Policy Research Working Paper 10133

# Integrating Mortality into Poverty Measurement through the Poverty Adjusted Life Expectancy Index

*Jean-Marie Baland*
*Guilhem Cassan*
*Benoit Decerf*

WORLD BANK GROUP
Development Economics
Development Research Group
July 2022

## Abstract

Poverty measures typically do not account for mortality, resulting in counter-intuitive evaluations. The reason is that they (i) suffer from a mortality paradox and (ii) do not attribute intrinsic value to the lifespan. The paper proposes the first poverty index that always attributes a positive value to lifespan and does not suffer from the mortality paradox. This index, called the poverty-adjusted life expectancy, follows an expected lifecycle utility approach a la Harsanyi and is based on a single normative parameter that transparently captures the trade-off between poverty and mortality. This indicator can be straightforwardly generalized to account for unequal lifespans. Empirically, we show that accounting for mortality substantially changes cross-country comparisons and trends. The paper also quantifies the fraction of these comparisons that are robust to the choice of the normative parameter.

This paper is a product of the Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at bdecerf@worldbank.org.

# Integrating mortality into poverty measurement through the Poverty Adjusted Life Expectancy index.*

Jean-Marie Baland[†], Guilhem Cassan[‡], Benoit Decerf[§]

Originally published in the Policy Research Working Paper Series on *July 2022*. This version is updated on *December 2023*.

To obtain the originally published version, please email prwp@worldbank.org.

**JEL**: D63, I32, O15.
**Keyworks**: Multidimensional poverty, Poverty, Mortality, Mortality paradox.

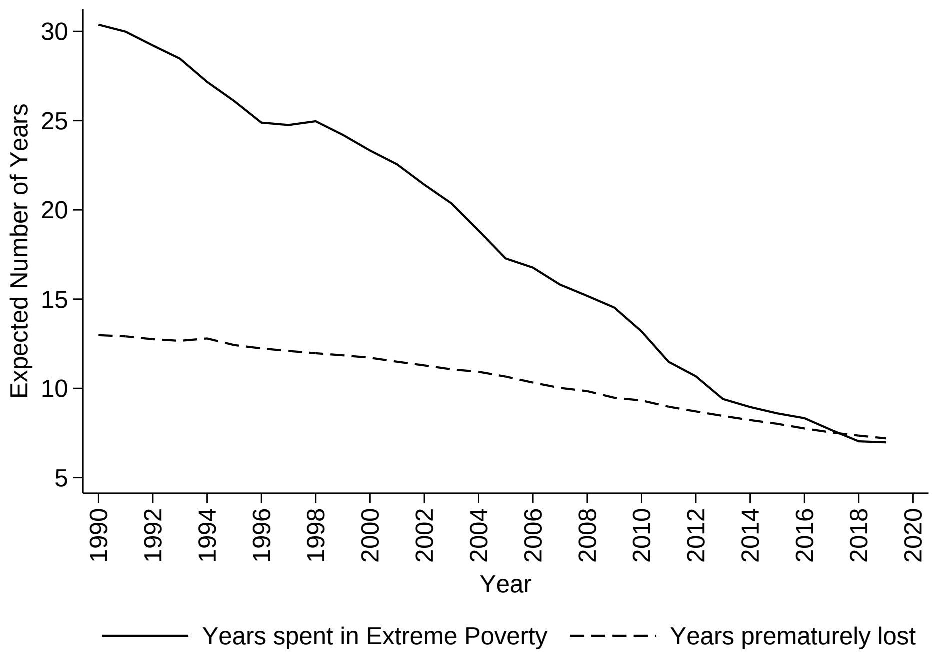
*Acknowledgments : We express all our gratitude to Kristof Bosmans, James Foster, Dilip Mookherjee and Jacques Silber for helpful discussions and suggestions. This work was supported by the Fonds de la Recherche Scientifique - FNRS under Grant n° 33665820 and Excellence of Science (EOS) Research project of FNRS n°O020918F. We are grateful to the audience at the World Bank seminar for providing insightful comments. All errors remain our own. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to members of its Board of Executive Directors or the countries they represent. The World Bank does not guarantee the accuracy of the data included in this paper and accepts no responsibility for any consequence of their use.

[†]CRED, DEFIPP, University of Namur.
[‡]CEPR, CRED, DEFIPP, University of Namur.
[§]World Bank, bdecerf@worldbank.org.

# 1 Introduction

Poverty measures are widely used for monitoring progress and guiding policies. However, most poverty measures do not account for the impact that mortality has on longevity.[1] The orders of magnitudes involved are staggering. As illustrated in Figure 1, in 2019, a new born expects to lose 7 years of life due to premature death[2] and to spend 7 years of life in poverty according to our Expected Deprivation index (which we define below). This represents, overall, 1.5 billion years of life either spent in extreme poverty or lost to premature death in that year.

**Figure 1:** Expected number of years spent in extreme poverty and prematurely lost for a newborn worldwide according to $ED_{1,70}$, 1990-2019.




Reading: in 1990, a newborn expected to spend 28 years in poverty and to lose 13 years due to premature death.

Mortality should be integrated into poverty measurement for several reasons. First, mortality reduces the lifespan of the deceased. As lifespan is a key resource, it should be attributed a positive *intrinsic* value. Second, mortality has a perverse *instrumental* impact on poverty measures. As observed by Kanbur and Mukherjee (2007), poverty measures face a "mortality paradox", since the death of poor individuals is measured as an improvement. Finally, an integrated indicator may be useful to guide policy decisions that require trading-off poverty and mortality. How to allocate a fixed budget between poverty alleviation and premature mortality reduction? How much should be spent on AIDS prevention programs? An integrated indicator that meaningfully reflects the relative impacts that poverty and mortality have on

[1] This remark also applies to measures of multidimensional poverty, which always ignore the impact of mortality on the deceased. While we refer to income poverty throughout the paper, our argument can be applied to these measures as well.

[2] We define a death as premature if it occured before 70 years old, the average life expectancy in 2019 in our data.

well-being may prove useful in guiding such choices.

However, integrating mortality in a meaningful way is not straightforward. Indeed, mortality reduces the *quantity* of life while all other forms of deprivation reduce the *quality* of life. The difficulty is that poverty – which reduces the quality of life – is typically measured *in a given year* while properly accounting for mortality – which reduces the quantity of life – requires taking a *life-cycle perspective*. As a result, an indicator that aggregates the poverty and the mortality outcomes for a given year should also reflect a life-cycle perspective. An added difficulty comes from the fact that mortality shocks generate long-lasting dynamic mechanical adjustments to population pyramids, which may blur normative comparisons. A related reason why integrating mortality into poverty measures requires a specific aggregation is that mortality necessarily excludes other forms of deprivation: individuals, once dead, cannot suffer from other forms of deprivation.

There are two views on the intrinsic value that poverty measures should attribute to lifespan and thus to mortality. The "minimalist" view holds that individuals whose death is too premature should be considered lifespan deprived. Hence, mortality matters in so far as it occurred below a given age threshold, which defines a minimally acceptable lifespan. The "maximalist" view holds that being alive is the most fundamental component of well-being. Therefore, death, no matter at which age, should always have a negative impact.[3]

In this paper, we propose a new index, the poverty-adjusted life expectancy ($PALE_\theta$), that meaningfully integrates the poverty and mortality observed in a given year under a "maximalist" view. We derive the conditions under which this index does not suffer from the mortality paradox. We then show that this index can be generalized to define a new poverty index consistent with the "minimalist" view, thereby encompassing the two different views on the integration of mortality into poverty measurement. We also study the conditions under which comparisons based on our indices are robust to all plausible values of their parameters. Our empirical application shows that our indicators substantially change poverty comparisons and quantifies the cases for which reversed comparisons are robust.

Our main indicator, $PALE_\theta$, is normatively grounded on the expected lifecycle utility, the measure of social welfare proposed by Harsanyi (1953).[4] $PALE_\theta$ normalizes the expected lifecycle utility of a newborn who *assumes* she will be confronted throughout her lifetime to the poverty and mortality prevailing in the current period.[5] This index simply counts the number of years that such newborn expects to live but weighs down the periods that she expects to live in poverty. Mathematically, our index is obtained by multiplying life expectancy at birth by a factor one minus the fraction of poor, with a lower weight being given to the latter. This (normative)

[3]Note how these two views, while conceptually different, may in practice differ only parametrically: a "minimalist" approach using a very large age threshold is in practice "maximalist".

[4]Following Harsanyi, social welfare in a given period can be understood as the lifecycle utility expected by a newborn when drawing at random a life that reflects the outcomes observed in that particular period.

[5]As we make clear later, our index is closely related to the concept of life expectancy, and its interpretation is based on similar assumptions. In particular, our index is not a forecast or a record of the actual average lifecycle utility of the cohort born in a particular period.

weight $\theta > 0$, which captures the trade-off between poverty and mortality, corresponds to the fraction of the period utility lost when poor. When being poor has no utility cost, $\theta$ takes the value zero and $PALE_0$ corresponds to life expectancy at birth. When being poor (for one year) is as bad as losing one year of life, $\theta = 1$ and our index $PALE_1$ then corresponds to the poverty-free life expectancy at birth (Riumallo-Herl et al., 2018), i.e. the number of years of life a newborn expects to live out of poverty. $PALE_\theta$ does not suffer from the mortality paradox as long as $\theta \leq 1$. Besides its theoretical properties, $PALE_\theta$ enjoys two practical advantages. First, its data-requirement are minimal as only the life-expectancy at birth and the poverty head-count ratio are necessary. Second, $PALE_\theta$ has a simple interpretation, as it measures the equivalent number of years of life spent out of poverty.

In a second step, we show how to generalize $PALE_\theta$ to account for the distribution of lifespans. More precisely, $PALE_\theta$ lends itself to the definition of a new indicator, which is consistent with the minimalist view. This requires the introduction of a normative age threshold $\hat{a}$, which corresponds to our definition of premature mortality. This new index, which we call the *expected* deprivation index ($ED_{\theta\hat{a}}$), is the weighted sum of the number of years that a newborn expects to lose prematurely or to spend in poverty, using the same weight as in $PALE_\theta$.[6] We call the index *expected* deprivation, given its proximity to the concept of life expectancy in its construction, interpretation and assumptions.

To measure the real world relevance of our indexes, we combine data sets provided by the World Bank data on income poverty (Poverty and Inequality Platform (World Bank, 2023)) and an internationally comparable data set on mortality data (Global Burden of Disease Collaborative Network, 2020) from 1990 to 2019. We show that mortality is growing in relative importance and substantially affects global poverty comparisons: during the 2005-2019 period, at least 34% of PALE's growth was due to the growth of life expectancy, as opposed to 17% from 1991 to 2004. For all possible values of $\theta$, $PALE_\theta$ is able to solve in 2019 about half of the between country and 40% of the within country comparisons when focusing on these comparisons for which life expectancy and headcount are conflicting.

**Literature review** Baland et al. (2021) proposed a way of integrating mortality into poverty measurement that is consistent with the minimalist view. They observe that an integrated indicator should at least meaningfully compare *stationary* societies, for which natality, mortality and poverty are constant over time. (In stationary societies, the outcomes observed in a given year completely reflect the life-cycle effects of their mortality.) They show that such indicators satisfy a set of basic axioms when based on a weighted sum of a number of years of life prematurely lost and a number of years of life spent in poverty.

We improve on $GD_{\theta\hat{a}}$, the main indicator proposed by Baland et al. (2021), along the following dimensions. First, given an age threshold $\hat{a}$, the death of a poor individual above the age of $\hat{a}$ is considered by $GD_{\theta\hat{a}}$ as an improvement. The "minimalist" view taken by this measure therefore implies a form a mortality paradox. Second, $GD_{\theta\hat{a}}$ is not straightforward to interpret, preventing its widespread diffusion in pub-

[6] Again, this implies that a newborn is expected to be exposed throughout her lifespan to the poverty and mortality observed in the current period.

lic debates. Third, the indicators they consider respond to mortality shocks with considerable inertia, reflecting long run adjustments in the population pyramid.[7] Inertia is not, in general, a desirable feature for poverty measures. Analyzing the relations between $ED_{\theta\hat{a}}$, $PALE_{\theta}$ and $GD_{\theta\hat{a}}$, we show that $ED_{\theta\hat{a}}$ is more reactive to shocks than $GD_{\theta\hat{a}}$ and lends itself to straightforward intepretations.

Our indicators differ from those proposed in the mortality paradox literature (Kanbur and Mukherjee (2007); Lefebvre et al. (2013)), which aimed at neutralizing the instrumental impact mortality has on poverty measurement (Decerf, 2023). By doing so, however, they do not attribute an intrinsic value to lifespan. As a result, an general increase in the life-cycle utility of a population is not necessarily considered as an improvement.

The poverty-adjusted life expectancy is reminiscent of several indicators proposed in health economics, like the quality-adjusted life expectancy (QALE) or the quality-adjusted life year (QALY).[8] Following Sullivan (1971), these two indicators account for the quality and quantity of life, by weighting down the quantity of life for periods with low quality. We show that they directly follow from the expected life-cycle utility approach in stationary societies, and correspond to well defined properties. Our index, however, accounts for a major dimension of well-being other than health, which is poverty.

The remainder of the paper is organized as follows. In Section 2, we shortly present a way to integrate poverty and mortality. In Section 3, we present the theory supporting our indicators. In Section 4, we present our global empirical application. Section 5 concludes by discussing a key limitation of our indicators, namely that neither $PALE_{\theta}$ nor $ED_{\theta\hat{a}}$ account for the unequal distribution of lifecycle utilities when the same individuals *cumulate* poverty and premature mortality.

# 2 Aggregating poverty and mortality as time units: a minimalist approach

Integrating mortality into poverty measurement is challenging. One reason is that poverty is measured in a given year while accounting for the direct impact of mortality requires a lifecycle perspective. Another difficulty is that, to be policy relevant, the combined index should account for the mortality that takes place *in the same year* as the year in which poverty is measured. In particular, the evaluation of current policies should not be affected by past mortality shocks. A solution, pioneered by Baland et al. (2021) is to express poverty and mortality outcomes in terms of years of human life, using time units to account for the life-cycle effects of mortality and poverty.[9]

[7] In particular, following a permanent mortality shock, they show that $GD_{\theta\hat{a}}$ may follow a non-monotonic trend.

[8] See for instance Whitehead and Ali (2010) for an economic interpretation of QALYs, or Heijink et al. (2011); Jia et al. (2011) for applications of the QALE index to comparisons of health outcomes across populations.

[9] Under the minimalist view, an alternative solution could be to weight the fraction of individuals who are poor in year $t$ with the fraction of individuals who will die prematurely given the mortality observed in year $t$. For instance, the Human Poverty Index is defined in this way (Watkins, 2006). Unfortunately, this seemingly natural solution suffers from the mortality paradox and yields counter-

Consider two stationary societies A and B, in which two individuals are born every year, one in the poor dynasty and one in the rich dynasty. Individuals born in the poor (resp. rich) dynasty remain poor (resp. non-poor) throughout their lives. The lifespan of a rich individual is four years. The only difference between societies A and B is the lifespan of a poor individual, which is one year in society A and three years in society B. These societies are stationary in the sense that natality is constant and the lifecycle outcomes of two individuals born in the same dynasty are the same.[10]

The relevant outcomes in society A and B are summarized in Table 1. Consider society A. In any arbitrary year $t$, one individual is poor (P) and four individuals are non-poor (NP). The head-count ratio, which we denote by $H$, is thus $1/5$. The minimalist view defines premature mortality using an age threshold ($\hat{a}$), which we assume is three years. A dead individual is considered prematurely dead (PD) if she is born less than 3 years before $t$ (and considered dead (D) if born at least 3 years before $t$).

**Table 1:** Comparison of stationary societies A and B under a lifecycle perspective.

| Age in year $t$<br>Birth year | 0<br>$t$ | 1<br>$t-1$ | 2<br>$t-2$ | 3<br>$t-3$ |
|---|---|---|---|---|
| Poor dynasty A | $P$ | $PD$ | $PD$ | $D$ |
| Non-poor dynasty A | $NP$ | $NP$ | $NP$ | $NP$ |
| Poor dynasty B | $P$ | $P$ | $P$ | $D$ |
| Non-poor dynasty B | $NP$ | $NP$ | $NP$ | $NP$ |

When mortality is ignored, the poverty comparison of societies A and B reveals a mortality paradox, since the head count ratio is larger in society B than in society A: $H(A) = 1/5$ and $H(B) = 3/7$. The longer lifespan of the poor dynasty in society B is thus recorded by $H$ as a worsening. The problem is that, in society A, $H$ does not take into account the two poor individuals born in $t-1$ and $t-2$ who died prematurely and miss year $t$.

Properly accounting for missing (dead) individuals allows for a more sensible comparison. Following the minimalist view, Baland et al. (2021) propose the inherited deprivation index ($ID_{\theta\hat{a}}$):

$$ID_{\theta\hat{a}} = \underbrace{\frac{\#PD}{\#P + \#NP + \#PD}}_{mortality\ term} + \theta \underbrace{\frac{\#P}{\#P + \#NP + \#PD}}_{poverty\ term}, \tag{1}$$

where $\#P$, $\#NP$ and $\#PD$ respectively denote the number of poor, non-poor and prematurely dead individuals in year $t$ and parameter $\theta > 0$ captures the normative trade-off between one poor individual and one prematurely dead individual. For $\theta = 1$, one poor individual contributes the same to $ID_{\theta\hat{a}}$ as one prematurely dead individual. Note that the reference population in the denominator accounts for the prematurely dead individuals, which prevents from making inconsistent trade-offs between a poor and a prematurely dead individual.

However, $ID_{\theta\hat{a}}$ is not policy relevant because this index depends on *past* mortality: it indeed measures the extent of deprivation *inherited* from the past. In our example

intuitive comparisons (e.g., concluding that society A is better off than society C, see below).

[10]Stationary societies are more formally defined in Appendix A.

above, the two prematurely dead individuals respectively died in years $t-1$ and $t-2$. One may prefer, given the need of policy relevance, indicators that only depend on current mortality – mortality in year $t$. One example is the generated deprivation index ($GD_{\theta\hat{a}}$) (Baland et al., 2021), which collects all the future years of life lost to premature deaths in year $t$ and attributes them to year $t$:

$$GD_{\theta\hat{a}} = \underbrace{\frac{YLL}{\#P + \#NP + YLL}}_{mortality\ term} + \theta \underbrace{\frac{\#P}{\#P + \#NP + YLL}}_{poverty\ term}, \tag{2}$$

where $YLL$ denotes the total number of years of life prematurely lost due to mortality in year $t$

$$YLL = \sum_{a=0}^{\hat{a}-2} N_a * \mu_a * (\hat{a} - (a+1)),$$

where $N_a$ is the number of alive individuals who have age $a$ and $\mu_a$ is the mortality rate observed for individuals who have age $a$. For example, in society A, the only premature death that takes place in year $t$ is that of the newborn in the poor dynasty. Her premature death implies that this newborn will prematurely lose two years of life, respectively in $t+1$ and $t+2$, and thus $YLL = 2$. Note that in society A, we have $YLL = \#PD$. That is, $GD_{\theta\hat{a}} = ID_{\theta\hat{a}}$ in stationary societies. This equality reflects the fact that each cell in Table 1 represents an individual as well as a unit of time.[11] The key difference is that $\#PD$ captures mortality before $t$ while $YLL$ captures mortality in $t$. By relying on units of time ($YLL$), $GD_{\theta\hat{a}}$ only depends on current mortality and compares the stationary societies A and B in a meaningful way.

As it is based on the minimalist view, $GD_{\theta\hat{a}}$ is still affected by the mortality paradox: any death of a poor occurring above the age threshold $\hat{a}$ is recorded as an improvement. To see this, consider the stationary society C whose only difference with society B is that the individuals born in the poor dynasty live for four periods. Hence, poor individuals live one year longer in C than in B. We have $GD_{\theta 3}(C) = 4\theta/8$ and $GD_{\theta 3}(B) = 3\theta/7$, which shows that $GD_{\theta\hat{a}}$ can improve with the death of a poor individual. In comparison, our approach also builds on an aggregation based on time units, considers only mortality in year $t$ but takes a "maximalist" view.[12]

# 3 Theory

In this section, we first define the poverty-adjusted life expectancy ($PALE_\theta$) index and relate it to the social welfare approach proposed by Harsanyi. We then introduce the expected deprivation ($ED_{\theta\hat{a}}$) index and characterize the conditions under which $ED_{\theta\hat{a}}$ avoids the mortality paradox. We then explore the connections between $ED_{\theta\hat{a}}$, $PALE_\theta$, $ID_{\theta\hat{a}}$ and $GD_{\theta\hat{a}}$ (Baland et al., 2021). Finally, we study the conditions

[11] Summing $YLL$ with $\#P$ and $\#NP$ in the denominator may seem strange until one realizes that all three terms capture years of human life, respectively prematurely lost, spent in poverty and spent out of poverty. Indeed, $\#P$ captures the number of individuals who spent one year – year $t$ – in poverty and thus $\#P$ is a number of poverty years. Each unit of time can be categorized as non-poor, poor, prematurely dead or dead, allowing for a proper account of the life-cycle effects of mortality. Under this approach, it is natural to compute the share of units of time spent in deprivation among the total amount of units one ought to live out of deprivation.

[12] Our measure will be compared to $ID_{\theta\hat{a}}$ and $GD_{\theta\hat{a}}$ in Subsection 3.2.

under which comparisons by $PALE_\theta$ and $ED_{\theta\hat{a}}$ are robust to all admissible values for their parameters.

## 3.1 The Poverty-Adjusted Life Expectancy index

**Definition of $PALE_\theta$**

The poverty-adjusted life expectancy index is defined as

$$PALE_\theta = LE(1 - \theta H). \tag{3}$$

where $\theta > 0$ captures the normative trade-off between one year spent in poverty and one year of life lost, $H$ denotes the poverty head-count ratio and $LE$ denotes life expectancy at birth, i.e., $LE = \sum_{a=0}^{a^*-1} \prod_{k=0}^{a-1}(1 - \mu_k^t)$ where $a^*$ denotes the maximal lifespan than can be reached and thus $\mu_{a^*-1} = 1$.

$PALE_\theta$ is the weighted sum of a number of years spent in poverty and a number of years spent out of poverty. Indeed, its mathematical expression can be written as $LE(1-H) + (1-\theta)LE\ H$. For a newborn who expects to face throughout her life the poverty and mortality observed in year $t$, the term $LE(1-H)$ captures the number of years she expects to live out of poverty and the term $LE\ H$ captures the number of years she expects to live in poverty.[13] Years out of poverty receive weight 1 and years in poverty receive weight $(1-\theta)$.

$PALE_\theta$ encapsulates the maximalist view according to which all death matters. Indeed, $PALE_\theta$ is based on life expectancy at birth, which depends on mortality rates at all ages. Its data requirements are therefore limited, as it is based simply on life expectancy at birth and the poverty head-count ratio. Two special cases are worth noting: $PALE_0$ corresponds to life expectancy at birth and $PALE_1$ corresponds to the Poverty Free Life Expectancy (PFLE), an indicator proposed by Riumallo-Herl et al. (2018).

**Relationship with social welfare a la Harsanyi**

We show that, under two assumptions, $PALE_\theta$ corresponds to social welfare a la Harsanyi. According to Harsanyi (1953), social welfare in a given year $t$ corresponds to the lifecycle utility expected by a newborn given the outcomes observed in year $t$. Behind the veil of ignorance, the newborn faces a lottery whereby she ignores whether and when she will be poor and for how long she will live. When evaluating her life-cycle utility,[14] she considers the life of a randomly drawn individual in that society. Following the formulation of Jones and Klenow (2016), her expected life-cycle utility is given by

$$EU = \mathbb{E} \sum_{a=0}^{a^*-1} \beta^a u(c_a) V(a), \tag{4}$$

[13]As we explain below, $PALE_\theta$ is not a forecast on the life of a newborn. Rather, its purpose is to jointly assess the mortality and poverty taking place in a given year.

[14]The rationality requirements of decision theory provide a structure on admissible life-cycle preferences. Rational preferences over streams of consumption have been axiomatized by Koopmans (1960) and later generalized by Bleichrodt et al. (2008). Such preferences must be represented by a discounted utility function, which aggregates these streams as a discounted sum of period utilities $U = \sum_{a=0}^{d} \beta^a u(c_a)$ where $d \in \mathbb{N}$ is the age at death, $\beta \in [0,1]$ is the discount factor, $c_a$ is consumption at age $a$ and $u$ is the period utility function.

where $\beta \in [0,1]$ is the discount factor, $c_a \geq 0$ is consumption at age $a$, $u$ is the period utility function, $V(a)$ is the (unconditional) probability that the newborn survives to age $a$, $a^*$ is the maximal lifespan one can reach and the expectation operator $\mathbb{E}$ applies to the uncertainty with respect to $c_a$. The period utility when being dead is normalized to zero, i.e., $u(D) = 0$. As a result, mortality is valued through its opportunity cost: death reduces the number of periods during which a newborn can enjoy consumption.

Under two assumptions, Eq. (4) simplifies into $PALE_\theta$. Assumption A1 is to ignore discounting, i.e. $\beta = 1$. Such assumption is necessary in order to assign equal weights to all individuals, regardless of their age.[15] Assumption A2 is to transform consumption into a binary variable, i.e., $c_a$ can be either being non-poor ($NP$) or being poor ($P$). This strong assumption implies that the impact on period utility of consumption differences within these two categories is ignored.[16] We denote the period utilities associated to being poor and being non-poor respectively by $u_P = u(P)$ and $u_{NP} = u(NP)$.

Proposition 1 shows that, under A1 and A2, in any stationary society, $PALE_\theta$ corresponds to expected life-cycle utility as expressed in Eq. (4).

**Proposition 1** (Correspondence between Harsanyi and $PALE_\theta$)**.**
*For any stationary society, assumptions A1 and A2 imply that*

$$\frac{EU}{u_{NP}} = LE\left(1 - \underbrace{\frac{u_{NP} - u_P}{u_{NP} - u_D}}_{\theta} H\right)$$

*and thus $PALE_\theta$ is ordinally equivalent to $EU$.*

*Proof.* The proof is provided in Appendix B. □

Proposition 1 calls for several remarks. First, this result provides a mathematical expression for parameter $\theta$. Parameter $\theta$ captures the fraction of her period utility that a non-poor individual looses when she becomes poor. This mathematical expression allows calibrating a value for parameter $\theta$ when selecting a period utility function and computing its value for the typical consumption of the poor and the non-poor, as we show in Appendix C.

Second, Proposition 1 holds even when mortality is selective, that is when mortality rates affect differently poor and non-poor individuals, as in the case of society A and B. The reason why $PALE_\theta$ is a simple normalization of $EU$ even when mortality is selective is that $EU$ is a *risk-neutral* social welfare function. Being risk-neutral, $EU$ is unaffected by the distribution across individuals of periods spent in poverty or lost to mortality. A social planner who cares for unequal lifespans will not evaluate welfare on the basis of Eq. (4), and may prefer $ED_{\theta\hat{a}}$, which we define below and accounts for unequal lifespans. In the conclusion, we discuss the more general case of a social planner who cares for unequal lifecycle *utilities*, when some individuals

[15] Indeed, Eq. (4) equates a society's welfare in a given period to the expected life-cycle utility of individuals born in that period. Clearly, the expected life-cycle utility of newborns is related to the society's welfare in a given period only when one assumes that their expected lives reflect at each age the outcomes observed for individuals of that age during the period considered. Discounting with a factor less than one would give less weight to the outcomes of older individuals.

[16] Assumption A2 allows us to use the head-count ratio, the simplicity of which largely explains its popularity. The headcount ratio remains however a crude indicator of poverty with well-known limitations (Sen, 1976).

combine poverty and premature mortality. Information on such individuals is often not available. As a result, a social planner who cares for unequal utilities may not do better than integrating mortality into poverty measurement through indicators like $PALE_\theta$ or $ED_{\theta\hat{a}}$.

Third, populations are in practice not stationary and we cannot in general interpret $PALE_\theta$ as the expected life-cycle utility of a newborn. Indeed, the poverty and mortality observed at birth are not necessarily good predictors of the future. Therefore, $PALE_\theta$ should *not* in general be interpreted as a projection or a forecast for $EU$. However, the validity of $PALE_\theta$ to evaluate a society in period $t$ does *not* rely on its capacity to correctly forecast the future. Indeed, our objective was to aggregate the mortality and poverty observed in period $t$ in a consistent manner, using a lifecycle perspective. This aggregation should not depend on the future evolutions of poverty and mortality.[17] Rather, one way to do so is to take the perspective of a newborn who *assumes* that she is born in a stationary society, i.e. that the poverty and mortality observed at the time of her birth remain unchanged during her whole life. It is worth noting that the same point can be made about life expectancy at birth ($LE$). In practice, this measure is derived from the mortality observed in a given period. As a result, this index does not correspond to the average lifespan of a cohort born in that period if the society is not stationary. However, life expectancy is widely accepted as a meaningful measure of period mortality.

## 3.2 The Expected Deprivation index

### Definition of $ED_{\theta\hat{a}}$

We define a new indicator – the expected deprivation index ($ED_{\theta\hat{a}}$) – which generalizes $PALE_\theta$ under a minimalist view. Under this view, one should only give a (negative) intrinsic value to the years of life lost *before* reaching a minimal age threshold $\hat{a}$. $ED_{\theta\hat{a}}$ accounts for mortality through the lifespan gap expectancy ($LGE_{\hat{a}}$), which measures the number of years that a newborn expects to lose prematurely.[18]

$$LGE_{\hat{a}} = \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a+1)) * \mu_a^t * \prod_{k=0}^{a-1} (1 - \mu_k^t),$$

We illustrate in Figure 2 the close connection between $LGE_{\hat{a}}$ and $LE$. The figure depicts for each age the fraction of newborns that are expected to still be alive at age $a$, assuming again that age-specific mortality rates are fixed. These fractions define a normalized counterfactual population pyramid. Indeed, the population pyramid of a stationary society confronted to these fixed mortality rates is obtained by multiplying these fractions by the fixed number of newborns.[19] In the left panel of Figure 2, $LE$ is proportional to the area below the normalized population pyramid. By contrast,

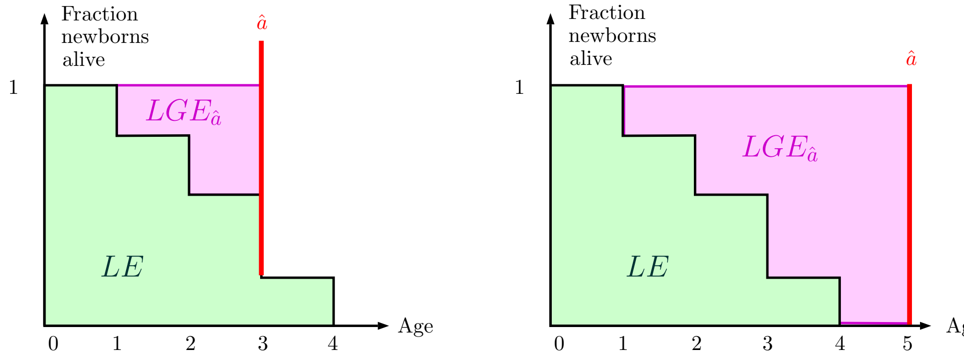
[17]For instance, a transitory mortality or poverty shock – due to war or to another disaster – does reduce current welfare, even if the country fully recovers in the next period. In contrast, the transitory nature of the shock implies that its consequences affect essentially the current generations. Its impact on the realized life-cycle utility of newborns can therefore be negligible, or nil if the shock did not affect the mortality rates of the newborns.

[18]$LGE_{\hat{a}}$ is a particular version of the Years of Potential Life Lost, an indicator used in medical research in order to quantify and compare the burden on society due to different causes of death (Gardner and Sanborn, 1990).

[19]In a stationary society, the current population pyramid can be obtained by successively applying the current age-specific mortality rates to each age group.

$LGE_{\hat{a}}$ is equal to the area between this normalized population pyramid and the age threshold. The right panel illustrates the property that, for large enough age thresholds, $LGE_{\hat{a}}$ is the complement of $LE$. Formally, when $\hat{a} \geq a^*$, where $a^*$ is the maximal lifespan, we have $LGE_{\hat{a}} = \hat{a} - LE$.

**Figure 2:** Life Expectancy and Lifespan Gap Expectancy




Note: In the Left panel, the light green area below the normalized "stationary" population pyramid is equal to $LE$ and the dark pink area is equal to $LGE_{\hat{a}}$.

The expected deprivation index ($ED_{\theta\hat{a}}$) aggregates the poverty and mortality observed in year $t$ by taking the perspective of a newborn who expects to be confronted, throughout her life-cycle, to the poverty and mortality prevailing at the time of her birth.

$$ED_{\theta\hat{a}} = \underbrace{\frac{LGE_{\hat{a}}}{LE + LGE_{\hat{a}}}}_{mortality\ term} + \theta \underbrace{\frac{LE * H}{LE + LGE_{\hat{a}}}}_{poverty\ term}, \tag{5}$$

with the same parameters $\theta > 0$ and $\hat{a} \geq 2$. The two normative parameters $\theta$ and $\hat{a}$ jointly define the respective importance attributed to poverty and mortality. Parameter $\theta$ determines the relative weights of being dead or being poor *for one period*. In contrast, parameter $\hat{a}$ determines the number of periods for which "being prematurely dead" is accounted for. Hence, $\hat{a}$ affects the relative size of the deprivation coming from mortality versus the deprivation coming from poverty.

Both terms have the same denominator, which measures a normative lifespan corresponding to the sum of $LE$ and $LGE_{\hat{a}}$. This normative lifespan can be interpreted as the (counterfactual) life expectancy at birth that would prevail if all premature deaths were postponed to the age threshold. It is at least as large as $LE$, and corresponds to $LE$ if the age threshold is equal to 1. The numerator of each term measures the expected number of years characterized by one of the two dimensions of deprivation, again assuming that the society is stationary. The numerator of the mortality term measures the number of years that a newborn expects to lose prematurely (when observing mortality in the period) given the age threshold, $\hat{a}$. The numerator of the poverty term measures the number of years that a newborn expects to spend in poverty.

### Relationship between $ED_{\theta\hat{a}}$ $ID_{\theta\hat{a}}$ and $GD_{\theta\hat{a}}$

Like $GD_{\theta\hat{a}}$, $ED_{\theta\hat{a}}$ only depends on current mortality and is thus policy relevant. Proposition 2 shows that $ED_{\theta\hat{a}}$ also compares stationary societies in the same way as $ID_{\theta\hat{a}}$ and $GD_{\theta\hat{a}}$.

**Proposition 2** ($ED_{\theta\hat{a}}$, $GD_{\theta\hat{a}}$ and $ID_{\theta\hat{a}}$ are identical in stationary societies)**.**
*For any stationary society, we have* $ED_{\theta\hat{a}} = GD_{\theta\hat{a}} = ID_{\theta\hat{a}}$.

*Proof.* See Appendix D. □

We now discuss more systematically the differences between $ED_{\theta\hat{a}}$ and $GD_{\theta\hat{a}}$. $ED_{\theta\hat{a}}$ and $GD_{\theta\hat{a}}$ rank non-stationary societies differently. The main difference between $ED_{\theta\hat{a}}$ and $GD_{\theta\hat{a}}$ comes from the way the two indices account for the number of years prematurely lost. $GD_{\theta\hat{a}}$ records the number of years prematurely lost over all premature deaths actually taking place in year $t$. $ED_{\theta\hat{a}}$ also counts the number of years prematurely lost but, instead of being computed on the actual population pyramid, $ED_{\theta\hat{a}}$ uses a counterfactual population pyramid, which is the one that would prevail in a stationary society characterized by the age-specific mortality rates observed in the period.

A major implication of this difference is that $ED_{\theta\hat{a}}$ is more reactive to policy changes than $GD_{\theta\hat{a}}$. Consider a permanent mortality shock. The population dynamics is such that a transition phase sets in during which the population pyramid slowly adjusts to the new mortality rates. This transition stops when a new stationary population pyramid is reached, typically after $a^*$ periods. $GD_{\theta\hat{a}}$ records each step of this transition and therefore exhibits inertia in its response to a permanent mortality shock.[20] By contrast, $ED_{\theta\hat{a}}$ immediately refers to the new stationary population pyramid and disregards the inertia caused by these transitory demographic adjustments. We provide an illustration of this property in Appendix E.

Finally, Baland et al. (2021) show that $GD_{\theta\hat{a}}$ is essentially the only index decomposable into subgroups to compare stationary societies in a way that satisfies some basic properties. As a result, $ED_{\theta\hat{a}}$ cannot be decomposable into subgroups.[21] This is no surprise given that $ED_{\theta\hat{a}}$ is based on life expectancy, which cannot be decomposed into subgroups. In Appendix F, we also show that $ED_{\theta\hat{a}}$ is the only index that is independent on the actual population pyramid (no inertia) and compares stationary populations in a way that respects basic properties.

### Relationship between $ED_{\theta\hat{a}}$ and $PALE_{\theta}$

We show that $PALE_{\theta}$ is a version of $ED_{\theta\hat{a}}$ that encapsulates the maximalist view. Indeed, as stated in Proposition 3, $ED_{\theta\hat{a}}$ ranks societies exactly in the same way as $PALE_{\theta}$ as long as its age threshold $\hat{a}$ is at least as large as the maximal lifespan $a^*$. For such values, the age threshold is not binding, and all deaths become relevant in terms of deprivation.

[20] For instance, assume that society A (see Table 1) undergoes a permanent mortality shock such that society A is subjected to her mortality vector in all years before $t$, namely $(\mu_0^A, \mu_1^A, \mu_2^A, \mu_3^A) = (1/2, 0, 0, 1)$, but from year $t$ onwards society A is subjected to the mortality vector of society B, namely $(\mu_0^B, \mu_1^B, \mu_2^B, \mu_3^B) = (0, 0, 1/2, 1)$. There is a mechanical adjustment to the population pyramid, such that only two poor individuals live in year $t+1$. Only in year $t+2$ does the population pyramid reach the new equilibrium, with three poor individuals. The inertia of $GD_{\theta\hat{a}}$ may be deemed undesirable because it may complicate the analysis. Baland et al. show that the mechanical adjustments following a permanent mortality shock may lead to a non-monotonic trend in $GD_{\theta\hat{a}}$.

[21] In other words, if decomposability into subgroups is seen as a key property, one should use $GD_{\theta\hat{a}}$. Indeed, this index yields the same ranking as $ED_{\theta\hat{a}}$ in stationary populations. In those populations, $GD_{\theta\hat{a}}$ thus yields the same ranking as $PALE_{\theta}$ when all deaths are normatively relevant ($\hat{a} \geq a^*$).

**Proposition 3** ($ED_{\theta\hat{a}}$ generalizes $PALE_\theta$)**.**
*For all* $\hat{a} \geq a^*$ *we have* $PALE_\theta = \hat{a}(1 - ED_{\theta\hat{a}})$*, which implies that, for any two societies A and B,*

$$PALE_\theta(A) \geq PALE_\theta(B) \Leftrightarrow ED_{\theta\hat{a}}(A) \leq ED_{\theta\hat{a}}(B).$$

*Proof.* See Appendix G. □

When the age threshold is binding (smaller than the maximal age $a^*$), the rankings obtained under $ED_{\theta\hat{a}}$ may not correspond to the rankings obtained under $PALE_\theta$. In Appendix H, we contrast the impact of mortality shocks on $PALE_\theta$ and $ED_{\theta\hat{a}}$.

### Escaping the mortality paradox with $PALE_\theta$

As observed in Section 2, $GD_{\theta\hat{a}}$ is not immune to the mortality paradox. A paradox-free index should record an improvement when a stationary society is obtained from another stationary society by *an increment to the lifespan of a poor person.*[22] For instance, in Table 1, stationary society B is obtained from stationary society A by two successive increments to the lifespan of individuals in the poor dynasty.

**Definition 1** (Paradox-free)**.**
$M_{\theta\hat{a}} \in \{ED_{\theta\hat{a}}, GD_{\theta\hat{a}}, ID_{\theta\hat{a}}\}$ *is paradox-free if for any two stationary societies A and B such that B is obtained from A by an increment to the lifespan of a poor person we have* $M_{\theta\hat{a}}(A) \geq M_{\theta\hat{a}}(B)$*.*

Proposition 4 identifies the values for the two parameters $\theta$ and $\hat{a}$ under which $ED_{\theta\hat{a}}$ is Paradox-free. First, all deaths should matter, which implies that the age threshold $\hat{a}$ should be at least as large as the maximal lifespan $a^*$. Second, one year of life prematurely lost should be at least as bad as one year of life spent in poverty, which implies that $\theta \leq 1$ and thus $u_{NP} - u_P \leq u_{NP} - u_D$.

**Proposition 4** ($ED_{\theta\hat{a}}$ and the mortality paradox)**.**
$M_{\theta\hat{a}} \in \{ED_{\theta\hat{a}}, GD_{\theta\hat{a}}, ID_{\theta\hat{a}}\}$ *is Paradox-free if and only if* $\theta \leq 1$ *and* $\hat{a} \geq a^*$*.*

*Proof.* See Appendix I. □

An immediate corollary for Propositions 3 and 4 follows: the only way for $ED_{\theta\hat{a}}$ to be Paradox-free is to be ordinally equivalent to $PALE_\theta$ with $\theta \leq 1$.

**Corollary 1** ($PALE_\theta$ and the mortality paradox)**.**
$ED_{\theta\hat{a}}$ *is Paradox-free if and only if* $ED_{\theta\hat{a}}$ *is ordinally equivalent to* $PALE_\theta$ *with* $\theta \leq 1$*.*

These results show that indicators that embody the minimalist view cannot avoid the mortality paradox. The mortality paradox can only be avoided when the deaths taking place at older age are also attributed negative intrinsic value. Corollary 1 thus shows that the mortality paradox provides a justification for $PALE_\theta$.

[22] We define more formally the notion of an increment to the lifespan of a poor person in this footnote. Following our formal framework presented in Appendix A, the life of an individual $i$ is a list of poverty statuses $l_i = (l_{i0}, \ldots, l_{id_i})$ that she experiences between age 0 and the age at which she dies $d_i \in \{0, \ldots, a^* - 1\}$, where $l_{ia} \in \{NP, P\}$. We say that stationary society B is obtained from stationary society A by an increment to the lifespan of a poor person when both societies have the same natality, $l_i^A = l_i^B$ for all individuals $i$ except for some individual $j$ such that $d_j^B = d_j^A + 1$, $l_{ja}^A = P$ for all $a \leq d_j^A$ and $l_{ja}^B = P$ for all $a \leq d_j^B$.

### 3.3 Robust comparisons

We study the conditions under which comparisons by $PALE_\theta$ are robust to the plausible values for parameter $\theta$.

By Corollary 1, $PALE_\theta$ is Paradox-free when $\theta \in (0, 1]$. Yet, the comparison of two societies with $PALE_\theta$ may depend on the particular value assigned to $\theta \in (0, 1]$. We show that a nontrivial part of these comparisons does not depend on the value for $\theta$ *even for some pairs not related by domination.* In other words, there exist pairs of societies such that one is poorer but the other has higher mortality that are *robustly* ranked by $PALE_\theta$, that is, in the same way for all values of $\theta \in (0, 1]$.

We illustrate this property in Figure 3. Without aggregation, domination alone allows comparing society A with the northwest quadrant (where societies have more poverty and more mortality) and the southeast quadrant (where societies have less poverty and less mortality). For any value of $\theta$, we can draw the iso-PALE$_\theta$ curves passing through A. The iso-PALE$_0$ curve (associated to $\theta = 0$) is a vertical line since poverty has no welfare costs and life expectancy is the sole determinant of welfare. However, the iso-PALE$_1$ curve (associated to $\theta = 1$) is *not* a horizontal line. This defines two additional areas for which welfare can be robustly compared with that of society A. The iso-PALE$_\theta$ curves associated to intermediate values of $\theta \in (0, 1]$ are indeed all located in the area between the iso-PALE$_0$ curve and the iso-PALE$_1$ curve. The area in the NE quadrant below the iso-PALE$_1$ curve yields an robustly higher social welfare than A, even though these societies have a higher poverty than A. The area in the SW quadrant above the iso-PALE$_1$ yields an robustly lower social welfare than A, even though these societies have a lower poverty than A. The size of these new areas depends on the marginal rate of substitution of $PALE_1$ at A. For society A and $PALE_1$, this marginal rate of substitution is given by $\frac{LE(A)(1-H(A))}{(LE(A))^2}$. If $LE(A) = 70$ and $H(A) = 20$, this marginal rate of substitution is equal to 0.011, meaning that one additional year of life is exactly compensated by an increase in the head-count ratio $H$ of 1.1% percentage points.These additional robust comparisons follow from (i) the fact that expected life-cycle utility sums period utilities and (ii) the assumption that a year of life spent in poverty is considered not worse than a year of life lost (i.e., $1 \geq \theta$, which is $u_P \geq u_D$).

As an illustration, Table 2 below reports the situation of Pakistan and Bangladesh in 2019. Note that Life Expectancy can trivially be decomposed into Poverty Expectancy (LE*H) and Poverty Fee Life Expectancy (LE*(1-H)). Pakistan has a lower headcount ratio than Bangladesh, but life expectancy is also lower in Pakistan. Therefore, it is a priori difficult to rank those two societies. Assuming that poverty and mortality remain unchanged, an individual born in Bangladesh can expect to spend 4.9 years of his life in poverty and 68.8 years out of poverty. In Pakistan, he can expect 2.8 years in poverty and 62.1 years out of poverty. Hence, a newborn in Bangladesh can not only expect to spend more years in poverty, but also more years out of poverty since the longer life expectancy there more than compensates for the higher poverty rate. As a result, $PALE_\theta$ ranks Bangladesh above Pakistan for all $\theta \in (0, 1]$.

In the absence of domination (NE and SW quadrants in Figure 3), ignoring mortality, i.e., comparing two societies based on $H$, may lead to robustly erroneous comparisons. This happens when the ranking provided by $PALE_\theta$ is robust but

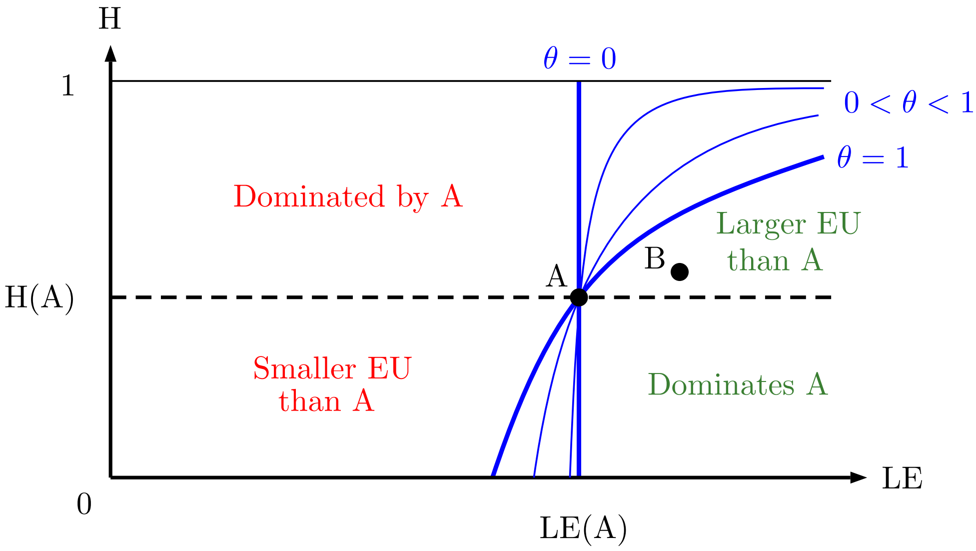



**Figure 3:** A and B are robustly ranked even though $H(A) < H(B)$ and $LE(A) < LE(B)$.

**Table 2:** An example of robust comparison: Pakistan and Bangladesh in 2019.

| | Headcount ratio | Life Expectancy | Poverty Expectancy ($LE * H$) | Poverty Free Life Expectancy $LE * (1 - H) = PALE_1$ |
|---|---|---|---|---|
| Pakistan | 4.3% | 64.8 | 2.8 | 62.1 |
| Bangladesh | 6.7% | 73.6 | 4.9 | 68.8 |

differs from the ranking provided by $H$. Proposition 5 describes the conditions under which $PALE_\theta$ comparisons are robust.

**Proposition 5.** *(Robust comparisons with $PALE_\theta$)*
*(i) For any two societies A and B, $PALE_\theta(A) \leq PALE_\theta(B)$ for all $\theta \leq 1$ if and only if*

$$PALE_0(A) \leq PALE_0(B) \text{ and } PALE_1(A) \leq PALE_1(B) \qquad \text{(Condition C1)}$$

*(ii) There exist societies A and B for which $PALE_\theta(A) \leq PALE_\theta(B)$ for all $\theta \leq 1$ even though $H(A) < H(B)$. These societies are such that $H(A) < H(B)$ and $LE(A) < LE(B)$.*

*Proof.* See Appendix J. □

In Appendix K, we study the conditions under which comparisons by $ED_{\theta\hat{a}}$ are robust to the plausible values for its parameters.

## 4 Real world implications

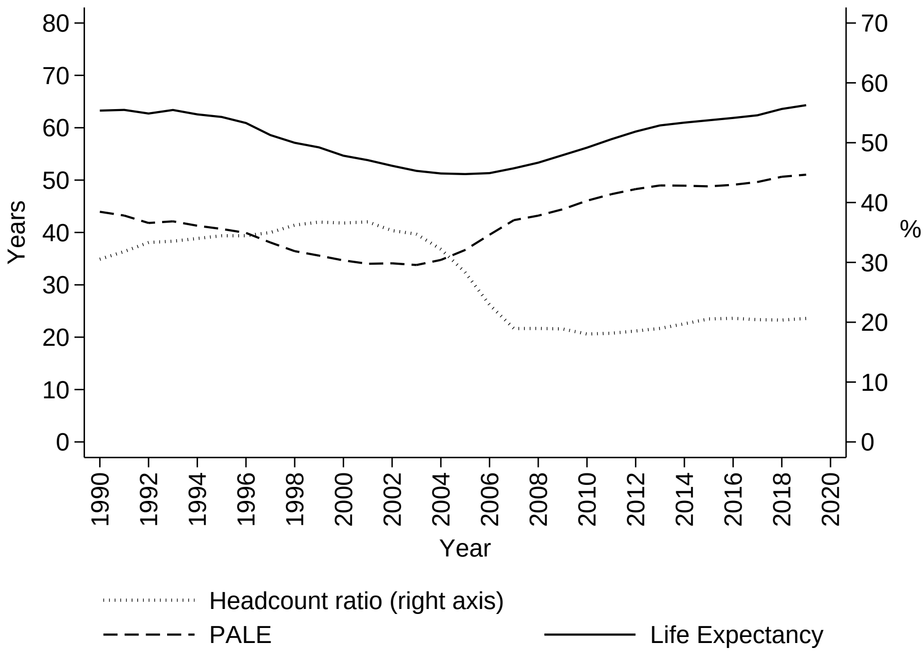
We now turn to data on poverty and life expectancy spanning the period 1990-2019. The data come respectively from the World Bank's Poverty and Inequality Platform (World Bank, 2023) and the Global Burden of Disease Project (Global Burden of

Disease Collaborative Network, 2020).[23] Appendix N presents a practictioner guide to the construction of our index.

## 4.1 A case study of South Africa

We first illustrate the relevance of our indices with the case of South Africa. Figure 4 reports the evolution of life expectancy, poverty rate and $PALE_1$ for South Africa from 1990 to 2019. From the perspective of poverty, the progress of South Africa is impressive, with poverty rates decreasing from 31% to 21% over the period. However, life expectancy shows a different pattern. Following the AIDS epidemic, life expectancy decreased from the mid 90s onwards, to revert back to the pre-AIDS levels after 2013. Thus, in 2007, poverty rates are low, at 19%, but life expectancy is also low, at only 53 years. How then do we compare South Africa in 2007 to South Africa in 1990 ? $PALE_1$ indicates that deprivation is higher in 2007 than in was in 1990. Indeed, $PALE_1$ is equal to 44 years in 1990 as opposed to 42 years in 2007. We discuss in Section 4.3 the sensitivity of the comparisons made under $PALE_\theta$ to the choice of $\theta$.

**Figure 4:** South Africa Evolution of $PALE_1$ and Life Expectancy, 1990-2019



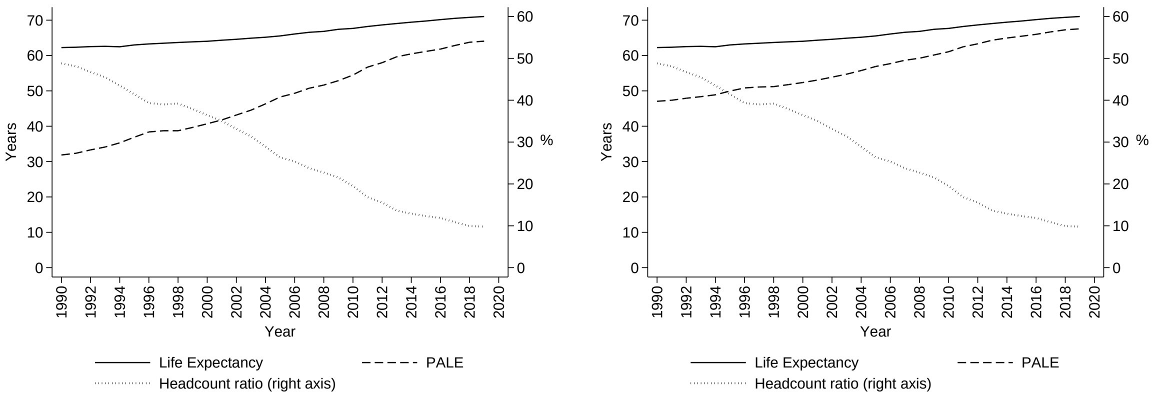

Reading: in 1990, life expectancy was 63 years, 31% of the population was living below the poverty line and poverty adjusted life expectancy was 44 years.

## 4.2 Life expectancy and poverty in the World, 1990-2019

At the world level, Figure 5 presents the evolution of life expectancy, the headcount ratio and $PALE_\theta$ between 1990 and 2019. Throughout this period, life expectancy

[23] See Appendix M for the list of countries in the database as well as their descriptive statistics for the year 2019.

**Figure 5:** Evolution of $PALE$ and Life Expectancy, 1990-2019




**(a)** $PALE_1$ **(b)** $PALE_{0.5}$

Reading: in 1990, Poverty-Adjusted Life Expectancy was about 30 years according to $PALE_1$ and 48 years according to $PALE_{0.5}$.

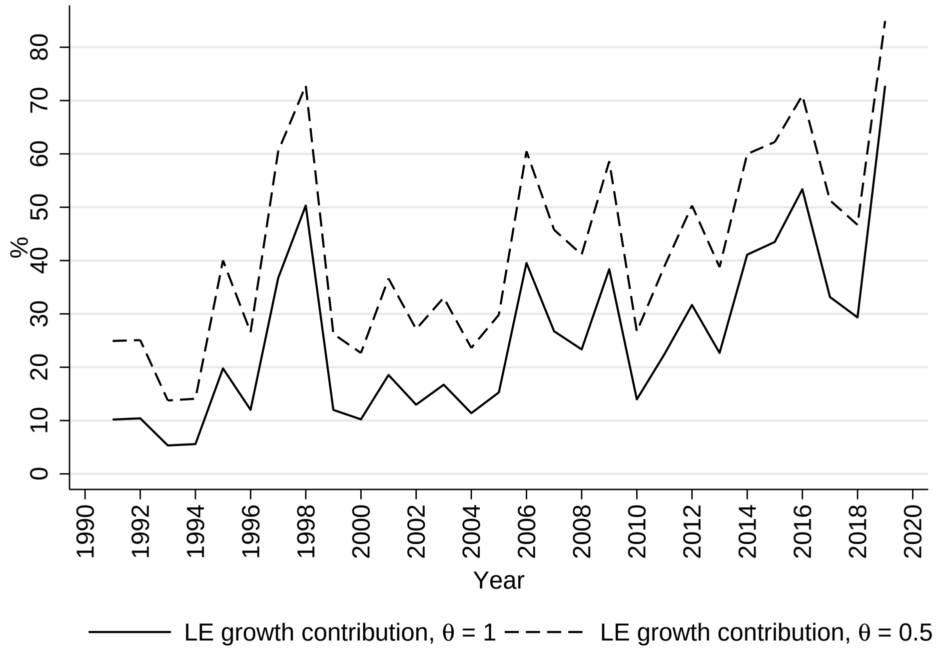
increased from 62 to 71 but the decrease in poverty expectancy is even more spectacular, from 30 years in 1990 to only 7 years in 2019. This decrease in poverty combined with an increase in life expectancy resulted in a large increase in $PALE_1$, from 32 in 1990 to 64 years in 2019. For $\theta < 1$, the corresponding $PALE_\theta$ curves all lie between life expectancy and the $PALE_1$ curve. We show $PALE_{0.5}$ in the right hand panel of Figure 5. By construction, $PALE_{0.5}$ is higher in absolute value (47 years in 1990). However, its evolution is much slower than that of $PALE_1$: from 1990 to 2019, $PALE_{0.5}$ increased by 44% as opposed to 101% for $PALE_1$. Indeed $\theta = 1$ implies that one year spent in poverty is equivalent to one year spent dead. When instead one assumes that a year spent in poverty is equivalent to half a year lost to death, life expectancy has more weight in PALE. As the progress on life expectancy have been much slower than those against poverty, $PALE_{0.5}$ growth is slower.

Note that PALE can not be directly decomposed into each of its components. However, it is possible to decompose its growth into the contribution of that of each of its component. Indeed, the growth rate of $PALE_\theta$ can be decomposed as follows:

$$\frac{\partial PALE}{\partial t} = \epsilon_{LE} * \frac{\partial LE}{\partial t} + \epsilon_H * \frac{\partial H}{\partial t}$$

where $\epsilon_{LE} = 1$ and $\epsilon_H = \frac{-\theta H}{(1-\theta H)}$ represent the elasticities of $PALE_\theta$ to life expectancy and poverty, respectively. Figure 6 shows the share of the growth of $PALE_\theta$ explained by changes in life expectancy, from 1991 to 2019. First, note that the choice of $\theta = 1$ is conservative: the contribution of life expectancy to PALE is on average 16 percentage point smaller when $\theta = 1$ than when $\theta = 0.5$. Second, irrespective of the precise value given to $\theta$, the contribution of life expectancy is growing over time. Life expectancy contributes on average to 17 (resp. 32) percent of PALE's growth from 1991 to 2004, as opposed to 34% (resp. 51%) from 2005 onwards. Even though mortality has not decreased as much as poverty, the changes in mortality still play a substantial role in the trend of $PALE_\theta$.

**Figure 6:** Share of the growth of LE in the growth of $PALE_1$ and $PALE_{0.5}$, 1990-2019




Reading: in 1991, the growth of life expectancy contributed to 10% of the growth of $PALE_1$ and to 25% to that of $PALE_{0.5}$.

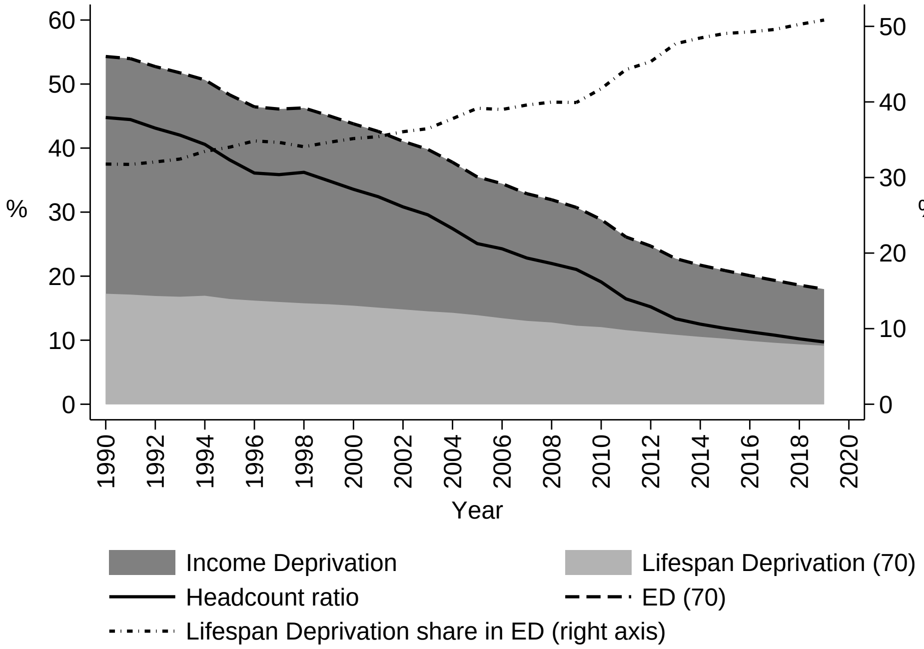
Turning now to deprivation as captured by the $ED_{\theta\hat{a}}$ index, Figure 7 reports the evolution of $ED_{1,70}$ in the world during the same period.[24] The evolution of $ED_{1,70}$ illustrates the vast progress made against deprivation at the world level. While a newborn expected to have 58% of its normative lifespan either lost to mortality or spent in poverty in 1990, this expected loss falls to 18% in 2019. Unlike PALE, ED can be directly decomposed into each of its components, as indicated in the graph. In particular, the share of the lifespan deprivation component increased over time: from 30% in 1990 to 51% in 2019. As a result, the neglect of mortality in most poverty measures amounts to missing an increasing and significant share of total deprivation.

## 4.3 Improving on comparisons based on the headcount only

Our indices also allow for comparing countries in which life expectancy and poverty evolve in opposite direction, as in the Pakistan-Bangladesh comparison presented in Table 2 or in the comparison between the years 1990 and 2007 in South Africa (Figure 4). However, the choice of $\theta$ may not be innocuous in these comparisons. We now focus on these cases. For these cases, we discuss the extent to which $PALE_\theta$ offers comparisons that are robust, that is, for which the ranking is not affected by the choice of $\theta$. Note that if the ranking proposed by $PALE_\theta$ is robust to the choice of $\theta$, this implies that $PALE_0$ and $PALE_1$ yield the same ranking. Since $PALE_0$ corresponds to life expectancy, a measure solely based on the headcount provides a

[24]We take 70 as the lifespan deprivation threshold since world life expectancy is 71 years in 2019. This threshold is therefore a reasonable choice given our maximalist perspective: dying before this age means dying below the average expected age of death at the world level.

**Figure 7:** Evolution of $ED_{\theta\hat{a}}$ and H, 1990-2019 (where $\theta = 1$ and $\hat{a} = 70$).




Reading: in 1990, a newborn expected to be deprived of 58% of its life: 17 percentage points lost because of lifespan deprivation and 41 because of income deprivation. Lifespan deprivation accounted for 30% of expected deprivation in 1990.

wrong ranking whenever life expectancy and headcount diverge (which is the case we focus on). The main interest of robust comparisons is to measure the extent to which PALE allows to improve on a ranking based on the headcount only: in all these situations, irrespective of the value given to $\theta$, the ranking under PALE contradicts the poverty ranking.

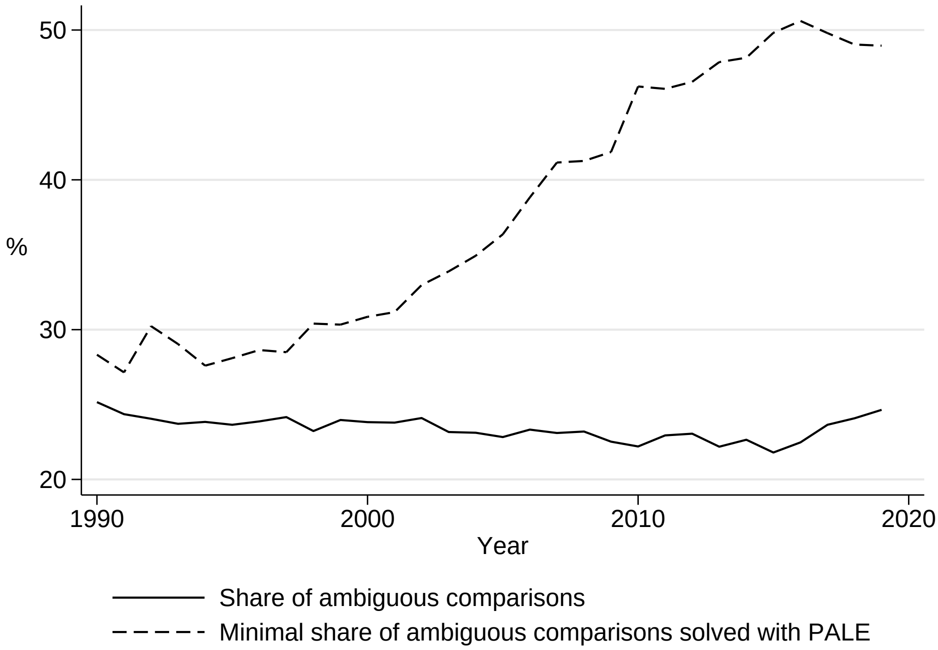
**Inter country comparisons**

To what extent does $PALE_\theta$ help in robustly ranking countries, as compared to a simple headcount? Figure 8 reports the proportion of all country-pairs comparisons whose ranking based on life expectancy and headcount ratio differs. There are 23% of them.[25] The share of these ambiguous cases for which $PALE_\theta$ provides a robust answer is equal to 37 percent, independently of the value given to $\theta$. In other words, 37% of these "ambiguous" cases are wrongly classified by the headcount ratio. Note also that the share of ambiguous comparisons that our index unambiguously solves strongly increases over time, owing to the falling incidence of absolute poverty in many countries.[26]

[25] These are the only situations in which PALE can offer a different ranking than the headcount.

[26] The falling incidence of absolute poverty implies that differences in H across countries in a given year become, on average, smaller over time. This explains why the share of ambiguous comparisons that our index unambiguously solves increases over time. This is easy to see when assuming that the differences in LE across countries in a given year remain constant over time. Indeed, a smaller difference in H can be "over-compensated" by a smaller difference in LE.

**Figure 8:** Evolution of the resolution of ambiguous inter-country comparisons, 1990-2019




Reading: in 1990, countries had on average 23% of ambiguous comparisons, out of which at least 26% were solved by the use of PALE.

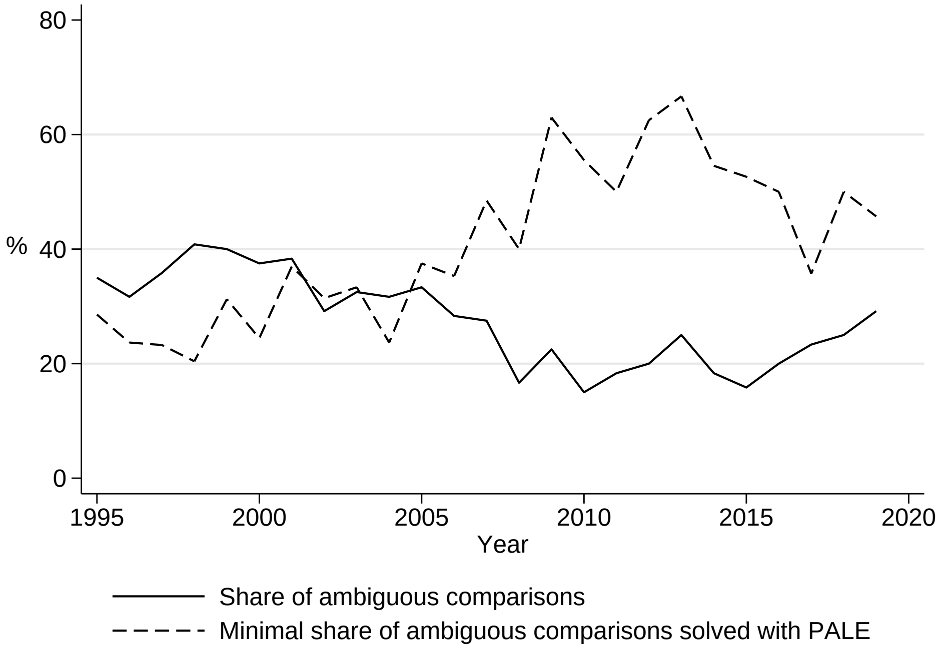
**Countries' trajectories**

We now turn to individual trajectories of all countries, such as the South African case discussed earlier. For each country in our data, we computed the growth rate of H and LE over each 5-years period. Figure 9 presents the evolution of the share of ambiguous intra-country comparisons as well as the share that is robustly resolved by $PALE_\theta$. Over the period, the share of ambiguous trajectories oscillates between 20 and 40% of all cases. The share of these cases that $PALE_\theta$ ranks unambiguously varies between between 20 and 40% for the period 1995-2005 up to 40 to 60% in the 2005-2015 period. As above, $PALE_\theta$ corrects an increasing share of the rankings proposed by the headcount. In Appendix O, we present each country's evolution for the period 1990-2019 and its resolution in a graphical format reminiscent of the theoretical Figure 3.

# 5 Concluding remarks

An important limitation of the two indices proposed in this paper, $PALE_\theta$ and $ED_{\theta\hat{a}}$, is that they account for the distribution of outcomes "dimension-by-dimension". More precisely, they account for the distribution of quality of life and for the distribution of quantity of life, but not for the distribution of life-cycle utilities. Indeed, our indices are insensitive to the allocation of years of life prematurely lost between the poor and the non-poor. This allocation may however have implications for the distribution of life-cycle utilities. When poor individuals die early, they cumulate low achievements

**Figure 9:** Evolution of the resolution of ambiguous countries' trajectories, 1990-2019




Reading: in the 1995, 35% of countries' trajectories was ambiguous. Among these, 29% can be assessed with PALE.

in the two dimensions and the difference between their life-cycle utility and that of non-poor individuals increases.

Accounting for the distribution of lifecycle utilities requires data that are typically not available. The necessary data include not only information on the correlation between poverty and premature mortality, but also information on mobility in and out of poverty. When such data is not available, $PALE_{\theta}$ and $ED_{\theta\hat{a}}$ can be used as a second-best solution, as they improve over the widespread practice of entirely ignoring the impact of mortality on longevity. This is particularly relevant for societies in which premature mortality is highly selective, affecting disproportionately poorer individuals. In particular, the premature mortality term of $ED_{\theta\hat{a}}$ essentially captures these negative outcomes.

If one cares about the distribution of life-cycle utilities and the necessary data is available, our indicators would need to be adjusted. Let us define as "life-cycle poor" the individuals whose life-cycle utility is smaller than that of a reference life, e.g., a life characterized by a lifespan of 40 years with no period of poverty. One index combining mortality and poverty that would account for the distribution of life-cycle utilities is the expected fraction of newborns who will be "life-cycle poor", again assuming constant poverty and mortality.[27]

Our paper calls for future research on the value that the normative parameter $\theta$ should take. Its mathematical expression based on social welfare a la Harsanyi allows calibrating its value, as we show in Appendix C. However, the calibrated

[27] Note that the indicators proposed in the literature on the mortality paradox are typically not appropriate to capture the distribution of life-cycle utilities, as they do not attribute an intrinsic value to the quantity of life. The may therefore miss improvements when the lifecycle utility of all individuals increase, for instance if the lifespan of all individuals is multiplied by a common factor.

values are highly sensitive to the parametric values selected for the period utility function. Survey-based estimates for $\theta$ may provide a firmer base for narrowing the plausible range of values for this central parameter.

# References


Baland, J.-M., Cassan, G., and Decerf, B. (2021). “too young to die”: Deprivation measures combining poverty and premature mortality. *American Economic Journal: Applied Economics*, 13(4):226–257.

Bleichrodt, H., Rohde, K. I., and Wakker, P. P. (2008). Koopmans constant discounting for intertemporal choice: A simplification and a generalization. *Journal of Mathematical Psychology*, 52(6):341–347.

Decerf, B. (2023). A preference-based theory unifying monetary and non-monetary poverty measurement. *Journal of Public Economics*, 222:104898.

Gardner, J. and Sanborn, J. (1990). Years of Potential Life Lost (YPLL): What Does it Measure? *Epidemiology*, 1:322–329.

Global Burden of Disease Collaborative Network (2020). Global burden of disease study 2019 (gbd 2019) results. *Seattle, United States: Institute for Health Metrics and Evaluation (IHME).*

Harsanyi, J. C. (1953). Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy*, 61(5):434–435.

Heijink, R., Van Baal, P., Oppe, M., Koolman, X., and Westert, G. (2011). Decomposing cross-country differences in quality adjusted life expectancy: the impact of value sets. *Population health metrics*, 9(1):1–11.

Jia, H., Zack, M. M., and Thompson, W. W. (2011). State quality-adjusted life expectancy for us adults from 1993 to 2008. *Quality of Life Research*, 20(6):853–863.

Jones, C. I. and Klenow, P. J. (2016). Beyond GDP? Welfare across countries and time. *American Economic Review*, 106(9):2426–2457.

Kanbur, R. and Mukherjee, D. (2007). Premature mortality and poverty measurement. *Bulletin of economic research*, (September):339–359.

Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica: Journal of the Econometric Society*, pages 287–309.

Lefebvre, M., Pestieau, P., and Ponthiere, G. (2013). Measuring poverty without the Mortality Paradox. *Social Choice and Welfare*, 40(1):285–316.

Riumallo-Herl, C., Canning, D., and Salomon, J. A. (2018). Measuring health and economic wellbeing in the sustainable development goals era: development of a poverty-free life expectancy metric and estimates for 90 countries. *The Lancet Global Health*, 6(8):e843–e858.

Sen, A. (1976). Poverty: an ordinal approach to measurement. *Econometrica: Journal of the Econometric Society*, pages 219–231.

Sullivan, D. F. (1971). A single index of mortality and morbidity. *HSMHA health reports*, 86(4):347.

Watkins, K. (2006). *Human Development Report 2006 - Beyond scarcity: Power, poverty and the global water crisis*, volume 28.

Whitehead, S. J. and Ali, S. (2010). Health outcomes in economic evaluation: the qaly and utilities. *British medical bulletin*, 96(1):5–21.

World Bank (2023). Poverty and inequality platform.

# Appendices

## A Notation and definition of stationary society

We present here the formal notation used for the proofs.

There is a discrete set of periods $\{\dots, t-1, t, t+1, \dots\}$. In each period, some individuals are born (at the beginning of the period) and some individuals die (at the end of the period). All alive individuals are assigned a consumption status for the period ($P$ or $NP$) . We define the life of an individual $i$ as the list of consumption statuses $l_i = (l_{i0}, \dots, l_{id_i})$ she enjoys between age 0 and age $d_i \in \{0, \dots, a^* - 1\}$ at which she dies, where $l_{ia} \in \{NP, P\}$. The set of lives is thus $L = \cup_{d\in\{0,\dots,a^*-1\}}\{NP, P\}^{d+1}$.

The number of newborns in period $t$ is denoted by $n_t$. The profile of lives for the cohort born in $t$ is denoted by $C_t = (l_i)_{i\in\{1,\dots,n_t\}}$, where $\{1, \dots, n_t\}$ is the set of newborns in $t$.

Let $n_t(a)$ denote the number of individuals born in period $t$ who are still alive when reaching age $a$. In particular, we have $n_t(0) = n_t$. Let $p_t(a)$ denote the number of individuals born in period $t$ who are poor at age $a$, with $p_t(a) \leq n_t(a)$. By definition, the probability that an individual born in $t$ survives to age $a$ is given by $V_t(a) = \frac{n_t(a)}{n_t}$, and the conditional probability that an individual born in $t$ will be poor when reaching age $a$ is $\pi_t(a) = \frac{p_t(a)}{n_t(a)}$. We denote the distribution on the set of lives that $C_t$ implicitly defines by $\Gamma_t : L \to [0, 1]$, with $\sum_{l\in L} \Gamma_t(l) = 1$.

In period $t$, we cannot observe the profile of lives for the cohort born in $t$. The only elements of $C_t$ that we observe in period $t$ are $n_t(0)$, $p_t(0)$ and $n_t(1)$. However, we also have information about the profile of lives of the cohorts born *before* $t$. Formally, let a **society** $S_t$ be the list of profiles of lives for all cohorts born during the $a^*$ periods in $\{t - (a^* - 1), \dots, t\}$, i.e. $S_t = (C_{t-a^*+1}, \dots, C_t)$. In period $t$, we observe (i) the number $N_t$ of individuals who are alive in $t$:

$$N_t = \sum_{a=0}^{a^*-1} n_{t-a}(a),$$

(ii) the fraction $H_t$ of alive individuals who are poor in $t$:

$$H_t = \frac{\sum_{a=0}^{a^*-1} p_{t-a}(a)}{\sum_{a=0}^{a^*-1} n_{t-a}(a)},$$

and (iii) the age-specific mortality vector $\mu^t = (\mu_0^t, \dots, \mu_{a*-1}^t)$ in period $t$ where for each $a \in \{0, \dots, a^* - 1\}$ we have

$$\mu_a^t = \frac{n_{t-a}(a) - n_{t-a}(a+1)}{n_{t-a}(a)},$$

with $\mu_{a^*-1}^t = 1$ (by definition of $a^*$).

The particularity of stationary societies is to have their natality, mortality and poverty constant over time, so that average outcomes in a given period are replicated over the next period. More formally, a society is stationary if both the distribution of lives and the size of generations are constant over the last $a^*$ periods.

**Definition 2** (Stationary Society)**.**
*A society $S_t$ is stationary if, at any period $t' \in \{t - a^* + 1, \dots, t\}$, we have*

- $\Gamma_{t'} = \Gamma_t$ *(constant distribution of lives),*
- $n_{t'} = n_t$ *(constant size of cohorts).*

It follows from this definition that $n_t(a) = n_{t-a}(a)$ and $p_t(a) = p_{t-a}(a)$ for all $a \in \{1, \dots, a^* - 1\}$.[28]

# B Proof of Proposition 1

The proof is based on Lemma 1, which shows that, in a stationarity society, the poverty and mortality observed in a given period completely reflects the life profile of newborns.

**Lemma 1.** *If society $S_t$ is stationary, then*

$$V_t(a) = \Pi_{k=0}^{a-1}(1 - \mu_k^t) \qquad \text{for all } a \in \{0, \dots, a^* - 1\}, \tag{6}$$

$$N_t = n_t * LE_t, \tag{7}$$

$$N_t * H_t = n_t * \sum_{a=0}^{a^*-1} V(a)\pi(a). \tag{8}$$

*Proof.* We first prove Eq (6). As $S_t$ is stationary, we have $n_t(k) = n_{t-k}(k)$ for all $k \in \{1, \dots, a^* - 1\}$ and $n_t(k+1) = n_{t-k}(k+1)$ for all $k \in \{0, \dots, a^* - 2\}$. Therefore, we have for all $a \in \{1, \dots, a^* - 1\}$ that

$$\begin{aligned} V_t(a) &= \frac{n_t(a)}{n_t}, \\ &= \Pi_{k=0}^{a-1} \frac{n_t(k+1)}{n_t(k)}, \\ &= \Pi_{k=0}^{a-1} \frac{n_{t-k}(k+1)}{n_{t-k}(k)}, \\ &= \Pi_{k=0}^{a-1}(1 - \mu_k^t). \end{aligned}$$

We then prove Eq (7). As $S_t$ is stationary, we have $n_t(a) = n_{t-a}(a)$ for all $a \in \{1, \dots, a^* - 1\}$. Recalling Eq (6) and $V_t(a) = \frac{n_t(a)}{n_t}$, we can successively write

$$\begin{aligned} LE_t &= \sum_{a=0}^{a^*-1} \Pi_{k=0}^{a-1}(1 - \mu_k^t), \\ &= \sum_{a=0}^{a^*-1} V_t(a), \\ &= \frac{\sum_{a=0}^{a^*-1} n_t(a)}{n_t}, \\ &= \frac{\sum_{a=0}^{a^*-1} n_{t-a}(a)}{n_t}, \\ &= N_t / n_t. \end{aligned}$$

[28] Clearly, a constant distribution of lives is not sufficient for these equalities, one also needs a constant size of cohorts.

Finally, we prove Eq. (8). As $S_t$ is stationary, we have $p_t(a) = p_{t-a}(a)$ for all $a \in \{1, \ldots, a^* - 1\}$. Given that $\pi_t(a) = \frac{p_t(a)}{n_t(a)}$ and $V_t(a) = \frac{n_t(a)}{n_t}$, we can successively write

$$\begin{aligned} H_t &= \frac{\sum_{a=0}^{a^*-1} p_{t-a}(a)}{\sum_{a=0}^{a^*-1} n_{t-a}(a)}, \\ &= \frac{\sum_{a=0}^{a^*-1} p_t(a)}{N_t}, \\ &= \frac{\sum_{a=0}^{a^*-1} \pi_t(a) V_t(a) n_t}{N_t}. \end{aligned}$$

□

We use Lemma 1 to prove Proposition 1.

The assumption that individuals only enjoy binary consumption statuses implies that $\mathbb{E}u(c_a) = \pi(a)u_P + (1 - \pi(a))u_{NP}$ where $u_{NP} = u(NP)$ and $u_P = u(P)$. By Eq. (6), life expectancy at birth can be written as $LE = \sum_{a=0}^{a^*-1} V(a)$. We can thus rewrite Eq. (4) as

$$EU = u_{NP} LE - (u_{NP} - u_P) \sum_{a=0}^{a^*-1} V(a)\pi(a). \tag{9}$$

The result follows directly when substituting Eq. (7) and (8) into Eq. (9).

# C Calibrating values for $\theta$

Proposition 1 shows that $\theta = \frac{u_{NP} - u_P}{u_{NP}}$, where $u_P$ and $u_{NP}$ respectively denote the Bernouilli utility of being poor and being non-poor (and the utility of being dead is normalized to zero). Consider the constant elasticity of substitution Bernouilli utility function defined as

$$u(c) = \frac{c^{1-\epsilon} - \hat{c}^{1-\epsilon}}{1 - \epsilon}, \tag{10}$$

where $\hat{c}$ denotes the subsistence consumption, for which $u(\hat{c}) = 0$, and $\epsilon$ is the coefficient of relative risk aversion that captures the curvature of utility function $u$. A parametric value for $\theta$ requires defining representative consumption for the (consumption) poor and non-poor statuses such that $u_P = u(c_{poor})$ and $u_{NP} = u(c_{non-poor})$. Typically, $c_{poor}$ and $c_{non-poor}$ could respectively be defined as mean or median consumption among the poor and non-poor.

Parametric values for $\theta$ are sensitive to the values selected for the parameters $\hat{c}$ and $\epsilon$. We illustrate this by providing values for $1/\theta$ for India in 2019 for different values of these parameters. We define poverty using the the International Poverty Line, whose value is $ 2.15 per person per day (2017 PPPs). We assume that $c_{poor}$ and $c_{non-poor}$ are defined as mean consumption among the poor and non-poor, which we extract from the Poverty and Inequality Platform from the World Bank.

# D Proof of Proposition 2

The proof builds on the framework presented in Appendix A.

**Table 3:** Parametric values for $1/\theta$ for India in 2019 using the International Poverty Line.

| | | $\hat{c}$ 0.5 ($ a day) | $\hat{c}$ 0.75 ($ a day) | $\hat{c}$ 1.0 ($ a day) | $\hat{c}$ 1.25 ($ a day) |
|---|---|---|---|---|---|
| $\epsilon$ | 2.5 | 7.8 | 4.1 | 2.6 | 1.8 |
| $\epsilon$ | 2.0 | 4.7 | 3.0 | 2.1 | 1.6 |
| $\epsilon$ | 1.5 | 3.0 | 2.2 | 1.7 | 1.4 |
| $\epsilon$ | 1.0 | 2.1 | 1.7 | 1.5 | 1.3 |

Note: According to the Poverty and Inequality Platform, mean consumption in India in 2019 was $ 5.13 per person per day. For the International Poverty Line, mean consumption among the poor is $ 1.75 and mean consumption among the non-poor is $ 5.51. The utility function considered is CES. $1/\theta$ can be interpreted as the number of years spent in poverty yielding the same well-being loss as one year of life lost.

By Proposition 2 in Baland et al. (2021) we have for any stationary society $S_t$ that $GD_{\theta\hat{a}}(S_t) = ID_{\theta\hat{a}}(S_t)$. Hence, we only need to prove that $GD_{\theta\hat{a}}(S_t) = ED_{\theta\hat{a}}(S_t)$ for any stationary society $S_t$. By definition, we have that $N_t = \#P + \#NP$, $N_t H_t = \#P$ and $n_{t-a}(a) = N_a$, i.e.,

$$GD_{\theta\hat{a}} = \frac{YLL_t}{N_t + YLL_t} + \theta \frac{N_t H_t}{N_t + YLL_t},$$

where

$$YLL_t = \sum_{a=0}^{\hat{a}-2} n_{t-a}(a) * \mu_a^t * (\hat{a} - (a+1)).$$

As society $S_t$ is stationary, Lemma 1 applies and $N_t = n_t LE_t$ (Eq. (7)). Substituting this expression for $N_t$ into the definition of $GD_{\theta\hat{a}}$ proves our result, provided $YLL_t = n_t LGE_{\hat{a}}$, which remains to be shown. As society $S_t$ is stationary, Lemma 1 applies and we have $\frac{n_{t-a}(a)}{n_t} = \prod_{k=0}^{a-1}(1-\mu_k^t)$ (Eq. (6)). Substituting this expression for $n_{t-a}(a)$ into the definition of $YLL_t$ gives:

$$YLL_t = n_t \sum_{a=0}^{\hat{a}-2} (\hat{a} - (a+1)) * \mu_a^t * \prod_{k=0}^{a-1}(1-\mu_k^t),$$

which shows that $YL_t = n_t LGE_{\hat{a}}$ (recall that $\hat{a} - (a+1) = 0$ when $a = \hat{a} - 1$), the desired result.

# E $ED_{\theta\hat{a}}$ and $GD_{\theta\hat{a}}$ under a transitory shock

We illustrate the difference between $ED_{\theta\hat{a}}$ and $GD_{\theta\hat{a}}$ in their reaction to a transitory mortality shock with the help of a simple example. Consider a population with a fixed natality $n_t(0) = 2$ for all periods $t$. At each period, all alive individuals are non-poor, implying that $H_t = 0$. For all $t < 0$, we assume a constant mortality vector $\mu^t = \mu^* = (0, 1, 1, 1)$, so that each individual lives exactly two periods. Let us assume $\hat{a} = 4$, so that an individual dies prematurely if she dies before her fourth period of life. Before period $t = 0$, the population pyramid is stationary, and the two indices are equal to 1/2 because there is no poor and individuals live for two periods instead of four. Consider now a permanent shock starting from period 0 onwards, such that half of the newborns die after their first period of life: $\mu^0 = (1/2, 1, 1, 1)$. The population pyramid returns to its stationary state in period 1, after a (mechanical) transition in period 0. This example is illustrated in Figure 10.

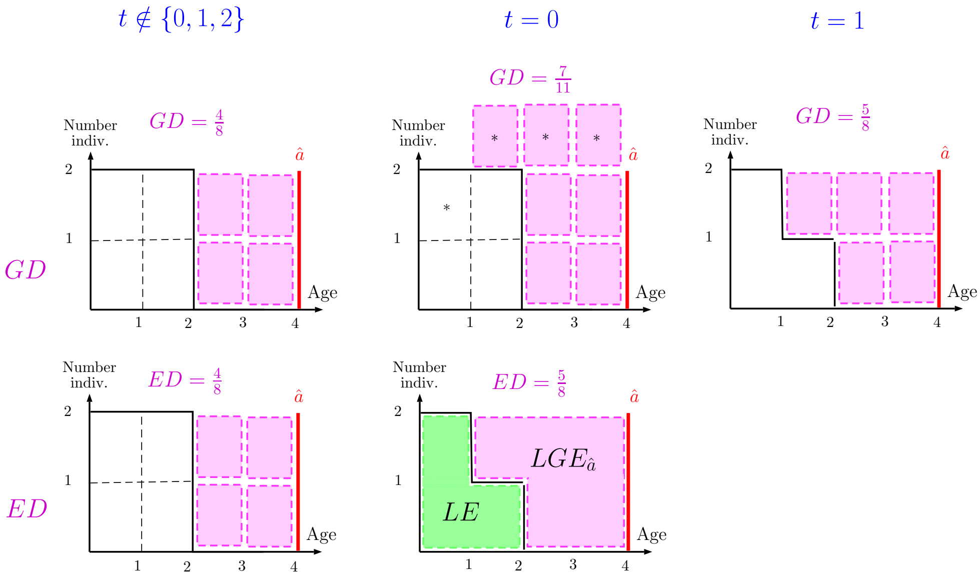



**Figure 10:** Response of $GD_{\theta\hat{a}}$ and $ED_{\theta\hat{a}}$ to a permanent mortality shock in $t = 0$. The years prematurely lost are shaded.

Consider first $GD_{\theta\hat{a}}$. In period 0, the actual population pyramid is not stationary because of the mortality shock. The premature death of one newborn leads to the loss of three years of life. Also, two one-year old individuals die in period 0, each losing two years of life. There are thus 7 years of life prematurely lost in period 0, and $GD_{\theta\hat{a}}$ takes value 7/11. In period 1, the population pyramid is stationary, and $GD_{\theta\hat{a}}$ is equal to 5/8 from then on.

We now turn to $ED_{\theta\hat{a}}$. Even if the actual population pyramid is not stationary in period 0, $ED_{\theta\hat{a}}$ is immediately equal to 5/8 since it records premature mortality as if the population pyramid had already reached its new stationary level. $ED_{\theta\hat{a}}$ focuses on the newborn and the one-year old who die prematurely, ignoring that there are two one-year old dying in the actual population pyramid in period 0 (which is a legacy of the past).

# F Characterization of the $ED_{\theta\hat{a}}$ index

We first introduce the set-up provided by Baland et al. (2021), which we will use to charcterize $ED_{\theta\hat{a}}$.

Each individual $i$ is associated to a birth year $b_i \in \mathbb{Z}$. In period $t$, each individual $i$ with $b_i \leq t$ is characterized by a **bundle** $x_i = (a_i, s_i)$, where $a_i = t - b_i$ is the age that individual $i$ would have in period $t$ given her birth year $b_i$, and $s_i$ is a categorical variable capturing individual status in period $t$, which can be either alive and non-poor ($NP$), alive and poor ($AP$) or dead ($D$), i.e. $s_i \in S = \{NP, AP, D\}$. In the following, we often refer to individuals whose status is $AP$ as "poor". We consider here that births occur at the beginning while deaths occur at the end of a period. As a result, an individual whose status in period $t$ is $D$ died before period $t$.[29]

[29] All newborns have age 0 during period $t$ and some among these newborns may die at the end

An individual "dies prematurely" if she dies before reaching the minimal lifespan $\hat{a} \in \mathbb{N}$. Formally, period $t$ is "prematurely lost" by any individual $i$ with $s_i = D$ and $a_i < \hat{a}$. A **distribution** $x = (x_1, \ldots, x_{n(x)})$ specifies the age and the status in period $t$ of all $n(x)$ individuals. Excluding trivial distributions for which no individual is alive or prematurely dead, the set of distributions in period $t$ is given by:

$$X = \{x \in \cup_{n \in \mathbb{N}} (\mathbb{Z} \times S)^n \mid \text{there is } i \text{ for whom either } s_i \neq D \text{ or } s_i = D \text{ and } \hat{a} > t - b_i\}.$$

Baland et al. (2021) show that the most natural *consistent* index to rank distributions in $X$ is the inherited deprivation index ($ID_{\theta\hat{a}}$). Let $d(x)$ denote the number of *prematurely dead* individuals in distribution $x$, which is the number of individuals $i$ for whom $s_i = D$ and $\hat{a} > t - b_i$, $p(x)$ the number of individuals who are poor and $f(x)$ the number of alive and non-poor individuals. The $ID_{\theta\hat{a}}$ index is defined as:

$$ID_{\theta\hat{a}}(x) = \underbrace{\frac{d(x)}{f(x) + p(x) + d(x)}}_{quantity\ deprivation} + \theta \underbrace{\frac{p(x)}{f(x) + p(x) + d(x)}}_{quality\ deprivation}, \tag{11}$$

where $\theta \in [0, 1]$ is a parameter weighing the relative importance of alive deprivation and lifespan deprivation. An individual losing prematurely period $t$ matters $1/\theta$ times as much as an individual spending period $t$ in alive deprivation.

We introduce additional notation for the mortality taking place in period $t$. Consider the population pyramid in period $t$, and let $n_a(x)$ be the number of *alive* individuals of age $a$ in distribution $x$, i.e. the number of individuals $i$ for whom $a_i = a$ and $s_i \neq D$. (The definition of $n_a(x)$ corresponds to $n_{t-a}(a)$ in the notation used in the main text of the paper. In this section, we adopt the notation of Baland et al. (2021), which does not require to mention period $t$.) The age-specific mortality rate $\mu_a \in [0, 1]$ denotes the fraction of alive individuals of age $a$ dying at the end of period $t$: the number of $a$-year-old individuals dying at the end of period $t$ is $n_a(x) * \mu_a$. Letting $a^* \in \mathbb{N}$ stand for the maximal lifespan (which implies $\mu_{a^*-1} = 1$), the **vector of age-specific mortality rates** in period $t$ is given by $\mu = (\mu_0, \ldots, \mu_{a^*-1})$. Vector $\mu$ summarizes mortality in period $t$, while distribution $x$ summarizes alive deprivation in period $t$ as well as mortality before period $t$. The set of mortality vectors is defined as:

$$M = \left\{ \mu \in [0, 1]^{a^*} \middle| \mu_{a^*-1} = 1 \right\}.$$

We consider pairs $(x, \mu)$ for which the distribution $x$ is a priori unrelated to vector $\mu$. We assume that the age-specific mortality rates $\mu_a$ must be feasible given the number of alive individuals $n_a(x)$. Given that distributions have finite numbers of individuals, mortality rates cannot take irrational values, i.e. $\mu_a \in [0, 1] \cap \mathbb{Q}$, where $\mathbb{Q}$ is the set of rational numbers. The set of pairs considered is given by:

$$O = \left\{ (x, \mu) \in X \times M \middle| \text{for all } a \in \{0, \ldots, a^*\} \text{ we have } \mu_a = \frac{c_a}{n_a(x)} \text{ for some } c_a \in \mathbb{N} \right\}.$$

Letting $d_a(x)$ be the number of *dead* individuals born $a$ years before $t$ in distribution $x$, the total number of individuals born $a$ years before $t$ is then equal to

of period $t$. This implies that $b_i = t \Rightarrow s_i \neq D$.

$n_a(x) + d_a(x)$. Formally, the **pair** $(x, \mu)$ is **stationary** if, for some $n^* \in \mathbb{N}$ and all $a \in \{0, \dots, a^*\}$, we have:

- $n_a(x) + d_a(x) = n^* \in \mathbb{N}$ (constant natality),
- $n_{a+1}(x) = n_a(x) * (1 - \mu_a)$ (identical population pyramid in $t+1$).

In a stationary pair, the population pyramid is such that the size of each cohort can be obtained by applying to the preceding cohort the current mortality rate. The pair associated to a stationary society (as defined in the main text) is stationary. An index is a function $P : O \times \mathbb{N} \to \mathbb{R}_+$. We simplify the notation $P(x, \mu, \hat{a})$ to $P(x, \mu)$ as a fixed value for $\hat{a}$ is assumed.

We now introduce the properties characterizing $ED_{\theta\hat{a}}$. $ID_{\theta\hat{a}}$ Equivalence requires that, as the current mortality (in period $t$) is the same as the mortality prevailing in the previous periods in stationary societies, any index defined on current mortality rates is equivalent to $ID_{\theta\hat{a}}$ in the case of a stationary pair:[30]

**Deprivation axiom 1** ($ID_{\theta\hat{a}}$ Equivalence). *There exists some $\theta \in (0, 1]$ and $\hat{a} \geq \underline{\hat{a}}$ such that for all $(x, \mu) \in O$ that are stationary we have $P(x, \mu) = ID_{\theta\hat{a}}(x)$.*

Independence of Dead requires that past mortality does not affect the index. More precisely, the presence of an additional dead individual in distribution $x$ does not affect the index:

**Deprivation axiom 2** (Independence of Dead). *For all $(x, \mu) \in O$ and $i \leq n(x)$, if $s_i = D$, then $P((x_i, x_{-i}), \mu) = P(x_{-i}, \mu)$.*

Independence of Birth Year requires that the index does not depend on the birth year of individuals, i.e. only their status matters. As Independence of Dead requires to disregard dead individuals, the only relevant information in $x$ is whether an alive individual is poor or not.

**Deprivation axiom 3** (Independence of Birth Year). *For all $(x, \mu) \in O$ and $i \leq n(x)$, if $s_i = s'_i$, then $P((x_i, x_{-i}), \mu) = P((x'_i, x_{-i}), \mu)$.*

Replication Invariance requires that, if a distribution is obtained by replicating another distribution several times, they both have the same deprivation when associated to the same mortality vector. By definition, a $k$-replication of distribution $x$ is a distribution $x^k = (x, \dots, x)$ for which $x$ is repeated $k$ times.

**Deprivation axiom 4** (Replication Invariance). *For all $(x, \mu) \in O$ and $k \in \mathbb{N}$, $P(x^k, \mu) = P(x, \mu)$.*

Proposition 6 shows that these properties jointly characterize the $ED_{\theta\hat{a}}$ index.

**Proposition 6** (Characterization of $ED_{\theta\hat{a}}$).
*$P = ED_{\theta\hat{a}}$ if and only if $P$ satisfies Independence of Dead, $ID_{\theta\hat{a}}$ Equivalence, Replication Invariance and Independence of Birth Year.*

[30] Recall that past mortality is recorded in distribution $x$ while current mortality is recorded in vector $\mu$. As vector $\mu$ is redundant in stationary pairs, in the sense that $\mu$ can be inferred from the population pyramid, the index can be computed on distribution $x$ only. See Baland et al. (2021) for a complete motivation for this axiom.

*Proof.* We first prove sufficiency. Proving that the $ED_{\theta\hat{a}}$ index satisfies Independence of Dead, Replication Invariance and Independence of Birth Year is straightforward and left to the reader. Finally, $ED_{\theta\hat{a}}$ index satisfies $ID_{\theta\hat{a}}$ Equivalence because $ED_{\theta\hat{a}}$ is equal to $GD_{\theta\hat{a}}$ in stationary populations (Proposition 2) and $GD_{\theta\hat{a}}$ satisfies $ID_{\theta\hat{a}}$ Equivalence (Proposition 2 in Baland et al. (2021)). (The pairs associated to stationary societies are stationary).

We now prove necessity. Take any pair $(x, \mu) \in O$. We construct another pair $(x''', \mu)$ that is stationary and such that $P(x''', \mu) = P(x, \mu)$ and $ED_{\theta\hat{a}}(x''', \mu) = ED_{\theta\hat{a}}(x, \mu)$. Given that $(x''', \mu)$ is stationary, we have by $ID_{\theta\hat{a}}$ Equivalence that $P(x''', \mu) = ID_{\theta\hat{a}}(x''', \mu)$ for some $\theta \in (0, 1]$. As $ID_{\theta\hat{a}} = GD_{\theta\hat{a}} = ED_{\theta\hat{a}}$ for stationary pairs, we have $P(x''', \mu) = ED_{\theta\hat{a}}(x''', \mu)$ for some $\theta \in (0, 1]$. If we can construct such pair $(x''', \mu)$, then $P(x, \mu) = ED_{\theta\hat{a}}(x, \mu)$ for some $\theta \in (0, 1]$, the desired result.

We turn to the construction of the stationary pair $(x''', \mu)$, using two intermediary pairs $(x', \mu)$ and $(x'', \mu)$. One difficulty is to ensure that the mortality rates $\mu_a$ can be achieved in the stationary population given the number of alive individuals $n_a(x''')$, that is $\mu_a = \frac{c}{n_a(x''')}$ for some $c \in \mathbb{N}$.

We first construct a $n'-$replication of $x$ that has sufficiently many alive individuals to meet this constraint. For any $a \in \{0, \ldots, a^* - 1\}$, take any naturals $c_a$ and $e_a$ such that $\mu_a = \frac{c_a}{e_a}$. Let $e = \prod_{j=0}^{a^*-1} e_j$, $n'_a = e \prod_{j=0}^{a-1}(1 - \frac{c_j}{e_j})$ and $n' = \sum_{j=0}^{a^*-1} n'_j$.[31] Let $x'$ be a $n'-$replication of $x$. Letting $n^x = \sum_{j=0}^{a^*-1} n_j(x)$ be the number of alive individuals in distribution $x$, we have that $x'$ has $n' * n^x$ alive individuals. We have $P(x', \mu) = P(x, \mu)$ by Replication Invariance.

We define $x''$ from $x'$ by changing the birth years of *alive* individuals in such a way that $(x'', \mu)$ has a population pyramid that is stationary. Formally, we construct $x''$ with $n(x'') = n(x')$ such that

- dead individuals in $x'$ are also dead in $x''$,
- alive individuals in $x'$ are also alive in $x''$ and have the same status,
- the birth year of alive individuals are changed such that, for each $a \in \{0, \ldots, a^* - 1\}$, the number of $a$-years old individuals is $n' * n^x * \frac{\prod_{j=0}^{a-1}(1-\frac{c_j}{e_j})}{\sum_{k=0}^{a^*-1}\prod_{j=0}^{k-1}(1-\frac{c_j}{e_j})}$.[32]

One can check that $(x'', \mu)$ has a population pyramid corresponding to a stationary population and that each age group has a number of alive individuals in $\mathbb{N}$. We have $P(x'', \mu) = P(x', \mu)$ by Independence of Birth Year.

Define $x'''$ from $x''$ by changing the number and birth years of *dead* individuals in such a way that $(x''', \mu)$ is stationary. To do so, place exactly $n_0(x'') - n_a(x'')$ dead individuals in each age group $a$. We have $P(x''', \mu) = P(x'', \mu)$ by Independence of Dead.

Together, we have that $P(x''', \mu) = P(x, \mu)$. Finally, by construction we have $H(x''') = H(x)$, which implies that $ED_{\theta\hat{a}}(x''', \mu) = ED_{\theta\hat{a}}(x, \mu)$. □

[31] These numbers imply that a constant natality of $e$ newborns leads to a stationary population of $n'$ alive individuals.

[32] Observe that $\sum_{k=0}^{a^*-1}\prod_{j=0}^{k-1}(1 - \frac{c_j}{e_j}) = LE$, implying that $e = \frac{n' * n^x}{\sum_{k=0}^{a^*-1}\prod_{j=0}^{k-1}(1-\frac{c_j}{e_j})}$.

# G Proof of Proposition 3

The proof builds on the framework presented in Appendix A.

We first show that $LE + LGE_{\hat{a}} = \hat{a}$ when $\hat{a} \geq a^*$. By definition, $LE$ and $LGE_{\hat{a}}$ only depend on the age-specific mortality vector $\mu^t$. Thus, the values for $LE$ and $LGE_{\hat{a}}$ do not depend on whether the society is stationary or not. Consider any stationary society $S_t$ whose constant mortality vector is $\mu^t$. We show for this stationary society $S_t$ that $LE + LGE_{\hat{a}} = \hat{a}$ when $\hat{a} \geq a^*$.

As society $S_t$ is stationary, Lemma 1 applies and we have $N_t = n_t * LE_t$ (Eq. (7)). As by definition $N_t = \sum_{a=0}^{a^*-1} n_t(a)$, we get

$$LE = \sum_{a=0}^{a^*-1} \frac{n_t(a)}{n_t}. \tag{12}$$

As society $S_t$ is stationary, Lemma 1 applies and we have $V_t(a) = \Pi_{k=0}^{a-1}(1-\mu_k^t)$ (Eq. (6)). Using the definition of age-specific mortality rate, namely $\mu_a^t = \frac{n_{t-a}(a)-n_{t-a}(a+1)}{n_{t-a}(a)}$, we can rewrite $LGE_{\hat{a}}$ as

$$LGE_{\hat{a}}(S_t) = \sum_{a=0}^{\hat{a}-1}(\hat{a}-(a+1)) * \frac{n_{t-a}(a)-n_{t-a}(a+1)}{n_{t-a}(a)} * V_t(a).$$

As society $S_t$ is stationary, we have that $n_t(a) = n_{t-a}(a)$ and $n_t(a+1) = n_{t-a}(a+1)$ for all $a \in \{0, \dots, a^*-1\}$. We can thus successively write

$$\begin{aligned}
LGE_{\hat{a}} &= \sum_{a=0}^{\hat{a}-1}(\hat{a}-(a+1)) * \frac{n_t(a)-n_t(a+1)}{n_t(a)} * \frac{n_t(a)}{n_t}, \\
&= \sum_{a=0}^{\hat{a}-1} \hat{a} * \frac{n_t(a)-n_t(a+1)}{n_t} - \sum_{a=0}^{\hat{a}-1}(a+1) * \frac{n_t(a)-n_t(a+1)}{n_t}, \\
&= \frac{1}{n_t}\left(\hat{a} * (n_t(0)-n_t(\hat{a})) - \sum_{a=0}^{\hat{a}-1} n_t(a) + \hat{a} * n_t(\hat{a})\right), \\
&= \hat{a} - \sum_{a=0}^{\hat{a}-1} \frac{n_t(a)}{n_t}.
\end{aligned}$$

By definition of $a^*$, we have $n_t(a) = 0$ for all $a \geq a^*$. When $\hat{a} \geq a^*$, this implies that $\sum_{a=0}^{\hat{a}-1} \frac{n_t(a)}{n_t} = \sum_{a=0}^{a^*-1} \frac{n_t(a)}{n_t}$. We have shown that $LGE_{\hat{a}} = \hat{a} - \sum_{a=0}^{a^*-1} \frac{n_t(a)}{n_t}$, which together with Eq. (12) proves that $LE + LGE_{\hat{a}} = \hat{a}$ when $\hat{a} \geq a^*$.

The fact that $LE + LGE_{\hat{a}} = \hat{a}$ implies that $PALE_\theta = \hat{a}(1 - ED_{\theta\hat{a}})$ because

$$\begin{aligned}
\hat{a}(1-ED_{\theta\hat{a}}) &= (LE + LGE_{\hat{a}})(1-ED_{\theta\hat{a}}) \\
&= LE(1-\theta H), \\
&= PALE_\theta.
\end{aligned}$$

Thus, when $\hat{a} \geq a^*$, $PALE_\theta$ is a linear function of $ED_{\theta\hat{a}}$ that depends negatively on $ED_{\theta\hat{a}}$. Therefore, these two indicators yields opposite ranking of any two societies $A$ and $B$, i.e. $PALE_\theta(A) \geq PALE_\theta(B) \Leftrightarrow ED_{\theta\hat{a}}(A) \leq ED_{\theta\hat{a}}(B)$.

## H Mortality shocks and the evolution of $ED_{\theta\hat{a}}$ and $PALE_\theta$

We briefly contrast the impact of mortality shocks on $PALE_\theta$ and $ED_{\theta\hat{a}}$, assuming that these mortality shocks are independent of the poverty status. Consider a mortality shock that equalizes individual lifespans across the age threshold $\hat{a}$ while keeping life expectancy $LE$ constant. This lower dispersion in mortality does not affect $PALE_\theta$, which only accounts for mortality through $LE$. By contrast, this shock reduces $ED_{\theta\hat{a}}$, since $LGE_{\hat{a}}$ is thereby reduced. It is indeed easy to check that $\frac{\partial ED_{\theta\hat{a}}}{\partial LGE_{\hat{a}}} > 0$ (for $\theta H < 1$).

Consider instead a mortality shock that reduces mortality above the age threshold $\hat{a}$. Such shock increases $LE$ but does not affect $LGE_{\hat{a}}$. As a result, $PALE_\theta$ mechanically increases. It is also easy to show that deprivation, as measured by $ED_{\theta\hat{a}}$, decreases: $\frac{\partial ED_{\theta\hat{a}}}{\partial LE} < 0$, for $\theta H < 1$. Moreover, $PALE_\theta$ is more sensitive to this kind of shock than $ED_{\theta\hat{a}}$, as the elasticity of $PALE_\theta$ to $LE$ is equal to 1 while the elasticity of $ED_{\theta\hat{a}}$ to $LE$ lies in $(-1, 0)$. If the mortality shock is such that it reduces mortality below the age threshold $\hat{a}$, this shock simultaneously increases $LE$ and reduces $LGE_{\hat{a}}$. Again, $PALE_\theta$ improves and $ED_{\theta\hat{a}}$ decreases since both $LE$ increases and $LGE$ decreases.

## I Proof of Proposition 4

The proof builds on the framework presented in Appendix A.

By Proposition 2, $M_{\theta\hat{a}}$ is Paradox-free if and only if $ID_{\theta\hat{a}}$ is Paradox-free.

First, we prove that $ID_{\theta\hat{a}}$ is Paradox-free only if $\theta \leq 1$ and $\hat{a} \geq a^*$. The proof is by contradiction.

Assume first that $\theta > 1$ and $\hat{a} \geq 2$. Consider two alternative stationary societies A and B that both feature only one newborn $i$ every year. The life of $i$ is respectively $l_i^A = (P, D)$ and $l_i^B = (P, P, D)$. Society B is obtained from A by a lifespan increment to the poor person $i$. However, we have $ID_{\theta\hat{a}}(A) = \frac{(\hat{a}-1)+\theta}{\hat{a}}$ and $ID_{\theta\hat{a}}(B) = \frac{(\hat{a}-2)+2\theta}{\hat{a}}$, which yields $ID_{\theta\hat{a}}(A) < ID_{\theta\hat{a}}(B)$, which shows that $ID_{\theta\hat{a}}$ is not Paradox-free.

Assume then that $\theta > 0$ and $\hat{a} < a^*$. Consider two alternative stationary societies A' and B' that both feature two newborns $i$ and $j$ every year. Their lives are respectively $l_i^{A'} = (P, \ldots, P, D)$ and $l_i^{B'} = (P, \ldots, P, P, D)$, where $i$'s lifespan is $\hat{a}$ years in society A' and $\hat{a}+1$ years in society B', while $l_j^{A'} = l_j^{B'} = (NP, \ldots, NP, D)$, where $j$'s lifespan is $\hat{a}$ years in both societies. Society B' is obtained from A' by a lifespan increment to the poor person $i$. However, we have $ID_{\theta\hat{a}}(A') = \frac{\theta(\hat{a})}{\hat{a}+\hat{a}}$ and $ID_{\theta\hat{a}}(B') = \frac{\theta(\hat{a}+1)}{\hat{a}+\hat{a}+1}$, which yields $ID_{\theta\hat{a}}(A') < ID_{\theta\hat{a}}(B')$, which shows that $ID_{\theta\hat{a}}$ is not Paradox-free.

We have thus proven that $ID_{\theta\hat{a}}$ is Paradox-free only if $\theta \leq 1$ and $\hat{a} \geq a^*$.

Second, we prove that $ID_{\theta\hat{a}}$ is Paradox-free if $\theta \leq 1$ and $\hat{a} \geq a^*$. Take any two stationary societies A" and B" such that B" is obtained from A" by a lifespan increment to the poor person $i$. Poor person $i$ dies prematurely in society A" because $\hat{a} \geq a^*$. The difference between societies A" and B" is thus that $i$ spends an additional year in poverty in society B", instead of prematurely loosing that year in society A". Thus, the following two equalities hold $\#P(A'') + \#NP(A'') + \#PD(A'') =$

$\#P(B'') + \#NP(B'') + \#PD(B'')$ and $\#PD(A'') - \#PD(B'') = \#P(B'') - P(A'') = 1$. The former equality implies that $ID_{\theta\hat{a}}(A'') \geq ID_{\theta\hat{a}}(B'')$ if and only if $\#PD(A'') + \theta\#P(A'') \geq \#PD(B'') + \theta\#P(B'')$. This inequality is equivalent to $\#PD(A'') - \#PD(B'') \geq \theta(\#P(B'') - \#P(A''))$, which further simplifies to $1 \geq \theta$ given the latter equality, which proves that $ID_{\theta\hat{a}}$ is Paradox-free.

# J Proof of Proposition 5

**Proof of (i)**. We start by the "only if" part. Assume to the contrary that $PALE_0(A) > PALE_0(B)$ or $PALE_1(A) > PALE_1(B)$. This directly implies that $PALE_\theta(A) > PALE_\theta(B)$ for some $\theta \in (0, 1]$ and therefore we cannot have $PALE_\theta(A) \leq PALE_\theta(B)$ for all $\theta \in (0, 1]$.

We now turn to the "if" part. By definition of the $\text{PALE}_\theta$ index, we have to show that

$$LE(B) - LE(A) \geq \theta * (LE(B)H(B) - LE(A)H(A)), \tag{13}$$

for all $\theta \in (0, 1]$. As $PALE_0(A) \leq PALE_0(B)$, we directly have that $LE(B) - LE(A) \geq 0$ because $PALE_0 = LE$. As $PALE_1(A) \leq PALE_1(B)$, we have $LE(B) - LE(A) \geq LE(B)H(B) - LE(A)H(A)$. It immediately follows that inequality (13) is verified for all $\theta \in (0, 1]$.

**Proof of (ii)**. From (i), proving (ii) only requires providing societies A and B with $H(A) < H(B)$ such that $PALE_0(A) \leq PALE_0(B)$ and $PALE_1(A) \leq PALE_1(B)$. If $H(A) = 0.2$, $H(B) = 0.4$, $LE(A) = 50$ and $LE(B) = 75$ we have $PALE_1(A) = 40$ and $PALE_1(A) = 45$, the desired result because $PALE_0 = LE$.

# K Robust $ED_{\theta\hat{a}}$ comparisons

$ED_{\theta\hat{a}}$ is Paradox-free when $\theta \in (0, 1]$ and $\hat{a} \geq a^*$. However, $ED_{\theta\hat{a}}$ no longer encapsulates the minimal view when $\hat{a} \geq a^*$. To ease the impossibility between paradox-freeness and the minimal view, we define a weaker notion of paradox-freeness. A deprivation index is *minimally paradox free* when the index does not record a worsening for increments to the lifespan of a poor person *who dies prematurely*, i.e., whose lifespan is smaller than $\hat{a}$. In that case, $ED_{\theta\hat{a}}$ is minimally paradox free when $\theta \in (0, 1]$.

We assume that the age threshold $\hat{a} \in \mathbb{N}_0$ must respect a lower-bound $\underline{\hat{a}} \in \mathbb{N}_0$, such that $\hat{a} \geq \underline{\hat{a}} \geq 0$. Clearly, the value for the lower bound $\underline{\hat{a}}$ influences the set of comparisons that are robust to the values selected for $\theta$ and $\hat{a}$. Proposition 7 provides the conditions under which the ranking by $ED_{\theta\hat{a}}$ is robust for all $\theta \in (0, 1]$ and all $\hat{a} \geq \underline{\hat{a}}$.

**Proposition 7** (Robust comparisons with $ED_{\theta\hat{a}}$)**.**
*(i) For any two societies A and B we have $ED_{\theta\hat{a}}(A) \geq ED_{\theta\hat{a}}(B)$ for all $\theta \leq 1$ and*

*all $\hat{a} \geq \underline{\hat{a}}$ if and only if*

$$ED_{0\hat{a}}(A) \geq ED_{0\hat{a}}(B) \text{ for all } \hat{a} \geq \underline{\hat{a}}, \text{ and}$$
$$ED_{1\hat{a}}(A) \geq ED_{1\hat{a}}(B) \text{ for all } \hat{a} \geq \underline{\hat{a}} \qquad \text{(generalized Condition C1)}$$

*(ii) For any $\underline{\hat{a}} \geq 2$, there exist societies A and B for which $ED_{\theta\hat{a}}(A) \geq ED_{\theta\hat{a}}(B)$ for all $\theta \leq 1$ and all $\hat{a} \geq \underline{\hat{a}}$ even though $H(A) < H(B)$. These societies are such that $LE(A) < LE(B)$.*

*Proof.* See Appendix L for the straightforward proof. □

We illustrate Proposition 7 in Figure 11.[33] The vertical axis represents the share of pairs of societies for which H and LE provide identical (at the top) or opposite rankings (at the bottom). By definition, rankings by H and LE are insensitive to the age threshold $\hat{a}$ considered. The horizontal axis represents all possible values of $\underline{\hat{a}}$, the lower bound on the age threshold.

The left panel describes the share of pairs for which $ED_{\theta\hat{a}}$ provides robust rankings as a function of $\underline{\hat{a}}$. Lower values of $\underline{\hat{a}}$ imply a fall in the share of cases that $ED_{\theta\hat{a}}$ can rank robustly. Indeed, a larger age interval of values of $\hat{a}$ over which $ED_{\theta\hat{a}}$ has to be computed implies a larger number of comparisons for ED. As a result, the number of pairs for which it can provide the same ranking for all age thresholds falls.[34] Second, if H and LE provide the same ranking, $ED_{\theta\hat{a}}$ provides the same ranking as H when $\underline{\hat{a}} = a^*$. Finally, as discussed above, when H and LE disagree, a larger value of $\underline{\hat{a}}$ implies that the share of cases for which H provides an robustly wrong ranking gets larger.

The right panel reports, for all values of $\underline{\hat{a}}$, the share of pairs of societies for which $PALE_\theta$ and $ED_{\theta\hat{a}}$ provide robust rankings. Since $PALE_\theta$ does not depend on the age threshold, it is able to rank a larger set of comparisons. As shown in Proposition 3, when $\underline{\hat{a}} = a^*$, the two indices are equivalent.

# L Proof of Proposition 7

We first prove the following: for any $\hat{a} \geq \underline{\hat{a}}$ and any two societies A and B, we have $ED_{\theta\hat{a}}(A) \geq ED_{\theta\hat{a}}(B)$ for all $\theta \in (0, 1]$ if and only if

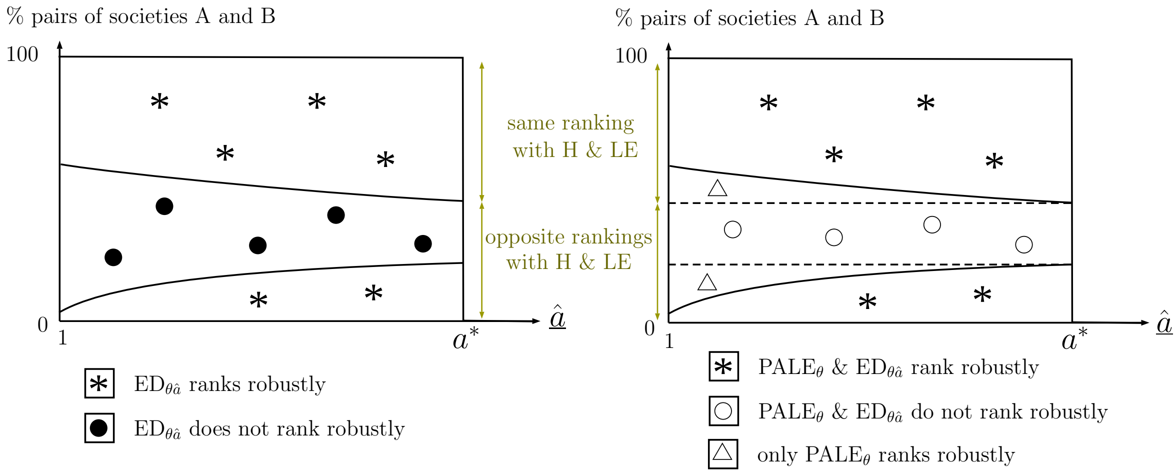
$$ED_{0\hat{a}}(A) \geq ED_{0\hat{a}}(B) \text{ and } ED_{1\hat{a}}(A) \geq ED_{1\hat{a}}(B).$$

We start with the "only if" part. Assume on the contrary that $ED_{0\hat{a}}(A) < ED_{0\hat{a}}(B)$ or $ED_{1\hat{a}}(A) < ED_{1\hat{a}}(B)$. This implies that $ED_{\theta\hat{a}}(A) < ED_{\theta\hat{a}}(B)$ for some $\theta \in (0, 1]$ and therefore we cannot have $ED_{\theta\hat{a}}(A) \geq ED_{\theta\hat{a}}(B)$ for all $\theta \in (0, 1]$.

[33] All graphs that follow are constructed using a lower bound on $\hat{a}$ equal to 1. Indeed, for $\theta = 0$ and $\hat{a} = 0$, $ED_{\theta\hat{a}}$ is equal to zero for all societies and cannot therefore deliver robust comparisons.

[34] It is not a sufficient condition that the rankings by H and LE are identical for the ranking by $ED_{\theta\hat{a}}$ to be robust. The reason is that, when $\underline{\hat{a}} < a^*$, $LE$ no longer contains all the relevant information on mortality: for instance, two societies can share the same life expectancy at birth but one with several deaths occurring below $\hat{a}$ while the other has all deaths occurring above $\hat{a}$.

**Figure 11:** Share of robust $ED_{\theta\hat{a}}$ comparisons as a function of $\underline{\hat{a}}$.




Reading: Left: The smaller the lower-bound $\underline{\hat{a}}$, the lower the share of societies pairs robustly ranked by $ED_{\theta\hat{a}}$. Right: The higher the lower-bound $\underline{\hat{a}}$, the higher the share of societies pairs robustly ranked by both $ED_{\theta\hat{a}}$ and $PALE_\theta$.

We turn to the "if" part. By definition of the $ED_{\theta\hat{a}}$ index, we have to show that

$$\frac{LGE_{\hat{a}}(A)}{LE(A)+LGE_{\hat{a}}(A)} - \frac{LGE_{\hat{a}}(B)}{LE(B)+LGE_{\hat{a}}(B)} \geq \theta\left(\frac{LE(B)*H(B)}{LE(B)+LGE_{\hat{a}}(B)} - \frac{LE(A)*H(A)}{LE(A)+LGE_{\hat{a}}(A)}\right) \text{ for all } \theta \in (0,1]. \tag{14}$$

As $ED_{0\hat{a}}(A) \geq ED_{0\hat{a}}(B)$, the left hand side of Eq. (14) is non-negative. As $ED_{1\hat{a}}(A) \geq ED_{1\hat{a}}(B)$, Eq. (14) holds for $\theta = 1$. As a result, inequality (14) holds for all $\theta \in (0,1]$.

**Proof of (i)**. This is an immediate implication of the statement proven above.

**Proof of (ii)**. Consider two societies A and B with $H(A) < H(B)$ for which the generalized condition C1 holds.

Society A is such that $H(A) = 0.4$ and all its individuals die in their first year of life, which implies that $LE(A) = 1$ and $LGE_{\hat{a}}(A) = \hat{a} - 1$. Therefore, society A is such that

- $ED_{0\hat{a}}(A) = \frac{\hat{a}-1}{\hat{a}}$ and $ED_{1\hat{a}}(A) = 1 - \frac{0.6}{\hat{a}}$ for all $\hat{a} \geq \underline{\hat{a}}$.

Society B is such that $H(B) = 0.5$ and all its individuals die at the maximal age $a^* - 1$, which implies that $LE(B) = a^*$ and

- $LGE_{\hat{a}}(B) = 0$ if $\hat{a} \in \{2, \ldots, a^*\}$,
- $LGE_{\hat{a}}(B) = \hat{a} - a^*$ if $\hat{a} > a^*$.

Therefore, society B is such that

- $ED_{0\hat{a}}(B) = 0$ and $ED_{1\hat{a}}(B) = 0.5$ for all $\hat{a} \in \{2, \ldots, a^*\}$,

- $ED_{0\hat{a}}(B) = \frac{\hat{a}-a^*}{\hat{a}}$ and $ED_{1\hat{a}}(B) = 1 - \frac{0.5a^*}{\hat{a}}$ for all $\hat{a} > a^*$.

By statement (i), we get $ED_{\theta\hat{a}}(A) \geq ED_{\theta\hat{a}}(B)$ for all $\theta \in (0, 1]$ and all $\hat{a} \geq \underline{\hat{a}}$ if we have $ED_{0\hat{a}}(A) \geq ED_{0\hat{a}}(B)$ and $ED_{1\hat{a}}(A) \geq ED_{1\hat{a}}(B)$ for all $\hat{a} \geq \underline{\hat{a}}$. Recalling that $\hat{a} \geq 2$, one can then easily check that we have $ED_{0\hat{a}}(A) \geq ED_{0\hat{a}}(B)$ and $ED_{1\hat{a}}(A) \geq ED_{1\hat{a}}(B)$ both for all $\hat{a} \in \{2, \ldots, a^*\}$ and for all $\hat{a} > a^*$.

# M Descriptive statistics

Table 4 lists all the countries present in our data set as well as the 2019 values of the main variables of interest.

**Table 4:** Countries used in the dataset and descriptive statistcs

| Country | H<br>% | LE<br>Years | $LGE_{70}$<br>Years | LE*H<br>Years | $PALE_1$<br>Years | $ED_{1,70}$<br>% |
|---|---|---|---|---|---|---|
| Albania | 0 | 77 | 3 | 0 | 77 | 4 |
| Algeria | 0 | 75 | 4 | 0 | 75 | 6 |
| Angola | 32 | 64 | 11 | 21 | 43 | 42 |
| Armenia | 1 | 75 | 4 | 1 | 74 | 6 |
| Azerbaijan | 0 | 70 | 6 | 0 | 70 | 8 |
| Bangladesh | 11 | 74 | 6 | 8 | 65 | 18 |
| Belarus | 0 | 73 | 5 | 0 | 73 | 6 |
| Belize | 18 | 73 | 6 | 13 | 60 | 24 |
| Benin | 19 | 63 | 12 | 12 | 51 | 32 |
| Bhutan | 0 | 72 | 6 | 0 | 72 | 8 |
| Bolivia | 2 | 71 | 6 | 1 | 70 | 10 |
| BosniaandHerzegovina | 0 | 76 | 3 | 0 | 76 | 4 |
| Botswana | 13 | 61 | 13 | 8 | 53 | 28 |
| Brazil | 5 | 75 | 5 | 4 | 71 | 12 |
| Bulgaria | 1 | 72 | 5 | 1 | 72 | 7 |
| BurkinaFaso | 31 | 61 | 14 | 19 | 42 | 44 |
| Burundi | 72 | 63 | 12 | 45 | 18 | 76 |
| CaboVerde | 3 | 73 | 5 | 2 | 71 | 9 |
| Cameroon | 23 | 62 | 13 | 14 | 48 | 36 |
| CentralAfricanRepublic | 65 | 51 | 21 | 33 | 18 | 75 |
| Chad | 31 | 59 | 15 | 18 | 41 | 45 |
| China | 0 | 77 | 3 | 0 | 77 | 4 |
| Colombia | 5 | 79 | 4 | 4 | 75 | 10 |
| Comoros | 18 | 68 | 9 | 12 | 56 | 27 |
| CostaRica | 1 | 79 | 3 | 1 | 78 | 5 |
| CotedIvoire | 11 | 63 | 12 | 7 | 56 | 25 |
| Djibouti | 18 | 66 | 10 | 12 | 54 | 29 |
| DominicanRepublic | 1 | 72 | 7 | 1 | 72 | 9 |
| Ecuador | 4 | 75 | 5 | 3 | 73 | 9 |
| ElSalvador | 1 | 75 | 6 | 1 | 74 | 8 |
| Eswatini | 34 | 57 | 16 | 19 | 38 | 48 |
| Ethiopia | 18 | 68 | 9 | 12 | 56 | 27 |

...continued from previous page

| Country | H % | LE Years | $LGE_{70}$ Years | LE*H Years | $PALE_1$ Years | $ED_{1,70}$ % |
|---|---|---|---|---|---|---|
| Fiji | 1 | 68 | 8 | 1 | 67 | 11 |
| Gabon | 2 | 67 | 9 | 2 | 65 | 14 |
| GambiaThe | 15 | 66 | 10 | 10 | 56 | 26 |
| Georgia | 5 | 72 | 5 | 4 | 69 | 11 |
| Ghana | 22 | 65 | 10 | 14 | 51 | 33 |
| Guatemala | 7 | 72 | 7 | 5 | 67 | 15 |
| Guinea | 13 | 60 | 14 | 8 | 52 | 30 |
| GuineaBissau | 21 | 60 | 14 | 13 | 47 | 36 |
| Guyana | 6 | 66 | 9 | 4 | 62 | 18 |
| Haiti | 25 | 63 | 12 | 16 | 47 | 37 |
| Honduras | 13 | 71 | 6 | 9 | 62 | 19 |
| India | 12 | 70 | 7 | 8 | 62 | 20 |
| Indonesia | 4 | 71 | 6 | 3 | 67 | 12 |
| Iraq | 0 | 72 | 5 | 0 | 72 | 7 |
| Jamaica | 1 | 75 | 5 | 1 | 74 | 7 |
| Jordan | 0 | 77 | 3 | 0 | 77 | 4 |
| Kazakhstan | 0 | 71 | 6 | 0 | 71 | 8 |
| Kenya | 32 | 66 | 10 | 21 | 45 | 41 |
| Kiribati | 2 | 60 | 13 | 1 | 59 | 19 |
| KyrgyzRepublic | 1 | 73 | 5 | 0 | 72 | 8 |
| Lebanon | 0 | 76 | 4 | 0 | 76 | 5 |
| Lesotho | 34 | 51 | 21 | 18 | 33 | 54 |
| Liberia | 31 | 65 | 11 | 20 | 45 | 40 |
| Madagascar | 79 | 65 | 11 | 51 | 14 | 82 |
| Malawi | 69 | 64 | 12 | 44 | 20 | 74 |
| Malaysia | 0 | 74 | 4 | 0 | 74 | 5 |
| Maldives | 0 | 78 | 3 | 0 | 78 | 4 |
| Mali | 15 | 61 | 15 | 9 | 52 | 31 |
| MarshallIslands | 1 | 65 | 10 | 1 | 64 | 14 |
| Mauritania | 5 | 70 | 8 | 4 | 66 | 14 |
| Mauritius | 0 | 74 | 5 | 0 | 74 | 6 |
| Mexico | 3 | 75 | 5 | 2 | 73 | 9 |
| Moldova | 0 | 73 | 5 | 0 | 73 | 7 |
| Mongolia | 1 | 67 | 8 | 0 | 67 | 11 |
| Montenegro | 3 | 75 | 3 | 2 | 73 | 7 |
| Morocco | 1 | 72 | 5 | 1 | 72 | 8 |
| Mozambique | 71 | 57 | 16 | 41 | 17 | 78 |
| Myanmar | 1 | 68 | 8 | 1 | 68 | 12 |
| Namibia | 17 | 64 | 11 | 11 | 53 | 29 |
| Nepal | 3 | 70 | 7 | 2 | 68 | 11 |
| Nicaragua | 4 | 74 | 4 | 3 | 71 | 9 |
| Niger | 50 | 61 | 14 | 31 | 31 | 59 |
| Nigeria | 31 | 63 | 13 | 20 | 44 | 43 |

...continued from previous page

| Country | H % | LE Years | $LGE_{70}$ Years | LE*H Years | $PALE_1$ Years | $ED_{1,70}$ % |
|---|---|---|---|---|---|---|
| NorthMacedonia | 3 | 74 | 4 | 2 | 72 | 7 |
| Pakistan | 5 | 65 | 11 | 3 | 62 | 19 |
| PapuaNewGuinea | 31 | 64 | 11 | 19 | 44 | 41 |
| Paraguay | 1 | 75 | 5 | 1 | 75 | 7 |
| Peru | 3 | 79 | 4 | 2 | 77 | 7 |
| Philippines | 5 | 71 | 7 | 3 | 68 | 13 |
| Romania | 2 | 75 | 4 | 2 | 73 | 7 |
| RussianFederation | 0 | 72 | 6 | 0 | 72 | 8 |
| Rwanda | 44 | 68 | 9 | 30 | 38 | 50 |
| Samoa | 1 | 70 | 7 | 1 | 69 | 10 |
| SaoTomeandPrincipe | 15 | 70 | 6 | 10 | 60 | 22 |
| Senegal | 9 | 67 | 9 | 6 | 61 | 20 |
| Serbia | 0 | 75 | 3 | 0 | 75 | 4 |
| SierraLeone | 25 | 61 | 14 | 15 | 46 | 39 |
| SolomonIslands | 25 | 58 | 14 | 14 | 44 | 39 |
| SouthAfrica | 21 | 64 | 12 | 13 | 51 | 33 |
| SriLanka | 1 | 76 | 4 | 1 | 76 | 6 |
| StLucia | 5 | 74 | 5 | 4 | 71 | 11 |
| Sudan | 23 | 69 | 8 | 16 | 53 | 31 |
| SyrianArabRepublic | 69 | 73 | 5 | 50 | 23 | 71 |
| TaiwanChina | 0 | 79 | 3 | 0 | 79 | 4 |
| Tajikistan | 4 | 69 | 7 | 3 | 66 | 12 |
| Tanzania | 43 | 66 | 10 | 29 | 38 | 51 |
| Thailand | 0 | 77 | 4 | 0 | 77 | 6 |
| Togo | 28 | 64 | 11 | 18 | 46 | 39 |
| Tonga | 1 | 72 | 6 | 1 | 71 | 9 |
| Tunisia | 0 | 77 | 3 | 0 | 77 | 4 |
| Turkmenistan | 1 | 70 | 7 | 1 | 69 | 10 |
| Tuvalu | 0 | 67 | 8 | 0 | 67 | 11 |
| Uganda | 42 | 65 | 11 | 27 | 38 | 50 |
| Ukraine | 0 | 69 | 7 | 0 | 69 | 10 |
| UnitedArabEmirates | 0 | 73 | 5 | 0 | 73 | 6 |
| Uzbekistan | 28 | 68 | 7 | 19 | 49 | 35 |
| Vanuatu | 9 | 65 | 10 | 6 | 59 | 21 |
| Vietnam | 1 | 74 | 5 | 1 | 73 | 7 |
| Zambia | 61 | 62 | 12 | 38 | 24 | 67 |
| Zimbabwe | 40 | 60 | 14 | 24 | 36 | 51 |

# N Building PALE and ED in practice

How should a practitionner build our different indices with available data ? Table 5 presents the different step required to build our PALE and ED indices, the data source as well as their 2019 value.

**Table 5:** PALE and Expected Deprivation in the developing world in 1990 and 2019, with $\hat{a} = 70$.

| | Unit | 1990 Value | 2019 Value | Computation |
|---|---|---|---|---|
| Life Expectancy (LE) | Years | 62.2 | 71.0 | Source: GBD (2019) |
| Poverty Headcount (H) | % | 48.8 | 9.8 | Source: Poverty and Inequality Platform |
| LE*H | Years | 30.4 | 7.0 | LE* H |
| $PALE_1$ | Years | 31.9 | 64.0 | LE-$\theta$H*LE |
| Life Gap Expectancy$_{70}$ | Years | 13.0 | 7.2 | See Section 3.2 |
| Expected Deprivation$_{1,70}$ | % | 57.6 | 18.1 | $\frac{LGE_{\hat{a}}}{LE+LGE_{\hat{a}}} + \theta \frac{LE*H}{LE+LGE_{\hat{a}}}$ |

# O Ambiguous countries' trajectories

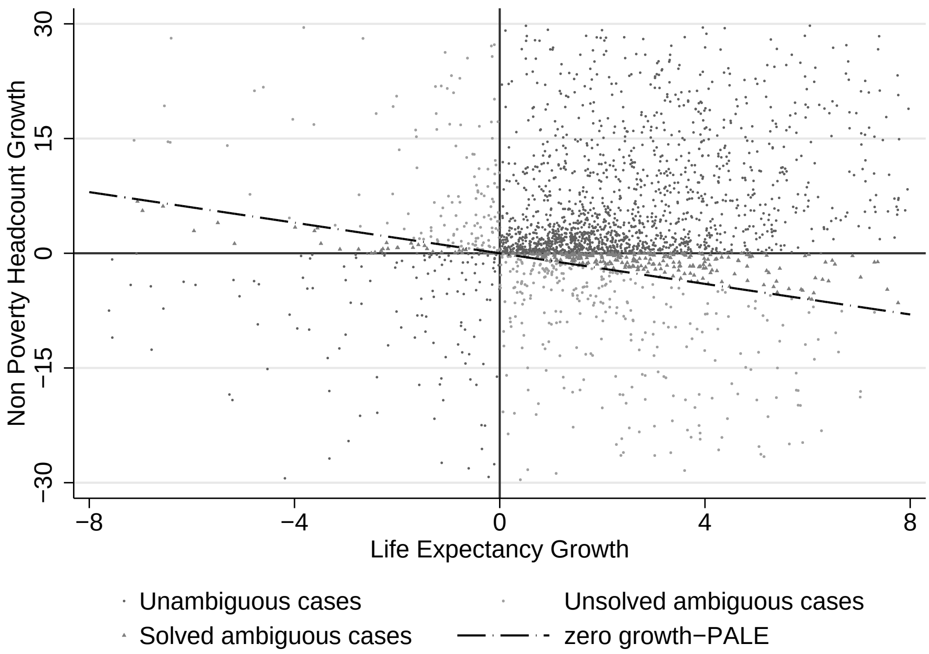
In Figure 12, we provide $PALE_\theta$ comparisons within countries between present and past situations. More precisely, for each year, we compare the situation in period t to the situation prevailing in the same country five years earlier. Given that each country's situation changed over time, we need to adapt our graphical presentation to represent the set of situations for which $PALE_\theta$ stays constant over time. We conservatively assume $\theta$ equal to one.

By definition, $PALE_1 = LE(1 - H)$, and thus $PALE_1$ increases if and only if $dLE/LE > d(1 - H)/(1 - H)$. This simple expression allows us to contruct a figure in the $(dLE/LE, d(1 - H)/(1 - H))$ plan, in which the rate of growth of $LE$ is measured on the horizontal axis, and the rate of growth of $(1 - H)$, which we refer to as the "Non-poverty Headcount", on the vertical axis. We define the "zero-growth $PALE_1$" curve, which represents all the combinations of the two growth rates such that $PALE_1$ remains unchanged: $dLE/LE = d(1 - H)/(1 - H)$ . Above this curve, $PALE_1$ increases and below this curve $PALE_1$ decreases.

The situations of interest are located in the northwest and in southeast quadrants in which the two indicators move in opposite directions. In these quadrants, there are two regions, one in the triangle below the curve in the northwest quadrant, and one in the triangle above the curve in the southeast quadrant for which $PALE_\theta$ is able to provide a clear welfare comparison. In these two areas, the shaded triangles represent situations in which, in a particular country, the situation either strictly improved (in the southeast quadrant) or deteriorated (in the northwest quadrant) compared to the situation prevailing in the same country five years earlier.[35]

[35] Again, if being dead is strictly worse than being poor, so that $\theta$ is always strictly lower than one, more situations can be strictly signed. They are located in the triangle above the "zero-growth $PALE_1$" in the NW quadrant, and in the triangle below the "zero-growth $PALE_1$" in the SE quadrant.

**Figure 12:** Resolution of ambiguous countries' trajectories, 1990-2019




Reading: Each dot represents a country-year. Countries located in the southwest (northeast) quadrant are worse (better) off than they were 5 years earlier. Countries' evolution located in the other quadrants can not be unambiguously assessed with a dashboard approach. Countries' trajectories located between the zero growth-PALE curve and the zero non poverty headcount growth line can be unambiguously assessed with PALE.
Note: for readibility, the graph only shows the points situated between a growth rate of +/- 30% in non poverty headcount and of +/-8% in life expectancy. These are 85% of all observations.