ACCOUNTING FOR THE INTERNATIONAL QUANTITY-QUALITY TRADE-OFF

Juan Carlos Córdoba [*], Xiying Liu [†], and Marla Ripoll [‡]

March, 2020

Abstract

We investigate what accounts for the observed international differences in schooling and fertility, in particular the role of TFP, age-dependent mortality rates and public education policies. We use a generalized version of the Barro-Becker model that: *(i)* includes accumulation of human capital; *(ii)* allows for separate roles for intertemporal substitution, intergenerational substitution, and mortality risk aversion; and *(iii)* considers intergenerational financial frictions. We calibrate the model to a cross-section of countries in 2013. We find that while differences in TFP account for a large fraction of the dispersion in schooling, fertility and income per capita, public education subsidies play a major role. Public education spending per pupil matters relatively more in explaining the dispersion of fertility, while both the amount spent per pupil and the duration (years) of the subsidy are important in accounting for the dispersion of schooling. Eliminating public education subsidies results in an increase in average fertility, a decrease in human capital and income per capita, and an increase in the dispersion of schooling, fertility and income.

Key words: public education subsidies, intergenerational financial frictions, fertility, mortality, schooling, parental altruism, TFP

JEL Codes: I25, J13, O50

1 INTRODUCTION

Fertility data from the World Bank and school enrollment data from UNESCO indicate that a woman in Niger is expected to have 7.62 children, and each child is expected to attend school for 5.3 years. In contrast, a woman in Finland is expected to have 1.75 children, and her children are expected to attend school for 19.6 years. Figure 1 illustrates this well-known international quantity-quality trade-off for a cross-section of 92 countries in 2013.[1] Around the world, one more child per woman is associated with an average of three fewer years of schooling.

[*]Department of Economics, Iowa State University. E-mail: cordoba@iastate.edu
[†]Economics and Management School, Wuhan University. E-mail: xiyingliu@whu.edu.cn
[‡]Department of Economics, University of Pittsburgh. E-mail: ripoll@pitt.edu

[1]The measure of schooling in Figure 1 corresponds to school life expectancy from UNESCO, which is the years of schooling a child is expected to attend given the current enrollment rates at all ages. More details on the data used in Figure 1 are explained in Section 3.

Many explanations for the high-fertility low-schooling trade-off have been identified in the literature: high child mortality risk; low wages and the associated low opportunity costs of allocating time to raise children; low returns to schooling; limited access to high-quality publicly provided education; and other social norms and cultural factors. While not an exhaustive list, examples of some of the empirical and theoretical papers exploring these explanations include: child mortality risk (Angeles, 2010; Canning *et al.* 2013; Wilson, 2015); wages and time cost of raising children (Barro and Becker, 1989; Becker and Barro, 1988; Becker and Lewis, 1973; Galor and Weil, 1996; Manuelli and Seshadri, 2009); returns to schooling (Becker *et al.*, 1990; Galor and Weil, 2000); and provision of public education (Breierova and Dufflo, 2004; Castro-Martin and Juarez, 1995; de la Croix and Doepke, 2004; Doepke, 2004; Kirk and Pillet, 1998; Pradhan and Canning, 2015; Ferreira *et al.*, 2019).

This paper proposes a unified microfounded framework to quantitatively assess the contribution of multiple factors in explaining the international evidence on schooling and fertility. We focus on the role of differences across countries in three types of variables: total factor productivity (TFP) or wages; age-dependent mortality rates; and the provision of public education subsidies in terms of spending per pupil and the number of years the subsidy is provided. While these factors have been analyzed separately in the literature, we study them within the same unified framework. More importantly, relative to the macroeconomics and development literature, our model provides a framework to analyze cross-country differences in public education subsidies and their role in affecting schooling and fertility choices.

Our paper has three distinct features relative to the literature. First, it uses a generalized version of the Barro and Becker (1989) model that: *(i)* includes accumulation of human capital as in Ben-Porath (1967); and *(ii)* allows for separate roles for intertemporal substitution, intergenerational substitution, and mortality risk aversion. As shown in Cordoba and Ripoll (2019), separating intergenerational from intertemporal substitution allows the model to be consistent with both the low intertemporal substitution typical of quantitative macro models, and the negative income-fertility relationship across countries. In addition, as shown in Cordoba and Ripoll (2017), separating intertemporal substitution from mortality risk aversion using preferences à la Epstein-Zin-Weil allows for the value of statistical life to be proportional to income, a desirable feature for cross-country comparisons. In particular, this proportionality eliminates the income effects introduced by the non-homothetic framework of Becker *et al.*(2005).

Second, intergenerational financial frictions play a central role in our model: parents have access to credit, but children fully depend on their parents' resources during schooling years. Parents have no control over their adult children's income, cannot borrow against their children's future income, and cannot enforce financial obligations on their children to compensate for the cost of raising them.[2] These financial frictions are at the core of the model's quantity-quality trade-off because parental income becomes a determinant of the number of children as well as the educational resources that can be invested on each of them. In large families income is diluted among the many children, each

[2]See Schoondbroot and Tertilt (2014) and Cordoba and Ripoll (2016) for a more detailed discussion on the relevance of these intergenerational credit constraints.

of whom will have access to less resources during childhood.

The third distinct feature of our paper is the role of public education subsidies. Since our focus is on cross-country differences on average schooling and fertility, we abstract from heterogeneity among dynasties in the same country, but only model within-country age heterogeneity. From this perspective, the provision of public education does not alter within-country inequality but directly changes the allocation of resources across generations. In particular, public education subsidies partially counteract the effects of intergenerational financial frictions. In the model governments tax parents in order to provide education subsidies, which are received directly by the child in the form of education spending and only when the child attends school. Since parents in the model cannot borrow against the future income of their children, nor they can impose debt obligations on them, even altruistic parents underinvest in the education of their children relative to the case with no intergenerational financial frictions. By taxing parents and providing education subsidies, governments guarantee a given level of educational investments in children. Although education subsidies do not optimally resolve intergenerational financial frictions, they do change the distribution of resources across generations and in favor of children.

More interestingly, in our model public education subsidies affect parents' incentives to privately invest in the education of their children. This is the case because public subsidies are provided for a limited number of years and there is complementarity in educational investments across ages. In most countries public education is subsidized starting in elementary school, when the child is six years old, and up until a number of years that varies across countries. Since parents are taxed but their children can only start receiving the education subsidy at age six, the complementarity in educational investments across ages raises the returns to privately investing in children before age six. Therefore, public education subsidies in the presence of this complementarity induce a quantity-quality trade-off –parents can spend more on the education of each child before age six when they have less children.

The calibration of our model features cross-country differences in key dimensions that turn out to be quantitatively relevant. Age-dependent mortality rates are calibrated to fit country-specific life tables. The provision of public education in each country reflects realistic heterogeneity in both the number of years of provision (extensive margin), as well as spending per pupil (intensive margin) as documented by UNESCO. Total factor productivity (TFP) levels are computed as residuals to match per capita GDP in each country in 2013. Parameters common to all countries are calibrated to match features of the international evidence. As we show, these basic forces go a long way in explaining the international schooling and fertility data.

Our analysis yields several insights. First, cross-country differences in TFP (wages), age-dependent mortality rates and public schooling policies account for a large fraction of the world distribution of schooling and fertility. The correlation between fertility in the model and the data is 76%, and that for schooling is 78%. Regarding the international quantity-quality trade-off, while in the data one extra child is associated with three fewer years of schooling, in the calibrated model one extra child is associated with 2.4 less schooling years.

The second main finding is that although TFP (wages) is quantitatively the most important exogenous variable in explaining the dispersion of schooling, fertility and income per capita, public education variables also play a major role. We find that equating TFP to the 90th percentile level in the sample reduces the standard deviation of schooling by 54%, that of fertility by 72% and that of income per capita by 76%. A major role for cross-country TFP differences is also found in Manuelli and Seshadri (2009). But interestingly, public education variables are also quantitatively important. Equating both the number of years of education subsidy provision and spending per pupil to the 90th percentile in the sample reduces the standard deviation of schooling by 47%, that of fertility by 62% and that of income per capita by 59%.

The third main insight of the paper regards the details of how public education provision affects schooling and fertility. Policy makers have long advocated for education as a key intervention to lower births per women and foster economic development. For example, according to the World Bank Group (2011) "... the development benefits of education extend well beyond work productivity ... to include better health [and] reduced fertility ..." (p. 13). One of the contributions of our paper is to evaluate this policy prescription within a microfounded model. We find that the extensive and intensive margins of public education provision have differential effects on schooling and fertility. Public education spending per pupil, the intensive margin, matters relatively more in explaining the dispersion of fertility, while both the amount spent per pupil and the years of public provision are important in accounting for the dispersion of schooling.

Equating public spending per pupil to the 90th percentile in the sample reduces the dispersion of fertility by 50%, but the reduction is only 14% when the duration of the subsidy is equalized across countries. There are two main reason why this occurs in the model: one is that higher human capital raises the time costs of having children; and the other is that human capital accumulation in the model features complementarity in education spending across ages. Everything else equal, higher public education subsidies increase human capital, increase the opportunity cost of time, and decrease fertility. In addition, since provision of public education subsidies in most countries starts in grade school around age six, parents respond to higher public spending per pupil by increasing the spending per child both in pre-school years and after the public subsidy ends. In contrast, on the extensive margin, adding more years of education subsidies does not decrease fertility by much. For instance, in poorer countries in which subsidies per pupil (PPP adjusted) are quite low, offering the same low subsidy for additional years does not increase human capital significantly. This "low quality" subsidization does not raise the time cost of raising children enough to generate a large drop in fertility and higher private education investments on each child. The main insight of this analysis is that providing higher educational subsidies or higher education "quality", even for a limited number of years, can become an effective way of reducing fertility levels, specially for poorer countries in which spending is low.

Fourth, we find that eliminating public education subsidies around the world would results in an increase in average fertility, a decrease in human capital and income per capita, and an increase in the dispersion of schooling, fertility and income. More importantly, eliminating public education has a differential effect in poorer versus richer countries. When public education subsidies are eliminated in poorer countries, schooling decreases and fertility increases. In poor countries, having some education subsidies that can only be claimed when the child goes to school changes parental incentives towards complementing educational investments and having less children. Without this incentive parents in low-wage countries with a low time cost of raising children maximize utility by using the "quantity" margin and having more kids. From this perspective, public education subsidies in poorer countries work as a sort of "big push" mechanism out of a high-fertility and low-human-capital trap. But the effects are different for richer countries. We find that when public education subsidies are eliminated in high-TFP countries, fertility tends to decrease and schooling increases. Without the education subsidy but in an environment with high wage per unit of human capital, parents substantially increase private education investments in children and have less kids. However, parents will not spend enough to fully compensate for the lost public subsidy, implying a fall in human capital. Here the mechanism is again the complementarity in spending across ages: without a government providing generous subsidies starting at age 6, parents spend less across all ages. Lower human capital results in lower per capita income in rich countries. Therefore, without public subsidies rich countries still remain high-schooling and low-fertility, but low private investments in education result in lower human capital and lower income. More interestingly, countries with higher TFP among the rich might experience a "demographic drag", as fertility rates fall below replacement levels. In these countries, the combination between lower human capital and a shrinking share of working-age population translates in non-trivial declines in per capita income.

The fifth main insight refers to the role of mortality. According to our analysis, reducing under-5 and retiree mortality rates to their respective 10th percentile values has small effects on schooling, fertility and income per capita. This result contrasts with the traditional emphasis demographers have placed on the role of child mortality in fertility choices, but it is not surprising given the observed drastic reductions in child mortality among poor countries. There is more quantitative action from the reduction in adult mortality rates. We find that decreasing mortality rates for those between ages 5 and 65 in all countries results in a drop of 24% of the schooling dispersion, and a drop of 20% of the fertility dispersion. In this case, a higher probability of surviving during working years increase the incentives to remain in school longer and invest in education, an effect that is stronger for countries with higher mortality rates. Higher human capital raises the time cost of having children and results in lower fertility rates.

Another insight of our analysis pertains to the importance of local counterfactuals in the cross-country context. While the macro-development literature has traditionally relied on global counterfactuals in evaluating the sources of cross-country dispersions, we find that in the case of public education policies, the existence of non-linearities makes local counterfactuals specially informative. For instance, we find that schooling has a non-monotonic relationship with TFP at high levels of educational subsidies: at very low TFP levels, high educational subsidies result in very high schooling. But as TFP increases, schooling falls sharply and then gradually increases. This non-monotonic behavior reflects the role of intergenerational financial frictions. At very low TFP levels, wages are so low that offering high education subsidies results in children attending school many years. The opportunity cost of staying at school is not high. While these children depend on parental resources to consume all these schooling years, it is still optimal to delay entering the labor force and having access to credit until enough human capital has been accumulated. But this trade-off quickly changes as TFP increases, in part because families also become smaller. Initially, as TFP rises the wage per unit of human capital is higher, and staying at school longer implies not only higher foregone wages, but also delaying the child's financial independence and direct access to credit. But for sufficiently high TFP, family size is small enough that parents are able to finance more children's consumption during schooling years, making it feasible for them to delay financial independence without incurring high welfare costs. This non-monotonic schooling pattern occurs at high levels of education subsidies. Therefore, global counterfactuals that equate public education subsidies to a high level are hard to interpret since changes in the standard deviation of schooling mask non-linear responses across countries.

When we perform local counterfactuals that increase the amount of the public education subsidy in each country by 10%, we find that while in almost all countries fertility drops and human capital increases, years of schooling tend to increase in poorer countries and decrease in richer countries. The mechanism behind this result is the complementarity of educational spending across ages. In fact, in response to a 10% higher educational subsidy, parents in almost all countries spend more in the education of their children. In poor countries where TFP is lower, additional private and public education spending results in students staying longer at school. But in rich countries where TFP and education subsidies are already high, the additional private and public spending allows students to stay at school less years, accumulate more human capital, and gain financial independence faster.

Interestingly, local responses are very different when the duration of the public education subsidy, the extensive margin, increases by 10%. In this case parents in almost all countries decrease private spending in education. The reason for this is twofold. First, in almost all countries in the world a typical child in 2013 was enrolled in school more years than the number of years public subsidies are provided. In this respect, except for a few countries, the duration of public education subsidies is not binding around the world. Second, our calibrated model predicts that in most countries in the world public education subsidies are higher than what parents optimally invest in education once the subsidy ends. Therefore by extending the duration of the subsidy by 10%, which is on average one extra year, children receive one more year of investments at a higher rate than what parents were providing. In this situation parents react by decreasing total private spending in education, but since the government subsidizes at a higher rate, human capital still increases in most countries. For the few poorer countries for which the duration of the public subsidy is binding, schooling increases one-to-one with the extended year of subsidy, and parents spend more in education to complement the subsidy. This local experiment underscores the special role public education subsidies play and the way they interact with the intergenerational allocation of private resources.

The final insight of our analysis pertains the role of demographics on the long-run predictions of per capita income across countries. Our model predicts that if TFP, mortality, fertility, school enrollment and public spending per pupil remain at their 2013 level, in the long run GDP per capita will increase in most countries, but proportionally more in poorer countries. Two opposing forces explain this result: on the one hand, the human capital of the working age population will be higher everywhere, more so in poorer countries where school enrollment and the expansion of public education policies has been significant relative to that of the working age population in 2013. On the other hand, there will be a "demographic drag" in richer countries, where high TFP levels and low fertility rates will shrink the share of the working age population in the long run. This demographic drag is pervasive not only among rich countries, but also among many middle-income countries. For instance, our calibration predicts that if it was not for human capital, the demographic drag alone would decrease GDP per capita in the long run by about 12% in the US, 20% in Japan, 25% in Italy and up about 35% in Korea. Many, but not all, countries with GDP per capita below $10,000 will not face this demographic drag. On balance, the human capital force prevails in most countries, counteracting the adverse effect from demographics. But a few countries like Hong-Kong, Korea and Hungary will experience a drop in GDP per capita in the long run. Our model provides a useful framework for long-run projections of the sort produced by the United Nations and other policy agencies, with the important distinction that ours are based on a microfounded model.

As mentioned, the literature on the international quantity-quality trade-off is vast. Our paper is most related to Manuelli and Seshadri (2009), Cordoba and Ripoll (2013), Doepke(2004) and Ferreira *et al.* (2019). Like Manuelli and Seshadri (2009), we study steady states for a cross-section of countries. TFP differences across countries are considered in both papers and turn out to be quantitatively important in both. But different from them, we consider the role of age-dependent mortality rates and public education policies. In this respect, our paper complements theirs. The systematic expansion of public education in developing countries, in part driven by initiatives from international organizations such as the United Nations in the context of the Millennium Development Goals, prompted us to analyze the potential effects of these policies. From this perspective, our analysis is novel and provides a microfounded framework to study the long-run effects of these programs.

Similar to Cordoba and Ripoll (2013), intergenerational financial frictions play a central role in our theory. While they analyze difference in cross-country schooling attainment with exogenous fertility, here we make fertility endogenous allowing demographics to respond to changes in the economy. As mentioned, intergenerational frictions stem from children's financial dependency from parents during childhood years. They also capture the fact that the almost universal introduction of compulsory schooling laws around the world has substantially curtailed the ways in which parents used to control the income generated by children. Although not in the context of cross-country comparisons, a few papers have examined the effect of these financial frictions, including Rangazas (2000) in the case of schooling and human capital investment in the US, and Schoonbroodt and Tertilt (2010, 2014) and Cordoba and Ripoll (2016) in the case of fertility choice. The main implication of these financial frictions is that both family-level income as well as the provision of public education subsidies play an important role in schooling and fertility choices. Relative to the cross-country literature on fertility and schooling, our theory is unique in that it allows us to evaluate the role on public education policies.

Our work is similar to Doepke (2004) and Ferreira *et al.* (2019) on the focus on public education policies. Both of these papers examine the impact of public education provision on fertility, schooling and growth, and calibrate their models to the cases of Brazil and South Korea. Our cross-country analysis for a sample of 92 countries complements their work. By using a full cross-section of countries our analysis can more directly speak to the universal policies of public education expansion around the world, as well as to the current debates on how demographic change may alter the world income distribution in the long run. In addition, our emphasis on intergenerational financial frictions allows us to provide a new perspective on the role of public education subsidies in changing the allocation of resources across generations.

The remainder of the paper is organized as follows. Section 2 sets up the model and derives the optimality conditions of consumption, savings, intergenerational transfers, education spending and fertility. Key steady state results are derived from the model. Section 3 describes the model's calibration, discussing in detail the cross-country data used in the analysis, and the model's performance. Global counterfactuals are presented in Section 4. Section 5 reports our main results on the role of public education, including the effects of eliminating it, the model's non-monotonic responses, and the local counterfactuals. Section 6 discusses further implications of the model and some robustness, including the role of demographics on the steady state distribution of per capita income; an out-of-sample prediction exercise for the US in 1900; and an exercise introducing income taxes in richer countries. Section 7 concludes.

2 MODEL

We model a representative dynasty in each country, with a parent who is altruistic towards his children. A representative individual in this economy faces a stochastic life span with the time-0 probability of surviving up to age a given by $\pi(a)$. Time is continuous. Prices are assumed to be actuarially fair. In particular, assume $q(a) = e^{-ra}\pi(a)$ is the age-contingent actuarially fair price, with r the interest rate.

The focus of the analysis is on the decisions of the individual over the life cycle, in particular schooling, educational investments, consumption, saving, fertility, and transfers to children. An individual is a student from age 6 until an endogenously chosen age s. Public subsidies for attending school are available in the economy from ages 6 to $\overline{s}$. After completing schooling at age s the individual becomes a worker until he retires at age R. At age $F > s$ he becomes a parent to n children. Children depend on parental resources for consumption and educational investments until they finish school.

2.1 Individual's problem

2.1.1 Preferences

We build on Cordoba and Ripoll (2017, 2019) who provide insights on how to extend life-cycle models to study fertility and mortality. In particular, consider the following generalized version of the Barro-Becker and Epstein-Zin-Weil preferences. The lifetime utility of the representative individual, V, is given by

$$V = \frac{C^{1-\eta}}{1-\eta} + \Phi(n)V', \quad \eta \in (0,1), \tag{1}$$

where

$$C = \begin{cases} \left[\rho \int_0^\infty e^{-\rho a} \pi\left(a\right)^{\frac{1-\sigma}{1-\theta}} c\left(a\right)^{1-\sigma} da\right]^{\frac{1}{1-\sigma}} + \underline{C} & \text{if } \sigma > 0 \text{ and } \sigma \neq 1 \\ \exp\left[\rho \int_0^\infty e^{-\rho a} \left(\frac{1}{1-\theta} \ln \pi\left(a\right) + \ln c\left(a\right)\right) da\right] + \underline{C} & \text{if } \sigma = 1 \end{cases}. \tag{2}$$

Let us explain each of the components of preferences in turn. In equation (1) C corresponds to selfish utility, the utility the individual derives from his own lifetime consumption. Absent children, C would be the only source of utility. According to equation (1), individuals also enjoy the utility of their children. The total utility derived from n children is given by $\Phi(n)V'$, where $\Phi(n)$ is an altruistic weight and V' is the lifetime utility of each of the children. Function $\Phi\left(\cdot\right)$ satisfies $\Phi\left(0\right) = 0$, $\Phi'\left(n\right) > 0$, $\Phi''\left(n\right) < 0$ and $\Phi\left(n\right) < 1$ for $n \in [0, \overline{n}]$ where $\overline{n}$ is the maximum feasible number of children. Parameter η controls the willingness to substitute consumption among parents and their children. Following Cordoba and Ripoll (2018), we call $1/\eta$ the elasticity of intergenerational substitution. The restriction $\eta \in (0,1)$ is required for children to be goods rather than bads.[3]

Equation (2) describes selfish utility C. Parameter ρ is the discount factor, $c(a)$ is consumption at age a, $1/\sigma$ is the elasticity of intertemporal substitution (EIS), $\theta \in (0,1)$ is the coefficient of risk aversion, in this case aversion to mortality risk, and $\underline{C} > 0$ is non-market consumption. The restriction on θ guarantees that longevity is a good rather than a bad.[4]

The following are three distinct aspects of selfish utility C. First, the formulation separates intertemporal substitution from (mortality) risk aversion. Parameter σ controls the former, while θ controls the latter. When $\sigma = \theta$ equation (2) reduces to the standard expected utility formulation, case in which marginal rates of substitution are linear in survival probabilities $\pi\left(a\right)$. Cordoba and Ripoll (2017) show that a model that separates σ from θ can more successfully account for evidence on the willingness to pay for longevity and other evidence from the medical literature.[5] In our context, mortality risk affects fertility decisions to the extent that it affects the longevity of the child. Here we adopt the same flexible representation with the value of σ determined from the

[3] Cordoba and Ripoll (2011, 2018) consider the general case $\eta \geq 0$.

[4] Cordoba and Ripoll (2017) consider the general case $\theta \geq 0$.

[5] Additional insights and technical details on the advantages of disentangling σ from θ can be found in Section 2 of Cordoba and Ripoll (2017). A cross-country comparison application is in their Section 4.1.

observed degree of consumption smoothing over the life cycle and θ determined from estimates of the value of statistical life in the health literature.

Second, the non-homotheticity due to non-market consumption $\underline{C}$ is introduced to create a quantitatively stronger link between fertility and a country's level of income. As discussed in Cordoba and Ripoll (2016), in the homothetic case both the marginal benefit and the marginal cost of having children are proportional to wages, eliminating the effect of income on fertility. As shown below, the presence of both lump-sum taxes and $\underline{C}$ break this proportionality. Both lump-sum taxes and $\underline{C} > 0$, together with the restriction $0 < \eta < 1$, allow the model to be consistent with the inverse relationship between income and fertility documented in the cross-country data.[6]

Third, our preferences also separate intertemporal from *intergenerational* substitution. From equations (1) and (2) it can be seen that when $\sigma = \eta$ the standard dynastic representation is obtained. Cordoba and Ripoll (2016, 2018) show that separating σ from η is important for dynastic models to be consistent both with the economic value of a child, and with the negative income-fertility relationship documented within and across countries.[7]

It is important to notice that our representation of preferences in (1) and (2) is quite general and flexible. It includes as a special case the expected utility version of the dynastic model, which is obtained when $\sigma = \eta = \theta$ and $\underline{C} = 0$. In this case the model reduces to the Becker and Barro (1988) framework under the additional assumption that $\Phi(\cdot)$ is isoelastic.

2.1.2 Human capital

The individual accumulates human capital by going to school and investing resources in education. Expenditures in education at age a, $e(a)$, are composed of a public subsidy, $e_p(a)$, and private education spending, $e_s(a) \geq 0$. We assume that the public subsidy is given by e_p between ages 6 and a maximum age of $\overline{s}$:

$$e_p(a) = \begin{cases} e_p & \text{if } 6 \leq a \leq \overline{s} \\ 0 & \text{otherwise} \end{cases} . \tag{3}$$

We allow for cross-country differences in $\overline{s}$ and e_p. As summarized in Lee and Barro (2001), there is evidence that differences in educational resources per pupil, which include higher teacher salaries and instructional materials, are important in explaining differences in student achievement across countries. Let $\mathbf{E}$ be a vector of educational expenditures for all ages. At the end of s years of schooling human capital is given by

$$h(s, \mathbf{E}) = \left(\int_0^s (d \cdot e(a))^\beta \, da \right)^{\gamma/\beta}, \tag{4}$$

[6] Notice from equation (2) that $\underline{C}$ is a reduced-form way of capturing non-market consumption. This term is only introduced to break the proportionality in the optimality condition for fertility. In a more general formulation, $\underline{C}$ should also be a function of the age-dependent probability of survival. Since $\underline{C}/C$ turns out to be small in the calibration, and since the more general formulation complicates the utility function without adding much to the analysis, we introduce $\underline{C}$ as a reduced-form parameter in (2).

[7] See Section 2 in Cordoba and Ripoll (2018).

where

$$e(a) = e_p(a) + e_s(a).$$

The human capital production function in equation (4) is a version of Ben-Porath (1967). Parameter $\beta \in (0,1]$ determines the degree of substitution among educational investments at different ages; $\gamma \in (0,1]$ determines the returns to scale; and d is the fraction of school non-repeaters. The restriction on β guarantees that $\partial h(s, \mathbf{E}) \partial s > 0$. Parameter d is introduced to account for differences in repetition rates across countries and to avoid overestimating human capital by double-counting expenditures.

Notice that we do not model public and private schools as separate entities, but instead assume that the public subsidy $e_p(a)$ is a perfect substitute for private spending $e_s(a)$ from ages 6 to $\overline{s}$. However, in our model public education subsides do not fully crowd out private spending because the public subsidy is not provided at all ages, particularly before age six, and also because the model features a complementarity of educational expenditures across ages, as we now turn to explain.

Self-productivity and complementarity To better understand our human capital production function, we follow Cunha *et al.* (2006) and characterize its properties in terms of self-productivity and complementarity. For this purpose, consider for a moment the discrete-time version of the time-derivate of equation (4). It satisfies the following version of Ben-Porath's (1967) formulation [8]

$$h(a+1) = z_h h(a)^{\gamma_1} (d \cdot e(a))^{\gamma_2} + (1-\delta_h)\, h(a) \equiv g(h(a), e(a)) + (1-\delta_h)\, h(a) \tag{5}$$

where $g(h(a), e(a)) \equiv z_h h(a)^{\gamma_1} (d \cdot e(a))^{\gamma_2}$ is the gross educational investment, and $(1-\delta_h)\, h(a)$ is undepreciated human capital. Our representation in (4) assumes $\delta_h = 0$ and normalizes the ability parameter $z_h = 1$.[9] Finally, the mapping between parameters in (4) and (5) is given by $\gamma_1 \equiv 1 - \beta/\gamma$ and $\gamma_2 \equiv \beta$.

Following Cunha *et al.* (2006), self-productivity corresponds to the notion that human capital at certain age raises human capital at later age. From this perspective, self-productivity arises when

$$\frac{\partial h(a+1)}{\partial h(a)} = \gamma_1 h(a)^{\gamma_1 - 1} (d \cdot e(a))^{\gamma_2} + 1 - \delta_h > 0.$$

Since we assume $\delta_h = 0$, self-productivity holds when $\gamma_1 \equiv 1 - \beta/\gamma > 0$, i.e., when $h(a)$ has a positive effect in gross educational investment $h(a)^{\gamma_1} (d \cdot e(a))^{\gamma_2}$. In the calibration we verify that $\beta/\gamma < 1$, confirming that our human capital production function exhibits self-productivity.

Complementarity captures the notion that early educational investments facilitate the produc-

[8] See derivation and discussion in Cordoba and Ripoll (2013).

[9] We normalize $z_h = 1$ because we consider a representative dynasty per country. As we discuss below in the calibration, introducing cross-country differences in students' ability z_h is not quantitatively relevant in explaining schooling differences.

tivity of later investments. Complementarity arises when

$$\frac{\partial^2 g(h(a),e(a))}{\partial h(a)\partial e(a)} = \gamma_1\gamma_2 h(a)^{\gamma_1-1}\left(d\cdot e(a)\right)^{\gamma_2-1} > 0,$$

in other words, there is complementarity when human capital stock $h(a)$ raises the marginal productivity of educational investments, $\partial g(h(a),e(a))/\partial e(a)$. Again, since $\gamma_2 \equiv \beta > 0$, then complementarity holds when $\gamma_1 \equiv 1-\beta/\gamma > 0$ or $\beta/\gamma < 1$, which we verify holds in our calibration.

Returns to schooling The returns to schooling implied by (4) are given by

$$r_s(s) = \frac{\partial \ln h\left(s,\mathbf{E}\right)}{\partial s} = \frac{\gamma}{\beta} h\left(s,\mathbf{E}\right)^{-\frac{\beta}{\gamma}}\left(d\cdot e(s)\right)^{\beta},$$

which are decreasing in $h\left(s,\mathbf{E}\right)$ and increasing in education expenditures at age s, $e(s)$.

It is instructive to consider for a moment the special case $e(a) = e$. In that case, equation (4) simplifies to $h\left(s,\mathbf{E}\right) = (d\cdot e)^{\gamma} s^{\gamma/\beta}$ which makes clear the role of γ and β: γ is the elasticity of human capital with respect to expenditures, while γ/β the elasticity with respect to years of schooling. Returns to schooling in that case are given by $r_s(s) = (\gamma/\beta)(1/s)$, which highlights the role of γ/β and s as its key determinants.

Returns to experience Beyond schooling years, human capital is also enhanced through experience. In particular, we assume that human capital at age $R \geq a \geq s$ is given by

$$h(a;s,\mathbf{E}) = h\left(s,\mathbf{E}\right)e^{\nu(a-s)}, \tag{6}$$

where ν are the returns to experience.

2.1.3 Lifetime income and labor supply

The present value of the individual's lifetime income, in age-0 prices, is given by

$$W\left(s,n,\mathbf{E}\right) = \int_s^R wh\left(s,\mathbf{E}\right)e^{\nu(a-s)} l(n,a)q(a)da,$$

where w is the wage per unit of human capital. Labor supply at age a is given by $l(n,a)$. It is a function of n, as parents incur time costs in raising children.

2.1.4 Budget constraints

There are two stages during the lifetime of an individual: schooling years and working years, including retirement. We assume that individuals fully depend on parental resources during the first stage of life. Let b_1 denote the present value of this parental support in age-0 prices. The budget constraint for the first stage of life reads

$$b_1 \geq \int_0^s \left(c\left(a\right) + e_s\left(a\right)\right) q\left(a\right) da. \tag{7}$$

The assumption that during schooling years individuals totally depend on parental resources is natural for the average school-age child in each country. In practice, the typical school-age child cannot access financial markets. Parents have access to financial markets but cannot substitute for banks, particularly as lenders, because children's debt obligations are not enforceable.[10] A key issue is whether altruistic parents will transfer enough resources to each child as to perfectly smooth their consumption between the student and the working periods.

The budget constraint for the second stage of life, which starts at age s, reads

$$W(s, n, \mathbf{E}) + q\left(s\right) b_2 \geq \int_s^\infty c\left(a\right) q\left(a\right) da + \bar{\tau} \int_s^R q\left(a\right) da + q\left(F\right) n b_1' + q\left(F\right) q(s')\, n b_2', \tag{8}$$

where b_2 is the present value (in age-s prices) of the transfers the (adult) child receives from the parent during the child's working years. In turn, b_1' and b_2' are the transfers the child will give to each of his own n children for their schooling and working years respectively. Finally, $\bar{\tau}$ is a lump-sum tax used to finance public education.[11]

Parental transfers are assumed to be non-negative. This restriction is not binding for b_1' since positive transfers to school-age children are the only way to guarantee positive consumption of children during school years, so we only write the constraint that

$$b_2' \geq 0. \tag{9}$$

This constraint prevents parents from endowing their adult children with debt. When the present value of the child's future income is larger than the cost of raising the child, altruistic parents would find it optimal to have the maximum number of children and endow them with debt to compensate for the costs incurred during schooling years, and to extract rents from them.[12] As we show below, in equilibrium constraint (9) binds and parents do not transfer enough resources to perfectly smooth the consumption of their children between their two stages of life.

[10] Others in the literature have emphasized the importance of this type of frictions in modeling intergenerational links. See Schoondbroot and Tertilt (2010, 2014) and Cordoba and Ripoll (2016).

[11] We model taxes in a lump-sum fashion because for most countries in the sample distortionary income taxes are a small fraction of total tax revenue collection. Income tax collection is relatively larger only for richer countries. In Section 6.3 we show that our counterfactual exercises would change little in richer countries if education was financed with proportional income taxes rather than lump-sum taxes.

[12] In contrast with the binding constraint in equation (9), Barro and Becker (1989) focus on unconstrained solutions. They avoid the situation in which parents would like to have the maximum number of children and endow them with debt by assuming that children are a net financial cost to parents (i.e., the cost of raising the child is larger than the present value of the child's future income). See Cordoba and Ripoll (2016) for a detailed discussion.

2.2 Optimal allocations

Given parental transfers, b_1 and b_2, taxes and prices, we describe the individual's problem recursively as follows

$$V\left(b_1, b_2\right) = \max_{[c(a)]_{t=0}^{\infty}, b_1', b_2', [e_s(a)]_{t=0}^{s}, s, n \in [0, \overline{n}]} \frac{1}{1-\eta} C^{1-\eta} + \Phi(n) V\left(b_1', b_2'\right),$$

subject to (7), (8), (4), and (9). We use superscript * to denote optimal solutions and focus the presentation on steady state situations, i.e., $b_1^* = b_1^{*\prime}$ and $b_2^* = b_2^{*\prime}$. The Appendix includes a detailed model solution.

2.2.1 Optimal consumption and parental transfers

Let λ_1 and λ_2 be the Lagrange multipliers associated to the budget constraints (7) and (8) respectively. Let $c^S(s)$ and $c^W(s)$ denote consumption at time s as a student and as a worker respectively. Optimal consumption over the life cycle satisfies the following pair of conditions

$$c^*(a) = \left[e^{(r-\rho)a} \pi(a)^{\frac{\theta-\sigma}{1-\theta}}\right]^{\frac{1}{\sigma}} c^*(0), \text{ for } a \leq s, \text{ and} \tag{10}$$

$$c^*(a) = \left[e^{(r-\rho)a} \pi(a)^{\frac{\theta-\sigma}{1-\theta}} G\right]^{\frac{1}{\sigma}} c^*(0) \text{ for } s \geq a \tag{11}$$

where

$$G \equiv \frac{\lambda_1}{\lambda_2} = \left(\frac{c^W(s^*)}{c^S(s^*)}\right)^{\sigma}. \tag{12}$$

G is a key measure of relative scarcity, or the shadow price of student-age resources relative to working-age resources. Equations (10) and (11) are standard Euler equations, except for two features. First, the survival probability term $\pi(a)^{\frac{\theta-\sigma}{1-\theta}}$ affects the growth rate of consumption. Notice that in the case of the expected utility model with $\theta = \sigma$, this term disappears. Here, if $\sigma > \theta$, which we find to be the case in the calibration, higher survival rates result in lower consumption growth, a prediction absent in the standard case with $\theta = \sigma$. Second, term G in (11) describes the extent of the credit frictions in the model, mainly the role of the non-negative bequest constraints. In the absence of credit frictions, $G = 1$. According to (12), G measures the extent of the consumption jump at age s when the student becomes a worker.

Parental transfers ultimately determine the degree of credit frictions. The optimality conditions for transfers, b_1' and b_2', are

$$\lambda_2^{parent} q(F) n^* = \Phi(n^*) \lambda_1^{child} \text{ and} \tag{13}$$

$$\lambda_2^{parent} q(F) q(s^*) n^* > \Phi(n^*) \lambda_2^{child} q(s^*), \tag{14}$$

where we have written (14) for the case in which (9) binds and $b_2' = 0$. In what follows we write the model solution assuming this is the case, and later verify it in the calibrated model. The left-hand side of (13) and (14) are the marginal costs of transfers while the right hand side are the marginal benefits, to the parents.

To gain some further understanding, (13) can be written in the steady-state as

$$c^*\left(F\right)^{-\sigma} = \frac{1}{e^{-\rho F}\pi\left(F\right)^{\frac{1-\sigma}{1-\theta}}}\frac{\Phi(n^*)}{n^*}c^*(0)^{-\sigma}.$$

This equation corresponds to the intergenerational version of the Euler equation, equalizing the marginal utilities of the parent at age F and the child at age 0. Notice how average altruism $\Phi(n^*)/n^*$ plays a key role weighting the marginal utility of the child.

In the steady state conditions (13) and (14) simplify to

$$G \equiv \left(\frac{c^W\left(s\right)}{c^S\left(s\right)}\right)^{\sigma} = G(n^*) = e^{-rF}\pi\left(F\right)\frac{n^*}{\Phi(n^*)}, \tag{15}$$

and

$$G\left(n^*\right) > 1, \tag{16}$$

so that $G\left(n^*\right) > 1$ is a sufficient condition for a binding transfer constraint.

The determination of $G\left(n^*\right)$ is described in equation (15). It depends directly on parameters r, $\pi(F)$ and the altruistic function $\Phi\left(\cdot\right)$. More importantly, it depends directly on the fertility choice and indirectly on the parameters determining n^*. $G\left(n^*\right) > 1$ is more likely to hold when n is large, the interest rate is low and/or average altruism, $\Phi(n)/n$, is low. This means that parental transfers, even from altruistic parents, may not be enough to fully smooth the children's consumption when family size is large, altruism is low, or when low interest rates make it optimal to consume earlier. In particular, the model predicts that, other things equal, countries with larger fertility will have larger credit frictions. In those countries, children would receive less parental transfers and experience a larger consumption jump at age s.

2.2.2 Educational expenditures

The optimality condition for private educational expenditures, $e_s(a)$, is given by

$$q\left(a\right) \geq \frac{1}{G\left(n^*\right)}\int_s^R w\frac{\partial h\left(s^*, \mathbf{E}^*\right)}{\partial e_s(a)}e^{\nu(a-s^*)}l(n^*, a)q(a)da, \tag{17}$$

which holds with equality if $e_s(a) > 0$. In an interior solution this expression equates the marginal cost of spending one unit of consumption goods in education at age a, $q(a)$, with the marginal benefit, which corresponds to the increase in human capital. This benefit is discounted by the rate $1/G$ because benefits are realized during the second stage of life when resources are less scarce, while the cost is paid in the first stage when resources are more scarce. From this perspective, in countries with higher fertility and higher G, the benefits to educational investments are reduced.

In addition to private educational expenditures, $e_s(a)$, public education subsidies e_p are provided yearly from age 6 to age $\bar{s}$. Since public and private investments are perfect substitutes in the human capital production function, then if e_p is higher than the optimal total educational investments, then $e_s(a) = 0$ and there is "pure public education." Let $\widehat{e}^*(a)$ be the optimal amount of *total* expenditure on education when $e_s^*(a) > 0$. Then the optimal educational investment $e^*(a)$ is the maximum between $\widehat{e}^*(a)$ and e_p as given by,

$$e^*(a) = \max\{\widehat{e}^*(a), e_p\} \quad \text{for } 0 < a < s. \tag{18}$$

Notice that during pre-school years and after age $\bar{s}$, educational investments are only private, or $e^*(a) = e_s^*(a)$. Figure 2 portrays function $e^*(a)$ and e_p, where e_p is the horizontal line between ages $\underline{s} = 6$ and $\bar{s}$ and zero otherwise.[13] The upward sloping curves correspond to different scenarios for $\widehat{e}^*(a)$. Notice that $\widehat{e}^*(a)$ increases with age because $q(a)$ is decreasing in age. It turns out that in the calibrated model case 2 in Figure 2 holds for most countries in the sample. Case 2 illustrates a case in which private spending includes pre-school and some years after $\bar{s}$ since optimal schooling, s_2, is larger than $\bar{s}$. Optimal total spending $\widehat{e}^*(a)$ in case 2 is lower than e_p, so between ages $\underline{s} = 6$ and $\bar{s}$ we have that $e^*(a) = e_p$.

As we explain in the calibration, cross-country differences in $\bar{s}$ are quite significant, varying from as little as age 10 to age 22 (i.e., a subsidy duration of 4 to 16 years). One of the objectives of this paper is to quantify the extent to which these large differences in provision of public education play a role in explaining the international quantity-quality trade-off.

2.2.3 Schooling

The optimality condition for schooling years is given by

$$e_s^*(s) + c^{*S}(s)\,\frac{G(n^*)^{1/\sigma-1} - 1}{1/\sigma - 1} = \frac{1}{q(s)\,G(n^*)} W_s(s^*, n^*, \mathbf{E}^*) + \frac{\bar{\tau}}{G(n^*)} \text{ for } \sigma \neq 1, \text{ or} \tag{19}$$

$$e_s(s^*) + c^S(s^*)\ln(G(n^*)) = \frac{1}{q(s^*)\,G(n^*)} W_s(s^*, n^*, \mathbf{E}^*) + \frac{\bar{\tau}}{G(n^*)} \text{ for } \sigma = 1. \tag{20}$$

The marginal cost of an extra year of schooling is given by the additional private education expenditures incurred, $e_s(s)$, plus the cost of waiting one extra year at a level of student consumption $c^S(s)$, which is lower than that of a worker when $G > 1$. In this respect, credit frictions increase the marginal cost of schooling. The marginal benefits of an extra year of schooling are given by the additional lifetime income $W_s(s^*, n^*, \mathbf{E}^*)$, plus the lump-sum tax payment avoided from not working that year, $\bar{\tau}$. Both components of the marginal benefit are discounted by G. As in the case of optimal educational expenditures, this discount captures the fact that more schooling increase resources in the second stage of life when resources are less scarce. Equation (19) describes a

[13] Figure 2 borrows from Cordoba and Ripoll (2013).

quantity-quality trade-off: ceteris paribus, countries with higher fertility will have higher G, which would tend to decrease their optimal level of schooling. This prediction stems from the role of credit frictions in the model.

Turning to other determinants of schooling, notice that higher probabilities of survival increase the marginal benefit of schooling through their effect on lifetime income $W(s, n, \mathbf{E})$. Higher wages, w, increase both the marginal cost and the marginal benefit of schooling: on the cost side, higher wages increase parental transfers to children, who will in turn spend more in education and consumption. On the benefit side, higher wages increase lifetime income $W(s, n, \mathbf{E})$ in a proportional way. The net effect will depend on the relative increase of marginal cost and benefits. We discuss this in Proposition 1 below after presenting the optimal fertility choice.

2.2.4 Fertility

Assume parameters are such that the solution for fertility is interior. We check that this is the case in the numerical results. In steady state, the optimality condition for fertility is given by

$$q(F)\, b_1^{*\prime} + q(F + s^*)\, b_2^{*\prime} - W_n(s^*, n^*, \mathbf{E}^*) = \Phi_n(n^*) \frac{V(b_1^{*\prime}, b_2^{*\prime})}{\lambda_2}, \tag{21}$$

where λ_2 is the marginal utility of parental consumption at age F as given by

$$\lambda_2 = C^{-\eta}(C - \underline{C})^{\sigma} \rho e^{(r-\rho)F} \pi(F)^{\frac{\theta-\sigma}{1-\theta}} c^*(F)^{-\sigma}.$$

Expression (21) equates the marginal costs and benefits of a child. The marginal costs are the resources parents transfer to the child, b_1' and b_2', plus the time costs of raising the child, which result in lower labor supply and lifetime income as given by $-W_n(s^*, n^*, \mathbf{E}^*)$. The marginal benefit corresponds to the lifetime utility of the child $V(b_1', b_2')$, weighted by marginal altruism toward the last child, $\Phi_n(n^*)$, and normalized by λ_2, which expresses the marginal benefit in terms of parental consumption units.

According to (21) the time costs of raising children are lower for parents with lower human capital. This is one of the mechanisms that generates a steady state with larger families and lower levels of schooling, a quantity-quality trade-off. Regarding other determinants of fertility, higher probabilities of survival increase the marginal benefit of children through their positive effect on lifetime utility $V(b_1', b_2')$, and also decrease the marginal cost of children as the time spent raising them becomes a smaller fraction of lifetime income $W(s, n, \mathbf{E})$.

2.2.5 Effects of wages on schooling and fertility

Of particular interest is the effect of wages, w, on schooling and fertility choices. In the general equilibrium, wages reflect total factor productivity levels. Equations (19) and (21) indicate that higher wages increase the marginal benefits as well as the marginal costs of schooling and of children.

The net effect depends on the model's features, particularly on the presence of non-homothetic utility and public education provision as we summarize in the following proposition.

Proposition 1. Optimal fertility and schooling are independent of wages if: (i) the utility function in (2) is homothetic, e.g. $\underline{C} = 0$; and *(ii)* there is no public education: $e_p = \bar{\tau} = 0$ for all a.

Proof. See Appendix.

Intuitively, schooling and fertility choices are independent of wages when the marginal costs and benefits are proportional to w. As discussed above, higher wages increase both the costs and benefits of schooling and fertility, but if the increase is proportional, the effect of wages cancels out. To see why both requirements in Proposition 1 eliminate the effect of wages on n and s, rewrite (19) and (21) as (see Appendix for details)

$$\frac{e_s\left(s\right)}{W^*} + \frac{c^S\left(s\right)}{W^*}\frac{G(n)^{1/\sigma-1}-1}{1/\sigma-1} = \frac{1}{G(n)}\frac{1}{q\left(s\right)}\frac{W_s\left(s^*,n^*,\mathbf{E}^*\right)}{W^*}, \tag{22}$$

and

$$\begin{aligned} & q(F)\frac{b_1^*}{W^*} - \frac{W_n(s^*,n^*,\mathbf{E}^*)}{W^*} \\ = \; & \frac{\Phi_n(n^*)}{1-\Phi\left(n^*\right)}\frac{G\left(n^*\right)/\rho}{1-\eta}\frac{\Omega_4\left(s^*,n^*\right)c^*\left(0\right)/W^*+\underline{C}/W^*}{\left(\Omega_4\left(s^*,n^*\right)\right)^{\sigma}}. \end{aligned} \tag{23}$$

where $\Omega_4\left(s^*,n^*\right)$ is a function only of s and n. Absent public education, expenditure variables such as $e_s\left(s\right)$, $c^S\left(s\right)$, b_1^* and $c^*\left(0\right)$, as well as W_s^* and W_n^* are all homogeneous of degree one in W^*. As a result, ratios $e_s\left(s\right)/W^*$, $c^S\left(s\right)/W^*$, b_1^*/W^*, $c^*\left(0\right)/W^*$, W_s/W^* and W_n/W^* are all homogenous of degree zero in W^*. If $\underline{C} = 0$ then W^*, and in particular wages, w, do not enter in the two equations above determining s and n. In other words, in the pure homothetic version of the model with pure private education, "time" variables s and n, are orthogonal to "money" variables.

The pure homothetic model with pure private education is unable to account for the negative fertility-income relationship suggested by the data. Our approach to recover such relationship is to introduce the non-homothetic term $\underline{C}$. According to equation (23), term $\underline{C}$ increases the marginal benefit of having children. This is because $\underline{C}$ acts as a public good that delivers utility to any alive person beyond private consumption $C - \underline{C}$. Moreover, what matters for fertility choice is $\underline{C}/W^*$. This means that the term is large for poorer countries, countries with low human wealth, but less significant for rich countries. Thus, the incentives to have children are stronger in poor rather than in rich countries whenever $\underline{C} > 0$.

Even when $\underline{C} = 0$, fertility could also depend indirectly on w through the schooling choice in the presence of public education. To see this, notice that when there is public provision so that $e_p > 0$ and $\bar{\tau} > 0$, then the marginal benefit on the right-hand-side of (20) is not proportional to wages because $\bar{\tau}$ does not directly depend on w. In our calibration, however, this effect is not strong enough.

2.3 Closing the model

2.3.1 Demographics

Consider a steady-state with constant population growth and a stationary distribution of people by age. Let g_n be the constant growth rate of population. The steady-state density of age-a people is given by

$$\tilde{n}(a) \equiv \frac{N(a)}{N} = \frac{e^{-g_n a}\pi(a)}{\int_0^\infty e^{-g_n a}\pi(a)da}, \tag{24}$$

where $N(a)$ is the population of age a and N the total population. Since birth rates are endogenous in the model, with n children born when the parent is age F, then population growth g_n must satisfy the following relationship

$$n\pi(F) = e^{g_n F}, \tag{25}$$

where recall that parents survives to age F with probability $\pi(F)$.

2.3.2 Government

The only role of the government in this model is to provide public education subsidies in the amount of e_p per pupil up to age $\overline{s}$. The government collects lump-sum taxes from workers in order to pay for education spending, so that the government budget constraint is given by

$$\bar{\tau}\int_s^R \tilde{n}(a)\,da = e_p \int_6^{\min(s,\bar{s})} \tilde{n}(a)\,da,$$

where e_p is exogenous, $\tilde{n}(a)$ is endogenous as it depends on the fertility rate through g_n, and $\bar{\tau}$ is computed as a residual to balance the budget.

2.3.3 Production

We assume each country is a small open economy facing an exogenous interest rate r. In this respect interest rates differentials play no role in our theory, consistent with the findings of Caselli and Feyrer (2007). The production function is a standard Cobb-Douglas of the form

$$Y = K^\alpha (AH)^{1-\alpha}, \tag{26}$$

with $0 < \alpha < 1$, where Y is output, K is the capital stock, A is TFP and H is aggregate human capital. The small-open economy assumption implies that the ratio $K/(AH)$ is equalized across countries, since

$$r = \alpha\left(\frac{K}{AH}\right)^{\alpha-1}. \tag{27}$$

The wage per unit of human capital is given by

$$w = (1-\alpha)\left(\frac{K}{AH}\right)^{\alpha} A = (1-\alpha)\,(\alpha/r)^{\frac{\alpha}{1-\alpha}}\,A, \tag{28}$$

so that differences in w across countries reflect differences in TFP, A.[14]

3 CALIBRATION

In this section we calibrate the model to international data. For this purpose, specific functional forms for the altruistic function, $\Phi(n)$, and labor supply, $l(n,a)$, are required. Following Cordoba *et. al.* (2016), we use an exponential function for $\Phi(\cdot)$ of the form[15]

$$\Phi(n) = e^{-\rho F}\pi(F)^{\frac{1-\eta}{1-\theta}}\,\psi(1-e^{-\chi n}). \tag{29}$$

This altruistic weight has three components: the first is time discounting, $e^{-\rho F}$, as all children are assumed to be born when the parent is age F. The second is the survival probability to age F, $\pi(F)^{\frac{1-\eta}{1-\theta}}$. As we show below, θ is a key parameter determining the model-implied value of statistical life.[16] The last component, function $\psi(1-e^{-\chi n})$, describes how the altruistic weight depends on the number of children: ψ is the level of altruism, and parameter χ controls the degree of diminishing altruism. This last component is analogous to the exponential time discount, except that now the discount is on children and it depends on how many are born.

As for the labor supply, we consider the following simple form

$$l(n,a) = \begin{cases} 1-\lambda n & \text{if } a > F \\ 1 & \text{otherwise} \end{cases},$$

[14] In computing w in the calibration we clean the data to properly measure the disposable income of the representative individual in each country. Since TFP is computed as a residual, controlling for differences across countries in income taxes results in a more accurate productivity measure.

[15] Cordoba *et. al.* (2016) compare the exponential formulation of $\Phi(n)$ with the Barro-Becker formulation of the form

$$\Phi(n) = e^{-\rho F}\pi(F)^{\frac{1-\eta}{1-\theta}}\,\frac{1}{1-\varepsilon}n^{\frac{1}{1-\varepsilon}}.$$

They favor (29) because it exhibits a stronger degree of diminishing altruism, helping the model better match the fertility data.

[16] The reader may wonder why in the altruistic function in equation (29) the exponent for $\pi(F)$ is given by $(1-\eta)/(1-\theta)$, while the exponent for $\pi(a)$ in equation (2) is given by $(1-\sigma)/(1-\theta)$. To understand why, it is useful to consider the following monotonic transformation of V in equation (1): $W = [(1-\eta)V]^{1/(1-\eta)}$. In this case, equation (1) can be written as:

$$W = \left[C^{1-\eta} + e^{-\rho F}\pi(F)^{\frac{1-\eta}{1-\theta}}\,\psi(1-e^{-\chi n})\,(V')^{1-\eta}\right]^{\frac{1}{1-\eta}}.$$

This CES representation of lifetime utility is analogous to the that of C in (2). As shown below, deriving the value of statistical life in the model requires taking the derivative of W with respect to π. In that derivative, the power $1/(1-\eta)$ on the bracket above cancels the term $(1-\eta)$ of the exponent on $\pi(F)$. As a result, the value of statitiscal life is determined by $(1-\theta)$. Similar argument applies to C in (2).

where λn is the time cost of having n children. This time cost function implies that when $n = 0$ the cost is zero.[17]

We calibrate the model to a sample of countries with the most recent available data, typically 2013. Data availability determines a sample size of 92 countries. Table 1 presents summary statistics for the main variables of interest in our sample. We will comment on Table 1 below along with our calibration strategy. We assume some parameters are country-specific, while others are common across countries as we now describe.

3.1 Country-specific parameters

Countries differ along key exogenous dimensions, allowing us to quantitatively evaluate the role of a number of factors on schooling and fertility choices. In particular, countries differ on mortality rates, $\{\pi(a)\}_0^\infty$, public education subsidies e_p, the age until which public provision is available, $\bar{s}$, school repetition rates, $(1-d)$, and TFP.

3.1.1 Mortality

In modeling the survival probabilities we are make a compromise between computational convenience and realism. We assume the following representation for $\pi(a)$,

$$\pi(a) = \begin{cases} e^{-p_1 a} & \text{for } a \leq a_1 \\ \pi(a_1)\, e^{-p_2(a-a_1)} & \text{for } a_1 \leq a \leq a_2 \\ \pi(a_2)\, e^{-p_3(a-a_2)} & \text{for } a_2 \leq a \end{cases} .$$

Since we are interested in evaluating the role of mortality on fertility and schooling, in the equation above we introduce three separate periods to allow for different hazard (mortality) rates for young children, students and workers, and retirees. We set $a_1 = 5$, so that p_1 is the hazard rate for young adults. We also set $a_2 = 65$ so that p_2 is the hazard rate for students and workers, and p_3 is that for retirees.

We calibrate p_1, p_2, and p_3 for each country in the sample. For this purpose we use the life tables from the World Population Prospects for the period 2010-2015 and extract the survival probabilities by 5-year age intervals. Parameters p_1, p_2, and p_3 are calibrated to survival probabilities $\pi(5)$, $\pi(65)$, and to life expectancy at birth. In particular, denoting by $LE(0)$ the life expectancy at birth we have.

$$LE(0) = \frac{1-e^{-p_1 a_1}}{p_1} + e^{-p_1 a_1}\frac{1-e^{-p_2(a_2-a_1)}}{p_2} + e^{-p_1 a_1 - p_2(a_2-a_1)}\frac{1}{p}.$$

We obtain reasonable survival profiles for all countries. Figure 3 plots survival probabilities by age for selected countries. While the calibration underpredicts survival in earlier years and overpredicts

[17] As in Becker and Barro (1988), time costs here are meant to capture the costs over the lifetime of the parent.

in later years, the overall fit is reasonable.

3.1.2 Schooling

We use UNESCO data to document differences in public education provision around the world. On this margin the two key variables are $\overline{s}$ and e_p. We also allow for differences in school repetition rates across countries, $1-d$, in order to measure human capital more accurately in (4). Before explaining how we measure this variables, notice that our measure of schooling in the model, which is used in Figure 1, is school life expectancy (SLE). SLE corresponds to the total number of years of schooling a child expects to receive in each country, assuming that the probability of being enrolled in school equals the current enrollment ratio for each age. Therefore, for a child of age 6 the SLE is given by

$$SLE_6 = \sum_{i=6}^{I} \frac{\text{enrollment}_i}{\text{population}_i} \times 100,$$

where I is a theoretical upper age-limit for schooling. Table 1 reports summary statistics for SLE in our sample, which has a mean of 13.86 years. Although there has been substantial increase in school enrollment rates in poor countries, large differences in SLE still remain, with the 90th percentile at 17.8 years and the 10th percentile at 10.4 years.

We measure $\overline{s}$ as the years of free education in each country from UNESCO.[18] When years of free education is not available for a country we use UNESCO's compulsory schooling variable. In addition, for the handful of countries for which none of these two measures are available, we construct a measure of $\overline{s}$ using the duration of primary, secondary and terciary education, as well as enrollment in public education by level. Figure 4 plots SLE as a function of our measure of $\overline{s}$. The plot suggests that for most countries $\overline{s}$ is not binding since SLE is above $\overline{s}$, but that for a few mostly African countries SLE is actually below $\overline{s}$. It also shows even for the same level of $\overline{s}$, SLE varies substantially across countries.

We compute public education subsidies e_p in each country as

$$e_p = \frac{\text{government educational expenditures}}{\text{pupils enrolled in public institutions}},$$

which measures the average subsidy received by pupils enrolled in all levels of public education. Since in the model we have a representative student per country, measuring e_p this way captures the average public subsidy available to this student in each school grade. Table 1 reports summary statistics for e_p in our sample. The cross-country dispersion is substantial: while the mean is \$6,601, the standard deviation is \$7,091 (PPP). Figure 5 plots e_p against GDP per capita in log-10 scale, suggesting a strong correlation (88.8%).[19]

[18] We exclude from this measure any number of years of free pre-school, since the prevalence of free pre-school is not universal, and this measure is missing for many countries.

[19] The UNESCO measure of government educational expenditures is PPP adjusted. To the extent that the education sector is labor intensive and nontradable, this PPP adjustment should make education spending comparable across countries. This PPP adjustment does not fully capture cross-country differences in the relative price of education

Last, school repetition rates across countries are constructed from UNESCO data. Data is available for the percentage of repeaters in primary and early secondary. An average measure of repetition is constructed using the country-specific duration of primary and secondary as a fraction of the SLE. Since repetition is measured only for "early" secondary, we half the duration of secondary to compute average repetition rates. Table 1 reports summary statistics for repetition rates: while the average is only 3.7%, the standard deviation is 5.6%. Repetition rates tend to be very high in Africa, as much as 30.4% in certain countries (Central African Republic).

3.1.3 Wages and TFP

In order to compute country-specific wages we use output per capita in the data for 2013, $y^{data} = Y/N$, and we construct a model-related measure of human capital for 2013, h^{data}. In particular, from equations (26) and (27) we can write wages in any given period as

$$w = \frac{(1-\alpha)y^{data}}{h^{data}}, \tag{30}$$

where we define h^{data} to be

$$h^{data} = \Theta^{data} h(s, \mathbf{E}) \left(s^{data}/s\right)^{\gamma/\beta}. \tag{31}$$

In equation (30) we measure y^{data} from the World Development Indicators (2013, PPP). As we explain below, α is set exogenously and is common across countries. Notice that while h^{data} in (30) is the level of human capital of the workers who produced the 2013 GDP, $h(s, \mathbf{E})$ in (31) is the model-implied steady-state human capital at age s. These are not the same values because the level of schooling of the workforce in 2013 is not the same as the SLE of the current school-age children. In fact, the data confirms a gap between the Barro-Lee schooling for the adult population (as reported in the PWT for 2010) and the 2013 SLE from UNESCO. The average Barro-Lee schooling in our sample is 8.41 years, while the average 2013 SLE is 13.86 years.

In equation (31) Θ^{data} captures the average experience of workers at the 2013 age distribution, and s^{data} is the Barro-Lee schooling for the adult population in 2013. We compute Θ^{data} as the weighted average of exponential functions $e^{\nu(a-s)}$, where ν are the returns to experience as in (6). The weights are given by the population shares reported in 5-year age intervals from the World Population Prospects for 2013, up to retirement age $R = 65$.

In order to understand our strategy for computing h^{data} notice that if the current adult workers had the same schooling as current students are expected to complete, or $s^{data} = s = SLE$, then $h^{data} = \Theta^{data} h(s, \mathbf{E})$. Exponent γ/β corresponds to the elasticity of human capital with respect to schooling in (4) when expenditures are constant. In this respect γ/β is a reasonable exponent to adjust for the gap s^{data}/s in order to obtain h^{data}.

goods. However, as it is well known, relative price of education indexes are not as reliable for cross-country comparisons due to the difficulty in measuring the quality of education. We use the UNESCO PPP numbers to avoid adding noise to the data.

Last, once wages are computed, TFP for each country obtained as a residual from (28) as,

$$A = \frac{w}{(1-\alpha)\,(\alpha/r)^{\frac{\alpha}{1-\alpha}}}. \tag{32}$$

3.2 Common parameters across countries

3.2.1 Exogenous parameters

The following parameters are assumed to be common across countries and are set exogenously: the EIS, $1/\sigma$; the interest rate, r; the rate of time preference, ρ; the capital share, α; the returns to experience, ν; the childbearing age, F; and the retirement age; R.

Table 2 summarizes the values for these parameters. We set $\sigma = 1$, a common value in the growth and business cycles literatures. We set $r = 2.5\%$, a standard value for a risk free rate. We assume $r = \rho = 2.5\%$, so that consumption growth over the life cycle is determined by the survival probabilities from equations (10) and (11). A capital share of $\alpha = 0.33$ is standard. Returns to experience is set to $\nu = 2\%$ implying that wages are multiplied by a factor of 2.23 after 40 years of experience, which is consistent with estimates from Bils and Klenow (2000).

We set $F = 28$. Recall that in the model all children are born at the same time, so F is the average childbearing age. According to the United Nations' World Fertility Patterns 2015, the average childbearing age in 2010-2015 was about 27.3 years in Asia and Latin America, 28.6 in North America, and slightly above 29 in Africa, Europe and Oceania. Setting $F = 28$ is a reasonable compromise.[20]

In the case of retirement, we set $R = 65$, a value that binds mostly for rich countries in the sample. This value allows us to address the concern that the positive effects of longer life expectancy in schooling may be overstated for rich countries, since individuals there do not necessarily have a longer working life span relative to poor countries.

3.2.2 Calibrated parameters

The following parameters are also assumed to be common across countries and are calibrated to targets from the data: the elasticity of intergenerational substitution, $1/\eta$; the parameter that determines mortality risk aversion, θ; non-market consumption, $\underline{C}$; returns to scale of human capital production, γ; degree of substitution among education expenditures at different ages, β; the level of altruism, ψ; the degree of child discounting, χ; and the level parameter of the time cost of raising children, λ.

Table 3 presents the calibration results. Although all parameters affect the targets jointly, some have relatively more quantitative impact in matching certain targets as we now explain. Parameter η is calibrated to match the average fertility in the sample. As shown above in equation (23), a lower η, or a higher intergenerational substitution, results in richer countries having a lower

[20] We check that in the model's calibration as well as in all counterfactuals $F = 28$ is not binding in the sense that school life expectancy is never larger than F.

number of children. In this respect η directly influences the level of fertility in the model. We obtain $\eta = 0.339$, which implies an elasticity of intergenerational substitution of 2.95. This is consistent with the findings in Cordoba and Ripoll (2019), who found values of this elasticity significantly larger than one.

Parameter θ is calibrated to match the value of statistical life (VSL) in the United States at age F. The VSL is defined in the literature as the willingness to pay to save one life by a large pool of identical individuals. In the model the VSL corresponds to the marginal rate of substitution between survival and consumption. In other words, the value of remaining life at age F is given by

$$VSL(F) = \frac{\partial V/\partial \pi\,(F)}{\partial V/\partial c\,(F)}.$$

As we show in the Appendix, θ has a first-order effect on $VSL(F)$. In particular, as $\theta \rightarrow 1$ then $VSL(F) \rightarrow \infty$. As the value of statistical life in the United States has been estimated to be between \$4 and \$9 million (Viscusi and Aldi, 2003), then θ must be well below one. We set a target for the VSL on the conservative end of \$4 million and obtain a calibrated $\theta = 0.535$.

As discussed above, the non-homotheticity introduced by $\underline{C}$ guarantees that wages affect fertility. We calibrate $\underline{C}$ to match the income elasticity of fertility. We set this target to -0.38, which is the value computed by Jones and Tertilt (2008) using historic Census data for the US. We obtain a calibrated $\underline{C} = 4,900$.

Regarding the human capital production function, we calibrate γ to match the average private educational expenditures as a fraction of GDP among OECD countries in the sample, and β to match the world mean of schooling.[21] As reported by the National Center of Education Statistics, in OECD countries private education spending was on average 0.9% of GDP in 2014. We obtain $\gamma = 0.335$ and $\beta = 0.172$. These parameters are similar to those obtained in Cordoba and Ripoll (2013), and consistent with the large human capital literature discussed therein.

The altruistic function, $\Phi(n) = \psi(1 - e^{-\chi n})$, plays a key role in determining the amount parents transfer to children for consumption and education expenditures during childhood. We then calibrate ψ to match the goods costs of raising a child (consumption and education) as a percentage of mean family lifetime income in the US. Using information in Lino (2012) on the costs of raising children from the US Department of Agriculture we set this target to 16.44%. We compute this target using information from families in the low-income bracket, whose upper-bound corresponds to the median family income in the US. Since our model includes college costs, we adjust Lino's (2012) cost computation by adding costs of attending public colleges. We obtain $\psi = 0.475$.

Regarding χ, which determines the degree of child discounting, we select as a target the standard deviation of fertility. Since χ drives marginal altruism, it plays a role in determining fertility. In this respect both η and χ are determinants of first and second moments of the fertility distribution.

[21] We use the average private educational expenditures as a fraction of GDP among OECD countries because the United States is somewhat atypical among rich countries in this dimension. While the OECD average is 0.9%, in the United States the corresponding number is 2%.

We obtain $\chi = 1.466$. Last, we calibrate λ, the level parameter of the time cost of children function to match the time cost of raising a child as a percentage of lifetime income in the US. We set this target to 17%, a conservative estimate. Using the same strategy as in Cordoba and Ripoll (2016, 2018), we compute the time costs of raising children to be between 60 and 75% of the total costs, which results in time costs being between 17 and 32% of family lifetime income in the US.[22]

3.3 Model's fit

Figure 6 portrays the fertility and schooling predictions of the model relative to the data. The overall model performance is quite good: the correlation between the data and the model is 75.8% for fertility and 77.9% for schooling. Table 4 reports untargeted moments to evaluate the model's performance along other dimensions. The model almost exactly replicates the standard deviation of schooling in the data. Regarding the main focus of this paper, the quantity-quality trade-off, the model explains roughly 80% of the negative correlation between schooling and fertility in the data, an substantial fraction. Figure 7 shows the quantity-quality trade-off in the data (as in Figure 1) and in the model. As shown, the model gets the downward slope well, but misses the flatter part at high schooling levels. As seen in Table 4, the model also does a relatively good job in replicating the maximum and minimum of both schooling and fertility in the data. Consistent with Figure 7, the model predicts a maximum school life expectancy of 18.3 years, while it is 20.4 in the data.

4 GLOBAL COUNTERFACTUALS

Table 5 reports the results of counterfactual exercises that equate one parameter at a time across countries. We create an artificial "rich country" whose TFP, educational policies and the fraction of school non-repeaters ($\overline{s}$, e_p and d) correspond to the 90th percentile of the sample, and whose mortality rates (p_1, p_2, p_3) are at the 10th percentile. The counterfactuals equate each parameter to its value in this artificial rich country. We also conduct counterfactuals for parameter groups: all mortality rates (p_1, p_2, and p_3) and the two education policies ($\overline{s}$ and e_p).[23]

Table 5 shows that equating TFP to the 90th percentile would reduce the standard deviation of schooling, fertility and income per capita substantially: in 54%, 72% and 76% respectively. Large TFP effects are typical in the cross-country literature with exogenous TFP. A major role for cross-country TFP differences in explaining schooling and fertility is also found in Manuelli and Seshadri (2009). But one of the main insights in Table 5 is that public education variables are also

[22] As explained in Cordoba and Ripoll (2016) this range depends on whether only active time taking care of children is taken into account, or also passive time (time spent in the presence of children surpervising, but not directly engaged). The range of costs also depend on whether hours are priced at the nanny's wage or the median wage. These different ways of computing time costs result in a range of 17 to 32% of family lifetime income in the US.

[23] It is common in the literature to equate values to US levels, but the US is somewhat of an outlier among the rich on both mortality and education variables. Creating the artificial rich country for the counterfactuals avoids values in outlier countries to influence the changes in the means.

quantitatively important.[24] As shown in the second panel, equating both the number of years of education subsidy provision ($\overline{s}$) and spending per pupil (e_p) to the 90th percentile in the sample reduces the standard deviation of schooling by 47%, that of fertility by 62% and that of income per capita by 59%. These results are novel in the literature.[25]

It is interesting to compare the separate role of $\overline{s}$ and e_p. On this regard, Table 5 shows that e_p matters relatively more in explaining the dispersion of fertility, while both e_p and $\overline{s}$ are important in accounting for the dispersion of schooling. Equating e_p to the 90th percentile in the sample, which is $17,179 per pupil per year, reduces the dispersion of fertility by 50%, but the reduction is only 14% when $\overline{s}$ is equalized across countries. There are two main mechanisms behind this result: the time cost of raising children, and the complementarity of educational expenditures across ages. Everything else equal, higher e_p increases human capital and the time cost of raising children, reducing fertility. In addition, since e_p is provided starting at age six, parents respond to higher e_p by increasing educational investments both in pre-school years and after year $\overline{s}$, when the subsidy ends. Notice how e_p also matters relatively more than $\overline{s}$ in explaining the dispersion of per capita income. This suggests that to the extent that cross-country differences in e_p capture differences in educational quality though teacher salaries and instructional materials, these differences are important in understanding the dispersion of per capita income. These findings speak to the literature that underscores the importance of improving educational quality in developing countries (Schoellman, 2012).

In contrast, on the extensive margin, adding more $\overline{s}$ does not decrease fertility significantly. What this finding suggests for poorer countries where e_p is low, is that just extending this low subsidy for additional years would not be an effective tool to reduce fertility and increase schooling. This "low quality" subsidization does not raise human capital and the time cost of raising children enough to generate a large drop in fertility and higher private education investments on each child. The lesson for developing countries is that providing higher e_p or higher education "quality", even for a limited number of years, is a more effective way of reducing fertility levels.

As mentioned, $\overline{s}$ has a relatively larger effect on the dispersion of schooling. This effect is completely driven by poorer countries. When $\overline{s}$ is equated everywhere to the 90th percentile, which corresponds to 13 years of subsidies, nothing changes for those countries that already had schooling beyond 13 years. As seen in Figure 4, this is the case for many countries. Under the counterfactual, $\overline{s}$ becomes binding for poorer countries, where schooling increases up to 13 years. Therefore, what occurs in poor countries with low e_p is that when the subsidy is extended for more years, students do go to school longer to claim the subsidy, but family size does not decrease by much.

Turning now to mortality rates, Table 5 shows that reducing either under-5 (p_1) or retiree (p_3) mortality rates to their respective 10th percentile values has small effects on schooling, fertility and income per capita. While demographers have emphasized the role of child mortality in fertility

[24] Results reported in Table 5 omit a few countries from the sample (seven). For these countries the government's budget constraint is violated in the counterfactual in which e_p is equated to the 90th percentile. We omit these countreis from all counterfactuals for comparison purposes.

[25] Although one may argue that educational policies are endogenous, this would also be true for TFP. In this analysis we follow the tradition of the development accounting literature and use the counterfactuals as a tool to guide the further study of variables that are quantitatively important in understanding cross-country differences.

choices, by 2013 child mortality rates had dropped substantially in developing countries. There is more quantitative action from the reduction in adult mortality rates (p_2). We find that decreasing mortality rates for those between ages 5 and 65 in all countries results in a drop of 24% of the schooling dispersion, and a drop of 20% of the fertility dispersion. In this case, a higher probability of surviving during working years increase the incentives to remain in school longer and invest in education, an effect that is stronger for countries with higher mortality rates. Higher human capital raises the time cost of having children and results in lower fertility rates.

5 THE ROLE OF EDUCATION SUBSIDIES

In this section we present additional exercises to further explore the role of public education subsidies. In our model intergenerational financial frictions play a central role in determining the investments parents made in the education of their children, as well as the number of children the parent has. Parents cannot pay for these educational investments by either borrowing against their children's future income, or by enforcing financial obligations on them. In the absence of contracts between parents and children that could improve the allocation of resources across generations, public education subsidies serve as an alternative mechanism to secure educational investments on children.

Public education subsidies are provided in virtually all countries, and they can only be received by children as long as they attend school. By design, these subsidies fundamentally alter parental's incentives. In particular, due to the complementary of investments in education across ages, public education subsidies affect the incentives parents have in investing on their children in years where no subsidies are provided. This effect reinforces the quantity-quality trade-off in the model. We now turn to examine how exactly these mechanisms play out by performing additional counterfactual exercises.

5.1 Eliminating public education

We first analyze the consequences of eliminating public education subsidies in the model. The bottom row in Table 5 summarizes the quantitative results. Eliminating public education subsidies in all countries increases average fertility by 14.8%, decreases average schooling by 0.1%, and decreases average GDP per capita by 22.3%. The effect of eliminating public education subsidies is heterogeneous across countries: while in poorer countries fertility increases and schooling decreases, the opposite occurs in richer countries. However, human capital and income per capita decrease everywhere. As we discuss next, the decrease in GDP per capita can be traced to a decrease in total educational investments and a decrease in human capital. Parents do not spend enough in the education of their children to fully compensate for the elimination of the public education subsidy. Eliminating public education increases the standard deviation of all schooling (35.6%), fertility (36.5%) and GDP per capita (27.5%).

More importantly, the effects of eliminating public education vary dramatically with the level of TFP. To show this, Figure 8 displays schooling (left panel) and fertility (right panel) choices as a function of TFP, for two scenarios: one with public education subsidies and one without. To construct Figure 8 we use the calibrated parameters of the model (Tables 2 and 3), as well as the country-specific parameters for the artificial "rich" country we constructed for the counterfactuals in Table 5.

Turning first to left panel of Figure 8, notice that in the presence of public education subsidies (offered at the level of the 90th percentile in the sample), schooling is non-monotonic in TFP: at very low TFP levels schooling is highest, then it drops sharply and it finally increases monotonically with TFP. At very low levels of TFP, it is optimal to send children to school to receive high education subsidies ($17,179 per pupil for 13 years). Not attending school and giving up the education subsidy is too costly relative to working at very low wages. Financial constraints are also at work here: at very low TFP levels, there would not be much of consumption jump between the schooling and working years, lowering the cost of staying at a highly subsidized school longer. As TFP levels start increasing, the trade-off between staying at school with high education subsides and cutting school shorter to start working changes in favor of the latter: schooling rapidly drops to about 16 years. Finally, the last portion of the schooling graph is more standard, with schooling increasing in TFP.

The left panel of Figure 8 suggests that in the presence of financial constraints, schooling choice interacts with both education spending and TFP in intricate ways. It also implies that caution is required when interpreting counterfactuals that involve equalization of TFP levels across countries. In contrast, in the absence of public education subsidies, schooling is monotonically increasing and concave in TFP: with no public education subsidies schooling is about 6.5 years at low TFP levels. Increasing schooling to 16 years requires multiplying TFP by a factor of 3. As can be seen on the right-panel in Figure 8, the quantity-quality trade-off explains this pattern: in the absence of public education subsidies, at low TFP levels fertility is as high as 6.3 children. Low-wage parents have lower time costs of raising children, have many of them, and invest less in the education of each.

One of the main messages of Figure 8 is that at low levels of TFP, public education has a potentially powerful effect increasing schooling and decreasing fertility. In contrast, at high levels of TFP fertility could be lower and schooling could be higher in the absence of public education subsidies. Is public education then undesirable at high TFP levels? As we show in Figure 9, this is actually not the case.

Figure 9 displays human capital upon finishing school (left panel) and steady-state GDP per capita (right panel) as a function of TFP, for the same two scenarios: one with public education subsidies and one without. As seen on the left panel, in the absence of public education subsidies human capital is lower than when subsidies are provided, except for very high TFP levels. In the calibrated model the 90th percentile of TFP is around 2, and the maximum TFP is 2.6, so there are no countries in our sample for the region in which human capital is higher with no public subsidies. What Figure 9 indicates is that in the absence of public education subsidies, even altruistic parents invest less in the human capital of their children. Take for instance the artificial "rich" country with TFP of 2, e_p of $17,179 per pupil per year, and $\overline{s}$ of 13 years: eliminating public education in that country results in a 10% decrease in human capital. Moreover, as the right panel in Figure 9 shows, GDP per capita in this country decreases by about 20% when public education is eliminated. This clearly illustrates the role of public education subsidies in the model and it interacts with intergenerational financial frictions: in low-TFP countries, public education promotes a strong quantity-quality trade-off, giving incentives to parents to reduce family size and complement educational investments during years in which no subsidy is provided. In this sense, public education in low-TFP countries acts as a sort of "big push" mechanism out of a high-fertility and low-human-capital trap. In high-TFP countries where family size is already low, public education is still useful as a mechanism to allocate resources across generations by making returns to complementary parental educational investments higher.

The right panel of Figure 9 also displays an interesting, non-monotonic behavior of GDP per capita in the absence of public education subsidies. At levels of TFP higher than 2.2, GDP per capita falls with TFP, a puzzling behavior. As we explore in more detail in Section 6, this drop in GDP per capita can be trace to a "demographic drag." As can be seen in the right panel of Figure 8, fertility at levels of TFP higher than 2 falls below replacement level, more so in the absence of public education subsidies. In the steady state of the model, this results in a drop in the fraction of the working-age population. Interestingly, this demographic drag does not occur in the presence of public education subsidies.

5.2 Local counterfactuals on education subsidies per pupil

So far we have analyzed the role of public education subsidies in the context of global counterfactuals (Table 5), or by eliminating the subsidies (Figures 8 and 9). Given the non-monotonic behavior of schooling at different TFP levels, it is natural to ask what would happen for local changes in educational policies. Although our results compare steady states, local changes to educational policies are still more informative for practical policy matters.

In this section we implement a 10% increase in subsidies per pupil e_p in each country. This local counterfactual simulates and increase in the "quality" of the public subsidy. The effects are presented in Figures 10 and 11. The left panel of Figure 10 shows the schooling semielasticity in each country versus their GDP per capita relative to the US. In most poor countries, 10% more spending per pupil increases schooling anywhere from less than a year to even up to 5.5 years. For middle-income countries the schooling semi-elasticity could be positive or negative, but for the richer countries, schooling decreases by up to one year. To understand these results, notice from the right panel of Figure 10 that for almost all countries human capital at age s increases, as much as 6.5% for poorer countries, and 1% for richer ones. This suggests that even though in response to higher education subsidies per pupil schooling falls in richer countries, human capital does increase. What explains this result?

To answer this question we turn to Figure 11, where the left panel shows the elasticity of total private education spending in response to a 10% increase in e_p in every country. Except for a few countries, this elasticity is positive, confirming the notion that when the size of the public education subsidy increases, parents complement these investments increasing total private spending in education by up to 10% in poorer countries and 2% in richer ones. We verify that most of these increases occur when children are ages 0 to 6, when returns to educational investments are the largest. This provides an answer to why in response to a 10% increase in e_p schooling drops in richer countries, while human capital increases: since private parental investments also increase, students in richer countries can better smooth consumption by ending school earlier. Financial constraints play a role here: resources are relatively scarce while children attend school, while access to credit markets is expanded once the child starts working.

As the right panel in Figure 11 shows, in almost all countries the semielasticity of fertility in response to a 10% increase in e_p is negative: fertility drops by up to 0.34 children in poorer countries. Altogether, Figure 11 illustrate how the quantity-quality trade-off is operative in response to a local increase in public education per pupil.

5.3 Local counterfactuals on duration of education subsidies

We now consider the effects of increasing the duration of the public education subsidy $\overline{s}$ by 10% in every country. To better understand this local counterfactual, Figure 12 displays some relevant properties of the calibrated model. The left panel shows the school life expectancy predicted by the model against $\overline{s}$ in the data. The figure shows that in most countries $\overline{s}$ is non-binding. Only for a handful of countries, about seven, $\overline{s}$ binds, while only for one country (Central African Republic) school life expectancy is less than $\overline{s}$. The right panel in Figure 12 portrays the optimal total education spending at age $\overline{s}$ against e_p in the data. As most points lie below the 45-degree line, then the model predicts that on the year the public education subsidy ends ($\overline{s}$) parents invest less than the public subsidy amount e_p. Therefore, Figure 12 suggests that a local counterfactual that increases $\overline{s}$ by 10%, which on average is about one extra year, automatically corresponds to offering students an additional year of school with higher educational resources than what the parents would privately invest in their children.

Figures 13 and 14 show the effects of the 10% increase in $\overline{s}$ in schooling (left panel Figure 13), human capital (right panel Figure 13), total private education spending (left panel Figure 14), and fertility (right panel Figure 14). First, as shown in Figure 13, for those poorer countries in which $\overline{s}$ binds, schooling increases by the full extra year additional provision, but it drops slightly for most other countries. More interestingly, as shown on the left panel in Figure 14, parents in almost all countries spend less on private education, on average 10% less. As predicted from Figure 12, extending the duration of public subsidies for one more year crowds out the parental private education spending that would have taken place that year. The only countries in which private education spending increases are those in which $\overline{s}$ is binding in the calibrated model.

The extent to which parents invest less than the educational subsidy can be seen on the right panel of Figure 13: human capital under this local counterfactual increases in almost every country, because students can get an extra year of schooling with higher investments than what parents would have spent.

6 FURTHER IMPLICATIONS AND ROBUSTNESS

In this section we discuss additional implications of our model, in particular the role of demographics in the cross-country steady-state distribution of GDP per capita. We also provide two robustness checks for our results: one is an out-of-sample prediction for schooling and fertility in the United States in 1990, and the other is the effect of introducing proportional taxes instead of lump-sum taxes in richer countries.

6.1 The role of demographics

This section uses the calibrated model to shed light on the separate role of human capital and demographics in explaining the cross-country GDP per capita distribution in the steady state of the model. Figure 15 displays steady-state GDP per capita relative to GDP per capita in 2013.[26] The figure shows in almost all countries, steady-state GDP per capita is larger than income per capita in 2013, more so for poorer countries, but not enough to close the income gaps with richer countries. Notice that while the model predicts that in the absence of further TFP growth, US steady-state GDP per capita is 10% higher than GDP per capita in 2013, in most northern European countries it is 50% higher. Countries like Hong Kong and South Korea, among few others, are notable exceptions in that steady state GDP per capita will be lower than in 2013. As we now turn to discuss, in these countries the "demographic drag" from the age-distribution of the population dominates the gains from higher human capital.

To better understand the determinants of steady-state GDP per capita in the model, y^*, use equations (26) and (28) to write

$$y^* = \frac{Y}{N} = A(\alpha/r)^{\frac{\alpha}{1-\alpha}} \frac{H}{N}, \tag{33}$$

where

$$\frac{H}{N} = \int_{s^*}^{R} h(a)\tilde{n}(a)da = \underbrace{h\left(s^*, \mathbf{E}^*\right)}_{\text{educational factors}} \underbrace{\int_{s^*}^{R} e^{\nu(a-s^*)} \frac{e^{-g_n a}\pi(a)}{\int_0^\infty e^{-g_n a}\pi(a)da} da}_{\text{demographic factors}}. \tag{34}$$

As explained before, since the model is calibrated to 2013 data, TFP (A) is computed to exactly match GDP per capita in 2013. Recall that r is a risk-free interest rate common across countries. Therefore equation (33) implies that if there is no TFP growth, differences between GDP per capita in 2013 and in the steady state of the model are driven by differences in human capital per person H/N, which is shown in detail in equation (34).

Steady-state human capital per person H/N is in turn determined by two main factors: educational factors (schooling and education spending) on the one hand, and demographic factors on the other hand. The first term on the right-hand-side of (34), $h\left(s^*, \mathbf{E}^*\right)$, corresponds to steady-state human capital upon finishing school at age s^*. Since the model is calibrated to match school life

[26] Since the calibrated bechmark model does not exactly match schooling and fertility for each single country, in Figure 15 we use the model's calibrated parameters and equations to compute the steady-state GDP per capita if the model exactly matched schooling and fertility in the data.

expectancy in 2013, s^* corresponds to the steady state schooling expected from the school enrollments observed from children in all age groups in 2013. These children will eventually be the adults of the model's steady state. Similarly, the vector of education spending at different ages, $\mathbf{E}^*$, includes the observed public education subsidies in 2013 and the model-implied private education spending. Notice that human capital $h\left(s^*, \mathbf{E}^*\right)$ is different than the human capital of the workers who produced the 2013 GDP. Therefore term $h\left(s^*, \mathbf{E}^*\right)$ captures the educational (schooling and spending) factors that would drive a wedge between steady-state GDP per capita and observed GDP per capita in 2013.

The last term on the right-hand-side of (34) corresponds to the "demographic" factors. This integral computes the returns to human capital experience, but weights each age by the share of people of that age in the population, $\tilde{n}(a)$. Although effectively a human capital term, it captures demographic factors through the population age profile. Recall that g_n represents of the stationary population growth rate as given in (25), which is determined by steady-state fertility n^*. Demographic factors drive the other wedge between steady-state GDP per capita and observed GDP per capita in 2013 because even though the model is calibrated to replicate the cross-country distribution of fertility in 2013, what matters for the steady-state GDP per capita is the stationary age-distribution implied by n^*. Notice that although we call the last integral in (34) the "demographic" factors, schooling s^* does affect the lower limit of the integral.

Figures 16 and 17 illustrate the separate effects of educational and demographic factors on steady-state GDP per capita in the model. Figure 16 isolates the role of educational factors by assuming that the age-structure of the population remains as in 2013, so that the wedge between steady-state and 2013 GDP per capita is solely driven by s^*, $\mathbf{E}^*$ and $h\left(s^*, \mathbf{E}^*\right)$. As shown in the figure, educational factors drive steady-state GDP per capita up in all countries, relative to GDP in 2013. This increase is relatively higher in some of the poorest countries, but it is sizeable across the whole distribution. Even in richer countries such as Finland, Denmark and Belgium, the ratio of steady-state GDP per capita to 2013 income per capita is about 1.7. Notably, for the US this ratio is much smaller, at 1.2.

Figure 17 isolates the role of demographic factors by assuming that the schooling of the workers remains as the Barro and Lee schooling in 2013, so that the wedge between steady-state and 2013 GDP per capita is solely driven by the age structure of the population in the steady state, term $\tilde{n}(a)$ for each a. As shown in the figure, demographic factors alone drag down steady-state GDP per capita in the majority of the countries, in particular for all countries with GDP per capita above \$13,000 in 2013. In fact, more than 50% of all countries in the sample already had fertility rates below replacement levels in 2013. Figure 17 suggests that if retirement is at age 65 and if mortality rates remain as in 2013, the already low levels of fertility rates in more than 50% of the countries will create a steady-state demographic drag through the shrinking of the working-age population. As Figure 15 shows, this demographic drag is ultimately counteracted in almost all countries by the higher rates of human capital accumulation implied by the higher levels of school life expectancy observed everywhere in the world in 2013.

6.2 Out-of-sample predictions

The analysis in this paper focuses on the model's steady state. Although the model is complicated to compute transitions, in this section we follow Manuelli and Seshadri (2009) and we provide an out-of-sample exercise as a robustness check. In particular, we use the calibrated model to predict school life expectancy and fertility in the US in 1900, which we assume to be another steady state. We can provide out-of-sample predictions only for the US because we do not have cross-country public education data (e_p and $\overline{s}$) in 1900. For this out-of-sample exercise we use the same model parameters from Tables 2 and 3, but we need to find the following US-specific parameters in 1900: age-specific mortality rates for children, adults and elder (p_1, p_2 and p_3); duration and amount of public education subsides ($\overline{s}$ and e_p); and TFP.

Table 6 reports the out-of-sample exercise for the US. As mentioned, the objective of the exercise is to use the model to predict fertility and school life expectancy in 1900 and compare these against the data. We obtain fertility data directly from Haines and Steckel (2001), as births per woman in 1900, which was 3.56 children for white women.[27] Since there is no data on school life expectancy in 1900, we compute it using available information on school enrollment by age. From the Historic Statistics of The United States (US Department of Commerce) we know that the total population enrolled in school as a fraction of 5 to 17 years old was 78.3% in 1900. Assuming the enrollment rate is the same for each of these years, school life expectancy is estimated to be 9.4 years in 1900.

In order to compute public expenditures per pupil in the US in 1900, we use data from the National Center of Education Statistics (1993), which reports public expenditures per pupil in elementary and secondary in current dollars. We convert this sum to 2013 dollars using the CPI estimate from the Federal Reserve Bank of Minneapolis. As shown in Table 6, in the US e_p was \$567 in 1900 (2013 dollars). For the duration of public education subsidies in 1900, we use $\overline{s} = 8$. This figure represents well the fact that in 1900 compulsory schooling laws were in effect in 34 states (four of them in the South), and in 30 states these laws required attendance until age 14.

We use the Life Tables for the United States Social Security Area 1900-2010 (Social Security Administration) to compute the mortality rate parameters p_1, p_2 and p_3 in 1900. For this purpose we use as targets the survival probabilities at ages 5 and 65, $\pi(5)$ and $\pi(65)$, and life expectancy at birth. As reported in Table 6, the survival rate at age 5 in 1900 was 80%, while that at age 65 was 39.2%. Life expectancy at birth was 47.7 years (average for men and women).

Last, to compute TFP in the US in 1900, we follow the same procedure as in the 2013 calibration. TFP in 1900 is computed to exactly match GDP per capita that year, which according to Maddison (2010) was \$7,118 (in 2013 dollars). For this purpose we also need to use equation (31), which requires information on Θ^{data} and s^{data}. In this case Θ^{data} captures the average experience of workers at the 1900 age distribution, and s^{data} is the schooling of the adult population in 1900. We obtain both the 1900 age distribution of workers and the median years of schooling completed by people above 25 years of age from the National Center of Education Statistics (1993).

As can be seen in Table 6, the model predicts a school life expectancy and fertility rates for the

[27] Fertility for black women in 1900 was higher, at 5.61 births per woman. Since our model does not incorporate race we compare the model's predictions against fertility for white women.

US that are very close to the 1900 data, providing a out-of-sample robustness check for the model.

6.3 Income taxes in richer countries

As mentioned, we model taxes in a lump-sum fashion because for most countries in the sample distortionary income taxes are a small fraction of total tax revenue collection. In fact, income tax collection is relatively larger only for richer countries. For instance, even among OECD countries, where income taxes on individuals are on average 23% of total tax revenue, there is substantial variation: it is as high as 52.9% in Denmark, 40.3% in Australia, and 38.7% in the United States; and it is as low as 14.2% in Hungary, and 9.7% in Chile (OECD, 2019). In fact, value added taxes (VAT) are a large fraction of tax revenues even in some OECD countries: 41.6% in Chile, 30.2% in New Zealand, and 27.7% in Estonia. Outside of the OECD, income taxes as a fraction of tax revenue are even lower: in 2017 they were on average 15% in Africa and 10% in Latin America (OECD, 2019). In contrast, VAT as a percent of tax revenue is high: 29% in Africa and 28% in Latin America.

Since income taxes tend to be relatively more important for richer countries, we now show that our counterfactual exercises would change little in richer countries if education was financed with proportional income taxes rather than lump-sum taxes in those economies. For this purpose we select all OECD countries with GDP per capita above $30,000 in 2015, which correspond to 18 rich countries. We now report the results of the global counterfactuals on TFP, $\overline{s}$ and e_p for both the lump-sum and the proportional income tax model for these 18 rich countries. For consistency, we equate TFP, $\overline{s}$ and e_p in all rich countries to the same values of the artificial "rich country" we used in the global counterfactual exercises reported in Section 3. We focus on how mean school life expectancy and mean fertility change relative to the benchmark. We find that although there are some quantitative differences, the average percentage changes are small and similar for the two different types of taxation in richer countries.

First, for the TFP counterfactual we find that equating TFP under the lump-sum tax model results in an increase of mean schooling of 1.22%, and a decrease of fertility of -3.01% among the 18 richer countries. The corresponding figures for the proportional income tax model are 0.27% and -1.46%. These results are consistent with the findings reported above, with schooling increasing in TFP and fertility decreasing in wages.

Second, for the $\overline{s}$ counterfactual we find that equating $\overline{s}$ to 13 years under the lump-sum tax model results in a decrease of mean schooling of -0.51%, and an increase of fertility of 0.35% among the 18 richer countries. The corresponding figures for the proportional income tax model are -2.02% and 1.98%. Consistent with the results explained above, extending the duration of the education subsidy in richer countries results in less schooling years and more fertility, although higher human capital.

Last, for the e_p counterfactual we find that equating e_p to $17,179 per pupil per year under the lump-sum tax model results in a decrease of mean schooling of -0.34%, and a decrease of fertility of -0.51% among rich countries. The corresponding figures for the proportional income tax model are -0.58% and -0.19%. These effects are overall small and have the same sign across both types of taxes for rich countries.

7 CONCLUDING COMMENTS

The provision of public education subsidies is universal. For decades, international organizations have been promoting the expansion of these subsidies in developing countries as a tool to increase human capital and lower fertility rates. The model we propose in this paper allows us to use a micro-founded framework to evaluate how cross-country differences in educational subsidies can explain the international quantity-quality trade-off observed in 2013. In addition, our analysis also incorporates differences in TFP and age-dependent mortality rate across countries, providing a more complete decomposition of the underlying forces behind schooling, fertility and income differences.

Some takeaways from the analysis can be highlighted. First, although differences in TFP (or wages per unit of human capital) play a quantitatively important role in explaining school life expectancy and fertility, the effect of public education subsidies is also sizeable. In particular, equating both the number of years of education subsidy provision and spending per pupil to the 90th percentile in the sample reduces the standard deviation of schooling by 47%, that of fertility by 62% and that of income per capita by 59%.

Second, we find that the design of public education subsidies matters, in particular the extensive (duration) and intensive margins (spending per pupil). Specifically, a significant decrease in fertility can be achieved by increasing spending per pupil rather than the duration of the subsidy. For instance, extending the length of public education by increasing compulsory schooling years, without increasing the educational resources per pupil may not generate significant drops in fertility rates in poor countries. These increases in length do result in higher schooling years, but at low levels of spending per pupil, the human capital gains are so minimal that there are no significant incentives for fertility rates to drop. Even if for a limited number of years, raising public educational resources per pupil in poorer countries could unleash a scenario where parents respond to the complementaries in human capital investments across ages by having less children and investing more in each of them. These insights speak to a literature that underscores the importance of improving educational quality in developing countries (Schoellman, 2012). While as summarized in Lee and Barro (2001) there is debate on how to achieve this, rethinking the characteristics of how human capital is produced at schools in the developing world is of first-order importance in understanding the international quantity-quality trade-off.

The third takeaway of the analysis is that eliminating public education subsidies results in an increase in average fertility, a decrease in human capital and income per capita, and an increase in the dispersion of schooling, fertility and income. While in poorer countries fertility increases and schooling decreases, the opposite occurs in richer countries. However, human capital and income per capita decrease everywhere. In the model parents cannot borrow against the future income of children, nor they can impose debt obligations on them. Due to this constraint to intergenerational transfers, when public education subsidies are eliminated, parents do not spend enough in the education of their children to fully compensate for the missing subsidy. From this perspective, public education plays the important role of ensuring investments in the human capital of children take place. This view of the role of public educational subsidies is novel, and our model captures its role in explaining the international quantity-quality trade-off.

Last, our exercise provides insights into the effect of demographic changes on GDP per capita the long run. We calibrate our model to the GDP per capita in 2013, but that output was produced by workers with lower schooling than the one we expect future workers to have given the observed enrollments in 2013 (school life expectancy). In addition, the age distribution of the workers in 2013 is on average younger than the one in the steady state of the model, since fertility rates in 2013 were lower than in the past. We find that while higher schooling in 2013 will result in higher steady-state GDP per capita everywhere, most countries will experience a demographic drag, which tends to lower steady-state GDP per capita. This demographic drag is reflected in the lower shares of the working-age population. Although on net steady-state GDP per capita will be larger than the GDP measured in 2013, the demographic drag is a reality for all rich countries as well as most middle-income and even poorer countries where fertility was already close-to or below replacement rates in 2013.

References

[1] Angeles, L. (2010). "Demographic transitions: analyzing the effects of mortality on fertility," *Journal of Population Economics*, 23, 99-120.

[2] Barro, R. and Becker, G. (1989). "Fertility choice in a model of economic growth," *Econometrica,* 57, 481-501.

[3] Becker, G. and Barro, R. (1988). "A reformulation of the economic theory of fertility," *Quarterly Journal of Economics,* 103, 1-25.

[4] Becker, G. and Lewis, G. (1973). "On the interaction between the quantity and quality of children," *Journal of Political Economy,* 81, S279–S288.

[5] Becker, G., Murphy, K. and Tamura, R. (1990). "Human capital, fertility, and economic growth," *Journal of Political Economy*, 98, S12-S37.

[6] Becker, G., Philipson, T. and Soares, R. (2005), "The Quantity and Quality of Life and the Evolution of World Inequality", *American Economic Review*, 95, 277-291.

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

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

[9] Breierova, L. and Duflo, E. (2004). "The impact of education on fertility and child mortality: Do fathers really matter less than mothers?" *NBER Working Paper* 10513.

[10] Castro-Martin, T. and Juarez, F. (1995). "The impact of women's education on fertility in Latin America: Searching for Explanations," *International Family Planning Perspectives*, 21, 52-57.

[11] Canning, D., Günther, I., Linnemayr, S. and Bloom, D. (2013). "Fertility choice, mortality expectations, and interdependent preferences: An empirical analysis," *European Economic Review*, 63, 273-289.

[12] Caselli, F. and Feyrer, J. (2007). "The marginal product of capital," *Quarterly Journal of Economics*, 122, 537-568.

[13] Cordoba, J. C., Liu, X. and Ripoll, M. (2016). "Endogenous fertility, social mobility and long-run inequality," *Journal of Monetary Economics,* 77, 103-124.

[14] Cordoba, J. C. and Ripoll, M. (2019). "The elasticity of intergenerational substitution, parental altruism and fertility choice," *Review of Economic Studies*, 86, 1935-1972.

[15] Cordoba, J. C. and Ripoll, M. (2017). "Risk aversion and the value of life," *Review of Economic Studies*, 84, 1472-1509.

[16] Cordoba, J. C. and Ripoll, M. (2016). "Intergenerational transfers and the fertility-income relationship," *Economic Journal*, 126, 949-977.

[17] Cordoba, J. C. and Ripoll, M. (2013). "What explains schooling differences across countries?" *Journal of Monetary Economics,* 60, 184-202.

[18] Cordoba, J. C. and Ripoll, M. (2011). "A contribution to the economic theory of fertility," Manuscript. Iowa State University Economics Working Papers No. 11011

[19] Cunha, F., Heckman, J., Lochner, L. and Masterov, D. (2006). "Interpreting the Evidence on Life Cycle Skill Formation," In Hanushek, E., and Welch, F. (Eds), *Handbook of the Economics of Education*, Volume 1, Chapter 12. Elsevier B.V.

[20] De la Croix, D. and Doepke, M. (2004). "Public versus private education when differential fertility matters," *Journal of Development Economics*, 73, 607-629.

[21] De la Croix, D. and Licandro, O. (1999). "Life expectancy and endogenous growth," *Economic Letters*, 65, 255-263.

[22] Doepke, M. (2004). "Accounting for Fertility Decline During the Transition to Growth," *Journal of Economic Growth,* 9, 347-383.

[23] Federal Reserve Bank of Minneapolis. *Consumer Price Index, 1800-*.

[24] Ferreira, P. Monge-Naranjo, A. and Torres, L. (2019). "From Malthus to Malthus." Manuscript.

[25] Galor, O., and Weil, D. (1996). "The Gender Gap, Fertility and Growth," *American Economic Review*, 86, 374-387.

[26] Galor, O., and Weil, D. (2000). "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *American Economic Review*, 90, 806-828.

[27] Haines, M. and Steckel, R. (eds) (2001). *A Population History of North America*. New York: Cambridge University Press.

[28] Hazan, M. (2012). "Life Expectancy and Schooling: New Insights from Cross-Country Data," *Journal of Population Economics*, 25, 1237-1248.

[29] Kirk, D. and Pillet, B. (1998). "Fertility levels, trends, and differentials in Sub-Saharan Africa in the 1980s and 1990s," *Studies in Family Planning*, 29, 1-22.

[30] Lee, J, and Barro, R. (2001). "Schooling quality in a cross-section of countries," *Economica*, 68, 465-488.

[31] Lino, M. (2012). "Expenditures on Children by Families, 2011", *USDA Miscellaneous Publication Number 1528-2011.*

[32] Maddison, A. (2010). Historical Statistics of the World Economy: 1-2008 AD. Groningen Growth and Development Centre. University of Groningen.

[33] Manuelli, R. and Seshadri, A. (2009). "Explaining international fertility differences," *Quarterly Journal of Economics,* 124, 771-807.

[34] National Center for Education Statistics. (1993). *120 Years of American Education: A Statistical Portrait*. United States Department of Education.

[35] OECD. (2019). *Revenue Statistics 2019: Tax Revenue Trends in the OECD.*

[36] OECD. (2019). *Revenue Statistics in Africa 2019.*

[37] Pradham, E. and Canning, D. (2015). "The effect of schooling on teenage fertility: Evidence from the 1994 education reform in Ethiopia," Working Paper Series, Program of the Global Demography of Aging, Harvard University.

[38] Rangazas, P. (2000). "Schooling and economic growth: A King-Rebelo experiment with human capital," *Journal of Monetary Economics*, 46, 397-416.

[39] Schoellman, T. (2012). "Education quality and development accounting," *Review of Economic Studies*, 79, 388-417.

[40] Schoonbroodt, A. and Tertilt, M. (2014). "Property rights and efficiency in OLG models with endogenous fertility," *Journal of Economic Theory*, 150, 551-582.

[41] Schoonbroodt, A. and Tertilt, M. (2010). "Who owns children and does it matter?" *NBER Working Paper # 15663*.

[42] United Nations (2015). *World Fertility Patterns 2015.* Department of Economic and Social Affairs, Population Division. - Data Booklet.

[43] United States Department of Commerce. *Historical Statistics of the United States: Colonial Times to 1970*. United States Bureau of the Census.

[44] Viscusi, W.K. and Aldy, J. (2003). "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World", *Journal of Risk and Uncertainty*, 27, 5-76.

[45] Wilson, N. (2015). "Child Mortality Risk and Fertility: Evidence from Prevention of Mother-to-Child Transmission of HIV," *Journal of Development Economics*, 116, 74-88.

[46] World Bank Group (2011). *Learning for All: Investing is people's knowledge and skills to promote development.* World Bank Education Strategy 2020, Washington, DC.

APPENDIX FOR ONLINE PUBLICATION

1 Model solution

1.1 Individual's problem

The problem of the representative agent is described recursively as:

$$V\left(b_1, b_2\right) = \max_{[c(a)]_{t=0}^{\infty}, [e_s(a)]_{t=0}^{s}, b_1', b_2', s, n} \frac{1}{1-\eta} C^{1-\eta} + \Phi\left(n\right) V\left(b_1', b_2'\right) \tag{1}$$

where

$$C = \left[\rho \int_0^{\infty} e^{-\rho a} \pi\left(a\right)^{\frac{1-\sigma}{1-\theta}} c\left(a\right)^{1-\sigma} da\right]^{\frac{1}{1-\sigma}} + \underline{C}, \text{ and} \tag{2}$$

$$\Phi(n) = e^{-\rho F} \pi\left(F\right)^{\frac{1-\eta}{1-\theta}} \psi\left(1 - e^{-\chi n}\right).$$

The maximization is subject to the following constraints:

$$b_1 \geq \int_0^{s} \left(c\left(a\right) + e_s\left(a\right)\right) q\left(a\right) da, \tag{3}$$

$$q(s) b_2 + W(s, n, \mathbf{E}) \geq \bar{\tau} \int_s^{R} q\left(a\right) da + \int_s^{\infty} c\left(a\right) q\left(a\right) da + q\left(F\right) n b_1' + q\left(F\right) q(s') n b_2', \tag{4}$$

$$\begin{aligned}
\left(\int_0^{s} \left(d\left(e_p\left(a\right) + e_s\left(a\right)\right) / p_E\right)^{\beta} da\right)^{\gamma/\beta} &\geq h\left(s, \mathbf{E}\right), \\
l\left(n, a\right) &= \begin{cases} 1 & \text{if } a \leq F \\ l\left(n\right) & \text{if } a > F \end{cases}, \text{ and} \\
b_2' &\geq 0 \text{ and } e_s(a) \geq 0 \text{ for } a \in [0, s].
\end{aligned}$$

where

$$W\left(s, n, \mathbf{E}\right) = w h\left(s, \mathbf{E}\right) \int_s^{R} e^{\nu(a-s)} l(n, a) q(a) da. \tag{5}$$

Prices and survival probabilities satisfy:

$$q(a) = e^{-ra}\pi(a) \text{ and } \tag{6}$$

$$\pi(a) = \begin{cases} e^{-p_1 a} & \text{for } a \leq a_c \\ \pi(a_c)\, e^{-p_2(a-a_c)} & \text{for } a_c < a \leq a_s \\ \pi(a_s)\, e^{-p_3(a-a_s)} & \text{for } a > a_s \end{cases}. \tag{7}$$

The associated Lagrangian can be written as:

$$\begin{gathered} V(b_1, b_2) = \tfrac{1}{1-\eta} C^{1-\eta} + \Phi(n) V(b_1', b_2') \\ +\lambda_1 \left[b_1 - \int_0^s (c(a) + e_s(a))\, q(a)\, da \right] \\ +\lambda_2 \left[q(s) b_2 + W(s, n, \mathbf{E}) - \int_s^\infty c(a)\, q(a)\, da - q(F)\, n b_1' - q(F) q(s')\, n b_2' - \bar{\tau} \int_s^R q(a)\, da \right] \\ +\lambda_3 \left[\left(\int_0^s (d(e_p(a) + e_s(a))/p_E)^\beta\, da \right)^{\gamma/\beta} - h(s, \mathbf{E}) \right] + \lambda_4 e_s(s) + \lambda_5 b_2'. \end{gathered}$$

The choice variables are $[c(a)]_{t=0}^{\infty}, b_1', b_2', [e_s(a)]_{t=0}^{s}, s, h(s, \mathbf{E})$, and $n \in [0, \overline{n}]$. Use (*) to denote optimal solutions. Let E^* be the present value of private expenditures in education defined as:

$$E^* \equiv \int_0^{s^*} e_s^*(a) q(a) da. \tag{8}$$

1.2 Optimal consumption

First order conditions with respect to $c(a)$ can be written as

$$\lambda_1 q(a) = C^{-\eta} (C - \underline{C})^\sigma \rho e^{-\rho a} \pi(a)^{\frac{1-\sigma}{1-\theta}} c^*(a)^{-\sigma} \text{ for } a \leq s^*, \text{ and} \tag{9}$$

$$\lambda_2 q(a) = C^{-\eta} (C - \underline{C})^\sigma \rho e^{-\rho a} \pi(a)^{\frac{1-\sigma}{1-\theta}} c^*(a)^{-\sigma} \text{ for } s^* \geq a. \tag{10}$$

Using (6), these equations become:

$$\begin{aligned} c^*(a) &= C^{-\frac{\eta}{\sigma}} (C - \underline{C}) \rho^{\frac{1}{\sigma}} \lambda_1^{-\frac{1}{\sigma}} e^{(r-\rho)\frac{a}{\sigma}} \pi(a)^{\frac{1}{\sigma}\frac{\theta-\sigma}{1-\theta}} \quad \text{for } a \leq s^* \\ c^*(a) &= C^{-\frac{\eta}{\sigma}} (C - \underline{C}) \rho^{\frac{1}{\sigma}} \lambda_2^{-\frac{1}{\sigma}} e^{(r-\rho)\frac{a}{\sigma}} \pi(a)^{\frac{1}{\sigma}\frac{\theta-\sigma}{1-\theta}} \quad \text{for } s^* \geq a. \end{aligned} \tag{11}$$

Let $c^S(s^*)$ and $c^W(s^*)$ denote consumption at time s^* as a student and as a worker respectively. Dividing (9) by (10) it follows that:

$$\frac{c^W(s^*)}{c^S(s^*)} = \left(\frac{\lambda_1}{\lambda_2} \right)^{\frac{1}{\sigma}}.$$

Define

$$G \equiv \frac{\lambda_1}{\lambda_2}. \tag{12}$$

Then

$$c^W(s^*) = c^S(s^*)\, G^{\frac{1}{\sigma}}. \tag{13}$$

Use (11), (12), and (13) to obtain:

$$c^*(a) = e^{\frac{r-\rho}{\sigma}a}\pi(a)^{\frac{1}{\sigma}\frac{\theta-\sigma}{1-\theta}}\, c^*(0) \quad \text{for } a \leq s^* \text{ and} \tag{14}$$

$$c^*(a) = e^{\frac{(r-\rho)a}{\sigma}}\pi(a)^{\frac{1}{\sigma}\frac{\theta-\sigma}{1-\theta}}\, G^{\frac{1}{\sigma}} c^*(0) \quad \text{for } a \geq s^*. \tag{15}$$

To solve for $c^*(0)$, substitute (14) into (3) to obtain:

$$c^*(0) = \frac{b_1^* - E^*}{\int_0^{s^*} e^{-\vartheta a}\pi(a)^{\frac{\theta}{\sigma}\frac{1-\sigma}{1-\theta}}\, da} \text{ where } \vartheta \equiv r - \frac{r-\rho}{\sigma}. \tag{16}$$

Substituting this result into (14) and (15):

$$c^*(a) = \frac{e^{\frac{r-\rho}{\sigma}a}\pi(a)^{\frac{1}{\sigma}\frac{\theta-\sigma}{1-\theta}}}{\int_0^{s^*} e^{-\vartheta a}\pi(a)^{\frac{\theta}{\sigma}\frac{1-\sigma}{1-\theta}}\, da}\left[b_1^* - E^*\right] \quad \text{for } a \leq s^* \text{ and} \tag{17}$$

$$c^*(a) = \frac{e^{\frac{(r-\rho)a}{\sigma}}\pi(a)^{\frac{1}{\sigma}\frac{\theta-\sigma}{1-\theta}}\, G^{\frac{1}{\sigma}}}{\int_0^{s^*} e^{-\vartheta a}\pi(a)^{\frac{\theta}{\sigma}\frac{1-\sigma}{1-\theta}}\, da}\left[b_1^* - E^*\right] \quad \text{for } a \geq s^*. \tag{18}$$

λ_1 and λ_2 can be solved in terms of $c^*(0)$, using (9) and (12), as

$$\lambda_1 = c^*(0)^{-\sigma} C^{-\eta} (C - \underline{C})^{\sigma} \rho \text{ and} \tag{19}$$

$$\lambda_2 = c^*(0)^{-\sigma} C^{-\eta} (C - \underline{C})^{\sigma} \rho / G. \tag{20}$$

1.3 Optimal transfers

First order conditions with respect to b_1' and b_2' are given by

$$\lambda_2 q(F) n^* = \Phi(n^*) V_1\left(b_1^{*\prime}, b_2^{*\prime}\right),$$

$$\lambda_2 q(F) q(s^*) n^* = \Phi(n^*) V_2\left(b_1^{*\prime}, b_2^{*\prime}\right) + \lambda_5,$$

while the corresponding envelope conditions are

$$V_1\left(b_1,b_2\right)=\lambda_1 \text{ and } V_2\left(b_1,b_2\right)=\lambda_2 q(s^*).$$

Then the optimality conditions for b_1' and b_2' can then be expressed as:

$$\lambda_2^{parent}q(F)n^*=\Phi(n^*)\lambda_1^{child}, \text{ and}$$

$$\lambda_2^{parent}q(F)q(s^*)n^*>\Phi(n^*)\lambda_2^{child}q(s^*),$$

where the latter has been written assuming $b_2^{*\prime}=0$, which we later verify. At steady state they become, using (6), (7), (12) and (13),

$$\left(\frac{c^W\left(s^*\right)}{c^S\left(s^*\right)}\right)^\sigma=G=G(n^*)\equiv e^{-rF}\pi\left(F\right)\frac{n^*}{\Phi(n^*)} \tag{21}$$

and

$$e^{-rF}\pi\left(F\right)\frac{n^*}{\Phi(n^*)}=G\left(n^*\right)>\frac{\pi\left(F\right)\pi\left(s^*\right)}{\pi\left(F\right)\pi(s^*)}=1.$$

If $p_1\geq p_2$, i.e., child mortality is larger than adult mortality, as is the case in the data, then a sufficient condition for the transfer constraint to bind is $G(n^*)>1$. In what follows we assume that parameters are such that the transfer constraint binds so that $b_2^{*\prime}=0$. We confirm that in all our calibrations, $G(n^*)>1$ for all countries in our sample.

To solve for $b_1^*=b_1^{*\prime}$, substitute (15), (6) and $b_2^*=0$ into (4) to obtain:

$$W^*\equiv W(s^*,n^*,\mathbf{E}^*)=\bar{\tau}\int_{s^*}^R e^{-ra}\pi\left(a\right)da+c^*\left(0\right)G(n^*)^{\frac{1}{\sigma}}\int_{s^*}^\infty e^{-\vartheta a}\pi\left(a\right)^{\frac{\theta}{\sigma}\frac{1-\sigma}{1-\theta}}da+q\left(F\right)n^*b_1^*$$

Using (16) to substitute for $c^*\left(0\right)$ and solving for b_1^* it transpires that:

$$b_1^*=\frac{W^*+G(n^*)^{\frac{1}{\sigma}}E^*\Omega_1(s^*)-\bar{\tau}\int_{s^*}^R e^{-ra}\pi\left(a\right)da}{\Omega_1(s^*)G(n^*)^{\frac{1}{\sigma}}+q\left(F\right)n^*}, \tag{22}$$

where

$$\Omega_1(s^*)\equiv\frac{\int_{s^*}^\infty e^{-\vartheta a}\pi\left(a\right)^{\frac{\theta}{\sigma}\frac{1-\sigma}{1-\theta}}da}{\int_0^{s^*} e^{-\vartheta a}\pi\left(a\right)^{\frac{\theta}{\sigma}\frac{1-\sigma}{1-\theta}}da}.$$

1.4 Optimal human capital, schooling and school expenditures

1.4.1 Human capital

First order condition with respect to $h(s,E)$ gives

$$\frac{\lambda_3}{\lambda_2} = \frac{W(s^*, n^*, \mathbf{E}^*)}{h\left(s^*, \mathbf{E}^*\right)} = \int_{s^*}^{R} we^{\nu(a-s^*)} q(a) l(n^*, a) da$$

Dividing $\frac{\lambda_1}{\lambda_2} = G(n^*)$ by $\frac{\lambda_3}{\lambda_2}$ to obtain:

$$\frac{\lambda_3}{\lambda_1} = \frac{W^*}{G(n^*)h^*} = \frac{1}{G(n^*)} \int_{s^*}^{R} we^{\nu(a-s^*)} q(a) l(n^*, a) da, \tag{23}$$

where $h^* \equiv h(s^*, E^*)$.

1.4.2 School expenditures

Now, the first order condition with respect to $e_s(a)$ is

$$\lambda_3 \frac{\partial h(s^*, \mathbf{E}^*)}{\partial e_s(a)} + \lambda_4 = \lambda_1 q\left(a\right).$$

When the solution is interior, $\lambda_4 = 0$, this expression reduces to, using (23):

$$q\left(a\right) = \frac{1}{G(n^*)} \int_{s^*}^{R} w \frac{\partial h(s^*, \mathbf{E}^*)}{\partial e_s(a)} e^{\nu(a-s^*)} q(a) l(n^*, a) da,$$

or

$$q\left(a\right) = \frac{W^*}{G(n^*)} p_E^{-\beta} \gamma d^\beta h^{*-\frac{\beta}{\gamma}} \hat{e}^*\left(a\right)^{\beta-1}, \tag{24}$$

where $\hat{e}^*\left(a\right)$ is the solution for $e^*\left(a\right) = e_s^*(a) + e_p(a)$ if $e_s^*(a) > 0$. This interior solution can be written as:

$$\hat{e}^*\left(a\right) = \hat{e}^*\left(0\right) q(a)^{-\frac{1}{1-\beta}}. \tag{25}$$

with

$$\hat{e}^*\left(0\right) = \left(\gamma d^\beta h^{*-\frac{\beta}{\gamma}} p_E^{-\beta} W^*/G(n^*)\right)^{\frac{1}{1-\beta}}. \tag{26}$$

Let $e^*\left(a\right)$ denotes the optimal solution for $e(a) = e_s(a) + e_p(a)$. Since $e_p(0) = 0$ then initial expenditures satisfy:

$$e^*\left(0\right) = \hat{e}^*\left(0\right). \tag{27}$$

The full solution for $e^*(a)$, allowing for corners, satisfy

$$e^*(a) = \begin{cases} e^*(0)\, q(a)^{-\frac{1}{1-\beta}} \text{ if } e_s(a) > 0 \\ e_p \text{ if } a \leq s^* \text{ and } e_s(a) = 0 \end{cases} \tag{28}$$

Figure 2 illustrates three possible solutions for $e^*(a)$. Case 1 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 post-public subsidy investments, since optimal schooling s_2 is larger than $\overline{s}$, but no private spending in the interval $[\underline{s}, \overline{s}]$. Finally in Case 3, optimal schooling is $s_3 > \overline{s}$ but now there is also some private spending in the interval $[\underline{s}, \overline{s}]$. In the calibration, we set $\underline{s}$ to be 6.

To describe more precisely the solution for $e^*(a)$, let $\widehat{s}$ be implicitly defined by the equation $\widehat{e}(\widehat{s}) = e_p$. Intuitively, $\widehat{s}$ is the age at which the individual stops relying fully in public education and start using some private funds. Using (25), (6) and (7), it follows that:

$$\widehat{s} = \begin{cases} \frac{1}{p_2+r}\left[(1-\beta)\ln\left(\frac{e_p}{\hat{e}^*(0)}\right) - p_1 a_c + p_2 a_c\right] & \text{if } e_p \geq \hat{e}^*(0) \\ 0 & \text{if } e_p \leq \hat{e}^*(0) \end{cases} \tag{29}$$

Now, it could happen that $\widehat{s} < 6$ or $\widehat{s} > s$, cases in which $\widehat{s}$ does not really represents the time at which full public education ends. An precise age for this to happen is defined by:

$$s_p^* \equiv \min\left\{s^*, \overline{s}, \max\left[\underline{s}, \widehat{s}\right]\right\}. \tag{30}$$

We now can characterize $e^*(a)$ more precisely as follows:

$$e^*(a) = \begin{cases} \widehat{e}^*(0)\, q(a)^{-\frac{1}{1-\beta}} & \text{for } a \leq \min(s^*, \underline{s}) \\ e_p & \text{for } \min(s^*, \underline{s}) \leq a \leq s_p^* \\ \widehat{e}^*(0)\, q(a)^{-\frac{1}{1-\beta}} & \text{for } s_p^* \leq a \leq s^* \end{cases} \tag{31}$$

where $\hat{e}^*(0)$ is given by (26). Private educational expense can then be obtained as:

$$e_s^*(a) = \begin{cases} e^*(a) & \text{for } a \leq \min(s^*, \underline{s}) \\ 0 & \text{for } \min(s^*, \underline{s}) \leq a \leq s_p^* \\ e^*(a) - e_p & \text{for } s_p^* \leq a \leq \min(s^*, \overline{s}) \\ e^*(a) & \text{for } \min(s^*, \overline{s}) \leq a \leq s^* \end{cases} \tag{32}$$

Plugging these results into (8) one obtains:

$$E^* = \hat{e}^*(0)\, \Omega_2\left(s^*, s_p^*\right) - e_p \int_{s_p^*}^{\min(s^*, \overline{s})} q(a)da \tag{33}$$

where

$$\Omega_2\left(s^*, s_p^*\right) = \left[\int_0^{\min(s^*,6)} q(a)^{-\frac{\beta}{1-\beta}} da + \int_{s_p^*}^{s^*} q(a)^{-\frac{\beta}{1-\beta}} da\right].$$

1.4.3 Human capital

Human capital at age s^*, $h^* = h\left(s^*, \mathbf{E}^*\right)$, can be written as

$$\begin{aligned} h^* &= \left(\int_0^{\underline{s}} \left(d\frac{e^*(a)}{p_E}\right)^{\beta} da + \int_{s_p}^{s^*} \left(d\frac{e^*(a)}{p_E}\right)^{\beta} da + \int_{\underline{s}}^{s_p} \left(d\frac{e_p}{p_E}\right)^{\beta} da\right)^{\frac{\gamma}{\beta}} \\ &= \left(\frac{de^*(0)}{p_E}\right)^{\gamma} \left(\Omega_3(s^*, \frac{e_p}{e^*(0)})\right)^{\gamma/\beta} \end{aligned} \tag{34}$$

where

$$\Omega_3(s^*, \frac{e_p}{e^*(0)}) \equiv \int_0^{\underline{s}} q(a)^{-\frac{\beta}{1-\beta}} da + \int_{s_p}^{s^*} q(a)^{-\frac{\beta}{1-\beta}} da + \left(\frac{e_p}{e^*(0)}\right)^{\beta} (s_p - \underline{s}).$$

Notice that:

$$h_s(s^*, \mathbf{E}^*) = \frac{\gamma}{\beta} h^{*1-\frac{\beta}{\gamma}} \left(\frac{de^*(s^*)}{p_E}\right)^{\beta}. \tag{35}$$

Plug (34) into (26),

$$\begin{aligned} e^*(0) &= \left(\gamma d^{\beta} p_E^{-\beta} W^*/G(n^*)\right)^{\frac{1}{1-\beta}} h^{*-\frac{\beta}{\gamma}\frac{1}{1-\beta}} \\ &= \left(\gamma d^{\beta} p_E^{-\beta} W^*/G(n^*)\right)^{\frac{1}{1-\beta}} \left(\frac{de^*(0)}{p_E}\right)^{-\frac{\beta}{1-\beta}} \left(\Omega_3\left(s^*, \frac{e_p}{e^*(0)}\right)\right)^{-\frac{1}{1-\beta}}. \end{aligned}$$

Solving for $e^*(0)$,

$$e^*(0) = \gamma W^* / \left(G(n^*)\Omega_3\left(s^*, e_p/e^*(0)\right)\right). \tag{36}$$

1.4.4 Schooling

The first order condition for s is given by

$$\begin{aligned} & C^{-\eta}\frac{1}{1-\sigma}\left(C-\underline{C}\right)^{\sigma} \rho e^{-\rho s^*} \pi\left(s^*\right)^{\frac{1-\sigma}{1-\theta}} \left[c^S\left(s^*\right)^{1-\sigma} - c^W\left(s^*\right)^{1-\sigma}\right] \\ = {} & \lambda_1\left(c^S\left(s^*\right) + e_s^*\left(s^*\right)\right) q\left(s^*\right) - \lambda_2\left[W_s\left(s^*, n^*, E^*\right) + c^W\left(s^*\right) q\left(s^*\right) + \bar{\tau} q\left(s^*\right)\right] \end{aligned} \tag{37}$$

where, according to (5),

$$\begin{aligned} W_s\left(s^*, n^*, \mathbf{E}^*\right) &= wh\left(s^*, \mathbf{E}^*\right)\left[-q(s^*)l\left(n^*, s^*\right) + \left(\frac{h_s\left(s^*, \mathbf{E}^*\right)}{h\left(s^*, \mathbf{E}^*\right)} - \nu\right)\left(\int_{s^*}^{R} e^{\nu(a-s^*)} q(a)l(n^*, a)da\right)\right] \\ &= W^*\left[\frac{h_s\left(s^*, \mathbf{E}^*\right)}{h\left(s^*, \mathbf{E}^*\right)} - \nu - \frac{q(s^*)l\left(n^*, s^*\right)}{\int_{s^*}^{R} e^{\nu(a-s^*)} q(a)l(n^*, a)da}\right]. \end{aligned} \tag{38}$$

According to (35) and (34):

$$\begin{aligned} \frac{h_s\left(s^*, \mathbf{E}^*\right)}{h\left(s^*, \mathbf{E}^*\right)} &= \frac{\gamma}{\beta} h\left(s^*, \mathbf{E}^*\right)^{-\frac{\beta}{\gamma}} \left(\frac{de^*(s^*)}{p_E}\right)^{\beta} \\ &= \frac{\gamma}{\beta}\left[\left(\frac{de^*\left(0\right)}{p_E}\right)^{\gamma}\left(\Omega_3\left(s^*, \frac{e_p}{e^*(0)}\right)\right)^{\gamma/\beta}\right]^{-\frac{\beta}{\gamma}} \left(\frac{de^*(s^*)}{p_E}\right)^{\beta} \\ &= \frac{\gamma}{\beta}\frac{1}{\Omega_3\left(s^*, \frac{e_p}{e^*(0)}\right)}\left(\frac{e^*(s^*)}{e^*\left(0\right)}\right)^{\beta}. \end{aligned} \tag{39}$$

In the case that $e^*(s^*)$ is interior, then $\frac{e^*(s^*)}{e^*(0)} = q\left(s^*\right)^{-\frac{1}{1-\beta}}$ according to (31).

Plugging (9) and (10) into (37), and combining terms results in:

$$\begin{aligned} \frac{1}{1-\sigma}\left[c^S\left(s^*\right)^{1-\sigma} - c^W\left(s^*\right)^{1-\sigma}\right] &= c^S\left(s^*\right)^{-\sigma}\left(c^S\left(s^*\right) + e_s^*\left(s^*\right)\right) \\ &\quad -c^W\left(s^*\right)^{-\sigma}\left[\frac{1}{q\left(s^*\right)}W_s\left(s^*, n^*, \mathbf{E}^*\right) + c^W\left(s^*\right) + \bar{\tau}\right] \end{aligned}$$

or

$$e_s^*\left(s^*\right) + c^S\left(s^*\right)\frac{G(n^*)^{1/\sigma-1} - 1}{1/\sigma - 1} = \frac{1}{G(n^*)}\frac{1}{q\left(s^*\right)}W_s\left(s^*, n^*, \mathbf{E}^*\right) + \frac{\bar{\tau}}{G(n^*)}. \tag{40}$$

When $\sigma \to 1$, this expression becomes

$$e_s^*\left(s^*\right) + c^S\left(s^*\right)\ln G(n^*) = \frac{1}{G(n^*)}\frac{1}{q\left(s^*\right)}W_s\left(s^*, n^*, \mathbf{E}^*\right) + \frac{\bar{\tau}}{G(n^*)}.$$

Lemma 1. Consider a pure private educational system. In particular, suppose $e_p = 0$ and $\bar{\tau} = 0$. Then, (40) is an equation in two unknowns: s^* and n^*. In particular, (40) is indepedent of w.

Proof. In that case, the first order condition with respect to schooling, Equation (40), simplifies to:

$$\frac{e_s^*\left(s^*\right)}{W^*} + \frac{c^S\left(s^*\right)}{W^*}\frac{G(n^*)^{1/\sigma-1} - 1}{1/\sigma - 1} = \frac{1}{G(n^*)}\frac{1}{q\left(s^*\right)}\frac{W_s(s^*, n^*, \mathbf{E}^*)}{W^*}. \tag{41}$$

We next show that all ratios in this equation depend only on s^* and n^*, none of them depend on w. Setting $e_p = 0$, the following equations follow from (31), (32), (33) and (36):

$$\frac{e_s^*(s^*)}{W^*} = \frac{\gamma}{q(s^*)^{\frac{1}{1-\beta}} G(n^*)\Omega_3(s^*,0)} \text{ and} \tag{42}$$

$$\frac{E^*}{W^*} = \frac{\gamma\Omega_2(s^*,6)}{G(n^*)\Omega_3(s^*,0)}. \tag{43}$$

Similarly, setting $\overline{\tau} = 0$, the following equations follow from (22) and (17):

$$\frac{b_1^*}{W^*} = \frac{1}{\Omega_1(s^*)G(n^*)^{\frac{1}{\sigma}} + q(F)n^*}\left[1 + G(n^*)^{\frac{1}{\sigma}}\Omega_1(s^*)\frac{E^*}{W^*}\right], \tag{44}$$

$$\frac{c^S(s^*)}{W^*} = \frac{e^{\frac{r-\rho}{\sigma}s^*}\pi(s^*)^{\frac{1}{\sigma}\frac{\theta-\sigma}{1-\theta}}}{\int_0^{s^*} e^{-\vartheta a}\pi(a)^{\frac{\theta}{\sigma}\frac{1-\sigma}{1-\theta}}da}\left(\frac{b_1^*}{W^*} - \frac{E^*}{W^*}\right). \tag{45}$$

According to (42), $\frac{e_s^*(s^*)}{W^*}$ only depends on s^* and n^*. Same result is obtained for $\frac{c^S(s^*)}{W^*}$ by substituing (43) and (44) into (45). In other words, the left hand side of (41) only depends on s and n. As for the right hand side of (41), according to (38) and (39),

$$\frac{W_s(s^*,n^*,\mathbf{E}^*)}{W^*} = \frac{\gamma}{\beta}\frac{q(s^*)^{-\frac{\beta}{1-\beta}}}{\Omega_3(s^*,0)} - \nu - \frac{q(s^*)l(n^*,s^*)}{\left(\int_{s^*}^R e^{\nu(a-s^*)}q(a)l(n^*,a)da\right)}$$

which depends only on s^* and n^*.

1.5 Fertility

First order condition with respect to fertility is:

$$q(F)b_1^{*\prime} + q(F+s^*)b_2^{*\prime} - W_n(s^*,n^*,\mathbf{E}^*) = \frac{\partial\Phi(n^*)}{\partial n}\frac{V(b_1^{*\prime},b_2^{*\prime})}{\lambda_2} \tag{46}$$

where

$$\begin{aligned} W_n(s^*,n^*,\mathbf{E}^*) &= wh(s^*,\mathbf{E}^*)\int_{s^*}^R e^{\nu(a-s^*)}l_n(n^*,a)q(a)da \\ &= W(s^*,n^*,\mathbf{E}^*)\frac{\int_{s^*}^R e^{\nu(a-s^*)}l_n(n^*,a)q(a)da}{\int_{s^*}^R e^{\nu(a-s^*)}l(n^*,a)q(a)da} \end{aligned} \tag{47}$$

The value function at steady state can be solved, from (1), as

$$V^* = \frac{\frac{1}{1-\eta}C^{1-\eta}}{1-\Phi\left(n^*\right)}, \tag{48}$$

while the term C can be solved, using (2), (14) and (15), as

$$\begin{aligned} C &= \left[\rho\int_0^\infty e^{-\rho a}\pi\left(a\right)^{\frac{1-\sigma}{1-\theta}}c^*\left(a\right)^{1-\sigma}da\right]^{\frac{1}{1-\sigma}} + \underline{C} \\ &= c^*\left(0\right)\Omega_4\left(s^*,n^*\right) + \underline{C}. \end{aligned} \tag{49}$$

where

$$\Omega_4\left(s^*,n^*\right) = \rho^{\frac{1}{1-\sigma}}\left[\int_0^{s^*} e^{-\vartheta a}\pi\left(a\right)^{\frac{\theta}{\sigma}\frac{1-\sigma}{1-\theta}}da + G(n^*)^{\frac{1-\sigma}{\sigma}}\int_{s^*}^\infty e^{-\vartheta a}\pi\left(a\right)^{\frac{\theta}{\sigma}\frac{1-\sigma}{1-\theta}}da\right]^{\frac{1}{1-\sigma}}.$$

Using $b_1^* = b_1^{*\prime}$, $b_2^{*\prime} = 0$, (48) and (49), (46) can be written as:

$$\begin{aligned} q(F)b_1^* - W_n(s^*,n^*,\mathbf{E}^*) &= \frac{\Phi_n(n^*)}{1-\Phi\left(n^*\right)}\frac{1}{1-\eta}\frac{C^{1-\eta}}{c^*\left(0\right)^{-\sigma}C^{-\eta}\left(C-\underline{C}\right)^{\sigma}\rho/G\left(n^*\right)} \\ &= \frac{\Phi_n(n^*)}{1-\Phi\left(n^*\right)}\frac{G\left(n^*\right)/\rho}{1-\eta}\frac{C}{\left(\left(C-\underline{C}\right)/c^*\left(0\right)\right)^{\sigma}} \\ &= \frac{\Phi_n(n^*)}{1-\Phi\left(n^*\right)}\frac{G\left(n^*\right)/\rho}{1-\eta}\frac{c^*\left(0\right)\Omega_4\left(s^*,n^*\right)+\underline{C}}{\left(\Omega_4\left(s^*,n^*\right)\right)^{\sigma}} \end{aligned} \tag{50}$$

Lemma 2. Consider a pure private educational system. In particular, suppose $e_p = 0$ and $\bar{\tau} = 0$. Furthermore suppose $\underline{C} = 0$. Then (50) is an equation in two unknowns: s^* and n^*. In particular, (50) is independent of w.

Proof. Equation (50) can be written as

$$\begin{aligned} &q(F)\frac{b_1^*}{W^*} - \frac{W_n(s^*,n^*,\mathbf{E}^*)}{W^*} \\ = &\frac{\Phi_n(n^*)}{1-\Phi\left(n^*\right)}\frac{G\left(n^*\right)/\rho}{1-\eta}\frac{\Omega_4\left(s^*,n^*\right)c^*\left(0\right)/W^*+\underline{C}/W^*}{\left(\Omega_4\left(s^*,n^*\right)\right)^{\sigma}}. \end{aligned} \tag{51}$$

According to (44) and (47), the left hand side of (51) only depends on s^* and n^*. According to (43), (44) and (17), $\frac{c^*(0)}{W^*}$ only depends on s^* and n^*. Therefore, the right hand side of (51) only depends on s^* and n^* if $\underline{C} = 0$.

Proposition 1. Optimal fertility and schooling are independent of wages if: (i) the utility function in (49) is homothetic, e.g. $\underline{C} = 0$; and *(ii)* there is no public education: $e_p = \bar{\tau} = 0$ for all a.

Proof. Follows from Lemma 1 and Lemma 2. Under the stated conditions, Equations (41) and (50) are two equation in two unknowns: s^* and n^*. Wages are not part of the two equations.

Notice that, according to (51), the marginal benefit of a child increases with $\frac{\underline{C}}{W^*}$. This means that a positive $\underline{C}$ increases the marginal benefit of children proportional more in poor countries where W is smaller.

1.6 Government's budget constraint

The revenue of government from every individual's taxes is $\bar{\tau}\int_{s^*}^{R}\tilde{n}\left(a\right)da$, where

$$\tilde{n}(a)=\frac{e^{-g_n a}\pi(a)}{\int_0^\infty e^{-g_n a}\pi(a)da}.$$

and the per capita government expenditure is $e_p\int_{\underline{s}}^{\min(s^*,\bar{s})}e^{-g_n a}\pi(a)da$. The government's budget constraint requires the lump-sum taxes $\bar{\tau}$ annually imposed on households satisfies

$$\bar{\tau}=\frac{\int_{\underline{s}}^{\min(s^*,\bar{s})}e_p\tilde{n}\left(a\right)da}{\int_{s^*}^{R}\tilde{n}\left(a\right)da}=\frac{e_p\int_{\underline{s}}^{\min(s^*,\bar{s})}e^{-g_n a}\pi(a)da}{\int_{s^*}^{R}e^{-g_n a}\pi(a)da}. \tag{52}$$

1.7 Steady state wage rate and human capital

Assume $Y=K^{\alpha}\left(AH\right)^{1-\alpha}$, where $K=kN$, $H=hN$. Then

$$y=A^{1-\alpha}k^{\alpha}h^{1-\alpha}=\frac{Y}{N}$$

Pre-tax wage per unit of human capital is

$$w=\frac{\partial Y}{\partial H}=\left(1-\alpha\right)A^{1-\alpha}K^{\alpha}H^{-\alpha}=\left(1-\alpha\right)A^{1-\alpha}k^{\alpha}h^{-\alpha}=\left(1-\alpha\right)\frac{y}{h}.$$

The steady state after tax wage is calculated according to

$$w=\left(1-\alpha\right)\frac{y^{data}}{h^{data}} \tag{53}$$

$$h^{data} = \Theta^{data} h(s^*, \mathbf{E}^*) \left(\frac{s^{data}(t)}{s^*} \right)^{\gamma/\beta}. \tag{54}$$

where Θ^{data} is the experience component as explained in the paper. The relationship of $h_t(s_t)$ and h_{ss} is motivated by the human capital formulation

$$h(s^*, \mathbf{E}^*) = (\hat{e}/p_E)^{\gamma} (s^*)^{\gamma/\beta}$$

when $\hat{e}(a)$ is a constant, $\hat{e}$.

2 Calibration targets

(1) Goods cost of raising a child as a percentage of lifetime income is $e^{-rF} b_1^* \pi(F) / W(s^*, n^*)$.

(2) Return to schooling:

$$\frac{h_s(s^*, \mathbf{E}^*)}{h(s^*, \mathbf{E}^*)} = \frac{\gamma}{\beta} h(s^*, \mathbf{E}^*)^{-\frac{\beta}{\gamma}} \left(\frac{d\hat{e}^*(s^*)}{p_E} \right)^{\beta}$$

(3) Private expenditures in education as a percentage of GDP, denoted by Epriv/y(US). Aggregate private expenditures in education, denoted by AE, are defined as the following form but taking into account the demographics in the economy.

$$Epriv/y(US) = \frac{AE}{Y_{ss}} = \frac{AE/N}{Y_{ss}/N}.$$

We will define the numerator and the denominator as follows. First the steady state density of age-a people is

$$\tilde{n}(a) = \frac{e^{-g_n a} \pi(a)}{\int_0^{\infty} e^{-g_n a} \pi(a) da},$$

where g_n is the steady state population growth satisfying

$$n^* \pi(F) = e^{g_n F}.$$

$$\frac{Y_{ss}}{N} = \frac{w}{(1-\alpha)} \frac{H_{ss}}{N}$$

It comes from

$$w = (1-\alpha) \frac{Y_{ss}}{H_{ss}},$$

where

$$\frac{H_{ss}}{N} = \int_{s^*}^{R} h(a) \tilde{n}(a) da = h(s^*, \mathbf{E}^*) \int_{s^*}^{R} e^{\nu(a-s^*)} \frac{e^{-g_n a} \pi(a)}{\int_0^{\infty} e^{-g_n a} \pi(a) da} da.$$

The last equality is because

$$h(a) = h(s^*)\, e^{v(a-s^*)}$$

$$\begin{aligned}
\frac{AE}{N} &= \int_0^{\min(s^*,\underline{s})} \hat{e}^*(a)\frac{N(a)}{N}da + \int_{s_p}^{s^*} \hat{e}^*(a)\frac{N(a)}{N}da - \int_{s_p}^{\min(s^*,\overline{s})} e_p\frac{N(a)}{N}da \\
&= \int_0^{\min(s^*,\underline{s})} \hat{e}^*(a)\tilde{n}(a)da + \int_{s_p}^{s^*} \hat{e}^*(a)\tilde{n}(a)da - \int_{s_p}^{\min(s^*,\overline{s})} e_p\tilde{n}(a)da \\
&= \widehat{e}^*(0)\int_0^{\min(s^*,\underline{s})} q(a)^{-\frac{1}{1-\beta}}\tilde{n}(a)da + \widehat{e}^*(0)\int_{s_p}^{s^*} q(a)^{-\frac{1}{1-\beta}}\tilde{n}(a)da - \int_{s_p}^{\min(s^*,\overline{s})} e_p\tilde{n}(a)da.
\end{aligned}$$

(4) Assume $\sigma = 1$, the value of statistical life at age-t is given by

$$\begin{aligned}
\frac{\partial c(t)}{\partial \pi(t)} &= \frac{\partial V/\partial \pi(t)}{\partial V/\partial c(t)} \\
&= \frac{C^{-\eta}(C-\underline{C})\,\rho\int_0^\infty e^{-\rho a}\left(\frac{1}{1-\theta}\frac{\partial\pi(a)/\partial\pi(t)}{\pi(a)}\right)da + \frac{1-\eta}{1-\theta}\pi(F)^{\frac{\theta-\eta}{1-\theta}}\frac{\partial\pi(F)}{\partial\pi(t)}e^{-\rho F}\phi(n^*)V}{C^{-\eta}(C-\underline{C})\,\rho\frac{e^{-\rho t}}{c(t)}} \\
&= \frac{c(t)e^{\rho t}}{1-\theta}\left[\begin{array}{c}\int_0^\infty e^{-\rho a}\left(\frac{\partial\pi(a)/\partial\pi(t)}{\pi(a)}\right)da \\ +\frac{1}{\rho}\pi(F)^{\frac{\theta-\eta}{1-\theta}}\frac{\partial\pi(F)}{\partial\pi(t)}\frac{e^{-\rho F}\phi(n^*)}{1-\pi(F)^{\frac{1-\eta}{1-\theta}}e^{-\rho F}\phi(n^*)}\left(1+\frac{\underline{C}}{C-\underline{C}}\right)\end{array}\right]
\end{aligned}$$

Consider the perpetual youth problem: $\pi(a) = \pi(t)e^{-m(a-t)}$. In that case, the previous expression reduces to:

$$\frac{\partial c(t)}{\partial \pi(t)} = \frac{c(t)/\pi(t)}{1-\theta}\frac{1}{\rho}\left[\frac{1+\pi(F)^{\frac{1-\eta}{1-\theta}}e^{-\rho(F-t)}\phi(n^*)\left(1-e^{-\rho t}+\frac{\underline{C}}{C-\underline{C}}\right)}{1-\pi(F)^{\frac{1-\eta}{1-\theta}}e^{-\rho F}\phi(n^*)}\right]$$

At $t = F$,

$$VSL(F) = \frac{\partial c(F)}{\partial \pi(F)} = \frac{c(F)/\pi(F)}{1-\theta}\frac{1}{\rho}\left[\frac{1+\pi(F)^{\frac{1-\eta}{1-\theta}}\phi(n^*)\left(1-e^{-\rho F}+\underline{C}/(C-\underline{C})\right)}{1-\pi(F)^{\frac{1-\eta}{1-\theta}}e^{-\rho F}\phi(n^*)}\right]$$

(5) Time cost of raising children as a percentage of lifetime income. By (5),

$$\frac{(1-l(n^*))\int_F^R e^{\nu(a-s^*)}q(a)da}{\int_{s^*}^F e^{\nu(a-s^*)}q(a)da + l(n^*)\int_F^R e^{\nu(a-s^*)}q(a)da}$$

(6) Income elasticity of fertility: OLS estimation of α from $log(n^*) = \alpha\log(W(s^*, n^*, \mathbf{E}^*)) + \varepsilon$.

$$W\left(s^*, n^*\right) = wh\left(s^*, \mathbf{E}^*\right)\left[\int_{s^*}^{F} e^{\nu(a-s^*)} q(a)da + l(n^*)\int_{F}^{R} e^{\nu(a-s^*)} q(a)da\right]. \tag{55}$$

3 Solution Algorithm

For each country, we solve the model by first assuming an initial set of values $\{s^*, n^*, e^*(0), \bar{\tau}\}$ given $e_p, \overline{s}, p_1, p_2, p_3$ and Θ^{data} obtained from the data, as well as other parameters given. With these initial values, we can obtain $G(n^*)$ by (21). The optimal total educational expenditure $\hat{e}^*\left(s^*\right)$ evaluated at age s^*, can be gotten by (25), the private educational expenditure $e_s^*\left(a\right)$ follows from (32), and $\widehat{s}$ is obtained by (29). After $\widehat{s}$ is solved, s_p, $h(s^*, \mathbf{E}^*)$, h^{data}, w, $\mathbf{E}^*$, $W(s^*, n^*, \mathbf{E}^*)$, b_1^*, $c^*\left(0\right)$, $c^S\left(s^*\right)$, $c^W(s^*)$, C, λ_1, λ_2 and V can be derived through (30), (34), (54), (53), (33), (5), (22), (16), (14), (13), (49), (19), (12), and (48) successively. After all these variables are available, we are able to update s^*, n^*, $e^*(0)$ and $\bar{\tau}$ by (40), (46), (27), and (52).

TABLE 1

Cross-country descriptive statistics - 2013

	Mean	Median	Standard deviation	Maximum	Minimum
GDP per capita (PPP)	$17,517	$12,668	$15,135	$63,483	$561
Education					
School life expectancy (years)	13.86	13.98	3.08	20.43	5.32
Free schooling years	10.6	12	2.4	16	4
Grade repetition rate (primary & secondary)	3.7%	1.2%	5.6%	30.4%	0.0%
Public education spending per pupil (PPP)	$6,601	$3,952	$7,091	$34,866	$61
Demographics					
Total fertility rate (number of births)	2.56	2.08	1.33	7.62	1.12
Life expectancy at birth (years)	72.17	74.10	8.66	83.83	48.94
Survival probability to age 5	0.97	0.98	0.03	1.00	0.85
Survival probability to age 65	0.75	0.78	0.13	0.91	0.37
Survival probability to age 85	0.29	0.29	0.14	0.55	0.07

Notes: Sample corresponds to 92 countries. GDP per capita (PPP), total fertility rate and life expectancy at birth are from the World Development Indicators. School life expectancy, free schooling years, grade repetition rates and public education spending per pupil are from UNESCO. Survival probabilities are from the life tables published by the World Population Prospects.

TABLE 2

Exogenous parameters

Parameter	Concept	Parameter value
σ	Inverse of EIS	1
r	Interest rate	2.5%
ρ	Rate of time preference	2.5%
α	Capital share	0.33
ν	Returns to experience	2%
F	Average childbearing age	28
R	Retirement age	65

Notes: The values of parameters σ, r, α and ν are standard in the quantitative macro literature. Setting $\rho = r$ implies that the growth rate of consumption over the life cycle is determined by the age-dependent mortality rate. Parameter F is consistent with the world average childbearing age from the United Nations' World Fertility Patterns 2015. Parameter R is set to be binding for richer countries.

TABLE 3
Calibrated parameters

Parameter	Concept	Target	Target value	Parameter value
η	Inverse of the elasticity of intergenerational substitution	World mean of fertility	2.56	0.339
θ	Mortality risk aversion	Value of statistical life at childbearing age in the US	$4 million	0.535
$\underline{C}$	Non-market consumption	Income elasticity of fertility	-0.38	4900
γ	Returns to scale human capital production function	Average private expenditures in education as % of GDP in OECD	0.9%	0.335
β	Degree of substitution education spending across ages	World mean of school life expectancy	13.86	0.172
ψ	Level of altruism	Goods cost of raising a child as % of lifetime income in US	16.44%	0.475
χ	Degree of diminishing altruism	World standard deviation of fertility	1.33	1.466
λ	Level time cost of raising children	Time cost of raising a child as % of lifetime income in US	17%	0.318

Notes: The value of statistical life for the US is from Viscusi and Aldi (2003). The income elasticity of fertility is from Jones and Tertilt (2008). Average OECD private educational expenditures as a % of GDP is from the National Center of Education Statistics. Goods and time costs of raising children are computed following Cordoba and Ripoll (2016, 2018). The remaining targets are computed using the sample of countries described in Table 1.

TABLE 4

Model's performance

Untargeted moments	Data	Model
World standard deviation of school life expectancy	3.08	3.09
World quantity-quality trade-off	-0.33	-0.41
World maximum school life expectancy	20.43	18.29
World minimum school life expectancy	5.32	6.61
World maximum fertility	7.62	6.29
World minimum fertility	1.12	0.96
Returns to schooling in the US	8.28%	8.68%

Correlations
Fertility in model and data = 75.8%
Schooling in model and data = 77.9%

Notes: Model is calibrated as in Tables 2 and 3. Returns to schooling for the US in the data are computed as in Bils and Klenow (2000). All other data moments are computed using the sample summarized in Table 1.

TABLE 5

Counterfactuals (% change)

	Mean			Standard deviation		
Parameter	Schooling	Fertility	Per capita income	Schooling	Fertility	Per capita income
Changes to individual parameters						
TFP	17.5	-39.3	59.8	-54.0	-72.2	-76.5
e_p	10.6	-24.2	27.9	-45.5	-50.3	-48.7
$\bar{s}$	4.5	-4.2	3.7	-36.4	-14.0	-12.9
p_1	1.1	-1.7	0.7	-5.7	-4.7	-1.8
p_2	7.7	-11.2	7.8	-24.1	-19.9	-12.4
p_3	1.1	-3.3	-3.4	-2.9	-5.3	-3.9
Changes to groups of parameters						
e_p, $\bar{s}$	9.3	-28.2	36.1	-46.8	-61.9	-58.9
p_1, p_2, p_3	10.4	-16.2	5.5	-32.8	-29.7	-19.1
Eliminating public education subsidies						
$e_p = 0$	-0.1	14.8	-22.3	35.6	36.5	27.5

Notes: Counterfactuals are computed equating each parameter to its value in an artificial country that has the 90th percentile of TFP, survival rates and public schooling policies. The model is calibrated as in Tables 2 and 3. The standard deviation of per capita income is computed over the log (10 base) of income. To preserve comparability across counterfactuals, results on this table omit 7 countries in the sample for which a public education subsidy per pupil of $17,179 (90th percentile of sample) cannot be implemented without violating the government's budget constraint.

TABLE 6

Out-of-sample predictions – The United States in 1900

Data for the US in 1900	
School life expectancy (years)	9.40
Fertility (births per woman)	3.56
Public education spending per pupil (2013 dollars)	$567
Duration of public education subsidies (years)	8
Survival probability to age 5	0.800
Survival probability to age 65	0.392
Life expectancy at birth	47.7
GDP per capita (2013 dollars)	$7,118

Model predictions for the US in 1900	
School life expectancy	9.56
Fertility	3.74

Notes: US data from 1900 is taken from various sources including the US Historic Statistics since Colonial Times to 1970 (US Department of Commerce); the Historic Summary of Public Elementary and Secondary School Statistics (National Center for Education Statistics); 120 Years of American Education (NCES); the CPI Estimate 1800- (Minneapolis Fed); the Life Tables for the US Social Security Area 1900-2010 (Social Security Administration); and the Historic Statistics for the World Economy (Angus Maddison).

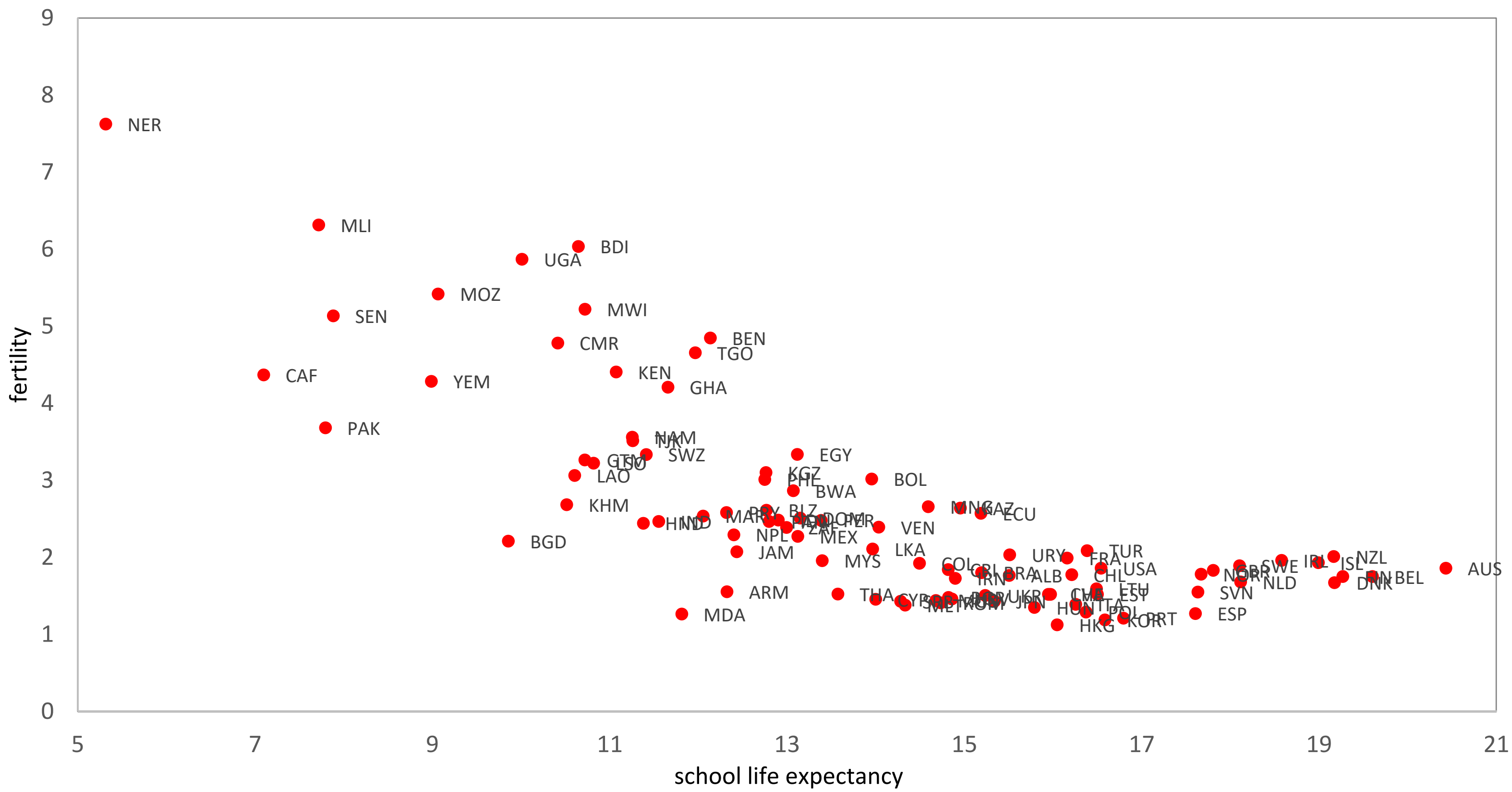

FIGURE 1
International quantity-quality trade-off in the data- 2013

Notes: Fertility is from the World Development indicators and it corresponds to the total fertility rate, or the number of children that would be born to a woman if she were to live to the end of her childbearing years and bear children in accordance with age-specific fertility rates. School life expectancy is from UNESCO and it corresponds to total number of years of schooling a child expects to receive assuming that the probability of being enrolled in school equals the current enrollment ratio for each age.

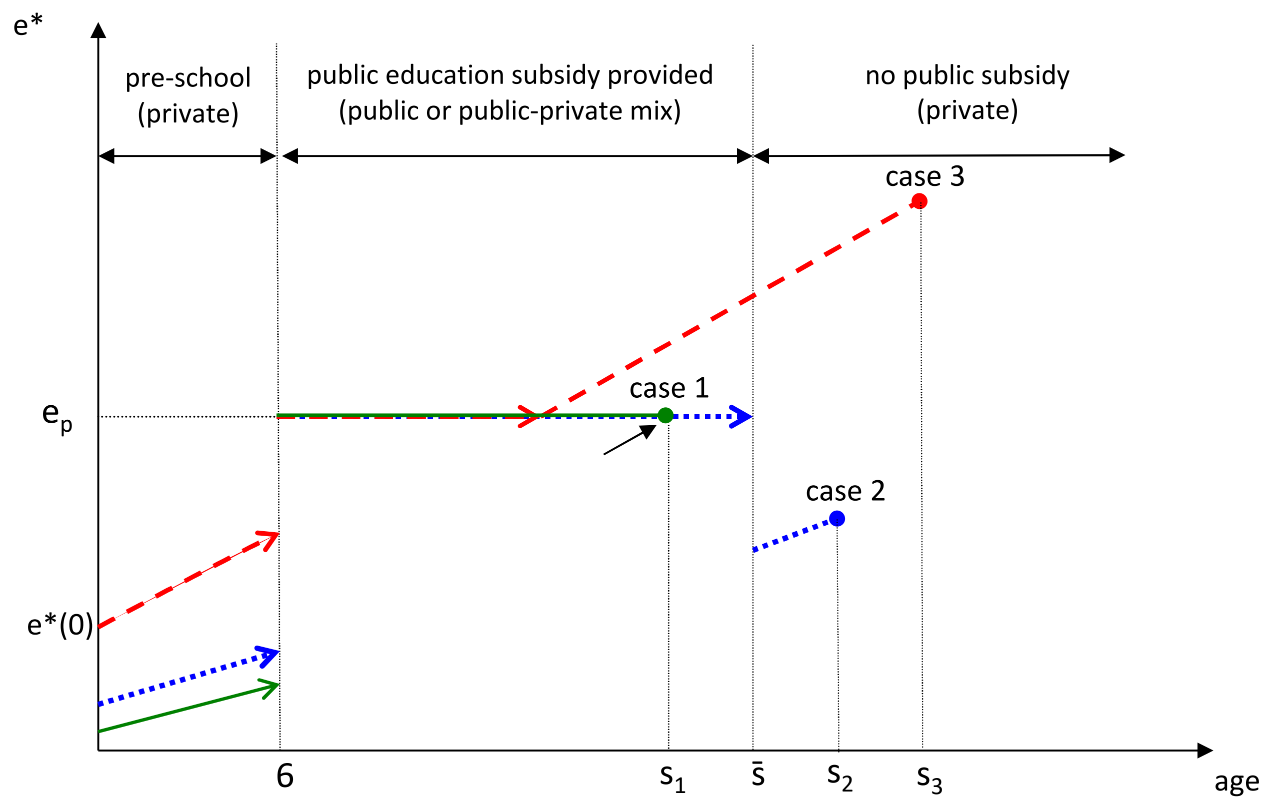

FIGURE 2
Optimal expenditures in education in the model - Total expenditures by age, e*(a)

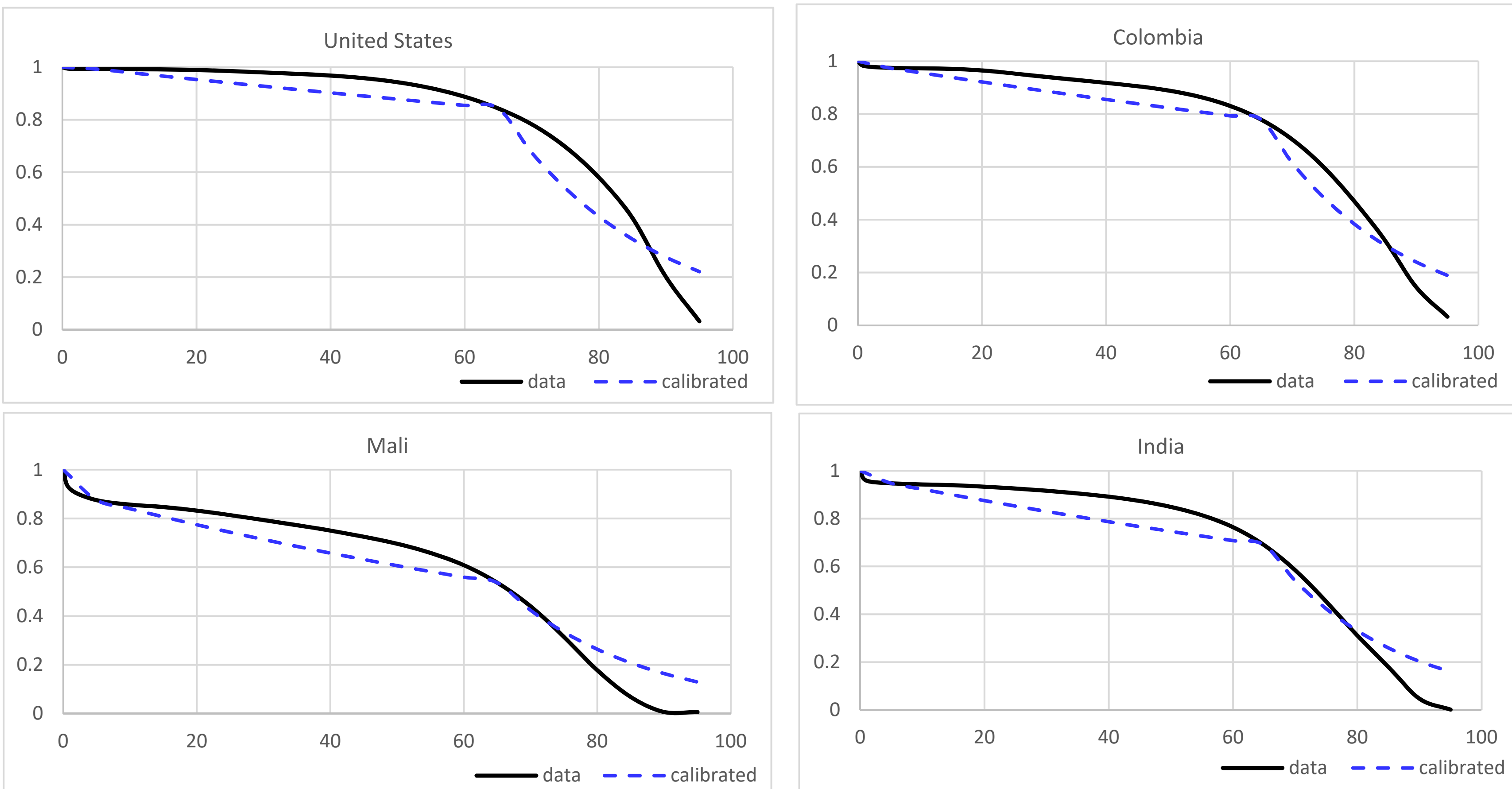

FIGURE 3
Survival probabilities by age in selected countries

Notes: Age-specific survival rates are calibrated for each country by assuming a survival probability process with distinct constant hazard rates before age 5, between ages 5 and 65, and after age 65. Hazard rates are calibrated to the survival probabilities at ages 5, 65 and 85 from the World Population Prospects data.

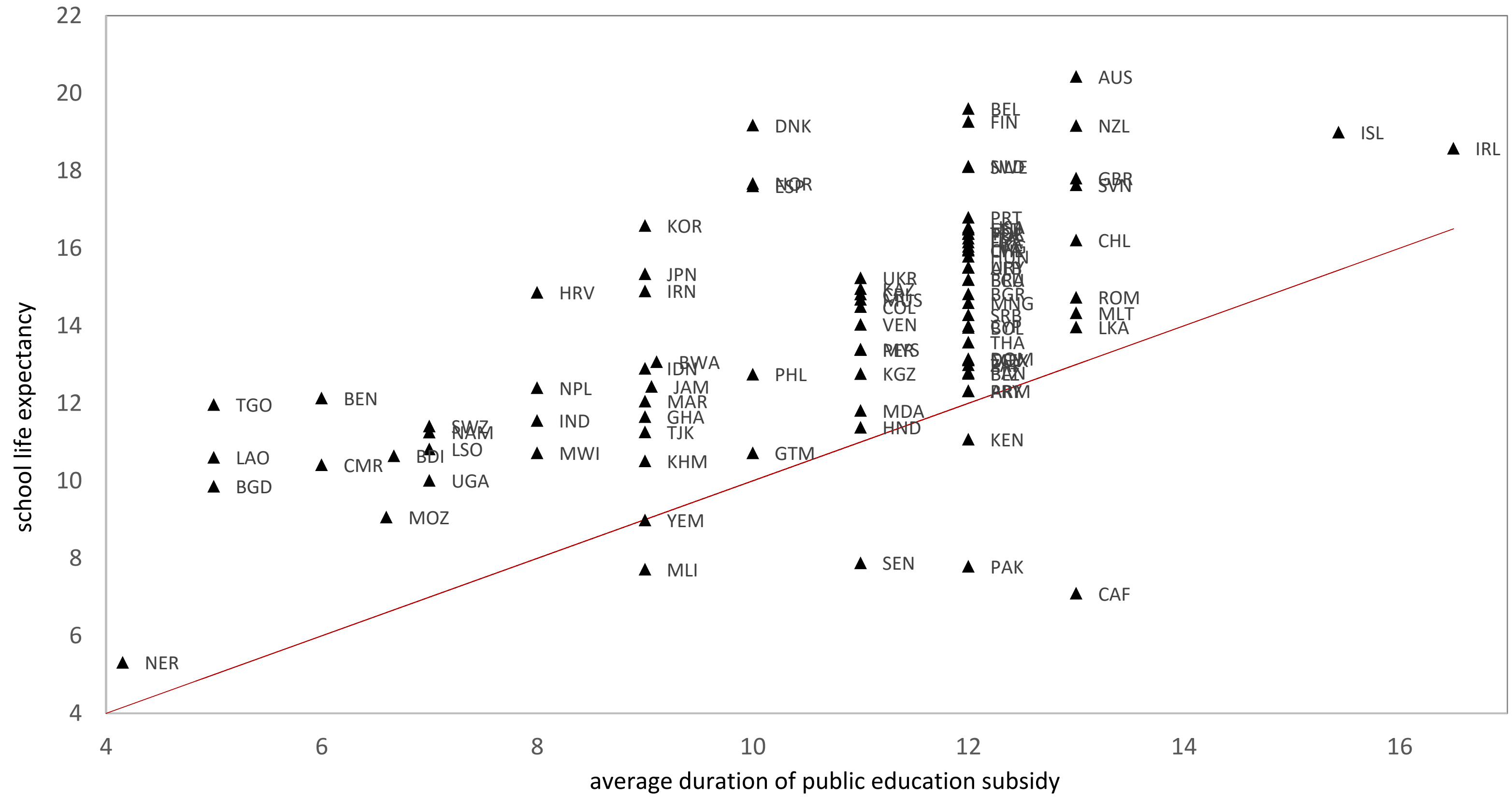

FIGURE 4
Duration of public education subsidy and school life expectancy- 2013

Notes: Duration of the public education subsidy corresponds to the number of years free education is available according to UNESCO. School life expectancy is from UNESCO as in Figure 1.

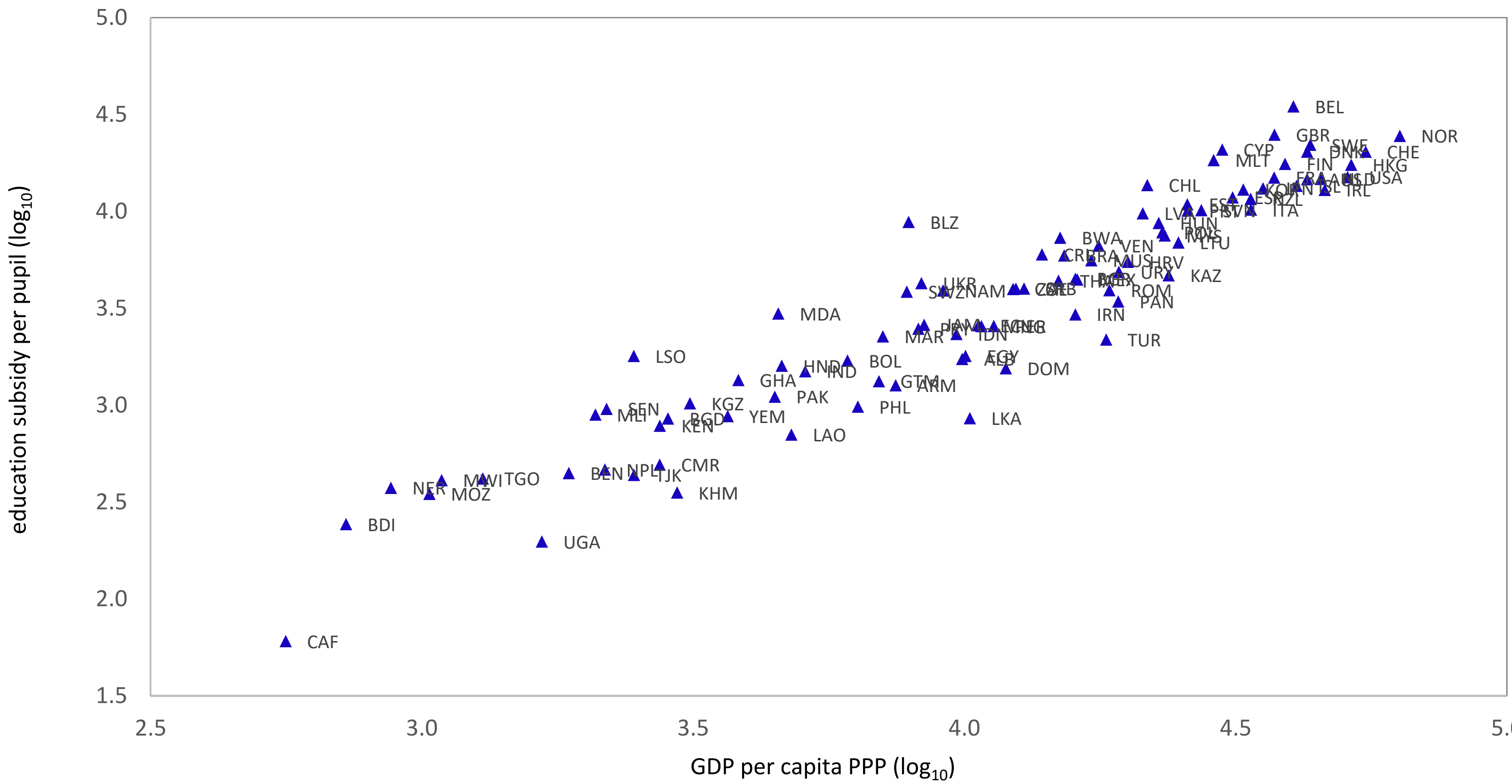

FIGURE 5

Public education subsidy per pupil and GDP per capita- 2013

Notes: Public education subsidies per pupil are constructed by dividing total government educational expenditures (PPP adjusted) by the total number of students enrolled in public institutions from UNESCO. GDP per capita (PPP adjusted) is from the World Development Indicators.

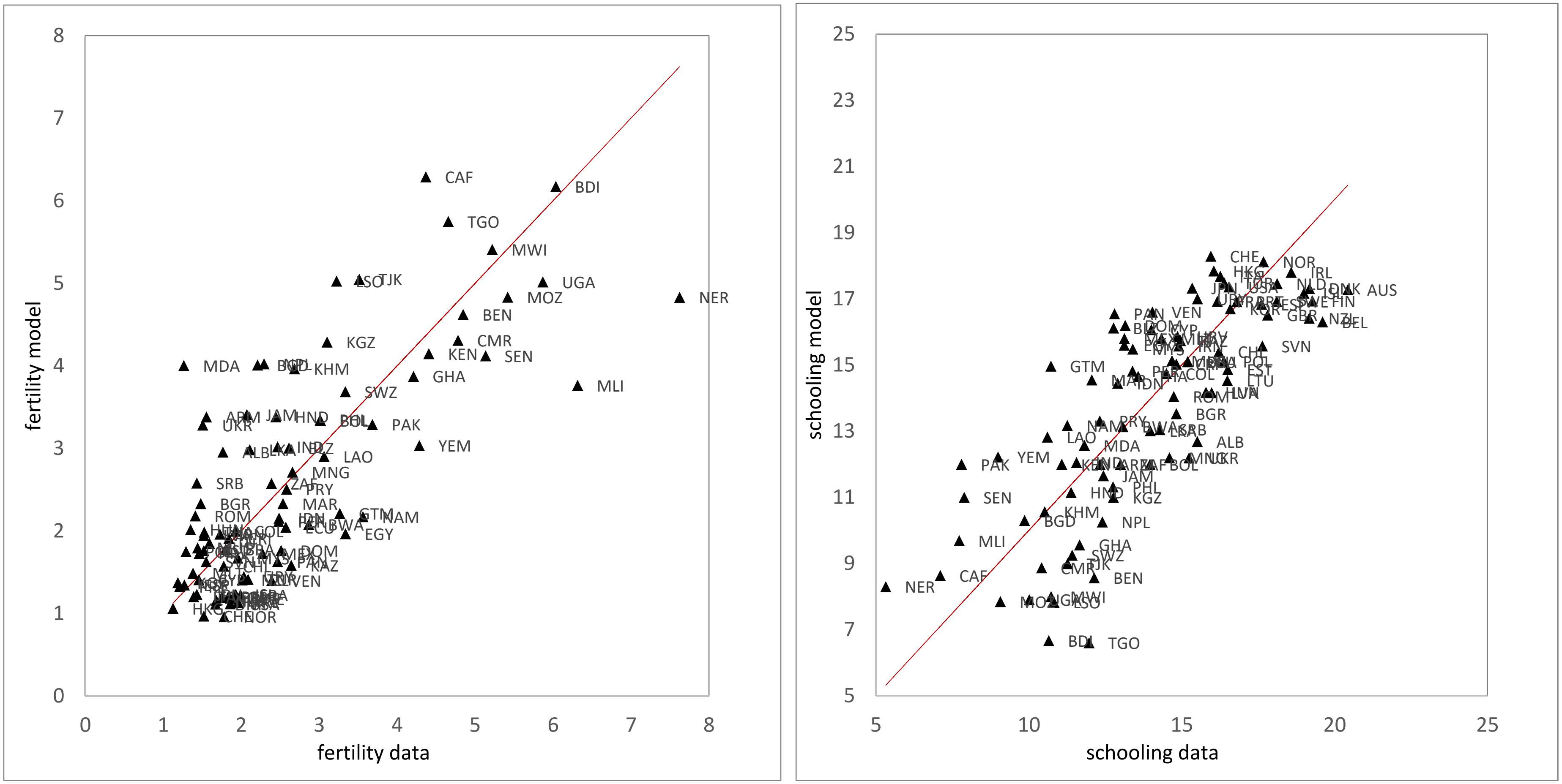

FIGURE 6
Fertility and schooling - Model versus data

Notes: Fertilty in the data corresponds to total fertility rate as in Figure 1. Schooling in the data corresponds to school life expectancy as in Figure 1. The model is calibrated as in Tables 2 and 3.

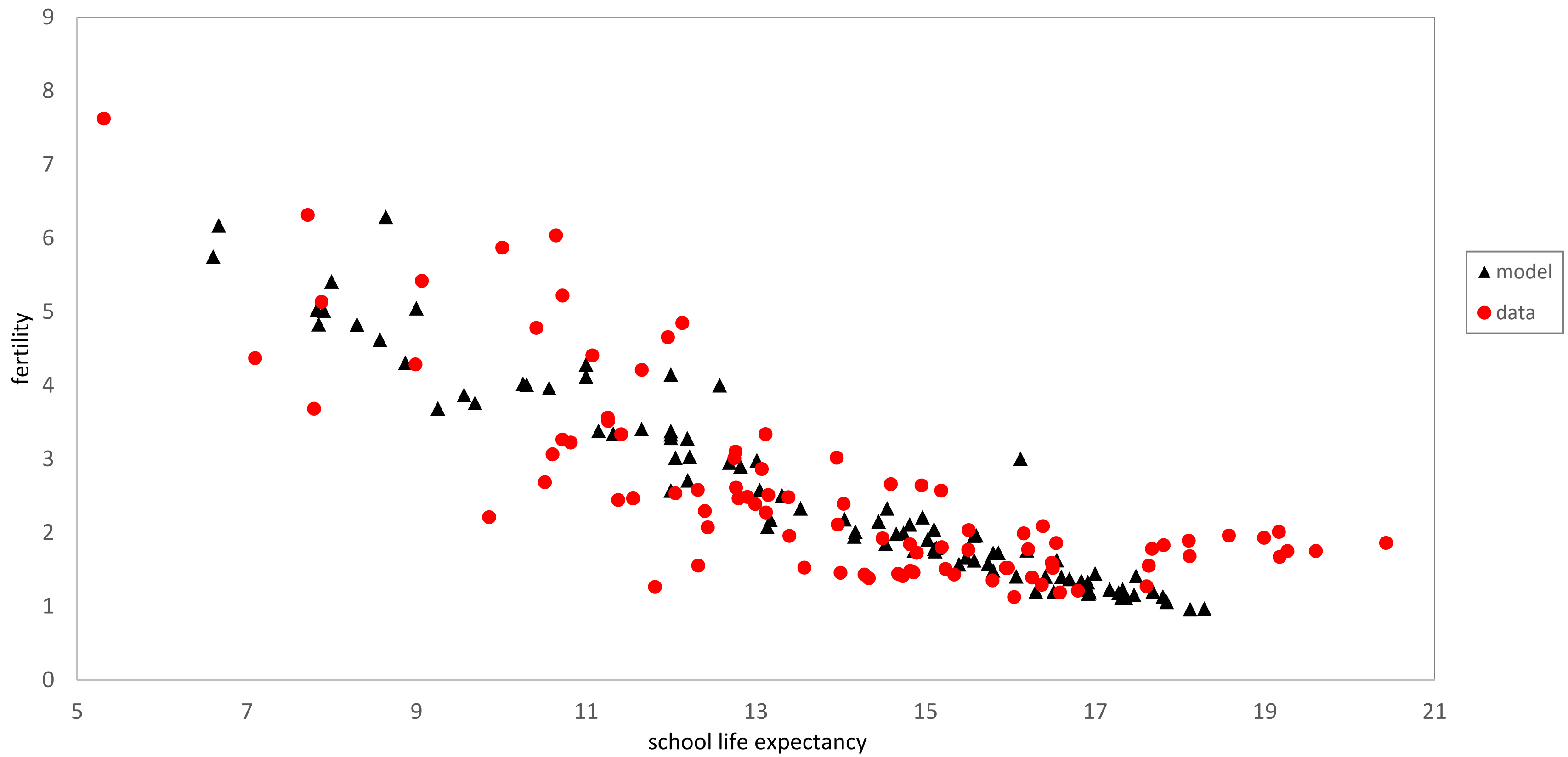

FIGURE 7
International quantity-quality trade-off in the model and the data

Notes: Same as in Figure 1. The model predictions correspond to those of the calibrated benchmark.

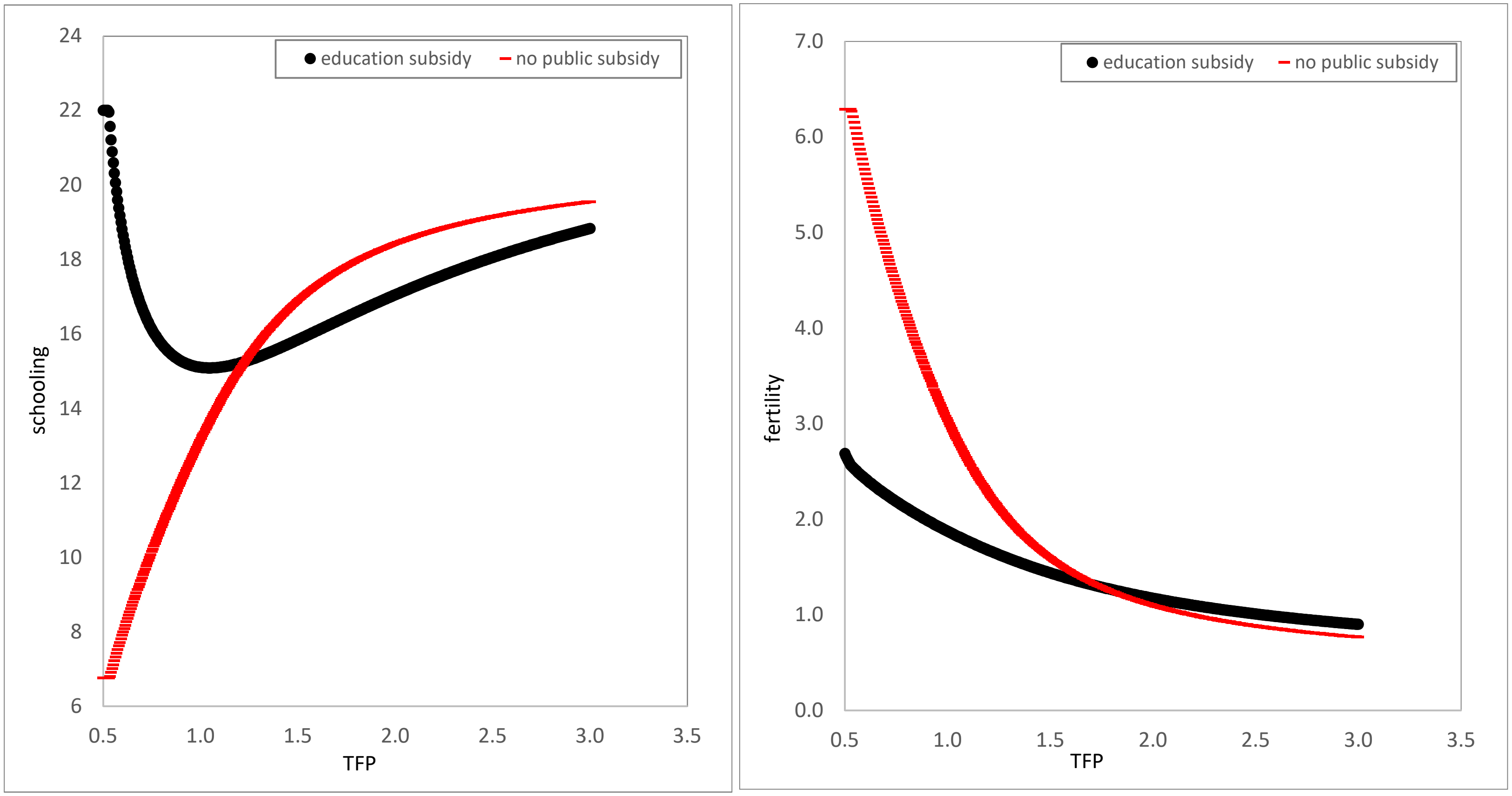

FIGURE 8
Schooling and fertility as a function of TFP

Notes: Fertility and schooling are shown as functions of hypothetical values of TFP. The figure uses the calibrated parameters for the benchmark model in Tables 2 and 3, as well as the country-specific parameters for an artificial "rich" country whose educational policies correspond to the 90th percentile of the sample, and whose mortality rates are at the 10th percentile.

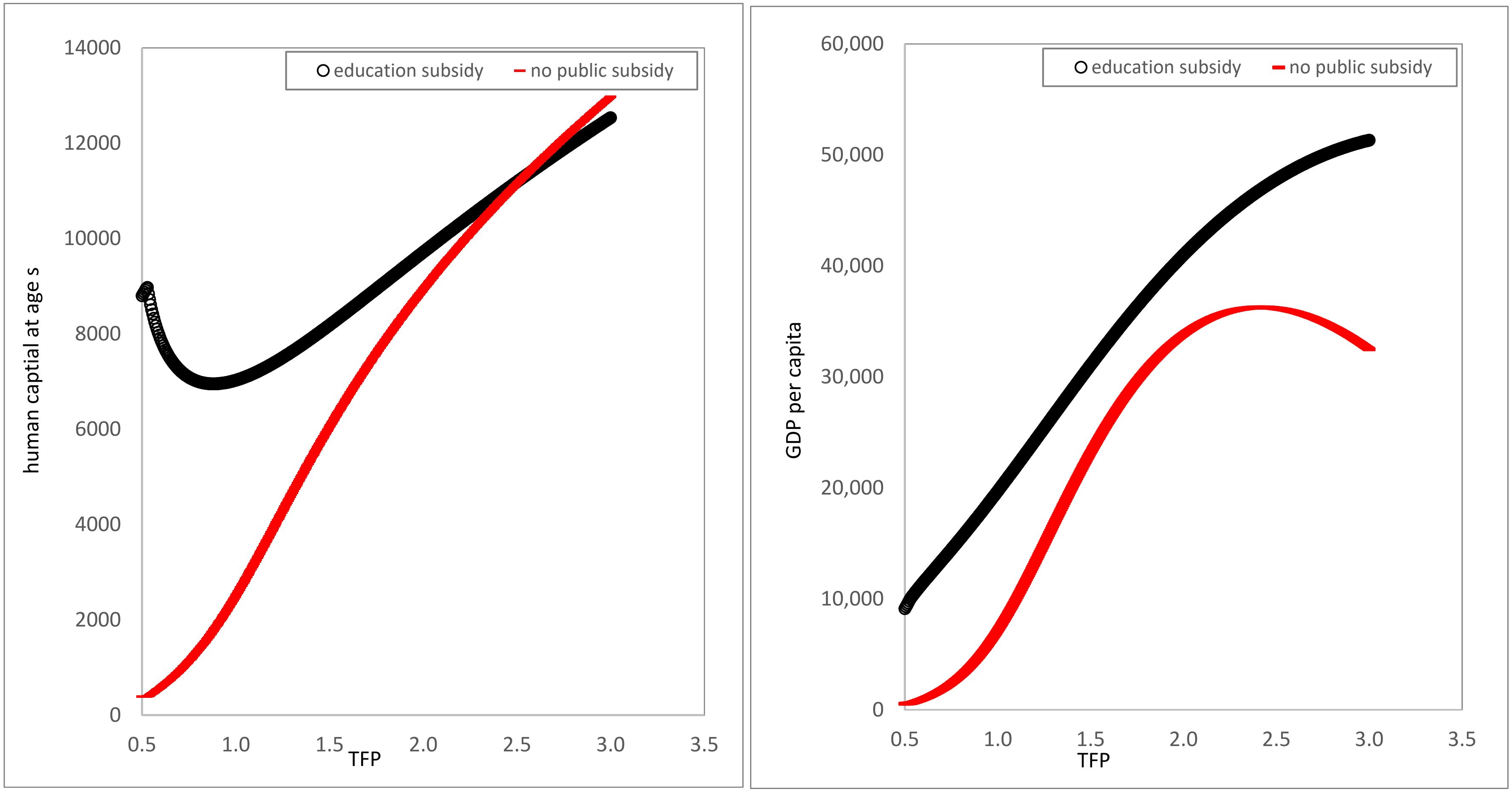

FIGURE 9
Human capital and GDP per capita as a function of TFP
Comparison between the case with education subsidiy and no public subsidy

Notes: Same as in Figure 8.

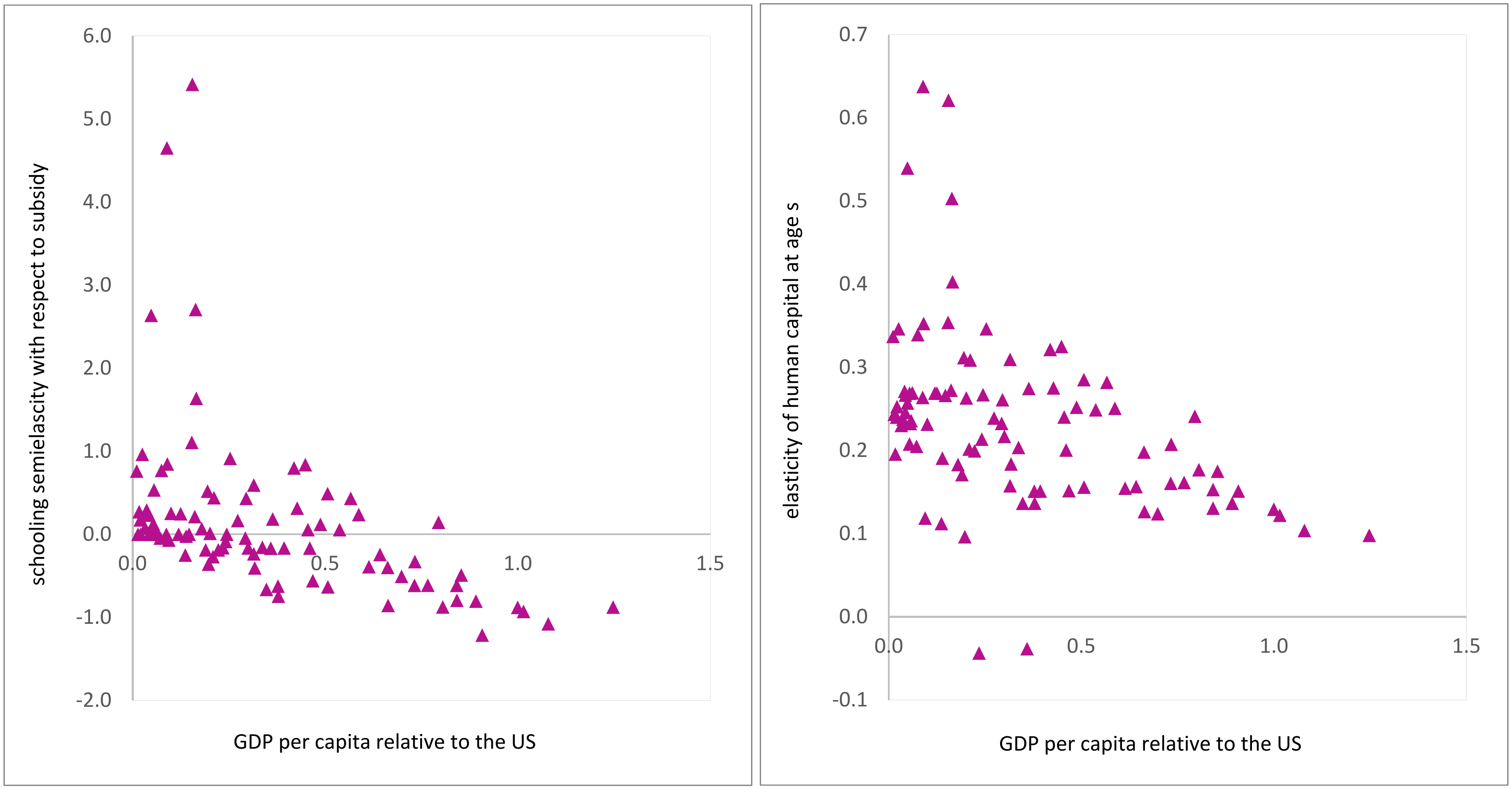

FIGURE 10
A 10% increase in public education spending - Effects on schooling and human capital at age s

Notes: GDP per capita in 2013 is from the World Development Indicators. Local counterfactuals are performed with the calibrated model.

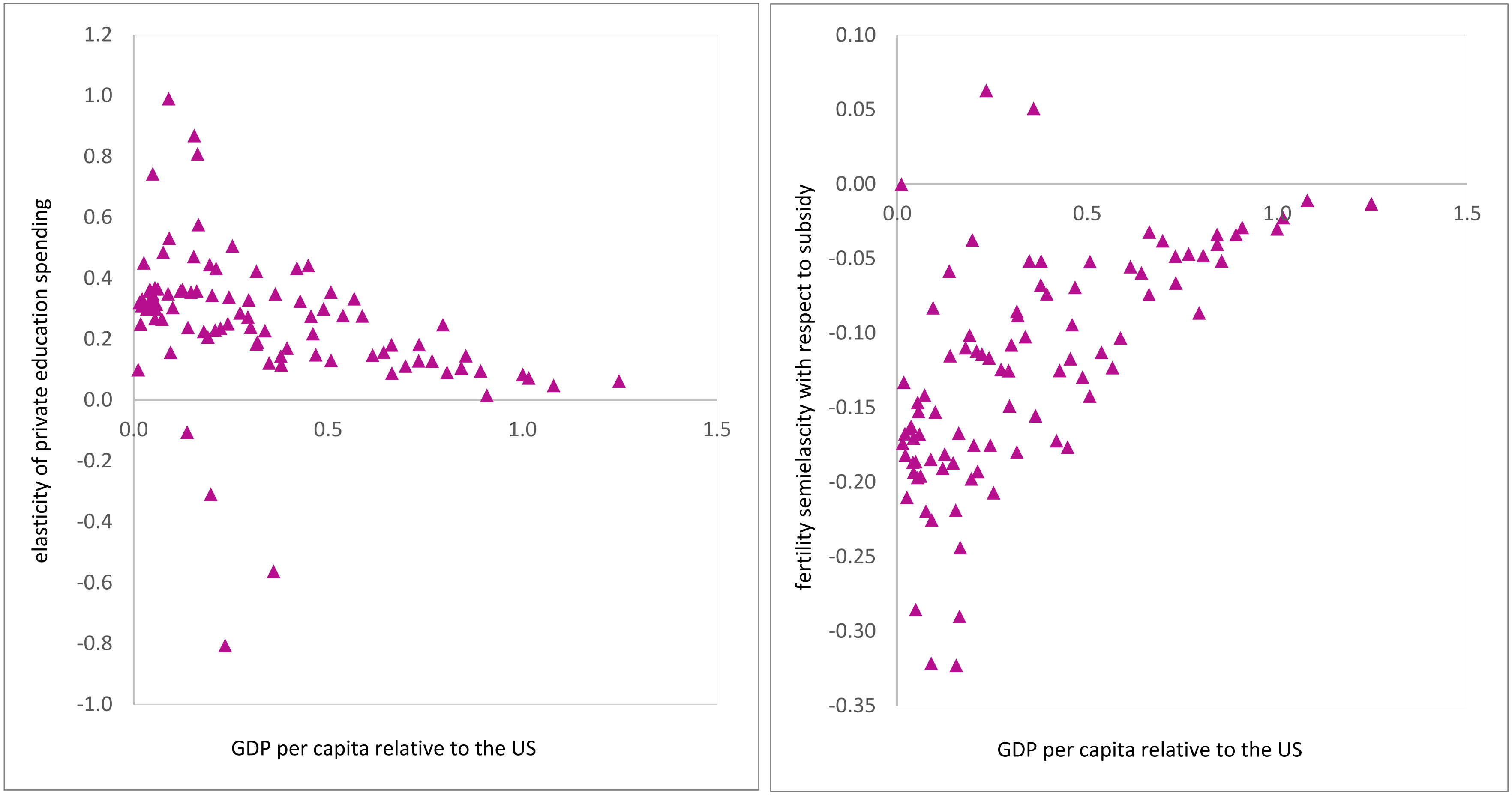

FIGURE 11
A 10% increase in public education spending - Effects on private education spending and fertility

Notes: Same as Figure 10.

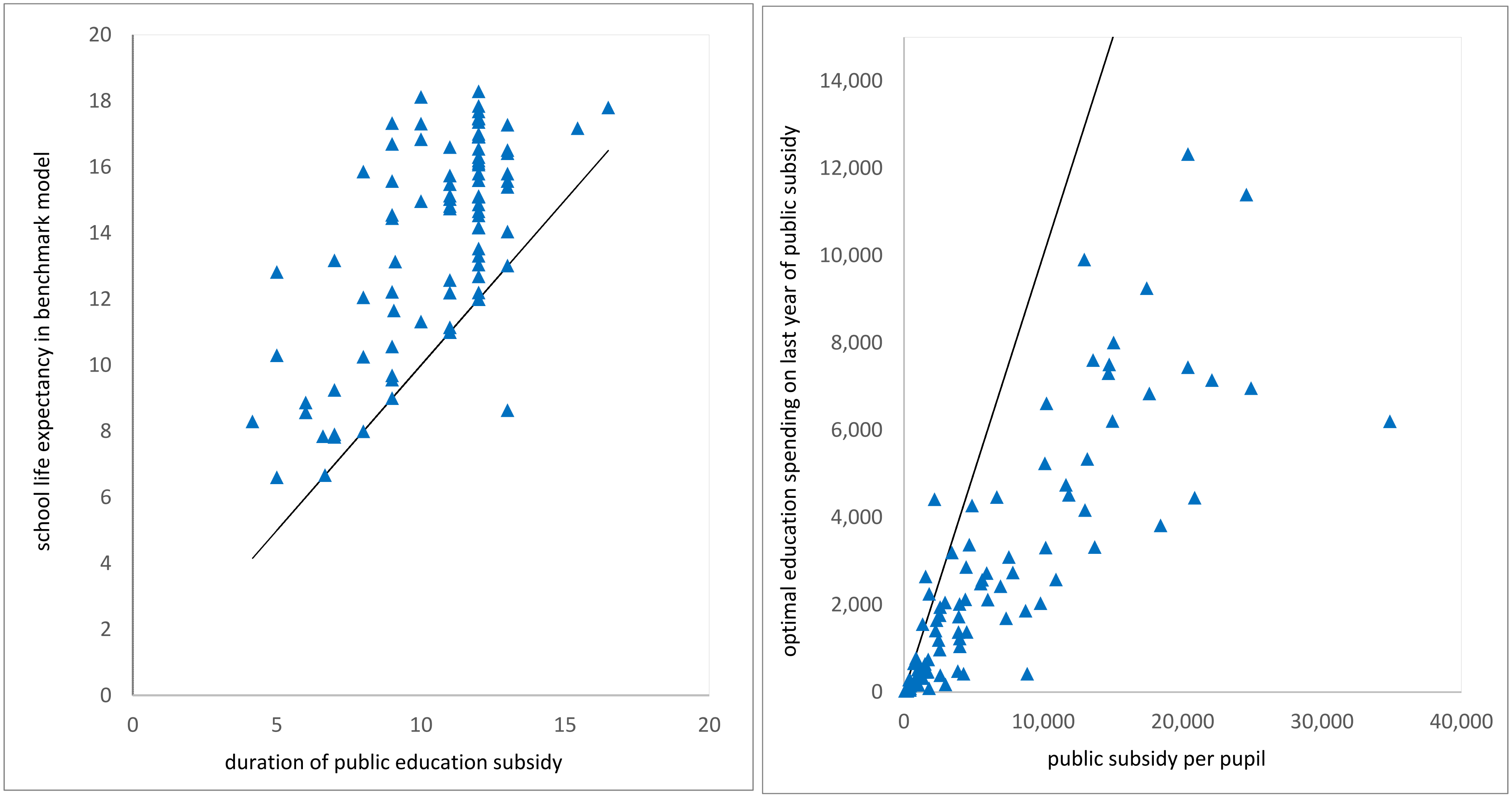

FIGURE 12

Public eduaction policies in calibrated model

Notes: Duration of the public education subsidy corresponds to the number of years free education is available according to UNESCO. School life expectancy is from UNESCO as in Figure 1. Public subsidy per pupil is as in Figure 4. Optimal education spending on the last year of public subsidy is computed using the calibrated model.

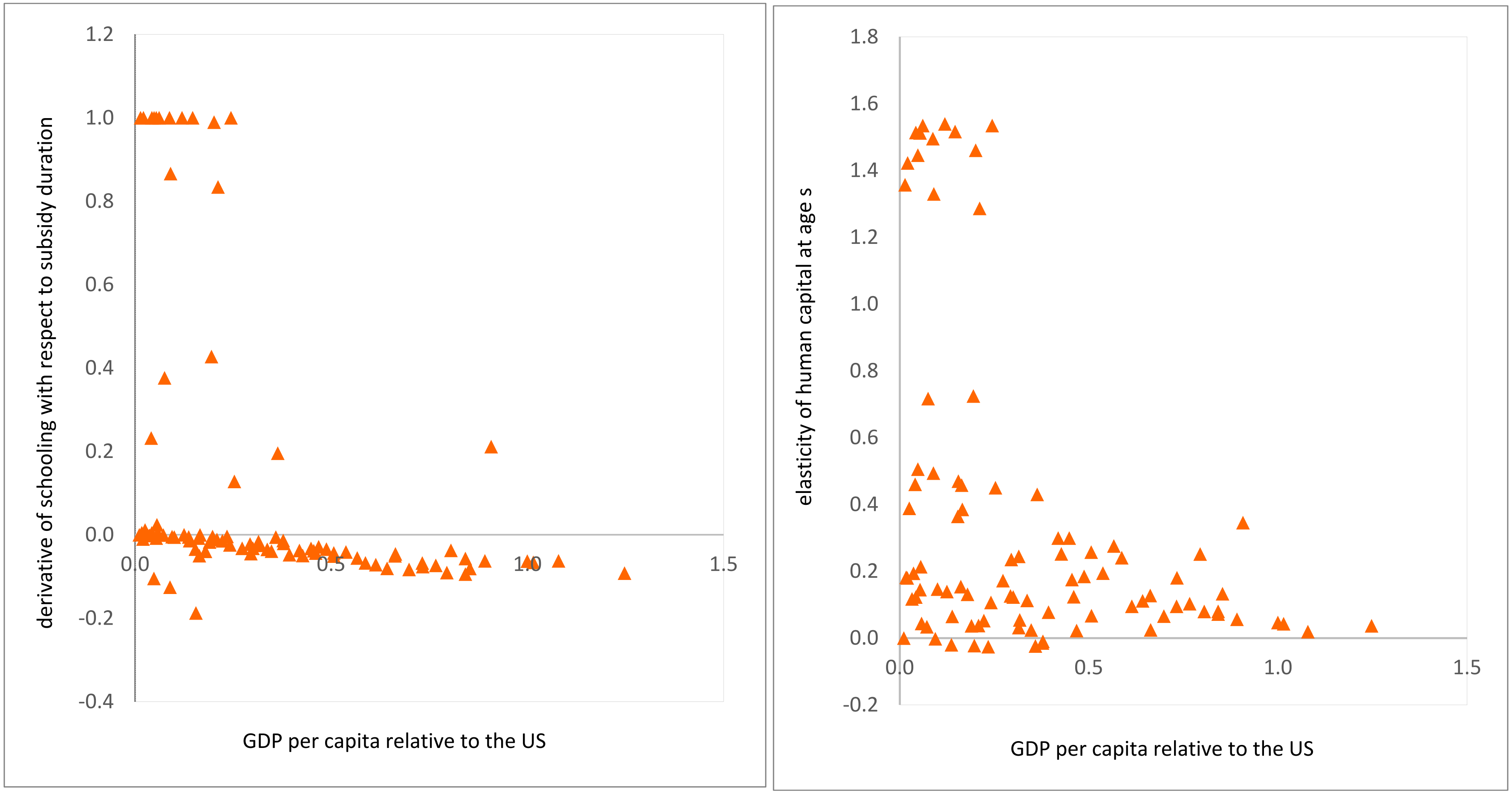

FIGURE 13

A 10% increase in the duration of public education subsidy - Effects on schooling and human capital at age s

Notes: Same as in Figure 10.

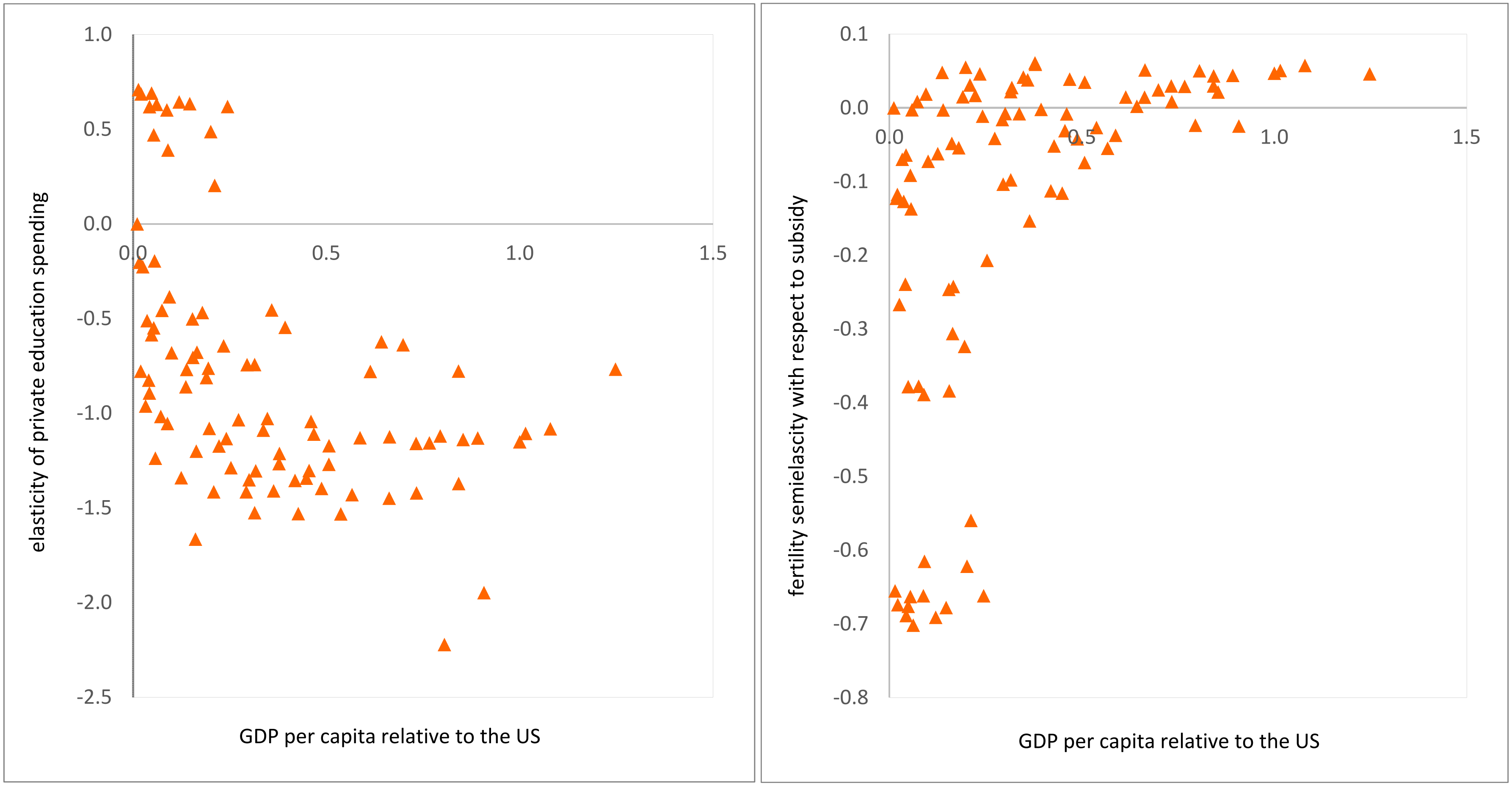

FIGURE 14
A 10% increase in the duration of public education subsidy - Effects on private education spending and fertility

Notes: Same as in Figure 10.

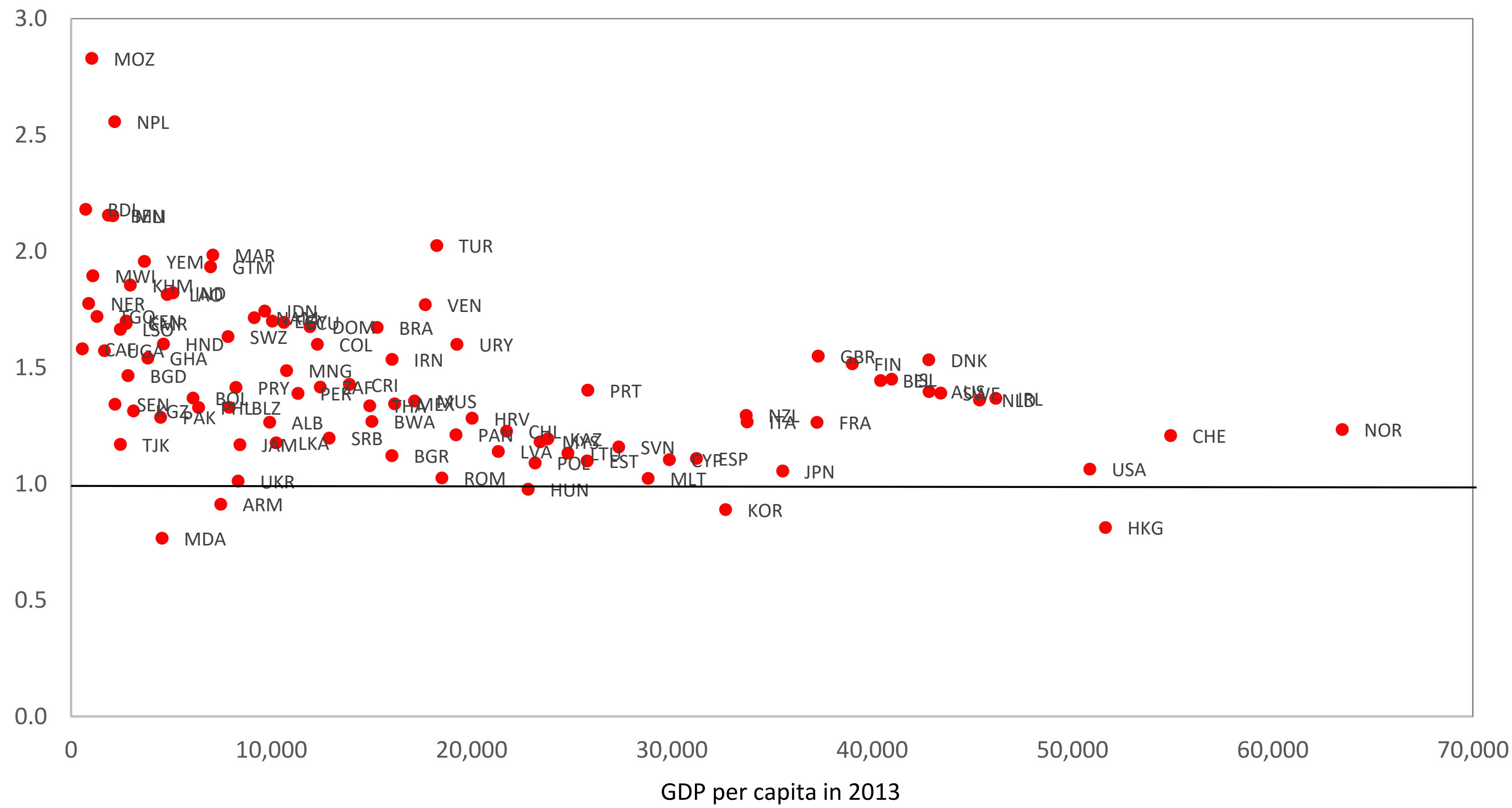

FIGURE 15

GDP per capita in the steady state relative to GDP per capita in 2013

Notes: GDP per capita in 2013 is from the World Development Indicators and GDP per capita in the steady state is as implied by the benchmark model if it exactly matched schooling life expecancy and fertility for every country.

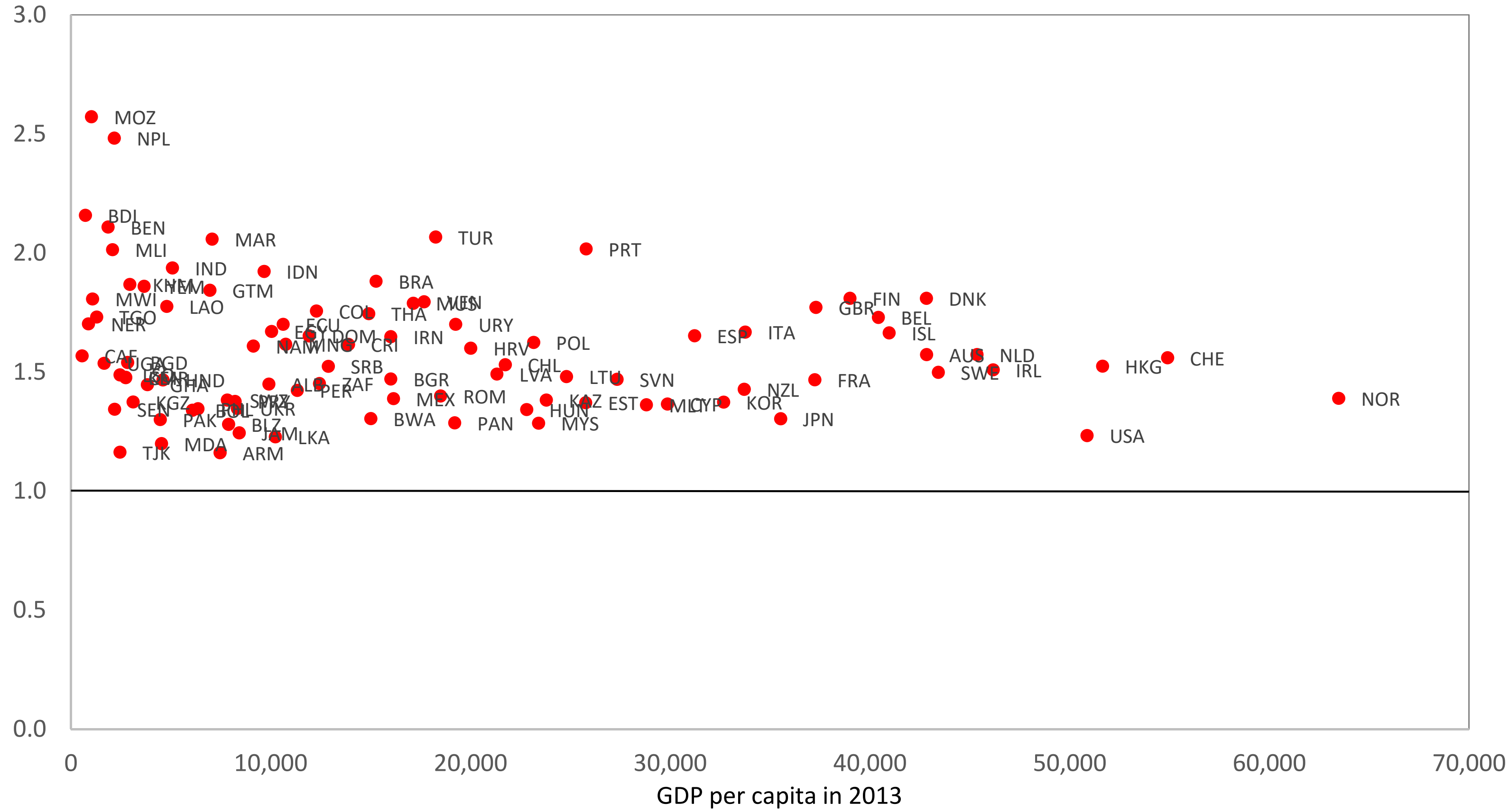

FIGURE 16
The role of human capital:
Counterfactual GDP per capita in the steady state relative to GDP per capita in 2013*

Notes: GDP per capita in 2013 is from the World Development Indicators. The computation of per capita income in the steady state assumes that the age structure of the population is as in 2013 and that the only change is that steady state workers have a schooling equal to the school life expectancy and the human capital of the model's steady state.

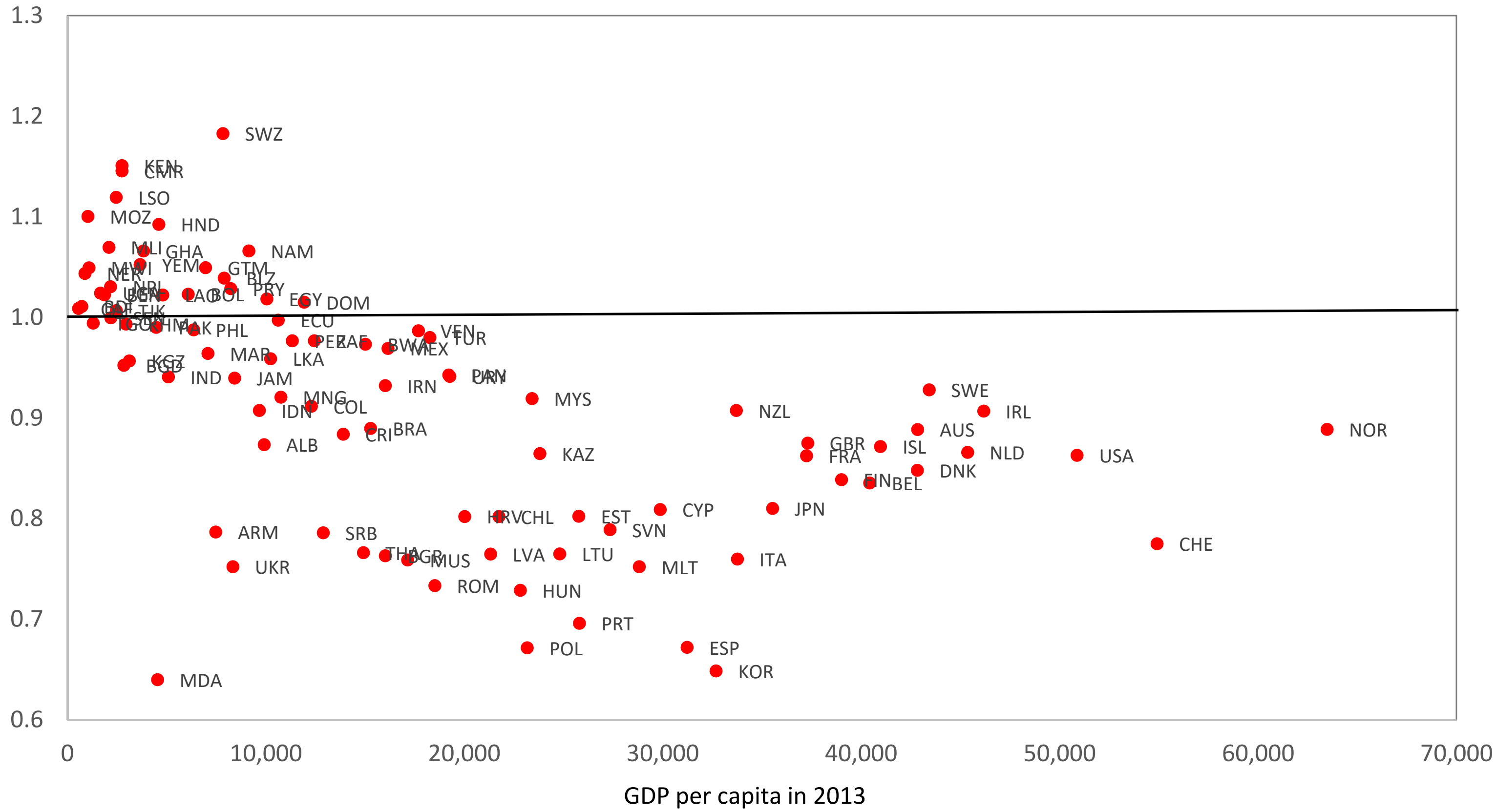

FIGURE 17

The role of demographics:

Counterfactual GDP per capita in the steady state relative to GDP per capita in 2013*

**Notes:* GDP per capita in 2013 is from the World Development Indicators. The computation of per capita income in the steady state assumes that the schooling of the workers remains as the Barro and Lee schooling in 2013, and that the only change is the age structure of the population in the steady state.