Longevity, Education, and Income: How Large is the Triangle?*

Hoyt Bleakley†

January 17, 2018

Abstract

While health affects economic development and wellbeing through a variety of pathways, one commonly suggested mechanism is a "horizon" channel in which increased longevity induces additional education. A recent literature devotes much attention to how much education responds to increasing longevity, while this study asks instead what impact this specific channel has on wellbeing (welfare). I note that death is like a tax on human-capital investments, which suggests the use of a standard public-economics tool: triangles. I construct estimates of the triangle gain if education adjusts to lower adult mortality. Even for implausibly large responses of education to survival differences, almost all of today's low-human-development countries, if switched instantaneously to Japan's survival curve, would place a value on this channel of less than 15% of income. Calibrating the model with well-identified micro- and cohort-level studies, I find that the horizon triangle for the typical low-income country is instead less than a percent of lifetime income. Gains from increased survival in the 20th-century are similarly sized.

Keywords: life expectancy, horizon, health, efficiency loss, Harberger triangles

*I thank Lucy Chen, Obeid Rehman, Patricia Padilla, and Robert Venyige for excellent research assistance. I also thank Barbara Anderson, Victoria Baranov, David Canning, Deon Filmer, Andrew Foster, Casper Worm Hansen, Adriana Lleras-Muney, Lars Lønstrup, Emily Oster, Elyce Rotella, Paul Rhode and seminar participants at Princeton (Center for Health and Wellbeing), University of Michigan (Robert Wood Johnson Workshop and Population Studies Center Brownbag), Michigan State University, the University of Connecticut, the PopPov2016 conference, Carnegie Mellon University, University of Pittsburgh, Northwestern University, the University of Buffalo, and the University of Copenhagen for helpful comments.

†Department of Economics and Population Studies Center, University of Michigan. Postal address: 611 Tappan Ave, Ann Arbor, MI 48109-1220. Electronic mail: hoytb at umich dot edu.

1 Introduction

In this world nothing can be said to be certain, except death and taxes.

(Benjamin Franklin, 1789)

In addition to their certainty, death and taxes have something else in common. Death acts like a tax on educational investments. The costs of education—both the direct cost of tuition, fees, et cetera, and the opportunity cost of foregone earnings—are incurred while one is in school. In contrast, the benefits of education accrue in the future through higher earnings after graduating. When deciding whether to continue in school, the student should take the costs for granted, but the expected value of benefits is a function of future survival. This argues for a "horizon mechanism" by which lower adult mortality can lead to more education, which in turn can lead to higher income (Preston, 1980, Schultz, 1981, *inter alia*).

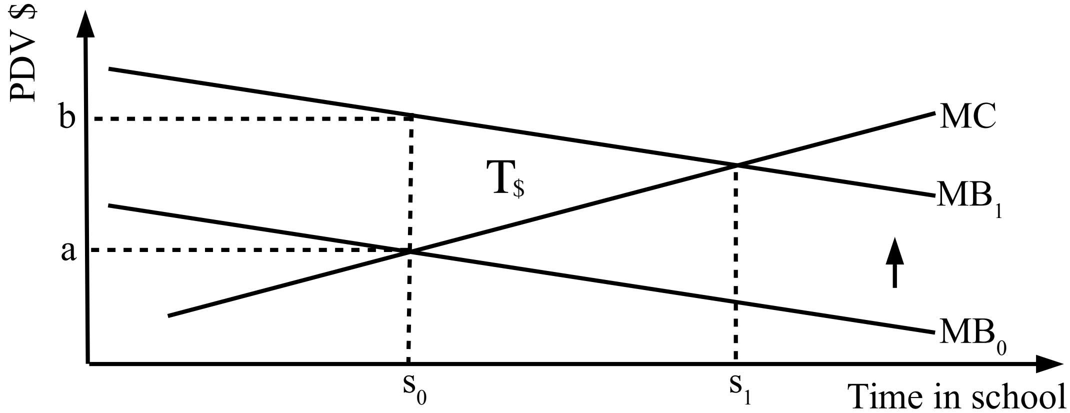
Indeed, life expectancy, education, and income per capita are strongly correlated both across countries and within countries along their paths of development. A question arises as to what are the various causal channels that give rise to these correlations, and indeed whether health improvements might represent high-return investments. (See Strauss and Thomas, 1998, Ashraf, Lester, and Weil, 2008, Behrman, 2009, and Bleakley, 2010a, for reviews.) The horizon mechanism is potentially one such channel, but how large is its contribution to welfare differences between rich and poor countries (or between unhealthy and healthy time periods)?

Much of the existing work on the horizon mechanism has focused on how much (or indeed if) longevity has affected education. One strand of this literature, summarized broadly, uses cross-sectional and/or time-series evidence combined with models using specific production technologies and preferences to estimate this magnitude; see Meltzer (1992), de la Croix and Licandro (1999), Kalemli-Ozcan, Ryder, and Weil (2000), Chakraborty (2004), Cervellati and Sunde (2005), Soares (2005), Hazan and Zoabi (2006), Lorensten, McMillan, and Wacziarg (2008), Hazan (2009, 2012), and Hansen and Lønstrup (2012). Another strand uses natural experiments—plausibly exogenous changes in mortality in this case—to estimate the response of education. Studies by Jayachandran and Lleras-Muney (2009), Fortson (2011), Dorsey, Oster, and Shoulson (2013), Hansen (2013), Stoler and Meltzer (2013), and Baranov and Kohler (2017) are noteworthy. To the extent that income gains are computed in these studies, however, they are specific to the model or based on *ad hoc* extrapolations using Mincer returns to schooling. Almost 20 studies are cited in this paragraph, with most of them published in top general-interest and field journals. With such manifest interest from the profession, one might surmise that the horizon mechanism, if true, would be associated with significant gains in wellbeing. However, I use a standard and transparent economic tool, combined with demographic and education data, to show that the horizon mechanism plays at best a small role in differences in economic welfare.

The present study adds to this literature by first noting the similarity between death and taxes. This suggests the use of a time-honored and transparent tool from public economics: triangles. Such triangles, which are familiar concepts in introductory microeconomics, are a second-order approximation to the efficiency loss caused by a tax that distorts behavior. An advantage of the triangle-based method is that it assumes very little about technology and preferences. Indeed, it only requires is that we evaluate the gains local to an interior solution for the optimal choices (and even this assumption can be relaxed in this application).

We can construct triangles for the horizon mechanism, if we think of the Grim Reaper as imposing a human-capital tax in the form of mortality. Consider a standard model of schooling; the optimum is determined by the intersection of marginal benefits (MB) and marginal costs (MC). Figure 1 illustrates such a model as well as the horizon mechanism. If adult mortality is lower, the marginal benefits of school, as perceived from youth, are higher, thus generating an upward shift of marginal benefits from MB_0 to MB_1, for example. This moves along the MC curve and optimal schooling goes from s_0 to s_1. Notice that expected lifetime income would rise even if we forced this person to choose s_0. But, because the MB at s_0 rises from a to b, there is an additional gain from increasing schooling. Yet diminishing marginal returns pinch off the gap between MB_1 and MC, which intersect at s_1. The gain from this reoptimization is the triangle (labeled $T_\$$). The remainder of this study is an attempt to gauge its size. (See Section 3 for description of the model and formal derivation

Figure 1: Graphical model of schooling and the triangle



Notes: this figure presents a simple graphical model of the schooling decision under high- and low-working-age mortality. The y axis measures present discounted value (PDV) and the x axis measures total time spent in (i.e., years of) school. MC denotes marginal costs of education. MB_0 measures the present value of marginal benefits from obtaining more schooling in the high-mortality scenario, while MB_1 does so in the low-mortality scenario. As shown in the text, a decrease in working-age mortality causes a shift up in the marginal-benefits curve; *e.g.*, from a to b at s_0. Thus MB_1 intersects with MC at a higher time in school than did MB_0. The area between MB_1 and MC between s_0 and s_1 is the triangle ($T_{\$}$) of gains from reallocating from s_0 to s_1. While the curves shown are linear, the results in the study do not require this, although they do require that the response of school to MB have a finite elasticity.

of the triangle.)

A thought experiment illustrates the methodology and shows that the gains from this horizon mechanism are likely small. Imagine taking a particularly unhealthy country to the global frontier in life expectancy. Suppose further that, *by the horizon mechanism alone*, this change also causes the country to reach the global frontier in education. This rather implausible scenario would result in an approximately 40% increase in life expectancy at age 15, and perhaps 10 additional years of schooling. (The data are presented in Section 2.) The triangle would be the product of these two numbers, divided by two because it is a triangle, and multiplied by the Mincer return to schooling (≈ 0.1) so as to be in units of lifetime income (see Section 3). These numbers multiply out to a gain of $\frac{1}{2} \times 0.4 \times 0.1 \times 10$, or 20% of lifetime income. This triangle is a slack upper bound, for two reasons. First, the change in marginal benefits is overstated because life expectancy omits interest-rate discounting and includes mortality reductions well past working ages. Second, it is likely that many factors impede education in developing economies besides just this horizon mechanism, and therefore the actual change in schooling would be less. Put another way, the elasticity of school to longevity implied by this calculation would be quite high ($\approx$ 5). In contrast, the cross-country relationship implies an elasticity of around 2.25, which itself is likely an overestimate. Indeed, recent, well-designed studies put this elasticity below one. The smaller response of education to longevity coupled with discounting brings the triangle to below 2%. A two-percent gain is, of course, better than nothing, but is dwarfed by the gap in income per capita across countries or, for example, the increase in wages possible for many after a two-hour flight from Port-au-Prince to Miami. (Upper bounds on the triangle are presented in Section 5, and the triangles computed using calibrated elasticities are seen in Section 6.)

The essential small-ness of this triangle is not sensitive to various assumptions that went into this calculation, as is shown in Section 7. There I present results for a variety of interest rates, truncation horizons (e.g., retirement), reference cohorts, and ages at which to anchor the survival curve. Note further that the triangle calculation is quite transparent. The Mincer coefficient and schooling elasticity both enter linearly and the proportional change in survival/discounting enters as a square. (The latter two features should be familiar; Harberger triangles are linear in elasticity and squared in the tax rate.) I also show, in Sections 7.3 and 7.4, that this triangle is small even allowing for misallocation, be it from rationing (*i.e.*, constraints), or from subsidies and externalities.

In Section 8, I also compute historical horizon triangles in a subset of countries with early-20th-century life tables. The first example is the US, for which the upper bound gain from the horizon mechanism is 1–3% of lifetime income (depending on discounting) and a unit-elastic schooling response would have generated a triangle of less than 1%. In contrast, the improved survival curve in India from 1933 to 2010 is associated with a triangle of no more than 21%, and more likely below 3%. (I obtain similar results for Egypt, Japan, and China.)

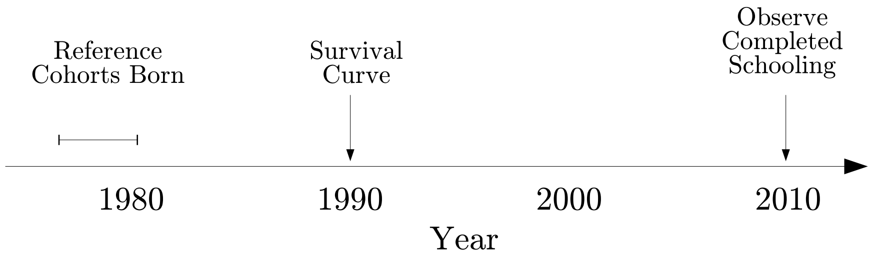
The triangles approximated in this study are dwarfed by two types of rectangles, as shown in Section 9. The first is the proportional increase in discounted life expectancy. This measures the potential increase in lifetime labor input, holding schooling fixed. (There may nevertheless be other factors that cause labor input to decline along the path of development; see Hazan, 2009, and Bloom, Canning, and Moore, 2014, for some historical evidence on this point.) Secondly, the triangles are a factor of 5 to 20 times smaller than a naive calculation that simply multiplies the change in schooling times the Mincer return. This arises because the Mincer return is not an internal rate of return on the investment but rather a later-life footprint of an earlier choice of time in school. At the point of deciding whether to continue in school, later-life benefits should be compared with current costs. This is what makes the triangle a more appropriate object than the 'rectangular' number of extra schooling times the Mincer return.

I then offer conclusions in Section 10. This study is relevant on several dimensions. First, this analysis adds to our understanding of how improvements in health can contribute to economic wellbeing, with a focus on a channel that has received significant attention on the quantity side, but little on the welfare dimension. Second, because this study demonstrates how triangles (and rectangles) can measure the benefits of changes that work through education, it can serve as a model for further applied welfare analysis of educational interventions, to complement existing *ad hoc* approaches.

2 Data

The study combines data on educational attainment and mortality across countries. While there is also heterogeneity in longevity and education within countries, this study focuses on the differences across countries. Information on educational attainment by country and five-year cohort is available from Barro and Lee (2013). These data are built from a combination of survey microdata and published flows of school enrollment and/or attendance over time. For a broad set of countries, cohorts born as early as the 1930s are available, but I focus on those born somewhat later for reasons explained in a moment. Data on life expectancy as well as full life tables are available for recent decades from the World Health Organization's (WHO) Global Health Observatory (2016). The main analysis is for males, but triangles are also small for females (see Section 7, Table 1).

Figure 2: Timeline for Measurement



Notes: this figure describes the timeline for measurement of longevity and education. The reference cohort was born in 1976-80. Health is measured with a period life table computed for circa 1990 (World Health Organization, 2016), and thus the reference cohort would be 10-15 years old at that time. Years of schooling are from the Barro and Lee (2013) database and refer to the year 2010. The reference cohort would be 30-35 years old at that time and thus very likely to have completed schooling.

For the main analysis, I align the education and life-table data based on several considerations. The relevant timeline is seen in Figure 2. The latest wave of education data refers to the year 2010. I focus on the cohort of ages 30-34 at that time to get the most recent cohort that has completed its educational investments. This puts them as being born in the late 1970s. (Below I show these results are not sensitive to using slightly older or slightly younger cohorts.) Someone from a low-income country would be most likely making the school-leaving decision in their teens, and the life table closest in time to such a decision comes from 1990. These data cover a nearly complete set of contemporary countries. (For the analysis over the course of the 20th century in a subset of countries, I present the additional data sources and the analysis in Section 8.)

A few other aspects of the life-table data bear mentioning. First, note that these are period, not cohort, life tables.[1] Life expectancies are generally on the rise, and therefore the period life tables will tend to understate marginal benefits of schooling, especially in unhealthy countries. Second, the demographic variable of interest is the survival curve, typically denoted by demographers as $\ell(t)$, which measures the fraction of the birth cohort alive at age t. Below this is renormalized to a time closer to the school-leaving age (s). Third, the WHO survival curves are specified at one year of age, and then at multiples of

[1] Cohort life tables will not be available for some time, as large fractions of those born in the late 1970s remain alive in all of the studied countries.

five years up to 100. I convert these to an annual frequency with a linear interpolation of the natural logarithm of the survival curve.[2]

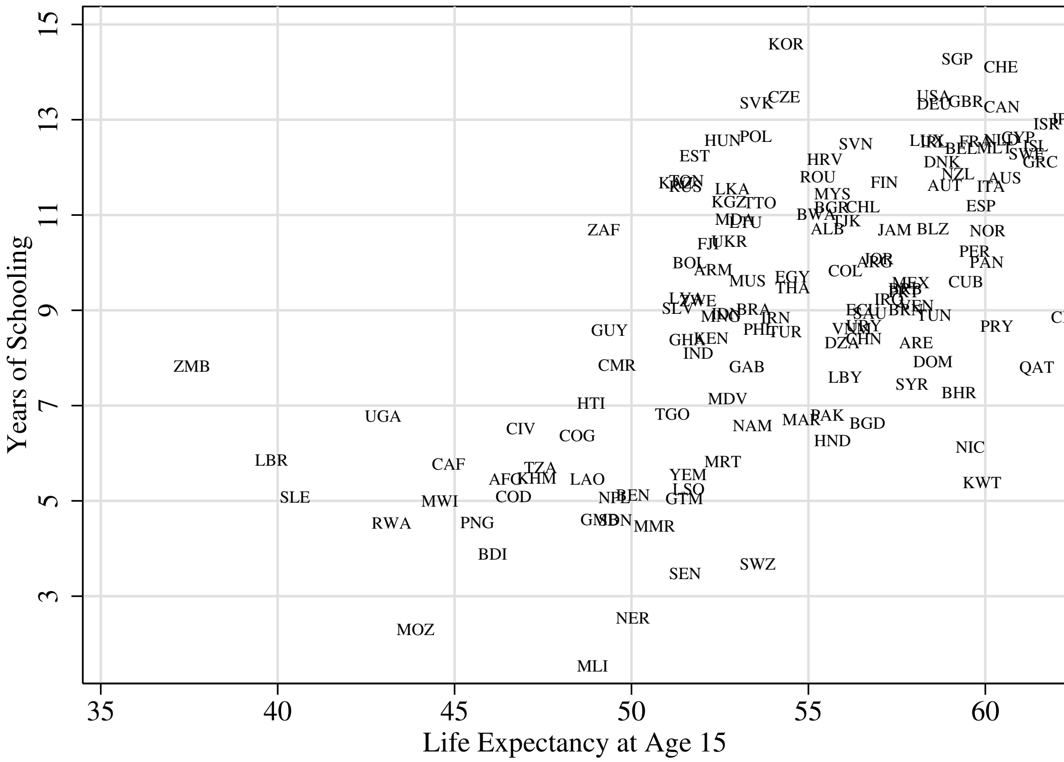
Figure 3 displays a scatterplot of life expectancy at age 15 and educational attainment for the reference cohort. Observations are labeled with World Bank country codes. These variables, denoted below as e_{15} and s, respectively, are strongly correlated. Note also the large range on both axes. There is close to a factor of five difference in s across countries, with some countries below three and others almost at 15. The measure of e_{15} ranges from around 40 to above 60. Taking this number literally, we would expect the average 15-year-old in an unhealthy country to be dead before an age at which a person in a rich country would typically retire. In contrast, with an e_{15} of 60, an average 15-year-old in a healthy country would expect to live well past typical working ages. The country with the highest e_{15} is Japan (JPN), although a number of countries are close on its heels with expectations of an additional 60+ years of life after age 15, even in a few countries with comparatively low education such as Costa Rica and Qatar. The country with the highest level of education in the Barro/Lee data is South Korea (KOR), with close to 15 years of schooling. Again, there are a number of countries not too far behind.

The lowest outcomes for life expectancy merit discussion. Circa 1990, some of these countries were in the grip of what we now know to be transitory shocks. For example, Zambia (ZMB) and Uganda (UGA) were in the worst years of the HIV/AIDS epidemic. Further, Liberia and Sierra Leone (LBR and SLE, respectively) were in the midst of civil wars. The presence of transitory factors in measures of survival will tend to overstate the differences in marginal benefits of schooling across countries insofar as it is the low-human-development countries that are more afflicted by such transitory negative shocks.[3] This will artificially inflate the size of triangles computed below.

[2]The so-interpolated survival curves are plotted for each country in Appendix Figure F. A linear and a bi-cubic interpolation of the log survival curve are highly correlated, with $\rho > 0.998$).

[3]Another point to note is that the 1990 life table actually draws on averages of mortality data from surrounding years, which explains why the civil war in Sierra Leone, which started in approximately 1991, would be influencing the 1990 period life table.

Figure 3: Years of Schooling versus Life Expectancy at Age 15



Notes: this figure contains a scatterplot of years of schooling (s) versus years of life expectancy at age 15 (e_{15}) for males in the sample of countries. Observations are labeled with World Bank country codes. Life expectancy is computed from a period life table measured circa 1990 (World Health Organization, Global Health Observatory, 2016). Years of schooling are from the Barro and Lee (2013) database and refer to the cohort born in 1976-80 and observed in 2010. (See Appendix Figure H for a log-log scatterplot of these data.)

3 Model

3.1 Preliminaries

In this section, I present a stylized model of the education decision and use it to motivate a triangle as an approximate measure of the efficiency gains from the horizon mechanism. The model is a simplified version of Ben-Porath (1967) and Mincer (1958), but augmented to allow for a general survival curve. It imagines an agent born at $t = 0$, in school until period s, and working in the labor market thenceforth. The probability of his survival to period t is $\ell(t)$. By definition, $\ell(0) = 1$, $\ell'(t) \leq 0$, and $\lim_{t\to\infty} \ell(t) = 0$. The interest rate is r and the growth rate of the economy is g.[4] Let the flow of direct costs of the sth year of school be $c(s)$, and let the wage, if he is working, be $f(s)$. For analytical convenience, I use Jacob Mincer's (1958) assumption/result of multiplicative separability of the wage from the discount term. (I abstract here from retirement and from post-school learning, although the latter might be proxied by g. These aspects are addressed in Section 7 below.)

Viewed from birth, the expected lifetime income, net of direct costs, is as follows:

$$Y \equiv \int_s^\infty f(s)e^{-(r-g)t}\ell(t)dt - \int_0^s c(t)e^{-(r-g)t}\ell(t)dt \tag{1}$$

The first term integrates income over the working lifetime and the second term integrates direct costs up to the point of leaving school.[5] The derivative of this object with respect to time in school is

$$-f(s)e^{-(r-g)s}\ell(s) - c(s)e^{-(r-g)s}\ell(s) + \int_s^\infty f'(s)e^{-(r-g)t}\ell(t)dt$$

The first term arises from moving the lower limit of integration. That is, slightly more time in school means slightly less time working. These foregone wages represent the opportunity cost of school. The second term gives us the additional direct costs incurred from spending slightly more time in school. Finally, the third term integrates the future discounted wage increases from that extra bit of schooling. The first two terms are marginal costs, and the second third term is the marginal benefit. If we re-evaluate the discounting (both financial

[4] These are treated as exogenous throughout the analysis. A fixed r can be justified by assuming the the country is a price-taker in international capital markets. But might the interest rate on student loans itself adjust to mortality risk? Theoretically, yes, however two factors mitigate this concern. First, typically student debts have a maturity shorter than the horizon over which survival gains will appear in the analysis below. Second, student loans (of any type) are still quite unusual in low-income countries. Spillovers from human capital, which might influence the growth of the economy on a transition path to higher average education, are discussed in Section 7.4. Finally, note the use of a constant interest rate. If instead the yield curve sloped up ($\dot{r} > 0$), then future benefits of increased survival would be discounted more heavily and both rectangles and triangles would be smaller than suggested below.

[5] While this exercise is cast as maximizing the monetary value (income net of costs), it would be simple to include psychic benefits and/or costs in this framework. I do not require direct observation of $c(s)$, which might well include non-pecuniary costs. Similarly, $f'(s)$ could be construed to include non-pecuniary benefits, although doing so would require increasing the calibrated β.

and demographic) from a viewpoint date of s, marginal costs are $MC = f(s) + c(s)$ and marginal benefits are

$$MB = \int_s^\infty f'(s)e^{-(r-g)(t-s)}\frac{\ell(t)}{\ell(s)}dt$$

Notice that, conditional on survival up to s, the discounting only affects marginal benefits.[6]

Further, for a fixed s, $f'(s)$ is a constant and can thus be factored out of the integral. Thus, MB reduces to $f'(s)\Delta(s)$ in which $\Delta(s) \equiv \int_s^\infty e^{-(r-g)(t-s)}\frac{\ell(t)}{\ell(s)}dt$ is a factor that inflates $f'(s)$ to account for the integration of future discounted flows. For zero mortality, for example, $\Delta(s)$ becomes $1/(r-g)$, the familiar factor that converts a constant flow over an infinite horizon to present discounted value. For non-zero mortality, $\frac{\ell(t)}{\ell(s)} < 1$, which pushes Δ below $1/(r-g)$.

The model lends itself naturally to graphical analysis, such as seen in Figure 1. The y-axis contains the present discounted value of money and the x-axis is time in school, s. The model can be distilled into two curves: the marginal benefit and marginal cost of time in school (MB and MC, respectively). Marginal direct costs typically increase with s, although this is not required. Opportunity costs almost certainly increase with s, for two reasons. First, even holding schooling fixed, children become more productive as they mature. Second, if MB > 0, this means that the opportunity wage is rising with s. Marginal benefits are positive if, as one would hope, time in school makes a student more productive in the labor market. Yet the marginal benefits may decline (at least relative to marginal costs) as students attend more school.[7] At the optimal level of schooling, $MB = MC$, which

[6] The triangles below will be defined in period-s units, but one could also re-value them in period-zero units. Two points are worth mentioning here. One, the triangles will have a lower value viewed from $t = 0$ because of financial discounting and because not all newborns will survive to period s. Two, while the analysis below contemplates decreases in mortality in $t > s$, these are, in the cross-country comparison, typically accompanied by decreases in mortality at younger ages as well. Lower mortality in $t < s$ would mean that a triangle viewed at period s would appear larger when viewed from $t = 0$. The triangle in period-s units need not change because it is conditional on surviving to s. (It would be strange to consider a period-s schooling decision if the person had died before period s.)

[7] Formally, declining marginal benefits are not required to obtain an interior solution. Rather, they must rise less quickly than marginal costs; which is required to satisfy the Second-Order Condition (SOC) for optimization. The graph in Figure 1 displays the case of declining MB, but the results of the study carry through as long as the SOC holds.

is the first-order condition or FOC.[8] The FOC, cast in period-s units, can also be written as

$$f'(s)\Delta(s) = f(s) + c(s). \tag{2}$$

As survival rates increase, future benefits are less heavily discounted. This causes a shift upwards in the marginal benefits curve, say from MB_0 to MB_1. Because the marginal costs are in the present, this change in survival does not affect the MC curve, and so the new optimal schooling moves along MC from s_0 to s_1.

3.2 Derivation of the triangle using the graph

The triangle is seen when comparing the curves and optima for high and low mortality. At the old optimum (s_0), the less-discounted marginal benefits are greater than marginal cost by an amount $(b - a)$ in Figure 1. This gap forms the left edge of the triangle. This gap also is what induces greater investment in education. However, the gap closes as s increases, with the two curves converging again at a new optimum, s_1. The two curves coming together to meet at s_1 form the other two edges of this triangle.

The triangle of gains is thus seen in the graph between s_0 and s_1 on the horizontal axis and a and b on the vertical. Let Δ_0 and Δ_1 be the discount/summation factors that hold at s_0 in the low- and high-survival environments, respectively. This means that, in Figure 1, $a = f'(s_0)\Delta_0$ and $b = f'(s_0)\Delta_1$. The triangle, being $\frac{1}{2}$ × height × base, is therefore

$$T_{\$} = \frac{1}{2}\left(\Delta_1 - \Delta_0\right) f'(s_0)\left(s_1 - s_0\right),$$

in dollars. I find it easier to think in proportions of lifetime income, which is $\Delta_0\, f(s_0)$ in the low-survival (high-mortality) scenario. The so-normalized triangle is

$$T = \frac{1}{2}\left(\frac{\Delta_1 - \Delta_0}{\Delta_0}\right)\frac{f'(s_0)}{f(s_0)}\left(s_1 - s_0\right). \tag{3}$$

[8] This rules out a corner solution at zero. Primary-school attendance rates have grown markedly in low-income countries during the past half century, and the vast majority of school-aged children attend at least some school.

What are the components of this equation? The $\frac{1}{2}$ is there because it is a triangle, not a rectangle. The $\left(\frac{\Delta_1 - \Delta_0}{\Delta_0}\right)$ term measures the proportional change in the discount/summation factor for future income flows. This change is positive if there is lower working-age mortality for the "1" than for the "0" case. In the sections that follow, I develop measures of this using data on survival curves. The $(s_1 - s_0)$ term reflects the increase in schooling *caused* by the longer horizon. Below, I provide use bounds and estimates of this response (in Sections 5 and 6, respectively). Finally the, (f'/f) term measures the proportional increase in wages from a small change in schooling. Another way to write this would be the $\frac{d \ln f(s)}{ds}$, which is what Mincer's β is meant to measure in the oft-seen regression of log wages on years of schooling. There is no shortage of cross-sectional estimates for this parameter, both in developed and developing countries. Psacharopolous and Patrinos (2002) and Montenegro and Patrinos (2014) provide accessible and thorough surveys of estimated Mincer βs for numerous countries and time periods. Commonly, recent estimates for Mincer's β are close to 10%. Although higher and lower estimates are found in the literature, they are generally not much higher than 15% and not much lower than 5%. One challenge is that the derivative represents a causal statement, while the typical coefficient comes from a conditional correlation. I will use the 10% number as a default, and the sensitivity of the triangle to this choice is discussed in Section 7 below.

The triangle measures the gains from the reallocation (from s_0 to s_1) in the lower-mortality environment. Intuitively, the triangle is collecting all of the gaps between marginal benefit and cost that accrue along the way to adjusting to the new s. Thus it provides a measure of the gain in productive efficiency if reallocating from s_0 to s_1, when exposed to $\Delta_1(s)$ instead of $\Delta_0(s)$.

Triangles are a time-honored technique, promoted extensively by Arnold Harberger, for computing the change in surplus associated with a distortion. (See Harberger, 1971, and Hines, 1999, for detailed reviews.) The triangle computation provides a quick and transparent method for approximating losses in surplus, yet its use has been displaced somewhat in recent decades by methods for computing so-called exact surplus (such as work building on Hausman, 1981). Such methods are exact if the model itself is exactly right, otherwise they should be viewed as approximations. An advantage of the triangles approach is that it does not require specification of the global properties of production functions (in this case the education production function, in addition to whatever technology firms use when they employ workers of various skill levels), but rather relies on generic properties of the decision problem at or near the optimum.[9]

3.3 The triangle as a second-order Taylor approximation

The triangle is derived formally as a second-order Taylor approximation to the objective function. Let lifetime income, net of costs, be $Y(s(\Delta), \Delta(s))$, as in equation 1. I consider its properties locally around s_0 and $\Delta_0(s_0)$. Subscripts denote derivatives in this expression of the second-order Taylor approximation for Y:

$$Y_s ds + Y_\Delta d\Delta + \frac{1}{2}\left[Y_{ss}(ds)^2 + 2Y_{s\Delta}(ds)(d\Delta) + Y_{\Delta\Delta}(d\Delta)^2\right] \tag{4}$$

The first term is zero, by the FOC. The second term is the rectangle of gains, call it $R_\$$, from increasing survival but holding s fixed. The remaining terms comprise the triangle, as measured in dollars. This can also be expressed as

$$T_\$ = \frac{1}{2}\left[Y_{ss}\left(\frac{ds}{d\Delta}\right)^2 + 2Y_{s\Delta}\left(\frac{ds}{d\Delta}\right) + Y_{\Delta\Delta}\right](d\Delta)^2$$

The components of $T_\$$ are simple to derive. For the remainder of the problem, I take the perspective of someone alive and in school at s_0 (rather than at 0). I therefore recast $\Delta(s)$

[9] A distinct issue in using triangles is how to deal with income effects on the consumer's demand curve (Hines, 1999). This issue is not relevant here because the choice of time in school is about productive efficiency (maximizing lifetime income, net of costs, w.r.t. s). At the FOC for s, the consumer's consumption decision can be separated from the education decision. How? Consider a consumption problem with a constrained decision for education. A shadow price for the education decision would appear in the FOC's for the consumption decisions. As the constrained value for years of education approaches the unconstrained optimum, those shadow prices go to zero, and thus disappear from the FOC. (See Section 7.3 for discussion of the problem if the FOC does not hold.)

as

$$\Delta(s) = \int_s^\infty e^{-(r-g)(t-s_0)} \frac{\ell(t)}{\ell(s_0)} dt$$

When taking a derivative with respect to s, the lower limit of the integral moves, but not the reference date. As such,

$$\frac{d\Delta}{ds} = -e^{-(r-g)(s_0-s_0)} \frac{\ell(s_0)}{\ell(s_0)} \ + \ 0$$

which equals -1. This reflects the assumption that the only mechanism by which time in school affects time worked is through the opportunity cost of time while a student. (See Appendix A for the derivation with $\frac{d\Delta}{ds} \neq -1$ and Section 7.2 for a quantitative evaluation.)

Lifetime income, net of costs, can be re-written as

$$Y(s, \Delta) = f(s)\Delta(s) + \int_0^s c(t) e^{-(r-g)(t-s_0)} dt.$$

The derivative of Y w.r.t. s is $Y_s = f'\Delta - (f + c)$. The remaining derivatives are as follows: $Y_{ss} = f''\Delta - (2f' + c')$; $Y_{s\Delta} = f'$; $Y_{\Delta\Delta} = 0$. Finally, full differentiation of the FOC ($Y_s = 0$) yields

$$\frac{ds}{d\Delta} = f'/(2f' + c' - f''\Delta) = -f'/Y_{ss}. \tag{5}$$

Substituting these expressions into $T_\$$ yields the following:

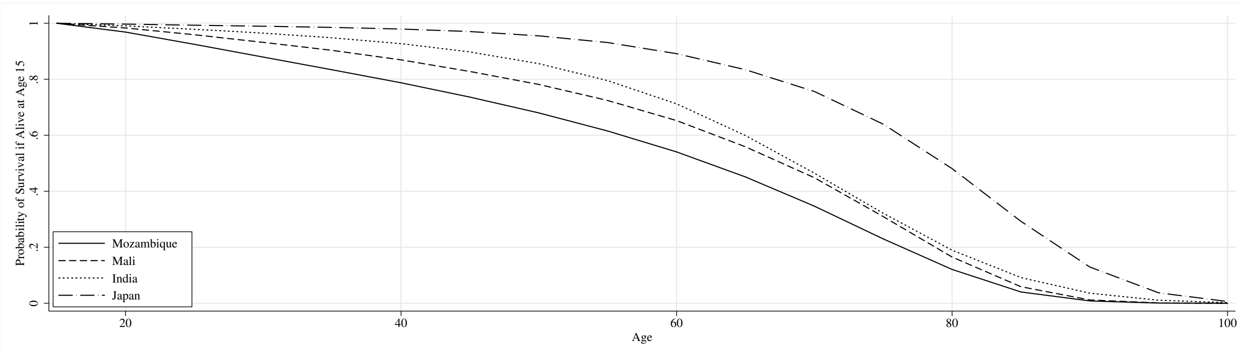
$$T_\$ = \frac{1}{2}\left[(Y_{ss})\left(\frac{f'}{-Y_{ss}}\right)^2 + 2(f')\left(\frac{f'}{-Y_{ss}}\right) + 0\right](d\Delta)^2 = \frac{1}{2} f' \frac{ds}{d\Delta}(d\Delta)^2. \tag{6}$$

I then normalize by lifetime income, $f(s_0)\Delta_0(s_0)$, to obtain T, which can be written as

$$T = \frac{1}{2}\left(\frac{d\Delta}{\Delta}\right)\beta ds, \tag{7}$$

which is the infinitesimal-change version of equation 3 above.

Figure 4: Survival Curve (Probability), Conditional on Surviving to Age 15 $\left(\frac{\ell(t)}{\ell(15)}\right)$



Notes: this figure displays survival curves for select countries. Data and computation are re-normalized to be conditional on surviving to age 15. The survival curves are drawn from the World Health Organization's Global Health Observatory (2016) and use the *circa*-1990 life table.

4 Constructing Δ

$\Delta(s)$ accumulates the flows of future discounted income, for each possible s. It requires two additional inputs: an interest rate (net of wage growth) for financial discounting and a life table for computing survival probabilities. I use a net interest rate, $r - g$, of 2% as a default, although I explore the sensitivity of the calculations to different values in Section 7. The survival curve is normalized to a base age, usually 15 years. Examples from four countries—Mozambique, Mali, India, and Japan—are seen in Figure 4. The highest curve, both in the graph and in the world, belongs to Japan, where a 15-year-old has an almost 50% probability of reaching age 80. The middle curves belong to Mali and India, for whom survival to age 80 is below 20%, conditional on being alive at age 15. The lowest curve, the solid line, belongs to Mozambique. According to this curve, a bare majority of 15-year-olds reach age 60 and less than 15% reach 80 years. The largest gaps between the Japanese and the other curves tend to occur at relatively older ages, in one's 70s, for example. There are, nevertheless, significant gaps that emerge at working ages.

It is instructive to compare the life expectancies and Δ's. I use the extreme examples from above: Mozambique versus Japan. According to the circa-1990 life table, Mozambique has an e_{15} of a bit more than 43 years, while Japan has a number closer to 62. (Life expectancies are defined from the perspective of the starting age, which means the expected age of death is $15+e_{15}$.) For an $r-g=0$, Δ is the same as e_{15}. However, even for a relatively small interest-rate like 1%, the Δ's drop by approximately one quarter. The discounted sum of life years would be closer to 34 for Mozambique and 46 for Japan. For a 2% interest rate, the numbers drop further to approximately 27 and 35 for Mozambique and Japan, respectively, and then to 19 and 22 if the interest-rate is as high as 4%. That being said, the roughly similar shapes of the survival curves mean that discounting principally compresses the differences between countries rather than causing great reshuffling amongst countries. Consider, for example, the regression of a Δ using $(r-g)=2\%$ on e_{15}, for $N=143$ countries.

$$\begin{array}{ccccccc} \Delta_i & = & .408 & e_{15,i} & + & 9.84 & + \ \epsilon_i \\ & & (.003) & & + & (0.19) & \end{array}$$

Yet the R^2 for this regression is 0.990. Even for a net interest rate of 20%, the R^2 remains higher than 75%, while the regression coefficient drops to less than .01.

5 Upper bound on the triangle

This section presents what is most likely an upper bound on the triangle associated with the horizon mechanism. I construct this bound (call it $\bar{T}$) using easily available numbers and a bare minimum of computation.

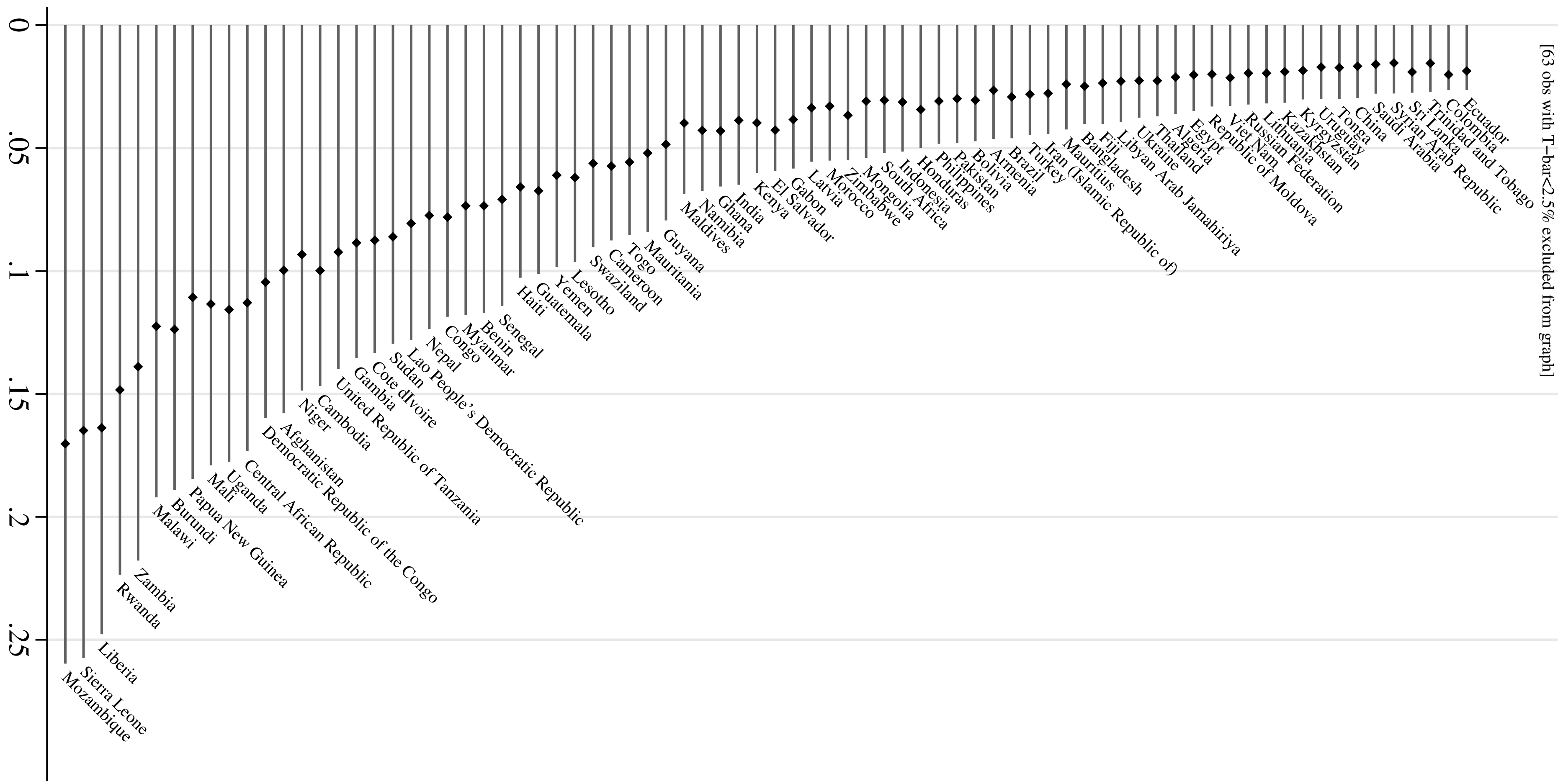
I start with a simple thought experiment: imagine taking unhealthy countries to the global frontier of health by giving them Japan's survival curve. This is, of course, an extreme exercise. (Consider the vast academic literature that puzzles over why a rich country like the US does not have a more Japan-like life table.) Nevertheless, this assumption gives us the height of a triangle. For starters, I use life expectancy at age 15 as a proxy. This generates an overstatement of the true gain because it ignores the interest-rate discounting, which would down-weight gains in later life. Alternatively, I use my computed Δ, evaluated at age 15 for a net interest rate of 2%.

Equation 3 above also uses the base of the triangle: $(s_1 - s_0)$, which measures the causal response of schooling to increased Δ. Were it not for the causal nature of this term, it would be easy to estimate with the observed differences in schooling over time and/or across space. But such differences almost certainly arise because of additional factors. That is, differences in the labor market, the school system, readiness to learn of the children, etc., would all contribute to higher school attendance even in the absence of a better survival curve. While this complicates obtaining a precise estimate of the model's $(s_1 - s_0)$, it does mean that the observed differences are likely an *upper bound* for the causal contribution of this horizon mechanism to the increase in schooling. For the reference cohort, in the Barro-Lee data, the country with the highest schooling is South Korea, at 14.6 years. I assume that $s_{Korea} \geq s_1$, which implies that $\bar{T} \geq T$.

Values of $\bar{T}$ are shown in Figure 5. The length of each country's vertical line denotes the size of the triangle computed with e_{15}. Because this exercise takes all countries to the global frontier in s and Δ, the values of $\bar{T}$ are strictly positive, with the exception of Japan and Korea, who are already at one of the frontiers. Black diamonds indicate $\bar{T}$ using Δ for $r - g = 2\%$. (The two measures of $\bar{T}$ have a correlation coefficient in excess of 0.99.) Country names appear at the end of each line.

The vast majority of $\bar{T}$ are quite small. The median upper bound using e_{15} is slightly above three percent. (Countries with $\bar{T}_{e_{15}} < 2.5\%$ were excluded to make the figure legible.) Over 90% of countries have a $\bar{T}_{e_{15}}$ below 15%. The largest triangle is 0.26, that of Mozambique. When using a modest, 2% discount rate, the upper bound on the triangle drops, typically by 40% of the value computed with e_{15}. Indeed, more than 9 out of 10 of countries have $\bar{T}_{\Delta, r-g=.02}$ below 10%. Further, only three countries have values above 15%: Liberia, Sierra Leone, and Mozambique, all countries for which Δ was temporarily low circa 1990.

Figure 5: Upper bound on the Triangle



Notes: this table presents estimates of $\bar{T}$ using equation 3. This calculation is based on taking all countries to the global frontier in both Δ and s. The line denotes the size of $\bar{T}$ computed with e_{15} (a Δ with zero discount rate) and the black diamonds denote the size of $\bar{T}$ for a discount rate of 2%. Country names are reported next to each line. This figure excludes the 63 observations that have a $\bar{T}$ less than 2.5%. The underlying data sources and variable definitions are described in the text.

6 Calibrate schooling response with elasticities

In this section, I use estimates of the response of schooling to Δ and find much smaller estimates of the triangle. The upper-bound exercise in the previous section, which imagined taking countries to the global frontier in both health and education, implicitly used very high elasticities for some countries, especially the poorest ones. (Appendix Figure G shows a scatterplot of this implied elasticity versus years of schooling, with countries labeled.) Most of the countries with less than eight years of schooling in the reference cohort have implied elasticities above 5. Mozambique, the extreme example from above, had an implied elasticity of more than 12. A comparable estimate from the cross-country data (from a log/log regression of s on e_{15}) is approximately 2.2, or 2.5 in the subsample below the median e_{15}, and even this estimated elasticity is presumably biased upwards.

The changes in schooling consistent with such elasticities seem implausible, but how much lower should the elasticity be? To answer this question, I turn to several well-identified (and well-published) studies of precisely this question at the micro and cohort levels. These studies analyze large changes in mortality and are careful in constructing a comparison/control group.

The first study is by Jayachandran and Lleras-Muney (2009), who examine the swift decline of maternal mortality in Sri Lanka circa 1950. The authors argue that the shock was induced by specific policy interventions combined with the rapid diffusion of new medication, such as antibiotics. They use the different exposures across cohorts and regions within the country to construct a difference in difference for women. They also use men as a further comparison group, as women were more directly affected by the drop in maternal mortality. They report an elasticity of 1.0, a far cry from the implied number for Mozambique in the upper-bound exercise. Their calculation uses un-discounted years of life expectancy between 15 and 65, but there is enough information to compute the proportional change in Δ for other net interest rates.[10] For $r - g = 2\%$, I compute an elasticity of 0.7 in that episode.

[10] In Jayachandran and Lleras-Muney (2009), Table II reports mortality rates and their Appendix Table 2 reports effects of the shock on mortality, by five-year categories. I use their mortality data to compute a 1946 survival curve for women and then use their regression estimates to construct a perturbed survival

The next study is by Fortson (2011) who examines the expansion of the HIV/AIDS epidemic through sub-Saharan Africa. She also uses a difference-in-difference design in which areas that received a greater rise in infection also experience greater across-cohort differences in the expected marginal benefits of schooling (through Δ). She finds statistically significant declines in school attendance amongst the most affected cohorts, and that such a response is consistent with an elasticity of school to marginal benefits of approximately 0.9.

Oster, Shoulson, and Dorsey (2014) examine a sample of people at risk to get Huntington's disease, which would significantly curtail their lifespans. They find a substantial drop in school attendance (in college, in this case) among the subsample that gets early confirmation that they have the disease. They report elasticities between 0.7 and 1.3 for, respectively, education w.r.t. life expectancy and w.r.t. "discounted lifetime return to investment" (page 1995, Table 6, first row). (There was no effect on high-school completion, consistent with the late timing of revelation of the disease.)

A few additional studies can be also used to estimate this elasticity for undiscounted life years.[11] Baranov and Kohler (2017) find a significant increase in schooling associated with increased regional exposure to anti-retro-viral therapy (used to combat HIV/AIDS) in Malawi. They do not report an elasticity, but their estimates are consistent with an elasticity of 0.9. Hansen (2013) examines within-country changes in life expectancy by using a shift-share instrument for disease burden. Although he does not report an elasticity, I can combine his point estimate for the effect of life expectancy at age 20 on schooling (0.3) with average values of s and e_{20} to compute an elasticity of approximately 0.8.[12] Cohen and Leker (2014) perform an exercise similar to Hansen's, but they did not report enough information to construct an elasticity. Stoler and Meltzer (2013) relate education to the timing of Huntington's onset status in a different sample from that in Oster *et al.* They report a slope similar to that in Jayachandran and Lleras-Muney's study, which translates

curve. I then use these to compute a Δ_1 and Δ_0.

[11] I thank Victoria Baranov and Casper Worm Hansen for providing a few additional numbers that were necessary to construct an elasticity with their estimates.

[12] It is noteworthy that his 0.3 coefficient is triple the slopes estimated in studies discussed above, yet the elasticity is similar.

into an elasticity of approximately 0.4 w.r.t. undiscounted working lifetime. Outside the realm of health, Bleakley and Hong (2014) argue that the post-Civil-War declines in returns to skill and in school attendance among white Southerners were consistent with an elasticity of 0.6-1.3.

These studies examine episodes widely separated in space and in quite different contexts, yet they obtain elasticities near or below unity. Conservatively, I take 1 as a point estimate for the elasticity and use it to calibrate the triangle. Because unity is on the high side of such estimates, the calculation remains something of an upper bound, albeit a tighter one than in Section 5 above. It is also an upper bound in the sense that (a) longevity gains past working ages are counted, (b) taking less developed countries to the Japanese life table might be an unrealistic assumption, and (c) labor-supply responses might attenuate the effect on lifetime hours (see Hazan, 2009, Hansen and Lønstrup, 2012, and Bloom, Canning, and Moore, 2014).

The use of an elasticity requires a small modification to the triangle formula. First, note that $s_1 \approx s_0 + (\Delta_1 - \Delta_0)\frac{ds^*}{d\Delta}$, by a first-order Taylor approximation. The triangle from equation 3 is therefore

$$T = \frac{1}{2}\left(\frac{\Delta_1 - \Delta_0}{\Delta_0}\right)^2 \frac{f'(s_0)}{f(s_0)}\Delta_0 \frac{ds^*}{d\Delta} \tag{8}$$

$$= \frac{1}{2}\left(\frac{\Delta_1 - \Delta_0}{\Delta_0}\right)^2 \beta s_0 \epsilon \tag{9}$$

in which $\epsilon \equiv \frac{ds^*}{d\Delta}\frac{\Delta}{s^*}$, the elasticity of optimal schooling w.r.t. Δ. Notice that T now has a form more familiar to those who use Harberger triangles. The elasticity appears linearly, and the distortion (a survival difference rather than a tax) appears in squared form.

Figure 6 displays a scatterplot of Δ versus s_0 in the sample of countries. Overlain on the scatter are a series of curves that trace out sets of points for which the triangle has a constant value (iso-triangle curves, if you will). Notice that these curves are associated with higher triangle values as one moves in a south-easterly direction on the graph. This is because such locations have both less healthy countries and a larger initial stock of schooling to be multiplied by the (constant) elasticity. The former means the height of the triangle is greater and the latter means that the base is greater. The vast majority of countries have triangles that are less than half a percent (to the northwest of the $[T = 0.5\%]$ line). Perhaps this is not altogether surprising in that many of these are developed countries, but many of them are not. Egypt and Colombia, for example, appear in the graph with approximately 10 years of school and 32 years of discounted lifetime (from age 15). This position in the graph gets them a horizon triangle of less than half a percent of income. Only five countries have triangles above 2%, and these are for countries whose life tables looked especially bad for transitory reasons. These exceptional countries notwithstanding, the average triangle for countries in the lowest quintile of e_{15} or s_0 is is only 1.5% or 1.0%, respectively.

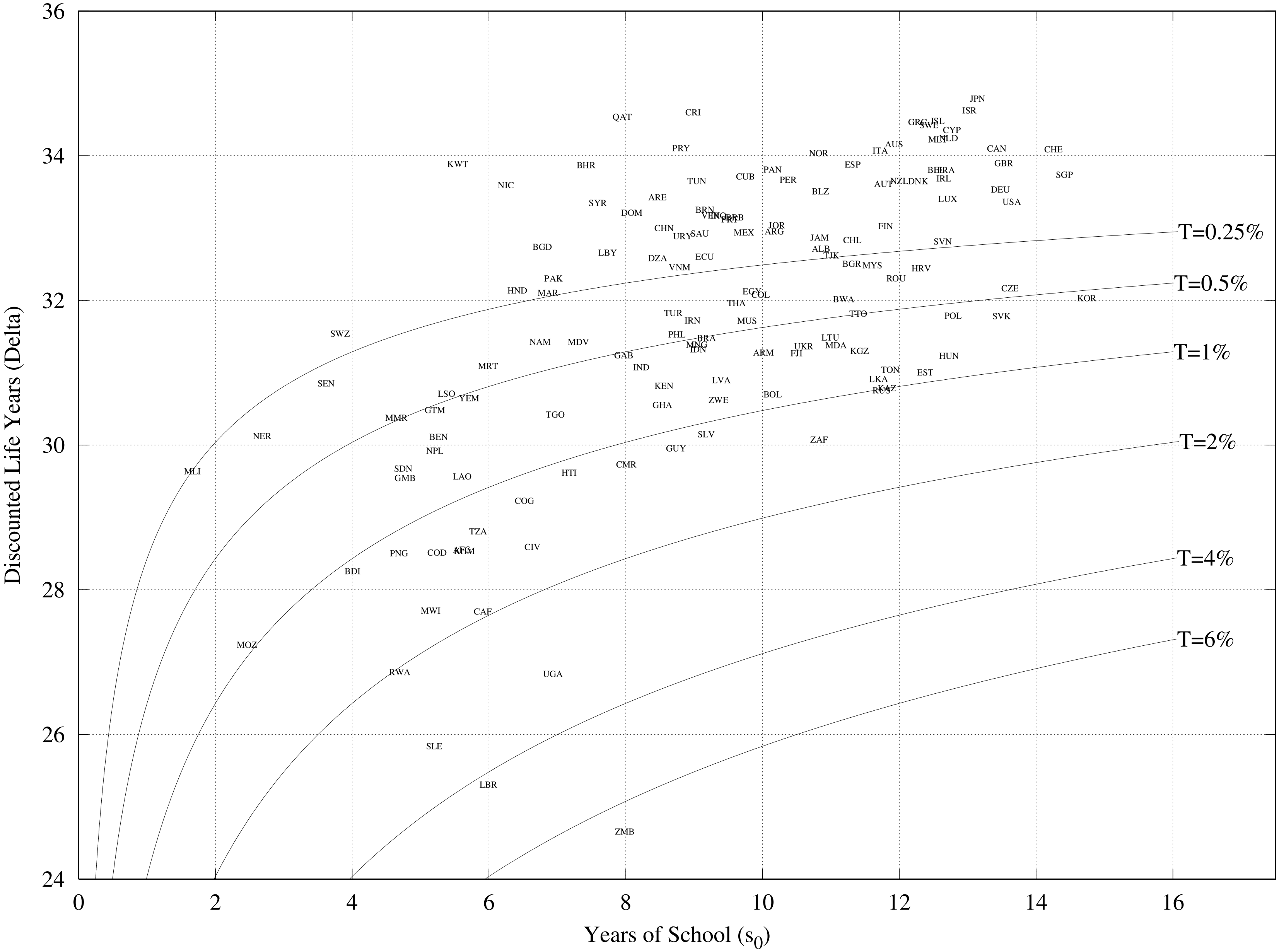
Note finally that using an assumption of constant slope rather than constant elasticity does not present a more favorable case for large welfare gains. (Appendix Figure I presents the same results for a constant-slope assumption instead, as in equation 8.) Jayachandran and Lleras-Muney (2009) estimated this slope ($\frac{ds^*}{d\Delta}$) to be approximately 0.11, and the other within-country studies cited above obtain similar numbers. The largest triangle in the sample is almost a factor of 3 smaller, and a regression of 'slope' triangles on 'elasticity' triangles has a coefficient of 0.44. By equation 8, triangles get larger only for Mali, Mozambique, and Niger, but only by a bit. If we instead use Hansen's slope of 0.3, approximately half of the triangles get bigger than the constant-elasticity version, with a regression coefficient of 1.21. Nevertheless, none exceed 6.25%, and only 6 of 141 are greater than 3%.

7 Sensitivity analysis

In this section, I check the sensitivity of the estimated triangles to alternate assumptions.

The transparency of the triangle calculation permits a degree of "do it yourself" sensitivity analysis. Consider equation 9 above. The triangle rises like the square of the proportional change in Δ. Thus, if the reader judges it more realistic to take the poorest countries only halfway to the Japanese survival curve, then the resulting triangle of gains would be smaller by a factor of 4.

Figure 6: Iso-Triangle Curves for Unit Elasticity of Schooling to Δ



Notes: this figure displays a scatter of discounted life years (Δ) versus initial years of school (s_0) for the sample of countries. Observations are denoted by their World Bank country codes. Overlain on the scatterplot are iso-triangle curves; these trace out constant values for the triangle (T) of reallocation gains. T is computed using equation 9. The underlying data sources and variable definitions are described in the text.

The triangle is also linear in the Mincer β. As mentioned above, 10% is a common number in the recent literature, although there are reasons to believe this number might be somewhat higher or lower. On the one hand, some readers might believe that estimates of this coefficient are biased upward if those who get more school are also more able along other dimensions. A lower true β would simply reduce the triangle proportionately. There do exist studies using instrumental variables that claim the true β might be higher than implied by cross-sectional estimates, but not massively higher. Readers are invited to plug in their preferred value for β.

The final parameter, also easy for the reader to change, is the elasticity of school with respect to Δ. Estimates above suggest this number is close to unity, perhaps a bit below. But perhaps this elasticity is 2, or perhaps it is 1/2. In either case, the triangle would change by the same factor. How far can tweaking this parameter inflate the triangle? Presumably the calculation becomes more implausible as we increase to an elasticity that would imply these countries jumping into the global education frontier if we did nothing but change the survival curve to Japanese levels. Thus the results in Section 5 are slack upper bounds on the triangles one might obtain by cranking up the elasticity.

Next, the triangle should be attenuated by standard general-equilibrium effects. A higher Δ would increase the supply of various types of labor, perhaps depressing wages. The accompanying increase in schooling would render labor more skilled on average. Consider separately the change in the base and height of the triangle. The height measures the gap between MB and MC of schooling if schooling is fixed at s_0. The height would shrink no more than proportionately with the wage.[13] By how much should the wage decline? According

[13] Consider the labor share of direct cost (c) associated with providing a unit of school at a given level. If this share is one (i.e., labor is the only expense), then the height shrinks proportionately with the reduced wage. If this share is less than one, then the height is reduced by less than the wage change, in proportional terms. Let w_i be an index of wages in scenario $i \in \{0,1\}$ and the wage at s_0 be $w_i f(s_0)$. Let the marginal direct cost of school be $c(s; w_i, \vec{p})$, in which $\vec{p}$ is the vector of prices for non-labor inputs to school, e.g. buildings. The height is now $\tilde{h} \equiv w_1(f'(s_0)\Delta_1(s_0) - f(s_0)) - c(s_0; w_1, \vec{p})$. Height reponds to wages as follows: $\frac{\partial \tilde{h}}{\partial w} = (f'(s_0)\Delta_1(s_0) - f(s_0)) - \frac{\partial c}{\partial w}$, which can be re-written as $h + \frac{c(s_0; w_0, \vec{p})}{w_0} - \frac{\partial c}{\partial w}$, by the '0' FOC. The second two terms are weakly positive iff $1 - \frac{\partial c}{\partial w} w_0 / c(s_0; w_0, \vec{p}) \geq 0$. If we apply Shephard's Lemma, we see that this expression is the share of direct costs associated with non-labor inputs in producing schooling at the s_0 level. This share must be non-negative, so $\tilde{h} < h$ if $w_1 < w_0$.

to Weil and Wilde (2009), a typical very-low-income country has a fixed-factor share of perhaps 0.15. In a Cobb-Douglas economy open to capital flows, the wage would therefore be 0.85 of its value above, and the height no smaller than that. Furthermore, for a fixed elasticity, the triangle drops like the square of the change in MB, for a reduction by a factor of $0.85^2 = 0.7225$ at most. The base instead measures the response of schooling to a given shock to marginal benefits. The additional supply of skill in this economy should reduce the price of skill, which would imply that the optimal choice of schooling is lower than in the partial-equilibrium analysis. Because the empirical studies above examined subnational differences in exposure to Δ, we should think of those elasticities as overestimates for the response that would obtain in general equilibrium. Thus the triangle would also have a smaller base.

7.1 Alternate versions of Δ

I now consider the sensitivity of the results in Section 6 to parameters that were used in constructing Δ. While the construction of Δ is straightforward, it is not quite so simple as to allow for do-it-yourself analysis as above. Thus, Table 1 includes calibrated values of the triangle for various assumptions and for various sets of countries. Countries are ranked according to indices of human development: years of schooling, life expectancy at age 15, and the product of the two. I then separate countries by quintiles according to these indices. Results for the first (lowest), second, and third (middle) fifths of each index are across the columns. Throughout this section, I use an elasticity of one for the response of school to Δ. If unit elasticity is not preferred, readers could simply multiply the values in Table 1 by their preferred elasticity.

For the first set of rows, I compute triangles for various discount rates. Recall that the default value used throughout the study is 2%, and results for this value are shown in row (5). In the lowest fifth of countries, the average triangle is between 1% and 1.5%, depending on the human-development indicator used. For the second fifth, average triangles are less than 1%, and close to 0.5% for the middle fifth. For higher interest rates (see Rows 6–7), triangles are significantly smaller, as should be expected. For example, if $r - g = 5\%$, the triangles in the lowest quintile are less than 0.04%. If the interest rate is lower, however, the triangles are larger. For $r - g = 0$, the average triangles in the lowest fifth are no greater than 4% of income, and significantly smaller in the higher fifths. Even in the extreme case of a negative discount rate ($r - g = -1\%$), the average triangles do not exceed 6% of income for the poorest countries.[14]

I also show that the result is not sensitive to the way in which life tables were mapped to cohorts. Recall that the default Δ is constructed starting at age 15 from the circa-1990 life table using cohorts that were born in the five years prior to 1980. Rows (8) and (9) present results with a Δ constructed starting at age 10 or age 20, respectively. In Row (10), I attempt to align the starting age for Δ more closely to the school-leaving age in each country. Specifically, I take the average schooling for the reference cohort in each country, round it up to the nearest integer, if necessary, and add 7 to account for a typical school-starting age. This becomes a country-specific starting age for the evaluation of Δ. Next, in Rows (11) and (12), I use cohorts that are born five years before or five years after the reference cohort. Then, in Row (13), I construct the triangles using the life table and schooling data for females rather than males. The resulting triangles in this analysis are similar in magnitude to the baseline, seen in Row (5).

Next, I consider the effect of changing the health frontier for the thought experiment. For Row (14), I imagine taking countries to the life table of the US rather than of Japan. From the distant perspective of a very unhealthy country, the US and Japan have quite similar health outcomes, but triangles are about 1/3 smaller when computed using the US survival curve. This is because the triangle rises like the square of the distortion, so that last bit of difference in Δ has an outsized impact on T. In Row (15), I conduct what is perhaps a more realistic exercise for health improvement: take unhealthy countries to the life table of

[14]Note the contrast with Preston's (1980) argument that life expectancy gains, once discounted, are too small to explain large differences in education across countries. Here, I instead show that, even for zero discounting, the triangles are limited in size. In the spirit of Preston's calculation, Appendix C presents analysis of the additional rate of return, conditional on survival, as a function of mortality and the net interest rate. I show that the required return on human capital is higher because of mortality, and that the spread for the best versus the worst life table is approximately 100 basis points.

Thailand. (Thailand is close to the median of both s_0 and e_{15} in these data.) The calibrated triangles are approximately 0.5 percent for the lowest quintile, about 2/3 lower than the baseline. Next, I consider truncating the horizon over which Δ is constructed. In Rows (16) and (17), I use the survival curve up to ages 80 and 65, respectively, and therefore ignore survival gains at higher ages. A significant fraction of the gains in longevity in the baseline exercise arise from increased survival late in life, and therefore the triangles are smaller, especially in the case of truncation at age 65. For Row (18), I truncate the calculation of Δ at age 65, but set $\ell(t)/\ell(s)$ equal to 1. This simulates taking the survival curve in these countries past the healthy Japanese one and all the way to the extreme of zero mortality during working ages. The triangles in this case nevertheless are smaller than the baseline. Why? Because the greater proportion of differences in Δ, when comparing unhealthy to healthy countries, occur at relatively old ages.

What if the truncation point (i.e., retirement age) is itself a response to the life table? Bloom, Canning, and Moore (2014) show that, in a calibrated model, increased longevity should postpone retirement, while increased income would have the opposite effect. They conclude that any secular trend towards earlier retirement is therefore driven by income rather than longevity effects. Because the thought experiment of the present study is to consider a change in the horizon, *ceteris paribus*, it is worth analyzing the effect of postponing retirement. Recall Row (17), in which I use a truncation horizon of 65 years of age, approximating a retirement age. In this calculation, the lowest-Δ country (for $r - g = .02$) is at 23.7 years, approximately 7.5 years behind Japan. Suppose that we were to add to this an additional five years of work. Fifty years earlier, the 2% discount rate would make this look like 1.84 years. This increases the $(\Delta_1/\Delta_0 - 1)$ term by a factor of 1.25, which squares to 1.57. Thus, the numbers in Row (17), for the lower fifth of countries, would inflate by a somewhat smaller fraction.

I then add more flexible estimates of returns to experience to the discount factor Δ. Above this was proxied by an exponential growth rate, g. Mincer's original work used instead a growth rate that was quadratic in potential experience, while more recent work has favored a quartic polynomial (Murphy and Welch, 1990, and Lemieux, 2006). I use estimates from Lemieux (2006) to calibrate a wage/experience profile.[15] I maintain the truncation at 65, because estimates of returns to experience are for working ages and it would make little sense to extrapolate such estimates out to centenarians. Results are shown in Row (19). The updated triangles are larger than their counterparts in in Row (17), but only by a bit. Note that, for this last calculation, I also maintain a 2% net discount rate $(r-g)$. This causes some double-counting in that g already captures some of the returns to experience. Finally, in Row (20), I adjust instead for age- and country-specific patterns of disability using the WHO's *Disability-Adjusted Life Years* (DALY). I use the fraction of time lost to disability to scale down the fraction surviving at a specific age $(\ell(t))$ and find that the triangles are little changed from the baseline. (See Appendix E for details on these data, calculations, and analysis as to why this adjustment has little effect.)

A final remark is in order on Mincer's assumption (used throughout) that the derivative of the wage w.r.t. schooling is multiplicatively separable from the rest of the discounting terms. Heckman *et al.* (2006) show that Mincer's separability result can be statistically rejected in recent decades (see Table 1, page 324). Two points about their results are worth noting. First, within the first 30 years of labor-market experience, the wage ratios between the college and high-school educated are relatively stable, with a maximum difference over time of no more than 0.02 per year of schooling. Second, the greatest departure from Mincer's result occurs at approximately 40 years of experience, when some convergence is seen in the college/high-school gap. So far out, such convergence matters less from a period-s perspective. For the present purposes, this finding would tend to decrease my estimate of the triangle based on an (f'/f) that is stable over one's career. Instead, the supposed health improvements tend to add additional survival probabilities in years with relatively low marginal benefit of schooling.

[15] For Lemieux's Figure 11.8, he uses *circa*-2000 US data and reports log wages growing with experience approximately like the following quartic:

$$\hat{g}(t-s) \equiv .09020(t-s)^1 - .005086(t-s)^2 + .000139(t-s)^3 - (1.445 \times 10^6)(t-s)^4.$$

For Table 1, Row (19), I add $\exp(\hat{g}(t-s))$ inside the integral for Δ.

Table 1: Sensitivity of Triangle to Assumptions in Constructing Δ

			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Quintile:	Bottom			Second			Middle		
		Variable used for ranking:	s_0	e_{15}	s_0 x e_{15}	s_0	e_{15}	s_0 x e_{15}	s_0	e_{15}	s_0 x e_{15}
(1)	Alternative (r-g)	-1%	3.58	5.55	4.98	3.01	2.80	1.79	1.31	1.54	1.48
(2)	"	0%	2.31	3.57	3.21	1.93	1.76	1.15	.83	.95	.92
(3)	"	0.5%	1.85	2.86	2.58	1.54	1.39	.92	.66	.74	.72
(4)	"	1%	1.49	2.29	2.07	1.23	1.09	.73	.52	.57	.56
(5)	"	2% (baseline)	.97	1.48	1.34	.79	.67	.47	.33	.35	.35
(6)	"	2.5%	.78	1.20	1.08	.63	.53	.37	.26	.27	.27
(7)	"	5%	.29	.44	.40	.22	.17	.13	.09	.08	.09
(8)	Initial ℓ(s)	10	.82	1.24	1.12	.65	.54	.39	.26	.28	.28
(9)	"	20	1.12	1.79	1.61	.98	.84	.56	.40	.43	.42
(10)	"	s_0	1.25	1.11	.77	.68	.44	.73	.37	.34	.31
(11)	Ages in 2010	25-29	1.02	1.56	1.39	.78	.70	.53	.36	.36	.37
(12)	"	35-39	.93	1.38	1.10	.63	.68	.61	.40	.34	.31
(13)	Females		.79	1.08	.86	.66	.61	.68	.34	.25	.34
(14)	'Frontier' country	USA	.58	.93	.85	.44	.26	.19	.11	.08	.11
(15)	"	Thailand	.32	.53	.49	.25	.05	.07	.06	.01	.06
(16)	Truncation Horizon	80	.80	1.24	1.13	.65	.52	.37	.25	.25	.26
(17)	"	65	.45	.71	.65	.36	.24	.19	.12	.10	.12
(18)	"	65,ℓ=1	.64	.97	.88	.54	.44	.34	.26	.25	.25
(19)	+ Experience profile	65	.53	.83	.77	.43	.27	.21	.14	.11	.13
(20)	+ Disability	65	.47	.74	.68	.38	.25	.20	.13	.11	.12

Notes: this table presents sensitivity analysis for estimates of the average triangle (T) for subsamples of countries. Except as noted, the calculations assume a net discount rate ($r - g$) of 2%, use a reference age of 15 years, take Japan as the frontier for Δ, and use as a reference cohort those born in the five years prior to 1980. These calculations are based on the unit-elastic response of schooling to Δ, as estimated in various of the studies cited in Section 6. Triangles are computed according to equation 3. Countries are grouped into fifths based on either initial years of schooling, life expectancy at age 15, or the product of the two. Results for the first (bottom), second, and third (middle) fifths of countries are reported in Columns 1–3, 4–6, and 7–9, respectively. Various assumptions for the construction of Δ are noted in the row headings. The underlying data sources and variable definitions are described in the text.

7.2 What if education reduces mortality?

How does the problem change if education itself has a causal effect reducing mortality? In many datasets (within a country), more education is associated with lower mortality, often many years after graduation. See, for example, Cutler and Lleras-Muney (2006) and Lleras-Muney (2005) and references therein. Many studies focus on mortality rates within specific age groups, typically older ones. For the present purposes, I require estimates that at least include working ages. Notably, Cutler and Lleras-Muney (2006) report a gain of approximately 0.2 undiscounted adult life expectancy per year of school, as does Sanchez-Gonzalez (2011, chapter 5).[16]

The modified triangle, $\tilde{T}$, that accounts for this effect is as follows:

$$\tilde{T} = \frac{1}{2}(s_1 - s_0)\beta \left[\left(\frac{\Delta_1 - \Delta_0}{\Delta_0} \right) + \frac{1}{\Delta_0 \beta} \left(\frac{d\Delta_1}{ds} - \frac{d\Delta_0}{ds} \right) \right]$$

in which the Δ concepts are evaluated at s_0. (See Appendix A for a derivation.) The second term in the square brackets is new. Notice that the additional marginal benefits depend on the *difference* between mortality scenarios in the survival/education gradient. In either scenario, a gradient would already be reflected in the FOC. Thus, an effect of education on survival increases the height of the triangle only if the survival/education effect is *stronger* in the low-mortality environment. The opposite seems at least as likely, especially at working ages, when mortality rates are relatively low for all education levels in rich countries. (The relevant units of the education/survival gradient are probabilities, not in odds ratios.) So, accounting for this effect might well shrink the triangle.

A plausible upper bound on this additional effect turns out to be small. I start with the assumption that $\frac{d\Delta_0}{ds} \geq -1$, because I know of no study finding a systematically negative education/survival gradient. The studies mentioned in the previous paragraph suggest that $\frac{d\Delta_1}{ds} \approx -0.8$, if we take the US as representative of the education/longevity gradient at the frontier. A country with very low human development (e.g. in the bottom left corner of

[16] Their estimates apply to the US. I am not acquainted with studies conducting a conducting a comparable exercise for Japan or for any of the least developed countries.

Figure 3) has $\Delta_0 \approx 45$, for $r - g = 0$. The β is assumed to be roughly 10%, as above. Putting these together reveals the new term to be $0.2/45/0.1 = .044$, at most. Compare this with the increase in marginal benefits through the original channel: Δ increases by roughly 40%, an order of magnitude larger. Thus, the height of the triangle would be 1.11 times the original estimate, which would increase the change in schooling by a similar fraction. In sum, the triangle would be at most a factor of 1.23 of those originally computed.

7.3 Misallocation I: Rationing/Constraints

Above I assumed that schooling was set optimally in both mortality scenarios. Indeed, this assumption is what makes the triangle a triangle. But what if schooling were below the optimum? The case of underinvestment is seen in Figure 7. The gray dashed lines denote the optimal choices of schooling in high- and low-mortality environments. For some reason, schooling levels are rationed below the respective optima, to s_0 and s_1. As the benefits of schooling occur in the future, the schooling decision could be distorted by psychological 'present' biases or imperfect intergenerational altruism.

In the low-mortality scenario (MB'), the gains of reallocating from s_0 to the optimum form a triangle composed of areas $a + b + d$. But a constraint prevents d of the gain from being realized. Further, the gain a arises only if a constraint is relaxed. Thus, only the area b is directly attributable to the horizon mechanism.

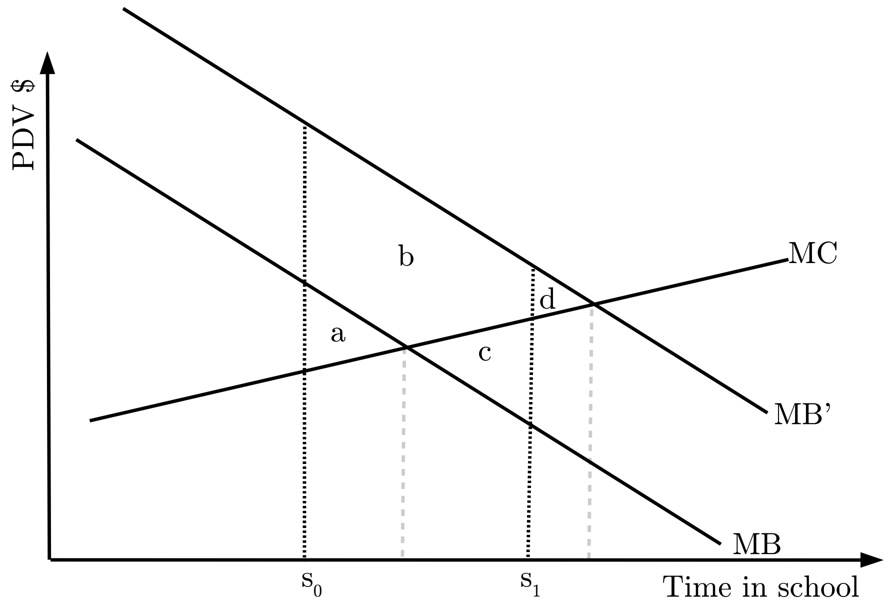
How large is the horizon polygon b? A point estimate would require knowledge of $c(s)$, $f''(s)$ and the optimal s to estimate the sizes of a and d. Instead, note that

$$(b + c) \;\approx\; (\Delta_1(s_0) - \Delta_0(s_0))f'(s_0)(s_1 - s_0)$$

The right-hand side is equal to $2T_{\$}$, which implies that the horizon-mechanism gain is $b \leq 2T_{\$}$. Thus, even in the presence of constraints on schooling, we can put an upper bound on the gains from reallocation because of the horizon mechanism by simply doubling our estimate of the triangle.[17] Even though this number would be twice as big as before, it

[17]This bound holds for all four combinations of rationing above or below in either mortality case. Indeed,

Figure 7: Under-investment in schooling and the horizon polygon (no longer a triangle)



Notes: this is a modification of the model in Figure 1 to allow for underinvestment in schooling in both scenarios.

would remain small relative to, e.g., cross-country gaps in income.

7.4 Misallocation II: The Price is Right?

In the absence of distortions, the triangle provides a complete (albeit approximate) accounting of gains from the horizon mechanism. As James Hines (1999) states in his *Journal of Economic Perspectives* piece on Harberger triangles,

> it is not necessary to take explicit account of spillover effects into undistorted markets in calculating deadweight loss, since price and quantity changes in undis-

if schooling is misallocated to be *too high* in the low-mortality environment, the horizon gains are strictly less than $T_{\$}$, rather than $2T_{\$}$. This upper bound would also apply to $b+a$, if the optimal schooling in the high-mortality scenario (call it s_0^*) is less than or equal to the midpoint between s_0 and s_1. Why? Because a and c are each a second-order Taylor approximation, which is symmetric about s_0^*. Therefore,

$$s_0^* \leq \left(\frac{s_0 + s_1}{2}\right) \Longleftrightarrow a \leq c.$$

This is also relevant in the next subsection, for which a is rightly considered part of the cost of misallocation.

> torted markets do not affect the efficiency of resource allocation—after all, in such markets, marginal values to consumers equal marginal costs of supply. (p. 178–9)

Yet treating the economy as otherwise undistorted is unrealistic. Indeed, distortions are commonplace, from business formalization to tortilla subsidies. Nevertheless, analysis of how education affects each of these distortions lies outside the scope of this study.

A modest goal is instead to analyze a distortion of the education decision itself. First, note that society's decision problem could be similarly modeled as the intersection of two curves (social marginal benefits and social marginal costs, in this case). Absent a distortion, society's optimization problem looks like Figure 1 and the model in Section 3. Second, if the private decision is misallocated (from the perspective of society) then we can simply use an upper bound of double the triangle, as in the prior section. This requires only one modification to the triangle formula, specifically that we use an f' that includes the external effects (in effect, society's rather than the individual's β). Fortunately, there are estimates in the literature using credible natural experiments to measure such externalities. Acemoglu and Angrist (2000) estimate externalities of around 1% per average year of education at the US-state level (with a confidence interval that reaches slightly below zero and almost to 3%). Moretti (2004) argues that this number (at the US-metro-area level) is somewhat higher, perhaps 5%. Duflo (2004), in contrast, estimates negative spillovers from education in Indonesia. (It was not evident to me how to convert Duflo's estimates into the units of Mincer's β.) Respectively, these studies imply a social β of 11%, 15%, or $< 10\%$.

8 Changes over the 20th century

For select countries, I repeat the analysis above but, rather than comparing across countries today, I compare within countries over the 20th century. Many countries experienced large gains in both life expectancy and schooling over this time, yet the horizon triangle is small.

The data sources and methods are largely as in prior sections. The Barro/Lee data also include information on cohorts born in the early part of the 20th century. The historical life tables are drawn from *Human Life-Table Database* (Max Planck Institute for Demographic Research, 2016). (Appendix D has further information about and analysis of the data.)

Table 2 shows the resulting triangles. The first several columns show the inputs to this calculation: the proportional change in Δ, the change in years of school, and initial years of school. I present calculations for three sets of assumptions for discounting: no discounting, 2% per year, and 2% per year but ignoring gains after 65 years of age. Once again, I compute both an upper bound on the triangle and a triangle that assumes unit-elastic response of schooling to Δ. I maintain the use of 10% for Mincer's β and of 0.5 for one half.

Table 2: Triangle Calculations for the 20th Century, Select Countries

	Data			Upper bound		Unit-elastic response	
	Proportional change in Δ	Change in school, (s_1-s_0)	Initial years of school (s_0)	Implied elasticity	Triangle	Fraction of school explained	Triangle
			Panel A: no discounting (Δ=e_{15})				
USA	.17	3.6	9.9	2.2	.031	.46	.014
Japan	.46	5.1	7.9	1.4	.117	.71	.082
China	.31	6.7	1.7	12.6	.103	.08	.008
Egypt	.50	8.6	1.1	15.8	.217	.06	.014
India	.64	6.4	1.7	6.0	.204	.17	.034
			Panel B: discount at 2%/year				
USA	.11			3.4	.020	.29	.006
Japan	.29			2.2	.075	.46	.034
China	.22	″		17.8	.073	.06	.004
Egypt	.37			21.3	.160	.05	.008
India	.44			8.6	.142	.12	.017
			Panel C: discount at 2%/year, truncate at 65 years				
USA	.07			5.4	.012	.18	.002
Japan	.20			3.2	.051	.31	.016
China	.17	″		22.4	.058	.04	.003
Egypt	.33			24.4	.140	.04	.006
India	.37			10.2	.120	.10	.012

Notes: this table presents estimates of the triangle based on the change in Δ within select countries. The Δ are computed using 1930 and 1990 life tables, while the school data refer to those born circa 1917 and 1977. The table presents both the upper bound ($\bar{T}$), which assumes all of the change in school and is due to the horizon mechanism, and also an estimate (T) assuming a unit-elastic response of schooling to Δ. Various assumptions about discounting are used for the panels. The underlying data sources and variable definitions are described in the text.

The results are in line with the cross-country evidence above. The upper bounds can be quite high. For example, the upper bounds with no discounting are approximately 20% for India and Egypt. Nevertheless, this would imply somewhat high elasticities, exceeding 20 in some cases. Using an elasticity of 1, I find triangles that are smaller, some almost negligible. The only numbers that exceed 4% are for Japan. However, because so much of the gain in the Japanese survival curve occurred at old ages, this result is only 1.6% when restricting our attention to gains at working ages. By this metric, the triangles for Egypt and China are substantially below 1%, and the triangle for India is only 1.2%.

For the US, the triangle gains are also small. Even at the upper bound and with no discounting, the triangle is approximately 3% of income. For more realistic discounting and ignoring survival gains during retirement years, the triangle is less than 0.02%. The US has smaller triangles in this calculation in part because its improvements in health and education were less concentrated in the 20th Century as compared to the other countries. Using instead David Hacker's (2010, Table 8) survival curve for white males in 1850 and an $r = 2\%$, I compute a Δ of 26.4 discounted years (and an e_{15} of 41.8). The proportional gain in Δ from 1850 to 1990 therefore is 0.26, which is comparable to the gains in China and Japan from 1930–90. Using data from Gould's (1869) sample of soldiers in the US Army, Bleakley, Costa, and Lleras-Muney (2014) report an average education of approximately 5.5 years for those born in the 1830s. If $\beta \approx 0.1$, as assumed above, then the triangle associated with giving the 1990 life table to the 1830 cohort is estimated to be 0.019. This is comparable to triangles for India and triple the estimate for the US for 1930–1990.

In all, this evidence indicates that the triangle associated with the horizon mechanism is quite small for these countries and cohorts. Furthermore, because the thought experiment involves instantaneously giving the earlier cohort the later life table, these triangles are much smaller than the actual gains experienced along the transition to the new survival curve.

One other feature is worth noting: the response of schooling to survival. The penultimate column of Table 2 shows the fraction of the increased schooling that is explained by the changing horizon. Notice that these numbers can be quite large: 18% and 31% for the USA and Japan, respectively, even in the relatively conservative case that discounts the future at 2% and ignores survival gains past age 65. Yet the associated triangles are an order of magnitude smaller than the fraction of schooling explained.

9 Rectangles versus Triangles

While this study treats the triangle of reallocation, there are two rectangles (i.e., first-order changes) that bear mentioning. Neither is a measure of the welfare gain, although the first is closer than the second. (Becker, Philipson, and Soares (2005) examine the direct welfare gain from shifting up the survival curve, but not any triangles of reallocation.)

The first rectangle R measures the proportional increase in discounted, expected life years (Δ). For the hypothetical transition to the Japanese survival curve, $R \equiv \frac{\Delta_{JAPAN}-\Delta_0}{\Delta_0}$ in which 0 denotes the baseline status for that country. R is related to wellbeing, although the full welfare calculation also depends on the value of life and the distribution over the life cycle of survival gains relative to that of borrowing.[18] Panel A of Figure 8 contains a plot of R versus the triangle T for the sample used above. This rectangle ranges from zero (for Japan, by construction) to over 40%, albeit with most of the observations below 15%.

[18] Consider the generic lifetime-consumption problem with time-separable utility, $u_t(c_t)$. The Lagrangian for this problem is as follows:

$$\mathcal{L} = \int_0^{\infty} e^{-\beta t} \left[\ell_t u_t(c_t) + (1-\ell_t)\hat{u}\right] dt + \lambda \left(\int_0^{\infty} e^{-rt} \ell_t y_t dt - \int_0^{\infty} e^{-rt} \ell_t c_t dt \right)$$

in which y_t is income, c_t is consumption, β is the subjective discount rate, $\hat{u}$ is the utility of not being alive, and λ is the shadow price on the budget constraint. Taking the derivative of $\mathcal{L}$ w.r.t. the survival probability in period t, I find that

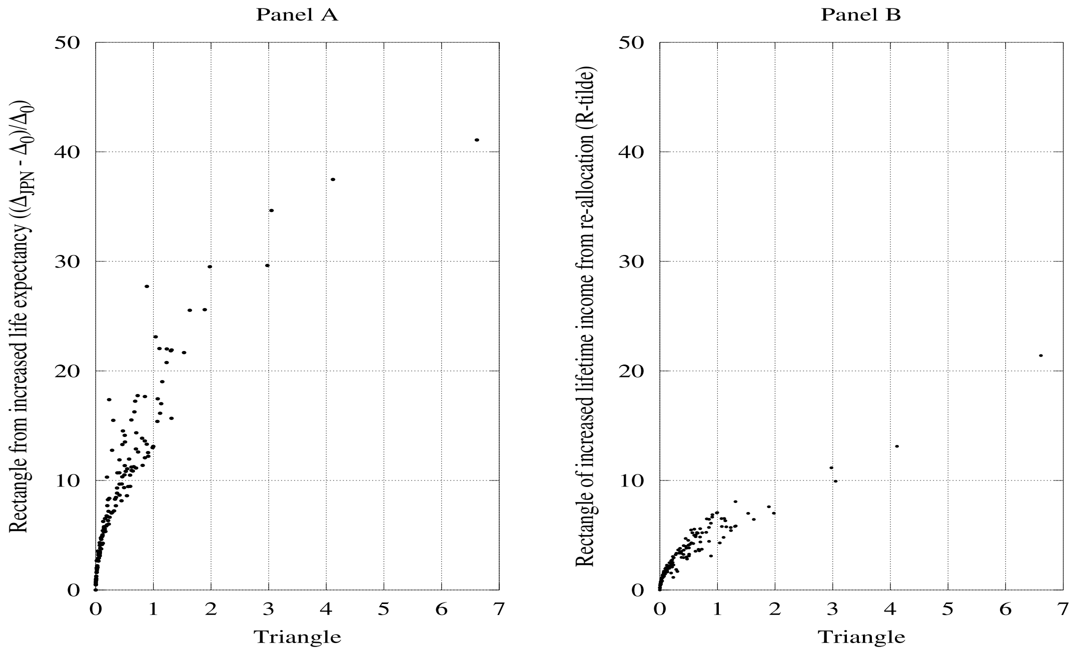
$$\frac{1}{\lambda} \frac{\partial \mathcal{L}}{\partial \ell_t} = \frac{e^{-\beta t}}{\lambda} \left(u_t(c_t) - \hat{u}\right) + e^{-rt} \left(y_t - c_t\right),$$

which is the gain in utility translated to period-0 dollars. The first term is the direct effect: increased survival in period t yields extra expected utility, if the utility of being alive exceeds the utility of being dead. The second term measures the effect of survival on lifetime consumption. Increased survival in periods with expected saving relaxes the lifetime budget constraint, while in periods of dissaving it contracts the budget constraint. Compare this with the derivative of Δ w.r.t. ℓ_t, evaluated from a $t = 0$ perspective:

$$\frac{\partial \Delta}{\partial \ell_t} = e^{-rt} y_t$$

in which only the present value of period-t income appears. Thus the change in Δ is part, but not all, of the change in welfare.

Figure 8: Two Rectangles, compared with the Triangle



Notes: this figure presents estimates of two rectangles: the direct gain from increased life expectancy R and the increase in discounted lifetime income due to re-optimization ($\tilde{R}$). Both are plotted against the triangle (T) as computed above (in Figure 6, for example). For the simulation, the frontier country for the survival curve is Japan and the elasticity of school to marginal benefits is one. Derivations of R and $\tilde{R}$ are found in Section 9. The underlying data sources, variable definitions, and parameters are described in the text.

The roughly parabolic relationship is as expected (see equation 3). Note as well that the rectangle dwarfs the triangle. This is also as expected; R is a first-order effect, while T is second order.

A second rectangle $\tilde{R}$ measures the increase in discounted lifetime income from the increase in s. In the presence of direct costs of schooling ($c(s) > 0$), an increase in schooling (local to the first-order condition) does raise lifetime income. This occurs because, in the problem above, we maximize lifetime income net of costs, not lifetime income itself. Yet we can compute $\tilde{R}$ by accumulating the gaps between marginal benefits ($\Delta_1(s)f'(s)$) and the marginal *opportunity* cost only, $f(s)$, as s shifts from s_0 to s_1. This object can be approximated as a trapezoid. Its width is $s_1 - s_0$. The vertical edges are the marginal gains in lifetime income, $\Delta_1(s)f'(s) - f(s)$, evaluated at s_0 and s_1. By the formula for the area of a trapezoid,

$$\tilde{R} \equiv (s_1 - s_0) \left(\frac{\Delta_1(s_0)f'(s_0) - f(s_0) + \Delta_1(s_1)f'(s_1) - f(s_1)}{2} \right) \left(\frac{1}{\Delta(s_0)f(s_0)} \right),$$

with the final term normalizing the object to be in units of lifetime income. Note that the shape of the survival curve implies that $\Delta(s_0) > \Delta(s_1)$ for plausible s. If β_t is constant, R defined as above, and $s_1 - s_0 = \epsilon s_0 R$, I can construct the following upper bound:

$$\tilde{R} < \frac{1}{2}\epsilon s_0 R \left(\frac{\Delta_1 \beta - 1}{\Delta_0} \right) (1 + \beta \epsilon s_0 R) \tag{10}$$

(See Appendix B for a detailed derivation.) Note that this is not a welfare gain, nor is it a proxy for one. Instead, the increase in lifetime income when moving from s_0 to s_1 is partially offset by the increase in direct cost along the way (and exactly offset at s_1). The scatter of this rectangle versus the triangle is seen in Panel B of Figure 8. The vertical scale is the same as in Panel A, which highlights $\tilde{R}$ being much smaller than R. The highest country has a number close to 20%, but almost all of the other countries are below 7.5%.

A final, somewhat rectangular concept is the increasing time in school itself. This number is too simple to be worth graphing. For a fixed elasticity, the proportional increase in s is just ϵR. For some of the least healthy countries, the thought experiment of going to the Japanese life table implies increases of dozens of percent in Δ. This would translate into a similar percentage increase in years of schooling, for unit elasticity. For not a few of the countries this would translate into several years of schooling.

10 Conclusion

Death is like a tax on human-capital investments. Commonly in public economics, (Harberger) triangles measure efficiency losses from taxes. In like manner, the triangle here measures the benefit of reoptimization through this horizon mechanism. An upper bound (that relies on an implausibly large responses of schooling to longevity) is below 10% for the typical low-income country. And that upper bound is pretty slack. Instead, if I use well-estimated elasticities, I find the triangle is less than a few percent. While it is always good to receive the equivalent of a few percent more income, this is but a drop in the bucket relative to the large gap in standards of living (as proxied by incomes per capita, e.g.) that separate rich and poor countries today. In contrast, the evidence suggests that increasing working-age longevity would promote large changes in years of schooling. Why the discrepancy between these two distinct aspects of the horizon channel? In fact, this mismatch is commonplace in a triangle calculation, which, were it a major motion picture, might be entitled "The Revenge of the Envelope Theorem." In optimization problems, a change in parameters (e.g. Δ) can induce a a large change in decision variables, but the resulting change in the decision only causes a second-order change in the objective function.

What limits the horizon-channel is that it works through a triangle, and triangles are shaped that way because of diminishing marginal returns to investment. Nevertheless, the longevity improvements themselves bring a 'rectangle' gain related to the proportional change in Δ. This is potentially large, but it is not the horizon channel. Also, this finding should not be construed to minimize the value of health through other mechanisms. Other studies that find childhood morbidity improves human capital and income (Behrman and Rosenzweig, 2004; Almond, 2006; Bleakley, 2007 & 2010b; Maccini and Yang, 2009; *inter alia*). Why is it different from the horizon mechanism? At some level, it is not. Morbidity also brings a rectangle and a triangle. That is, even if you held fixed the investments (time spent in school), nevertheless the reduced childhood morbidity could make you more productive through various mechanisms. Early-life changes in child health and nutrition should achieve the greatest effect if working through rectangle-like mechanisms (Bleakley, 2010ab).

While this article is focused on welfare and the horizon mechanism, the methodology can be extended to other education questions. Various shocks and interventions increase the marginal benefits of and/or the time in school. Such increases are inputs to computing rectangles and triangles. (See Bleakley, 2017, for an additional example.) Thus, the results above serve as a demonstration of applied welfare analysis in education questions. This would complement analysis of *ad hoc* extrapolations using Mincer's β or of cost effectiveness (e.g., Dhaliwal *et al.*, 2011), the latter of which normalizes increased time in school per dollar spent by an intervention. Canning (2013) provides a rigorous welfare-theoretic foundation for the cost-effectiveness analysis of health interventions by treating the time alive as the numeraire consumption good. However, it seems less reasonable to apply this logic to time in school, not the least because it is an input to human-capital production rather than strictly a consumption good. Therefore, while these latter methodologies have the advantage of being more straightforward to use, they have the disadvantage of being less well connected to welfare measurement than the rectangle-and-triangles framework.

References

Acemoglu, Daron and Joshua Angrist (2000). "How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws." in B. S. Bernanke and K. Rogoff, eds., *NBER Macroeconomics Annual*, NBER MIT Press Cambridge, Massachusetts November pp. 9–58.

Almond, Douglas (2006). "Is the 1918 Influenza Pandemic Over? Long-term Effects of In Utero Influenza Exposure in the Post-1940 U.S. Population." *Journal of Political Economy.* August **114** (4), 672–712.

Ashraf, Quamrul H., Ashley Lester, and David N. Weil (2009). "When Does Improving Health Raise GDP?" *NBER Macroeconomics Annual.*

Baranov, Victoria and Hans-Peter Kohler (2018). "The Impact of AIDS Treatment on Savings and Human Capital Investment in Malawi." *American Economic Journal: Applied Economics.* **10** (1), 266–306.

Barro, Robert J. and Jong Wha Lee (2013). "A new data set of educational attainment in the world, 1950–2010." *Journal of Development Economics.* September **104**, 184–198.

Becker, Gary S., Tomas J. Philipson, and Rodrigo R. Soares (2005). "The Quantity and Quality of Life and the Evolution of World Inequality." *American Economic Review.* March **95** (1), 277–291.

Behrman, Jere R. (2009). "Early Life Nutrition and Subsequent Education, Health, Wage, and Intergenerational Effects." in M. Spence and M. Lewis, eds., *Health and Growth*, Washington, DC: Commission on Growth and Development and the World Bank, chapter 6, pp. 167–183.

——— and Mark R. Rosenzweig (2004). "Returns to Birthweight." *Review of Economics and Statistics.* **86** (2), 586–601.

Ben-Porath, Yoram (1967). "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy.* **75** (4), 352–365.

Bleakley, Hoyt (2007). "Disease and Development: Evidence from Hookworm Eradication in the American South." *The Quarterly Journal of Economics.* February **122** (1), 73–117.

——— (2010a). "Health, Human Capital, and Development." *Annual Review of Economics.* **2** (1), 283–310.

——— (2010b). "Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure." *American Economic Journal: Applied Economics.* April **2** (2), 1–45.

——— (2017). "A Nudge to School: Triangulating the Gains." Technical Report, University of Michigan, Ann Arbor, MI.

——— and Sok Chul Hong (2013). "When the Race between Education and Technology Goes Backwards: The Postbellum Decline of White School Attendance in the Southern US." Technical Report, Sogang University.

———, Dora Costa, and Adriana Lleras-Muney (2014). "Health, Education, and Income in the United States, 1820–2000." in L. P. Boustan, C. Frydman, and R. A. Margo, eds., *Human Capital in History: The American Record*, National Bureau of Economic Research, chapter 4.

Bloom, David E., David Canning, and Michael Moore (2014). "Optimal Retirement with Increasing Longevity." *Scandinavian Journal of Economics.* **116** (3), 838–858.

Canning, David (2013). "Axiomatic Foundations for Cost-Effectiveness Analysis." *Health Economics.* **22** (12), 1405–1416.

Cervellati, M. and Uwe Sunde (2005). "Human capital formation, life expectancy, and the process of development." *American Economic Review.* **95**, 1953–1672.

Cohen, Daniel and Laura Leker (2014). "Health and Education: Another Look with the Proper Data." Technical Report, Paris School of Economics February.

Cutler, David M. and Adriana Lleras-Muney (2006). "Education and Health: Evaluating Theories and Evidence." Working Paper 12352, National Bureau of Economic Research July.

de la Croix, David and Omar Licandro (1999). "Life expectancy and endogenous growth." *Economics Letters.* November **65** (2), 255–263.

Dhaliwal, Iqbal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch (2011). "Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education." Technical Report, Abdul Latif Jameel Poverty Action Lab (J-PAL), MIT August.

Duflo, Esther (2004). "The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia." *Journal of Development Economics.* June **74** (1), 163–197.

Goldin, Claudia and Lawrence F. Katz (2010). *The Race between Education and Technology*, Belknap Press.

Gould, Benjamin Apthorp (1869). *Investigations in the Military and Anthropological Statistics of American Soldiers*, New York: Hurd and Houghton for the US Sanitary Commission.

Hacker, J. David (2010). "Decennial Life Tables for the White Population of the United States, 1790–1900." *Historical Methods.* **43** (2), 45–79.

Hansen, Casper Worm (2013). "Life expectancy and human capital: Evidence from the international epidemiological transition." *Journal of Health Economics.* December **32** (6), 1142–1152.

——— and Lars Lonstrup (2012). "Can higher life expectancy induce more schooling and earlier retirement?" *Journal of Population Economics.*

Harberger, Arnold C. (1971). "Three basic postulates for applied welfare economics." *Journal of Economic Literature.* September **9** (3), 785–797.

Hausman, Jerry A. (1981). "Exact consumer's surplus and deadweight loss." *American Economic Review.* September **71** (4), 662–676.

Hazan, Moshe (2009). "Longevity and Lifetime Labor Supply: Evidence and Implications." *Econometrica.* **77** (1829–1863), November.

——— (2012). "Life Expectancy and Schooling: New Insights from Cross-Country Data." *Journal of Population Economics.* October **25**, 1237–1248.

Heckman, James J., Lance J. Lochner, and Petra E. Todd (2006). "Earnings Functions, Rates of Return and Treatment Effects: the Mincer Equation and Beyond." in E. A. Hanushek and F. Welch, eds., *Handbook of the Economics of Education*, Vol. 1, Elsevier B.V., chapter 7, pp. 307–458.

Hines, James R. (1999). "Three Sides of Harberger Triangles." *Journal of Economic Perspectives.* June **13** (2), 167–188.

Jayachandran, Seema and Adriana Lleras Muney (2009). "Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines." *Quarterly Journal of Economics.* February **124** (1), 349–397.

Kalemli-Ozcan, Sebnem, Harl E. Ryder, and David N. Weil (2000). "Mortality decline, human capital investment, and economic growth." *J. Dev. Econ.* **62** (1), 1–23.

Lemieux, Thomas (2006). "The Mincer Equation Thirty Years after *Schooling, Experience, and Earnings*." in S. Grossbard-Shechtman, ed., *Jacob Mincer, A Pioneer of Modern Labor Economics*, Springer Verlag.

Lleras-Muney, Adriana (2005). "The Relationship between Education and Adult Mortality in the United States." *Review of Economic Studies.* January **72** (1), 189–221.

Lorentzen, Peter, John McMillan, and Romain Wacziarg (2008). "Death and development." *Journal of Economic Growth.* June **13** (2), 81–124.

Maccini, Sharon and Dean Yang (2009). "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall." *American Economic Review.* June **99** (3), 1006–26.

Max Planck Institute for Demographic Research (2016). "The Human Life-Table Database." Online at http://www.lifetable.de. Accessed March 30, 2016.

Mincer, Jacob (1958). "Investment in Human Capital and Personal Income Distribution." *Journal of Political Economy.* **66** (4), 281–302.

Montenegro, Claudio E. and Harry Anthony Patrinos (2014). "Comparable estimates of returns to schooling around the world." Policy Research Working Paper Series 7020, The World Bank September.

Moretti, Enrico (2004). "Human Capital Externalities in Cities." in J. V. Henderson and c. T. Jacques-Fran eds., *Handbook of Regional and Urban Economics*, Vol. 4, Amsterdam: North-Holland, chapter 51, pp. 2243–2291.

Murphy, Kevin M. and Finis Welch (1990). "Empirical Age-Earnings Profiles." *Journal of Labor Economics.* April **8** (2), 202–229.

Preston, Samuel H. (1980). "Causes and Consequences of Mortality Declines in Less Developed Countries during the Twentieth Century." in R. A. Easterlin, ed., *Population and Economic Change in Developing Countries*, University of Chicago Press, pp. 289–360.

Psacharopoulos, George and Harry Anthony Patrinos (2002). *Returns to investment in education : a further update* Policy research working papers, Washington, D.C.: World Bank, Education Sector Unit, Latin America and the Caribbean Region. "September 2002.".

Sanchez Gonzalez, Yuri S. G. (2011). "The Longevity Gains of Education." Doctoral dissertation, University of Chicago.

Schultz, Theodore W. (1981). *Investing in people: the economics of population quality* Royer lectures (1980), Berkeley, Calif.: University of California Press.

Soares, Rodrigo R. (2005). "Mortality Reductions, Educational Attainment, and Fertility Choice." *American Economic Review.* June **95** (3), 580–601.

Stoler, Avraham and David Meltzer (2013). "Mortality and Morbidity Risks and Economic Behavior." *Health Economics.* **22** (2), 132–143.

Strauss, John and Duncan Thomas (1998). "Health, Nutrition, and Economic Development." *Journal of Economic Literature.* **36** (2), 766–817.

Weil, David N. and Joshua Wilde (2009). "How Relevant Is Malthus for Economic Development Today?" *American Economic Review.* May **99** (2), 255–60.

World Health Organization (2016a). *Global Health Estimates 2015: Disease burden by Cause, Age, Sex, by Country and by Region, 2000-2015*, Geneva: WHO. Accessed August 18, 2017.

——— (2016b). "Global Health Observatory." Online at http://apps.who.int/gho/data/. Accessed March 9, 2016.

Appendices for Online Publication

A Derivation of the Triangle with a Generic $\frac{d\Delta}{ds}$

In this appendix, I re-derive the triangle T for alternative assumptions for the dependency of Δ on s. I provide bounds for the extent to which extreme relaxations of this assumption can affect the triangle. In Section 3, I derive an expression for T if $\frac{d\Delta}{ds} = -1$. In other words, this assumption is that the only effect of time in school on time available to work is the inability to work while a student. Yet a large literature argues that more education might cause greater longevity; see Cutler and Lleras-Muney (2006), Lleras-Muney (2005), Sanchez-Gonzalez (2011), for instance, and references therein.[19] This suggests that $\frac{d\Delta}{ds}$ might be greater than -1, if indeed the foregone work-time while in school is partially offset with increased survival during working ages. Indeed, those authors report approximately 0.2 extra years of adult life expectancy, which suggests that $\frac{d\Delta}{ds} > -0.8$ for $r - g > 0$.

Consider a generic $\frac{d\Delta}{ds}$, not necessarily equal to -1. This relaxation affects the problem in two distinct ways: (1) the level of marginal benefits and (2) the response of schooling to what is now a two-part change in MB. Lifetime income, net of costs, Y is as defined above:

$$Y(s, \Delta) = f(s)\Delta(s) + \int_0^s c(t)e^{-(r-g)(t-s_0)}dt.$$

The Y_s for the original and generic $\frac{d\Delta}{ds}$ are as follows:

	$\frac{d\Delta}{ds} = -1$	Generic $\frac{d\Delta}{ds}$
Y_s :	$f'\Delta - (f + c)$	$f'\Delta + f\frac{d\Delta}{ds} - c$

(The dependencies on s, e.g. $f(s)$, are suppressed for legibility of the equations.) It is

[19] Sanchez-Gonzalez uses the *National Longitudinal Mortality Study* to estimate education/mortality gradients for US populations aged 25 and up. He also combines information on differences in survival by education with estimates of the value of statistical life. He takes the extra longevity experienced by the more educated and discounts it back to the school-leaving age. He finds the value of such gains to be important, but nevertheless substantially smaller in magnitude than the Mincer return. Based on his calculations, he argued that the mortality-adjusted Mincer-like returns should be perhaps one percentage point higher (*e.g.* 12-13% instead of 10%). Note that he estimates the mortality/education gradient using both cross-sectional comparisons and standard instrumental-variables strategies.

instructive to re-write Y_s for the generic case as follows:

$$\left(f'\Delta + (1 + \tfrac{d\Delta}{ds})f\right) - (f + c)$$

The second term in parentheses is the same as marginal cost in the baseline case. There remains an opportunity cost of time associated with going to school. Whatever additional working time is created by more years of schooling now appears in the first term, as an addition to marginal benefits. These MB are in the first parenthetical term. Within the MB, there is a level effect, $f'(s)\Delta(s)$, as above and newly a slope effect, $(1 + \frac{d\Delta}{ds})f$.

The height of the triangle continues to measure the gap between marginal benefits in the '1' and '0' scenarios. However, it now contains two terms, because there are two things about marginal benefits that might change with more schooling: wages and survival. Let $\tilde{h}_0$ be the height, in dollars, for a generic $\frac{d\Delta}{ds}$.

$$\tilde{h}_0 = f'(s_0)\left(\Delta_1(s_0) - \Delta_0(s_0)\right) + f(s_0)\left(\frac{d\Delta_1(s_0)}{ds} - \frac{d\Delta_0(s_0)}{ds}\right)$$

or $d\tilde{h}_0 = f'(s_0)d\Delta(s_0) + f(s_0)d\Delta'(s_0)$ for small changes.[20] If $\frac{d\Delta}{ds} = -1$, then the second term disappears and the height is as above. Otherwise, the second term accounts for the extra marginal survival gain from schooling. Notice that it is the difference in $d\Delta/ds$ that matters; any pre-existing survival gains from schooling were already built in to the decision problem. Thus the modified triangle, $\tilde{T}$, is as follows:

$$\begin{aligned}\tilde{T} &= \frac{1}{2}(s_1 - s_0)\tilde{h}_0/(\Delta_0(s_0)f(s_0)) \\ &= \frac{1}{2}(s_1 - s_0)\left[\beta_0\left(\frac{\Delta_1 - \Delta_0}{\Delta_0}\right) + \frac{1}{\Delta_0}\left(\frac{d\Delta_1}{ds} - \frac{d\Delta_1}{ds}\right)\right]\end{aligned}$$

in which all of the Δ concepts are evaluated at s_0.

For a generic $\frac{d\Delta}{ds}$, the response of schooling to survival implied by the model becomes more complicated, because there are now both level and slope effects. Consider full differentials

[20] $d\Delta'(s)$ is a small change in the slope of Δ w.r.t. s. I use Δ' and $\frac{d\Delta}{ds}$ interchangeably in this appendix.

of the new FOC, $Y_S = 0$, expressed compactly as follows:

$$ds \, Y_{ss} \; + \; d\Delta \; f' \; + \; d\Delta' \; f \; = \; 0,$$

which I re-write as

$$ds = -\frac{f'}{Y_{ss}} \left(d\Delta + \frac{d\Delta'}{\beta} \right). \tag{11}$$

Let $\Omega \equiv \left(d\Delta + \frac{d\Delta'}{\beta} \right)$, which has the same units as Δ, but includes the slope effect. Note that

$$\frac{ds}{d\Omega} = -\frac{f'}{Y_{ss}}$$

has the same right-hand side as equation 5, which expressed $\frac{ds}{d\Delta}$. Thus the sensitivity of time in school to marginal benefits is in some sense the same, as long as we can account for the slope effect ($d\Delta'$) as well.

Theoretically, there exists the possibility that the effect of mortality reductions on schooling might be negative, although this is unlikely. For this to happen, we would need the survival gains to be very strong for the low educated not so much for the more educated. This could make the slope effect negative, and perhaps with a magnitude large enough to overwhelm the level effect. (There would still be a triangle of gains from reoptimization in this case, but it would involve less rather than more schooling.) The extant empirical literature, however, finds positive effects of mortality reductions on education. This suggests that, in the episodes studied, the slope effect did not offset the level effect. Let us take an extreme case to see why. Suppose that, in unhealthy countries the survival benefits of education were so large that an additional year of education generated an additional year of (discounted) survival, for a $\Delta'_0 \approx 0$. This would be markedly different from any estimate with which I am familiar. Compare this with estimates for the US, which imply $\Delta'_1 \approx -.8$ for $r - g = 0$. Consider the shock to marginal benefits, expressed as a proportion of Δ_0:

$$\frac{d\Omega}{\Delta_0} = \left(\frac{d\Delta}{\Delta_0} + \frac{d\Delta'}{\Delta_0 \beta} \right).$$

For the least healthy countries, going to the Japanese survival curve implies a proportional increase in life expectancy of about 0.4, which I take as an approximation for $\frac{d\Delta}{\Delta_0}$, the first term. How large is the second term, the slope effect? In this extreme thought experiment, $d\Delta'$, the change in the longevity gains per year of schooling, is -0.8. The denominator has $\beta \approx 0.1$ and $\Delta_0 \approx 45$ in undiscounted units. These combine for a slope effect of around -0.18. Even though $\Delta'_0 \approx 0$ is an implausibly large starting point for the longevity gains of education, this nevertheless implies a slope effect that remains at less than half the magnitude of the level effect.

On the other side of things, the slope effect could get stronger as we transition from an unhealthy to a healthy life table. This would increase the size of the triangle, both because marginal benefits are higher and because the induced response of schooling is larger. But by how much? Again, consider what is likely an extreme case: zero effect of education on longevity ($\Delta'_0 = -1$) in the unhealthy scenario versus with the US gradient discussed above, ($\Delta'_1 = -0.8$). This gives a slope effect, expressed in the same units as above, of $0.2/45/10 = .0044$, which is an order of magnitude smaller than the level effect. The increase of marginal benefits, $d\Omega$, would be about a factor 1.11 larger than a baseline that only includes the level effect. The response of schooling would be larger by a similar fraction, as per equation 11. These combine to give a triangle about 1.23 times larger than the baseline (with no slope effect). Such an increase does not materially alter the conclusions offered in Section 6.

B Additional details in the derivation of a bound on $\tilde{R}$

In Section 9, equation 10 presents an upper bound on the increase in discounted lifetime income, $\tilde{R}$. This appendix presents a detailed derivation of this bound. As noted above, the starting point for $\tilde{R}$ is the area of a trapezoid:

$$\tilde{R} \equiv (s_1 - s_0)\left(\frac{\Delta_1(s_0)f'(s_0) - f(s_0) + \Delta_1(s_1)f'(s_1) - f(s_1)}{2}\right)\left(\frac{1}{\Delta(s_0)f(s_0)}\right),$$

I shift the denominators one term to the left and use that $f'(s_t) = \beta_t f(s_t)$,

$$\tilde{R} = \frac{(s_1 - s_0)}{2}\left[\frac{\Delta_1(s_0)\beta_0 - 1}{\Delta_0(s_0)} + \frac{\Delta_1(s_1)\beta_1 - 1}{\Delta_0(s_0)}\left(\frac{f(s_1)}{f(s_0)}\right)\right]$$

If education does not bring very large increases in discounted longevity ($\frac{d\Delta}{ds} > 0$), the shape of the survival curve implies that $\Delta(s_0) > \Delta(s_1)$. (For a very high infant mortality, survival from age zero to one might bring a greater remaining life expectancy, but mortality rates at typical school-leaving ages are not so extreme.) Thus, we can replace $\Delta_1(s_1)$ with $\Delta_1(s_0)$ to find

$$\tilde{R} < \frac{(s_1 - s_0)}{2}\left[\frac{\Delta_1(s_0)\beta_0 - 1}{\Delta_0(s_0)} + \frac{\Delta_1(s_0)\beta_1 - 1}{\Delta_0(s_0)}\left(\frac{f(s_1)}{f(s_0)}\right)\right]$$

For β_t constant, we can re-write this upper bound as

$$\tilde{R} < \frac{(s_1 - s_0)}{2}\left(\frac{\Delta_1(s_0)\beta_0 - 1}{\Delta_0(s_0)}\right)\left(1 + \frac{f(s_1)}{f(s_0)}\right)$$

If the β is indeed constant, then $f''(s)$ must be positive.[21] This in turn implies that $\frac{f(s_1)}{f(s_0)} < \beta(s_1 - s_0)$. Thus, $\tilde{R}$ is bounded from above by

$$\tilde{R} < \frac{(s_1 - s_0)}{2}\left(\frac{\Delta_1(s_0)\beta_0 - 1}{\Delta_0(s_0)}\right)(1 + \beta(s_1 - s_0))$$

With the definition of R and ϵ, the inequality can be rewritten as in equation 10.

[21]Consider logarithmic differentiation of $\beta(s)$: $\frac{dln\beta}{ds} = \frac{f''(s)}{f'(s)} - \frac{f'(s)}{f(s)}$. By assumption, f and f' are positive. For β to be constant, $f''(s) = \beta^2 f(s) > 0$.

C Spreads in the required return to human capital

How does mortality increase the required rate of return on human capital, conditional on survival? In the absence of mortality, $\Delta(s) = (r - g)^{-1}$. Thus, $\left(\frac{1}{\Delta} - (r - g)\right)$ is a spread over the interest rate that has to be built into the return to human capital.

Appendix Figure C–1 shows these spreads in the four example countries' life tables for age ≥ 15 and for different assumptions about $r - g$. The curves have an opposite order as in Figure 4, with Japan having the lowest spread and Mozambique having the highest. The curves all slope downwards, as they should. As the financial discounting gets extreme, later-life differences in survival might as well not exist because they occur so far in the future. Yet the spreads do not go to zero even at a high interest rate.

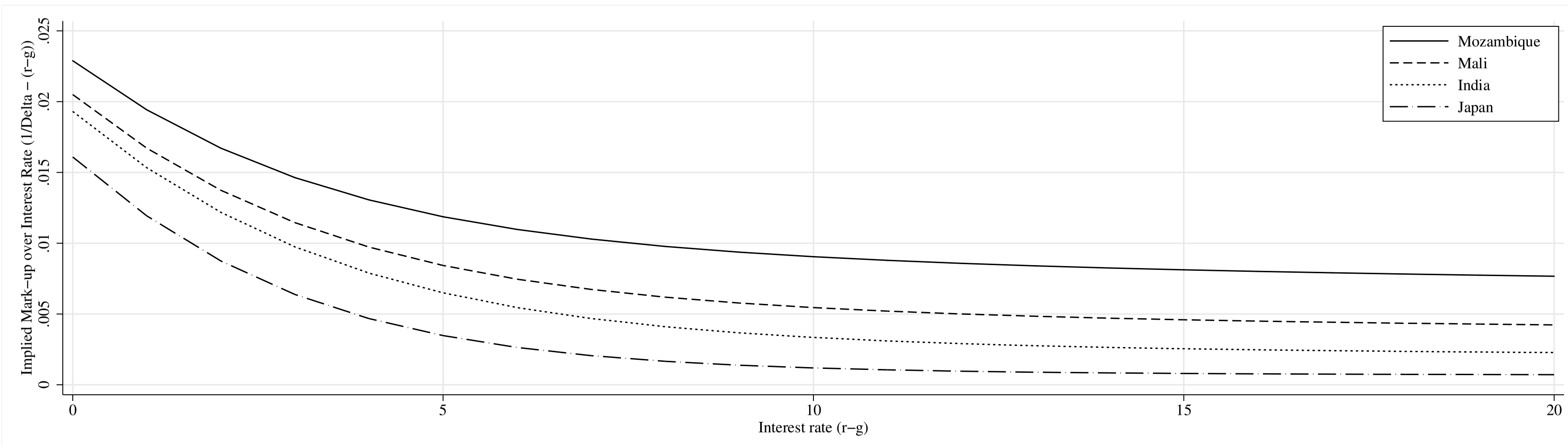
For low interest rates and even for a healthy country like Japan, there is a spread. This spread is at its largest for $r - g = 0$, yet it is only 1.6 percentage points. For the developing countries, this markup is higher, but the gap with Japan is not so great; Mozambique's curve is consistently less than 100 basis points higher than Japan. The e_{15} numbers given above serve as a quick check on the difference between the curves at the intercept, where $(r - g = 0) \Longrightarrow (\Delta(15) = e_{15})$. The markup for Mozambique over Japan, if $r - g = 0$, is

$$\left(\frac{1}{e_{15,MOZ}} - \frac{1}{e_{15,JPN}}\right) \approx \left(\frac{1}{43} - \frac{1}{62}\right) \approx 0.7\%.$$

This result gives us another insight as to why the triangles are small. On the one hand, the return to human capital in Mozambique would be higher if it were valued using the Japanese survival curve. On the other hand, the extra rate of return in that "if" is simply not very big. Furthermore, the rate-of-return differential diminishes as one responds to it by increasing time in school. (In other words, the reallocation generates a triangle, not a rectangle.)

Note that this result is similar to Samuel Preston's (1980) well-known calculation. He used life tables for Mexico to compute returns to schooling in the early and mid 20th century. He showed that this made a small difference in the rate of return. Preston then used this result to argue that the response of education to longevity must be small, at least as a practical matter. He did not attempt to compute welfare gains. As seen in the main text, estimating such gains requires credible estimates of the elasticity of school to Δ, which are only recently available.

Appendix Figure C– 1: Extra Required Return for Human Capital Because of Mortality $\left(\frac{1}{\Delta} - (r-g)\right)$, Select Countries



Notes: this figure shows implied markups on the human-capital rate of return for select countries and various interest rates. Data and computation are re-normalized to be conditional on surviving to age 15. The underlying survival curves (shown in Figure 4) are drawn from the World Health Organization's Global Health Observatory (2016). The Δ are computed as described in the text. The graph displays the difference between the inverse of Δ and the net interest rate $(r-g)$, thus measuring a spread between the required return to human capital and the net interest rate.

D Additional details for the comparison across 20th Century

These data are online at lifetable.de. I use circa-1930 survival curves from India, Egypt, and China, the only still-developing countries available from the *Human Life-Table Database* at the time of writing. (An early-20th-century life table for South Africa was also available, but only for whites.) For comparison purposes, I also use circa-1930 life tables from the United States, more or less the frontier country at that time, and from Japan, a country that was relatively poor in the early 20th Century, but clearly on the path of development. The precise years for the early-20th-Century life tables are as follows:

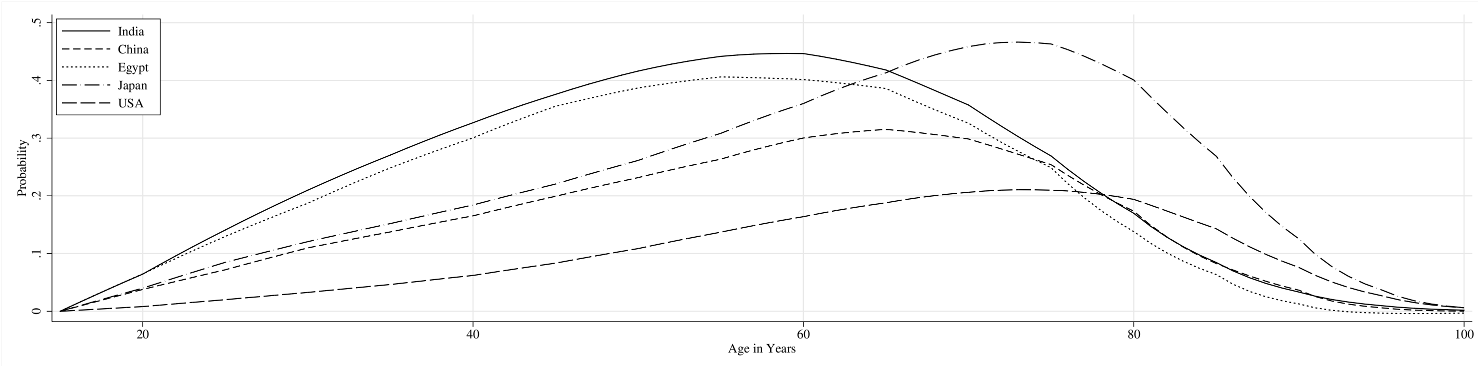
USA	Japan	Egypt	China	India
1930	1921-1925	1927-1937	1929-1931	1921-1931

The data contain survival curves reported at five-year intervals, except for the Egyptian data, which have ten-year intervals. I interpolate the log survival curves, as above, to obtain intermediate years' data.

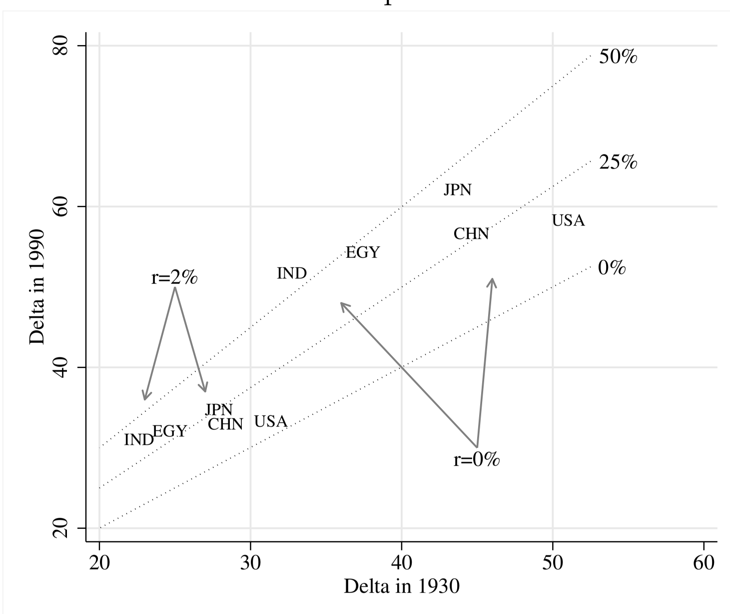
Appendix Figure C–2 shows some key features of these countries' health and school performance in the 20th century. Panel A shows the over-20th-century *differences* in the survival curves for these countries. (As above, these curves are all normalized to age 15.) Japan stands out as having the largest difference between curves at any single age: an over 45% increase in the survival to age 70 when comparing the 1930 and 1990 life tables. India and Egypt also stand out for very large differences in the survival curves at working ages, with differences that peak over 40% just below 60 years old. China and the US have smaller, but still impressive improvements in their survival curves. Panel B is a scatterplot of the Δ for each country/period. I show two sets of Δ: one computed with no discounting, another computed with a 2% discount rate. The dotted lines are level curves denoting proportional increases for reference. The countries besides the USA show large proportional gains in Δ, approximately 25-50% if no discounting is applied, and closer to 20-30% using a 2% discount rate. Panel C is a scatterplot of schooling for the cohorts aged 10-15 at the reference date for each life table. All of the countries show large gains in years of schooling across cohorts.

Appendix Figure C– 2: Changes in Survival and Schooling over the 20th Century, Select Countries

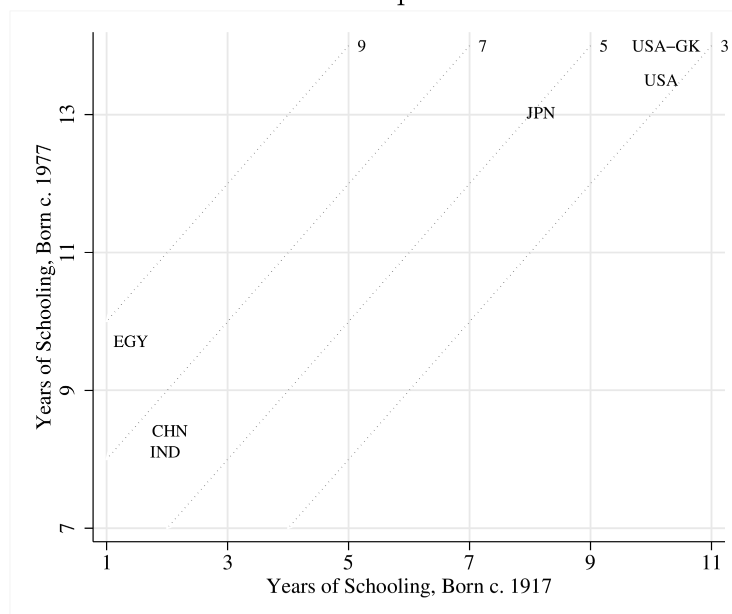
Panel A: Changes in Survival Curves, c. 1930–1990



Panel B: Comparison of Δ



Panel C: Comparison of s



Notes: this figure displays, for select countries, the change in the survival curve (Panel A), a scatterplot of Δ in the early and later life tables, (Panel B), and the scatterplot of years of schooling for early and later reference cohorts (Panel C). Scatter plots use World Bank country codes as labels. (In Panel C, an alternate value for the US is presented ("USA-GK") using estimates from Figure 1.4 of Goldin and Katz, 2010, page 20.) Panel B uses displays early and late deltas for two sets of discounting assumptions (zero and two percent). Dotted lines in the scatters are level curves indicating the degree of difference between the early and late values. The underlying data sources and variable definitions are described in the text.

E Adjustment for Disability: Hello, *DALY*

While the present study is focused on valuing the effect of mortality on education, I show here that accounting for a standard measure of disability makes little difference for the results above. The World Health Organization (2016) uses the concept of a disability-adjusted life year (DALY), a measure which combines years lost to mortality and to disability. For mortality, the lost life expectancy is computed relative to a frontier life table. For disability, the WHO reports the 'years of healthy life lost due to disability' or YLDs for non-fatal health conditions. Conditions are scaled according to a disability weight. Such weights are meant to translate the intensity of each condition into utility equivalents of death. This is an inexact match for the present study: a condition could be very unpleasant yet not impede work, or *vice versa*. A condition might also impede the formation or use of human capital in a way that is not measured well with standard instruments.[22] Nevertheless, these are, to my knowledge, the only data comprehensive enough to be suitable for adjusting the county- and age-specific survival curves. This version of the estimated disability burden uses a prevalence-based approach. As such, the reported numbers refer to conditions occurring in the specific age bracket.

I combine the disability-by-age tables for 2000 with the 1990 life table used above to produce a disability-adjusted measure of discounted life years, $\tilde{\Delta}$:

$$\tilde{\Delta}(s) \equiv \int_{s}^{\infty} e^{-(r-g)(t-s)} \frac{\ell(t)}{\ell(s)} \left(\frac{\text{YLD}_t}{N_t} \right) dt$$

I normalize the years lost to disability (YLD) by population (N, also reported in WHO, 2016) to produce a probability of disability (among those alive) for each age. The DALYs and YLDs are reported in fairly broad age categories: during working ages, 15-29, 30-49, 50-59, and 60-69. The final age category (70+) is so coarse that I restrict my analysis with DALYs to working ages. I apply the average value for each bracket to each age within it.

[22] See, for example, Hotez PJ, Alvarado M, Basanez MG, Bolliger I, Bourne R, *et al.* "The global burden of disease study 2010: interpretation and implications for the neglected tropical diseases." *PLoS Negl Trop Dis.* 2014; 8: e2865.

Because the prevalence of disability tends to increase with age, this will overstate the burden of discounted disability.

How does $\tilde{\Delta}$ compare with Δ? The two measures are very highly correlated, as shown in Appendix Figure E–1, Panel A. This panel plots the proportional jump that would be required to reach Japan for both measures of discounted life years. The points, one for each country, are all very close to the diagonal, which indicates that the disability adjustment makes very little difference for the thought experiment in which I take a country to Japan's level of health.

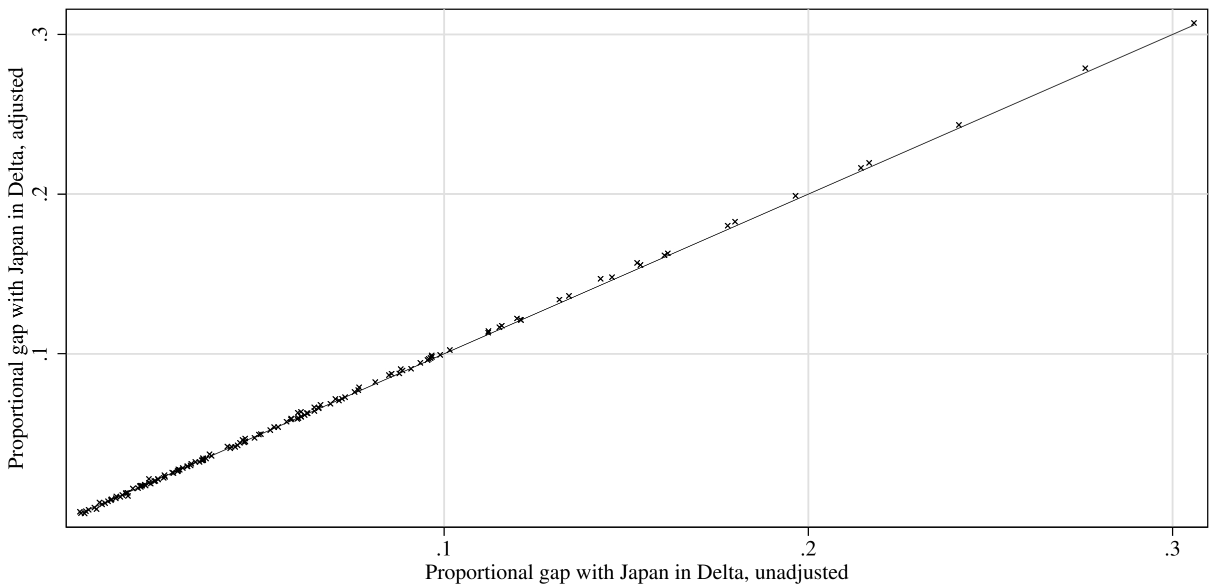
Why do these adjustments have so little effect on Δ? For one, the disability weights are generally much less than 1, thus reducing the importance of disability relative to death. For another, the adjustment is perhaps surprisingly neutral across the global distribution of Δ. Because no country is disease-free, the $\tilde{\Delta} < \Delta$ for all observations. For Japan, this adjustment reduces discounted life years by 0.20. Yet the jump to the Japanese level can get larger or smaller with the adjustment. For the median country, $(\Delta - \tilde{\Delta})$ is just 0.0015 larger than Japan's. For Zambia (ZMB), the country with the lowest Δ, the adjustment makes less of a difference than it did for Japan. Appendix Figure E–1, Panel B. contains a scatterplot of the following object versus the baseline Δ:

$$(\tilde{\Delta}_i - \tilde{\Delta}_{JPN}) - (\Delta_i - \Delta_{JPN})$$

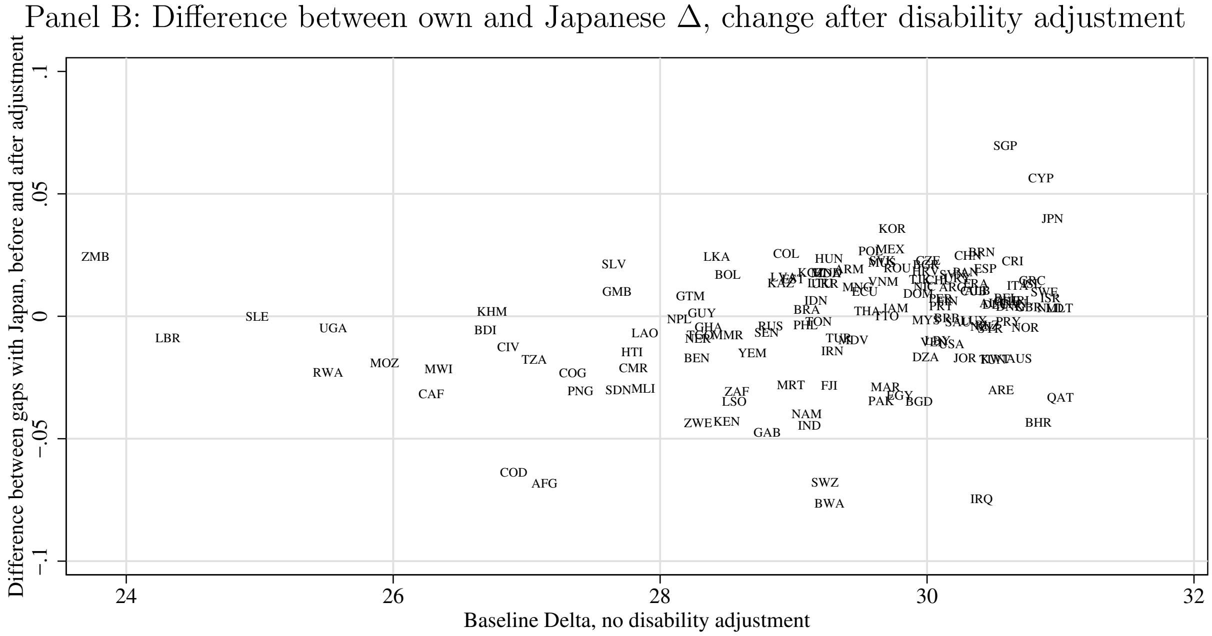
for each country i. The y axis therefore measures a difference in difference, which is what would alter the computation of the triangle. As can be seen in the graph, the disability adjustment delivers a very small change in the gap in discounted life years with Japan. Thus, the triangles presented in the main text also change very little.

Appendix Figure E– 1: Effect on Δ of Adjustments for Disability

Panel A: Percentage difference in Δ with Japan, adjusted for disability versus baseline

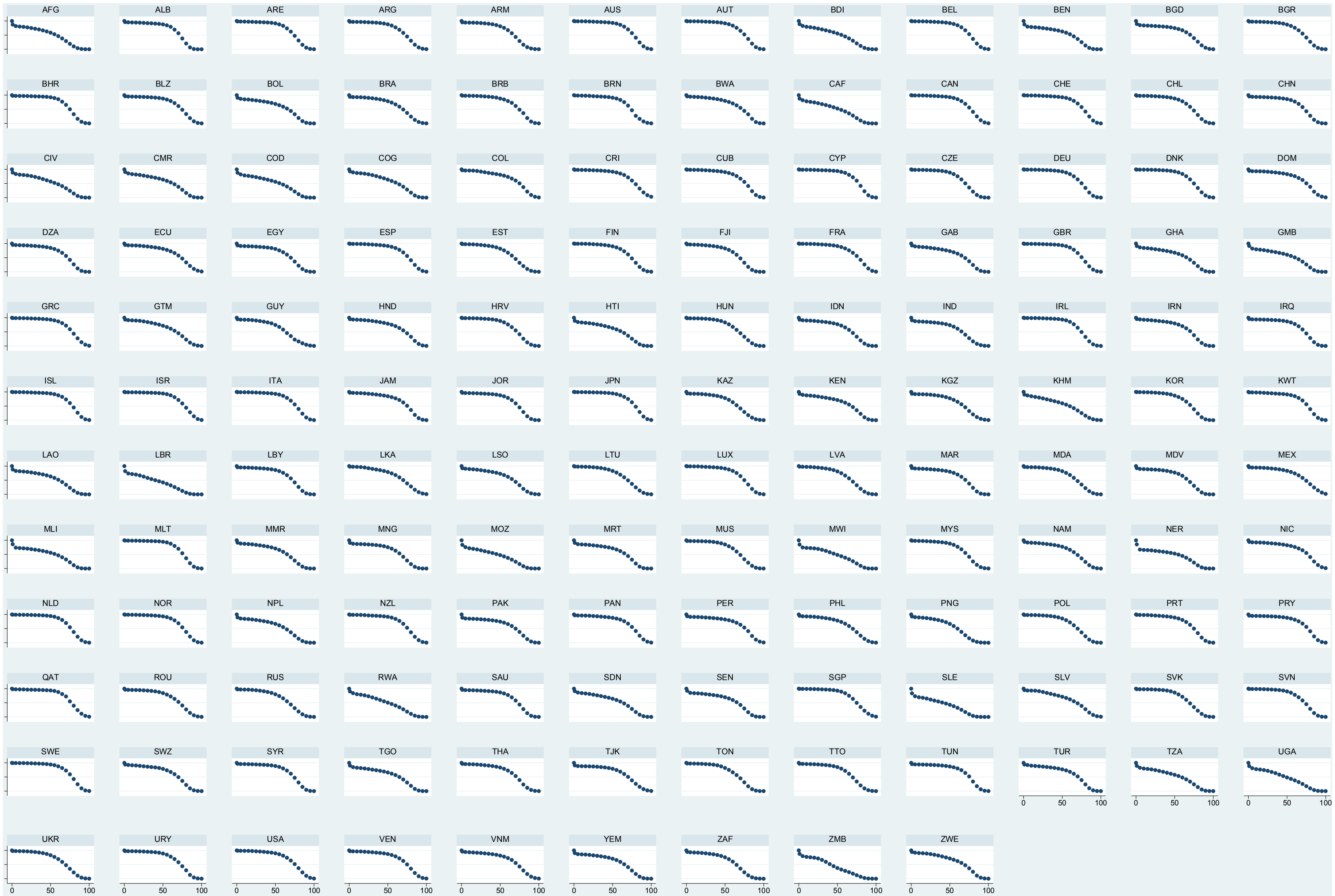


Panel B: Difference between own and Japanese Δ, change after disability adjustment



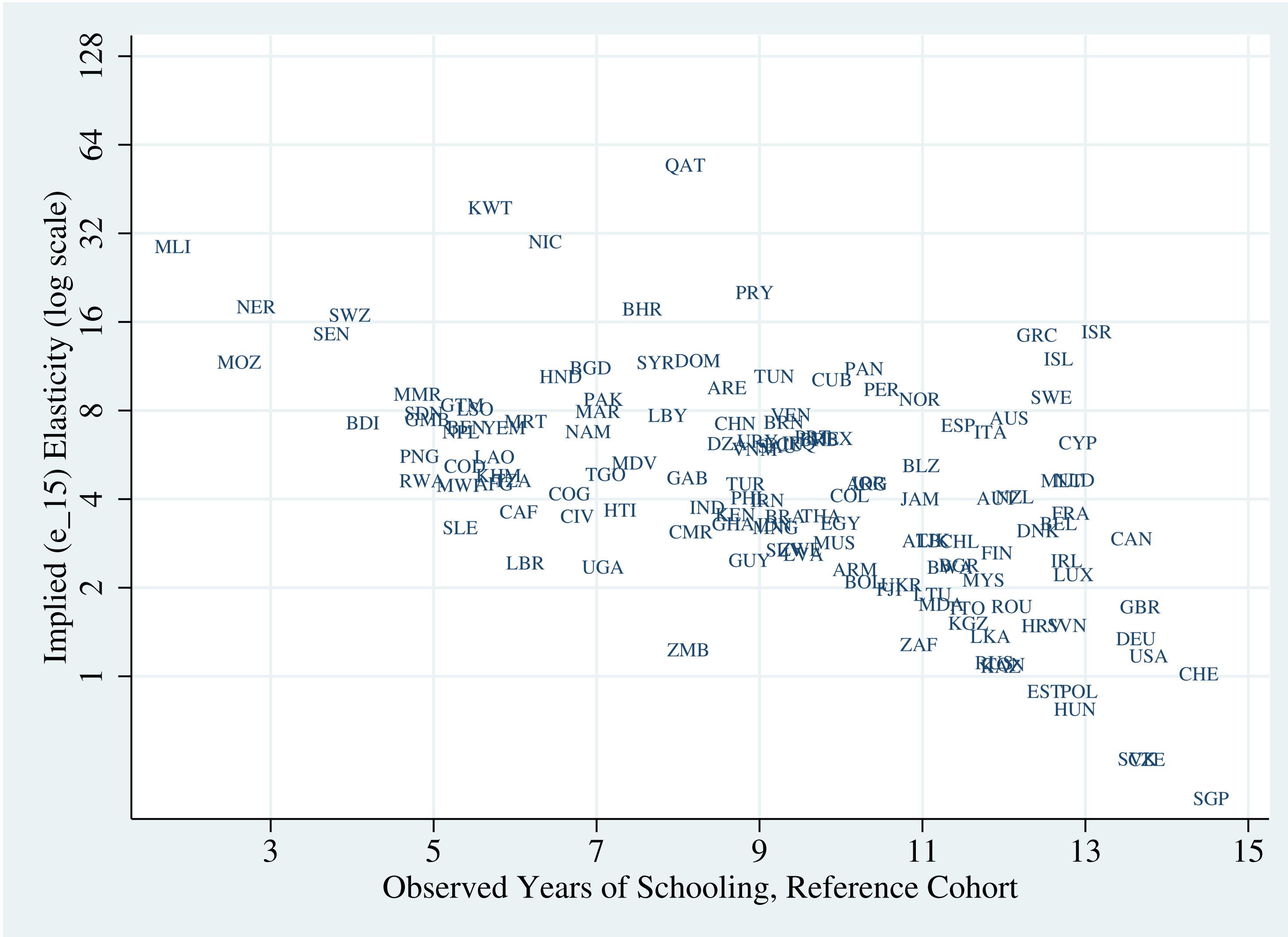
Note: this figure presents comparisons of Δ with $\tilde{\Delta}$, a version of discounted life years that is adjusted for disability using DALY measures (WHO, 2016). See Appendix E for variable sources and definitions. The proportional gaps with Japan in Panel A are $\left(\frac{\tilde{\Delta}_i-\tilde{\Delta}_{JPN}}{\tilde{\Delta}_i}\right)$ and $\left(\frac{\Delta_i-\Delta_{JPN}}{\Delta_i}\right)$, for each country i. The solid line in Panel A is the 45-degree (diagonal) line. Three-letter codes in Panel B are World Bank country codes. The y axis in Panel B measures the change (before and then after disability adjusted) in the difference in discounted life years with Japan. This object is a double-difference equaling $(\tilde{\Delta}_i - \tilde{\Delta}_{JPN}) - (\Delta_i - \Delta_{JPN})$.

Appendix Figure F: Survival Curves, 1990, for Countries in Main Sample



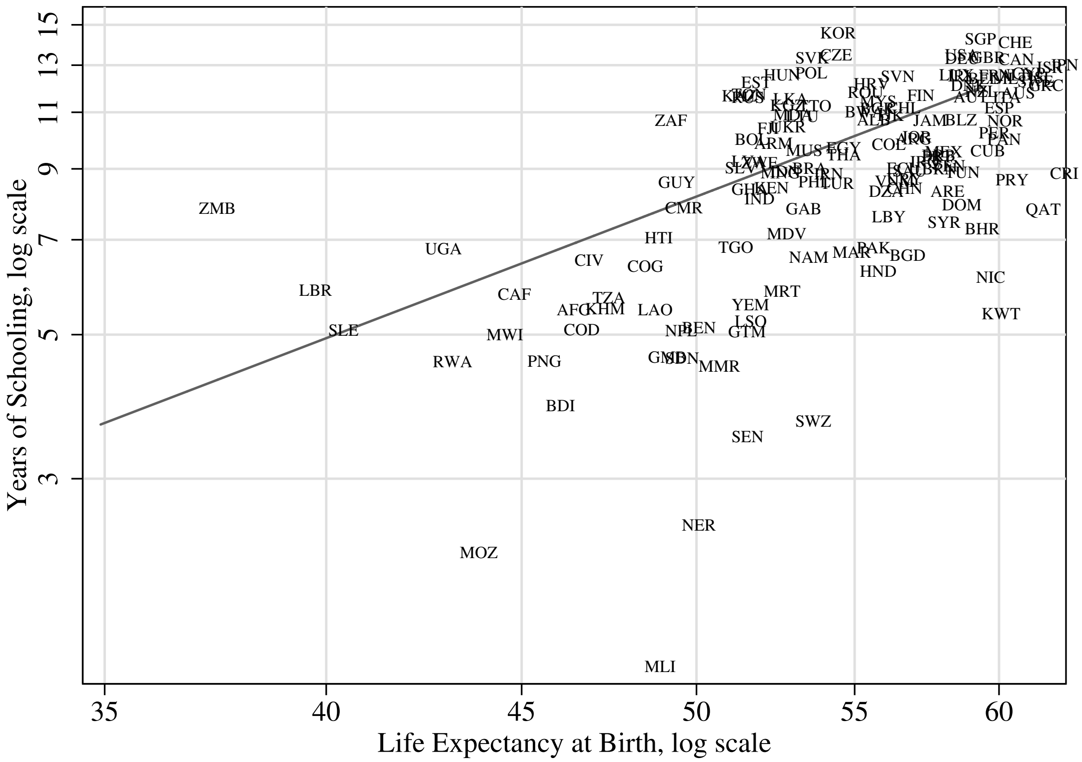
Note: this figure displays survival curves in 1990 for all countries in the sample. The x axis age. The y axis, which ranges from 0 to 1, denotes the fraction alive at that age. The data source is the World Health Organization, Global Health Observatory, 2016. Subplot titles denote World Bank country codes. Values between reporting intervals are log-linearly interpolated as described in the text.

Appendix Figure G: Implied Elasticities from the Upper-Bound Calculation



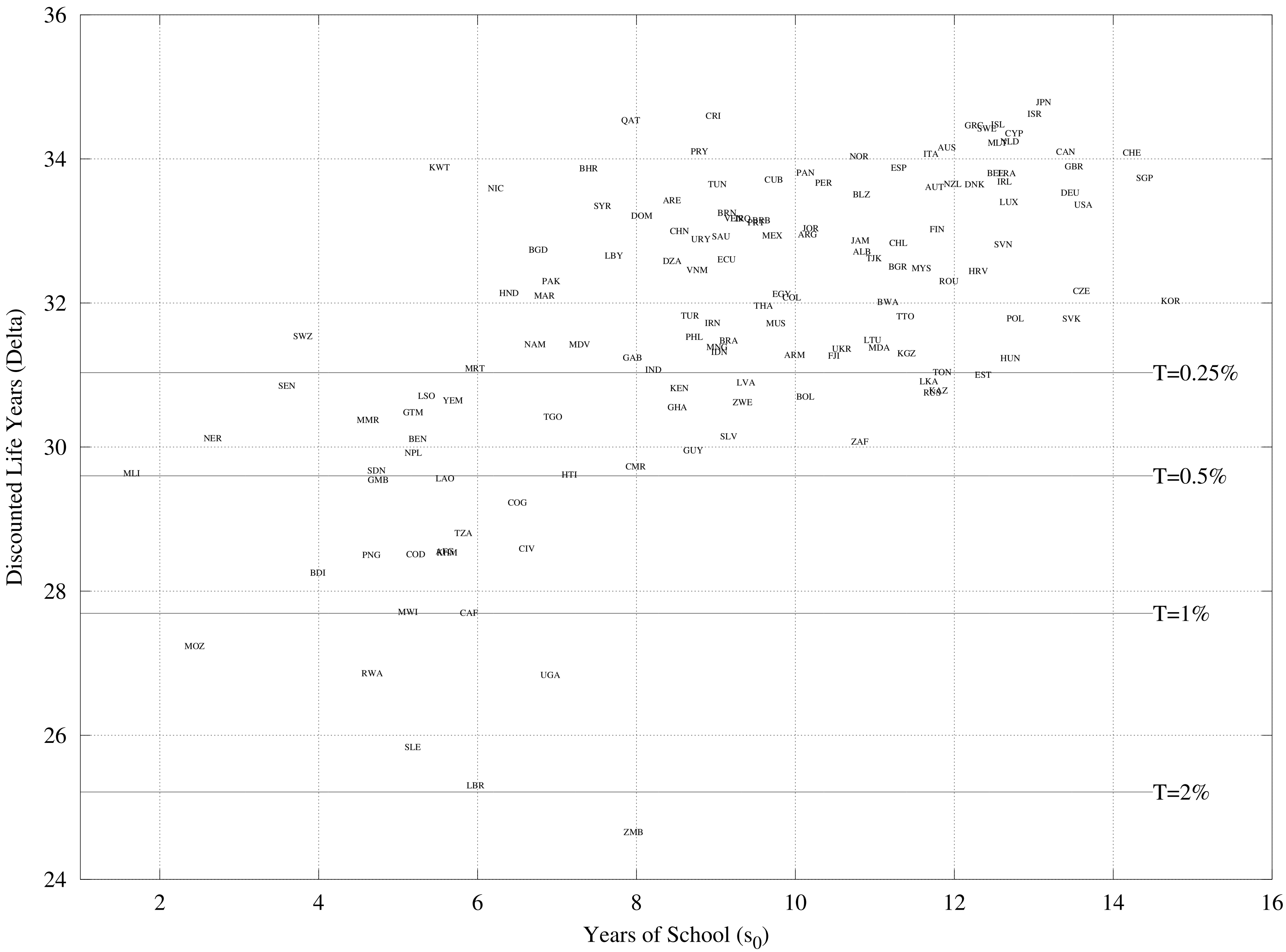
Note: This figure presents the implied elasticities from the upper-bound calculation in Section 5, scattered against years of school (s_0). Three-letter codes are World Bank country codes.

Appendix Figure H: Years of Schooling versus Life Expectancy at Age 15, Log Scales



Notes: this figure contains a log-log scatterplot of years of schooling (s) versus years of life expectancy at age 15 (e_{15}) for males in the sample of countries. Observations are labeled with World Bank country codes. Life expectancy is computed from a period life table measured circa 1990 (World Health Organization, Global Health Observatory, 2016). Years of schooling are from the Barro and Lee (2013) database and refer to the cohort born in 1976-80 and observed in 2010. The best-fit line has a elasticity of 2.25, with a standard error of 0.32. In the subsample at or below the median e_{15}, the estimated elasticity is 2.48 (0.69 standard error). (See Figure 3 for a scatterplot of these data in levels.)

Appendix Figure I: Iso-Triangle Curves for Constant Slope of Schooling to Δ, $\left(\frac{ds^*}{d\Delta} = 0.11\right)$



Notes: this figure displays a scatter of discounted life years (Δ) versus initial years of school (s_0) for the sample of countries. Observations are denoted by their World Bank country codes. Overlain on the scatterplot are iso-triangle curves; these trace out constant values for the triangle (T) of reallocation gains. T is computed using equation 8 and uses a constant slope ($\frac{ds^*}{d\Delta}$) of 0.11. This number is drawn from Jayachandran and Lleras-Muney (2009) and is consistent with the other micro- and cohort-level studies. The underlying data sources and variable definitions are described in the text.