

ACCOUNTING FOR THE INTERNATIONAL QUANTITY-QUALITY TRADE-OFF

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

January, 2019

Abstract

We investigate what accounts for the observed international differences in schooling and fertility, and draw lessons for the underlying sources of cross-country income differences. For this purpose, we extend a life-cycle dynastic model to include features relevant for schooling and fertility choices. Our approach allows for country-specific human capital technologies in addition to differences in TFP, public education policies, and demographic factors. We find that differences in human capital production functions, specifically in the degree of complementarity of educational investments, are key to match schooling data, and result in novel estimates of human capital stocks and TFP levels. According to the model, differences in TFP, public education spending per pupil and retiree survival rates are the most important factors explaining the international dispersion of fertility. Differences in the number of years of public education provision and working-age survival rates are key determinants of the schooling dispersion. Our model suggests that human capital policies are key for development.

Key words: public education, 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 women is associated with an average of three fewer years of schooling.

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

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

[†]Economics and Management School, Wuhan University, Wuhan, Hubei, 430072, China, e-mail: xiyingliu@whu.edu.cn.

[‡]Department of Economics, University of Pittsburgh, 4532 W.W. Posvar Hall, Pittsburgh PA 15260, e-mail: ripoll@pitt.edu.

¹The 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.

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).

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); age-dependent mortality rates; and the provision of public education in terms of the number of years provided and spending per pupil. While these factors have been analyzed separately in the literature, we study them within the same unified framework. This unified framework uncovers novel interactions and amplification effects among different determinants of schooling and fertility choices.

Our model features altruistic parents who make fertility choices in a version of the Barro and Becker (1989) model, and who finance the consumption and private educational expenditures of their children during schooling years. Human capital technologies are of the Ben-Porath type. A publicly provided education subsidy is available for a number of years and financed through lump-sum taxes. Parents face both time and goods costs of raising children. Financial frictions play a role in our theory, since 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 generated 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 of whom will have access to less resources during childhood. To the extent that parental income plays a role in fertility and educational choices, the provision of public education becomes a first-order determinant of educational and fertility outcomes.

The calibration of our model is to a large extent standard, but it 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. 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

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

long way in explaining the international schooling and fertility data.

In order to perform precise accounting and counterfactual exercises, we enrich the benchmark model so that it exactly matches the schooling and fertility data. For this purpose, we allow for country-specific human capital technologies and altruism. We find that in countries such as Niger and Mali, low child discounting helps explain their particularly high fertility levels, beyond what the benchmark predicts, while a high discount rate explains the unusually low fertility in Moldova. Residual high levels of schooling in the benchmark model, which occur for example in countries known for having efficient educational systems such as the Scandinavian countries, are explained in the model by higher complementarity in educational investments in the sense of Cunha *et al.* (2006). Complementarity refers to property of the human capital production function according to which early educational investments enhance the productivity of later investments.

Interestingly, we find significant differences in human capital production functions even among similar countries. Two countries with similar education spending per pupil may have different schooling levels due to differences in their human capital production functions. For example, Niger and Burundi are among the poorest countries in our sample, with similar levels of output per-capita and educational expenditures per pupil, but expected years of schooling are around twice as large in Burundi. Our model rationalizes these schooling difference as arising from differences in human capital technologies. Uncovering human capital production differences results in novel estimates of human capital stocks and in a re-evaluation of the role TFP differences.

Our analysis yields five main insights. First, cross-country differences in TFP, age-dependent mortality rates and public schooling policies go a long way in accounting for the overall international quantity-quality trade-off. We find that jointly these differences can explain 81.6% of the standard deviation of schooling and 77.7% of the fertility dispersion. These variables also can explain 80% of the international quantity-quality trade-off illustrated in Figure 1. In the calibrated benchmark model, one extra child is associated with 2.5 fewer years of schooling. As mentioned, the residual dispersion in schooling and fertility can be rationalized by introducing cross-country differences in human capital production technologies and altruism.

The second main insight refers to the role of TFP in explaining the quantity-quality trade-off. We find that while TFP is quantitatively the most important factor in accounting for the world dispersion of fertility, its role in schooling dispersion is less clear. Equating TFP to the 90th percentile in all countries results in a reduction of the standard deviation of fertility of 55.7%. The key mechanism at work here is the time cost of raising children, which is increasing in TFP (wages) for most high fertility countries. But there are poor countries, such as Niger and Pakistan, where fertility under the counterfactual remains high, at 4.68 children per woman. The reason is that, according to our model, the poverty of Niger and Pakistan is rooted more in its low human capital than in its low TFP. Niger and Pakistan's human capital production functions exhibit a relatively low degree of complementarity, which results in only a marginal increase of schooling under the counterfactual. In this respect, our TFP counterfactual suggests an important qualification to the common view that development, in the form of higher TFP (wages), is the most important tool to reduce fertility. What we find is that higher wages reduce fertility the most when complementarities

in the production of human capital are higher.

In contrast with Niger and Pakistan, there are other poor countries like Burundi, Togo and Malawi, which exhibit very low TFP levels, but whose schooling levels that are quite close to the sample average. The model rationalizes these countries as having relatively more complementarity in human capital production: children stay in school longer in these poorer countries because early education spending somewhat increases the productivity of later investments by more than in other similarly low-TFP countries. Under the TFP counterfactual effects are quite large for these countries: fertility drops and schooling increases, both in large magnitudes. These large schooling increases result in substantial gains in average schooling levels under the counterfactual. In fact, raising TFP in all countries results in the largest average schooling increase of all counterfactuals (19.9%), but it also results in a slight increase of the standard deviation (1.7%). The latter result is surprising, but it reflects the heterogeneity in human capital production across countries, particularly those with low TFP levels. This suggests that local, rather than global TFP counterfactual experiments may be more informative for policy analysis.

The third main insight of the paper regards to the role of public education provision in explaining the international quality-quantity trade-off. 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 while differences in the number of years of public education provision (extensive margin) is important to explain the schooling dispersion across countries, differences in public spending per pupil are relatively more important to explain the dispersion of fertility. First, when we equate the number of years of public provision to 13.8 years in all countries, the 90th percentile value, the dispersion of schooling is reduced significantly, by 38.9%, but the effect on fertility rates is minor. According to the model, children in most low schooling countries fully take advantage of this extended period of public education except in certain countries, such as Niger, where human capital accumulation is particularly inefficient. Second, when we equate public spending per pupil in all countries to its 90th percentile value, the dispersion of fertility is reduced by 22.6% while the dispersion of schooling falls slightly. Overall, large educational subsidies, even if only for a few years, induce further parental investment in education and longer schooling duration due to the complementary properties of the human capital production function, an effect that is stronger among poorer countries. Enhanced human capital increases the time cost of children and reduces fertility. School dispersion does not decrease much due to the strong effect of this policy on the schooling of poor countries with efficient human capital accumulation. The main insight of this counterfactual is that providing higher levels of educational subsidies, even for a limited number of years, can become an effective way of reducing fertility levels, specially for countries that can better transform this spending into higher human capital.

The fourth main insight refers to the role of mortality. According to our analysis, reducing working-age and retiree mortality rates, to their respective 10th percentile values, is quantitatively

more important in reducing cross-country differences in schooling and fertility than reducing child mortality. We find that decreasing mortality rates for those between ages 5 and 65 in all countries results in a drop of 13.6% of the schooling dispersion, and a drop of 14.6% of the fertility dispersion. In this case, a higher probability of surviving during working years increase the incentives to remaining in school longer and invest in education, an effect that is stronger for countries with higher mortality rates.

More interestingly, we find that reducing the mortality rates of retirees (above age 65) reduces the standard deviation of schooling by 9.9% and that of fertility by 19.7%. The latter result is particularly intriguing and underscores the role of financial frictions in our model. If retirees live longer, this raises welfare and the marginal benefit of having children, but it also raises the value of schooling since living longer requires financing more years of consumption. In the presence of financial frictions the second effect dominates, increasing human capital and the time cost of raising children, and reducing fertility.³ We also find that since the most important differences in mortality rates across countries are among adults and elders, reducing mortality rates for children under five results in only marginal changes to fertility and schooling. This result contrasts with the traditional emphasis demographers have placed on the role of child mortality in fertility choices.

Our final insight refers to the sources of cross-country income differences. We find a strong role for public education variables, a lesser role for mortality variables and a less clear role for productivity variables. Equating public educational variables to its 90th percentile results in a 44% reduction of the cross-sectional standard deviation of log output, while equating mortality rates to its 10th percentile results in a 9% reduction. Surprisingly, equating TFP to its 90th percentile increases this standard deviation. This result echoes the one found for the dispersion of schooling. While TFP differences is an important factor to explain why many countries are poor or rich in our sample, there are many countries for which differences in human capital technologies are more important. For the latter group of countries, a large increase in TFP interacts with the human capital technology, amplifying the TFP effect. The lesson here is not that TFP is not a culprit of income differences but that due to non-linearities, heterogeneity, and potential amplification effects, more local or alternative experiments could be more informative about the role of TFP. For example, we conduct an alternative experiment in which all variables are equated across countries except for TFP. This experiment avoids compounding and amplification effects but it is less informative for policy analysis because it does not allow for interactions with other local characteristics. In this case we find that TFP can produce up to 66% of the dispersion observed in the data.

In addition to the quantitative findings on the international quantity-quality trade-off, our paper also contributes to the application of innovative theoretical frameworks to cross-country comparisons. The theory in the paper builds on the altruistic fertility model of Barro and Becker (1988, 1989), and the cross-country mortality analysis of Becker *et al.* (2005). We generalize both of these building blocks in order to better fit the data. First, as in Cordoba and Ripoll (2018), we extend the standard altruistic Barro-Becker model to introduce the elasticity of *intergenerational* substitu-

³Our model is consistent with Hazan's (2012) evidence that higher life expectancy does not increase life-time labor supply, since retirement is exogenous. But contrary to Hazan's model, increases in life expectancy do increase human capital. The key reason is the presence of credit frictions, which are absent in Hazan's model.

tion as a different parameter from the standard intertemporal substitution. The intergenerational substitution parameter controls the willingness of parents to substitute consumption across generations. As we show below, this extension allows the model to better fit the data on the economic value of a child, or the total cost of raising children, as well as the cross-country negative income-fertility relationship depicted in Figure 1. Second, as in Cordoba and Ripoll (2017), we introduce state non-separable preferences à la Epstein-Zin-Weil to disentangle mortality risk aversion from intertemporal substitution. The mortality risk aversion parameter is calibrated to match the value of statistical life. As we discuss later, disentangling mortality risk aversion from intertemporal substitution allows 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).

Another distinct feature of our theory is the explicit introduction of intergenerational financial frictions. As mentioned, these 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, 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 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 of public education policies. On this dimension, our paper extends the cross-country schooling analysis in Cordoba and Ripoll (2013) to include endogenous fertility choice.

The remainder of the paper is organized as follows. Section 2 sets up the benchmark 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. It also discusses the model’s performance and the need to introduce additional heterogeneity across countries in order to perfectly fit the data for the purpose of counterfactuals. Section 4 presents all our main results including counterfactual exercises, developing accounting, and some robustness checks. Section 5 concludes.

2 BENCHMARK 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 $\bar{s}$. 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, 2018) 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

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

where

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

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 $\Phi(\cdot)$ satisfies $\Phi(0) = 0$, $\Phi'(n) > 0$, $\Phi''(n) < 0$ and $\Phi(n) < 1$ for $n \in [0, \bar{n}]$ where $\bar{n}$ 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.⁴

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.⁵

The following are three distinct aspects of selfish utility C . First, the formulation separates

⁴Cordoba and Ripoll (2011, 2018) consider the general case $\eta \geq 0$.

⁵Cordoba and Ripoll (2017) consider the general case $\theta \geq 0$.

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 $\pi(a)$. 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.⁶ 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 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 introduced by non-market consumption $\underline{C}$ is necessary to create a 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. The presence of $\underline{C}$ breaks this proportionality. As shown below, the presence of $\underline{C} > 0$ together with the restriction $0 < \eta < 1$ allows the model to be consistent with the inverse relationship between income and fertility documented in the cross-country data.⁷

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.⁸

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 $\bar{s}$:

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

⁶ 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.

⁷ 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).

⁸ See Section 2 in Cordoba and Ripoll (2018).

We allow for cross-country differences in $\bar{s}$, e_p , as well in the relative price of education goods, p_E . As summarized in Lee and Barro (2001), there is evidence that cross-country differences in educational resources per pupil, which include higher teacher salaries and instructional materials, are important in explaining differences in student achievement. 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

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

where

$$e(a) = \frac{e_p(a) + e_s(a)}{p_E}.$$

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.

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) formulation⁹

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

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 $z_h = 1$.¹⁰ 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.

⁹See derivation and discussion in Cordoba and Ripoll (2013).

¹⁰We 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.

Complementarity captures the notion that early educational investments facilitate the productivity of later investments. Complementarity arises when

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

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 in our calibration.

The notions of self-productivity and complementarity underscore the role of parameter β in (4). The lower the value of β , or as $\beta \rightarrow 0$, the higher the degrees of self-productivity and complementarity. We exploit this insight below when we explore a version of our model in which countries differ in β . We find that differences in β are necessary to quantitatively explain cross-country schooling differences. We also find that countries known to be good at producing human capital, like Scandinavian countries, exhibit lower calibrated β s.

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

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

which are decreasing in $h(s, \mathbf{E})$ and increasing 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 $h(s, \mathbf{E}) = (d \cdot e)^\gamma s^{\gamma/\beta}$ 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

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

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

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

where w is the after-tax 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

$$b_1 \geq \int_0^s (c(a) + e_s(a)) q(a) da. \quad (7)$$

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.¹¹ 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

$$W(s, n, \mathbf{E}) + q(s) b_2 \geq \int_s^\infty c(a) q(a) da + \bar{\tau} \int_s^R q(a) da + q(F) n b'_1 + q(F + s') n b'_2, \quad (8)$$

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.

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

$$b'_2 \geq 0. \quad (9)$$

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

¹¹Others in the literature have emphasized the importance of this type of frictions in modeling intergenerational links. See Schoonbroot and Tertilt (2010, 2014) and Cordoba and Ripoll (2016).

¹²In 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

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.

2.2 Optimal allocations

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

$$V(b_1, b_2) = \max_{[c(a)]_{t=0}^{\infty}, b'_1, b'_2, [e_s(a)]_{t=0}^s, s, n \in [0, \bar{n}]} \frac{1}{1-\eta} C^{1-\eta} + \Phi(n) V(b'_1, b'_2),$$

subject to (7), (8), (4), and (9). We use superscript $*$ to denote optimal solutions and focus the presentation on steady state situations, i.e., $b_1^* = b'_1$ and $b_2^* = b'_2$. 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

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

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

where

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

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

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.

for transfers, b'_1 and b'_2 , are

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

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

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

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

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

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

and

$$G(n^*) > \frac{\pi(F) \pi(s^*)}{\pi(F + s^*)}. \quad (16)$$

In a perpetual youth model, $\pi(F) \pi(s) = \pi(F + s)$ so that, in that case, a binding transfer constraint is equivalent to $G(n^*) > 1$ according to (16). Our calibrated model below allows for higher child mortality than parental mortality so that $\pi(F + s) > \pi(F) \pi(s)$. In that case, $G > 1$ is a sufficient condition for a binding transfer constraint.

The determination of $G(n)$ is described in equation (15). It depends directly on parameters r , $\pi(F)$ and the altruistic function $\Phi(\cdot)$. More importantly, it depends directly on the fertility choice and indirectly on the parameters determining n^* . $G(n^*) > 1$ 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

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

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 $\hat{e}^*(a)$ 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 $\hat{e}(a)$ and e_p as given by,

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

Notice that during pre-school years and after age $\bar{s}$, educational investments are only private, or $e^*(a) = e_s^*(a)$. As we explain in the calibration, cross-country differences in $\bar{s}$ are quite significant, varying from as little as 11 to 22. 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

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

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

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 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

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

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)

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

and

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

where $\Omega_4(s^*, n^*)$ is a function only of s and n . Absent public education, expenditure variables such as $e_s(s)$, $c^S(s)$, b_1^* and $c^*(0)$, as well as W_s^* and W_n^* are all homogeneous of degree one in W^* . As a result, ratios $e_s(s)/W^*$, $c^S(s)/W^*$, b_1^*/W^* , $c^*(0)/W^*$, 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

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

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

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

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 $\bar{s}$. 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

$$Y = K^\alpha (AH)^{1-\alpha}, \quad (25)$$

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

$$r = \alpha \left(\frac{K}{AH} \right)^{\alpha-1}. \quad (26)$$

The after-tax wage per unit of human capital is given by

$$w = (1 - \tau)(1 - \alpha) \left(\frac{K}{AH} \right)^{\alpha} A = (1 - \tau)(1 - \alpha) (\alpha/r)^{\frac{\alpha}{1-\alpha}} A, \quad (27)$$

so that differences in w across countries reflect differences in TFP, A , as well as differences in tax rates, τ . Exogenous tax rate τ , which is different from $\bar{\tau}$, is introduced for two purposes: first, since tax rates vary across countries, we want to use after-tax labor income in the model to properly measure the actual disposable resources for representative individual in each country. Second, since TFP is computed as a residual, controlling for differences across countries in τ results in a more accurate productivity measure.

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 form¹³

$$\Phi(n) = e^{-\rho F} \pi(F)^{\frac{1-\eta}{1-\theta}} \psi(1 - e^{-\chi n}). \quad (28)$$

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.¹⁴ The last component, function $\psi(1 - e^{-\chi n})$, describes how the altruistic weight

¹³Cordoba *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 (28) because it exhibits a stronger degree of diminishing altruism, helping the model better match the fertility data.

¹⁴The reader may wonder why in the altruistic function in equation (28) 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

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 simple form

$$l(n, a) = \begin{cases} 1 - \lambda \left[(n + \kappa)^\xi - \kappa^\xi \right] & \text{if } a > F \\ 1 & \text{otherwise} \end{cases},$$

where $\lambda \left[(n + \kappa)^\xi - \kappa^\xi \right]$, with $0 < \xi < 1$, is the time cost of having n children. This time cost function implies that when $n = 0$ the cost is zero, and that the marginal time cost is decreasing on the number of children as governed by ξ .¹⁵

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, $(1 - d)$, relative price of education, p_E , tax rates, τ , and TFP.

3.1.1 Mortality

In modeling the survival probabilities we 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.

¹⁵As in Becker and Barro (1988), time costs here are meant to capture the costs over the lifetime of the parent.

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 $\pi(85)$. As can be seen in Table 1, the cross-country dispersion on survival probabilities across countries increases with age.

We obtain reasonable survival profiles for all countries. Figure 2 plots survival probabilities by age for selected countries. While the calibration underpredicts survival in earlier years and overpredicts 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 $\bar{s}$ 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). The measure of $\bar{s}$ in the data should capture the number of years a representative child in each country receives public education subsidies. We then construct $\bar{s}$ using information on the percentage of students enrolled in public education institutions at different levels, primary, secondary and tertiary, as well as the duration of each of these levels, as follows,

$$\begin{aligned}\bar{s} = & \text{duration primary} \times \text{net public enrollment primary rate} \\ & + \text{duration secondary} \times \text{net public enrollment secondary rate} \\ & + (SLE - \text{duration primary} - \text{duration secondary}) \times \text{net public enrollment tertiary net},\end{aligned}$$

where SLE is school life expectancy, our measure of schooling in the model. UNESCO measures the SLE as 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.

Our measure of $\bar{s}$ captures the extent to which children enroll in public schools. In constructing $\bar{s}$ we take into account the cases of some countries in the sample, like Belgium and Chile, where subsidies to religious private schools or the existence of universal vouchers make public school enrollment an inaccurate measure of the fraction of students who receive education subsidies. In these cases we use UNESCO's compulsory schooling to construct $\bar{s}$. Figure 3 plots our measure of $\bar{s}$

against SLE. All countries fall below the 45-degree line indicating that a representative child in each country receives public education subsidies only for a fraction of the SLE years. Figure 3 suggests a positive correlation between $\bar{s}$ and SLE and a large dispersion across countries in the provision of public education. The average $\bar{s}$ in the sample is 9.64 years, while the standard deviation is 3.26 years.

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 4 plots e_p against GDP per capita in log-10 scale, suggesting a strong correlation (88.8%).

The UNESCO measure of government educational expenditures is PPP adjusted, but does not capture differences in the relative price of education good across countries, p_E . In order to check whether differences in the relative price of education goods play a role, we proxy p_E with the 2011 relative prices of government goods from the Penn World Tables (PWT). Relative prices of education goods are available from the 1996 PWT benchmark data, but for a smaller sample of countries. Since the correlation of education and government prices in the benchmark data is high, in the order of 80%, using government prices to proxy education prices is reasonable.¹⁶ Table 1 reports summary statistics for our proxy measure of p_E . It turns out that in the data the correlation between p_E and GDP per capita is -34.6%, suggesting that education is relatively more expensive than consumption goods in poorer countries. Therefore the dispersion of education expenditures per pupil across countries is even larger than what Figure 4 suggests.

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,

¹⁶We check whether our results below are sensitive to our proxy of p_E by directly using the price of education goods from the 1996 PWT benchmark data for a subsample of countries. We verify our results are robust.

from equations (25) and (26) we can write wages in any given period as

$$w = \frac{(1 - \tau)(1 - \alpha)y^{data}}{h^{data}}, \quad (29)$$

where we define h^{data} to be

$$h^{data} = \Theta^{data} h(s, \mathbf{E}) \left(s^{data}/s \right)^{\gamma/\beta}. \quad (30)$$

In equation (29) we measure y^{data} from the World Development Indicators (2013, PPP) and calculate τ as government spending as a percent of GDP (2013, World Development Indicators) net of government education expenditures (UNESCO). As we explain below, α is set exogenously and is common across countries.

Notice that while h^{data} in (29) is the level of human capital of the workers who produced the 2013 GDP, $h(s, \mathbf{E})$ in (30) 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 (30) 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} .

Last, once wages are computed, TFP for each country obtained as a residual from (27) as,

$$A = \frac{w}{(1 - \tau)(1 - \alpha) (\alpha/r)^{\frac{\alpha}{1-\alpha}}}. \quad (31)$$

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. We set $F = 28$ as a compromise and check the robustness of the results to this value.

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 ; the level parameter of the time cost of raising children, λ ; the shift parameter of the time cost, κ ; and the elasticity of the time cost, ξ .

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 schooling in the sample. As shown above in equation (23), a lower η , or a higher intergenerational substitution, results in richer countries having a lower number of children and providing more transfers to each of them, which allows them to stay at school longer. In this respect, η influences both the level of schooling and fertility in the model. We obtain $\eta = 0.529$, which implies an elasticity of intergenerational substitution of 1.81. This is consistent with the findings in Cordoba and Ripoll (2018), 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 $\theta = 0.55$.

As discussed above, the non-homotheticity introduced by $\underline{C}$ guarantees that wages affect fertility. We calibrate $\underline{C}$ to match the correlation between income and fertility in the data. We compute this correlation to be -0.3 in our sample and obtain a calibrated $\underline{C} = 4,500$.

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 returns to schooling in the US.¹⁷ 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 $\gamma = 0.282$ and $\beta = 0.164$. 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 the Lino (2012) cost computation by adding costs of attending public colleges. We obtain $\psi = 0.54$.

Regarding χ , which determines the degree of child discounting, we select as a target the standard deviation of schooling. Since χ drives marginal altruism, it plays a role in determining both fertility and schooling. In this respect both η and χ are determinants of first and second moments of the fertility and schooling distributions. We obtain $\chi = 2.36$, although the calibrated model cannot fully account for the standard deviation of schooling.

Last are the parameters in the time cost of children function, $\lambda \left[(n + \kappa)^\xi - \kappa^\xi \right]$. Parameter λ is a key determinant of the level of fertility, so we calibrate it to match average fertility in the sample. We obtain $\lambda = 2.9$. Parameters κ and ξ are calibrated to other moments of the distribution of schooling and fertility: the minimum fertility and the maximum schooling in the sample. We obtain $\kappa = 1.84$ and $\xi = 0.2$.

3.3 Model's fit

As can be seen in Table 3, the calibrated benchmark model is able to match most targets quite well, but it cannot fully account for the standard deviation of schooling as well as the minimum fertility and the maximum schooling. Figure 5 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 78.8% for fertility and 82.1% for schooling. Similarly, the benchmark model

¹⁷We 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%.

explains 81.6% of the standard deviation of schooling and 77.7% of the fertility dispersion.

Table 4 reports untargeted moments to evaluate the model’s performance along other dimensions. 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. At the same time, and as seen in Figure 5, the model predicts more schooling than the minimum observed and less fertility than the maximum observed. Last, the model predicts that the time cost of raising children is within the range of that in the data. 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 19 and 32% of family lifetime income in the US.¹⁸ The calibrated model implies this number to be 25.4%, within the interval in the data.

3.4 Full model

For the purpose of counterfactual exercises we introduce additional cross-country heterogeneity so that the model exactly matches schooling and fertility for each country in the sample. Specifically, we extend our benchmark model by introducing cross-country heterogeneity in χ to exactly match fertility, and in β to exactly match schooling for each country in the sample. We call this the full model. We maintain the same calibrated parameters of Tables 2 and 3, and we compute the deviations of χ and β from the calibrated benchmark that would be needed for each country to exactly fit the data. Specifically, for a country i , $\chi_i = \chi_{benchmark} \times \underline{\chi}_i$ and similarly $\beta_i = \beta_{benchmark} \times \underline{\beta}_i$. It is important to emphasize that these χ_i s and β_i s are introduced to explain the residual difference between the benchmark model and the data, and can only be interpreted in this context.¹⁹

To understand the role of country-specific χ_i recall that this parameter plays an important role in determining the marginal benefit of children in equation (23). Therefore, countries for which the benchmark model predicts lower fertility than in the data in Figure 5, such as the US, Switzerland, Australia, Mexico, Pakistan, Niger, and Mali, among others, would require $\underline{\chi}_i < 1$ so that the marginal benefit of a additional child does not fall as fast and the model is able to predict a larger number of children exactly as in the data. In contrast, countries like Moldova, Bangladesh and Lesotho require $\underline{\chi}_i > 1$.

The country-specific β_i s introduce differences in the human capital production function across countries. As discussed above, countries with lower β_i would be better at producing human capital,

¹⁸As 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 supervising, 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 19 to 32% of family lifetime income in the US.

¹⁹It is not possible to exactly match the data by introducing cross-country heterogeneity in parameters that would only shift the human capital and altruistic functions. For instance, cross-country differences in the level parameter of the altruistic function, ψ , in equation (28) do not deliver an exact match of the fertility data. Similarly, introducing a level parameter in the human capital production function in (4) to capture overall efficiency is not enough to match the residual schooling data. This is the case because this level effect would be subsumed by TFP.

as every extra year of schooling would raise human capital by more than in countries with higher β_i . Recall that countries with lower β_i exhibit a higher degree of complementarity and self-productivity in human capital production in the sense of Cunha *et al.* (2006). Therefore, countries for which the benchmark model predicts lower schooling than in the data in Figure 5, such as Belgium, Denmark, Finland, Norway, Bolivia, Burundi, and Togo, among others, would require $\beta_i < 1$ so that by raising the degree of complementarity in human capital production, the model is able to correctly predict schooling as in the data. In contrast, countries like Switzerland, Thailand, Panama and Peru require $\beta_i > 1$.

The full model with country-specific χ_i s and β_i s has reasonable implications in a number of dimensions and provides interesting insights. For instance, Figure 6 plots the returns to schooling implied by the full model, which vary between about 13% and 6% and are decreasing in SLE, all reasonable features. Figure 7 plots the measure of human capital $h(s, E)$ in the full model relative to the benchmark model (in log 10 scale). The figure underscores the relatively higher degree of complementarity in educational investments among some of the rich countries. The largest upward adjustments in the human capital measures are observed in Australia, New Zealand, Belgium, Ireland, Iceland, Great Britain, Netherlands, and all the Scandinavian countries, which are particularly known for their high-quality educational systems. For the rest of the countries in the sample, the full model suggests important differences in human capital production functions, even among countries with similar per capita income. In other words, even among countries with similar educational inputs, differences in their complementarity in the production of skills result in different schooling and human capital levels.

More interestingly, the introduction of differences in human capital production functions across countries in the full model results in a re-evaluation of cross-country TFP differences.²⁰ Figure 8 displays TFP in the full model relative to the benchmark (in log 10 scale). TFP is adjusted downwards precisely in those countries in which human capital was adjusted upwards in Figure 7, and the other way around. This is the case because in the absence of country-specific β_i s (benchmark model) any differences in human capital production technologies are attributed to differences in TFP. We find that while in the benchmark model the correlation between TFP and GDP per capita is 85%, it is 25% in the full model. Most the drop of this correlation is explained by the fact that human capital in richer countries is substantially higher in the full model than in the benchmark.

Table 5 reports some moments implied by the full model. Except for the time cost of raising children, all other moments are targeted in the calibration (see Table 3). Overall the full model matches these moments quite well.

4 RESULTS

In this section we use our model to evaluate the role of different exogenous variables in explaining the world dispersion of schooling, fertility and per capita income. As we show, these counterfac-

²⁰Notice that the presence of country-specific χ_i s and β_i s requires re-calculating wages in (29) for each country.

tual exercises uncover the main model's mechanisms, in particular a non-monotonic relationship between schooling and TFP. The exercises also highlight the role of heterogeneity in human capital production functions across countries.

4.1 Counterfactuals

Table 6 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 ($\bar{s}$, e_p and d) correspond to the 90th percentile of the sample, and whose mortality rates and education prices (p_1 , p_2 , p_3 and p_E) 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 ($\bar{s}$ and e_p).²¹ As we now discuss, there are four main insights from Table 6. First, the number of years of public education provision ($\bar{s}$) is the most important determinant of schooling dispersion, followed by working-age mortality rates (p_2). Second, TFP (wages) is the most important determinant of the international dispersion of fertility, followed by public education spending per pupil (e_p) and retiree mortality rates (p_3). Third, when taken as a group, mortality rates (p_1 , p_2 , and p_3) are the second most important determinant of schooling dispersion. Last, the second most important determinant of fertility dispersion is almost a tie between mortality rates as a group (p_1 , p_2 , and p_3), and education policies ($\bar{s}$ and e_p).

4.1.1 TFP

Under the counterfactual that equates TFP in every country to the 90th percentile value, mean schooling increases while mean fertility decreases. In fact, as seen in Table 6, these changes in the means of schooling and fertility are the largest of any counterfactual: mean schooling increases by 19.9% and fertility decreases by 36.8%. In addition, the reduction of the standard deviation of fertility is the largest among all counterfactuals, in the order of 55.7%, while the standard deviation of schooling slightly increases by 1.7%.

To better understand these results, Figure 9 plots the predicted fertility and schooling in the model under the TFP counterfactual. Recall that the full model is calibrated to exactly match fertility and schooling for each country, so all the observations in Figure 9 can be interpreted as deviations from the plotted 45-degree line (calibrated model). Fertility drops in all countries, with particularly large falls in countries like Burundi, Malawi, Uganda, Bolivia, Kenya and Ghana. These countries exhibit particularly low TFP (and wages) in the calibrated full model, so raising TFP significantly increases the time cost of raising children, resulting in lower fertility. Interestingly, these are also the countries for which schooling increases the most under the counterfactual, a strong quantity-quality trade-off. Notice also from Figure 7 that these are among the poor countries for

²¹It 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 results.

which human capital is adjusted upwards in the full model, and from Figure 8 that their TFP is adjusted down. In other words, among the poor, these are countries that exhibit relatively lower TFP and higher complementarity in the production of human capital. The message of this exercise is very clear and suggests an important qualification of the common view that development in the form of higher wages is the most important tool to reduce fertility. What we find is that a high degree of complementarity of educational investments is key to reinforce the effect of higher wages on the reduction of fertility among poor countries.

Notice that the standard deviation of schooling increases slightly under the TFP counterfactual (1.7%). Although puzzling, this occurs precisely because of the heterogeneous effect among poor countries. While schooling increases substantially in countries with higher complementarity in human capital production, it increases relatively less in those countries in which educational investments have a lower impact in skill formation. This result suggests the importance of understanding differences in the human capital production functions among poorer countries.

The muted effect of cross-country TFP differences in explaining the dispersion of schooling should be interpreted carefully. As we now turn to explain, our model features a number of compounded and non-linear effects which alter aggregate measures such as the standard deviation, but hide interesting insights regarding the model's mechanisms.

4.1.2 Non-monotonic schooling and model's mechanisms

Before turning to the counterfactuals regarding public education policies, and to better understand the role of TFP in our model, we now report two exercises to uncover some of the model's mechanisms. Figure 10 plots the relationship between fertility and TFP (left panel) and schooling and TFP (right panel). The figure is constructed by using the model's calibration of the US economy and varying TFP across wide range of values. Notably, mortality rates, public education policies, β_i and χ_i are at their US values in the calibrated full model. The left panel shows that fertility is monotonically decreasing in TFP, with fertility dropping sharply at low TFP levels, and flattening out at higher TFP levels. The main mechanism at work is the time cost of raising children, which increases with wages.

In contrast with the monotonic behavior of fertility, schooling is non-monotonic in TFP. At very low levels of TFP schooling is highest, then it drops sharply and it finally increases monotonically with TFP. The reason for this monotonic behavior can be traced to the public education policies in the US, which are kept at their 2013 levels. In the US, public education spending per pupil is $e_p = \$15,040$ and years of public provision are $\bar{s} = 12.94$. At very low levels of TFP, it is optimal to send children to school to receive the comparatively high education subsidy. 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. Notice that optimal schooling at very low TFP levels is 22 years, which means students attend school beyond the years of the public subsidy coverage. Parents find it optimal to

finance private education spending beyond $\bar{s}$.

As TFP levels start increasing, the trade-off between staying at school with high education subsidies and cutting school shorter to start working changes in favor of the latter. As can be seen in Figure 10, schooling sharply drops to levels below 16 years. Notice that even at this level, students are still enjoying the benefit of the high levels of public spending for the full length of the subsidy (12.94 years). The main change is that parents are now financing fewer years of fully private education spending. Finally, the last portion of the schooling graph is more standard, with schooling increasing in TFP.

The main message of Figure 10 is that non-monotonicities require caution in interpreting counterfactuals that involve equalization of TFP levels across countries. In particular, in the presence of financial constraints, schooling choice interacts with both education spending and TFP in intricate ways. To further understand this point in the context of cross-country analysis, we perform an additional counterfactual exercise displayed in Figure 11.

The exercise in Figure 11 contrasts with that of Figure 9 in that rather than equating TFP across countries, the counterfactual in Figure 11 equates all country-specific parameters to that of the artificial "rich country", except for TFP. In other words, all countries have $\bar{s}$, e_p and d equal to the 90th percentile of the sample; p_1 , p_2 , p_3 and p_E equal to the 10th percentile; and their own TFP level. Although this counterfactual is not as informative for policy analysis, it eliminates some of the compounding effects present in Figure 9. What Figure 11 implies is that if countries only differed in TFP levels, one could still explain about 66% of the schooling dispersion observed in the data. What is remarkable is the heterogeneity among countries whose years of schooling in the data is below the 13.86 sample average. In a first group of countries, including Burundi, Mozambique, Togo, Malawi and Uganda, schooling choice is the highest. This is the case because even though TFP is very low there, if government education subsidies were as high as \$17,179 in all these countries, students would optimally choose to stay at school longer. This intuition parallels the mechanism explaining the behavior of schooling at very low TFP levels in Figure 10.

As seen on the right panel of Figure 11, there is a second group of countries, including Niger, Mali and Pakistan, where schooling increases under the counterfactual, but not nearly as much as in the first group. TFP in this second group of countries is relatively higher than in the first. Students in these countries stay at school fewer years because even though they enjoy a high education subsidy under the counterfactual, wages are not as low as in the first group of countries. These countries can be rationalized by the decreasing portion of the schooling function on the right panel of Figure 10.

In addition to clarifying the relationship between schooling and TFP in the model, Figures 10 and 11 also shed light on the relationship between public education variables in the model ($\bar{s}$ and e_p) in the model, as we now turn to explain.

4.1.3 Public schooling policies

One of the contributions of this paper is to use a microfounded model to evaluate the notion that education is the best tool to lower fertility rates in the developing world. Our model allows us to separately analyze the role of differences in the amount of public education subsidies per pupil (intensive margin) as well as the number of years of public school provision (extensive margin). First, as seen in Table 6, equating the amount of the public education subsidy per pupil e_p to the 90th percentile value results in an increase of the schooling mean, and a decrease of the fertility mean as well as the standard deviations of schooling and fertility. The largest effect of this counterfactual is the reduction of the standard deviation of fertility by 22.6%. Recall from Table 1 that the mean e_p in the sample is \$6,601, while the median is \$3,952. The 90th percentile value in the sample is \$17,179. Everything else equal, an increase in e_p means that students can accumulate more human capital for each year they stay at school. For most countries this results in increased years of schooling and lower fertility rates, as shown in Figure 12.

Most of the action of this counterfactual comes from poorer countries, which exhibit the largest drops in fertility and increases in schooling. In fact, in some poorer countries schooling increases so much as to hit the constraint $s < F$. Notice that some of these are among the poorer countries with relatively higher degree of complementarity in educational investments (Burundi, Togo, Malawi, and Uganda). Again, this can be seen in Figure 7 from the set of poor countries for which the full model adjusts human capital upwards. The main message from this counterfactual is that even if governments provide educational subsidies for only a few years, investing more resources per pupil is the key to drop fertility levels and increase schooling years. This is particularly true for poorer countries in which educational investments exhibit higher complementarity in the production of skills. Notice that under the counterfactual the government's budget is balanced, so these higher educational resources are financed through lump-sum taxation.

Turning now to the extensive margin of public education provision, equating the number of years of public education provision ($\bar{s}$) to the 90th percentile level has similar qualitative effects as equating the amount of the subsidy (e_p), but very different quantitative effects. The average $\bar{s}$ in the sample is 9.6 years, the maximum value is 16.5 years (Ireland), and the 90th percentile of $\bar{s}$ is 13.8 years. The largest quantitative effect of this counterfactual is the drop of the standard deviation of schooling by 38.9%. In fact, this is the counterfactual that generates the largest drop in schooling dispersion. Most of the action in this counterfactual comes from schooling in poorer countries. In fact, as seen in Figure 13, schooling remains virtually unchanged in high-schooling countries, while for almost all low-schooling countries schooling is at least equal to the new $\bar{s}$ of 13.8 years. In the latter countries students take full advantage of the larger $\bar{s}$ and stay at school for at least that long, resulting in a large decrease of the standard deviation of schooling.

Perhaps the most interesting insight of the $\bar{s}$ counterfactual is that fertility drops little in poor countries (Table 6). Recall that in these countries governments tend to offer modest education subsidies per pupil (e_p). The counterfactual suggests that keeping public schools open for more grades while spending little per pupil may not decrease fertility levels. Parents under this counterfactual send their children to school as long as the government offers the subsidy, but not beyond that point.

They take advantage of the small education subsidy and finance their children’s consumption for more schooling years, but this does not provide enough incentives for them to reduce the number of children. This counterfactual echoes the concern in policy circles that extending the provision of public education may not result in decreased fertility levels in developing countries. But as we now turn to explain, the effective tool to decrease fertility levels is to improve the provision of public education in both the extensive and intensive margins. The extensive margin alone is not enough.

In fact, when both e_p and $\bar{s}$ are equalized to their respective 90th percentile values across countries, there is a large drop of 39.6% of the standard deviation of fertility (Table 6). In addition, a nontrivial increase in mean schooling is observed (13.8%), together with a drop in mean fertility (17.7%). Fertility rates fall below 4 children in all poorer countries. As indicated above, it is the simultaneous increase in the duration and the magnitude of the public education subsidy what is key to reducing the mean and standard deviation of fertility across countries. The increase in human capital achieved through public schools that offer not only more grade levels, but that also invest more per pupil, are large enough to induce drops in fertility rates. The channels at work here are the larger time costs of raising children for those with high human capital, together with the larger goods costs generated by the additional private education spending per child.

All in all, this counterfactual underscores the importance of public education provision in explaining the cross-country dispersion of fertility and schooling, an insight that has been widely discussed in policy circles, but not evaluated in a carefully calibrated micro-founded model. The main message is that improving the quality of schools in poorer countries, as measured by spending per pupil, is key to decrease fertility levels.

Last, Table 6 also reports the counterfactuals for school non-repeating rates (d) and the relative price of education (p_E). We find that these do not have a quantitative significant effect on the dispersion of either schooling or fertility.

4.1.4 Mortality rates

As can be seen in Table 6, equating each of the mortality rates to the 10th percentile value has the same qualitative effect: mean schooling increases, while mean fertility and the standard deviations of fertility and schooling decrease. Higher survival rates imply a higher working-life span, increasing the incentives to remain in school longer. In the steady-state equilibrium longer schooling years requires a longer period for parents to finance consumption and education spending for their children. The quantity-quality channel is again present here, resulting in lower fertility rates.

Notice from Table 6 that the decrease in students/ workers’ mortality (p_2) has the largest effect on the standard deviation of schooling, while it is the decrease in retiree’s mortality (p_3) that has the largest effect on the dispersion of fertility. The first result is intuitive –the higher the probability of surviving the working years, the higher the incentives to remain in school longer, which will have a larger impact among the poor countries. The second result is more intriguing, but it captures interesting mechanisms in the model. The direct effect of a decrease in p_3 is to increase the length of life, increasing both welfare and the marginal benefit of having children. However, the marginal

cost of children also increases due to the presence of the binding transfer constrained captured by $G > 1$. Holding s , n and $e(a)$ constant, an decrease in p_3 implies lower consumption for workers and retirees because the same lifetime income needs to be spread over a longer lifetime. Lower c^W implies lower G and therefore a higher marginal benefit of going to school and spending more in education. As a result, $h(s)$ increases and the time cost of raising children increases. This higher marginal cost of raising children dominates, inducing fertility to fall and schooling, expenditures in education, and $h(s)$ to further increase. This is a novel channel, mostly absent in the literature on fertility. Demographers have traditionally emphasized the role of child mortality (p_1) on fertility choice, while our channel refers to retiree mortality (p_3). The novel insight of our model in this respect is that higher survival at elder ages induces both higher schooling and early investments in human capital, as well as lower fertility. This channel is absent in models without credit frictions, but it arises naturally in models in which children financially depend on their parents and parents cannot leave debt obligations to them.

Another notable result from Table 6 is that children's mortality has a relatively small effect on the dispersion of schooling and fertility. This is not surprising to the extent that in 2013 the cross-country dispersion of children's mortality was relatively smaller (Table 1).

Figure 14 portrays the case in which all mortality rates, for children, workers and retirees, are equated to their respective 90th percentile values. As seen in the figure, fertility decreases across the board, with the drop being higher in poorer than in richer countries, resulting in a fall of the standard deviation of 38.5%, a quantitatively strong effect (Table 6). The standard deviation of schooling also falls significantly, by 25.4%. Figure 14 shows that while schooling is mostly unchanged for rich countries, it increases for all poorer countries. Altogether, the mortality counterfactuals largely support the view held by demographers on the importance of mortality in explaining differences in schooling and fertility across countries, although with the qualification that adult mortality has become quantitatively more important in recent years.

4.2 Development accounting

The discussion of counterfactual exercises has so far been centered on the international quantity-quality trade-off. But as reported in Table 6, our model has also implications for development accounting, providing insights about the sources of cross-country per capita income differences. The main message from Table 6 is that the largest drops in the standard deviation of per capita income are attributed to the public education variables. Equating e_p across countries to the 90th percentile value results in a drop of 26.6% in the standard deviation of per capita income. If both e_p and $\bar{s}$ are equalized, then the drop is 44%.

Compared to public education policies, mortality plays a quantitatively lesser role. Equating all p_1 , p_2 , and p_3 across countries results in a drop of the standard deviation of per capita income of 9.2%. But what is perhaps more surprising is that equating TFP to its 90th percentile increases this standard deviation by 11.4%. Although surprising relative to the development accounting literature, this result echoes the one found for the dispersion of schooling (Figure 9) and highlights

the importance of non-monotonicities in our model, as well as the role of cross-country differences in human capital technologies. What this result says is that while TFP differences are important to explain why many countries are poor or rich in our sample, there are also many countries for which differences in human capital technologies are more important. As explained above, raising TFP will lead to higher gains in years of schooling in those countries with higher degree of complementarity in human capital production. For these countries, the human capital technology amplifies the TFP effect. But there are other similarly poor countries for which this amplification does not occur due to the lower complementarity in the production of human capital.

The main lesson from the development accounting exercises is not that TFP is unimportant, but that the efficiency in the production of human capital among similarly poor countries mediates the ultimate effects of a TFP increase on cross-country income differences. Even though we have so far emphasized differences in human capital production functions among poor countries, our paper also documents important differences among the rich. For example, our full model implies a higher degree of complementarity in human capital production in Scandinavian countries, and a lower one in countries such as Spain and Italy. These differences also result in varying responses to increases in TFP. In this respect, a message from our development accounting exercise is the caution that should be taken in interpreting global TFP counterfactuals, and the need to better understand the differences in the efficiency of output production and human capital production across countries.

4.3 Robustness analysis

In this section we report some robustness checks for our results. We concentrate on two relevant features of our full model. One is the role of the country-specific χ_i s and β_i s, and the other is the disentangling of θ from σ in the utility function.

4.3.1 Counterfactuals under the benchmark model

Table 7 repeats the full set of counterfactuals in Table 6, but rather than using our full model, it uses the benchmark. Recall that under the benchmark model human capital production parameter β and altruistic parameter χ are the same across countries. Recall also that it is the full model the one that provides a more accurate decomposition of the mechanisms at work, since the benchmark does not perfectly fit the schooling and fertility data.

Except for the TFP counterfactual, Tables 6 and 7 are by in large very similar. In fact, the main difference is on the effect of TFP on the dispersion of schooling and per capita income, as the effect on fertility is also similar in both tables. Two conclusions can be drawn from this comparison. First, introducing differences across countries in altruism (χ) does not have major quantitative effects on the model's predictions. Second, differences in human capital production parameter β do alter the qualitative properties of the model because they induce a re-evaluation of both human capital stocks and TFP levels, both of which are unobservable variables. As shown in Figure 6, the returns to schooling in the full model are quite reasonable, and as shown in Figures 7 and 8, the full model does change our measurement of human capital and TFP.

The comparison of Tables 6 and 7 suggest the need to more systematically explore the differences in human capital productions across countries, a task we leave for future research. We introduced enough exogenous variation across countries in the form of mortality rates and public education policies, but these together with the implied TFP differences were not enough to fully explain the observed schooling dispersion. While our β_i s rationalize the residual differences, a pressing issue is to further investigate the origin of these differences. As the task of the development accounting literature has been in part to unbundle the large cross-country TFP differences, what we learn from our analysis is that differences in human capital production functions induce not only a re-evaluation of TFP levels, but also interesting trade-offs on the schooling and fertility choices of individuals.

4.3.2 Other robustness

As explained above, our model provides novel insights into the role of retiree mortality (p_3), particularly in explaining fertility differences. One may wonder the extent to which our utility function specification drives this result, and any other results regarding the role of mortality variables. Recall that our utility function disentangles the mortality aversion coefficient, θ , from the elasticity of intertemporal substitution $1/\sigma$. This contrasts with the standard expected utility version that assumes $\theta = \sigma$.

We tested the effect of disentangling θ from σ by exogenously increasing θ to make it closer to $\sigma = 1$. Of course this requires omitting the VSL as one of the calibration targets. We find that as $\theta \rightarrow 1$ the effect of retiree mortality under the counterfactual exercises becomes even stronger. In fact, as $\theta \rightarrow 1$ the VSL in the model explodes to infinity. For instance, raising θ from the calibrated benchmark of 0.55 to $\theta = 0.75$ implies a VSL of \$7 million in the US, a value at the high end of available estimates. This higher VSL in the presence of financial frictions amplifies the effect of living longer on fertility choices. We conclude that our results regarding the importance of adult mortality are not exaggerated. Quite the opposite, they are quantitatively important without implying an unreasonably high VSL.

5 CONCLUDING COMMENTS

International data reveals a persistent and significant dispersion of schooling and fertility rates across countries and an overall quantity-quality trade-off: one more child per women is associated with around three fewer years of schooling. We investigate the underlying determinants of this trade-off. Our model incorporates some of the key mechanisms behind schooling and fertility decisions stressed by the existing literature into a single unified theory. We incorporate enough heterogeneity into the model as to exactly match the data using a carefully calibrated version of the model. The result is a more complete and compelling decomposition of the underlying forces behind schooling, fertility and income differences.

Four takeaways from the analysis can be highlighted. First, the notion that development and

the associated higher wages are the best recipe to lower fertility levels is plausible, but it requires the qualification that this channel works best in countries in which educational investments are more complementary in producing human capital. In other words, the quantity-quality trade-off is stronger in countries that are better able to use educational resources to generate higher human capital.

The second takeaway is that extending the length of public education, as has been seen with the increase in 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. This takes us to the third takeaway of the analysis: even if for a limited number of years, raising public educational resources per pupil in poorer countries could unleash a virtuous cycle, particularly for countries with higher complementarity in the human capital production function. In these countries, higher public education subsidies result not only in lower fertility, but also in higher complementary private educational investments, and higher schooling attainment.

The fourth takeaway is that while demographers have emphasized the role of child mortality in fertility choice, we find that increases in retiree survival rates are more important. For one, by 2013 the cross-country dispersion of child mortality is lower than that of elderly mortality. What we learn from the model is that living longer requires spreading consumption over more years, which raises the incentives of increasing schooling and human capital, raises the time cost of having children, and lowers fertility rates.

Our 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.

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] Bils, M., Klenow, P. (2000). "Does Schooling Cause Growth?" *American Economic Review*, 90, 1160-1183.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] Caselli, F. and Feyrer, J. (2007). "The marginal product of capital," *Quarterly Journal of Economics*, 122, 537-568.
- [12] Cordoba, J. C., Liu, X. and Ripoll, M. (2016). "Endogenous fertility, social mobility and long-run inequality," *Journal of Monetary Economics*, 77, 103-124.
- [13] Cordoba, J. C. and Ripoll, M. (2018). "The elasticity of intergenerational substitution, parental altruism and fertility choice," *Review of Economic Studies*, forthcoming.
- [14] Cordoba, J. C. and Ripoll, M. (2017). "Risk aversion and the value of life," *Review of Economic Studies*, 84, 1472-1509.
- [15] Cordoba, J. C. and Ripoll, M. (2016). "Intergenerational transfers and the fertility-income relationship," *Economic Journal*, 126, 949-977.
- [16] Cordoba, J. C. and Ripoll, M. (2013). "What explains schooling differences across countries?" *Journal of Monetary Economics*, 60, 184-202.
- [17] Cordoba, J. C. and Ripoll, M. (2011). "A contribution to the economic theory of fertility," Manuscript. Iowa State University Economics Working Papers No. 11011
- [18] 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.
- [19] De la Croix, D. and Doepke, M. (2004). "Public versus private education when differential fertility matters," *Journal of Development Economics*, 73, 607-629.

- [20] De la Croix, D. and Licandro, O. (1999). "Life expectancy and endogenous growth," *Economic Letters*, 65, 255-263.
- [21] Doepke, M. (2004). "Accounting for Fertility Decline During the Transition to Growth," *Journal of Economic Growth*, 9, 347-383.
- [22] Galor, O., and Weil, D. (1996). "The Gender Gap, Fertility and Growth," *American Economic Review*, 86, 374-387.
- [23] 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.
- [24] Hazan, M. (2012). "Life Expectancy and Schooling: New Insights from Cross-Country Data," *Journal of Population Economics*, 25, 1237-1248.
- [25] 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.
- [26] Lee, J, and Barro, R. (2001). "Schooling quality in a cross-section of countries," *Economica*, 68, 465-488.
- [27] Lino, M. (2012). "Expenditures on Children by Families, 2011", *USDA Miscellaneous Publication Number 1528-2011*.
- [28] Manuelli, R. and Seshadri, A. (2009). "Explaining international fertility differences," *Quarterly Journal of Economics*, 124, 771-807.
- [29] 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.
- [30] Rangazas, P. (2000). "Schooling and economic growth: A King-Rebelo experiment with human capital," *Journal of Monetary Economics*, 46, 397-416.
- [31] Schoellman, T. (2012). "Education quality and development accounting," *Review of Economic Studies*, 79, 388-417.
- [32] Schoonbroodt, A. and Tertilt, M. (2014). "Property rights and efficiency in OLG models with endogenous fertility," *Journal of Economic Theory*, 150, 551-582.
- [33] Schoonbroodt, A. and Tertilt, M. (2010). "Who owns children and does it matter?" *NBER Working Paper # 15663*.
- [34] United Nations (2015). *World Fertility Patterns 2015* partment of Economic and Social Affairs, Population Division. - Data Booklet. De

- [35] 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.
- [36] 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.
- [37] World Bank Group (2011). *Learning for All: Investing in 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:

$$V(b_1, b_2) = \max_{[c(a)]_{t=0}^\infty, [e_s(a)]_{t=0}^s, b'_1, b'_2, s, n} \frac{1}{1-\eta} C^{1-\eta} + \Phi(n) V(b'_1, b'_2) \quad (1)$$

where

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

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

The maximization is subject to the following constraints:

$$b_1 \geq \int_0^s (c(a) + e_s(a)) q(a) da, \quad (3)$$

$$q(s)b_2 + W(s, n, \mathbf{E}) \geq \bar{\tau} \int_s^R q(a) da + \int_s^\infty c(a) q(a) da + q(F) nb'_1 + q(F + s') nb'_2, \quad (4)$$

$$\begin{aligned} \left(\int_0^s (d(e_p(a) + e_s(a)) / p_E)^\beta da \right)^{\gamma/\beta} &\geq h(s, \mathbf{E}), \\ l(n, a) &= \begin{cases} 1 & \text{if } a \leq F \\ l(n) & \text{if } a > F \end{cases}, \text{ and} \\ b'_2 &\geq 0 \text{ and } e_s(a) \geq 0 \text{ for } a \in [0, s]. \end{aligned}$$

where

$$W(s, n, \mathbf{E}) = wh(s, \mathbf{E}) \int_s^R e^{\nu(a-s)} l(n, a) q(a) da. \quad (5)$$

Prices and survival probabilities satisfy:

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

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

The associated Lagrangian can be written as:

$$\begin{aligned} V(b_1, b_2) = & \frac{1}{1-\eta} C^{1-\eta} + \Phi(n) V(b'_1, b'_2) \\ & + \lambda_1 [b_1 - \int_0^s (c(a) + e_s(a)) q(a) da] \\ & + \lambda_2 \left[q(s) b_2 + W(s, n, \mathbf{E}) - \int_s^\infty c(a) q(a) da - q(F) n b'_1 - q(F + s') n b'_2 - \bar{\tau} \int_s^R q(a) da \right] \\ & + \lambda_3 \left[\left(\int_0^s (d(e_p(a) + e_s(a)) / p_E)^\beta da \right)^{\gamma/\beta} - h(s, \mathbf{E}) \right] + \lambda_4 e_s(s) + \lambda_5 b'_2. \end{aligned}$$

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 $n \in [0, \bar{n}]$. Use $(*)$ to denote optimal solutions. Let E^* be the present value of private expenditures in education defined as:

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

1.2 Optimal consumption

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

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

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

Using (6), these equations become:

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

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

$$G \equiv \frac{\lambda_1}{\lambda_2}. \quad (12)$$

Then

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

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

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

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

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

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

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

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

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

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

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

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

1.3 Optimal transfers

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

$$\lambda_2 q(F) n^* = \Phi(n^*) V_1(b_1^*, b_2^*),$$

$$\lambda_2 q(F + s^*) n^* = \Phi(n^*) V_2(b_1^*, b_2^*) + \lambda_5,$$

while the corresponding envelope conditions are

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

Then the optimality conditions for b'_1 and b'_2 can then be expressed as:

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

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

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

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

and

$$e^{-rF} \pi(F) \frac{n^*}{\Phi(n^*)} = G(n^*) > \frac{\pi(F) \pi(s^*)}{\pi(F + s^*)} = e^{-(p_1 - p_2)a_c}.$$

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 $b_2^{*'} = 0$. We confirm that in all our calibrations, $G(n^*) > 1$ for all countries in our sample.

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

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

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

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

where

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

1.4 Optimal human capital, schooling and school expenditures

1.4.1 Human capital

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

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

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

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

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

1.4.2 School expenditures

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

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

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

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

or

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

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

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

with

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

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

$$e^*(0) = \hat{e}^*(0). \quad (27)$$

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

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

Figure 1a 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 $\bar{s}$. Case 2 illustrates a case in which private spending includes pre-school and some college, since optimal schooling s_2 is larger than $\bar{s}$, but no private spending in the interval $[\underline{s}, \bar{s}]$, say during primary and secondary. Finally in Case 3, optimal schooling is $s_3 > \bar{s}$ but now there is also some private spending in the interval $[\underline{s}, \bar{s}]$. In the calibration, we set $\underline{s}$ to be 6.

To describe more precisely the solution for $e^*(a)$, let $\hat{s}$ be implicitly defined by the equation $\hat{e}(\hat{s}) = e_p$. Intuitively, $\hat{s}$ 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:

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

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

$$s_p^* \equiv \min \{s^*, \bar{s}, \max [\underline{s}, \hat{s}]\}. \quad (30)$$

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

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

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

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

Plugging these results into (8) one obtains:

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

where

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

1.4.3 Human capital

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

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

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:

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

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 $e^*(0)$,

$$e^*(0) = \gamma W^* / (G(n^*) \Omega_3(s^*, e_p/e^*(0))). \quad (36)$$

1.4.4 Schooling

The first order condition for s is given by

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

where, according to (5),

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

According to (35) and (34):

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

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

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

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

or

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

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

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

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 independent of w .

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

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

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):

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

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

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

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

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

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 substituting (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),

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

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

1.5 Fertility

First order condition with respect to fertility is:

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

where

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

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

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

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

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

where

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

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

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

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

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

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 equations 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 $\frac{C}{W^*}$. 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 $\bar{\tau} \int_{s^*}^R \tilde{n}(a) da$, 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

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

1.7 Steady state wage rate and human capital

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

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

The steady state after tax wage is calculated according to

$$w = (1 - \tau) (1 - \alpha) \frac{y^{data}}{h^{data}} \quad (53)$$

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

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 $E_{priv}/y(US)$. Aggregate private expenditures in education, denoted by AE , are defined as the following form but taking into account the demographics in the economy.

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

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

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

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

It comes from

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

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

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

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

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

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

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 $t = F$,

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

(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 $\log(n^*) = \alpha \log(W(s^*, n^*, \mathbf{E}^*)) + \varepsilon$.

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

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 $e_p, \bar{s}, p_1, p_2, p_3$ 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 $\hat{e}^*(s^*)$ evaluated at age s^* , can be gotten by (25), the private educational expenditure $e_s^*(a)$ follows from (32), and $\hat{s}$ is obtained by (29). After $\hat{s}$ is solved, $s_p, h(s^*, \mathbf{E}^*), h^{data}, w, \mathbf{E}^*, W(s^*, n^*, \mathbf{E}^*), b_1^*, c^*(0), c^S(s^*), c^W(s^*), C, \lambda_1, \lambda_2$ 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
Compulsory schooling years	10	10	2	15	5
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
Relative price government goods	1.17	1.09	0.47	2.93	0.54
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, compulsory 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
v	Returns to experience	2%
F	Average childbearing age	28
R	Retirement age	65

Notes: The values of parameters σ , r , α and v 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 – Benchmark model

Parameter	Concept	Target	Target value	Target in the model	Parameter value
η	Inverse of the elasticity of intergenerational substitution	World mean of schooling	13.86	13.52	0.529
θ	Mortality risk aversion	Value of statistical life at childbearing age in the US	\$4 million	\$4.1 million	0.55
$\underline{C}$	Non-market consumption	Income elasticity of fertility	-0.30	-0.32	4500
γ	Returns to scale human capital production function	Average private expenditures in education as % of GDP in OECD	0.9%	0.87%	0.282
β	Degree of substitution education spending across ages	Returns to schooling in the US	8.28%	7.48%	0.164
ψ	Level of altruism	Goods cost of raising a child as % of lifetime income in US	16.44%	18.84%	0.54
χ	Degree of child discounting	World standard deviation of schooling	3.08	2.52	2.36
λ	Level time cost of raising children	World mean of fertility	2.56	2.67	2.9
κ	Shift parameter time cost	World minimum fertility	1.12	1.44	1.84
ξ	Elasticity time cost	World maximum schooling	20.43	17.20	0.20

Notes: Most targets are computed using the sample of countries described in Table 1. The value of statistical life for the US is from Viscusi and Aldi (2003). Average OECD private educational expenditures as a % of GDP is from the National Center of Education Statistics.

TABLE 4
Model's performance – Benchmark model

Untargeted moments	Data	Model
World quantity-quality trade-off	-0.33	-0.40
World minimum schooling	5.32	7.55
World maximum fertility	7.62	5.83
World standard deviation of fertility	1.33	1.04
Time cost of raising children as % of lifetime income in US	19 to 32%	25.6%
Correlations		
Fertility in model and data = 82.1%		
Schooling in model and data = 78.8%		

Notes: Model is calibrated as in Tables 2 and 3. All data moments are computed using the sample summarized in Table 1. Time cost of raising children is computed following Cordoba and Ripoll (2016, 2018) and it corresponds to about 60% of the total cost of raising children ages 0 to 17.

TABLE 5
Moments in full model

Moments	Data	Model
Value of statistical life at childbearing age in the US	\$4 million	\$3.99 million
Income elasticity of fertility	-0.30	-0.30
Average private expenditures in education as % of GDP in OECD	0.9%	0.99%
Returns to schooling in the US	8.28%	7.95%
Goods cost of raising a child as % of lifetime income in US	16.44%	15.05%
Time cost of raising children as % of lifetime income in US	19 to 32%	31.2%

Notes: The full model is calibrated as in Tables 2 and 3, but it also includes country-specific β and χ parameters. These country-specific β and χ are computed so that the model exactly matches schooling and fertility in each country.

TABLE 6
Counterfactuals (% change) – Full model

Parameter	Mean			Standard deviation		
	Schooling	Fertility	Per capita income	Schooling	Fertility	Per capita income
<i>Changes to individual parameters</i>						
TFP	19.9	-36.8	379.2	1.7	-55.7	11.4
e_p	17.7	-11.4	19.9	-4.8	-22.6	-26.6
$\bar{s}$	8.1	-3.5	5.5	-38.9	-11.0	-10.4
p_1	0.7	-1.5	0.3	-2.9	-4.3	-0.4
p_2	6.4	-7.5	5.7	-13.6	-14.6	-7.3
p_3	4.1	-10.0	-7.3	-9.9	-19.7	-0.9
d	0.2	-0.7	0.7	-0.9	-1.8	-2.0
p_E	3.0	-5.2	13.9	-2.0	-4.2	-3.5
<i>Changes to groups of parameters</i>						
$e_p, \bar{s}$	19.7	-17.4	34.7	-11.4	-39.9	-44.0
p_1, p_2, p_3	11.6	-18.6	-1.9	-25.4	-38.5	-9.2

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 full model is calibrated as in Tables 2 and 3, but it also includes country-specific β and χ parameters. These country-specific β and χ are computed so that the model exactly matches schooling and fertility in each country. The standard deviation of per capita income is computed over the log (10 base) of income.

TABLE 7
Counterfactuals (% change) – Benchmark model

Parameter	Mean			Standard deviation		
	Schooling	Fertility	Per capita income	Schooling	Fertility	Per capita income
<i>Changes to individual parameters</i>						
TFP	12.7	-25.8	67.8	-43.6	-62.1	-73.2
e_p	22.6	-11.4	22.3	-0.8	-23.2	-28.9
$\bar{s}$	8.5	-3.2	5.9	-53.6	-10.7	-11.6
p_1	0.8	-1.3	0.3	-4.0	-4.2	-0.3
p_2	6.2	-7.0	5.8	-22.1	-15.9	-7.3
p_3	3.9	-9.6	-4.1	-12.9	-19.9	-4.8
d	0.2	-0.6	0.8	-1.1	-1.9	-2.0
p_E	3.0	-5.3	15.7	-2.0	-4.0	-2.7
<i>Changes to groups of parameters</i>						
$e_p, \bar{s}$	24.1	-16.9	38.0	0.5	-39.6	-47.1
p_1, p_2, p_3	11.3	-17.5	1.4	-40.3	-39.6	-13.7

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 benchmark model is calibrated as in Tables 2 and 3. In the benchmark model β and χ are the same across countries. The standard deviation of per capita income is computed over the log (10 base) of income.

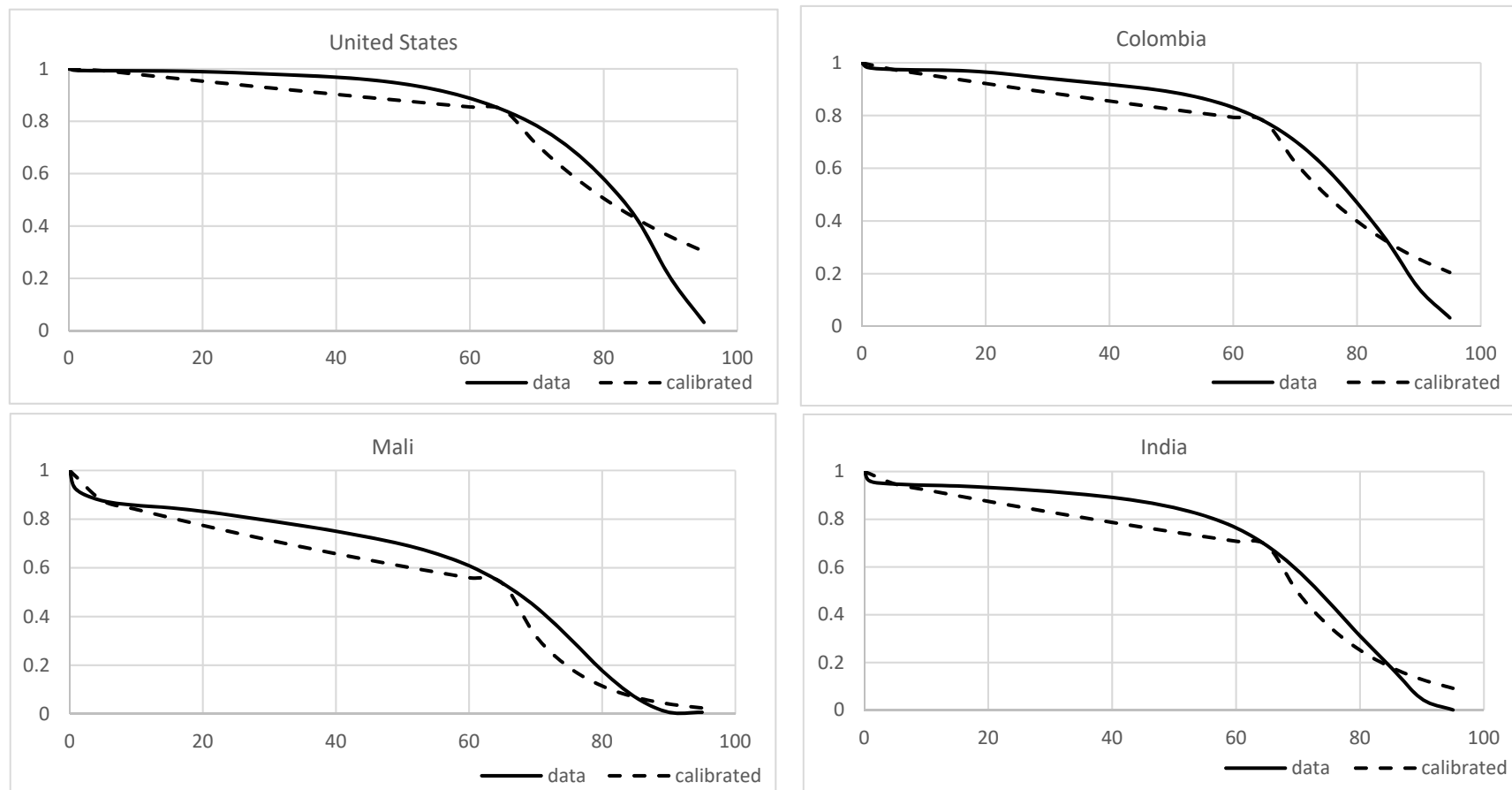


FIGURE 2
Survival probabilities by age in selected countries

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.

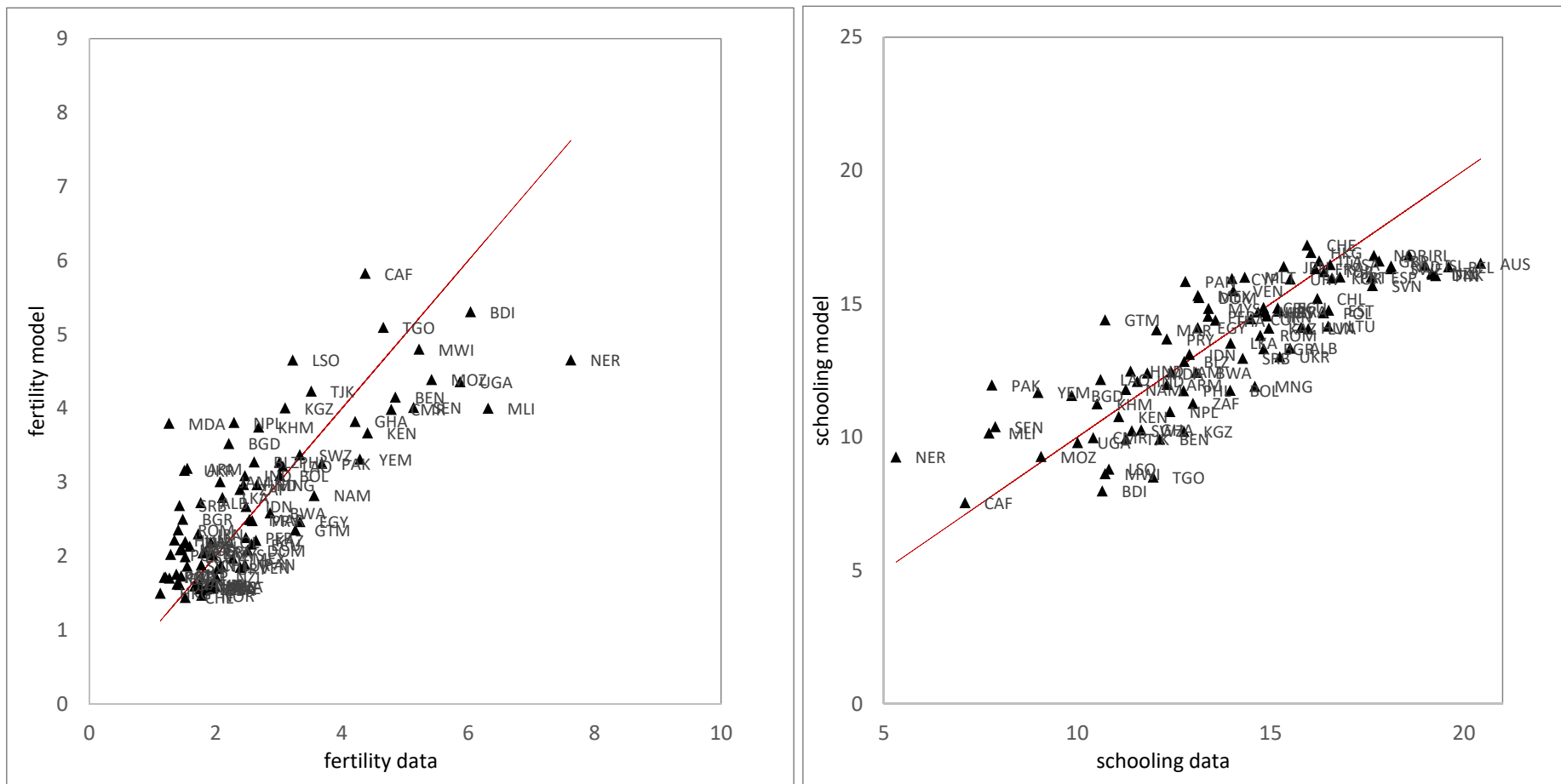


FIGURE 5
Fertility and schooling - Benchmark model versus data

Notes: Fertility 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 benchmark model is calibrated as in Tables 1 and 2, with countries having the same β and χ parameters.

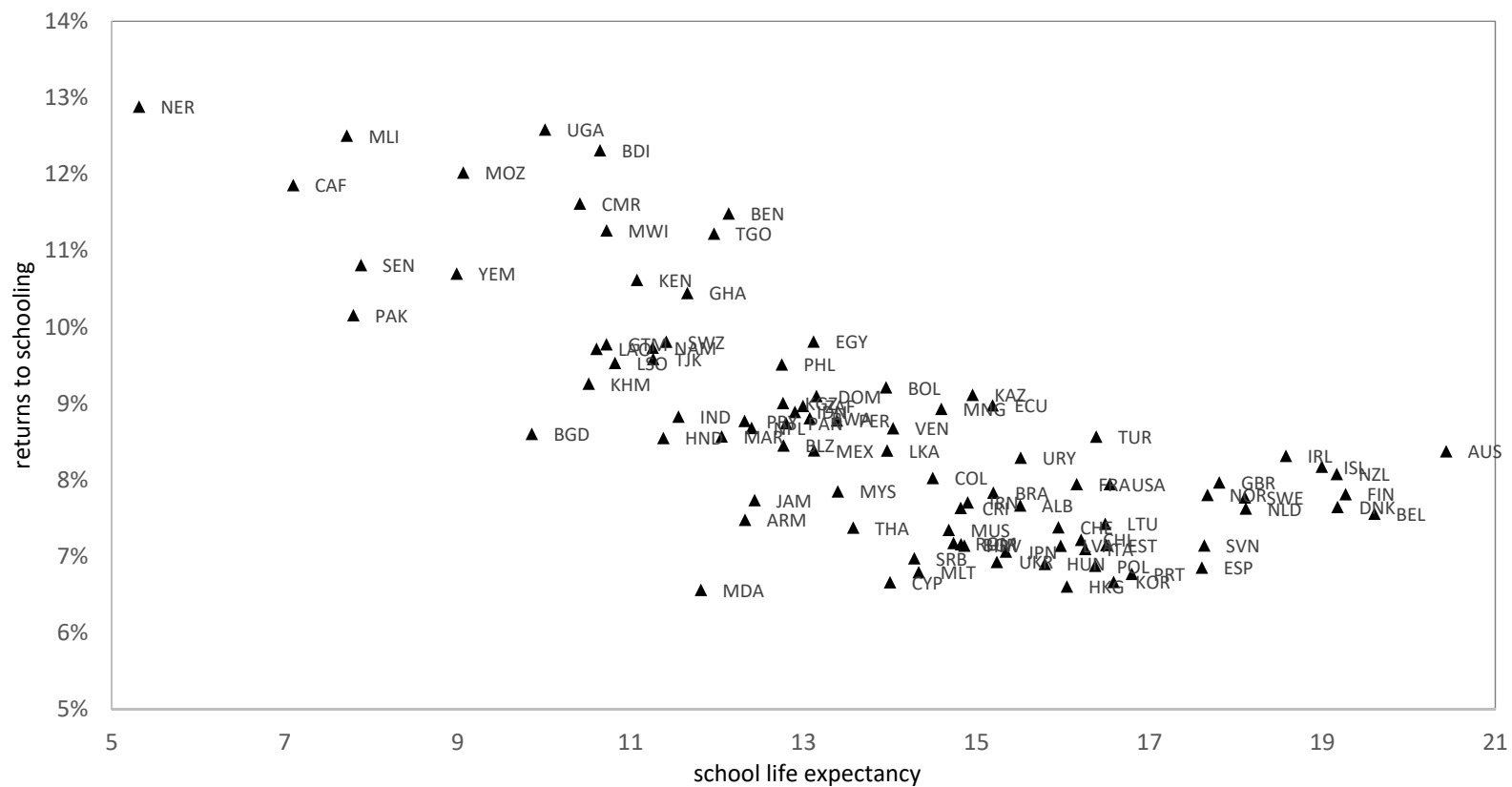
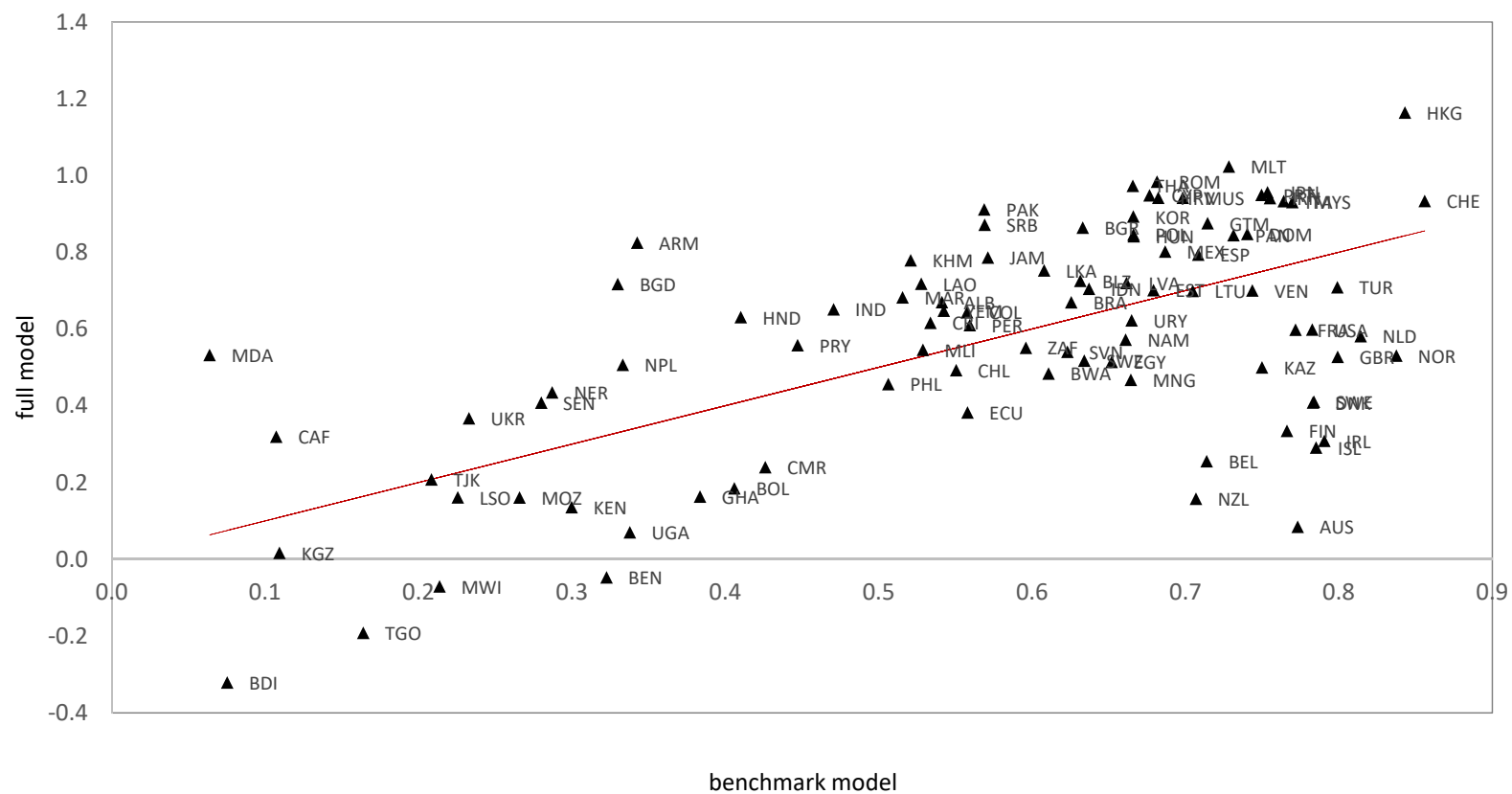


FIGURE 6
Returns to schooling and school life expectancy- Full model

Notes: The full model exactly matches the data by introducing two additional country-specific parameters. Altruistic parameter χ is calibrated to each country so that fertility in the model is the same as in the data. Human capital production function parameter β is calibrated to each country so that schooling in the model is the same as in the data.



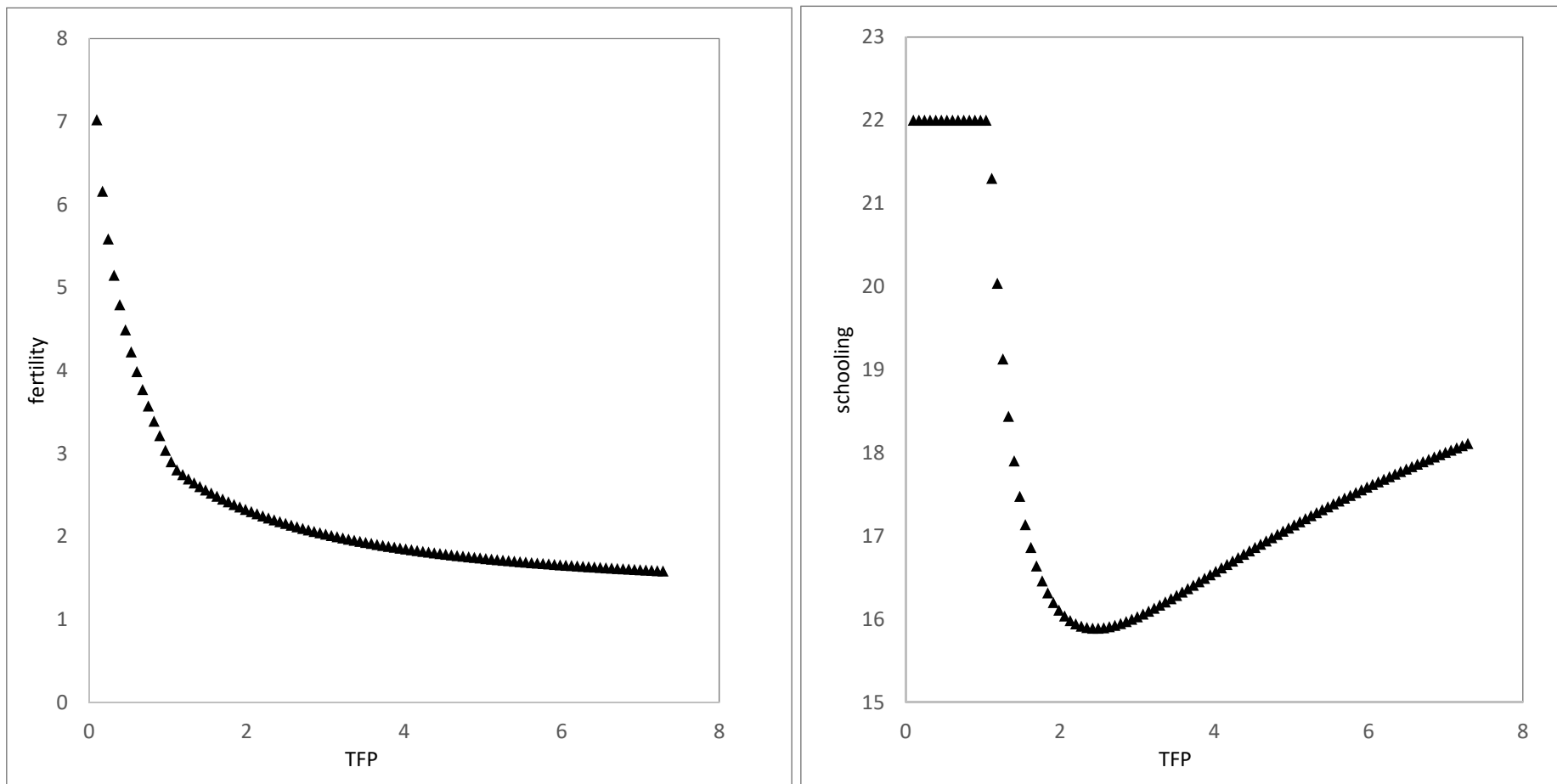
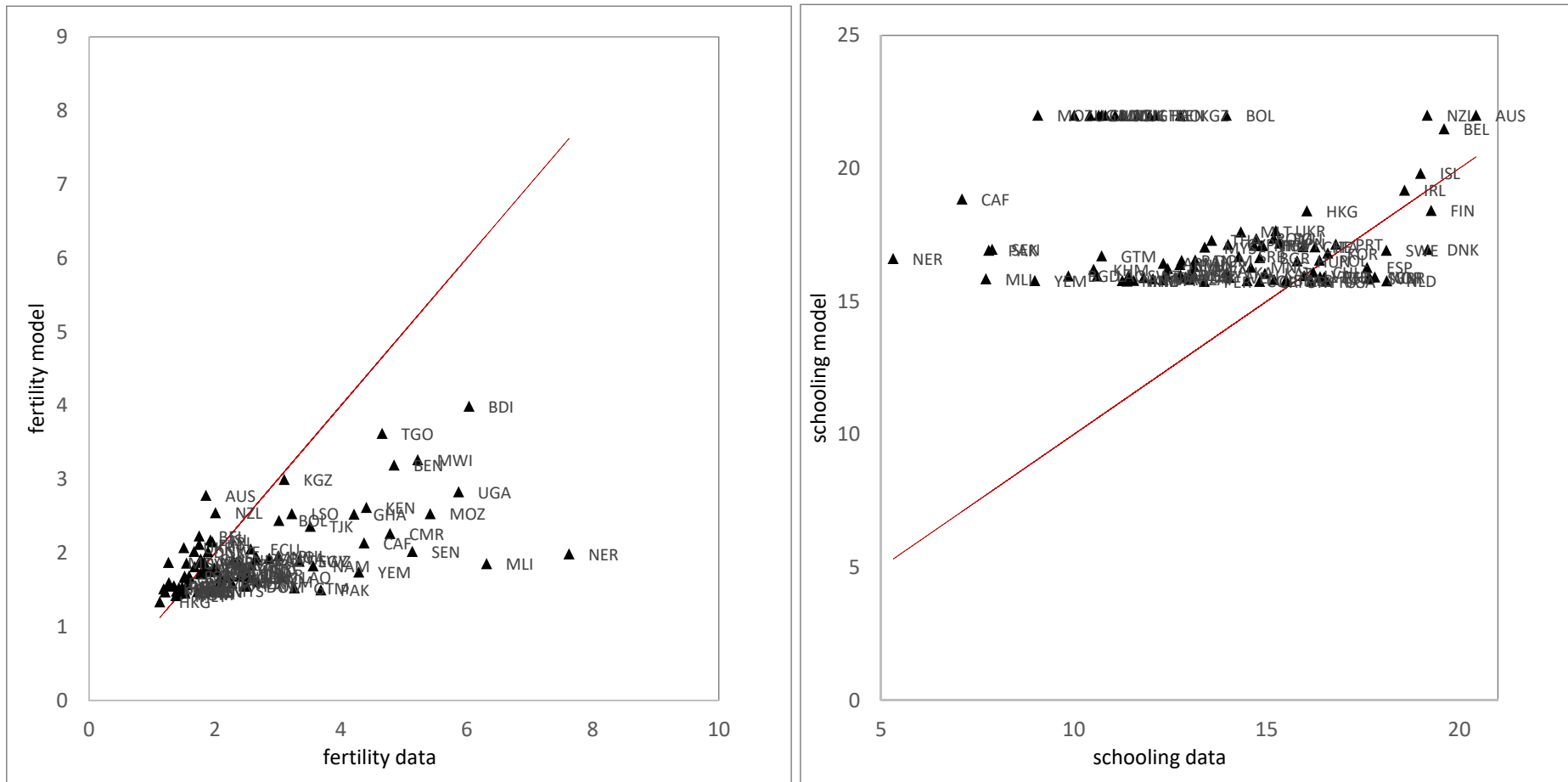


FIGURE 10
Fertility and schooling as a function of TFP - United States

Notes: Fertility and schooling are shown as functions of hypothetical values of TFP. The figure uses the calibrated parameters for the full model in Tables 2 and 3, as well as the country-specific parameters for the United States, including the US-specific β and χ .



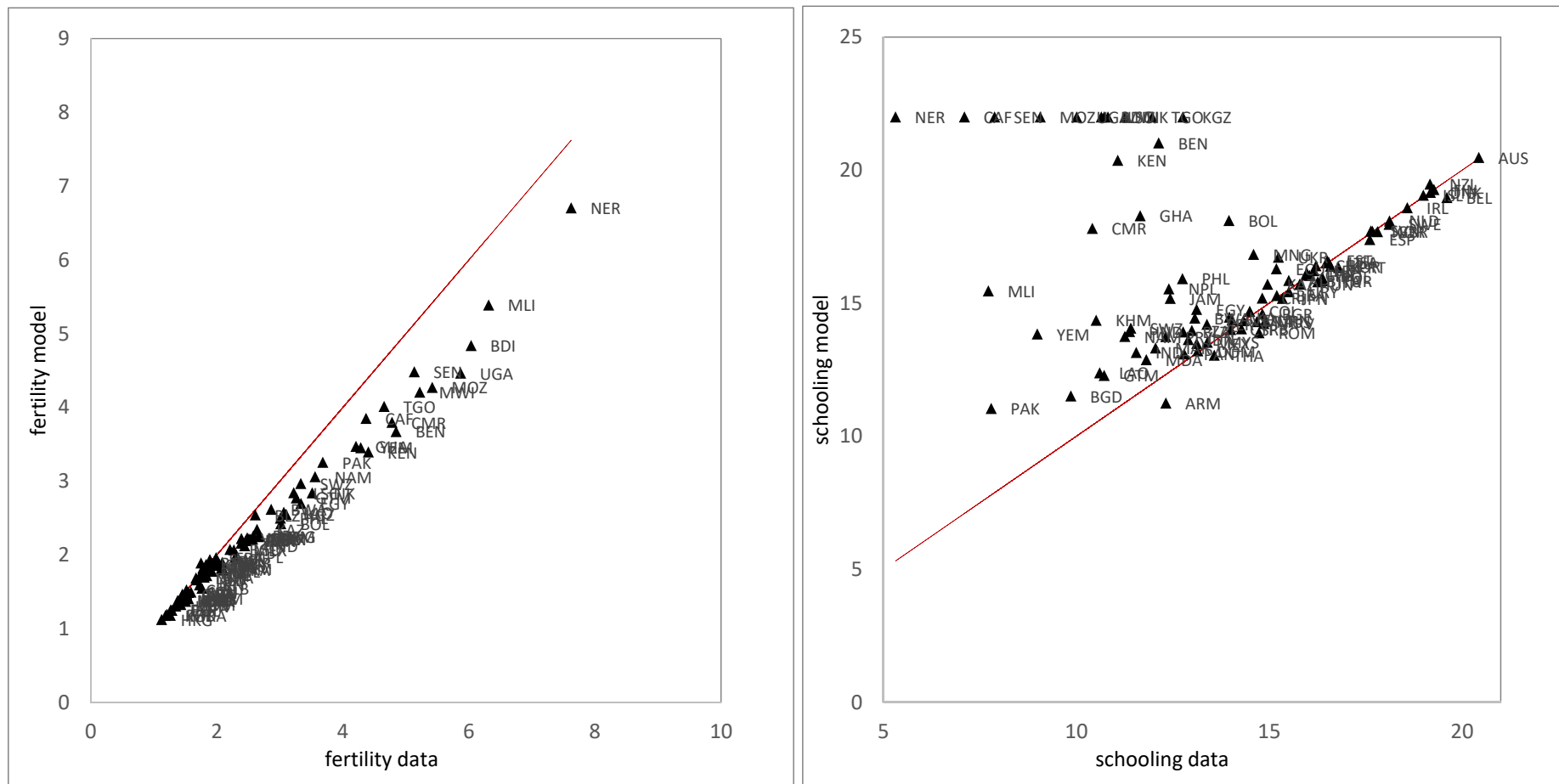


FIGURE 12

Equating amount of public education subsidy per pupil to the 90th percentile value - Fertility and schooling under the counterfactual

Notes: Same as in Figure 9.

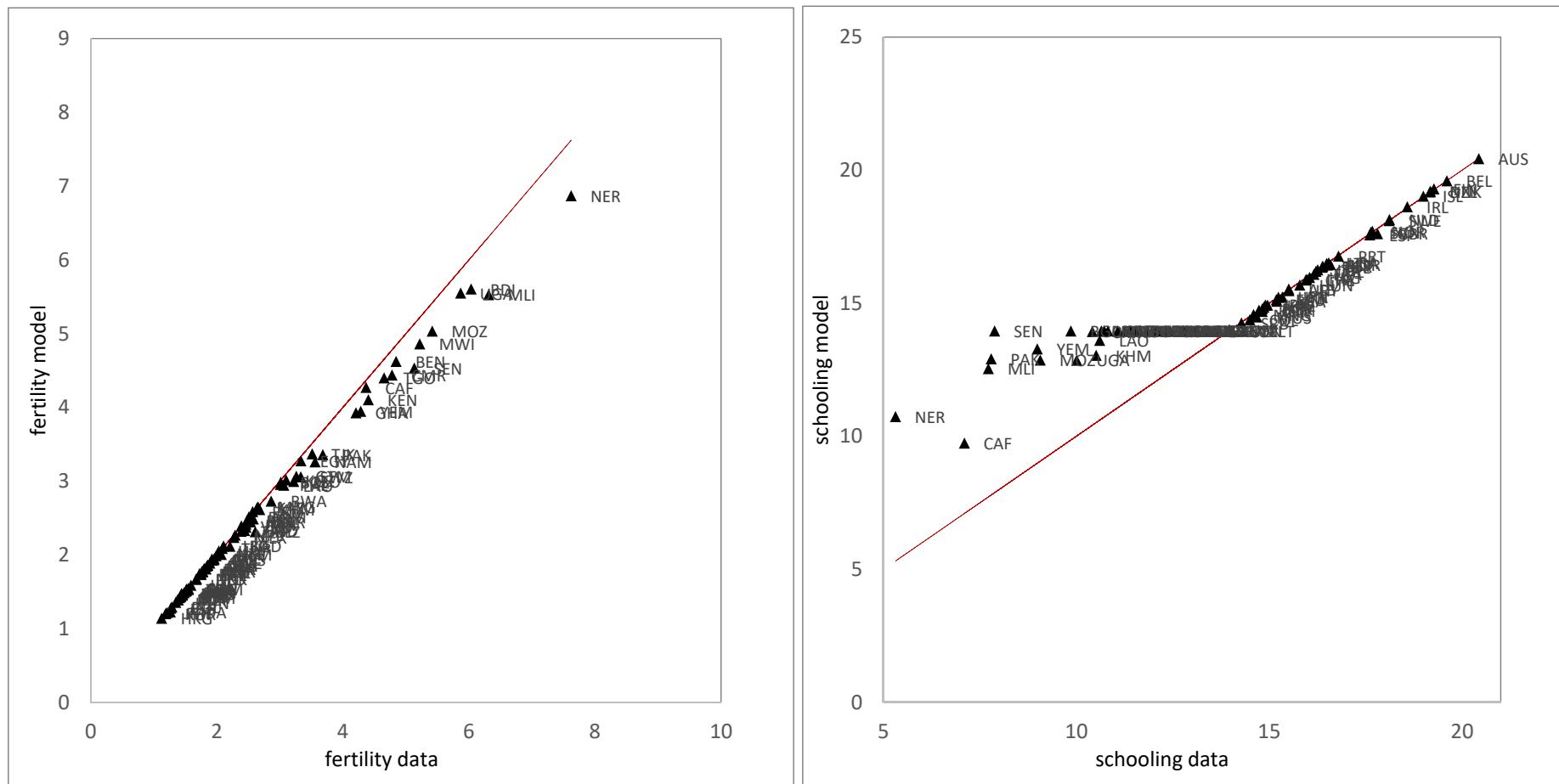


FIGURE 13

Equating duration of public education subsidy to the 90th percentile value - Fertility and schooling under the counterfactual

Notes: Same as in Figure 9.

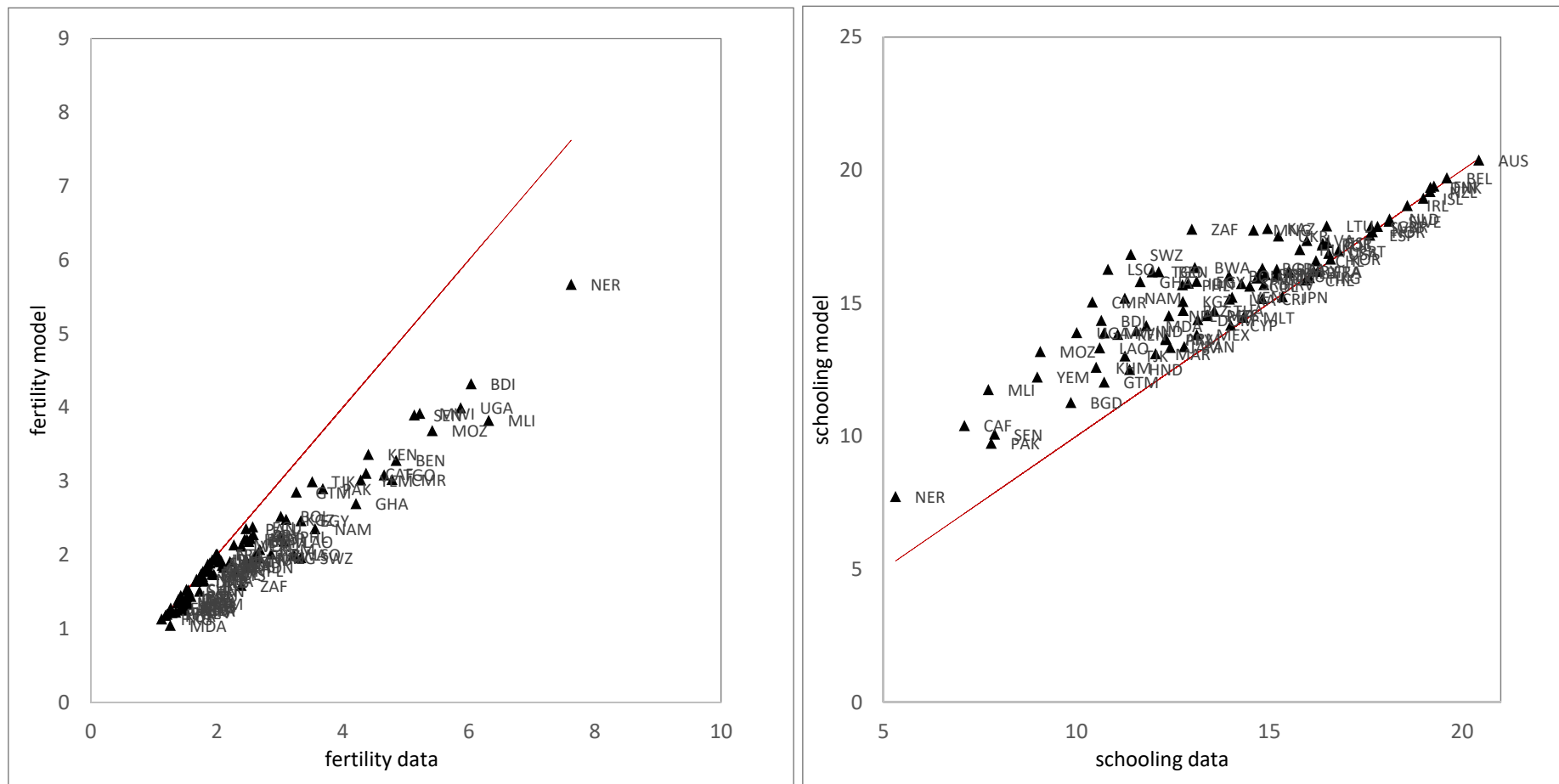


FIGURE 14

Equating all mortality rates to the 10th percentile value - Fertility and schooling under the counterfactual

Notes: Same as in Figure 9.