DISCUSSION PAPER SERIES

DP14059

"TOO YOUNG TO DIE". DEPRIVATION MEASURES COMBINING POVERTY AND PREMATURE MORTALITY

Jean-Marie Baland, Benoit Decerf and Guilhem Cassan

DEVELOPMENT ECONOMICS

PUBLIC ECONOMICS

ISSN 0265-8003

"TOO YOUNG TO DIE". DEPRIVATION MEASURES COMBINING POVERTY AND PREMATURE MORTALITY

Jean-Marie Baland, Benoit Decerf and Guilhem Cassan

Discussion Paper DP14059
Published 16 October 2019
Submitted 02 October 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

“TOO YOUNG TO DIE”. DEPRIVATION MEASURES COMBINING POVERTY AND PREMATURE MORTALITY

Abstract

Most measures of deprivation concentrate on deprivation among the living population and, thus, ignore premature mortality. This omission leads to a severe bias in the evaluation of deprivation. We propose three different measures that combine information on poverty and premature mortality of a population in a meaningful manner. These indices are consistent and satisfy a number of desirable properties unmet by all other measures combining early mortality and poverty. Moreover, these measures are readily computable with available data and easily interpretable. We show that omitting premature mortality leads to an underestimation of total deprivation in 2014 of at least 30% at the world level. We also show that the ranking in terms of deprivation of countries is substantially changed with our measures, and that our nderstanding of the evolution of countries' deprivation may be reversed when taking premature mortality into account.

JEL Classification: D63, I32, O15

Keywords: Deprivation Measurement, Premature Mortality, Composite Indices

Jean-Marie Baland - jean-marie.baland@fundp.ac.be
University of Namur and CEPR

Benoit Decerf - benoit.decerf@unamur.be
university of namur

Guilhem Cassan - guilhem.cassan@unamur.be
university of namur

"Too young to die". Deprivation measures combining poverty and premature mortality*

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

July 17, 2019

Abstract

Most measures of deprivation concentrate on deprivation among the living population and, thus, ignore premature mortality. This omission leads to a severe bias in the evaluation of deprivation. We propose three different measures that combine information on poverty and premature mortality of a population in a meaningful manner. These indices are consistent and satisfy a number of desirable properties unmet by all other measures combining early mortality and poverty. Moreover, these measures are readily computable with available data and easily interpretable. We show that omitting premature mortality leads to an underestimation of total deprivation in 2014 of at least 30% at the world level. We also show that the ranking in terms of deprivation of countries is substantially changed with our measures, and that our understanding of the evolution of countries' deprivation may be reversed when taking premature mortality into account.

JEL: D63, I32, O15.

Keyworks: Deprivation Measurement, Premature Mortality, Composite Indices.

*Acknowledgments : We express all our gratitude to Kristof Bosmans, Francois Maniquet, Erwin Ooghe, Paolo Piacquadio, Gregory Ponthiere and Martin Van der Linden for helpful discussions. We are grateful to the audiences at UNamur, INED, Paris School of Economics, KULeuven, U. Saint-Louis, EDEEM Berlin, EUDN conference in Bergen and Tinbergen Institute for providing insightful comments. We thank Yiyang Luo for her excellent research assistance. This work was supported by the Fonds de la Recherche Scientifique - FNRS under Grant n° 33665820. All errors remain our own.

[†]CRED, DEFIPP, University of Namur. jean-marie.baland@unamur.be

[‡]CRED, DEFIPP, University of Namur. guilhem.cassan@unamur.be

[§]CRED, DEFIPP, University of Namur. benoit.decerf@unamur.be

1 Introduction

"No winning words about death to me, shining Odysseus!
By god, I'd rather slave on earth for an other man-
some dirt-poor tenant farmer who scrapes to keep alive-
than rule down here over all the breathless dead."
Achilles' ghost to Odysseus, Homer, Odyssey.

Consider the evolution of Botswana at the end of the last century. In 1990, life expectancy in Botswana was 63.6 years while 33.6 % of its population was considered as extremely poor. In 2000, life expectancy was 45.6 years, while the proportion of extremely poor people had dropped to 29.5%.[1] Over a decade in Botswana, extreme (income) poverty has decreased, but people also lived a shorter life. The question we raise in this paper is how to evaluate, in a simple, meaningful and unambiguous manner, the evolution of total deprivation in Botswana between 1990 and 2000.

Deprivation is widely recognized as a multidimensional phenomenon (Alkire and Foster, 2011). The dimensions most often discussed, e.g. income, education or health, can only affect individuals when they are alive. We instead consider here *premature* mortality, which involves taking into account dead individuals. As such, dying is not a form of deprivation: everyone is mortal while being deprived means falling short of a minimal standard in a given resource. However, an individual dying *too young* is deprived in the sense that she will not live a number of years considered as minimally acceptable. Therefore, from the perspective of premature mortality, the resource of interest is the number of years spent alive, i.e. the lifespan.[2] As the lifespan is an important resource for well-being (Sen, 1998; Deaton, 2013), we argue that lifespan deprivation is a serious form of deprivation and should therefore be taken into account when evaluating total deprivation.

In this paper, we develop measures of deprivation that explicitly take lifespan deprivation into account. In this respect, the measures of deprivation proposed so far in the literature are unsatisfactory either because, as most poverty indices, they simply ignore lifespan deprivation or because, as most composite indices, they account for it in an questionable fashion. More precisely, the main weakness of simple composite indices is that they do not hold constant the trade-off (i.e. the exchange rate) between alive deprivation and lifespan deprivation. To illustrate this point, consider the example given in Table 1, which compares three societies. In all societies, two individuals are born every year and no individual lives for more than two years. In society A, the two newborns are non-deprived and the two 1-year-old are (income) poor. As we assume the age threshold defining lifespan deprivation to be 2 years, no individual is lifespan deprived in society A. Society B is identical to society A except for the status of a 1-year-old individual: she is prematurely dead instead of being poor. Similarly, society C is identical to society B except for a 1-year-old individual who is prematurely dead instead of poor.

[1] We present our databases below.

[2] This way of accounting for premature mortality is different from the missing poor approach followed by Lefebvre et al. (2013) and from the missing women approach (Anderson and Ray, 2010), where individuals dying in excess to a death rate are considered missing (see Section 3). We take an absolute deprivation approach to mortality, while the missing poor and missing women approaches take a counterfactual approach based on reference mortality rates. In our view, any individual dying early is deprived (and therefore, "missing"), while in the missing women and missing poor approaches, individuals dying at an early age may or may not be considered "missing".

Table 1: Composite indices are not consistent.

	0 year old	1 year old	HC	LD	$P_{0.5}$
Society A	Non-Poor, Non-Poor	Poor, Poor	0.5	0	0.25
Society B	Non-Poor, Non-Poor	Poor, Dead	0.33	0.25	0.29
Society C	Non-Poor, Non-Poor	Dead, Dead	0	0.5	0.25

The age threshold defining lifespan deprivation is 2 years

Income poverty is measured by the head-count ratio (HC), i.e. the fraction of alive individuals who are poor, which is 0.5, 0.33 and 0, respectively in society A, B and C. Lifespan deprivation is measured by the fraction of individuals who are lifespan deprived (LD), which is 0, 0.25 and 0.5 respectively, in society A, B and C. A typical composite index of total deprivation simply aggregates the two dimensions by weighing them:

$$P_w = wHC + (1 - w)LD$$

where $w \in [0, 1]$ is the weight parameter w. Assuming for instance $w = 0.5$, we get that total deprivation as measured by $P_{0.5}$ is *smaller* in society A than in society B but *larger* in society B than in society C. Yet, comparing society B to A, or C to B, the only difference between those societies is that a single individual changed status, from being poor to being dead. We therefore call judgments based on simple composite indices "inconsistent" as they do not satisfy a basic separability property and typically imply that being poor is worse than being prematurely dead in some situations but better in other situations. This inconsistency arises because the two measures, HC and LD, are based on different reference populations. The measures of deprivation we propose below do not suffer from this undesirable property.

In the first part of this paper, we develop, characterize axiomatically and compare three new indices of total deprivation that explicitly combine alive deprivation and lifespan deprivation. Two main lessons can be drawn from this exercice. First, when aggregating different dimensions of deprivation, lifespan deprivation should be treated separately. The fundamental reason lies in the *exclusive* nature of this dimension: individuals, once dead, cannot be considered as deprived along another dimension. This implies that, to measure total deprivation, a lifespan deprivation component can be added to an alive deprivation component. Second, when measuring total deprivation in a given year, the lifespan deprivation component should be measured in *time units*, i.e. the number of years prematurely lost due to early death. The reason is twofold. First, the alive deprivation component is also measured in time units since it records the number of (alive) individuals who are poor *in a given year*, that is, the number of years spent in poverty by members of the population in the given year. Second, as we show below, the number of individuals who should be alive today but died prematurely in the past corresponds logically to the number of person-years prematurely lost due to current mortality (they are actually identical in stationary populations).

Our theory provides the foundations for a particular aggregation of alive deprivation and lifespan deprivation based on time units. More precisely, our indices aggregate person-years in alive deprivation (PYADs) with person-years prematurely lost (PYPLs), given an age threshold below which dying is considered as premature. The three indices we propose, Inherited Deprivation (ID), Generated Deprivation (GD) and Expected Deprivation at birth (ED), are consistent and satisfy a number of desirable properties unmet by all other measures combining deprivation among the living and premature mortality. ID is based on past mortality, and records individuals who, in period t, should be alive given the age threshold but have died prematurely. The two other indices are based on current age-specific mortality rates, which makes them more sensitive to contemporaneous changes in the society. GD is based on the actual number of years prematurely lost by individuals who die prematurely in period t, while ED is based on the number of years a newborn expects to lose prematurely, given the current mortality rates. This last measure has the lowest inertia (i.e. reacts instantaneously to mortality shocks), is easily interpretable and requires less information (as only age-specific mortality rates in period t are necessary to compute the lifespan deprivation component). Moreover, we show that ED judgments are under some conditions equivalent to decisions made under the veil of ignorance à la Harsanyi (1953). These three indices divide the sum of years spent in alive deprivation and years prematurely lost by the counterfactual number of years obtained when no one dies before reaching the age threshold (which can be interpreted as the minimum number of years an individual should normally live).

Some may question the value added of indices aggregating different deprivation dimensions, like alive and lifespan deprivation, since such aggregation relies on a normative weight given to one dimension over the other. The choice of such normative weight may appear arbitrary. In the following, however, we rely on the relatively mild normative assumption according to which one year prematurely lost is at least as bad as one year spent in alive poverty. This is in sharp contrast with most composite indices, for which no normative assumption can be brought upon to justify a particular weight.

Several measures have been proposed in the literature to combine basic welfare with mortality indicators into a single index. The first approach is to use *composite* indices such as the Human Development Index. This simple indicator of well-being aggregates mortality with income information as a weighted sum of its mortality and income components, typically using equal weights. As discussed in Ravallion (2011), this type of aggregation hides underlying trade-offs between the dimensions being aggregated. More fundamentally, as shown above, composite indices of deprivation fail to satisfy a basic separability property. By contrast, our proposed indices are based on an explicit weighing parameter, which measures the marginal rate of substitution between the two dimensions in the relevant space. The value of this weighing parameter can therefore be chosen normatively, in a meaningful and transparent way.

The second approach is to use preference-based indicators that aggregate the quality and quantity of life by assuming or calibrating a particular inter-temporal utility function, unique across time and space (Gary S. Becker and Soares, 2005; Grimm and Harttgen, 2008; Jones and Klenow, 2016). In contrast, our indicators aggregate these two aspects without relying on a particular representation of the preferences. From the perspective of the practitioner, they are therefore more workable as they require selecting values for only two transparent normative parameters: the age threshold and the weighing parameter. Moreover, they are much less data demanding, while providing easily interpretable indicators.

The third approach is to aggregate both aspects while keeping an exclusive focus on poverty. As discussed in Kanbur and Mukherjee (2007), differences in mortality rates across income groups lead to serious mis-measurement of income poverty. Indeed, higher mortality rates among the poor lead to a "mortality paradox", whereby poor who died early are ignored in most measures of deprivation. As a result, alive deprivation is, in this sense, underestimated. They therefore propose to assign fictitious incomes to the prematurely dead individuals, in order to provide a more accurate measure of deprivation (see in particular Kanbur and Mukherjee (2007), Lefebvre et al. (2013, 2017)). The validity of these approaches relies on the assumptions made in the construction of these counterfactual, "fictitious" incomes. Our approach is fundamentally different. We do not interprete the premature death of poor individuals as the source of a downward bias in the measurement of alive deprivation. We rather consider premature death as a form of deprivation in itself, that differs from alive deprivation. Moreover, we constrain ourselves to an information set-up in which mortality rates are not known for different income groups (we discuss that constraint in Section 3).

In the second part of the paper, we investigate the importance, the distribution and the evolution of total deprivation (income poverty and lifespan deprivation) in the developing world over the period 1990-2014, combining data sets on income deprivation (PovCalNet) and on mortality (Global Burden of Disease). Under conservative assumptions, we show that lifespan deprivation is not negligible as compared to income poverty, and that its relative importance increases over time. The omission of lifespan deprivation leads to an underestimation of global total deprivation of at least 20 to 25% during the whole period. In 2014, there were 680 millions income poor individuals (PYADs) and premature mortality in the same year caused the loss of 390 millions person-years (PYPLs). Moreover, the relative importance of lifespan deprivation in total deprivation has been increasing over time: the omission of premature mortality from deprivation measures leads to an increasing bias.

At the country level, important differences arise between alive deprivation and total deprivation, and the evolution of total deprivation regularly contradicts that of income poverty for several countries and periods. Thus, for more than 7% of the country-periods considered, total deprivation evolves in the opposite direction as income deprivation. Deprivation assessments ignoring premature mortality are therefore seriously biased, and may lead to flawed policy evaluations.

The remainder of the paper is organized as follows. The three indices are presented, characterized and discussed in Section 2. The differences between our indices and alternative approaches are discussed in Section 3. Our empirical results at aggregated level and at country level are presented in Section 4 and 5 respectively. We conclude in Section 6.

2 Three families of total deprivation measures

2.1 Basic framework

In this section, we propose three measures of total deprivation that incorporate in a single index alive deprivation and lifespan deprivation. We first present and characterize an index based on past mortality. We refer to this index as the inherited deprivation index (ID).

In the next few paragraphs, we only introduce the minimal elements that are necessary to define our ID index. The remainder of our framework is presented at the beginning of Section 2.3.

In period t, each individual i is characterized by a bundle $x_i = (b_i, s_i)$, where $b_i \in \mathbb{Z}$ is her birth year with $b_i \leq t$ 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. $s_i \in S = \{NP, AP, D\}$. To keep terminology short, we often write that an indiviual whose status is AP is "poor".

This three-status framework is intentionally restrictive in order to focus our attention on the aggregation of lifespan deprivation with other forms of deprivation. Our results can easily be extended to richer structures where individual achievements are measured in multiple dimensions. These achievements could perfectly well be measured using continuous variables rather than with categorical variables.[3]

For simplicity, we assume that births occur at the beginning of a period 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.[4]

Let $a_i = t - b_i$ be the age that individual i would have in period t given her birth year b_i. Measuring lifespan deprivation requires the definition of a normative lifespan threshold $\hat{a} \in \mathbb{N}$. The introduction of this age-threshold is in line with the methodology used in the literature on multidimensional poverty measurement (Alkire and Foster, 2011; Pattanaik and Xu, 2018). This literature assumes dimension-specific thresholds in order to define dimension-specific deprivation status. We consider the lifespan of individuals as one important dimension in which individuals can be deemed deprived if their lifespan is too short. Threshold $\hat{a}$ defines when a lifespan is too short. In our terminology, an individual is "prematurely dead" if she died before reaching this minimal lifespan. We say that period t is prematurely lost by any individual i with $s_i = D$ and $a_i < \hat{a}$. This lifespan threshold is assumed independent on the distribution, which corresponds to an "absolute" definition of lifespan deprivation.

A distribution $x = (x_1, \ldots, x_{n(x)})$ specifies the birth year 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 denoted as:

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

This framework extends the one used in the traditional poverty measurement literature in two ways: to all individuals is attached a birth year and some individuals may be dead. A total deprivation index ranks all distributions in the set X as a function of the deprivation that they contain. Formally, it is a function $P : X \times \mathbb{N} \to \mathbb{R}_+$, where $P(x, \hat{a}) \geq P(x', \hat{a})$ means that x has weakly more deprivation than x'

[3] In such richer framework, we would need to define dimension-specific deprivation thresholds, impose a series of classical axioms that would constrain how to aggregate the continuous achievements in these multiple dimensions, and ultimately obtain a classification of individuals into those who are multidimensionally deprived and those who are not multidimensionally deprived. The first category could be described by a continuous multidimensional poverty score, in the vein of Alkire and Foster (2011). In order to simplify the exposition, we directly assume this score to be zero or one.

[4] All newborns have age 0 during period t and some among these newborns may die at the end of period t. This implies that $b_i = t \Rightarrow s_i \neq D$.

and strictly more if $P(x, \hat{a}) > P(x', \hat{a})$. For expositional purpose, we simplify the notation $P(x, \hat{a})$ to $P(x)$ in most of what follows since $\hat{a}$ is assumed fixed.

2.2 The inherited deprivation index

Our extended framework reveals that classical deprivation indices are not sensitive to lifespan deprivation. Consider the following distribution in period t with three individuals

$$x = ((young, NP), (young, D), (old, D)),$$

where a birth year more distant than $\hat{a}$ years before t is noted as *old* (*young* otherwise). Because she is young and dead, individual 2 is prematurely losing period t. We contrast distribution x with two alternative distributions x' and x'' in period t that are both obtained from x by changing the status of individual 2. In x', individual 2 is alive and non-poor, while in distribution x'' individual 2 is alive and poor, i.e.

$$x' = ((young, NP), (young, NP), (old, D))$$
$$x'' = ((young, NP), (young, P), (old, D)).$$

These three distributions are compared in Table 2.

Table 2: Comparing distributions using the head-count ratio and the Inherited Deprivation index

	(young, P)	(young, NP)	(young, D)	(old, D)
Distribution x	0	1	1	1
Distribution x'	0	2	0	1
Distribution x''	1	1	0	1

In these three distributions, no individual is alive and poor, except individual 2 in distribution x''. Therefore, the head-count ratios (HC) of distributions x ($HC = \frac{0}{1}$) and x' ($HC = \frac{0}{2}$) are identical: it is zero and is lower than that of x'' ($HC = \frac{1}{2}$).[5] However, distribution x' is arguably better than distribution x, since individual 2 is not prematurely dead in the former. It is also not clear that distribution x'' is worse than distribution x: individual 2 is poor in x'' but prematurely dead in x. Whether distribution x is doing better than distribution x'' is a judgment based on how one compares spending period t in poverty versus prematurely loosing period t. In our epigraph, for example, Achilles clearly states that spending a year in poverty is much preferable than spending a year in lifespan deprivation: Achilles would consider that distribution x is much worse than distribution x''.

Our inherited deprivation indices would consider distribution x more deprived than distribution x'. The comparison between society distribution x and x'' would depend on a weighting parameter γ which transparently weights a year spent in alive deprivation compared to a year spent in lifespan deprivation.

The HC is not able to capture the difference between distributions x and x' because dead individuals do not matter in its computation. Formally, classical de-

[5] The comparison of distribution x'' to distribution x is an example of the "mortality paradox": the reason why the HC of x'' is higher than that of x is because the poor individual of distribution x'' is dead in distribution x.

privation indices, such as HC, satisfy the property of Independence of Dead, which requires that the presence of an additional dead individual does not affect them.

Deprivation axiom 1 (Independence of Dead). *For all $x \in X$ and $i \leq n(x)$, if $s_i = D$, then $P(x_i, x_{-i}) = P(x_{-i})$.*

This property is too strong for total deprivation indices, because it implies that prematurely dead individuals are irrelevant. Rather, the inherited deprivation index satisfies a weaker property which requires that the presence of an additional dead individual does not affect it only when this individual is born at least $\hat{a}$ years before period t:

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

Weak Independence of Dead defines the reference population relevant for total deprivation indices. A priori, a distribution x contains all individuals that ever lived in a particular society. Weak Independence of Dead implies that two types of individuals are irrelevant in period t: those who died above the age threshold and those who died below the age threshold but too far away in the past. Among the dead individuals, only those who died prematurely and whose birth year is less than $\hat{a}$ years before t enter the reference population.

We can now introduce the inherited deprivation index. 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 index is defined as

$$P_\gamma^{ID}(x) = \underbrace{\frac{p(x)}{f(x) + p(x) + d(x)}}_{alive\ deprivation} + \gamma \underbrace{\frac{d(x)}{f(x) + p(x) + d(x)}}_{lifespan\ deprivation}, \tag{1}$$

where $\gamma > 0$ is a parameter weighing the relative importance of alive deprivation and lifespan deprivation. An individual losing prematurely period t matters γ times as much as an individual spending period t in alive deprivation.

Index P_γ^{ID} has an alive deprivation component (poverty) and a lifespan deprivation component (premature mortality). The alive deprivation component counts the number of persons who are poor in period t, and the lifespan deprivation component records the number of persons who were born less than $\hat{a}$ years before t but have already died. The denominator of both components is identical and equal to the number of individuals in the reference population. This reference population includes all individuals born less than $\hat{a}$ years before t as well as all older individuals who are still alive in t. This reference population may include dead individividuals born less than $\hat{a}$ years before t. However, whether the individuals who were born less than $\hat{a}$ years before t died or not does not change our reference population.

We are now able to further illustrate the main differences between the ID index and a classical deprivation measure, such as HC. Consider again distributions x, x' and x''. As required by Weak Independence of Dead, P_γ^{ID} compares these distributions by focusing on young individuals, no matter whether they are alive or not, and old individuals who are alive. In all three distributions, the relevant population is composed of three individuals. As individual 2 is prematurely dead in distribution x whereas she is alive and non-poor in distribution x', $P^{ID}_{\gamma}(x) > P^{ID}_{\gamma}(x')$. In addition, as individual 2 is prematurely dead in distribution x whereas she is alive and poor in x'', $P^{ID}_{\gamma}(x) \geq P^{ID}_{\gamma}(x'')$ when $\gamma \geq 1$. In that case, the larger premature mortality in x more than compensates for the larger alive deprivation in x'', and the ID index contradicts HC. ID therefore provides a more comprehensive picture of total deprivation in period t than HC.

We now show that the ID index is characterized by a small number of desirable properties. First, Least Deprivation requires that being non-poor is better than being either poor or prematurely dead. This weak axiom compares distributions with a unique individual, i.e. individual 1. Recall that we assumed for distributions with a unique individual that, if the individual is dead, then she is prematurely dead.

Deprivation axiom 3 (Least Deprivation)**.**
$P(b_1, NP) < P(b_1, AP)$ *and* $P(b_1, NP) < P(b_1, D)$.

The second property, Weak Independence of Birth Year, requires that the particular year of birth of an individual in the reference population is irrelevant, only her status matters. The birth year is only relevant in order to distinguish prematurely dead from other dead individuals.[6] This property implies that each person-year lost due to premature death contributes equally to the index.

Deprivation axiom 4 (Weak Independence of Birth Year)**.** *For all* $x \in X$ *and* $i \leq n(x)$*, if* $s_i = s'_i$ *and* $d(x_i, x_{-i}) = d(x'_i, x_{-i})$*, then* $P(x_i, x_{-i}) = P(x'_i, x_{-i})$.

Weak Independence of Birth Year requires that one person-year prematurely lost matters equally in the index, independently of the particular age of the individual who died. Thus, if $\hat{a}$ is equal to 50, the death of a newborn in $t-1$ is equivalent to the death of a 48 years old in $t-1$ in the computation of the ID index at period t. However, the death of the younger individual will be recorded in the ID indices for several periods following her death, while the death of the 48 years old individual will be accounted for only once (in the period t following her death). In that sense, the death of the younger individual matters proportionally more.

Finally, we impose a standard separability property, Subgroup Consistency. This axiom requires that, if deprivation decreases in a subgroup while remaining unchanged in the rest of the distribution, overall deprivation must decline.[7]

Deprivation axiom 5 (Subgroup Consistency)**.** *For all* $(x, x'), (x, x'') \in X$*, if* $P(x') > P(x'')$ *and* $f(x')+p(x')+d(x') = f(x'')+p(x'')+d(x'')$*, then* $P((x, x')) > P((x, x''))$.

To be complete, three auxiliary properties are also needed. First, the name of individuals should not influence the deprivation index (Anonymity). Second, if a distribution is obtained by replicating another distribution several times, they both

[6]This is why Weak Independence of Birth Year has the precondition $d(x_i, x_{-i}) = d(x'_i, x_{-i})$, which holds the number of prematurely dead constant: the birth year b'_i can be different from b_i, but if $s_i = D$, then individual i is either prematurely dead in both x_i and x'_i, or in none of these two bundles.

[7]The precondition $f(x') + p(x') + d(x') = f(x'') + p(x'') + d(x'')$ ensures that distributions x' and x'' have a relevant population with the same size. The additive separability result of Foster and Shorrocks (1991), which rationalizes the use of additive indices, is based on a stronger version of Subgroup Consistency with the additional precondition $f(x') + p(x') = f(x'') + p(x'')$.

have the same deprivation (Replication Invariance). Finally, the deprivation index evolves "continuously" on its domain. Given that this domain is discrete, the index should satisfy a particular continuity property as proposed by Young (1975) (Young Continuity).[8]

Proposition 1 fully characterizes the ID index, which implies that any deprivation index satisfying our properties ranks distributions in exactly the same way as the ID index.

Proposition 1 (Characterization of ID)**.**
P is ordinally equivalent to P^{ID}_{γ} if and only if P satisfies Weak Independence of Dead, Least Deprivation, Weak Independence of Birth Year, Subgroup Consistency, Anonymity, Replication Invariance and Young Continuity.

Proof. See Appendix 7.1. □

Proposition 1 is a stepping stone for our main results in Section 2.3. Yet, this proposition supports the first message coming out of our theory. It implies that alive deprivation and lifespan deprivation enter the index in an additive way, so that computing the ID index amounts to a very basic accounting exercise. The fundamental intuition underlying this additive separability is that an individual cannot simultaneously be "prematurely dead" and "poor": these two statuses are mutually exclusive, which allows us to sum the number of prematurely dead individuals with the number of individuals affected by alive deprivation. In contrast, non-exclusive dimensions of alive deprivation such as (say) material deprivation and health deprivation would not be additively separable. These two dimensions imply *three* different statuses, namely being materially deprived, health deprived and deprived in both dimensions. If these three statuses can still be aggregated in an additive way, it is clear that the two dimensions are not additively separable because of the presence of the third mixed status (deprived in both dimensions).

A number of remarks are in order. To begin with, the implementation of the ID index involves two important normative choices. The first one is the choice of $\hat{a}$, the age threshold below which the death of an individual contributes to total deprivation. The second one is the value of γ, the parameter weighing the relative importance of alive deprivation and premature mortality.

Then, our definition of the individual status is agnostic to the particular definition of alive deprivation, and could as well capture income deprivation, as in our empirical application, or multidimensional poverty (Alkire and Foster, 2011). Proposition 1 can easily be extended to a framework in which alive deprivation is measured as a continuous variable such as an income deprivation score or a multidimensional poverty score, provided that the axioms are duly adapted (see Foster and Shorrocks (1991)).

Finally, our definition of a distribution does not simultaneously contain information about an individual deprivation status and on her chances of survival. This particular assumption, which we discuss more carefully at the end of Section 2, is consistent with the data constraints we face in real world measurements, in which it is rare to find comparable data sets that simultaneously contain information about lifetime duration and deprivation status at the individual level. The absence of this

[8] Formal definitions of these traditional axioms can be found in Appendix 7.1.

information implies that all our measures are indifferent about the repartition across individuals of periods spent in alive deprivation and periods prematurely lost. At the end of Section 2, we discuss how to adapt our main index for applications for which the relevant information is available.

The ID index suffers from two limitations arising from the fact that total deprivation in t depends on *past* mortality (before period t). First, computing this index requires detailed information on mortality of each age cohort for all $\hat{a}$ years preceding t, and can therefore only be computed for situations for which such data exist. Second, the ID index exhibits inertia, which may be undesirable, for instance when used to evaluate the impact of public policies. The impact of a mortality shock, whether permanent or temporary, takes decades to be fully accounted for, as the impact of a shock continues to matter for the $\hat{a} - (a_i + 1)$ years following the death of i. For instance, today's ID index for Rwanda's still accounts for children who died during the genocide of 1994. One can consider that past mortality shocks which occurred decades ago are not particularly relevant to current state of a society. The two indices that we propose below still account for premature mortality while improving on these limitations.

2.3 The generated and expected deprivation at birth indices

While the ID index is an intuitive and straightforward manner to include premature mortality in deprivation measures, its limitations make its empirical implementation difficult. We therefore propose two total deprivation indices who can easily be computed with available datasets and have less inertia. They are based on mortality rates in period t instead of on mortality before period t, but rely on the same intuition as the ID index. In particular, these indices offer the same diagnostic as the ID index when evaluating stationary populations (see the "ID equivalence" axiom, which we formalize below). As an alternative foundation, we also show that the way they aggregate lifespan deprivation to alive deprivation is closely linked to decision making under the veil of ignorance à la Harsanyi (1953).

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 $s_i \neq D$. These numbers $n_a(x)$ entirely define the population pyramid in period t. Let $d_a(x)$ 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 $n_a(x) + d_a(x)$. The age-specific mortality rate μ_a denotes the fraction of *alive* a-year-old individuals dying at the end of period t. Hence, the number of a-year-old individuals dying at the end of period t is $n_a(x) * \mu_a$. We have that $\mu_a \in \mathcal{M} = [0,1] \cap \mathbb{Q}$, where $\mathbb{Q}$ is the set of rational numbers.[9] Letting $a^* \in \mathbb{N}$ stand for the maximal lifespan (which implies $\mu_{a^*} = 1$), the vector of age-specific mortality rates in period t is given by $\mu = (\mu_0, \ldots, \mu_{a^*})$. Vector μ summarizes mortality in period t, while distribution x summarizes alive deprivation in period t

[9] Given that distributions have finite numbers of individuals, mortality rates cannot take irrational values.

and mortality before period t.[10] The set of mortality vectors is given by

$$M = \left\{ \mu \in \mathcal{M}^{a^*+1} | \mu_{a^*} = 1 \right\}.$$

We consider pairs $(x, \mu) \in X \times M$, for which the distribution x is a priori unrelated to vector μ. We only impose that the age-specific mortality rates μ_a is feasible given the number of alive individuals $n_a(x)$, which is $\mu_a = \frac{c}{n_a(x)}$ for some $c \in \mathbb{N}$. Our next total deprivation indices are defined on domain O, which is the subset of pairs in $X \times M$ that meet the above restriction. Formally, an index is a function $P : O \times \mathbb{N} \to \mathbb{R}_+$. Again, we simplify the notation $P(x, \mu, \hat{a})$ to $P(x, \mu)$ since a fixed value for $\hat{a}$ is considered.

We now focus on indices that do not depend on mortality before period t and, hence, satisfy instead Independence of Dead* (where the asterisk denotes that Independence of Dead is adapted to domain O, see Appendix 7.3). Rather than embarking on a long axiomatic analysis of the two indices, we provide a characterization that builds upon our definition of the ID index. Given that *current* mortality (in period t) need not be the same as *past* mortality (before period t), the distribution and mortality vector defining a pair are in general unrelated. Yet, current mortality corresponds exactly to past mortality in *stationary populations*. A population is stationary if the number of newborns and the mortality vectors are constant over time. In such case, the population pyramid in period $t + 1$ is the same as the population pyramid in period t. In other words, the population pyramid exactly mirrors the mortality vector. We say that a pair (x, μ) 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),
- $n_{a+1}(x) = n_a(x) * (1 - \mu_a)$ (constant population pyramid).

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. This population pyramid corresponds to the one prevailing in the long run if mortality and natality rates in period t remain constant over time (see for instance Preston et al. (2000)).

In a stationary pair, past and current mortality coincide and the mortality vector μ does not convey any information that cannot be inferred from the distribution x. Indeed, the population of prematurely dead individuals in x directly reflects μ: the number of prematurely dead individuals in x can be computed from μ (and n^*). Conversely, the mortality vector μ can be computed from the population pyramid associated to x. As a result, the deprivation index can be computed from distribution x only. Proposition 1 shows that, when measuring deprivation based on distribution x only, one should use the ID index. Therefore, we require for stationary pairs that the deprivation index is equal to the ID index.

Deprivation axiom 6 (ID Equivalence). *For all $(x, \mu) \in O$ and some $\gamma > 0$, if (x, μ) is stationary, then $P(x, \mu) = P_\gamma^{ID}(x)$.*

[10] Observe again that this framework is consistent with our data-constraint. A pair (x, μ) does not simultaneously contain the information about an individual's deprivation and the information about her chances of survival.

Such equivalence is a minimal requirement for deprivation indices based on mortality in period t. Indeed, assuming constant natality, a permanent change in mortality rates affects the long-run distribution. ID Equivalence requires that the index agrees with the ID index on the long-run consequences of such a change. As we show below, this requirement allows for two different indices.

We first define the generated deprivation index (GD) as follows:

$$P_\gamma^{GD}(x,\mu) = \underbrace{\frac{p(x)}{f(x)+p(x)+d^{GD}(x,\mu)}}_{alive\ deprivation} + \gamma \underbrace{\frac{d^{GD}(x,\mu)}{f(x)+p(x)+d^{GD}(x,\mu)}}_{lifespan\ deprivation} \tag{2}$$

where d^{GD} counts the number of person-years prematurely lost generated by deaths occuring in period t:

$$d^{GD}(x,\mu) = \sum_{a=0}^{\hat{a}-1} n_a(x) * \mu_a * (\hat{a} - (a+1)).$$

According to this definition, GD is closely related to ID, as they both sum up an alive deprivation component, recording the number of person-years in alive deprivation (PYADs) and a lifespan deprivation component. The lifespan deprivation component of GD differs from that of ID, as it records the number of person-years prematurely lost (PYPLs) generated by deaths occuring in period t. By contrast, the ID index records the number of PYPLs inherited in period t, which were generated by deaths occuring before period t. When an individual dies at age $a < \hat{a}$, she prematurely loses the $\hat{a} - (a+1)$ periods following her death. The GD index records these $\hat{a} - (a+1)$ PYPLs and assigns this number to the year during which the death occurs. The denominator of GD is analogous to that of ID, as it simply adds the number of alive individuals in period t to the number of PYPLs.

Second, the expected deprivation at birth index (ED) is based on expectations given the poverty and mortality rates prevailing in period t. It again combines an alive deprivation and a lifespan deprivation component in a additive way:

$$P_\gamma^{ED}(x,\mu) = \underbrace{\frac{LE(\mu) * HC(x)}{LE(\mu) + LGE_{\hat{a}}(\mu)}}_{alive\ deprivation} + \gamma \underbrace{\frac{LGE_{\hat{a}}(\mu)}{LE(\mu) + LGE_{\hat{a}}(\mu)}}_{lifespan\ deprivation}, \tag{3}$$

where $HC(x) = \frac{p(x)}{p(x)+f(x)}$ is the head-count ratio and $LE(\mu)$ is life expectancy at birth:[11]

$$LE(\mu) = \sum_{a=0}^{a^*} \prod_{l=0}^{a-1} (1-\mu_l).$$

We interpret the term $LE(\mu) * HC(x)$ as the expected number of years that a newborn will spend in alive deprivation, given the mortality rates and head-count ratio in period t. It represents the expected person-years spent in alive deprivation for such a newborn. The second term, $LGE_{\hat{a}}$, is the *lifespan gap expectancy* relative to the

[11] Life expectancy measures the expected lifespan of an individual facing throughout her life the age-specific mortality rates reported in vector μ.

age threshold:

$$LGE_{\hat{a}}(\mu) = \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a+1)) * \mu_a * \prod_{l=0}^{a-1} (1 - \mu_l).$$

$LGE_{\hat{a}}$ measures the number of years that a newborn expects to lose prematurely if confronted to the mortality rates of vector μ throughout her $\hat{a}$ first years of life.[12] The ED index therefore takes the viewpoint of a newborn and compute the expected proportion of her *augmented life expectancy* – a normative lifetime defined as the sum of life expectancy and lifespan gap expectancy – that she expects to lose prematurely or spend in alive deprivation. A more precise interpretation will be presented at the end of Section 3.

As should be clear from their definitions, the essential difference between GD and ED indices is that GD is based on the actual population pyramid prevailing in period t, while ED, by taking the viewpoint of a newborn, is actually based on an abstract, counterfactual population pyramid defined by the mortality rates in period t. This observation invites two remarks. First, GD indirectly depends on past natality and mortality, which shape the current population pyramid on which it is defined. This implies that GD exhibits some inertia by partly reflecting deaths that occurred in the past, even if the magnitude of this inertia is smaller than that of ID. Second, unlike ED for which the only information needed is that on mortality rates in period t, GD also requires information on the relative size of each cohort in the population.

Lemma 1 shows that these two indices meet our requirement.

Lemma 1. *Both P_{γ}^{GD} and P_{γ}^{ED} satisfy ID Equivalence.*

Proof. See Appendix 7.2. □

ID and GD indices are identical in a *stationary population* because d^{GD} coincides with d in that case. The intuition for this equivalence is illustrated in Figure 1. The left panel shows that d counts "vertically" the number of individuals who are younger than $\hat{a}$ years and died before period t. The right panel shows that d^{GD} counts "horizontally", for each age group below $\hat{a}$, the number of person-years prematurely lost by individuals in that age group who die at the end of period t. When the mortality rates of the young correspond to the population pyramid, the two shaded areas coincide.

The intuition for the equivalence between ED and GD indices in stationary pairs can be illustrated by graphical representations of LE and $LGE_{\hat{a}}$, as shown in the left panel in Figure 2 for a stationary pair. The green area below the population pyramid represents life expectancy, while the lifespan gap expectancy corresponds to the pink area between the young part of the population pyramid and the age threshold. As long as the proportion of individuals of each generation corresponds to the current mortality vector, GD and ED indices provide identical measures of total deprivation.[13] When compared with the right panel in Figure 1, the right panel

[12] Note that $LGE_{\hat{a}}$ 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 death causes (Gardner and Sanborn, 1990).

[13] Observe that $LGE_{\hat{a}}$ is inversely related to LE: when the age threshold is larger than a^*, the two indicators move in opposite ways on any two mortality vectors.

in Figure 2 reveals that ED and GD take equal values in stationary pairs; only the interpretation of period-years lost slightly differs.

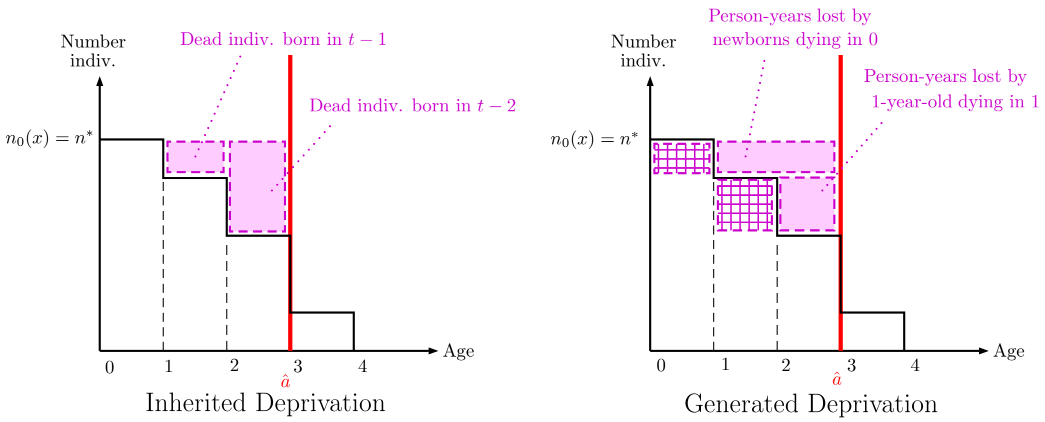


Figure 1: Left panel: The pink area above the population pyramid represents $d(x)$. Right panel: The pink area above the population pyramid represents $d^{GD}(x, \mu)$.

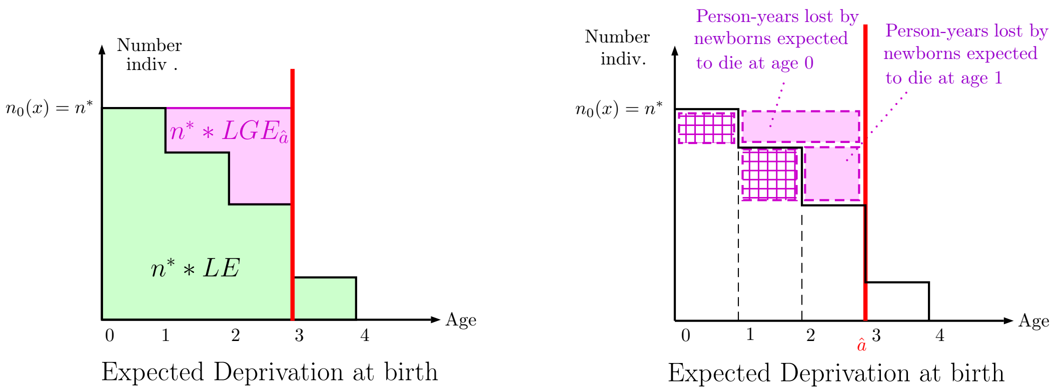


Figure 2: For the stationary pair (x, μ), the green area corresponds to a multiple of life-expectancy (LE) and the pink area to a multiple of the lifespan gap expectancy ($LGE_{\hat{a}}$).

Lemma 1 supports the second message coming out of our theory. When aggregating total deprivation in a given period t, the lifespan deprivation component must be measured in number of periods prematurely lost, at least when this component is computed from the mortality vector in t. The reason is that, in stationary populations, counting the number of individuals who prematurely miss period t due to mortality before period t is equivalent to counting the number of person-years lost due to premature mortality in period t.

Because of their equivalence for stationary pairs, these two indices share many similarities with the ID index. In particular, the person-years lost due to "mature" deaths do not enter in the reference population of PYs (as implied by Weak Independence of Dead for ID indices), all PYPLs have the same weight (in the spirit of Weak Independence of Birth Year) and the weight γ given to a PYPL relative to a

PYAD is constant (as implied by Subgroup Consistency). However, the properties of GD and ED are different, as our next characterizations reveal.

Proposition 2 characterizes the ED index. In particular, this index does not depend on the birth year of individuals (Independence of Birth Year).[14]

Proposition 2 (Characterization of ED)**.**
$P = P^{ED}_{\gamma}$ if and only if P satisfies Independence of Dead, ID Equivalence, Replication Invariance* and Independence of Birth Year.*

Proof. See Appendix 7.3. ☐

Because ED satisfies Independence of Birth Year, it does not depend on the population pyramid in period t, and therefore avoids the inertia associated with the demographic evolution of population pyramids. This advantage comes at a price, as ED cannot be decomposed additivitely between subgroups (unless the mortality vector is the same in the subgroups). This non-decomposability is intrinsic to the concept of life-expectancy, which underlies the ED index.

In contrast to ED, the GD index violates Independence of Birth Year but satisfies Additive Decomposibility, a strengthening of Subgroup Consistency. This last property implies that, if deprivation decreases in a subgroup while remaining unchanged in the rest of the population, overall deprivation declines.

Deprivation axiom 7 (Additive Decomposibility)**.** *For all $(x', \mu'), (x'', \mu'') \in O$, if $x = (x', x'')$ and $\mu_a = \frac{n_a(x')*\mu'_a + n_a(x'')*\mu''_a}{n_a(x')+n_a(x'')}$ for all $a \in \{0, \ldots, a^*\}$, then*

$$P(x, \mu) = \frac{\eta(x', \mu') * P(x', \mu') + \eta(x'', \mu'') * P(x'', \mu'')}{\eta(x', \mu') + \eta(x'', \mu'')}, \tag{4}$$

where the "size" function $\eta : O \to \mathbb{N}_0$ is such that $\eta(x, \mu) = \eta(x', \mu') + \eta(x'', \mu'')$.

Proposition 3 (Characterization of GD)**.**
$P = P^{GD}_{\gamma}$ if and only if P satisfies Independence of Dead, ID Equivalence and Additive Decomposibility.*

Proof. See Appendix 7.4. ☐

Given that GD and ED indices are different, Propositions 2 and 3 together imply that the five axioms involved are jointly incompatible. Either the index is ED and it cannot be decomposed in subgroups, or the index is GD and it exhibits some inertia by relying on the actual population pyramid.

Alternative foundation

It is worth noting that a welfare approach à la Harsanyi provides an alternative foundation for the additive separability of the lifespan deprivation and the alive deprivation components, once the unit of account is based on time units, e.g. years.[15] More precisely, there is an important link between ED and the welfare approach developed in Harsanyi (1953). According to the latter, behind the veil of ignorance, each individual faces a lottery whereby she ignores whether she will be in alive

[14] See Appendix 7.3 for the formal definitions of these axioms.

[15] We thank Dilip Mookherjee as well as Kristof Bosmans for raising this point in a discussion of an early version of this paper.

deprivation and for how long, or whether she will be the victim of a premature death. When evaluating her welfare, she considers drawing at random the life of any individual in that society. Following the formulation of Jones and Klenow (2016), her expected lifetime utility is given by

$$EU(x,\mu) = \mathbb{E}\sum_{a=0}^{a^*} \beta^a u(s_a) S(a,\mu), \tag{5}$$

where β is the discount factor, $S(a,\mu)$ is the probability an individual survives to age a given the mortality vector, the expectation operator applies to the uncertainty with respect to individual status at age a (if alive) and $u(NP) > u(AP)$ and $u(NP) > 0$. Proposition 4 shows that, if the discount factor takes value one and the age threshold is large enough, Harsanyi's welfare evaluation and ED always yield identical comparisons.

Proposition 4 (Connection between ED and Harsanyi's welfare)**.**
If $\beta = 1$ and $\hat{a} > a^*$, then there exists $\gamma > 0$ such that

$$EU(x,\mu) \geq EU(x',\mu') \Leftrightarrow P_\gamma^{ED}(x,\mu) \leq P_\gamma^{ED}(x',\mu')$$

for all $(x,\mu),(x',\mu') \in O$.

Proof. See Appendix 7.5. □

The equivalence demonstrated above relies on a three-status framework and does not survive in richer structures where individual achievements are measured by continuous variables. The reason is that, by nature, deprivation measures attribute the same value to all individuals who are non-deprived, regardless of their exact achievements.

We now investigate the relation between our three indices by contrasting their dynamic responses to mortality shocks.

2.4 Dynamic behavior of the three indices

Actual populations are typically not stationary. Permanent and transitory mortality shocks regularly affect population pyramids, which take decades to adjust to these shocks. In this section we compare the three indices for pairs that are not stationary, by investigating their reactions to different kinds of mortality shocks.

For non-stationary pairs, GD and ED indices are not equivalent as they weigh current mortality rates in young age in a different way. Proposition 5 shows that the GD index relies on the current population pyramid to weight mortality rates while the ED index uses the counterfactual population pyramid generated by the current mortality vector.

Proposition 5 (P^{ED} and P^{GD} weigh μ with different population pyramids)**.**
Take any pair $(x,\mu) \in O$ for which x has a monotone population pyramid, i.e. $n_{a+1}(x) \leq n_a(x)$ for all $a \in \{0,\dots,a^*-1\}$. Let μ^x be the mortality vector for which (x,μ^x) is a stationary pair. If $\gamma \geq 1$, then we have $P_\gamma^{ED}(x,\mu) \leq P_\gamma^{GD}(x,\mu)$ if and

only if

$$\frac{LGE_{\hat{a}}(\mu)}{LE(\mu) * HC(x) + LGE_{\hat{a}}(\mu)} \leq \frac{d^{GD}(x,\mu)}{p(x) + d^{GD}(x,\mu)} \tag{6}$$

if and only if

$$\frac{\sum_{a=0}^{\hat{a}-1}(\hat{a}-(a+1)) * \frac{n_a}{n_0}(\mu) * \mu_a}{LE(\mu)} \leq \frac{\sum_{a=0}^{\hat{a}-1}(\hat{a}-(a+1)) * \frac{n_a}{n_0}(\mu^x) * \mu_a}{LE(\mu^x)} \tag{7}$$

where $\frac{n_a}{n_0}(\nu) = \prod_{l=0}^{a-1}(1-\nu_l)$ *denotes the proportion of newborns expected to survive until age* a *given mortality vector* ν*.*

Proof. See Appendix 7.6. □

Inequality (6) shows that when index GD is larger than index ED, the former attributes a larger part of total deprivation to lifespan deprivation than the latter, i.e. emphasizes lifespan deprivation more. Then, inequality (7) shows the condition under which this happens: when the age-cohorts with a large value of $(\hat{a}-(a+1))*\mu_a$ have a larger relative size in the actual population pyramid than in the counterfactual population pyramid associated with μ.[16] As the weight of a given age-cohort is given by its mortality rate multiplied by its distance to the age threshold, each death in a younger age cohort is then associated with a larger number of PYPLs.

Virtually all our empirical results have GD larger than ED. A simple explanation can be found in inequality (7). Consider for instance a growing population characterized by a constant mortality vector and a high child mortality. The actual population pyramid has relatively more children than the counterfactual population pyramid and GD is larger than ED.

Transitory mortality shocks

We first investigate the response of our three indices to a transitory mortality shock. In the Online Appendix 1, we formally show that, in a stationary population affected by a series of transitory mortality shocks, GD and ID indices compute the same number of PYPLs, but distribute these PYPLs over different periods of time. By contrast, as the example below illustrates, ED may record a different number of PYPLs. This is again related to the fact that the latter uses a counterfactual population pyramid to weigh mortality rates.

We consider a population with a fixed natality $n_0(x) = n^* = 1$ for all period t. At each period, all alive individuals are non-poor, implying that $HC(x) = 0$. For all $t \neq 0$, we assume a constant mortality vector $\mu = \mu^* = (0,0,1)$, so that each individual lives exactly three periods. Let us fix the normative parameters at $\gamma = 1$ and $\hat{a} = 3$, so that an individual dies prematurely if she dies before her third period of life. Before period $t = 0$, the population is stationary, and the three indices are equal to zero since there is no poor and no premature deaths. Let us now consider a one period shock in period 0, such that all individuals die: $\mu^0 = (1,1,1)$. After the shock, mortality rates directly come back to their initial value and the population

[16] In the particular case $\gamma < 1$, whether inequality (7) holds or not depends on the level of alive deprivation. For values of γ smaller than the level of alive deprivation, the intuition provided in last paragraph is reversed.

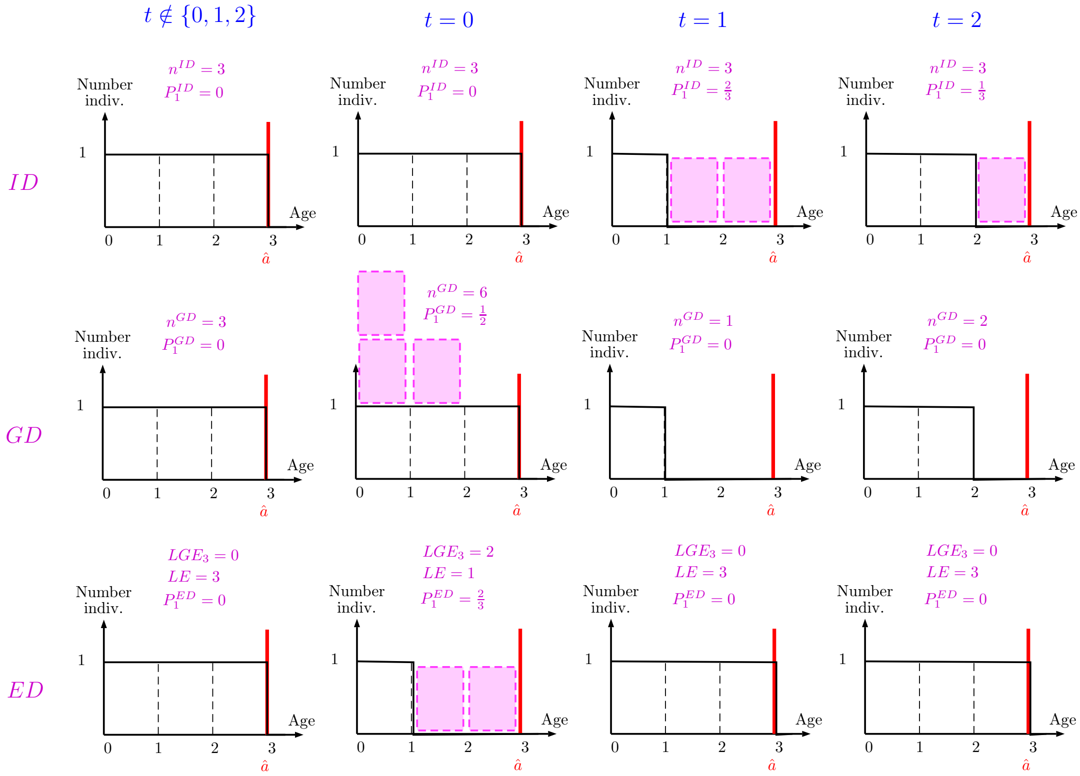


Figure 3: Response of ID, GD and ED indices to the transitory mortality shock in $t^* = 0$. The person-years that are prematurely lost are in pink. The population pyramids are drawn. For ED indices, these population pyramids are counterfactual.

pyramid returns to its stationary state in period 3, after a (mechanical) transition in period 1 and 2 during which the newborns of period 1 and 2 grow up. This example is illustrated in Figure 3.

Let us first consider the ID index. In period 0, no premature deaths are recorded, since they all happen at the end of period 0. The number of person-years prematurely lost recorded by the ID index is equal to 2 in period 1, 1 in period 2 and 0 afterwards, as illustrated by the shaded areas in the first row of Figure 3. Given that one individual is born in every period and $\hat{a} = 3$, the relevant population given by $n^{ID} = 3$ in all periods. Therefore, the ID index is equal to $\frac{2}{3}$ in period 1 and $\frac{1}{3}$ in period 2.

The GD index records the shock immediately in period 0. The newborn who dies in period 0 produces 2 PYPLs and the individual aged 1 in that period produces 1 PYPL. To compute the GD index in period 0, we consider a total of 6 person-years and the GD index is equal to $\frac{1}{2}$ in period 0. Since the newborn in period 1 does not die in period 1 and is the only individual alive, GD indices records one PY with no deprivation and no PYPL. For period 2, there are 2 individuals alive, but no deprivation, and the GD index is equal to 0 in periods 1 and 2.[17]

The ED index also records the shock in period 0. Mortality rates in period 0 are such that $LE(\mu^0) = 1$ and $LGE_{\hat{a}}(\mu^0) = 2$. The ED index is therefore equal to $\frac{2}{3}$ in period 0. In period 1 and 2 the counterfactual population pyramid, computed from

[17]Note that the fact that the index returns to its inital value after one period is a particularity of this simple example. If instead we had $n^* = 4$, $\mu^0 = (\frac{1}{2}, 1, 1)$ and $\mu^* = (0, \frac{1}{2}, 1)$, the index would not return to its stationary value in period 1.

μ^*, corresponds to the stationary population, with $LE(\mu^*) = 3$ and $LGE_{\hat{a}}(\mu^*) = 0$, and the ED index is equal to 0. In contrast with the ID index (and the GD index in general), the ED index features no inertia and is equal to its stationary value as soon as the mortality vector μ^* returns to its stationary value.

Finally, note that the ED index counts a smaller number of PYPLs than the ID or the GD index. The reason for this difference is that the mortality rate $\mu_1^0 = 1$ is given a lower weight in the ED index than in the GD and ID indices. (As a matter of fact, given that $\mu_0^0 = 1$, the newborn does not expect to survive the first period, so that the mortality rate $\mu_1^0 = 1$ is irrelevant for the ED index.)

Permanent mortality shocks

We now investigate the consequences of a permanent mortality shock on a stationary population. After a mortality shock, a transition phase sets in during which the population pyramid adjusts to the new mortality vector, before reaching a new long run equilibrium. This transition takes several decades and is particularly long in the case of a mortality shock on young age individuals. During this transition, the three indices are not equivalent.

We use simulations in order to illustrate the relative inertia of the three indices for different types of permanent shocks. The results of these simulations can be found in Figure 4. We compare indices P_1^{ID}, P_1^{GD} and P_1^{ED}, and assume that natality is constant, and there is no alive deprivation. The age threshold is $\hat{a} = 50$ and the maximal age is $a^* = 100$. Before the shock, the population is stationary when considering mortality rates equal to 1% for each age before 100. We simulate three different types of shocks: (1) the mortality rates are increased from 1 to 2% for all ages, (2) the mortality rate is increased from 1 to 2% only at age 40 and (3) the mortality rate is increased from 1 to 2% only at age 10. Figure 4 shows the result of our simulation experiments.

The upper graph illustrates the consequences of the uniform mortality shock, the middle graph of the mortality shock at age 40 and the bottom graph of the mortality shock at age 10. The three indices evolve very differently over the transition period. In all scenarios, the ED index jumps immediately to the value corresponding to the new long run equilibrium, and remains constant over the whole transition. By contrast, the ID index remains unaffected during the period of the shock, but adjusts in a smooth monotonic way afterwards. The GD index follows an intermediary evolution, as it jumps discretely in the period of the shock, and continues to slowly evolve, along with the long run transformation of the population pyramid. In the new equilibrium, the three indices are equal.

Importantly, these simulations show that the evolution of GD is not necessarily monotonic during the transition. This occurs because the relative size of young age cohorts in the current population pyramid does not evolve in a monotonic way. This property of the GD index is not necessarily desirable, as it indicates changes in total deprivation that follow from the mechanics of the demographic evolution. We now provide a stylized illustration of this questionable property.

Consider a stationary population with one individual born every year who lives exactly for 4 periods, with a mortality rate at age 3 equal to 1. The mortality vector is thus $\mu = (0, 0, 0, 1, \dots)$. We assume that the age threshold, $\hat{a}$, is equal to 12, and

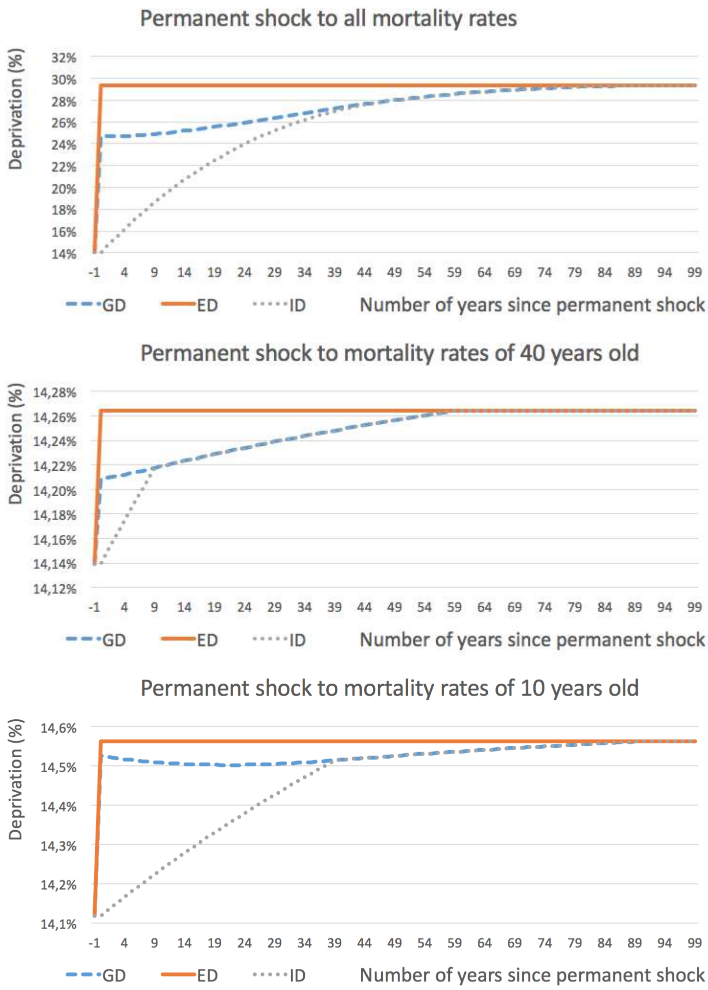


Figure 4: Simulation of permanent mortality shocks on a stationary population

$\gamma = 1$. There is no alive deprivation. The GD index for this situation is equal to 8/12, and is equal to the ID and ED indices. In period t^s, there is a permanent mortality shock such that the new mortality rate at age 1 is equal to 1. The new mortality vector is thus $\mu^s = (0, 1, 0, 1, \ldots)$. Table 3 summarizes the evolution of this population after this permanent shock.

Table 3: Non-monotonicity of GD indices after permanent mortality shock.

period	0	1	2	3	4	...	11	GD
$t < t^s$	NP	NP	NP	NP	D	...	D	$\frac{8}{12} = 0.66$
t^s	NP	NP	NP	NP	D	...	D	$\frac{18}{22} = 0.82$
$t^s + 1$	NP	NP	D	NP	D	...	D	$\frac{18}{21} = 0.86$
$t^s + 2$	NP	NP	D	D	D	...	D	$\frac{10}{12} = 0.83$

Two individuals die at the end of period t^s and the GD index records 18 PYPLs. Given that four individuals lived in period t^s, GD is equal to 18/22. In period $t^s + 1$, there is no individual of age 2, and one individual of age 0, 1 and 3. The GD index records again 18 PYPLs, but given that only three individuals were alive, GD is equal to 18/21. Two periods after the shock, the new stationary population is such that there are only two individuals alive, of age 1 and 2 respectively. There are 10 PYPLs, out of a total of 12, so that the GD index is equal to 10/12. Because of the mechanical adaptation of the population pyramid, the GD index increases in $t^s + 1$ but decreases in $t^s + 2$. By contrast, the ED index remains constant and equal to 10/12 over these three periods.

How should we think about the non-monotonic behavior of the GD index? This behavior reflects the evolution of the population. Indeed, the presence of the 3-years old individual in t^s+1 implies that the mortality vector μ^s does create more PYPLs in $t^s + 1$ than in $t^s + 2$. So the GD index conveys correct information about actual deprivation. However, a fixed mortality vector μ^s is related to fundamentals for a population's health situation. One should therefore not necessarily conclude from the evolution of the GD index that these fundamentals have necessarily changed. The non-monotonicity of the GD index creates a risk of misinterpretation of the evolution in the fundamentals. This example illustrates the undesirable consequences of the non-monotonicity: a situation in which one more person is prematurely dead instead of alive is considered better, subject to less deprivation, according to the GD index.

To sum up, when compared to the ED index, the GD index can be misinterpreted and requires additional information related to the current distribution of individuals by age. However, it also has a very interesting property, as it is decomposable in subgroups. The GD index measured on a set of individuals can always be calculated as the weighted sum of the same index measured on any partition of this set, where the weight attributed to a subset is the fraction of its reference population divided by the total reference population. This is an important property, whose relevance matters if one wishes to compare the relative deprivation of different groups in a society, such as men and women, black and white, old and young, rural and urban, south and north, etc. By contrast, the ED index is not decomposable, since life expectancy is itself non-decomposable.

3 Comparison with alternative approaches

In this section we compare our deprivation indices to the alternative measures proposed in the literature. This allows us to discuss some of the important assumptions underlying the construction of our indices.

Composite indices

As we show in the Introduction, composite indices are inconsistent with a basic separability property. We discuss this inconsistency. To do so, we define the mortality statistic $LD_{\hat{a}}(\mu) = \frac{LGE_{\hat{a}}(\mu)}{\hat{a}}$, which measures the fraction of the age threshold that a newborn expects not to live.[18] $LD_{\hat{a}}$ is an indicator of *lifespan deprivation*, and satisfies the basic properties of a deprivation index, when these properties are adapted to lifespan distributions. In particular, $LD_{\hat{a}}$ is unaffected by changes in mortality in old ages (above $\hat{a}$) but strictly decreases when mortality in young age is reduced. Also, $LD_{\hat{a}}$ decreases when the death of a lifespan deprived individual is postponed by one year.[19]

[18]See footnote 12 for references on $LGE_{\hat{a}}$.

[19]A formal proof is given in the Online Appendix 1. We show in the Appendix that both the GD index and the ED index violate one of the two basic properties of a lifespan deprivation indicator,

We define a *composite* index of total deprivation by simply weighing alive deprivation as measured by HC and lifespan deprivation as measured by $LD_{\hat{a}}$:

$$P_w^C(x, \mu) = w * HC(x) + (1 - w) * LD_{\hat{a}}(\mu),$$

where weight $w \in [0, 1]$.[20] As illustrated in the Introduction, this index is not consistent as it does not attribute a fixed relative weight to one PYPL compared to one PYAD. In other words, when comparing stationary pairs, the index violates a weakening of Subgroup Consistency that we call Separability of status comparisons. This axiom requires that, for a given birth year, the comparison of two alternative statuses does not depend on the remaining part of the distribution. The precondition on birth years guarantees that i belongs to the reference population.[21]

Deprivation axiom 8 (Separability of status comparisons)**.** *For all $x, x' \in X$, if $n(x) = n(x')$, $b_i = b_i' > t - \hat{a}$, then $P(x_i, x_{-i}) \geq P(x_i', x_{-i}) \Leftrightarrow P(x_i, x_{-i}') \geq P(x_i', x_{-i}')$.*

Consider the impact on P_w^C of a reform that increases the lifespan of individuals who would otherwise die prematurely, such that the additional time alive is spent in alive deprivation. Depending on the fraction of the alive population that is poor, a composite index P_w^C may increase or decrease.[22] The same inconsistency also characterizes the Human Deprivation Index, a composite index which aggregates both premature mortality and alive deprivation using the head-count ratio in both dimensions (Watkins, 2006).[23] By contrast, P_1^{ED} is never affected by such reform. Indeed, assuming $\gamma = 1$ implies weighing equally one PYAD and one PYPL. A higher value of γ always involves a decreasing value of our index after such reform.

The intuition behind the inconsistency of composite indices is that their two components consider different reference populations, i.e. different numbers of person-years. Their deprivation component divides the number of PYADs by the number of PY spent *alive* while their mortality component divides the number of PY spent alive by the normative lifespan. The implicit weight that a composite index attaches to one PYAD over one PYPL therefore depends on the levels of alive deprivation and life expectancy. The root of the problem is that composite indices first normalize each component using different reference populations and then take a weighed sum. In contrast, our total deprivation indices add the number of PYADs with the number of PYPLs *before* normalizing by the same reference population. As a result, the relative weight attributed to one PYPL over one PYAD remains fixed, as required by Separability of status comparisons.[24]

i.e. *Current Mortality Focus* or *Current Mortality Monotonicity below* $\hat{a}$.

[20] This index satisfies the basic properties of a lifespan deprivation indicator. These properties, *Current Mortality Focus* and *Current Mortality Monotonicity below* $\hat{a}$, are defined in the online Appendix 1.

[21] Even if index P_w^C is defined on domain O, we provide an axiom for indices defined on X for ease of exposition. The parallel axiom on O is: For all $((x_i, x_{-i}), \mu), ((x_i', x_{-i}), \mu'), ((x_i, x_{-i}'), \mu''), ((x_i', x_{-i}'), \mu''') \in O$ that are stationary, if $n(x) = n(x')$ and $b_i = b_i' > t - \hat{a}$, then $P((x_i, x_{-i}), \mu) \geq P((x_i', x_{-i}), \mu') \Leftrightarrow P((x_i, x_{-i}'), \mu'') \geq P((x_i', x_{-i}'), \mu''')$.

[22] This problem does not depend on the value of the parameter parameter w. For all possible values of the $w \in (0, 1)$, one can always find situations under which the composite index P_w^C is not consistent.

[23] The premature mortality dimension of the human deprivation index is measured by the probability to die before reaching 40 years in developing countries and 60 years in developed countries.

[24] Another difference between our indices and composite indices is that our total deprivation indices are a generalization of the alive deprivation index. In the absence of premature mortality,

Preference-based indicators and weight γ

We now compare our indices to the preference-based indicators used in Gary S. Becker and Soares (2005); Jones and Klenow (2016). In contrast to well-being indicators, deprivation indices disregard the actual achievements of non-poor individuals in order to focus on the fate of deprived individuals. Beyond this conceptual difference, the two approaches are not equivalent when it comes to applications. From the practitioner's point of view, applying our indices rely on fewer normative assumptions, as they only require selecting an age threshold and a value for the weight γ. By contrast, preference-based indicators are based on explicit preferences and, therefore, on a particular utility function.

Moreover, focusing on deprivation instead of well-being naturally leads us to use a fixed weight γ. Consider HDI, a composite index measuring well-being. It is well-known that the monetary value of one extra year of life implicitly attached by the HDI is higher for richer countries (Ravallion, 2011). This non-separability of the HDI is not problematic per se as it reflects the higher opportunity cost of dying in richer countries (a similar remark also holds for lifetime utility approaches (Gary S. Becker and Soares, 2005; Jones and Klenow, 2016).) Things are very different when the two dimensions being compared are deprivations (PYPLs and PYADs) rather than achievements (GDP and LE). Indeed, deprivations are assumed to involve similar trade-offs and, therefore, carry the same relative weight in all countries. There is indeed a priori no reason to trade-off differently these dimensions according to the observed levels of alive deprivation and life expectancy, and Separability of status comparisons is a natural requirement.

In the empirical application presented below, we shall assume $\gamma = 1$, as we believe that $\gamma = 1$ is a reference value of particular interest. First, it is a conservative choice if we believe that one PYPL is at least as bad as one PYAD, which requires $\gamma \geq 1$. A revealed preference argument supports $\gamma \geq 1$ given that committing suicide is an outside option (plausibly) available. More generally, this inequality is relevant as long as the fraction of "young" individuals who prefer to be dead instead of poor is quantitatively negligible. Finally, if we believe that $\gamma \geq 1$, then any disagreement between HC and our indices when $\gamma = 1$ is robust to taking a larger value of γ (see Lemma 2 in the Online Appendix).

Second, when $\gamma = 1$, one PYPL and one PYAD have the same weight. As a result, computing the index is a simple accounting exercise, which consists in measuring the fraction of person-years that are either spent in alive deprivation or prematurely lost. The interpretation of the ED index becomes straightforward as we can then write it as:[25]

$$P_1^{ED}(x, \mu) = \frac{LE(\mu) * HC(x) + LGE_{\hat{a}}}{LE(\mu) + LGE_{\hat{a}}(\mu)}.$$

This index takes the perspective of a newborn that would be confronted throughout

our total deprivation indices is identical to alive deprivation, as measured by HC.

[25] It may seems odd that the ED index combines a head-count ratio for alive deprivation, which captures the incidence of income deprivation, with a form of Poverty Gap ratio for lifespan deprivation, which captures the depth of lifespan deprivation. However, the unit used is not defined in terms of individuals, but in terms of person-years. An individual who is poor in period t loses one person-year while an individual who dies prematurely in period t loses as many person-years as the difference between her age and the age threshold.

her life to the mortality rates and alive deprivation rates observed in period t. The first term of the numerator measures the number of years that a newborn may expect to spend in alive deprivation. The second term measures the number of years that a newborn expects to lose prematurely. The denominator measures the *augmented life expectancy* of a newborn, which is the life expectancy of a newborn who is confronted to a mortality vector $\mu^{\hat{a}}$ constructed from μ by postponing all premature deaths in μ to the age threshold.[26] Index P_1^{ED} measures the fraction of years in her augmented life expectancy that a newborn expects to lose to deprivation given the alive deprivation and mortality observed in period t.

Deprivation measures improving on the mortality paradox

As noted in the Introduction, our approach differs from the literature on the mortality paradox (Kanbur and Mukherjee, 2007; Lefebvre et al., 2013, 2017) which proposes various methods to assign fictitious incomes to missing individuals. One such method assigns fictitious incomes regardless of the pre mortem income of missing individuals (e. g. Lefebvre et al. (2013, 2017)). This idea can be applied in our constrained information setup. However the definition of a missing poor used there is conceptually very different from ours as it relies on a reference mortality vector, corresponding to that of the most affluent societies such as Norway or the US. In this perspective, the missing population is defined as those individuals who died *in excess* with respect to the reference mortality vector. As a result, not all individuals dying early are considered as missing individuals and an 80-year-old individual dying in excess can be considered missing while this may not be true for a 5-year-old child (as long as the reference society also present some form of child mortality). Our deprivation approach does not rely on such a reference mortality vector.

The fictitious incomes assigned may also depend on the incomes earned before dying. Thus, Kanbur and Mukherjee (2007) attribute to rich individuals dying prematurely fictitious incomes that are above the deprivation threshold. In our approach, we do not distinguish between the premature mortality affecting the poor and that affecting the non-poor. As noted in the Introduction, the necessary information on the mortality rates of different income groups is not always available. More fundamentally, the availability of such information is not sufficient to solve the underlying normative issue, which was raised by the literature on multidimensional poverty (Alkire and Foster, 2011): there is more overall poverty when the same individuals concentrate several dimensions of deprivation. In this respect, the premature mortality of poor individuals constitutes such a non-desirable concentration of deprivations, and to address this question, we should distinguish mortality rates of poor and non poor individuals. To make our ED index sensitive to concentration, we can define an individual as being in *total poverty* if she spends more than k person-years in deprivation, either in the form of PYPLs or PYADs. Our indices can therefore be accommodated to allow for this type of approach.[27] However, to compute such concentration-sensitive indices of total poverty, we not only need mortality rates by

[26] Mortality vector $\mu^{\hat{a}}$ is constructed from μ by letting $\mu_a^{\hat{a}} = 0$ for all $a \in \{0, \ldots, \hat{a}-2\}$, $\mu_{\hat{a}-1}^{\hat{a}} = 1 - \Pi_{a=0}^{\hat{a}-1}(1-\mu_a)$ and $\mu_a^{\hat{a}} = \mu_a$ for all $a \in \{\hat{a}, \ldots, a^*\}$.

[27] Such a definition of total poverty is consistent with the definition of multidimensional poverty proposed by (Alkire and Foster, 2011): an individual is multidimensionally poor if she is deprived in at least k dimensions.

income groups but also *information on mobility in and out alive deprivation across consecutive periods*, a type of information which is typically not available.[28]

4 Aggregate deprivation

4.1 The data

We apply our indices of total deprivation to real world data. In the process, we compare our results to those using more conventional deprivation measures, such as the HC.[29] The definition of ED requires a value for the age threshold $\hat{a}$ and the weight γ. As already discussed, the latter will be set conservatively at 1, so that one person-year prematurely lost is equivalent to one person-year spent in income deprivation. Choosing a higher value for γ, by increasing the weight given to the mortality component, would simply magnify the difference between the ED index and more traditional deprivation measures.

The choice of the age threshold is essentially similar to the choice of an income threshold used for income deprivation. It is ultimately a normative choice about the minimum number of years of life that a society judges essential for its members. In the following, we use a threshold $\hat{a} = 50$ years, which is much lower than the median age at death observed in our data (64 years old). Of course, a higher age threshold would inflate our indices and their difference with income deprivation measures. The robustness of our empirical findings was tested by using alternative thresholds (40 and 60 years), and we present these alternative results in the Online Appendix of the paper. In our code, available online, the reader can herself chose her own threshold of $\hat{a}$ as well as for γ.

The computation of the ED index requires information on alive deprivation as well as information on mortality and population by age. Ideally, this information should be comparable across countries and over time. In the following, we make use of two publicly available data sets to construct our measures of deprivation. The data on population and mortality by country, age group and year comes from the Global Burden of Disease database (2016 version of the data). It is available for the 1990-2016 period and is, to our knowledge, the most comprehensive mortality data available for international comparison. To construct this database, population and mortality data are systematically recorded across countries and time from various data sources (from official vital statistics data, to fertility history data as well as to data sources compiling deaths from events such as wars and other catastrophic events). These primary data are then converted into data at the age group, year and country level using various interpolations and inference methods.[30] Details on the method used are given in the appendix of Mortality and of Death Collaborators

[28] Note that, when mobility is very low and premature mortality is mostly concentrated on poor individuals, our indices approximately count the number of person-years lost to deprivation by the poor.

[29] Remember that to construct ID measures, we would need information on the number of death by age in the past $\hat{a}$ years. Such information does exist, for example via the Human Mortality Database (https://www.mortality.org/), but the countries and years available in this database are very different from those for which comparable alive deprivation data is available.

[30] Therefore, the number of deaths in each cell is an estimate and comes with a confidence interval. Following the convention in the literature, we do not use these confidence intervals, and only consider the point estimate of the number of death. (See also Hoyland et al. (2012) for a critique of this approach).

(2016).[31]

Our data on alive deprivation come from the PovcalNet website which provides internationally comparable estimates of income deprivation level. This data set is based on income and consumption data from more than 850 representative surveys carried out in 127 low- and middle-income countries between 1981 and 2014.[32] Each country's income deprivation level in PovCalNet is computed on a three yearly basis, so that the yearly data used below were obtained by linear interpolation of the income deprivation estimates across years. A complete description of the data set is given in Chen and Ravallion (2013).[33] In our empirical application, we follow the World Bank's definition of extreme income deprivation, corresponding to the 1.9$ a day threshold (Ferreira et al., 2016).

To compute the GD and ED indices, we merged the two databases at the year and country level. The Global Burden of the Disease data are only available since 1990 and the PovCalNet data for low and middle income country, until 2014. As a result, we focus in the following on the 1990-2014 period for a total of 124 low- and middle-income countries, representing 79% of the World population in 2014 (see Online Appendix 2 for a list of those countries).

4.2 World deprivation

We begin with the GD approach. Remember that a feature of this approach is that it relies on the computation of the total number of years of deprivation generated in a given year. Table 4 presents this computation and decomposes it into lifespan and income deprivation for the year 2014. *In 2014, 1,070 billion person-years of deprivation have been generated, 680 million from income deprivation and 390 from lifespan deprivation. That is, 17.6% of the person-years of 2014 were lost to deprivation.*

Table 4: Generated Deprivation in the developing world in 2014, with $\hat{a} = 50$.

	$PYADs$ $person - years$ $(millions)$	$PYPLs$ $person - years$ $(millions)$	$PYADs + PYPLs$ $person - years$ $(millions)$	P_1^{GD} %
World	680	390	1,070	17.6

The ED approach takes the perspective of a newborn: how many years of live does a newborn expect to spend deprived? Table 5 presents this approach by computing the ED index for the World in 2014. That year, 11.9% of the world's population was extremely poor and a newborn had a life expectancy of 69.1 years. Hence, such a newborn expected to spend 8.2 years in extreme poverty. The same newborn had a lifespan gap expectancy of 3.8 years. *In total, in 2014, newborns expected to spend 12*

[31] Moreover, the mortality information is given into 5 year age brackets (except for the 0-5 years group, for which the information is decomposed into 0-1 and 1-5). When necessary, we transform the data into age groups of one year by assuming a uniform death rate within an age category. Finally, the older age group is "95 and above". As we do not know the precise age of death of individuals in that category, we assume that 95 is the maximum age they can reach. This last assumption is of no consequence here, since our age threshold $\hat{a}$ is well below 95.

[32] The website address is http://iresearch.worldbank.org/PovcalNet/povOnDemand.aspx.

[33] Clearly, these transformations may matter for the empirical analysis, as they tend to smooth the evolution of income deprivation across years. In particular, in the case of catastrophic events such as earthquakes or tsunami, income deprivation appears as less reactive then lifespan deprivation, which may be due to the interpolated nature of the data. In the following, we therefore do not analyze these events.

years either in alive deprivation or in lifespan deprivation. Therefore, the ED index is 16.5%, which is the share of her "augmented" life expectancy that a newborn expects to spend in deprivation.

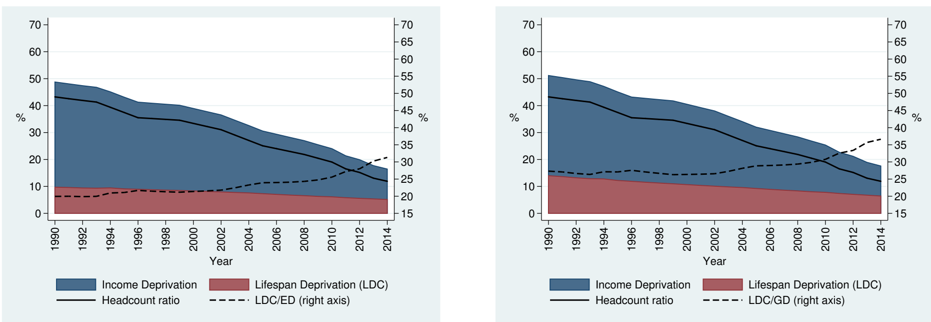
Table 5: Expected Deprivation in the developing world in 2014.

	HC %	LE *years*	$E(PYADs)$ *years*	LGE_{50} *years*	$E(PYADs) + LGE_{50}$ *years*	P_1^{ED} %
World	11.9	69.1	8.2	3.8	12	16.5

Figure 5 presents the evolution of world's total deprivation, as measured by GD and ED, and of their two components, the alive deprivation component (income deprivation, in our case) and the lifespan deprivation component (LDC). We also report the HC for comparison purposes. A first major point is that the lifespan deprivation component is far from negligible.

In 1990, according to ED, a newborn expected to spend 10% of his "augmented" life expectancy in lifespan deprivation, as compared to 39% in income deprivation: lifespan deprivation thus represented 20% of total deprivation. In other words, if one were to focus on alive deprivation only, it would result in an underestimation of total deprivation of 20% in 1990[34]. Also, the relative importance of lifespan deprivation increased over time: its share in total deprivation increased from 20% in 1990 to more than 30% in 2014. Therefore, focusing on alive deprivation only leads to an underestimation of total deprivation which is both substantial and growing. Note that given our conservative choice of parameters, these estimates can be considered as lower bounds: in 2014, total deprivation is underestimated by *at least* 30% if lifespan deprivation is neglected. A similar result is obtained when using the GD index, for which the share of lifespan deprivation in total deprivation is even larger and increased from 27% to 37%. This increase in the share of lifespan deprivation indicates that much more progress has been made against alive deprivation as against lifespan deprivation over the past 25 years. One can only wonder if that would have been the case had premature mortality systematically been taken into account in deprivation measures.

Figure 5: Decomposition of Total Deprivation with $\hat{a} = 50$, World Level



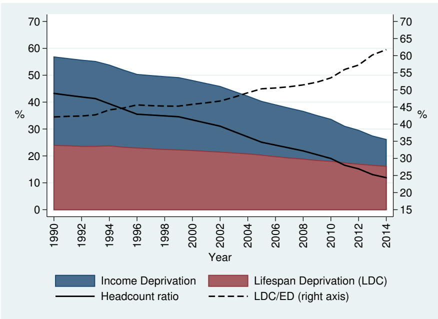
Expected Deprivation at birth with $\hat{a} = 50$ — Generated Deprivation with $\hat{a} = 50$

A second point to note is that these three measures follow parallel trends. They do

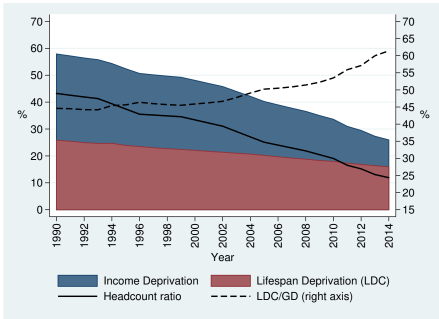
[34] Note that HC and ED can not be directly compared, since their denominator is different. However, comparing the lifespan deprivation component of ED to ED enables to rigorously evaluate how large is the underestimation of deprivation when premature mortality is not taken into account.

not offer a different diagnostic about the evolution of world deprivation in the last 25 years. World deprivation fell dramatically between 1990 and 2014: while a newborn in 1990 could expect to spend 49% of his "augmented" life expectancy in deprivation, this proportion fell down to 16% in 2014. Note that these results derive from the very conservative choice of 50 years as the lifespan deprivation threshold. One may wonder how total deprivation would evolve if a more ambitious threshold had been set. In Figure 6, we reproduce Figure 5, but using 80 instead of 50 as a threshold. It is striking to see that with such a threshold, lifespan deprivation constitutes a very important component of total deprivation even in the early 1990's, representating more than 40% in total deprivation at the beginning of our period and more than 60% at the end of our period. By construction, deprivation rates are also much higher, but note as well how they decrease less rapidly over time.

Figure 6: Decomposition of Total Deprivation with $\hat{a} = 80$,World Level



Expected Deprivation at birth with $\hat{a} = 80$



Generated Deprivation with $\hat{a} = 80$

Finally, both in terms of levels and in terms of trends, GD and ED offer very similar diagnostics. For the sake of simplicity, we will therefore focus in the following on ED. We reproduce the graphs presented in the following sections using GD instead of ED in the Online Appendix. The main difference between GD and ED is that GD gives more weight to premature mortality because the number of newborns in the developing world has been increasing over time. This implies that the counterfactual population pyramids considered by ED have relatively smaller cohorts affected by premature mortality than the actual population pyramids considered by GD. The choice of ED for the results presented in the main text can therefore be considered as conservative.

4.3 Regional deprivation

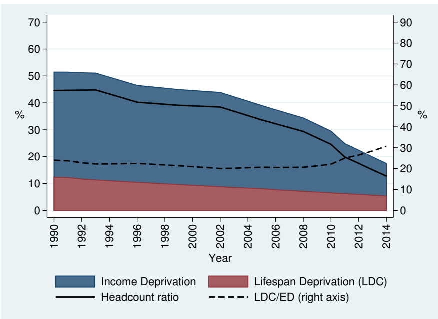
Figure 7 compares the evolution of deprivation for the six World Bank regions as measured by HC and the ED index. The regional diagnostic is similar to that of the global one. Indeed, focusing on alive deprivation leads to a large underestimation of deprivation in all regions. This is particularly the case of the Middle East and North Africa and of East Asia and the Pacific, for which alive deprivation has been almost eradicated while total deprivation remained non negligible.[35] In the latter region for instance, the share of lifespan deprivation in total deprivation increased substantially from 10% in 1990 to 46% in 2014. In addition, for all regions but Sub Saharan Africa,

[35]In our data, the Middle East and North Africa region only includes Djibouti, Iran, Morocco and Tunisia.

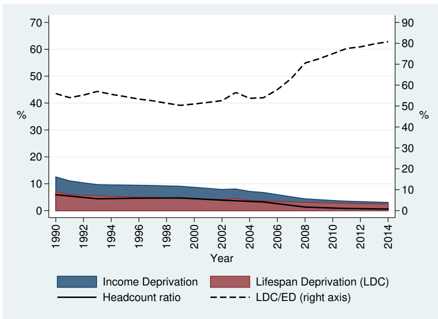
this underestimation is growing over time since the 2000's (Middle East and North Africa, Latin America and Caribbean, Europe and Central Asia) or the 2010's (South Asia, East Asia and Pacific). For Sub Saharan Africa, the underestimation, while substantial (around 20%) remain stable throughout. More generally, while lifespan deprivation decreases smoothly across all regions, the evolution of alive deprivation varies much more across regions and across periods. This latter finding parallels and complements the well-documented fact that GDP inequalities across countries are larger than inequalities in life expectancy (see for instance Jones and Klenow (2016)).

In addition, note that the relative position of each region as measured with ED is similar to that mesured with HC: the income-poor regions are also the most deprived ones.

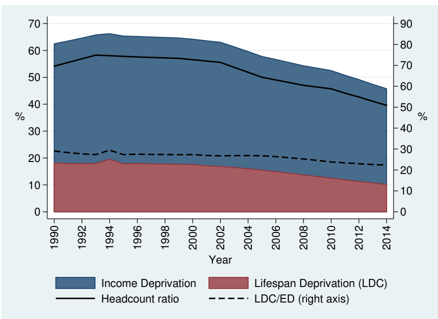
Figure 7: Decomposition of Total Deprivation, Continental Level. ED with $\hat{a} = 50$



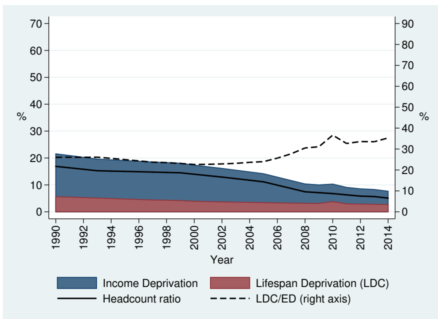
South Asia



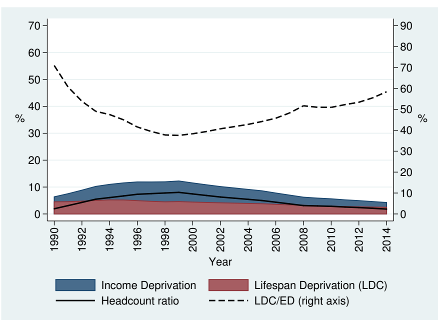
Middle East and North Africa



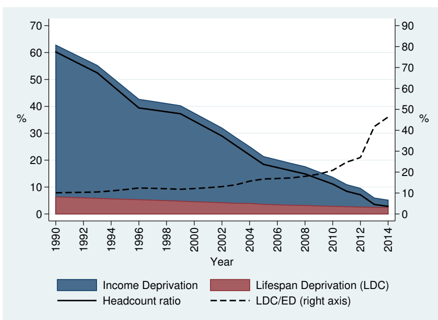
Sub Saharan Africa



Latin American and Caribbean



Europe and Central Asia



East Asia and Pacific

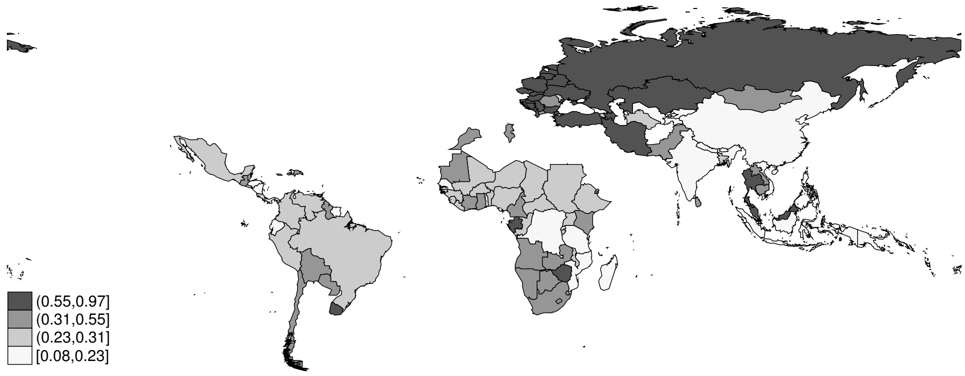
5 Countries' deprivation

5.1 Countrie's levels of deprivation

We now investigate deprivation for individual countries. The extent and the evolution of deprivation at the country level measured by ED can substantially differ from those described with more traditional measures of deprivation. We will show that focusing on alive deprivation biases our understanding of individual countries' deprivation: in terms of level (some countries are much more deprived than we thought), in terms of trajectories (countries which we thought were doing better in terms of deprivation may actually be doing worse) and in terms of international ranking.

Figure 8 maps the median value[36] of each country's yearly share of the lifespan deprivation component in ED over 1990-2014, that is, the extent to which omitting premature mortality leads to an underestimation of deprivation in that country. This underestimation is particularly pronounced in the ex-USSR countries as well as in Turkey or Iran, where lifespan deprivation represents at least 55% of ED. In these countries, premature mortality is an essential component of deprivation, and its omission leads to a very severe underestimation of deprivation.

Figure 8: Map of LD/ED ratio, 1990 to 2014 ($\hat{a} = 50$)



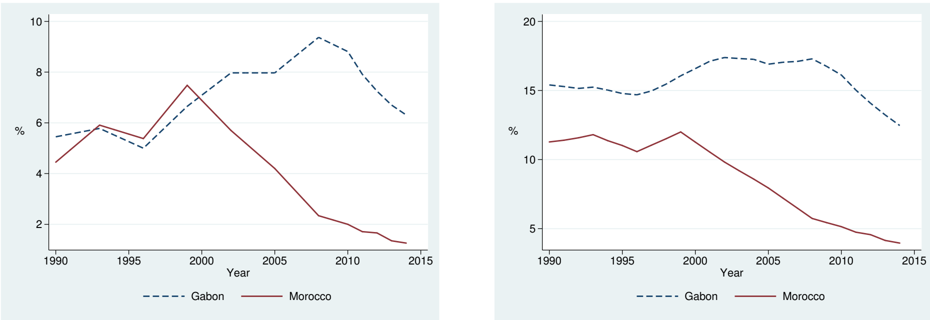
Median LD/ED yearly ratio,1990 and 2014.
Reading: the yearly LD/ED ratio in the Russian Federation was at least 55% at least 50% of the time in the 1990-2014 period.

Individual countries are often compared and ranked according to a measure of deprivation (Hoyland et al., 2012). In the following, we investigate the extent to which the use of our measures of deprivation changes our understanding of the ranking of countries. Figure 9 provides the example of Morocco and Gabon. According to HC, throughout the 1990s, Gabon and Morocco are virtually at the same level of deprivation. However, deprivation was much higher in Gabon once lifespan deprivation is taken into account. Ranking countries according to deprivation (starting from the least poor), Gabon was ranked 32^{nd} and Morocco 33^{rd} in 1993. When the ranking is based on ED, Gabon was 46^{th} and Morocco 38^{th}. Table 6 decomposes the sources of this re-ranking in 1993. While both countries have a similar level of alive deprivation (a HC close to 6%), their mortality differs widely. Indeed, in 1993, the life expectancy at birth in Morocco was 67 years, against 59 years in Gabon. Hence, a newborn in Morocco expects to live $67{*}5.9\% = 3.9$ years in alive deprivation against

[36] We use the median rather that the mean to prevent extreme values to bias the general picture.

3.4 in Gabon. However, as early mortality is higher in Gabon than in Morocco, a newborn in Gabon expects to loose 6.6 years of life, as against 4.5 in Morocco. Total expected years lost to deprivation therefore amount to 8.4 in Morocco against 10 in Gabon, while augmented life expectancy is respectively 71.5 and 65.6. A newborn in Morocco is therefore expected to lose 11.8% of her "augmented" life expectancy to deprivation, as against 15.2% in Gabon.

Figure 9: Examples of re-rankings: Gabon and Morocco. HC and ED with $\hat{a} = 50$



Head-count ratio

Total Deprivation

Table 6: Decomposition of re-rankings: Gabon and Morocco in 1993.

Country	HC %	LE *years*	$E(PYADs)$ *years*	LGE_{50} *years*	$E(PYADs) + LGE_{50}$ *years*	P_1^{ED} %
Morocco	5.9	67	3.9	4.5	8.4	11.8
Gabon	5.8	59	3.4	6.6	10	15.2

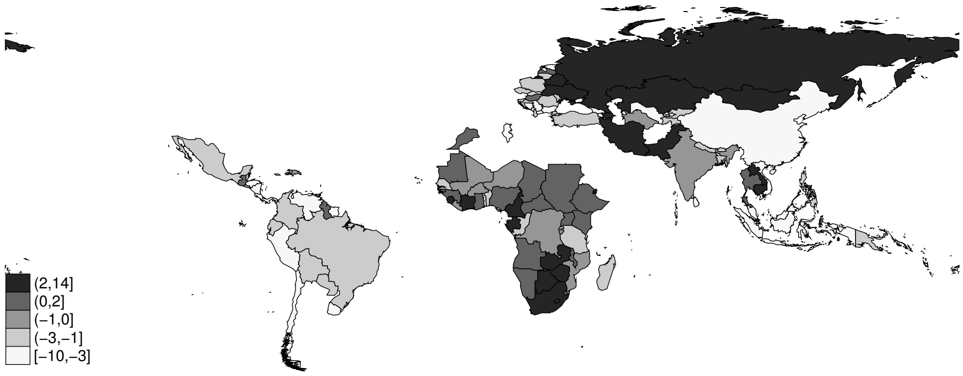
More generally, our measures lead to substantial re-rankings across countries. Indeed, throughout the period, the average change in ranking across all countries is equal to 3.4 ranks. How are these changes distributed across countries and time? Figure 10 reports each country's median change in rank during the period. These changes can be particularly important: countries of the ex-USSR and a few African countries lose up to 14 ranks while some Latin American countries improve their ranking substantially.

How do these re-rankings evolve over time? Figure 11 reports for each year the box plot of the absolute value of the change in the ranking of all countries when it is based on ED instead of HC. The figure clearly indicates that re-ranking between countries is more frequent and larger over time. This is due to the increasing importance of premature mortality in total deprivation, and implies that the relevance of rankings based on total deprivation instead of income poverty increases over time.

5.2 Countrie's trajectory of deprivation

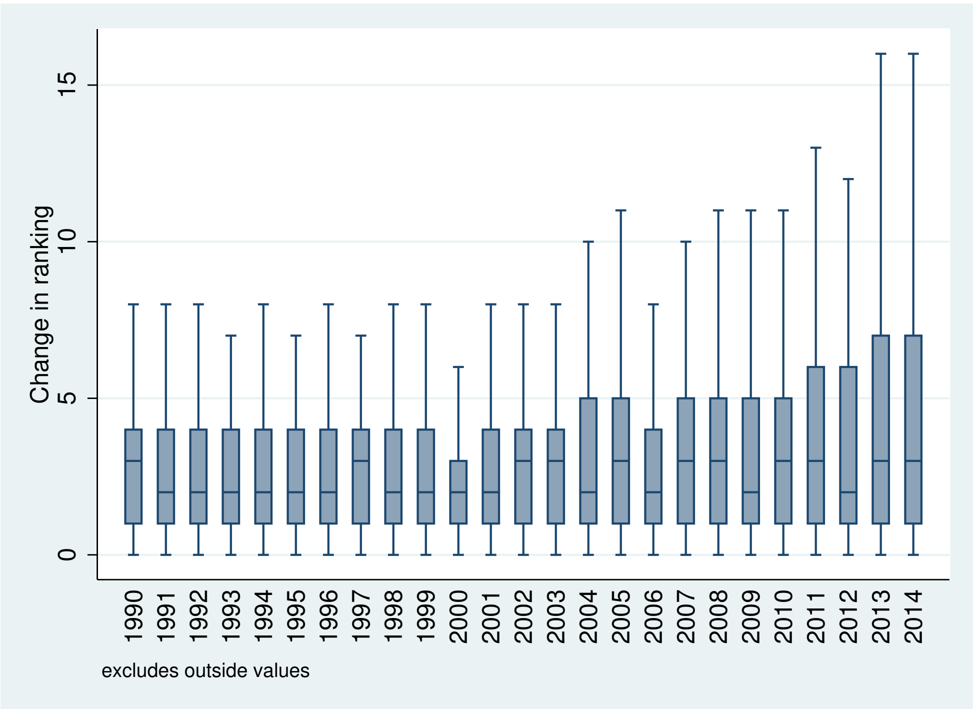
Our indices also change our assessment of the evolution of deprivation in a given country. Let us take, for instance, the cases of the Comoros and of Botswana. Figure 12 presents the evolution of HC and ED for these two countries. In the Comoros throughout the period, HC increased while ED decreased, due to the important progress made against premature mortality which more than compensates for the in-

Figure 10: Map of re-rankings of countries. ED with $\hat{a} = 50$ vs HC



Median absolute rank change between 1990 and 2014.
Reading: the Russian federation is ranked 2 to 14 ranks lower with ED than with HC at least 50% of the time in the 1990-2014 period.

Figure 11: Evolution of re-rankings of countries: 1990-2014. ED with $\hat{a} = 50$

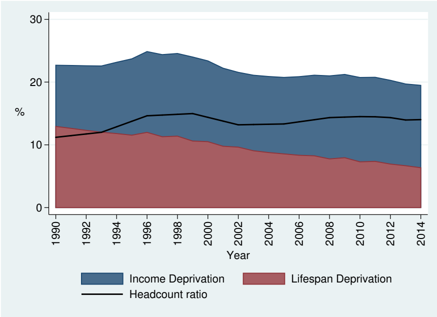


Reading: in 1990, the 25th percentile was a change in 1 rank, the 75th in 4 ranks and the median 3 ranks, the upper adjacent value was 7 and the lower adjacent value 0.

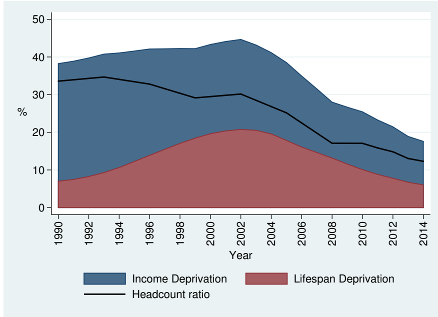
crease in alive deprivation. By contrast, in Botswana during the 1990s, HC decreased substantially, while ED increased dramatically due to the HIV epidemics.

In table 7, we present a more detailed analysis of these cases. According to HC, deprivation increased in the Comoros by 25% between 1990 and 2014. However, according to ED, total deprivation fell by 14% during that same period. Focusing on alive deprivation hides the large progress made in lifespan deprivation: life ex-

Figure 12: Differences in trends between ED and HC: Comoros and Botswana. HC and ED with $\hat{a} = 50$



Comoros



Botswana

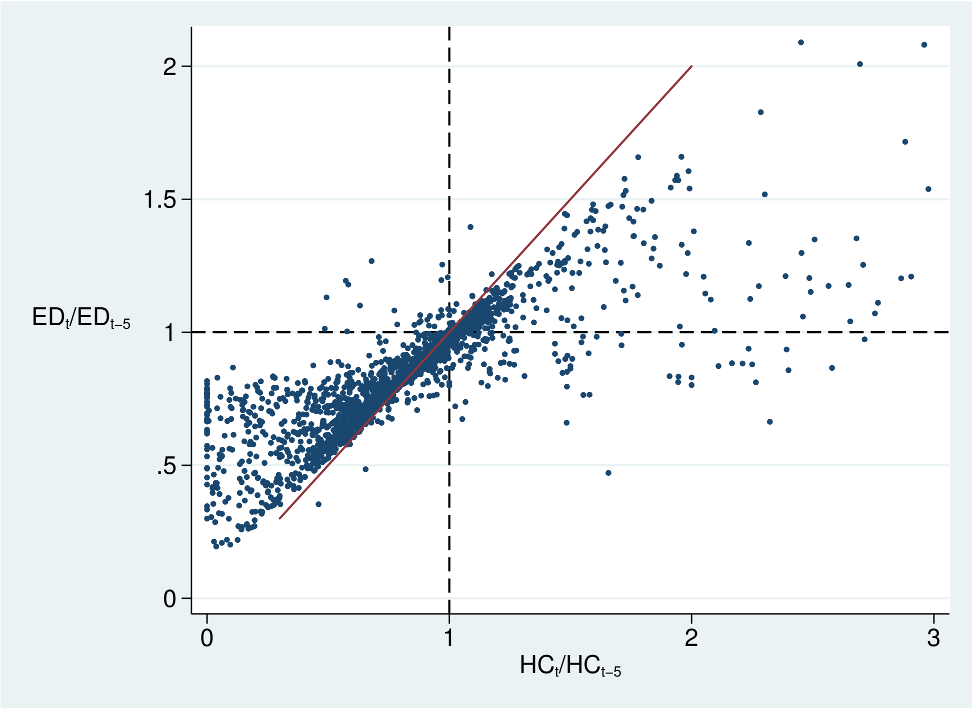
pectancy increased from 57 to 66 years and the lifespan gap expectancy decreased from 8.5 to 4.5 years. Conversely, in 1990, a newborn would expect to spend 6.4 years in alive deprivation, as against 9.2 years in 2014. Overall, the number of years spent in deprivation decreased from 14.9 to 13.7 and the augmented life expectancy increase from 65.5 to 70.5 years. Botswana evolved very differently between 1990 and 2000: HC decreased by 12% during that period while ED increased by 13%.

Table 7: Example of evolution reversals: Comoros and Botswana

Year	HC %	LE *years*	$E(PYADs)$ *years*	LGE_{50} *years*	$E(PYADs) + LGE_{50}$ *years*	P_1^{ED} %
				Comoros		
1990	11.2	57	6.4	8.5	14.9	22.7
2014	14.0	66	9.2	4.5	13.7	19.5
	+ 25%	+ 12.6	+2.8	-4.0	-1.2	- 14%
				Botswana		
1990	33.6	63.6	21.5	4.8	26.3	38.2
2000	29.5	45.6	13.5	11.1	24.6	43.3
	- 12%	- 11	-8	+ 6.3	-1.7	+ 13%

How often do these opposite diagnostics arise in the last 25 years? In Figure 13, we plot the ratio of the value of ED in year t relative to its value in t-5 for each country in our sample against that for HC. As indicated by the figure, overall, the two measures generally agree. For most countries and periods, a decrease (increase) in HC is accompanied by an increase (decrease) in ED. Note that the relation between the two measures is flatter than the 45° line, which indicates that HC varies more than ED, owing to the greater inertia of lifetime deprivation. However, the two measures do not always agree, as attested by the large number of points located in the North-West and in the South-East quadrants. These points represent 7.4% of the comparisons made: in these cases, the diagnostic of deprivation based on deprivation among the living is so biased that the sign of its evolution is wrong. Note that this result relies on the conservative assumption $\gamma = 1$. This percentage tends to 26.2% as $\gamma \to \infty$.

Figure 13: Deprivation trends. HC and ED with $\hat{a} = 50$, t to (t-5) ratios.



Countries for which the deprivation rate increased by more than 200% are dropped from this figure.[37]

6 Concluding remarks

Most measures of poverty or deprivation ignore premature mortality. In this paper, we propose three measures of "total deprivation" that combine meaningfully information on income poverty and early mortality in a population, by adding time units spent in income poverty and time units of life lost due to premature mortality. This additive approach follows from the exclusive nature of the two dimensions considered, income poverty and premature death. We characterize our proposed measures, show that they satisfy a number of desirable properties, and contrast their implications with existing multidimensional indices, such as the MPI index of the World Bank.

Our aggregation method allows placing an explicit and meaningful lower bound on the normative trade-off (the weight γ) between premature mortality and poverty. This lower bound is based on the view that being prematurely dead is no better than being in alive deprivation ($\gamma \geq 1$). Using this conservative approach, our empirical results show that ignoring premature mortality regularly leads to biased evaluations in the level and in the evolution of deprivation. Their frequency is increasing over time due to the relative importance of premature mortality. Therefore, our results suggests that lifespan deprivation should be integrated in deprivation measusures.

References

Alkire, S. and Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics*, 95(7-8):476–487.

Anderson, S. and Ray, D. (2010). Missing women age and disease. *Review of Economic Studies*, 77:1262–1300.

Chen, S. and Ravallion, M. (2013). More Relatively-Poor People in a Less Absolutely-Poor World. *Review of Income and Wealth*, 59(1):1–28.

Deaton, A. (2013). *The great escape: health, wealth, and the origins of inequality.*

Ferreira, F. H., Chen, S., Dabalen, A., Dikhanov, Y., Hamadeh, N., Jolliffe, D., Narayan, A., Prydz, E. B., Revenga, A., Sangraula, P., Serajuddin, U., and Yoshida, N. (2016). A global count of the extreme poor in 2012: data issues, methodology and initial results. *Journal of Economic Inequality*, 14(2):141–172.

Foster, J. and Shorrocks, A. (1991). Subgroup Consistent Poverty Indices. *Econometrica*, 59(3):687–709.

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

Gary S. Becker, T. J. P. and Soares, R. R. (2005). American Economic Association The Quantity and Quality of Life and the Evolution of World Inequality. *The American Economic Review*, 95(1):277–291.

Grimm, M. and Harttgen, K. (2008). Longer life, higher welfare? *Oxford Economic Papers*, 60:193–211.

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

Hoyland, B., Moene, K., and Willumsen, F. (2012). The tyranny of international index rankings. *Journal of Development economics*, 97(1):1–14.

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.

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

Lefebvre, M., Pestieau, P., and Ponthiere, G. (2017). FGT Old-Age Poverty Measures and the Mortality Paradox: Theory and Evidence. *Review of Income and Wealth*, 33(0):0–30.

Mortality, G. . and of Death Collaborators, C. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specifi c mortality for 249 causes of death, 1980-2015: a systematic analysis for the global burden of disease study 2015. 388:1459–1544.

Pattanaik, P. and Xu, Y. (2018). On measuring multdimensional deprivation. *Journal of Economic Literature*, 56(2):657–672.

Preston, S., Heuveline, P., and Guillot, M. (2000). *Demography: measuring and modeling population processes.*

Ravallion, M. (2011). Mashup indices of development. *The World Bank Observer*, 27(1):1–32.

Sen, A. (1998). Mortality as an indicator of economic success and failure. *Economic Journal*, 108(January):1–25.

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

Young, H. P. (1975). Social choice scoring functions. *Journal of Applied Mathematics*, 28(4):824–838.

7 Appendix 1: Proofs

7.1 Proof of Proposition 1

We first provide the formal definition of the three axioms left undefined in the main text.

Deprivation axiom 9 (Anonymity)**.** *For all $x \in X$, if $n(x') = n(x)$ and x' is obtained from x by a permutation of the index set $\{1, \ldots, n(x)\}$, then $P(x) = P(x')$.*

For any $k \in \mathbb{N}$, we denote by x^k the k-**replication** of x, which is the distribution such that $n(x^k) = kn(x)$ and $x^k = (x, x, \ldots, x)$.

Deprivation axiom 10 (Replication Invariance)**.** *For all $x \in X$ and $k \in \mathbb{N}$, $P(x^k) = P(x)$.*

Deprivation axiom 11 (Young Continuity)**.** *For all $x, y, z \in X$, if $P(x) > P(y)$ and $n(z) = 1$, then for $k \in \mathbb{N}$ sufficiently large we have $P(x^k, z) > P(y)$ and $P(x) > P(y^k, z)$.*

It is easy to check that the ID index satisfies the seven axioms, so that the proof of necessity is omitted. Herebelow, we concentrate on the proof of sufficiency.

Let $\mathbb{Q}_+$ denote the set of non-negative rational numbers. Consider Δ, the 2-simplex on rational numbers, i.e. $\Delta = \{v \in \mathbb{Q}^3_+ \mid v_1 + v_2 + v_3 = 1\}$.

Step 1: Construct a mapping $m : X \to \Delta$ such that $m(X) = \Delta$ and for any two $x, x' \in X$, if $m(x) = m(x')$ then $P(x) = P(x')$.

We construct mapping m as the composition of four mappings, i.e. $m(x) = m^4 \circ m^3 \circ m^2 \circ m^1(x)$.[38]

First, mapping m^1 removes individuals who are not in the reference population. Let X^* be the subset of distributions that do not have any individual who is born at least $\hat{a}$ years before t and is dead, i.e. $X^* = \{x \in X \mid b_i > t - \hat{a}$ for all i for whom $s_i = D\}$. Let mapping $m^1 : X \to X^*$ return for any $x \in X$ the image $x^* = m^1(x)$ with $n(x^*) = f(x) + p(x) + d(x)$ and for any $i \leq n(x^*)$ the i^{th} component of x^* is defined as $x^*_i \equiv x_j$, where j is the i^{th} individual in x for whom either $s_i \neq D$ or $s_i = D$ and $b_i > t - \hat{a}$. By the definition of mapping m^1, we have for all $x \in X^*$ that $m^1(x) = x$. Hence, $m^1(X) = X^*$. Also, any two $x, x' \in X$ for which $m^1(x) = m^1(x')$ are such that $P(x) = P(x')$ by Weak Independence of Dead and Anonymity.

Second, mapping m^2 removes the birth year of individuals. Recalling that $S = \{NP, AP, D\}$, let mapping $m^2 : X^* \to \cup_{n \in \mathbb{N}} S^n$ return for any $x \in X^*$ the image $o = m^2(x)$ with $n(o) = n(x)$ and for any $i \leq n(o)$ the i^{th} component of o is defined from $x_i = (b_i, s_i)$ as $o_i \equiv s_i$. By construction of m^2, we have $m^2(X^*) = \cup_{n \in \mathbb{N}} S^n$. By construction of m^1, all dead individuals in a distribution $x \in X^*$ are *prematurely* dead. Hence, if $o_i = D$, then i is *prematurely* dead. Therefore, any two $x, x' \in X^*$ for which $m^2(x) = m^2(x')$ are such that $P(x) = P(x')$ by Weak Independence of Birth Year.

[38] The composite mapping m is defined a $m(x) = m^4(m^3(m^2(m^1(x))))$.

Third, mapping m^3 counts the number of individuals exhibiting each status. Consider the set $\mathbb{N}_0^3\backslash_{(0,0,0)}$, which contains all triplets of numbers in $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ except the nul triplet $(0, 0, 0)$. Let the mapping $m^3 : \cup_{n\in\mathbb{N}} S^n \to \mathbb{N}_0^3\backslash_{(0,0,0)}$ return for any $o \in \cup_{n\in\mathbb{N}} S^n$ the image $w = m^3(o)$ such that $w_1 \equiv \#\{i \leq n(o) \mid o_i = NP\}$, $w_2 \equiv \#\{i \leq n(o) \mid o_i = AP\}$ and $w_3 \equiv \#\{i \leq n(o) \mid o_i = D\}$.[39] By construction, we have $m^3 \circ m^2(X^*) = \mathbb{N}_0^3\backslash_{(0,0,0)}$. Also, any two $x, x' \in X^*$ for which $m^3 \circ m^2(x) = m^3 \circ m^2(x')$ are such that $P(x) = P(x')$ by Anonymity and Weak Independence of Birth Year.

Fourth, mapping m^4 computes the fraction of individuals exhibiting each status. Let mapping $m^4 : \mathbb{N}_0^3\backslash_{(0,0,0)} \to \Delta$ return for any $w \in \mathbb{N}_0^3\backslash_{(0,0,0)}$ the image $v = m^4(w)$ is defined as

$$v = (v_1, v_2, v_3) \equiv \left(\frac{w_1}{w_1 + w_2 + w_3}, \frac{w_2}{w_1 + w_2 + w_3}, \frac{w_3}{w_1 + w_2 + w_3}\right),$$

where v_1 is the fraction of non-poor, v_2 is the fraction of poor and v_3 is the fraction of prematurely dead. Let mapping $m : X \to \Delta$ be defined as $m(x) = m^4 \circ m^3 \circ m^2 \circ m^1(x)$.

First, we show that for any $v \in \Delta$ there exists a $x \in X^*$ such that $m(x) = v$. As $v \in \Delta$, there exist $c_1, c_2, c_3, e_1, e_2, e_3 \in \mathbb{N}$ such that $(v_1, v_2, v_3) = (c_1/e_1, c_2/e_2, c_3/e_3)$. Consider any distribution x with $n(x) = e_1e_2e_3$, where $c_1e_2e_3$ individuals are non-poor, $c_2e_1e_3$ individuals are poor, and $c_3e_1e_2$ individuals are prematurely dead. As $v_1 + v_2 + v_3 = 1$, we have that $c_1e_2e_3 + c_2e_1e_3 + c_3e_1e_2 = e_1e_2e_3$. All individuals in x who are dead are prematurely dead, hence, $x \in X^*$. By construction of x, we have $m(x) = v$.

There remains to show that for any two $x, x' \in X$ such that $m(x) = m(x')$ we have $P(x) = P(x')$. We have shown above that if $m^3 \circ m^2 \circ m^1(x) = m^3 \circ m^2 \circ m^1(x')$, then $P(x) = P(x')$. There remains to show that if $m^3 \circ m^2 \circ m^1(x) \neq m^3 \circ m^2 \circ m^1(x')$ and $m(x) = m(x')$, we have $P(x) = P(x')$. To do so, we show that for any two $w, w' \in \mathbb{N}_0^3\backslash_{(0,0,0)}$ such that $m^4(w) = m^4(w')$, there exist $y, y' \in X$ such that $m^3 \circ m^2 \circ m^1(y) = w$, $m^3 \circ m^2 \circ m^1(y') = w'$ and $P(y) = P(y')$. By construction of mapping m^4, any two $w, w' \in \mathbb{N}_0^3\backslash_{(0,0,0)}$ for which $m^4(w) = m^4(w')$ are such that for $k = w'_1 + w'_2 + w'_3$ and $k' = w_1 + w_2 + w_3$ we have a $w'' \in \mathbb{N}_0^3\backslash_{(0,0,0)}$ such that $w'' = kw = k'w'$. Then, there exist $y, y', y'', y''' \in X^*$ with $m^3 \circ m^2 \circ m^1(y) = w$, $m^3 \circ m^2 \circ m^1(y') = w'$, $m^3 \circ m^2 \circ m^1(y'') = m^3 \circ m^2 \circ m^1(y''') = w''$ such that y'' is a k-replication of y and y''' is a k'-replication of y'. By Replication Invariance, we have that $P(y) = P(y'')$ and $P(y') = P(y''')$. As $m^3 \circ m^2 \circ m^1(y'') = m^3 \circ m^2 \circ m^1(y''')$, we have $P(y'') = P(y''')$. Together, $P(y) = P(y')$.

Step 2: Using mapping m, define an ordering $\succeq$ on Δ from the ordering on X represented by P.

Let $\succeq$ be an ordering on Δ defined such that for any two $v, v' \in \Delta$ we have $v \succ v'$ (resp. $v \sim v'$) if there exist $x, x' \in X$ such that $v = m(x)$ and $v' = m(x')$ and $P(x) < P(x')$ (resp. $P(x) = P(x')$). We showed at the end of Step 1 that there always exist $x, x' \in X$ such that $v = m(x)$ and $v' = m(x')$, which shows that

[39] For any set A, we denote the cardinality of A by $\#A$.

ordering $\succeq$ is complete. Moreover, any two $x, x' \in X$ with $m(x) = m(x')$ are such that $P(x) = P(x')$, which shows that ordering $\succeq$ is well-defined. Together, we have that for any two $x, x' \in X$ and $v, v' \in \Delta$ with $v = m(x)$ and $v' = m(x')$, we have

$$P(x) \leq P(x') \Leftrightarrow v \succeq v'. \tag{8}$$

Step 3: Identifying the appropriate value for γ.

First, we show that $\succeq$ satisfies the following **convexity property**: for any two $v, v' \in \Delta$ with $v \succ v'$ and any rational $\lambda \in (0, 1)$ we have $v \succ \lambda v + (1-\lambda)v' \succ v'$. Take any two $x, y \in X^*$ such that $v = m(x)$ and $v' = m(y)$. Using Replication Invariance, these two distributions can be taken such that $n(x) = n(y)$, which we assume henceforth. By (8), we have $P(x) < P(y)$. By definition of λ, there exists $c, e \in \mathbb{N}$ such that $\lambda = c/e$. Let x^c be a c-replication of x, $x^{(e-c)}$ be a $(e-c)$-replication of x, y^c be a c-replication of y and $y^{(e-c)}$ be a $(e-c)$-replication of y. By construction, we have $n(x^c) = n(y^c)$ and $n(x^{(e-c)}) = n(y^{(e-c)})$. By Replication Invariance, we have

$$P(x^c) = P(x^{(e-c)}) = P(x^c, x^{(e-c)}) < P(y^c) = P(y^{(e-c)}) = P(y^c, y^{(e-c)}).$$

As all these distributions belong to X^*, we have by Subgroup Consistency that $P(x^c, x^{(e-c)}) < P(x^c, y^{(e-c)})$ and that $P(x^c, y^{(e-c)}) < P(y^c, y^{(e-c)})$. Now, we constructed these replications such that $v = m(x^c, x^{(e-c)})$, $v' = m(y^c, y^{(e-c)})$ and also $\lambda v + (1-\lambda)v' = m(x^c, y^{(e-c)})$. This yields the desired result by (8).

Second, we derive the value $\gamma > 0$ for which P is ordinally equivalent to P_γ. Let the three vertices $(1,0,0), (0,1,0), (0,0,1) \in \Delta$ be respectively denoted by v^{100}, v^{010} and v^{001}. By Least Deprivation and (8), we have that $v^{100} \succ v^{010}$ and $v^{100} \succ v^{001}$. There are three cases.

- **Case 1:** $v^{010} \sim v^{001}$.

  Take $\gamma = 1$.

- **Case 2:** $v^{010} \succ v^{001}$.

  Consider the edge connecting vertices v^{100} and v^{001}, which we denote by $E^{100}_{001} = \{v \in \Delta \mid v_2 = 0\}$. As $v^{100} \succ v^{001}$, the convexity property implies that for any $v, v' \in E^{100}_{001}$, if $v_1 > v'_1$ then $v \succ v'$ and if $v_1 < v'_1$ then $v \prec v'$. Let $\Delta^{\mathbb{R}_+}$ be the 2-simplex on the set of real numbers. As $v^{100} \succ v^{010} \succ v^{001}$, there exists a $v^* \in \Delta^{\mathbb{R}_+}$ on the edge connecting the two vertices v^{100} and v^{001} such that for any $v \in E^{100}_{001}$, if $v_1 > v^*_1$ then $v \succ v^{010}$ and, if $v_1 < v^*_1$ then $v \prec v^{010}$. Moreover, if $v^* \in \Delta$, then $v^* \sim v^{010}$ (see proof below). As $\mathbb{Q}$ is dense in $\mathbb{R}$, there is always a rational between two irrationals. Therefore, v^* is the unique element of $\Delta^{\mathbb{R}_+}$ with these properties.

  We show that if $v^* \in \Delta$, then $v^* \sim v^{010}$. Consider the contradiction assumption that $v^* \in \Delta$ and $v^* \succ v^{010}$.[40] We construct a $v' \in E^{100}_{001}$ such that $v'_1 < v^*_1$ and $v' \succ v^{010}$. Such v' is in contradiction with the definition of v^*, which requires that for any $v' \in E^{100}_{001}$ with $v'_1 < v^*_1$ we have $v' \prec v^{010}$. We construct $v' \in E^{100}_{001}$

[40]The alternative contradiction assumption for which $v^* \in \Delta$ and $v^* \prec v^{010}$ also leads to an impossibility.

as follows. Take any two distributions $x, y \in X^*$ such that $v^{010} = m(x)$ and $v^* = m(y)$. As $v^* \succ v^{010}$, we have by (8) that $P(x) > P(y)$. Let $z \in X^*$ be a distribution with $n(z) = 1$ and whose unique individual is prematurely dead. By Young Continuity, there exists some k such that $P(x) > P(y^k, z)$. Consider $v' = m(y^k, z)$. By (8), we have $v' \succ v^{010}$. As $v^* \in E^{100}_{001}$, we have by construction that $v' \in E^{100}_{001}$ and $v'_1 < v^*_1$, the desired result.

We take $\gamma = \frac{1}{v^*_3}$. We have $v^*_3 \in (0, 1)$ because $v^{010} \succ v^{001}$ and $v^{100} \succ v^{010}$ respectively imply that $v^* \neq v^{001}$ and $v^* \neq v^{100}$. As $v^*_3 \in (0, 1)$, we have $\gamma > 1$.

- **Case 3:** $v^{010} \prec v^{001}$.

  The construction of γ is similar to that proposed in Case 2. We find the unique element $v^{**} \in \Delta^{\mathbb{R}_+}$ that splits the edge from v^{100} to v^{010} between elements v for which $v \succ v^{001}$ and elements v' for which $v' \prec v^{001}$. We take $\gamma = v^{**}_2$ and have $\gamma \in (0, 1)$.

We assume henceforth that Case 2 applies, i.e. $v^{010} \succ v^{001}$. We omit the proof for Case 1 that is is simpler and the proof for Case 3 that is very similar.

Step 4: Show that P_γ is ordinally equivalent to P.

Let function $F : \Delta^{\mathbb{R}_+} \to \mathbb{R}_-$ be defined by $F(v) = -(v_2 + \gamma v_3)$. By construction of mapping m and the definition of P_γ, for any $v \in \Delta$ and any $x \in X$ such that $v = m(x)$ we have that $F(v) = -P_\gamma(x)$. If we show that F represents the ordering $\succeq$ on Δ, then we get from (8) that P_γ is ordinally equivalent to P, the desired result.

First, we show that for any $v \in \Delta$ we have $v \succeq v^{010}$ if and only if $F(v) \geq F(v^{010})$. By definition of F, we have that $F(v^{100}) = 0$, $F(v^{010}) = -1$ and $F(v^{001}) = -\gamma$. Partition Δ into three subsets, i.e. $\Delta = \Delta^{100} \cup \Delta^{010} \cup \Delta^{001}$ defined as $\Delta^{010} = \{v \in \Delta \mid F(v) = -1\}$, $\Delta^{100} = \{v \in \Delta \mid F(v) > -1\}$ and $\Delta^{001} = \{v \in \Delta \mid F(v) < -1\}$.[41] We need to show that any $v \in \Delta^{100}$ is such that $v \succ v^{010}$, any $v \in \Delta^{010}$ is such that $v \sim v^{010}$ and any $v \in \Delta^{001}$ is such that $v \prec v^{010}$. In order to avoid repetitions, we only prove that any $v \in \Delta^{100}$ is such that $v \succ v^{010}$. To do so, we show that $v = \lambda v^{010} + (1 - \lambda)v'$ for some rational $\lambda \in [0, 1)$ and some v' on the edge E^{100}_{001} with $v'_1 > v^*_1$. This construction is illustrated in Panel A of Figure 14. Given that any $v' \in E^{100}_{001}$ for which $v'_1 > v^*_1$ is such that $v' \succ v^{010}$, the convexity property of $\succeq$ then implies that $v \succ v^{010}$. Take $v' \equiv \left(1 - \frac{v_3}{1-v_2}, 0, \frac{v_3}{1-v_2}\right)$. As $v \in \Delta$, the definition of v' is such that $v' \in E^{100}_{001}$. Let $v'' \equiv \lambda v^{010} + (1 - \lambda)v'$ where $\lambda \equiv v_2 \in [0, 1)$ since $v \in \Delta^{100}$. We have $v'' = v$ since, by construction of v', we have $v''_2 = \lambda = v_2$ and $v''_3 = (1 - \lambda)v'_3 = v_3$. There remains to show that $v'_1 > v^*_1$. Last inequality is equivalent to $1 - \frac{v_3}{1-v_2} > 1 - \frac{1}{\gamma}$, which simplifies to $\frac{1}{\gamma} > \frac{v_3}{1-v_2}$. This inequality holds because, as $v \in \Delta^{100}$, we have $F(v) > -1$, which simplifies to the same inequality.

[41] We have defined F and γ such that $F(v^*) = F(v^{010})$. From a geometric perspective, the set of elements $v \in \Delta^{\mathbb{R}_+}$ for which $F(v) = -1$ is the segment connecting v^{010} with v^*. Observe that if $v^* \notin \Delta$, then the only element in this segment belonging to Δ is the vertex v^{010} and, therefore, Δ^{010} degenerates to $\{v^{010}\}$. The subset Δ^{100} contains vertex v^{100} and all elements of Δ that are on v^{100}'s side of the segment connecting v^{010} with v^*. In turn, Δ^{001} contains vertex v^{001} and all elements of Δ that are on v^{001}'s side of the segment connecting v^{010} with v^*.

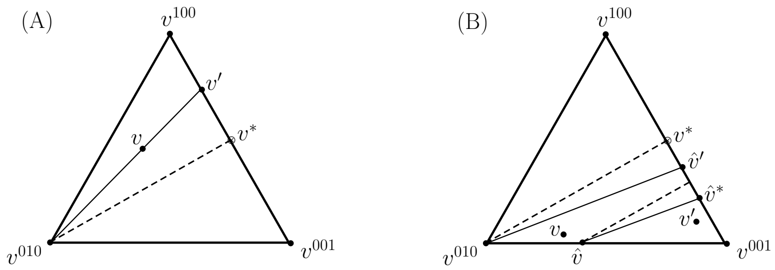


Figure 14: Panel A: construction used in order to show that $v \succ v^{010}$ when $F(v) > F(v^{010})$. Panel B: construction used in order to show that $v \succ v'$ when $F(v) > F(v')$. Iso-F lines are dashed.

Geometrically, we have just shown that the segment connecting v^{010} to v^* is an "implicit" indifference curve of $\succeq$.[42] The intuition for the rest of the proof is that all parallel segments are also "implicit" indifference curves of $\succeq$.

Take any two $v, v' \in \Delta$ with $F(v) \geq F(v')$, we show that $v \succeq v'$. If $F(v) \geq -1 \geq F(v')$, then the previous argument directly yields the result. We focus on the particular case $-1 > F(v) > F(v')$ and show that $v \succ v'$ (the proofs for the other cases are similar). The construction is illustrated in Panel B of Figure 14. This case is such that there exists a $\hat{v} = (0, \hat{v}_2, 1 - \hat{v}_2) \in E^{010}_{001}$ with $F(v) > F(\hat{s}) > F(v')$, because $F(v^{010}) = -1$ and $F(v^{001}) = \min_{v'' \in \Delta} F(v'')$. By the convexity property of $\succeq$, our assumption $v^{010} \succ v^{001}$ implies that $v^{010} \succ \hat{v} \succeq v^{001}$. Therefore, there exists a unique $\hat{v}^* \in \Delta^{\mathbb{R}_+}$ on the edge connecting vertices v^{100} and v^{001} such that for any $v'' \in E^{100}_{001}$, if $v''_1 > \hat{v}^*_1$ then $v'' \succ \hat{v}$, if $v''_1 < \hat{v}^*_1$ then $v'' \prec \hat{v}$ and if $\hat{v}^* \in \Delta$, then $\hat{v}^* \sim \hat{v}$ (the omitted proof for this claim follows the argument provided in Step 3 Case 2).

First, we show that the segment connecting $\hat{v}$ to $\hat{v}^*$ is parallel to the segment connecting v^{010} to v^*. Formally, this is equivalent to showing that $\hat{v}_2 = \frac{\hat{v}^*_1}{v^*_1}$. Consider the contradiction assumption for which $\hat{v}_2 > \frac{\hat{v}^*_1}{v^*_1}$.[43] Assume that $\hat{v}^* \in \Delta$.[44] Consider now $\hat{v}^{'} = (\frac{\hat{v}^*_1}{\hat{v}_2}, 0, 1 - \frac{\hat{v}^*_1}{\hat{v}_2}) \in E^{100}_{001}$. By the contradiction assumption, we have $\hat{v}^{'}_1 < v^*_1$ and, hence, $v^{010} \succ \hat{v}^{'}$. By construction, for the rational $\lambda = \hat{v}_2$ we have:

$$\hat{v} = \lambda v^{010} + (1 - \lambda) v^{001} \qquad \text{and} \qquad \hat{v}^* = \lambda \hat{v}^{'} + (1 - \lambda) v^{001}.$$

We use that $v^{010} \succ \hat{v}^{'}$ in order to show that $\hat{v} \succ \hat{v}^*$, a contradiction to the definition of $\hat{v}^*$. Take any three distributions $x, y, z \in X^*$ such that $v^{010} = m(x)$, $\hat{v}^{'} = m(y)$ and $v^{001} = m(z)$. By (8), we have $P(x) < P(y) < P(z)$. Using Replication Invariance, these three distributions can be taken such that $n(x) = n(y) = n(z)$, which we assume henceforth. As $\lambda = \hat{v}_2$, there exist $c, e \in \mathbb{N}$ such that $\lambda = c/e$. Let x^c be a c-replication of x, y^c be a c-replication of y and $z^{(e-c)}$ be a $(e-c)$-replication of z. By Replication Invariance, we have $P(x^c) < P(y^c) < P(z^{(e-c)})$. Thus, by Subgroup Consistency, we have that $P(x^c, z^{(e-c)}) < P(y^c, z^{(e-c)})$. Now, we constructed these replications such that $\hat{v} = m((x^c, z^{(e-c)}))$ and $\hat{v}^* = m((y^c, z^{(e-c)}))$. By (8), we obtain $\hat{v} \succ \hat{v}^*$, the desired contradiction.

Second, we use the previous result to show that $v \succ v'$. Partition Δ into three

[42] We call this indifference curve "implicit" because it is defined in $\Delta^{\mathbb{R}_+}$ rather than in Δ.

[43] The alternative contradiction assumption for which $\hat{v}_2 < \frac{\hat{v}^*_1}{v^*_1}$ also leads to an impossibility.

[44] If $\hat{v}^* \notin \Delta$, then replace $\hat{v}^*$ by a nearby $\tilde{v}^* \in E^{100}_{001}$ for which $\tilde{v}^*_1 > \hat{v}^*_1$ and $\hat{v}_2 > \frac{\tilde{v}^*_1}{v^*_1}$. As $\tilde{v}^*_1 > \hat{v}^*_1$, we have $\tilde{v}^* \succ \hat{v}$.

subsets, i.e. $\Delta = \Delta^{100'} \cup \Delta^{\hat{v}} \cup \Delta^{001'}$ defined as $\Delta^{\hat{v}} = \{v'' \in \Delta \mid F(v'') = F(\hat{v})\}$, $\Delta^{100'} = \{v'' \in \Delta \mid F(v'') > F(\hat{v})\}$ and $\Delta^{001'} = \{v'' \in \Delta \mid F(v'') < F(\hat{v})\}$. We have by construction that $v \in \Delta^{100'}$ and $v' \in \Delta^{001'}$. We can show that $v' \prec \hat{v}$ using the same proof technique as above, i.e. show that v' is on a segment connecting $\hat{v}$ to a v'' on the edge E^{100}_{001} with $v''_1 < \hat{v}^*_1$ and, hence, such that $v'' \prec \hat{v}$. By the convexity property of $\succeq$, this yields in turn $v' \prec \hat{v}$. Similarly, we can show that $v \succ \hat{v}$ by showing that v is on a segment connecting $\hat{v}$ to a v''' that is *either* on the edge E^{100}_{001} with $v'''_1 > \hat{v}^*_1$ and, hence, such that $v''' \succ \hat{v}$ *or* on the edge E^{100}_{010} and, as $v^{100} \succ v^{010} \succ \hat{v}$, such that $v''' \succ \hat{v}$. This implies in both cases that $v \succ \hat{v}$.

7.2 Proof of Lemma 1

We prove in Proposition 6 a slightly more general result, which requires to refine the definition of a constant population pyramid.

Definition 1 (Constant population pyramid up to a')**.**
The pair $(x, \mu) \in X \times M$ has a constant population pyramid up to a' if we have for all $a \in \{0, \dots, a'\}$ that

$$n_{a+1}(x) = n_a(x) * (1 - \mu_a). \tag{9}$$

The pair (x, μ) has a **constant population pyramid** if it has a constant population pyramid up to $a^* - 1$, implying that (9) holds for all $a \in \{0, \dots, a^* - 1\}$. Observe that a constant population pyramid does not require that past mortality rates and natality were constant.

Proposition 6 (Equivalence between ID, GD and ED indices in equilibrium)**.**
If the pair $(x, \mu) \in X \times M$ has a constant population pyramid, then we have that

$$P^{ED}_\gamma(x, \mu) = P^{GD}_\gamma(x, \mu).$$

If in addition $n_a(x) + d_a(x) = n^* \in \mathbb{N}$ for all $a \in \{0, \dots, \hat{a} - 1\}$,

$$P^{ED}_\gamma(x, \mu) = P^{GD}_\gamma(x, \mu) = P^{ID}_\gamma(x).$$

Proof. We prove both equalities in turn.

First, we show that $P^{ED}_\gamma(x, \mu) = P^{GD}_\gamma(x, \mu)$, where

$$P^{GD}_\gamma(x, \mu) = \frac{(p(x) + f(x)) * HC(x)}{p(x) + f(x) + d^{GD}(x, \mu)} + \gamma \frac{d^{GD}(x, \mu)}{p(x) + f(x) + d^{GD}(x, \mu)}.$$

Given that (x, μ) has a constant population pyramid, we have for all $a \in \{0, \dots, a^*\}$ that

$$n_a(x) = n_0(x) * \prod_{l=0}^{a-1} (1 - \mu_l). \tag{10}$$

Using (10), the definition of $d^{GD}(x, \mu)$ may be rewritten

$$\begin{aligned} d^{GD}(x,\mu) &= \sum_{a=0}^{\hat{a}-1} n_a(x) * \mu_a * (\hat{a} - (a+1)), \\ &= n_0(x) \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a+1)) * \mu_a * \prod_{l=0}^{a-1} (1-\mu_l), \\ &= n_0(x) LGE_{\hat{a}}(\mu). \end{aligned}$$

Similarly, given that (x, μ) has a constant population pyramid, we may rewritte $p(x) + f(x)$ using (10) as

$$\begin{aligned} p(x) + f(x) &= \sum_{a=0}^{a^*} n_a(x), \\ &= n_0(x) \sum_{a=0}^{a^*} \prod_{l=0}^{a-1} (1-\mu_l), \\ &= n_0(x) LE(\mu). \end{aligned}$$

Replacing $d^{GD}(x, \mu)$ and $p(x) + f(x)$ in the definition of $P_\gamma^{GD}(x, \mu)$ yields

$$\begin{aligned} P_\gamma^{GD}(x,\mu) &= \frac{n_0(x)LE(\mu)HC(x)}{n_0(x)LE(\mu) + n_0(x)LGE_{\hat{a}}(\mu)} + \gamma \frac{n_0(x)LGE_{\hat{a}}(\mu)}{n_0(x)LE(\mu) + n_0(x)LGE_{\hat{a}}(\mu)}, \\ &= \frac{LE(\mu) * HC(x)}{LE(\mu) + LGE_{\hat{a}}(\mu)} + \gamma \frac{LGE_{\hat{a}}(\mu)}{LE(\mu) + LGE_{\hat{a}}(\mu)} = P_\gamma^{ED}(x,\mu). \end{aligned}$$

Second, we show that $P_\gamma^{GD}(x, \mu) = P_\gamma^{ID}(x)$. As the pair (x, μ) has a constant population pyramid and for all cohorts $a \in \{0, \ldots, \hat{a}-1\}$ we have that $n_a(x)+d_a(x) = n^*$, Lemma 2 applies and we have $d(x) = d^{GD}(x, \mu)$. Therefore, the definition of $P_\gamma^{GD}(x, \mu)$ becomes

$$\begin{aligned} P_\gamma^{GD}(x,\mu) &= \frac{(p(x)+f(x)) * HC(x)}{p(x)+f(x)+d^{GD}(x,\mu)} + \gamma \frac{d^{GD}(x,\mu)}{p(x)+f(x)+d^{GD}(x,\mu)}, \\ &= \frac{(p(x)+f(x)) * HC(x)}{p(x)+f(x)+d(x)} + \gamma \frac{d(x)}{p(x)+f(x)+d(x)} = P_\gamma^{ID}(x). \end{aligned}$$

□

Lemma 2 (Equivalence between d and d^{GD} in stationary pairs)**.**
If the pair $(x, \mu) \in O$ has a constant population pyramid up to $\hat{a}-2$ and for all cohorts $a \in \{0, \ldots, \hat{a}-1\}$ we have $n_a(x) + d_a(x) = n^* \in \mathbb{N}$, then we have $d(x) = d^{GD}(x, \mu)$.

Proof. The proof is direct. Take any pair $(x, \mu) \in X \times M$ that has a constant population pyramid up to $\hat{a} - 2$ and such that for all cohorts $a \in \{0, \ldots, \hat{a} - 1\}$ we have $n_a(x) + d_a(x) = n^* \in \mathbb{N}$. By definition, the number of prematurely dead individuals in period t counted by the inherited deprivation approach is

$$d(x) = \sum_{a=1}^{\hat{a}-1} d_a(x).$$

Given that the number of newborns is assumed constant in earlier periods,

$$d(x) = \sum_{a=1}^{\hat{a}-1} \left(n^* - n_a(x)\right).$$

As $n_0(x) = n^*$, we may rewritte the previous equation as

$$d(x) = \sum_{a=1}^{\hat{a}-1} \left(\sum_{a'=0}^{a-1} \left(n_{a'}(x) - n_{a'+1}(x)\right) \right),$$

and developing the sums, we get

$$\begin{aligned} d(x) &= (\hat{a}-1)(n_0(x) - n_1(x)) + (\hat{a}-2)(n_1(x) - n_2(x)) + \cdots + (\hat{a} - (\hat{a}-1))(n_{\hat{a}-2}(x) - n_{\hat{a}-1}(x)), \\ &= \sum_{a=0}^{\hat{a}-2} (n_a(x) - n_{a+1}(x))(\hat{a} - (a+1)), \end{aligned}$$

and given that $\hat{a} - ((\hat{a}-1)+1) = 0$, this is equivalent to

$$d(x) = \sum_{a=0}^{\hat{a}-\mathbf{1}} (n_a(x) - n_{a+1}(x))(\hat{a} - (a+1)).$$

Finally, as the pair (x, μ) has a constant population pyramid up to $\hat{a}-2$ we have for all $a \in \{0, \ldots, \hat{a}-2\}$ that $n_{a+1}(x) = n_a(x) - n_a(x) * \mu_a$ and therefore we have

$$d(x) = \sum_{a=0}^{\hat{a}-1} n_a(x) * \mu_a * (\hat{a} - (a+1)) = d^{GD}(x, \mu).$$

□

7.3 Proof of Proposition 2

First, we provide the formal definition of the axioms not defined in the text.

Deprivation axiom 12 (Independence of Dead*). *For all $(x, \mu) \in O$ and $i \leq n(x)$, if $s_i = D$, then $P((x_i, x_{-i}), \mu) = P(x_{-i}, \mu)$.*

Deprivation axiom 13 (Independence of Birth Year). *For all $(x, \mu) \in O$ and $i \leq n(x)$, if $s_i = s_i'$, then $P((x_i, x_{-i}), \mu) = P((x_i', x_{-i}), \mu)$.*

Deprivation axiom 14 (Replication Invariance*). *For all $(x, \mu) \in O$ and $k \in \mathbb{N}$, $P(x^k, \mu) = P(x, \mu)$.*

Proving that the ED index satisfies Independence of Dead*, Replication Invariance* and Independence of Birth Year is straightforward and left to the reader. Finally, Lemma 1 shows that the ED index satisfies ID Equivalence.

We prove sufficiency. Take any pair $(x, \mu) \in O$. We construct another pair (x''', μ) that is stationary and such that $P(x''', \mu) = P(x, \mu)$ and $P^{ED}(x''', \mu) = P^{ED}(x, \mu)$. Given that (x''', μ) is stationary, we have by ID Equivalence that $P(x''', \mu) = P^{ED}(x''', \mu)$. The characteristics of (x''', μ) then imply that $P(x, \mu) = P^{ED}(x, \mu)$, the desired result.

We turn to the construction of the stationary pair (x''', μ). One difficulty is to ensure that the mortality rates μ_a are feasible in a constant population pyramid given the number of alive individuals $n_a(x''')$, which is $\mu_a = \frac{c}{n_a(x''')}$ for some $c \in \mathbb{N}$. We first construct a $k-$replication of x that has sufficiently many alive individuals to meet this constraint. By definition, mortality rates can be expressed as $\mu_a = \frac{c_a}{e_a}$ where $c_a, e_a \in \mathbb{N}$. Let $e = \prod_{j=0}^{a^*-1} e_j$, $n'_a = e\prod_{j=0}^{a-1}(1-\frac{c_j}{e_j})$ and $n' = \sum_{j=0}^{a^*} n'_j$. Let x' be a $n'-$replication of x. Letting $n^x = \sum_{j=0}^{a^*} n_j(x)$ be the number of alive individuals in distribution x, we have by construction 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 age of alive individuals in such a way that (x'', μ) has a constant population pyramid. 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 $a \in \{0, \dots, a^*\}$, the number of a-years old individuals is $n' * n^x * \frac{\prod_{j=0}^{a-1}(1-\frac{c_j}{e_j})}{\sum_{k=0}^{a^*}\prod_{j=0}^{k-1}(1-\frac{c_j}{e_j})}$.

One can check that (x'', μ) has a constant population pyramid 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 age of dead individuals in such a way that (x''', μ) 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, $HC(x''') = HC(x)$, which implies that $P^{ED}(x''', \mu) = P^{ED}(x, \mu)$.

7.4 Proof of Proposition 3

We first prove necessity. Proving that the FP index satisfies Independence of Dead* is straightforward and left to the reader. Lemma 1 shows that the GD index satisfies ID Equivalence. Finally, the GD index satisfies Additive Decomposibility when the size function is defined as $\eta(x, \mu) = f(x) + p(x) + d^{GD}(x, \mu)$. We show that this function is indeed such that $\eta(x, \mu) = \eta(x', \mu') + \eta(x'', \mu'')$. Given that $f(x', x'') + p(x', x'') = f(x') + p(x') + f(x'') + p(x'')$, we must show that $d^{GD}((x', x''), \mu) = d^{GD}(x', \mu') + d^{GD}(x'', \mu'')$. We have

$$\begin{aligned}
d^{GD}((x', x''), \mu) &= \sum_{a=0}^{\hat{a}-1} n_a(x', x'') * \mu_a * (\hat{a} - (a+1)) \\
&= \sum_{a=0}^{\hat{a}-1} (n_a(x') + n_a(x'')) * \frac{n_a(x') * \mu'_a + n_a(x'') * \mu''_a}{n_a(x') + n_a(x'')} * (\hat{a} - (a+1)) \\
&= d^{GD}(x', \mu') + d^{GD}(x'', \mu'').
\end{aligned}$$

It is then straighforwd to verify (4) by replacing P and η by their expressions.

We now prove sufficiency. Take any pair $(x', \mu) \in O$. Consider the distribution x obtained from x' by removing all dead individuals in x. We have $P(x, \mu) = P(x', \mu)$ by Independence of Dead* and also $P_\gamma^{GD}(x, \mu) = P_\gamma^{GD}(x', \mu)$.

The proof requires to define, for each $a \in \{0, \dots, a^*\}$, two counterfactual pairs (x_a^*, μ_a^*) and (x_a^0, μ_a^0), which are illustrated in the center and right panels of Figure 15.

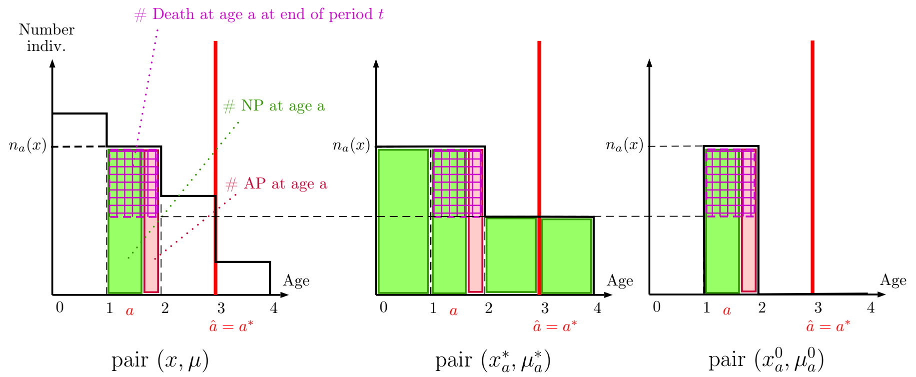


Figure 15: Left panel: pair (x, μ). Center panel: stationary pair (x_a^*, μ_a^*), where dead individuals are not shown. Right panel: degenerated pair (x_a^0, μ_a^0).

The counterfactual pair (x_a^*, μ_a^*) is stationary. The vector μ_a^* is such that mortality rates are zero except for two cases: $\mu_a^* = \mu_a$ and $\mu_{a^*}^* = 1$, which is $\mu_a^* = (0, \dots, 0, \mu_a, 0, \dots, 0, 1)$. We now turn to the construction of the distribution x_a^*. At all ages $a' \leq a$, there are exactly $n_a(x)$ alive individuals (i.e. $n_{a'}(x_a^*) = n_a(x)$); for ages $a' > a$ we have $n_{a'}(x_a^*) = n_a(x) * (1 - \mu_a)$. At all ages $a' \leq a$, there are no dead individuals; for ages $a' > a$, this number is $n_a(x) * \mu_a$. There are no poor (AP) individuals in x_a^* except at age a, where this number is equal to the number of a-years old individuals in x whose status is AP, i.e. $\#\{i \leq n(x) | s_i = AP \text{ and } b_i = t - a\}$.

The counterfactual pair (x_a^0, μ_a^0) is not stationary and all its alive individuals are a-years old. The vector $\mu_a^0 = \mu_a^*$, which is $\mu_a^0 = (0, \dots, 0, \mu_a, 0, \dots, 0, 1)$. We now turn to the construction of distribution x_a^0. At all ages $a' \neq a$, there are no alive individuals (i.e. $n_{a'}(x_a^0) = 0$); and we have $n_a(x_a^0) = n_a(x)$. There are no dead individuals. The number of AP individuals in x_a^0 is equal to the number of a-years old individuals in x who are AP, i.e. $\#\{i \leq n(x) | s_i = AP \text{ and } b_i = t - a\}$.

By iterative application of Additive Decomposibility, we have that

$$P(x, \mu) = \frac{\sum_{j=0}^{a^*} \eta(x_j^0, \mu_j^0) * P(x_j^0, \mu_j^0)}{\sum_{j=0}^{a^*} \eta(x_j^0, \mu_j^0)}. \tag{11}$$

Expression (11) holds in particular for the stationary pair (x_a^*, μ_a^*):

$$P(x_a^*, \mu_a^*) = \frac{\sum_{j=0}^{a^*} \eta((x_a^*)_j^0, (\mu_a^*)_j^0) * P((x_a^*)_j^0, (\mu_a^*)_j^0)}{\sum_{j=0}^{a^*} \eta((x_a^*)_j^0, (\mu_a^*)_j^0)}, \tag{12}$$

where the pair $((x_a^*)_j^0, (\mu_a^*)_j^0)$ is the degenate pair associated to stationary pair (x_a^*, μ_a^*) at age j. In particular, the mortality vector $(\mu_a^*)_j^0 = \mu_a^*$ for $j = a$ and $(\mu_a^*)_j^0 = (0, \dots, 0, 1)$ for $j \neq a$; and for $j = a$ we have $((x_a^*)_j^0, (\mu_a^*)_j^0) = (x_a^0, \mu_a^0)$.

For all $j \neq a$ we show that $P((x_a^*)_j^0, (\mu_a^*)_j^0) = 0$. Recall that $(\mu_a^*)_j^0 = (0, \dots, 0, 1)$ and that $(x_a^*)_j^0$ has zero j-years old individuals whose status is AP. Consider the stationary pair (x''', μ''') such that $\mu''' = (0, \dots, 0, 1)$ and such that distribution x'''

has zero AP individual and zero dead individual. By ID Equivalence, we have that $P(x''', \mu''') = P^{GD}_{\gamma}(x''', \mu''') = 0$. Provided that $n(x''') = n_j(x^*_a) * (a^* + 1)$, we have that $P((x^*_a)^0_j, (\mu^*_a)^0_j)$ appears in the decomposition (11) applied to (x''', μ'''). Given that P does not yield negative images, we must have that $P((x^*_a)^0_j, (\mu^*_a)^0_j) = 0$.[45]

As (x^*_a, μ^*_a) is stationary, we have from ID Equivalence that $P(x^*_a, \mu^*_a) = P^{GD}_{\gamma}(x^*_a, \mu^*_a)$. Given that $P((x^*_a)^0_j, (\mu^*_a)^0_j) = 0$ for all $j \neq a$, and $\sum_{j=0}^{a^*} \eta((x^*_a)^0_j, (\mu^*_a)^0_j) = \eta(x^*_a, \mu^*_a)$, (12) may be rewritten as

$$P((x^*_a)^0_a, (\mu^*_a)^0_a) = \frac{\eta(x^*_a, \mu^*_a) * P^{GD}_{\gamma}(x^*_a, \mu^*_a)}{\eta((x^*_a)^0_a, (\mu^*_a)^0_a)}.$$

As $((x^*_a)^0_a, (\mu^*_a)^0_a) = (x^0_a, \mu^0_a)$, this last identity becomes

$$P(x^0_a, \mu^0_a) = \frac{\eta(x^*_a, \mu^*_a) * P^{GD}_{\gamma}(x^*_a, \mu^*_a)}{\eta(x^0_a, \mu^0_a)}.$$

Inserting this last expression in (11), where $\sum_{j=0}^{a^*} \eta(x^0_j, \mu^0_j) = \eta(x, \mu)$, yields

$$P(x, \mu) = \frac{\sum_{j=0}^{a^*} \eta(x^*_j, \mu^*_j) * P^{GD}_{\gamma}(x^*_j, \mu^*_j)}{\eta(x, \mu)}. \tag{13}$$

Equation (13) holds for all pairs in O. If we have that function η is defined as $\eta(x, \mu) = f(x) + p(x) + d^{GD}(x, \mu)$, then (13) simplifies to $P(x, \mu) = P^{GD}_{\gamma}(x, \mu)$ and the proof is complete. We now show that the function η is indeed expressed as $\eta(x, \mu) = f(x) + p(x) + d^{GD}(x, \mu)$. Equation (13) holds in particular for any stationary pair (x', μ'). Therefore, by ID Equivalence we have

$$P^{GD}_{\gamma}(x', \mu') = \frac{\sum_{j=0}^{a^*} \eta((x')^*_j, (\mu')^*_j) * P^{GD}_{\gamma}((x')^*_j, (\mu')^*_j)}{\eta(x', \mu')}, \tag{14}$$

where the pair $((x')^*_j, (\mu')^*_j)$ is the stationary pair associated to the pair (x', μ') at age j. Then, last expression only holds if function η has the appropriate expression.

7.5 Proof of Proposition 4

Take any $(x, \mu) \in O$. For the sake of notation simplicity, we henceforth use u_{s_a} to denote $u(s_a)$. Consider the affine transformation of EU defined as $EU'(x, \mu) = \frac{1}{\hat{a}} EU(x, \mu) - u_{NP}$. As $\beta = 1$, we can write

$$EU'(x, \mu) = \frac{1}{\hat{a}} \left[\mathbb{E} \sum_{a=0}^{a^*} u(s_a) S(a, \mu) - \sum_{a=0}^{a^*} S(a, \mu) u_{NP} - \left(\hat{a} - \sum_{a=0}^{a^*} S(a, \mu) \right) u_{NP} \right].$$

Letting $\pi_a(x)$ denote the *fraction* of the alive individuals of age a who are poor in distribution x, we have $\mathbb{E} u(s_a) = \pi_a(x) u_{AP} + (1 - \pi_a(x)) u_{NP}$ and therefore

$$EU'(x, \mu) = \frac{1}{\hat{a}} \left[\sum_{a=0}^{a^*} \pi_a(x) \left(u_{AP} - u_{NP} \right) S(a, \mu) - \left(\hat{a} - \sum_{a=0}^{a^*} S(a, \mu) \right) u_{NP} \right].$$

[45] In order to be complete, there remains to show that $\eta((x^*_a)^0_j, (\mu^*_a)^0_j) > 0$ when $n_j(x^*_a) > 0$. If not, one can derive a contradiction with the requirement that $\eta(x, \mu) = \eta(x', \mu') + \eta(x'', \mu'')$.

As $\sum_{a=0}^{a^*} S(a,\mu) = \sum_{a=0}^{a^*} \prod_{l=0}^{a-1}(1-\mu_l) = LE(\mu)$, the previous expression becomes

$$EU'(x,\mu) = \frac{1}{\hat{a}}\left[LE(\mu)HC(x)\left(u_{AP} - u_{NP}\right) - \left(\hat{a} - \sum_{a=0}^{a^*} S(a,\mu)\right)u_{NP}\right].$$

When $\hat{a} > a^*$, we have that $\hat{a} = LE(\mu) + LGE_{\hat{a}}(\mu)$ and therefore $\hat{a} - \sum_{a=0}^{a^*} S(a,\mu) = LGE_{\hat{a}}(\mu)$, hence

$$EU'(x,\mu) = \frac{1}{LE(\mu) + LGE_{\hat{a}}(\mu)}\left[LE(\mu)HC(x)\left(u_{AP} - u_{NP}\right) + LGE_{\hat{a}}(\mu)(-u_{NP})\right].$$

As $(u_{AP} - u_{NP}) < 0$ we have for any two $(x,\mu), (x',\mu') \in O$ that

$$EU'(x,\mu) \geq EU'(x',\mu') \Leftrightarrow \frac{EU'(x,\mu)}{u_{AP} - u_{NP}} \leq \frac{EU'(x',\mu')}{u_{AP} - u_{NP}},$$

the desired result since $P_\gamma^{ED} = \frac{EU'}{u_{AP}-u_{NP}}$ when $\gamma = \frac{-u_{NP}}{u_{AP}-u_{NP}}$ and the affine transformation preserves the ranking of all pairs.

7.6 Proof of Proposition 5

The definition of $P_\gamma^{ED}(x,\mu)$ is

$$P_\gamma^{ED}(x,\mu) = \frac{LE(\mu) * HC(x)}{LE(\mu) + LGE_{\hat{a}}(\mu)} + \gamma\frac{LGE_{\hat{a}}(\mu)}{LE(\mu) + LGE_{\hat{a}}(\mu)}.$$

Let $n(x) = p(x) + f(x)$ be the number of alive individuals in x. When multiplying the numerator and denomitor by $\frac{n(x)}{LE(\mu)}$, we get

$$P_\gamma^{ED}(x,\mu) = \frac{n(x) * HC(x) + \gamma * \frac{LGE_{\hat{a}}(\mu) * n(x)}{LE(\mu)}}{n(x) + \frac{LGE_{\hat{a}}(\mu) * n(x)}{LE(\mu)}},$$

Therefore inequality $P_\gamma^{ED}(x,\mu) \leq P_\gamma^{GD}(x,\mu)$ becomes

$$\frac{n(x) * HC(x) + \gamma * \frac{LGE_{\hat{a}}(\mu) * n(x)}{LE(\mu)}}{n(x) + \frac{LGE_{\hat{a}}(\mu) * n(x)}{LE(\mu)}} \leq \frac{n(x) * HC(x) + \gamma * d^{GD}(x,\mu)}{n(x) + d^{GD}(x,\mu)}. \tag{15}$$

When $\gamma \geq 1$, each of the two fractions compared in inequality (15) is monotonically increasing in the factor multiplying γ, respectively $LGE_{\hat{a}}(\mu) * n(x)/LE(\mu)$ and d^{GD}.[46] Therefore, inequality (15) is equivalent to

$$\frac{LGE_{\hat{a}}(\mu) * n(x)}{LE(\mu)} \leq d^{GD}(x,\mu). \tag{16}$$

[46]For example, P_γ^{GD} is increasing in d^{GD} as we have by chain derivation that

$$\frac{\partial P_\gamma^{GD}}{\partial d^{GD}} = \frac{\gamma(n(x) + d^{GD}) - (n(x) * HC(x) + \gamma * d^{GD})}{(n(x) + d^{GD})^2} = \frac{n(x) * (\gamma - HC(x))}{(n(x) + d^{GD})^2},$$

where $\gamma \geq HC(x)$ when $\gamma \geq 1$ as there is at least one individual who is non-poor.

Last inequality is equivalent to

$$\frac{p(x)}{p(x) + n(x) * \frac{LGE_{\hat{a}}(\mu)}{LE(\mu)}} \geq \frac{p(x)}{p(x) + d^{GD}(x, \mu)},$$

and recalling that $HC(x) = \frac{p(x)}{n(x)}$, we obtain

$$\frac{LE(\mu) * HC(x)}{LE(\mu) * HC(x) + LGE_{\hat{a}}(\mu)} \geq \frac{p(x)}{p(x) + d^{GD}(x, \mu)},$$

which is equivalent to inequality (6).

Then, replacing d^{GD} and $LGE_{\hat{a}}$ in inequality (16) by their definitions leads to

$$\frac{n(x)}{LE(\mu)} * \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a+1)) * \mu_a * \prod_{l=0}^{a-1} (1 - \mu_l) \leq \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a+1)) * \mu_a * n_a(x).$$

As μ^x is the mortality vector for which (x, μ^x) has a constant population pyramid, we have for all $a \in \{0, \ldots, a^*\}$ that $n_a(x) = n_0(x) * \prod_{l=0}^{a-1} (1 - \mu_l^x)$. Replacing $n_a(x)$ in last inequality leads to

$$\frac{n(x)}{LE(\mu)} * \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a+1)) * \mu_a * \prod_{l=0}^{a-1} (1 - \mu_l) \leq n_0(x) * \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a+1)) * \mu_a * \prod_{l=0}^{a-1} (1 - \mu_l^x).$$

where $n(x) = \sum_{a=0}^{a^*} n_0(x) * \prod_{l=0}^{a-1} (1 - \mu_l^x) = n_0(x) * LE(\mu^x)$, yielding

$$\frac{\sum_{a=0}^{\hat{a}-1} (\hat{a} - (a+1)) * \prod_{l=0}^{a-1} (1 - \mu_l) * \mu_a}{LE(\mu)} \leq \frac{\sum_{a=0}^{\hat{a}-1} (\hat{a} - (a+1)) * \prod_{l=0}^{a-1} (1 - \mu_l^x) * \mu_a}{LE(\mu^x)},$$

the desired result.

7.7 Lemma 3

Lemma 3 ($LGE_{\hat{a}}$ equivalent to LE for large $\hat{a}$)**.**
For any two mortality vectors $\mu, \nu \in M$, if we have $\hat{a} > a^*$, then we have

$$LGE_{\hat{a}}(\mu) \leq LGE_{\hat{a}}(\nu) \quad \Leftrightarrow \quad LE(\mu) \geq LE(\nu)$$

Proof. Take any $\mu, \nu \in M$. If we show for all $\mu' \in M$ that if $\hat{a} > a^*$ we have

$$LE(\mu') = \hat{a} - LGE_{\hat{a}}(\mu'), \tag{17}$$

then $LE(\mu) \geq LE(\nu)$ may be rewritten as

$$\hat{a} - LGE_{\hat{a}}(\mu) \geq \hat{a} - LGE_{\hat{a}}(\nu),$$

which is equivalent to $LGE_{\hat{a}}(\mu) \leq LGE_{\hat{a}}(\nu)$, the desired result.

There remains to prove that (17) holds for all $\mu' \in M$. For all $\mu' \in M$ we have

$$\begin{aligned} LE(\mu') &= \sum_{a=0}^{a^*} \prod_{l=0}^{a-1} (1-\mu'_l), \\ &= \sum_{a=0}^{a^*} \left(\sum_{a'=a}^{a^*} \mu'_{a'} * \prod_{l=0}^{a'-1} (1-\mu'_l) \right), \end{aligned}$$

and developing the sums, we get

$$\begin{aligned} LE(\mu') =& (a^*+1) \left(\mu'_{a^*} * \prod_{l=0}^{a^*-1} (1-\mu'_l) \right) + (a^*) \left(\mu'_{a^*-1} * \prod_{l=0}^{a^*-2} (1-\mu'_l) \right) + \\ & \cdots + 2 \left(\mu'_1 * \prod_{l=0}^{0} (1-\mu'_l) \right) + \mu'_0, \\ &= \sum_{a=0}^{a^*} (a+1) * \mu'_a * \prod_{l=0}^{a-1} (1-\mu'_l) \\ &= \sum_{a=0}^{a^*} (\hat{a} - \hat{a} + (a+1)) * \mu'_a * \prod_{l=0}^{a-1} (1-\mu'_l) \\ &= \sum_{a=0}^{a^*} \hat{a} * \mu'_a * \prod_{l=0}^{a-1} (1-\mu'_l) - \sum_{a=0}^{a^*} (\hat{a} - (a+1)) * \mu'_a * \prod_{l=0}^{a-1} (1-\mu'_l) \end{aligned}$$

and because we have that $\sum_{a=0}^{a^*} \mu'_a * \prod_{l=0}^{a-1} (1-\mu'_l) = 1$ since $\mu'_{a^*} = 1$, last equation becomes

$$LE(\mu') = \hat{a} - \sum_{a=0}^{a^*} (\hat{a} - (a+1)) * \mu'_a * \prod_{l=0}^{a-1} (1-\mu'_l).$$

For all $a \geq a^*+1$ we have that $\prod_{l=0}^{a-1} (1-\mu'_l) = 0$ since $\mu'_{a^*} = 1$ as $\mu' \in M$. Therefore, as $\hat{a} > a^*$ we have that

$$\begin{aligned} LE(\mu') &= \hat{a} - \sum_{a=0}^{\hat{\mathbf{a}}} (\hat{a} - (a+1)) * \mu'_a * \prod_{l=0}^{a-1} (1-\mu'_l), \\ &= \hat{a} - LGE_{\hat{a}}(\mu'), \end{aligned}$$

the desired result. ☐