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

Hoyt Bleakley†

February 10, 2017

Abstract

While health affects economic development 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, but this study asks instead what impact this specific channel has on well-being. 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 would gain less than 15% of income through this channel if switched to Japan’s survival curve. Calibrating the model with well-identified micro- and cohort-level studies, I find that the horizon triangle for the typical low-income country is less than a percent of lifetime income. Similarly, increased survival in the 20th-century USA generates a triangle less than 1% of initial income.

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

*Preliminary and incomplete. Comments welcome. 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, Paul Rhode and seminar participants at Princeton (Center for Health and Wellbeing), Michigan (Robert Wood Johnson workshop), Michigan State, the University of Connecticut, the PopPov2016 conference, Carnegie Mellon University, University of Pittsburgh, and Northwestern University for helpful comments.

†Department of Economics and Population Research Center, University of Michigan. Postal address: 611 Tappan St, 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.

Indeed, life expectancy, education, and income per capita are strongly correlated both across countries and within countries along their paths of development. The question arises as to what are the various causal channels that give rise to these correlations. (See Strauss and Thomas, 1998, or Bleakley, 2010, for reviews.) The horizon mechanism is potentially one such channel, but how large is its contribution to differences in well-being 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 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 (2006), and Hazan (2009). 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), and Dorsey, Oster, and Shoulsen (2013) are particularly 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. Based on the prominence of the literature asking whether...
longevity affects education, one might surmise that the horizon mechanism, if true, would be associated with significant gains in well-being. Nevertheless, standard and transparent economic tools, combined with demographic and education data, indicate that the horizon mechanism plays at best a small role in economic development, as I demonstrate below.

The present study adds to this literature by first noting the similarity between death and taxes. This allows us to apply a time-honored and transparent tool from public economics: Harberger 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 I show that 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 in the MB curve. 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 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’ and MC, which intersect at \( s_1 \). The gain from this reoptimization is the triangle (labeled \( T_3 \)). The remainder of this study is an attempt to gauge its size. (See Section 3 for description of the model and derivation 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 causes the country to go to the global frontier in education. This rather implausible
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. $MC$ denotes marginal costs of education. $MB$ measures the present value of marginal benefits from obtaining more schooling in the high-mortality scenario, while $MB'$ 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 $MB'$ intersects with $MC$ at a higher time in school than did $MB$. The area between $MB'$ and $MC$ between $s_0$ and $s_1$ is the triangle ($T_s$) of gains from reallocating from $s_0$ to $s_1$. While the curves shown are globally 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.

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 ($\approx 0.1$) 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 an 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 ($\approx 5$). But recent well-identified studies, notably Jayachandran and Lleras-Muney (2009), Fortson (2011), and Dorsey, Oster, and Shoulson (2013), put this elasticity closer to one. This smaller response of education to longevity coupled with discounting brings the triangle to below two percent. A two-percent gain is, of course, better than nothing, but is dwarfed by the gap in income per capita across
countries, for example. (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 7.2 and 7.3, 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 for which
early-20th-century life tables exist. 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. 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,
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 present costs. This is what makes the triangle a more appropriate object than the
‘rectangular’ number of extra schooling times the Mincer return.

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 micro data 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 Global Health Observatory (2016). The main analysis is for males, but triangles are also small for females (see Section 7).

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.\footnote{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.} Life expectancies are generally on the rise, and therefore the period life tables will tend to understate future marginal benefits of schooling, especially
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.

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 five years up to 100. I convert these to 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 close to 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

\(^2\) The so-interpolated survival curves are plotted for each country in Appendix Figure A. A linear and a bi-cubic interpolation of the log survival curve are highly correlated, with \( \rho > 0.998 \).
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. This will artificially inflate the size of triangles computed below.

3 Model

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 \( g \). Let the flow of direct costs of the \( s \)th year of school be \( c(s) \), and let the wage, if he is working, be \( f(s) \). (For analytical convenience, I abstract from retirement and from post-school learning here, but these aspects are addressed in Section 7.)

---

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 so-called 1990 life table.

4 These are treated as exogenous throughout the analysis. A fixed \( r \) can be justified by assuming 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 7.3.
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.
Viewed from birth, the expected lifetime income, net of direct costs, is as follows:

\[ Y = \int_{s}^{\infty} f(s)e^{-(r-g)t} \ell(t)dt - \int_{0}^{s} c(t)e^{-(r-g)t} \ell(t)dt \]  

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 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 \(\Delta(s) = \int_{s}^{\infty} e^{-(r-g)(t-s)} \frac{\ell(t)}{\ell(s)}dt\) is a factor that inflates

---

\(^5\)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. This point is addressed in Section 7.

\(^6\)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 \(s\).
\( f'(s) \) to account for integration of future discounted flows. For zero mortality, for example, 
\( \Delta(s) \) becomes \( 1/(r-g) \), the familiar factor that converts an infinite, constant flow to present discounted value. For non-zero mortality, \( \frac{e^t}{\ell(s)} < 1 \), which pushes \( \Delta \) 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 most 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 is the first-order condition or FOC.\(^8\) The FOC, cast in period-\( s \) units, can also be written as

\[
f'(s)\Delta(s) = f(s) + c(s).
\]

As survival rates increase, future benefits are less heavily discounted. This causes a shift upwards in the marginal benefits curve, say from MB to MB'. 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 \( s_1 \).

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 \( (b-a) \). 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

\(^7\)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.

\(^8\)Assuming an interior solution is an approximation, but not a bad one. 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.
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 \( \Delta_0 \) and \( \Delta_1 \) be the discount/summation factors that hold at \( s_0 \) in the low and high-survival environments, respectively. This means that, in the Figure 1, \( a = f'(s_0)\Delta_0 \) and \( b = f'(s_0)\Delta_1 \). The triangle, being \( \frac{1}{2} \times \text{height} \times \text{base} \), is therefore

\[
T_s = \frac{1}{2} \left( \Delta_1 - \Delta_0 \right) f'(s_0) (s_1 - s_0).
\]

The dollar-sign subscript on \( T_s \) signifies its units of measurement: the \( \Delta \)'s are unit free, the \( f' \) is money per year, and the \( s \) is years; therefore \( T_s \) measures a monetary gain in period-\( s \) dollars. While describing the triangle in money terms is meaningful, I find it easier to think in proportions of lifetime income. By the above assumptions, lifetime income is \( \Delta_0 f(s_0) \) in the low-survival (high-mortality) scenario. The so-normalized triangle is

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

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 “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 \( \frac{\text{dln} f(s)}{\text{ds}} \).) This is what Mincer’s \( \beta \) is meant to measure in the oft-seen regression of log wages on years of schooling. There are no shortage of cross-sectional estimates for this parameter, both in developed and developing entries. Psacharopolous (1985), Psacharopolous and Patrinos (2002) provide useful surveys of estimated Mincer \( \beta \)'s for numerous countries and time periods. Commonly,
recent estimates for Mincer’s $\beta$ 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 problem 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, made popular 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 only exact if the model itself is exactly right, and otherwise should be viewed as approximations as well. 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 with skill), but rather relies on generic properties of the decision problem at or near the optimum.9

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 $\Delta_0(s_0)$.10 Subscripts denote derivatives in this expression

9A 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 essentially about productive efficiency (maximizing lifetime income, net of costs, w.r.t. $s$). This means that, at the FOC for $s$, the consumer’s consumption decision can be separated from the education decision.

10One might treat $\Delta$ as a parameter (ignoring its local dependence on $s$) in a standard derivation of a Harberger triangle, in which the parameters are exogenous prices and taxes. (See Harberger, 1971, page 788, for example.) Here I endogenize $\Delta$. Either approach yields an identical triangle formula. If $\Delta$ is parametric, $Y_{ss}$ is different ($Y_{ss} = f'' \Delta - (f' + c')$). But it remains the case that $\frac{ds}{d\Delta} = -f'/Y_{ss}$. Thus, a different $Y_{ss}$...
of the second-order Taylor approximation for $Y$:

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

(4)

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

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

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

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

The continued presence of $s_0$ in this expression is not a typo; rather it indicates that we have fixed the viewpoint date at $s_0$ for discounting. When taking a derivative with respect to $s$, the lower limit of the integral moves, but not the reference date:

$$\frac{d\Delta}{ds} = -e^{-(r-g)(s_0-s_0)} \frac{\ell(s_0)}{\ell(s_0)} + 0 = -1$$

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 $\frac{ds}{d\Delta} = f'/(2f' + c' - f''\Delta)$, or $-f'/Y_{ss}$. Substituting these expressions into $T_s$ yields drops into Equation 5 below, but yields the same formula when stated in terms of $\frac{ds}{d\Delta}$. 

14
the following:

\[ T_s = \frac{1}{2} \left[ (Y_{ss}) \left( \frac{f'}{-Y_{ss}} \right)^2 + 2(f') \left( \frac{f'}{-Y_{ss}} \right) + 0 \right] (d\Delta)^2 = \frac{1}{2} f' \frac{ds}{d\Delta} (d\Delta)^2. \quad (5) \]

I then normalize by lifetime income, \( f(s_0)\Delta \), to obtain \( T \). After some manipulation, this can be written as

\[ T = \frac{1}{2} \left( \frac{d\Delta}{\Delta} \right) \beta ds = \frac{1}{2} \left( \frac{d\Delta}{\Delta} \right)^2 \beta s_0 \left( \frac{ds}{d\Delta} \Delta s_0 \right). \]

The first expression for \( T \) is the infinitesimal-change version of equation 3 above; the second corresponds to equation 7 below.

4 Constructing \( \Delta \)

\( \Delta(s) \) summarizes the integration of discounted income flows in the future for each possible \( s \). It requires two 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, Panel A. 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 \( \Delta \)’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
Figure 4: Survival Curves and Mortality Markups, Select Countries

Panel A: Survival Curve (Probability), Conditional on Surviving to Age 15 \( \frac{\ell(t)}{\ell(15)} \)

Panel B: Extra Required Return for Human Capital Because of Mortality \( \frac{1}{\Delta} - (r-g) \)

Notes: this figure displays survival curves (Panel A) and implied markups on the human-capital rate of return (Panel B) for select countries and various interest-rate assumptions. 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). The \( \Delta \) are computed as described in the text for various interest-rate assumptions. Panel B displays the difference between the inverse of \( \Delta \) and the net interest rate \( (r-g) \), thus measuring a spread between the required return to human capital and the net interest rate.
are defined from the perspective of the starting age, which means the expected age of death is $15 + e_{15}$. For an $r - g = 0$, $\Delta$ is the same as $e_{15}$. However, even for a relatively small interest-rate like 1%, the $\Delta$’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 shape 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 $\Delta$ using $(r - g) = 2\%$ on $e_{15}$, for $N = 143$ countries.

$$\Delta_i = .408 \ e_{15,i} + 9.84 + \epsilon_i$$

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.

### 4.1 Spreads in the required return to human capital

Instead of saying that mortality depresses the return to human capital, we could ask a related question: how does mortality increase the required rate of return on human capital, conditional on survival? In the absence of mortality, $\Delta = (r - g)^{-1}$. Thus, $\left(\frac{1}{\Delta} - (r - g)\right)$ is a markup / spread over the interest rate that has to be built into the return to human capital. Panel B of Figure 4 shows these spreads in the four example countries’ life tables for age $\geq 15$ and for different assumptions about $r - g$. The curves have an opposite order as before, 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. But for low interest rates, even a healthy country like Japan, there is a spread, at its largest for $r - g = 0$. Yet this spread is only as large as 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.\textsuperscript{11}

This result gives us our first 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.)

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 $\Delta$, evaluated at age 15 for a net interest rate of 2%.

Equation 3 above also requires the base of the triangle: $(s_1 - s_0)$, which measures the causal response of schooling to increased $\Delta$. 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

\textsuperscript{11}The $e_{15}$ numbers given above serve as a quick check on the difference between the curves at the intercept, where $(r - g = 0) \Rightarrow (\Delta(15) = e_{15})$. The markup for Mozambique over Japan, if $r - g = 0$, is

$$\left( \frac{1}{e_{15, MOZ}} - \frac{1}{e_{15, JPN}} \right) \approx \left( \frac{1}{43} - \frac{1}{62} \right) \approx 0.7\%.$$
is, differences in the labor market, changes in the school system, changes in the readiness to
learn of the children, etc., would have all contributed 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 \(T \geq \bar{T}\).

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 \(\Delta\), 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 \(\Delta\) 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 Mozam-
bique. When using a modest, 2% discount rate, the upper bound on the triangle drops,
typically by forty percent of the value computed with \(e_{15}\). Indeed, more than nine out of ten
of countries have \(\bar{T}_{\Delta,r-g=0.02}\) below 10%. Further, only three countries have values above 15%:
Liberia, Sierra Leone, and Mozambique, all countries for which \(\Delta\) was temporarily low.

6 Calibration of the schooling response using elastici-
ties

In this section, I use estimates of the response of schooling to \(\Delta\) 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 B shows
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 $\Delta$ and $s$. The line denotes the size of $\bar{T}$ computed with $\epsilon_{15}$ (a $\Delta$ 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.
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 five. Mozambique, the extreme example from above, had an implied elasticity of more than twelve.

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 exposure across cohorts and regions within the country to construct a difference in difference. They can also use men as a comparison group, as women are more directly exposed to 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 $\Delta$ for other net interest rates.\footnote{Their Table II reports mortality rates and their Appendix Table 2 reports effects of the shock on mortality, both 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 curve. I then use these to compute a $\Delta_1$ and $\Delta_0$.} For $r - g = 2\%$, I compute an elasticity of 0.7 in that episode.

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 rising infection also experience greater across-cohort differences in the expected marginal benefits of schooling (through $\Delta$). She finds statistically significant declines in school attendance amongst the most affected cohorts, and that such 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 Hunting-
ton’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 college years 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 used to estimate this elasticity. Baranov and Kohler (2017) find a significant increase in schooling associated with increased regional exposure to anti-retroviral therapy (used to combat HIV/AIDS) in Malawi, which essentially reverses the quasi-experiment studied by Fortson. They do not report an elasticity, but their estimates are consistent with an elasticity of 0.9.13 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 on schooling (0.17) with average values of $s$ and $e_0$ (reported in his Table 1) to compute an elasticity of approximately 1.6.14 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 all obtain elasticities close to unity.15 I use this to calibrate the triangles. Note

---

13 I thank Victoria Baranov for providing the additional numbers to construct an elasticity.

14 His 0.17 coefficient is also somewhat higher than the comparable slopes estimated in studies discussed above. Note further that he reports a slope of school w.r.t. $e_{20}$ of approximately 0.3, but he does not report the average $e_{20}$ within his sample, which precludes constructing an elasticity for working-age life expectancy.

15 This leads me to use constant-elasticity, rather than constant-slope, assumption for the response of $s$ to $\Delta$ in the calibration. That said, I did not see an obvious reason from the theory above as to why the elasticity should be constant, let alone equal to one. The elasticity, $\epsilon$, is

$$\epsilon = \frac{ds}{d\Delta} \frac{s}{s} = \left[ \frac{f'}{2f'' + c'} - f'' \right] \frac{\Delta}{s}.$$

A constant $\epsilon$ therefore implies that $f(s)$ is governed by the following differential equation:

$$f''(s) = \left( \frac{2}{\Delta(s)} - \frac{1}{se} \right) f'(s) + \frac{c'(s)}{\Delta(s)}.$$

Intuitively, curvature of the single-period production function ($f$) with respect to schooling should depend
that this is no longer an upper bound in that an elasticity of one is the (approximate) point estimate for the response of schooling to mortality reduction. (It might nevertheless be an upper bound in that (a) longevity gains past working ages are counted and (b) taking less developed countries to the Japanese life table might be an implausible assumption.) 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

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

\[
= \frac{1}{2} \left( \frac{\Delta_1 - \Delta_0}{\Delta_0} \right)^2 \beta s_0 \epsilon
\]

in which \( \epsilon \equiv \frac{ds^* \Delta}{d\Delta s^*} \), the elasticity of optimal schooling w.r.t. \( \Delta \). 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. (The \( \beta s_0 \) help transform the units into a fraction of lifetime income.)

Figure 6 displays a scatterplot of \( \Delta \) 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.

7 Sensitivity analysis

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

The transparency of the triangle calculation permits a degree of “do it yourself” sensitivity analysis. Consider equation 7 above. The triangle rises like the square of the proportional change in $\Delta$. 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 four.

The triangle is also linear in the Mincer $\beta$. 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 $\beta$ would simply reduce the triangle proportionately. (There do exist studies using instrumental variables that claim the true $\beta$ might be higher than implied by cross-sectional estimates, but not massively higher.) Readers are invited to plug in their preferred value for $\beta$.

A related issue is that there may be non-pecuniary benefits to education. These are easy to incorporate if we can put a dollar value on them. The most pertinent of such benefits might be the longevity gains that might be caused by increasing education. (See, for example, Cutler, Deaton, and Lleras-Muney, 2009, and Lleras-Muney, 2005.) In this vein, Sanchez-Gonzalez (2011, chapter 5) combines information on differences in survival by education with estimates of the value of statistical life. When he takes the extra longevity experienced by
Figure 6: Iso-Triangle Curves for Unit Elasticity of Schooling to $\Delta$

Notes: this figure displays a scatter of discounted life years ($\Delta$) versus initial years of school ($s_0$) for the sample of countries. Observations are denoted by their World Bank country codes. Overlaid on the scatterplot are iso-triangle curves; these trace out constant values for the triangle ($T$) of reallocation gains. $T$ is computed using equation 7. The underlying data sources and variable definitions are described in the text.
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. 11% instead of 10%). That being the case, the inflating the $\beta$ parameter to 11% does not materially alter the conclusions offered above.

The final parameter, also easy for the reader to change, is the elasticity of school with respect to $\Delta$. Estimates above suggest this number is close to unity, perhaps a bit below. But perhaps this elasticity is two, or perhaps it is one half. 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 changed the survival curve to Japanese levels. Thus the results in Section 5 are most likely slack upper bounds on the triangles one might obtain by cranking up the elasticity.

### 7.1 Alternate versions of $\Delta$

I now consider the sensitivity of the results in Section 6 to parameters that were used in constructing $\Delta$. While the construction of $\Delta$ 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 into quintiles according to these indices. Results for the first (lowest), second, and third (middle) quintiles of each index are seen in Panels A, B, and C. Throughout this section, I use an elasticity of one for the response of school to $\Delta$.

For the first set of columns, 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

---

Sanchez-Gonzalez estimates the mortality/education gradient using both cross-sectional comparisons and standard instrumental-variables strategies.
Column (5). In the lowest quintile, the average triangle is between 1% and 1.5%, depending on the human-development indicator used. For the second quintile, average triangles are less than 1%, and close to half a percent for the third quintile. For higher interest rates (see Cols. 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 four tenths of a percent of income. If the interest rate is lower, however, the triangles are larger. In this case, the average triangles in the lowest quintile are no greater than 4% of income, and significantly smaller than that in the higher quintiles. 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.

I also show that the result is not sensitive to the way in which life tables were mapped to cohorts. Recall that the default \( \Delta \) is constructed starting at age 15 from the circa-1990 life table using cohorts that were born in the quinquennium prior to 1980. Columns (8) and (9) present results with a \( \Delta \) constructed starting at age 10 or age 20, respectively. In Column (10), I attempt to align the starting age for \( \Delta \) 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 seven to account for a typical school-starting age. This becomes a country-specific starting age for the evaluation of \( \Delta \). Next, in Columns (11) and (12), I use cohorts that are born five years before or five years after the reference cohort. Then, in Column (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 column (5).

Finally, I consider the effect of changing the health frontier for the thought experiment. For Column (14), I imagine taking countries to the life table of the USA 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 perhaps 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 \( \Delta \) has an outsized impact on the \( T \). In Column (15), I conduct what is perhaps a more realistic exercise for health improvement: take unhealthy
countries to the life table of 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 two-thirds lower than the baseline. Next, I consider truncating the horizon over which $\Delta$ is constructed. In Columns (16) and (17), I use the survival curve up to age 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. In the final exercise, I truncate the calculation of $\Delta$ at age 65, but set $\ell(t)/\ell(s)$ equal to one. This simulates taking the survival curve in these countries past the healthy Japanese one, 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 $\Delta$, when comparing unhealthy to healthy countries, occur at relatively old ages.

### 7.2 Misallocation I: Rationing/Constraints

Above I assumed that schooling was set optimally in both mortality scenarios. Indeed, this is what makes the triangle a triangle. But what if schooling were below the optimum? The case of underinvestment in both scenarios 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 dysfunctional credit markets, for example.

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 the constraint prevents $d$ of the gain from being realized. Further, the gain $a$ arises only if the underlying constraint (e.g., a credit market problem) is relaxed. Thus, only the area $b$ is directly attributable to the horizon mechanism.

How large is the horizon polygon ($b$) in this case? Getting an approximate point estimate
Table 1: Sensitivity of Triangle to Assumptions in Constructing $\Delta$

<table>
<thead>
<tr>
<th>Variable used for ranking:</th>
<th>(r-g)</th>
<th>Initial $I(s)$</th>
<th>Ages in 2010</th>
<th>Female</th>
<th>'Frontier'</th>
<th>Horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1%</td>
<td>0%</td>
<td>0.5%</td>
<td>1%</td>
<td>2%</td>
<td>2.5%</td>
</tr>
<tr>
<td>$s_0$</td>
<td>0.036</td>
<td>0.023</td>
<td>0.019</td>
<td>0.015</td>
<td>0.010</td>
<td>0.008</td>
</tr>
<tr>
<td>$e_{15}$</td>
<td>0.056</td>
<td>0.036</td>
<td>0.029</td>
<td>0.023</td>
<td>0.015</td>
<td>0.012</td>
</tr>
<tr>
<td>$s_0 \times e_{15}$</td>
<td>0.050</td>
<td>0.032</td>
<td>0.026</td>
<td>0.021</td>
<td>0.013</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Panel A: Lowest Quintile

<table>
<thead>
<tr>
<th>Variable used for ranking:</th>
<th>(r-g)</th>
<th>Initial $I(s)$</th>
<th>Ages in 2010</th>
<th>Female</th>
<th>'Frontier'</th>
<th>Horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1%</td>
<td>0%</td>
<td>0.5%</td>
<td>1%</td>
<td>2%</td>
<td>2.5%</td>
</tr>
<tr>
<td>$s_0$</td>
<td>0.030</td>
<td>0.019</td>
<td>0.015</td>
<td>0.012</td>
<td>0.008</td>
<td>0.006</td>
</tr>
<tr>
<td>$e_{15}$</td>
<td>0.028</td>
<td>0.018</td>
<td>0.014</td>
<td>0.011</td>
<td>0.007</td>
<td>0.005</td>
</tr>
<tr>
<td>$s_0 \times e_{15}$</td>
<td>0.018</td>
<td>0.011</td>
<td>0.009</td>
<td>0.007</td>
<td>0.005</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Panel B: Second Quintile

<table>
<thead>
<tr>
<th>Variable used for ranking:</th>
<th>(r-g)</th>
<th>Initial $I(s)$</th>
<th>Ages in 2010</th>
<th>Female</th>
<th>'Frontier'</th>
<th>Horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1%</td>
<td>0%</td>
<td>0.5%</td>
<td>1%</td>
<td>2%</td>
<td>2.5%</td>
</tr>
<tr>
<td>$s_0$</td>
<td>0.013</td>
<td>0.008</td>
<td>0.007</td>
<td>0.005</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>$e_{15}$</td>
<td>0.015</td>
<td>0.009</td>
<td>0.007</td>
<td>0.006</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>$s_0 \times e_{15}$</td>
<td>0.015</td>
<td>0.009</td>
<td>0.007</td>
<td>0.006</td>
<td>0.003</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Panel C: Third (Middle) Quintile

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 $\Delta$, 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 $\Delta$, as estimated in various of the studies cited in Section 6. Triangles are computed according to equation 3. Countries are grouped into quintiles based on either years of schooling, initial years of schooling, life expectancy at age 15, or the product of the two. Results for the first (lowest), second, and third (middle) quintile are reported in Panels A, B, and C, respectively. The underlying data sources and variable definitions are described in the text.
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 the high- and low-mortality scenarios.

would require also measuring the costs of misallocation from the constraints. This would in turn 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 would remain small relative to, e.g., cross-country gaps in income.

\(^{17}\)This bound holds for all four combinations of rationing above or below in either mortality case. Indeed, if schooling is misallocated to be too high in the low-mortality environment, the horizon gains are strictly less than \( T_\$ \), rather than \( 2T_\$ \).
7.3 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 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 undistorted markets do not affect the efficiency of resource allocation—after all, in such markets, marginal values to consumers equal marginal costs of supply. (Hines, 1999, pages 178–9)

Even in the presence of distortions in other markets, the triangles computed above still describe the benefit as it pertains to the individual’s education choice, but they might fail to capture fully the impact on the rest of society.

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 far outside the scope of this study. A more feasible approach would be to analyze a distortion in the education decision itself. Suppose that there is a wedge between what the worker produces and what he/she earns. Suppose further that this wedge is a fraction $\tau > 0$ of the worker’s marginal product. The Greek letter $\tau$ is meant to evoke a tax, although it might also be an externality, which is also something the worker produces but does not receive as income. This would reduce marginal benefits to $(1 - \tau)\Delta(s)f'(s)$ but also the marginal opportunity cost to $(1 - \tau)f(s)$. But taxation is hardly the government’s only intervention in the education decision; all modern governments also subsidize schooling, and most do so heavily. Call this fractional subsidy $\alpha$. The private, marginal direct cost is thus $(1 - \alpha)c'(s)$.

The resulting first-order condition (FOC), as seen by the individual, is as follows:

$$(1 - \tau)\Delta(s)f'(s) = (1 - \tau)f(s) + (1 - \alpha)c'(s)$$
and society’s FOC corresponds to the original, undistorted version above (equation 2, e.g.). Observe that, if $\alpha = \tau$, then the individual and society have the same optimality condition. In other words, social marginal benefits and costs are equated at the privately optimal education decision. In this case, the triangle measures the proportional gains to society from the horizon mechanism.

If $\alpha < \tau$, schooling is distorted to be too low relative to the social optimum. Thus the horizon triangle would be accompanied by an external rectangle (or perhaps trapezoid) of benefits.$^{18}$ If, in contrast, $\alpha > \tau$, the privately optimal decision would imply too much schooling, from society’s perspective.

While some readers might have strong priors that positive externalities from schooling exist, ultimately this is a quantitative, not an existence, question. This idea is expressed by James Heckman in a rejoinder to a comment on one of his own articles:

> My paper does not claim that *de novo* there are no externalities in education. The issue I discuss is whether the externalities at current levels of spending warrant further subsidies. Invoking unmeasured externalities to support government interventions is a standard rhetorical device in a variety of fields of economics. It is a flabby argument which is also a potentially dangerous one—a pretext for increased government activity without any substantive basis for the intervention.

(Heckman, 2000, page 81)

Thus the question is, which is larger: $\alpha$ or $\tau$?

Throughout the world, primary schools are very heavily subsidized, by close to 100% in most cases. High- and middle-income countries often afford equally generous subsidies to secondary school, and some to tertiary education as well.$^{19}$ Let me be conservative and assume an $\alpha$ of 75%. (Highly paid professors with children in private school will dispute this number, but average parents avail themselves of subsidized schools in most countries.)

---

$^{18}$This would be, approximately, $(s_1 - s_0)(\tau(\Delta(s_0)f'(s_0) - f(s_0)) - \alpha c(s_0))$.

$^{19}$It bears mentioning that there is an issue of the quality of such schooling, but the present study is about allocation of time over the life course, in schools of a given quality.
What about the wedge, \( \tau \)? A reasonable value for an income tax(es) in high-income countries might be one third, with typically lower numbers in developing economies (Gordon and Li, 2009). A number for externalities is more nebulous, however. 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 a Indonesia. For a Mincer beta of 10%, the above estimates are like giving away one eleventh or one third, respectively, of one’s marginal output from an additional year of school. Probably the tax and spillover numbers should not simply be combined, because the tax revenue may fund public goods that raise productivity and thus look like a positive spillover. But we can construct an upper bound on the distortion/wedge by simply multiplying the two together, \((1 - \frac{1}{3}) \times (1 - \frac{1}{3})\) or four ninths, if we take the high values for taxes and externalities. As long as the subsidy for education is greater than \((1 - \frac{4}{3}) \approx 55\%\), then we are in the \((\alpha > \tau)\) case, in which the extra schooling induced by the longer horizon is a drag on the economy. This would mean that the triangle is an upper bound for the gain to society. That said, the brief treatment in this subsection is not meant to be definitive. Readers are again invited to plug in their preferred values, this time for \(\alpha\) and \(\tau\).

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. Nevertheless, I find the triangle of horizon gains to be 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. Life tables from developing countries in the early 20th century are sparse. I use circa-1930 survival curves
from India, Egypt, and China, the only still-developing countries whose data I could locate.  

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 historical life tables are drawn from *Human Life-Table Database* (Max Planck Institute for Demographic Research, 2016). Cohorts and data for the later period are as above.

Figure 8 shows some key features of these countries’ health and school performance in the 20th century. Panel A shows the 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.

How large are the resulting triangles? Table 2 shows these results. 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

---

20 An early life table for South Africa was also available, but only for whites.

21 These data are online at lifetable.de. The precise years for the early-20th-century life tables as follows:

<table>
<thead>
<tr>
<th>Country</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>1930</td>
</tr>
<tr>
<td>Japan</td>
<td>1921-1925</td>
</tr>
<tr>
<td>Egypt</td>
<td>1927-1937</td>
</tr>
<tr>
<td>China</td>
<td>1929-1931</td>
</tr>
<tr>
<td>India</td>
<td>1921-1931</td>
</tr>
</tbody>
</table>

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.

34
Figure 8: Changes in Survival and Schooling over the 20th Century, Select Countries


Panel B: Comparison of ∆

Panel C: Comparison of s

Notes: this figure displays, for select countries, the change in the survival curve (Panel A), a scatterplot of ∆ in the early and later life tables, (Panel B), and the scatterplot of years of schooling for different reference cohorts 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.
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 $\Delta$. For these computations, I maintain the use of 10% for Mincer’s $\beta$ and of 0.5 for one half.

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

<table>
<thead>
<tr>
<th>Data</th>
<th>Upper bound</th>
<th>Unit-elastic response</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportional change in $\Delta$</td>
<td>Change in school, $(s_1-s_0)$</td>
</tr>
<tr>
<td>Panel A: no discounting ($\Delta=e^{0.15}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>.17</td>
<td>3.6</td>
</tr>
<tr>
<td>Japan</td>
<td>.46</td>
<td>5.1</td>
</tr>
<tr>
<td>China</td>
<td>.31</td>
<td>6.7</td>
</tr>
<tr>
<td>Egypt</td>
<td>.50</td>
<td>8.6</td>
</tr>
<tr>
<td>India</td>
<td>.64</td>
<td>6.4</td>
</tr>
<tr>
<td>Panel B: discount at 2%/year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>.11</td>
<td>3.6</td>
</tr>
<tr>
<td>Japan</td>
<td>.29</td>
<td>2.2</td>
</tr>
<tr>
<td>China</td>
<td>.22</td>
<td>&quot;</td>
</tr>
<tr>
<td>Egypt</td>
<td>.37</td>
<td>&quot;</td>
</tr>
<tr>
<td>India</td>
<td>.44</td>
<td>&quot;</td>
</tr>
<tr>
<td>Panel C: discount at 2%/year, truncate at 65 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>.07</td>
<td>5.4</td>
</tr>
<tr>
<td>Japan</td>
<td>.20</td>
<td>3.2</td>
</tr>
<tr>
<td>China</td>
<td>.17</td>
<td>&quot;</td>
</tr>
<tr>
<td>Egypt</td>
<td>.33</td>
<td>&quot;</td>
</tr>
<tr>
<td>India</td>
<td>.37</td>
<td>&quot;</td>
</tr>
</tbody>
</table>

Notes: this table presents estimates of the triangle based on the change in $\Delta$ within select countries. The $\Delta$ 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 $\Delta$. 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 one, in accord with the evidence cited in Section 6, I find triangles that are smaller, some so small as to be almost ignorable. 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 India 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 two tenths of one percent. 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 David Hacker’s (2010, Table 8) survival curve for white males in 1850 and an $r = 2\%$, I compute a $\Delta$ of 26.4 discounted years (and an $e_{15}$ of 41.8). The proportional gain in $\Delta$ 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 three times larger than the estimate for the US for 1930–1990 and comparable to the triangle for India above.

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 of these results is worth noting. The earlier literature focuses attention mainly on the response of schooling to survival. The penultimate column of Table 2 uses the unit elasticity to compute the fraction of the observed increased schooling that is explained through the horizon mechanism. 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. 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 such problems, a change in parameters (e.g. $\Delta$) can induce a large change in decision variables, but the resulting change in the decision only causes a second-order change in the objective function.

9 Rectangles versus Triangles

While this study treats mainly the triangle of reallocation, there are a few rectangles that bear mentioning. As in Harberger’s work, these rectangles are first-order approximations to differences arising from a change in the survival curve. Neither is a measure of the welfare gain, although the first is closer than the second. (See Becker, Philipson, and Soares (2005), for direct treatment of the welfare gain associated with a better survival curve.)

The first rectangle $R$ measures the proportional increase in discounted, expected life years ($\Delta$). 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 the path of saving.\(^{22}\) Panel A of Figure 9 contains a

\(^{22}\)Consider the generic lifetime-consumption problem with time-separable utility, $u_t(c_t)$. The Lagrangian for this problem is as follows:

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

in which $y_t$ is income, $c_t$ is consumption, $\beta$ is the subjective discount rate, $\hat{u}$ is the utility of not being alive, and $\lambda$ 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

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

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. (The lapse of time in period $t$ is infinitesimal, so I ignore the change in $\lambda$.) Compare this with
plot of $R$ versus the triangle $T$ for the sample used above. This rectangle ranges from zero (in Japan, by construction) to over 40%, albeit with most of the observations below 15%. 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.

Figure 9: 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.

A second rectangle measures the increase in discounted lifetime income. This is, in fact, more trapezoid than rectangle, but I call it $\tilde{R}$ both to indicate that it is a first-order concept the derivative of $\Delta$ w.r.t. $\ell_t$, 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 $\Delta$ is part, but not all, of the change in welfare.
and to avoid confusion with the triangle $T$. 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 $\tilde{R}$ using the logic of maximization, namely 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 $\beta_t$ is constant, $R$ defined as above, and $s_1 - s_0 = \epsilon s_0 R$, I construct the following upper bound:

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

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 9. The linear aspect of the relationship between the two seems dominant in the data. 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%.

### 10 Conclusion

Death is like a tax on human-capital investments. We know what to do with taxes in public economics: we make Harberger triangles to measure efficiency losses. Here we do perhaps
the reverse. Improving the survival curve allows for an efficiency gain. The triangle 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 have a few percent more income, this is but a drop in the bucket relative to the large gap in incomes that separate rich and poor countries today (or countries today versus a century ago). Through this specific channel, it appears that health’s impact on development is small. 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. It bears mentioning that the longevity improvements themselves bring a ‘rectangle’ gain that is simply the proportional change in \( \Delta \). This is potentially large, but it is not the horizon-extension channel.

References


Appendices for Online Publication
Appendix Figure A: Survival Curves, 1990, for Countries in Main Sample

Note: this figure displays survival curves in 1990 for all countries in the sample. The data source is the World Health Organization, Global Health Observatory, 2016. Subplot titles denote World Bank country codes. Values in between reporting intervals are log-linearly interpolated as described in the text.
Appendix Figure B: 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 ($s_0$).