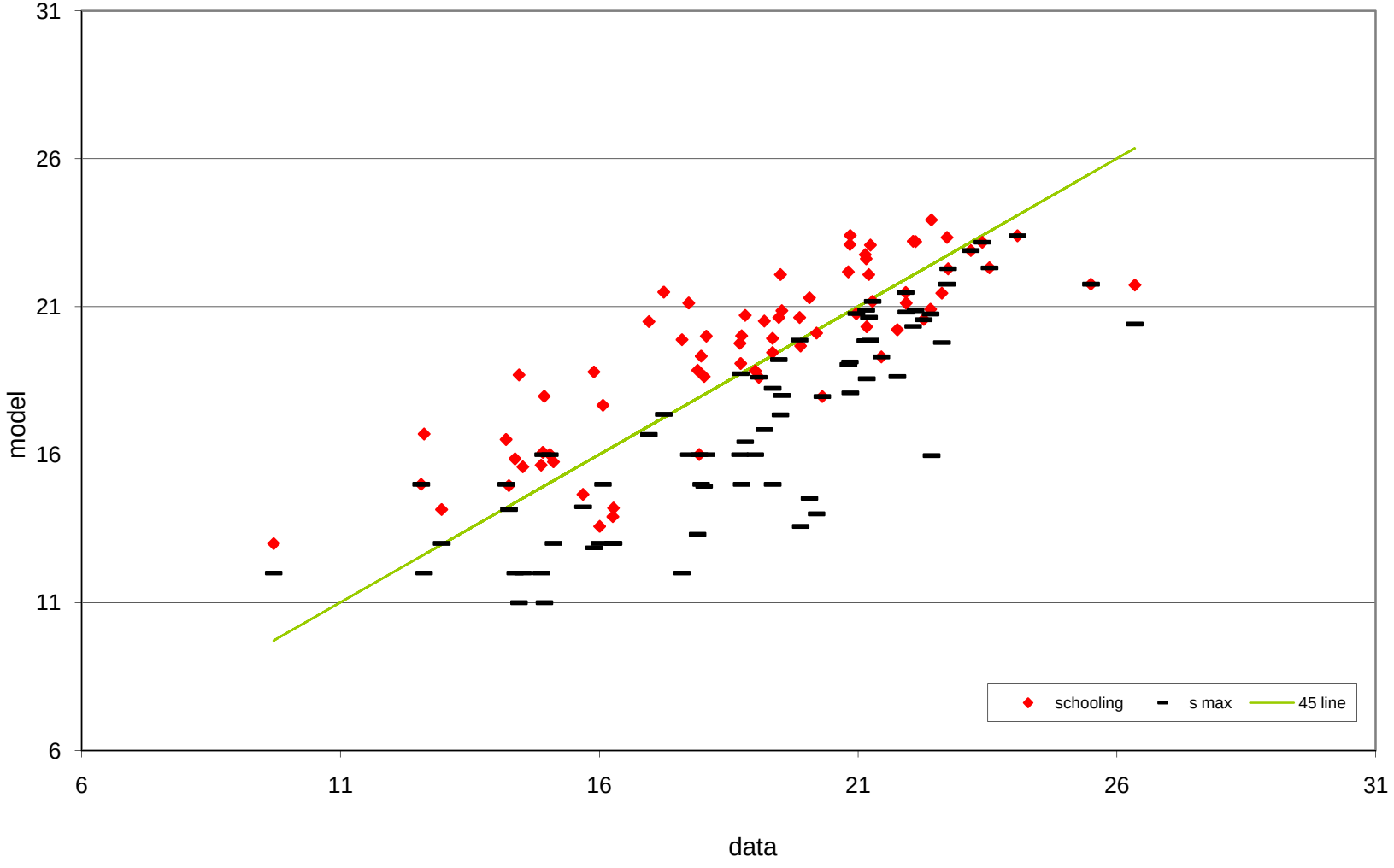


Figure 4. Returns to schooling
Model versus BK Estimates

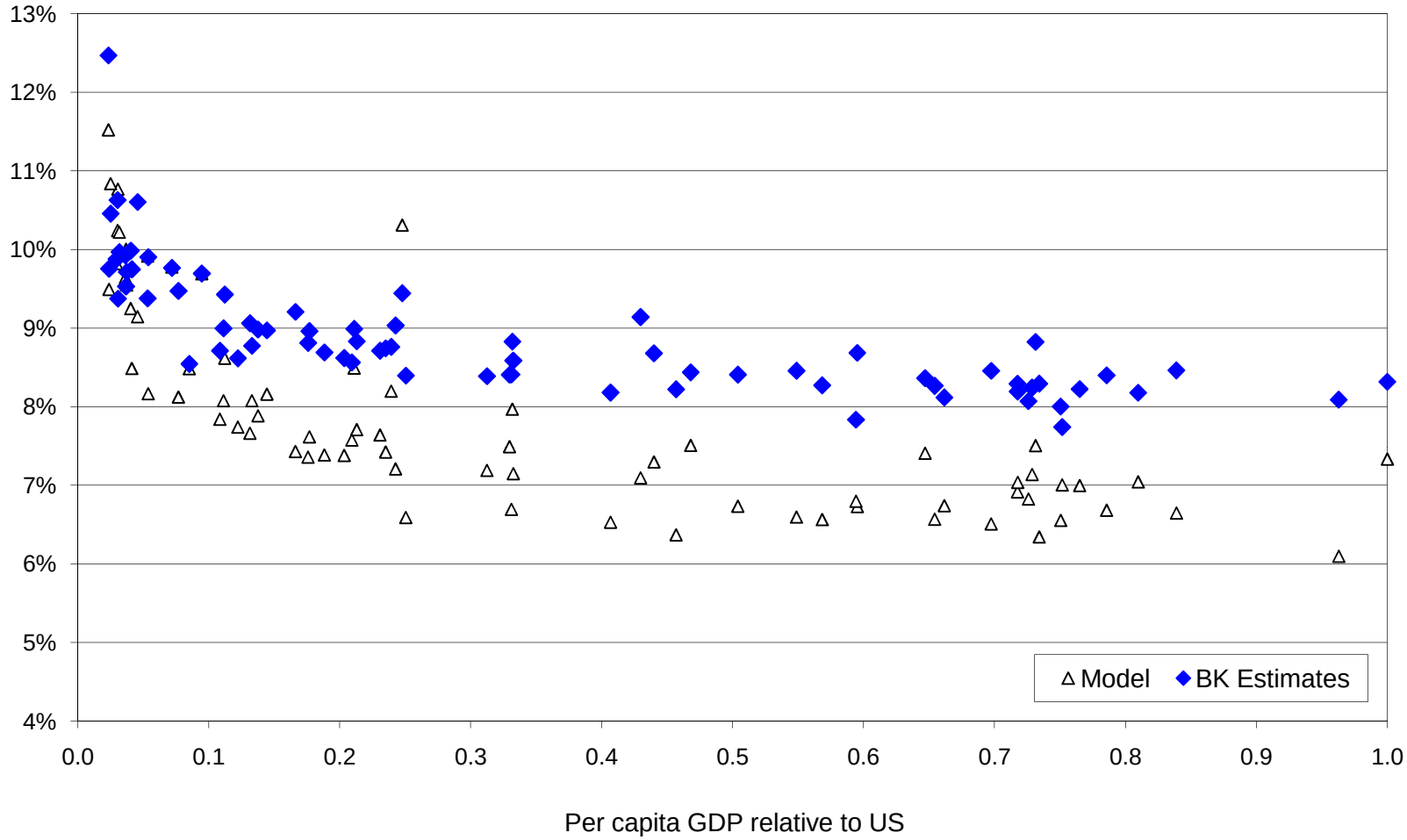


Figure 5. Private expenditures in education as a % of GDP
Model versus Data - Subset of countries

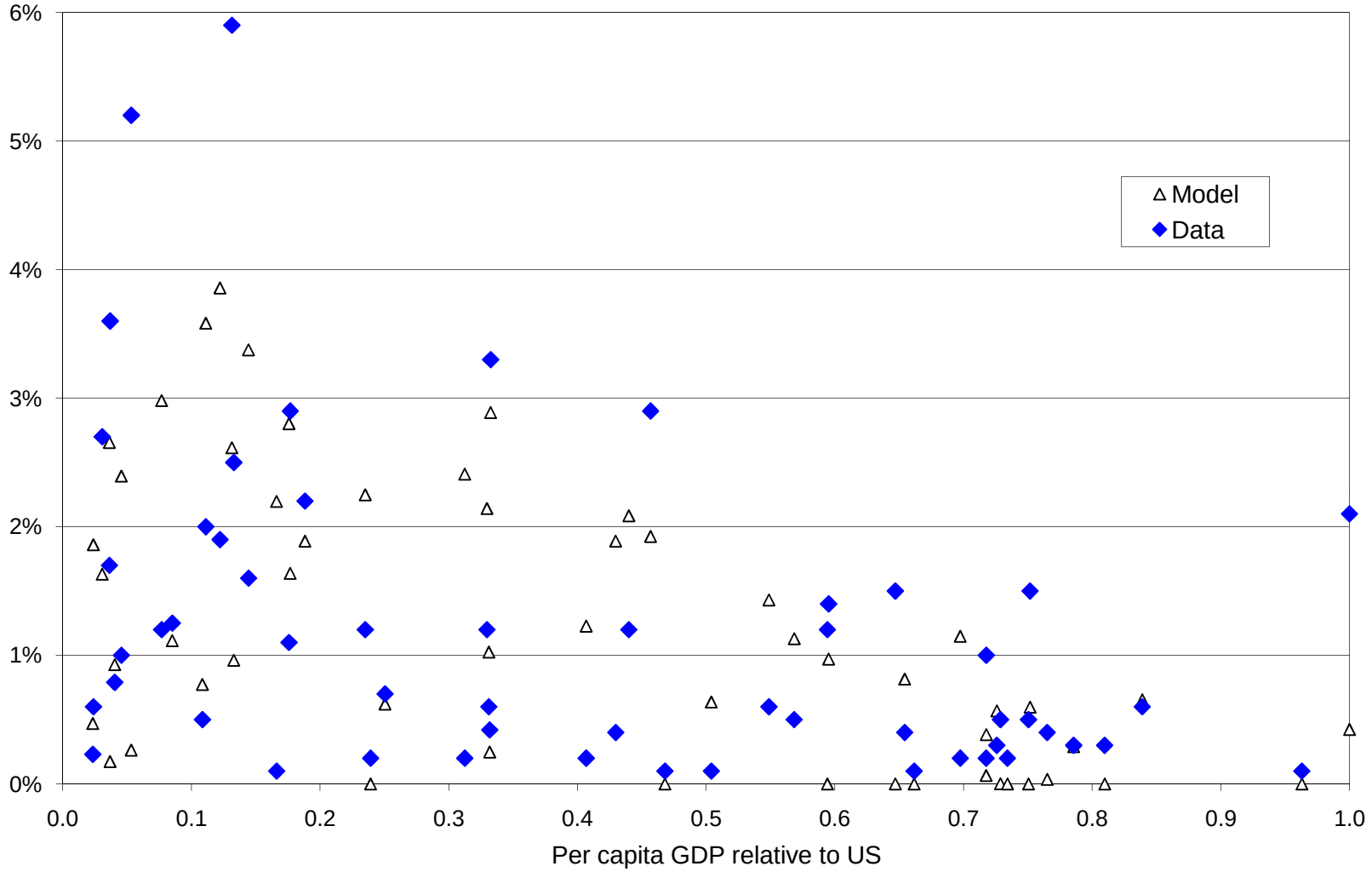


Figure 6. Quality of human capital

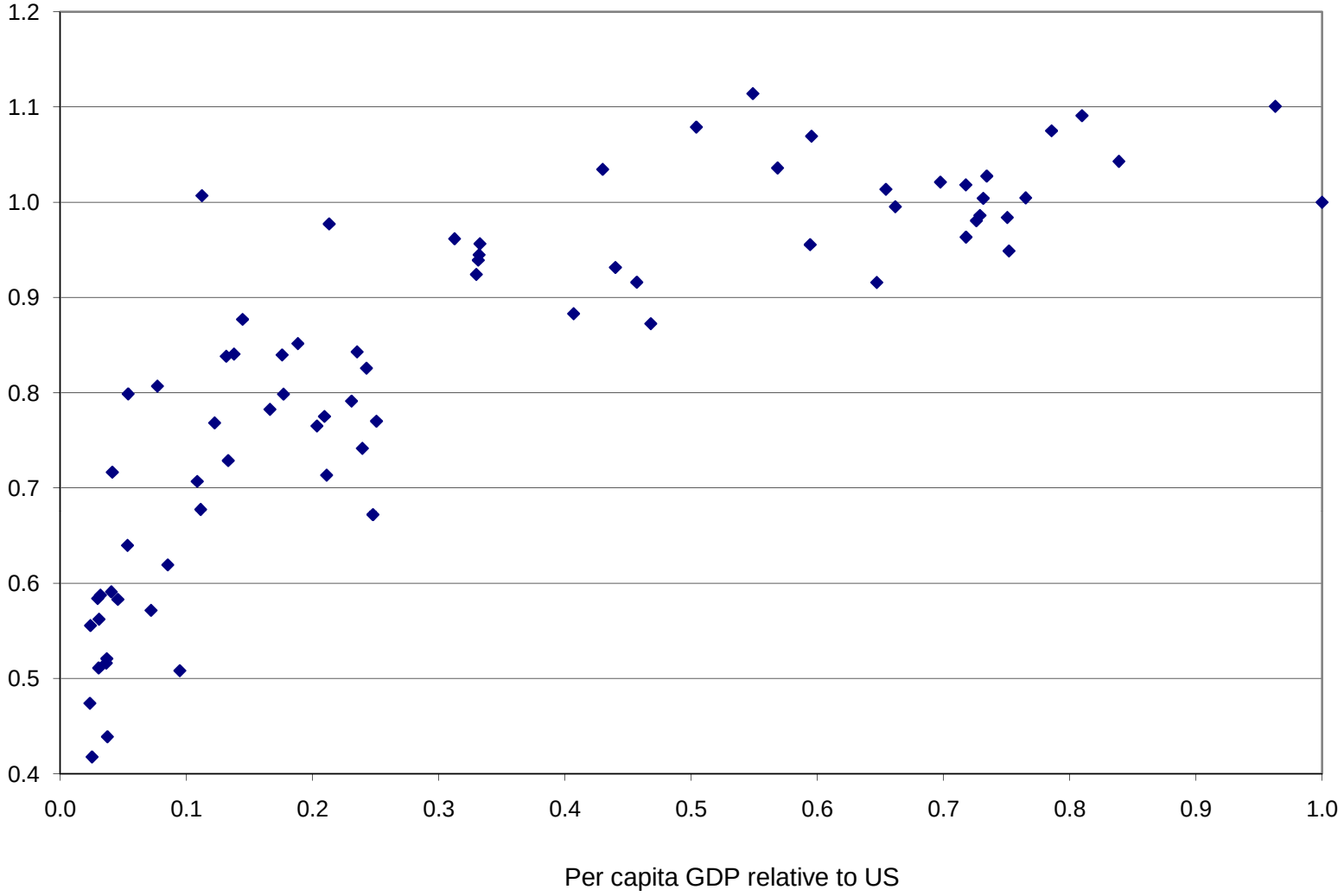


Figure 7. Borrowing constraints: cW/cS

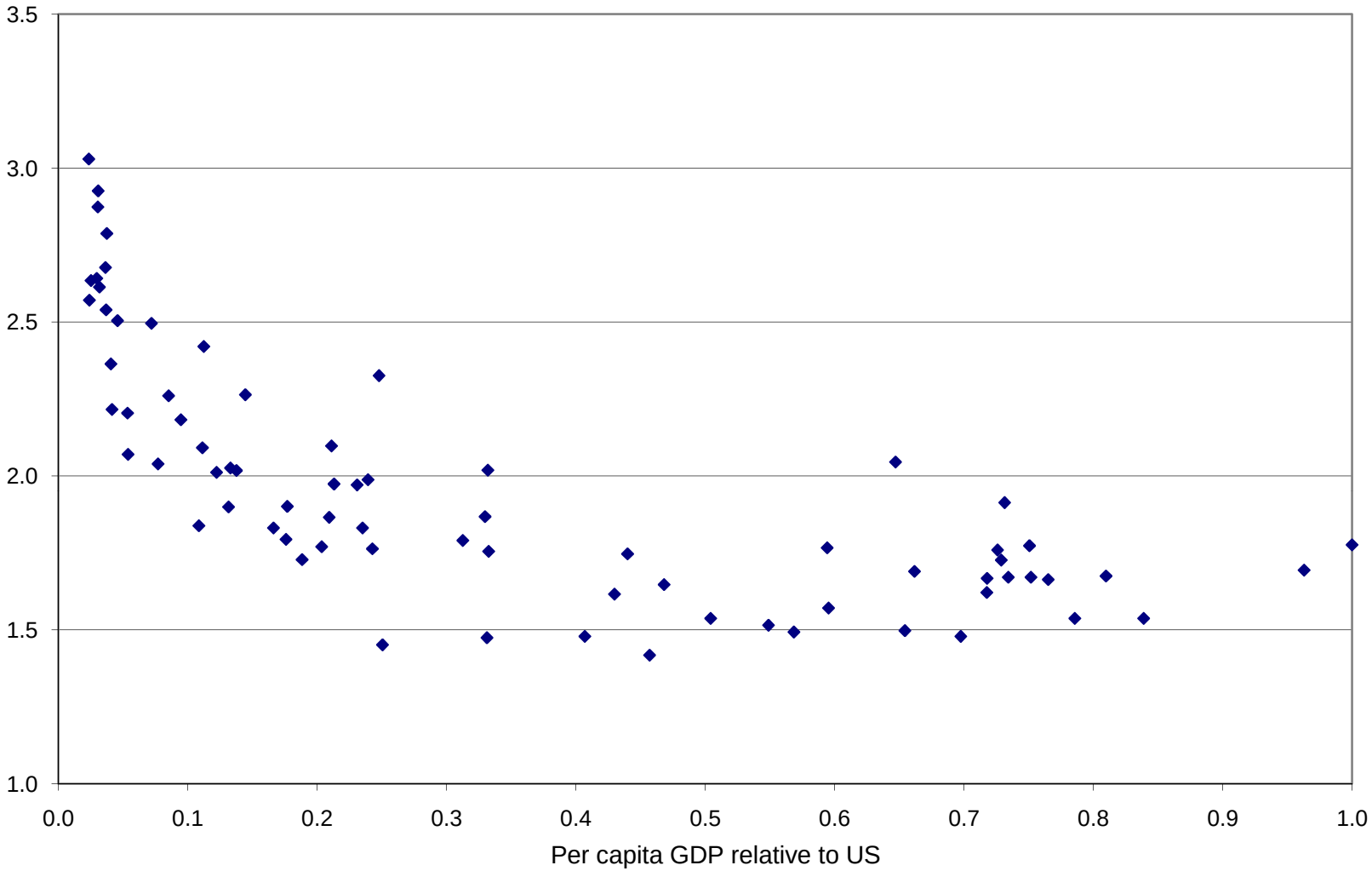


Figure 8. Parental transfers

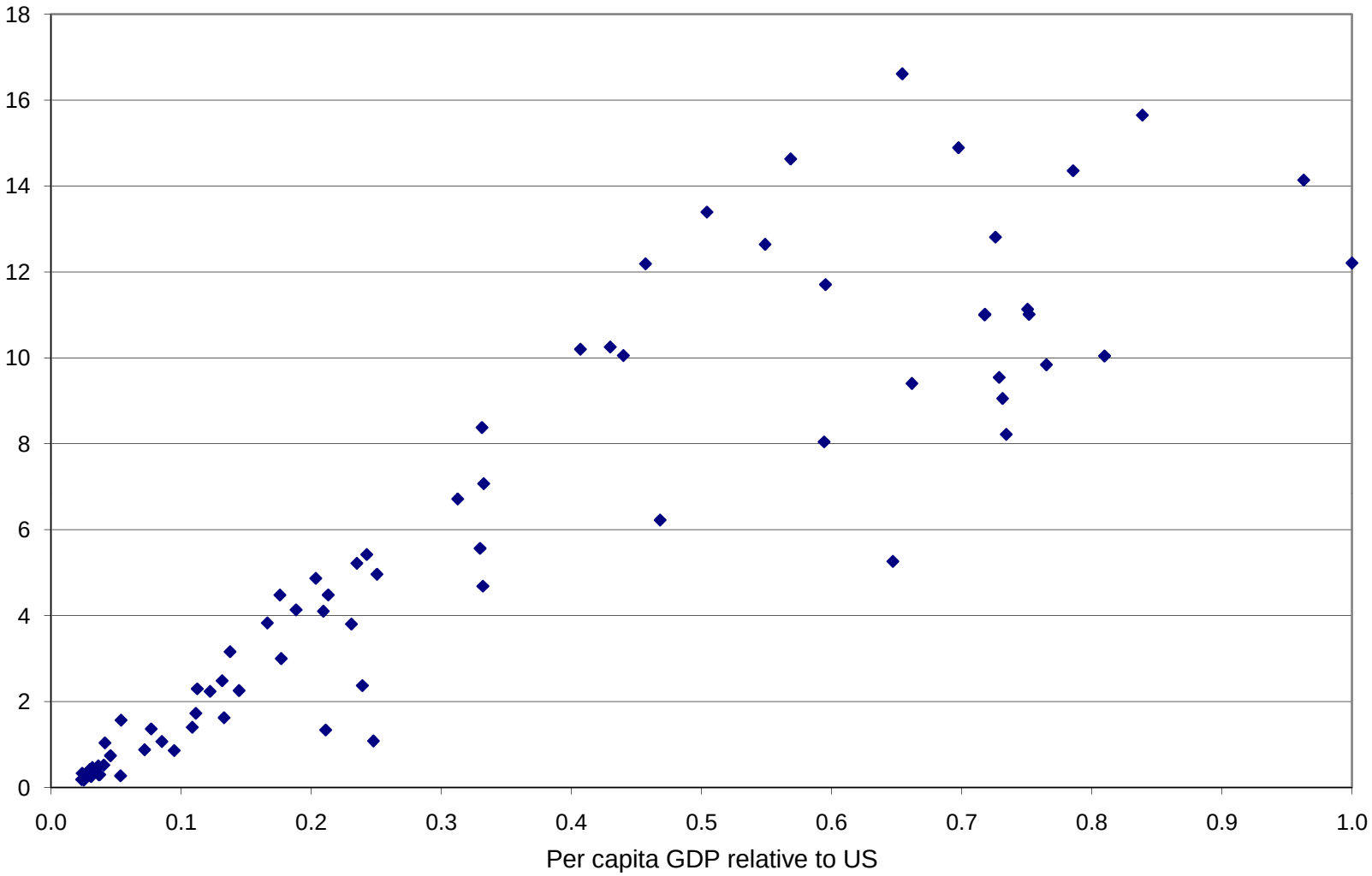


Figure 9. Schooling: benchmark and counterfactual

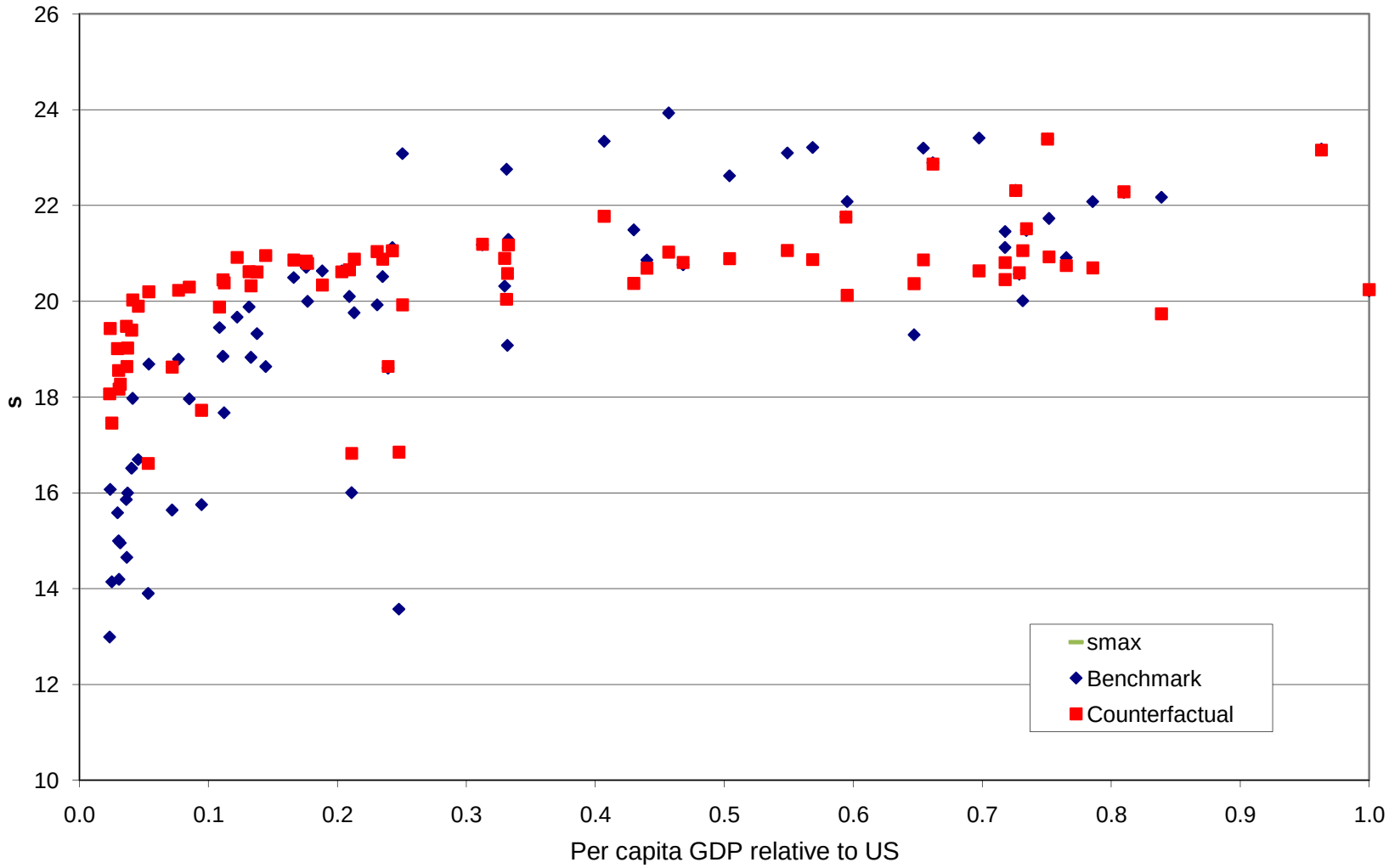


Figure 10. Parental transfers: benchmark and counterfactual

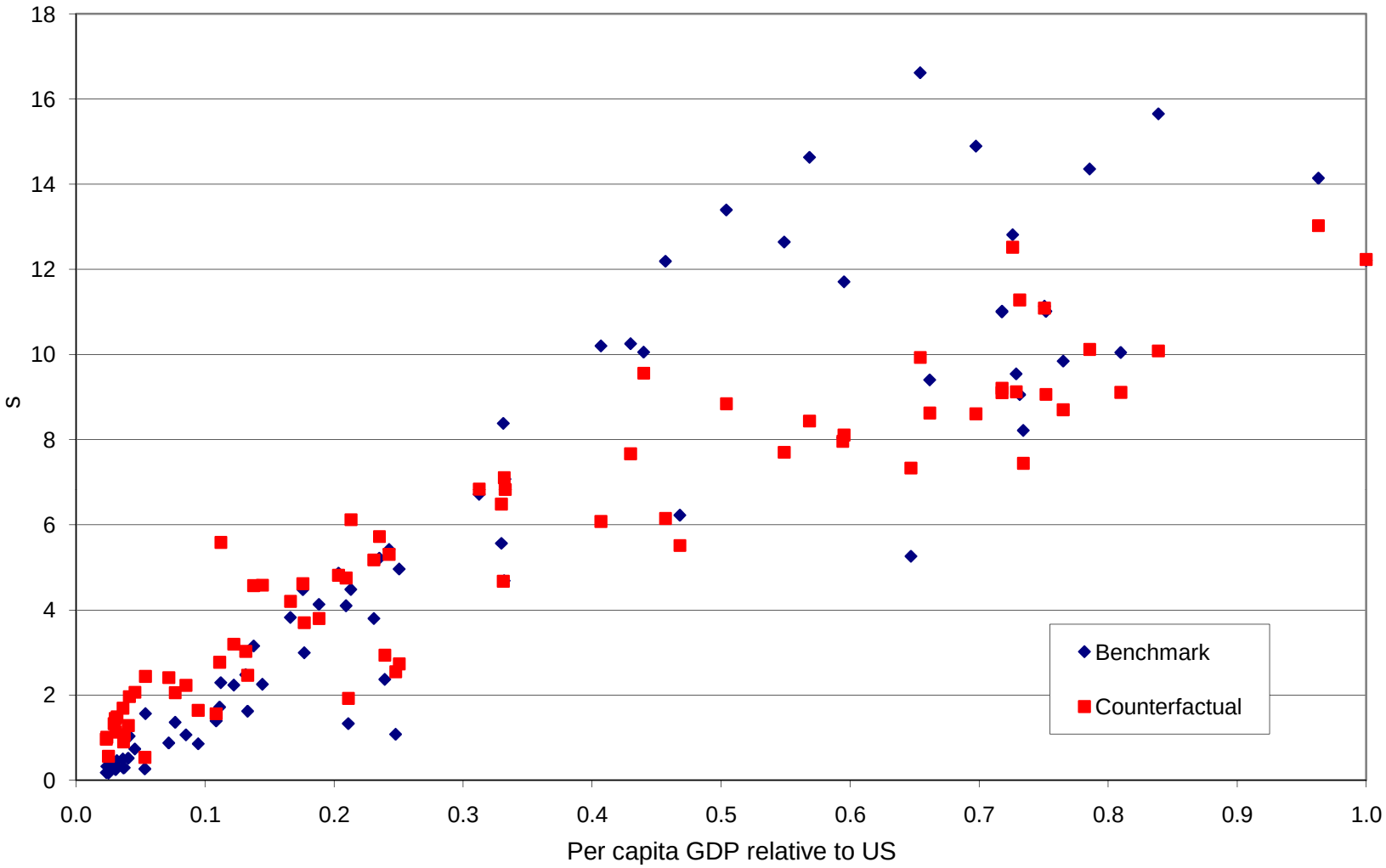


Figure A1. Years of Schooling - 2005
Data versus Bils and Klenow model - Subsample

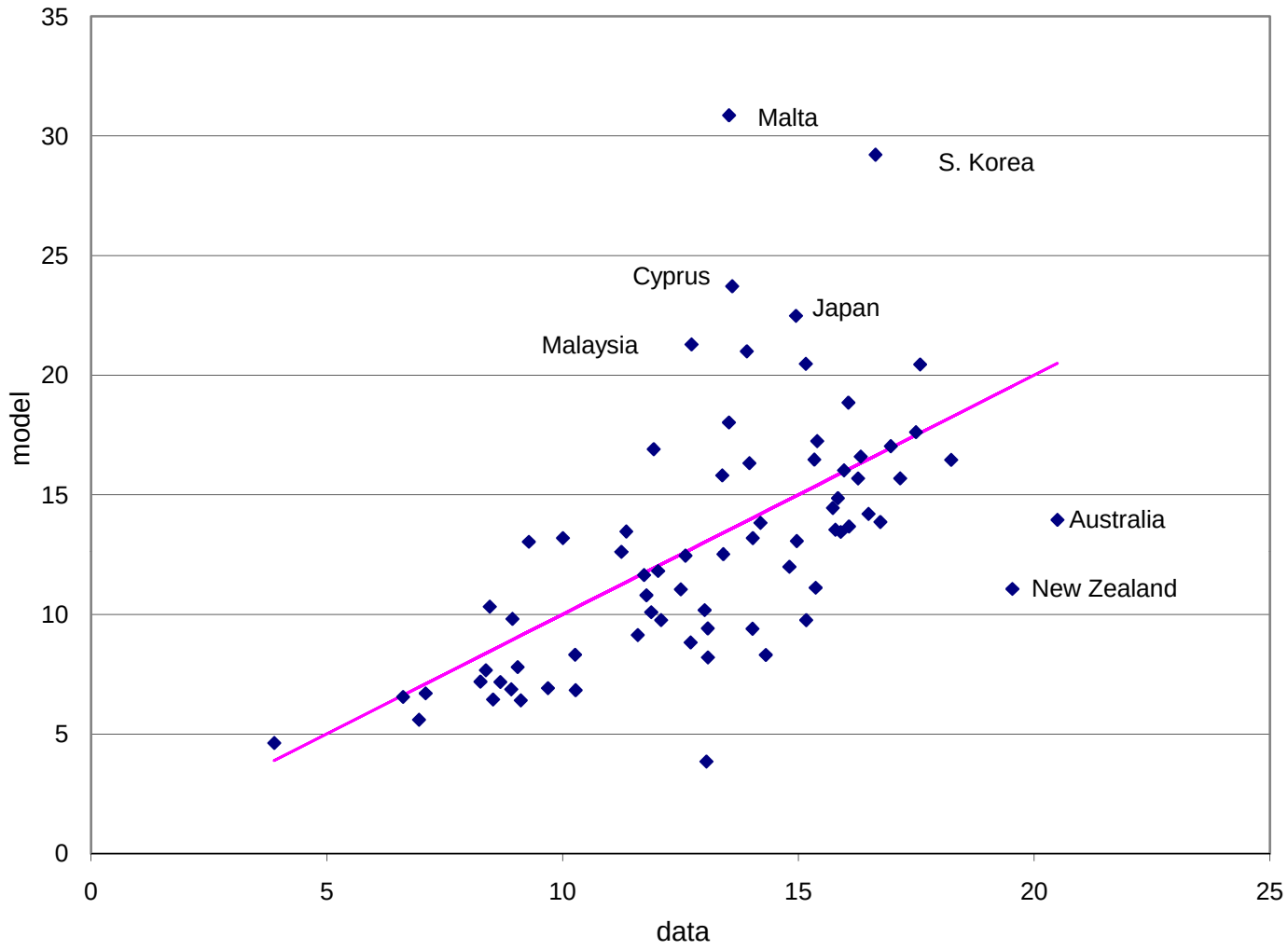
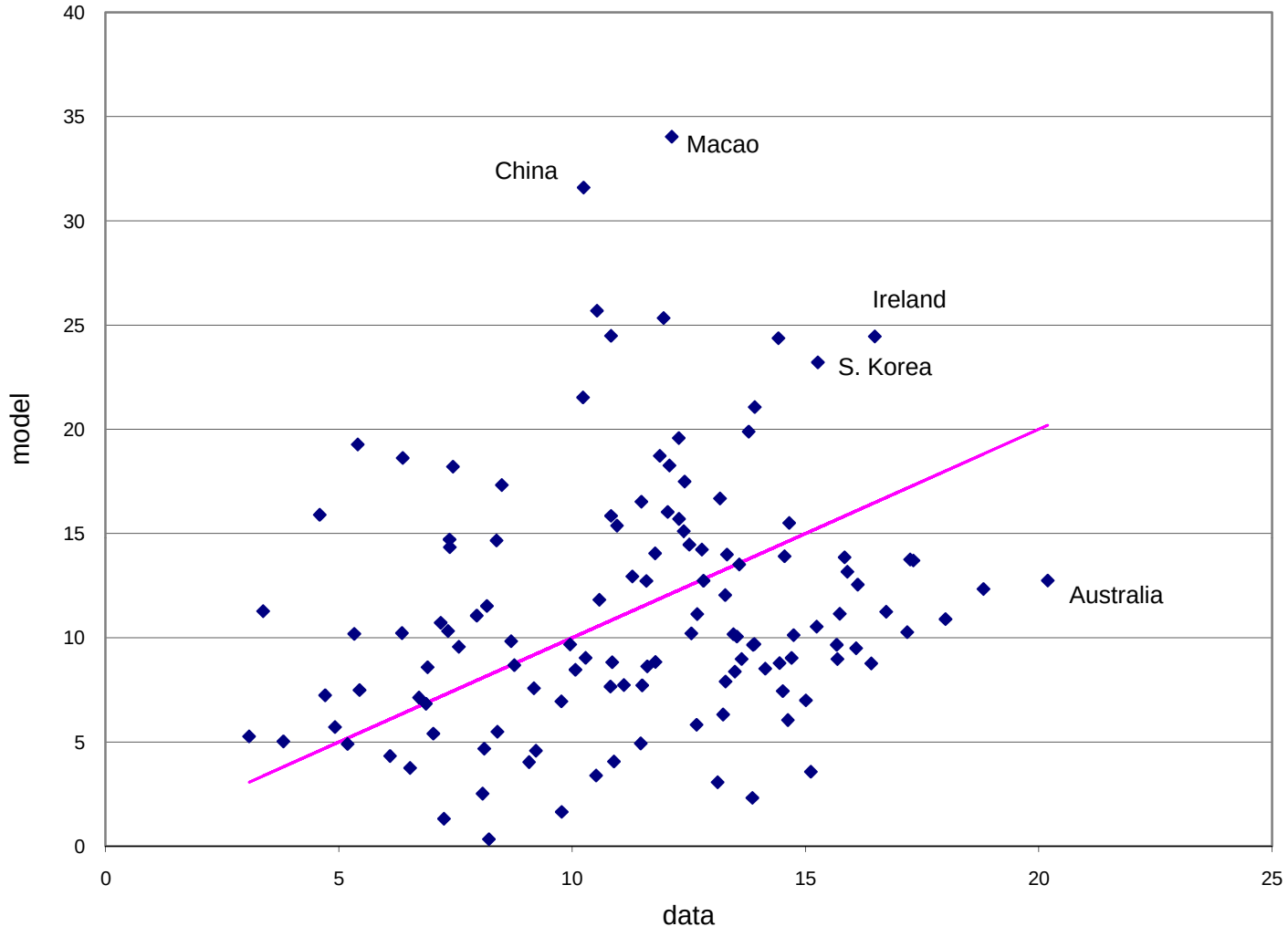



Figure A2. Years of Schooling -1999
Data versus Bils and Klenow Model - Subsample