Deprivation measures combining poverty and premature mortality

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

January 11, 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 25% 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, 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.
‡CRED, DEFIPP, University of Namur.
§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-
then rule down here over all the breathless dead.”
Achilles’ ghost to Odysseus, Homer, Odyssey.

Premature mortality is a major determinant of well being (Sen, 1998; Deaton, 2013). However, most measures of deprivation concentrate on deprivation among the living population and, thus, ignore premature mortality. This omission has important implications both to assess the level and evolution of deprivation as well as to guide and evaluate policies.\(^1\)

In this paper, we propose three measures of “total deprivation” that combine information on poverty and 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 mortality and deprivation among the living. Our theoretical results show that, when aggregating different dimensions of deprivation, premature mortality must be treated separately. The fundamental reason lies in the exclusive nature of this dimension: individuals cannot be dead and simultaneously suffering another form of deprivation. As a result, a premature mortality component must be added to an alive deprivation component. Moreover, we show that the premature mortality component must be measured in time units, i.e. the number of years lost due to premature death. The reason is that alive deprivation is also measured in time units since alive deprivation measures the number of (alive) individuals who are poor in a given year, i.e. a number of person-years. These findings run counter the construction of alternative multidimensional poverty indices that account for premature mortality, such as the MPI index of the World Bank.

Being deprived is generally understood as falling short of a minimal standard in a given resource. Since humans are mortal, being dead is not by itself a form of deprivation, but dying too young is. In this respect, the resource of interest is the number of years spent alive, i.e. the lifespan. An individual dying in young age is deprived in the sense that she will not live a number of years considered as minimally acceptable. She is victim of premature mortality and should be considered as lifespan deprived.\(^2\) In Section 2, we provide the theoretical 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 (PYAD) with person-years prematurely lost (PYPL). Given that these two forms of deprivation are mutually exclusive, it is natural to aggregate both dimensions by summing them. Table 1 illustrates the construction of our indices.

---

\(^1\)Quantified deprivation reduction targets have for example recently been adopted as major policy goals by the World Bank or the European Commission (European Commission, 2015; World Bank, 2015).

\(^2\)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: Example with age threshold equal to 50 years.

<table>
<thead>
<tr>
<th></th>
<th>Headcount</th>
<th>Lifespan</th>
<th>E(PY AD)</th>
<th>E(PYPL)</th>
<th>ALE</th>
<th>Deprivation Index</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Society A</strong></td>
<td>20%</td>
<td>60 years</td>
<td>12 years</td>
<td>0 years</td>
<td>60</td>
<td>20%</td>
</tr>
<tr>
<td><strong>Society B</strong></td>
<td>12.5%</td>
<td>40 years</td>
<td>5 years</td>
<td>10 years</td>
<td>50</td>
<td>30%</td>
</tr>
</tbody>
</table>

In Society A, 20% of the alive population is poor (as measured by the headcount ratio) and all individuals live for 60 years. A newborn in this group therefore expects to spend 12 years in alive deprivation (E(PY AD)=12). Taking 50 years as the age threshold defining lifespan deprivation, this newborn does not expect to lose prematurely any year of life (E(PYPL)=0). In Society B, 12.5% of the living population is poor and all individuals live for 40 years. A newborn in this society expects to spend 5 years in alive deprivation and to lose prematurely 10 years of her life. She therefore expects a total deprivation of 15 years of her life, which exceeds the deprivation in Society A (12 years), even though the latter is more affected by alive deprivation. As this example illustrates, total deprivation can be lower in situations characterized by a larger alive deprivation.

We propose, characterize axiomatically and compare three measures of total deprivation in a given calendar year: Inherited Deprivation (ID), Generated Deprivation (GD) and Expected Deprivation at birth (ED). ID is based on past mortality, and records individuals who, at time t, should be alive given the age threshold but have died prematurely. The two other indices are based on current 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 at time t, while ED is based on the number of years a newborn expects to lose prematurely, given the current mortality rates. As we shall show when comparing those three measures in non-stationary populations, this last measure has the lowest inertia (i.e. reacts the fastest to mortality shocks), is easily interpretable and requires less information. The three indices divide the sum of years spent in alive deprivation and years prematurely lost by a counterfactual number of years obtained when all premature deaths are postponed to the age threshold. In the case of ED, we refer to this counterfactual lifetime as the augmented life expectancy (ALE). In the above example, the ALE is 60 years in Society A and 50 years in Society B. Our indices (which in this particular example are identical) report the fraction of years in her ALE that a newborn expects to spend in deprivation or to lose prematurely. Hence, the share of ALE lost to deprivation is 20% in Society A and 30% in Society B.

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, proposed by Sen and Mahbub ul Haq. 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, we show that composite indices also fail to satisfy a simple separability property, which implies that composite indices are not consistent when comparing one person-year spent

---

3The concept of stationary population is defined in Subsection 2.3.
in deprivation and one person-year lost due to premature death. 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 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 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 consider that premature death masks an otherwise larger “true” alive deprivation. Rather, we consider premature death as a form of deprivation 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).

We provide an analysis of the evolution of total deprivation in the world over the past 25 years. In order to do so, we combine datasets on income deprivation (PovCalNet) and on mortality (Global Burden of Disease) for the period 1990 to 2014. Assuming that being prematurely dead is no better than being in income deprivation, which is arguably a conservative hypothesis, as illustrated by Achille’s opinion in the epigraph, we find that the omission of lifespan deprivation leads to an underestimation of global total deprivation of at least 20 to 25% during the period. 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, we show important differences between alive deprivation and total deprivation. We document that for more than 7% of country-years, total deprivation in fact evolves in the opposite direction as income deprivation. In addition, inter-country comparisons are changed when lifespan deprivation is taken into account: in 2014, the median variation in rank is 3. Deprivation assessments ignoring premature mortality are therefore seriously biased, and may lead to flawed policy evaluations.
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 premature mortality. We first present and characterize an index based on past mortality. We refer to this index as the inherited deprivation index (ID).

In period \( t \), each individual \( i \in \{1, \ldots, n\} \) 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) \( s_i \in S = \{NP, AP, D\} \). To keep terminology short, we often write that an individual in alive deprivation is “poor”. 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 \( t \).

Let \( a_i = t - b_i \) be the age that individual \( i \) would have in period \( t \) given her birth year \( b_i \). Premature mortality requires the definition of a normative lifespan threshold \( \hat{a} \). 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 \( x = (x_1, \ldots, x_n) \) specifies the birth year and the status at time \( t \) 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:

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

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 \( P : X \times \mathbb{N} \rightarrow \mathbb{R}_+ \), where \( P(x, \hat{a}) \geq P(x', \hat{a}) \) means that \( x \) has weakly more deprivation than \( x' \) 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 with three individuals

\[ x = ((\text{young}, NP), (\text{young}, D), (\text{old}, D)), \]

where a birth year more distant than \( \hat{a} \) years before \( t \) is noted as \( \text{old} \) (\( \text{young} \) otherwise). Because she is young and dead, individual 2 is prematurely losing period \( t \). We contrast distribution \( x \) with two distributions \( x' \) and \( x'' \) 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. These three distributions are compared in Table 2.

### Table 2: Comparing distributions using the headcount ratio and the Inherited Deprivation index

<table>
<thead>
<tr>
<th></th>
<th>(young, P)</th>
<th>(young, NP)</th>
<th>(young, D)</th>
<th>(old, D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Distribution</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Distribution</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

No individual is alive and poor, except individual 2 in distribution \(x''\). Therefore, the headcount ratios (HC) of distributions \(x\) (HC = \(\frac{0}{4}\)) and \(x'\) (HC = \(\frac{0}{7}\)) 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 \(\gamma\) 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 deprivation 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 \(P(x_i, x_{-i}) = P(x_{-i})\).

The inherited deprivation index captures premature mortality, and therefore cannot satisfy this property. Rather, it 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. Typically, 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.

\(^5\)The comparison of distribution \(x''\) to distribution \(x\) would be an example of the “mortality paradox”: if the poor individual of distribution \(x''\) was dead, the HC of distribution \(x\) and \(x''\) would be equal.
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

\[
P^\gamma_{ID}(x) = \frac{p(x)}{f(x) + p(x) + d(x)} + \gamma \frac{d(x)}{f(x) + p(x) + d(x)},
\]

where \( \gamma > 0 \) is a parameter weighing the relative importance of alive deprivation and premature mortality. An individual losing prematurely year \( t \) matters \( \gamma \) times as much as an individual spending year \( t \) in alive deprivation.

Index \( P^\gamma_{ID} \) 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 year \( t \), and the lifespan deprivation component records the number of persons who have died before and were born less than \( \hat{a} \) years before \( t \). The denominator of both components is identical and equal to the number of individuals in the reference population, which includes alive and prematurely dead individuals.

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, \( P^\gamma_{ID} \) 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 \( x' \), \( P^\gamma_{ID}(x) > P^\gamma_{ID}(x') \). In addition, as individual 2 is prematurely dead in distribution \( x \) whereas she is alive and poor in \( x'' \), \( P^\gamma_{ID}(x) \geq P^\gamma_{ID}(x'') \) if \( \gamma \geq 1 \). 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 born in year \( b_1 \) and 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 (we discuss this below). More formally,

\(^6\)This is why Weak Independence of Birth Year has a precondition holding 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.
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 $P(x, x_{-i}) = P(x', x_{-i})$.

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 $(x, x') \in X$, 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 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 (Characterization of ID).

$P$ is ordinally equivalent to $P_{ID}$ 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. \(\square\)

First, Proposition 1 fully characterizes the ID index, which implies that any deprivation index satisfying our properties ranks distribution in exactly the same way as the ID index. Second, this proposition implies that alive deprivation and premature mortality 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 individuals of different status across different states.\(^9\)

A number of remarks are in order. First, 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 $\gamma$, the parameter weighing the relative importance of alive deprivation and premature mortality. Second, 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.

\(^7\)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 $f(x') + p(x') = f(x'') + p(x'')$.

\(^8\)Formal definitions of these traditional axioms can be found in Appendix 7.1.

\(^9\)This is in contrast to sub-dimensions of alive deprivation, which can simultaneously affect the same individual.
the period \( t \) following her death). In that sense, the death of a younger individual matters more.

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). Moreover, we defined alive deprivation as a categorical variable for the sake of simplicity. 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 information implies that all our measures are indifferent about the repartition across individuals of periods spent in alive deprivation and periods prematurely lost.

The ID index suffers from two limitations arising from the fact that total deprivation in \( t \) depends on past mortality. First, computing this index requires detailed information on mortality of each age cohort for all \( a \) 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 \( a - (a_i + 1) \) 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 rely on the same intuition as the ID index, but are based on mortality rates in period \( t \) instead of on past mortality. Because they rely on the same intuition, and are meant to represent the same underlying phenomena, all three indices must offer the same diagnostic in the long run: they follow the “ID equivalence” axiom, which we formalize below. However, given their different construction, they exhibit different short run dynamics, as discussed in Subsection 2.4.

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 \( s_i \neq D \). 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 $\mu_a$ 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 $n_a(x) \ast \mu_a$. We have that $\mu_a \in \mathcal{M} = [0, 1] \cap \mathbb{Q}$, where $\mathbb{Q}$ is the set of rational numbers.\footnote{Given that distributions have finite numbers of individuals, mortality rates cannot take irrational values.} 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 $\mu$ summarizes current mortality, while distribution $x$ summarizes current alive deprivation and past mortality.\footnote{Observe again that this framework is consistent with our data-constraint. A pair $(x, \mu)$ does not simultaneously contain the information about an individual’s deprivation and the information about her chances of survival.} The set of mortality vectors is given by

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

We consider pairs $(x, \mu) \in O = X \times M$, in which the distribution $x$ is a priori unrelated to vector $\mu$. The only constraint is that the age specific mortality rates $\mu_a$ must be 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 an extension of domain $X$. 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 past mortality and 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. If the distribution and mortality vector defining a pair are in general unrelated, there are particular pairs whose distribution exactly mirrors the mortality vector. Such pairs arise for stationary populations. A population is stationary if the number of newborns and the mortality vectors are constant over time. We say that a pair $(x, \mu)$ is stationary if, for some $n^* \in \mathbb{N}$ and all $a \in \{0, \ldots, a^*\}$, we have:

- $n_{a+1}(x) = n_a(x) \ast (1 - \mu_a)$ (demographic equilibrium),
- $n_a(x) + d_a(x) = n^* \in \mathbb{N}$ (constant natality).

In a demographic equilibrium, 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 current mortality and natality rates remain constant over time (see for instance Preston et al. (2000)). In a stationary pair, past and current mortality coincide and the mortality vector $\mu$ does not convey any information that cannot be inferred from the distribution $x$. Indeed, the population of prematurely dead individuals in $x$ directly reflects $\mu$: the number of prematurely dead individuals in $x$ can be computed from $\mu$ (and $n^*$) and, conversely, the mortality vector $\mu$ can be computed from $x$. As a result, for stationary pairs, a deprivation index based on the mortality rates must be identical to a deprivation index computed on the actual number of dead in the distribution. For stationary pairs, we therefore require that a deprivation index is equal to the ID index.
Deprivation axiom 6 (ID Equivalence). For all \((x, \mu) \in O\) and some \(\gamma > 0\), if \((x, \mu)\) is stationary, then \(P(x, \mu) = P^{ID}(x)\).

Such equivalence is a minimal requirement for deprivation indices based on current mortality. Indeed, a permanent change in mortality rates affects the long-run distribution, assuming constant natality. 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:

\[
P^{GD}_\gamma(x, \mu) = \frac{p(x)}{f(x) + p(x) + d^{GD}(x, \mu)} + \gamma \frac{d^{GD}(x, \mu)}{f(x) + p(x) + d^{GD}(x, \mu)},
\]

where \(d^{GD}\) counts the number of person-years prematurely lost generated by deaths occurring in year \(t\):

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

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 occurring in year \(t\). By contrast, the ID index records the number of PYPLs inherited in year \(t\), which were generated by deaths occurring before year \(t\). 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 PYADs 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 year \(t\). It again combines an alive deprivation and a lifespan deprivation component in an additive way:

\[
P^{ED}_\gamma(x, \mu) = \frac{LE(\mu) \ast HC(x)}{LE(\mu) + LGE_\gamma(\mu)} + \gamma \frac{LGE_\gamma(\mu)}{LE(\mu) + LGE_\gamma(\mu)},
\]

where \(HC(x) = \frac{p(x)}{p(x) + f(x)}\) is the headcount ratio and \(LE(\mu)\) is life expectancy at birth: \(^{12}\)

\[
LE(\mu) = \sum_{a=0}^{\gamma} a^{-1} \prod_{l=0}^{a-1} (1 - \mu_l).
\]

We interpret the term \(LE(\mu) \ast HC(x)\) as the expected number of years a newborn in year \(t\) will spend in alive deprivation, given the current mortality rates and headcount ratio. It represents the expected person-years spent in alive deprivation for such a

\(^{12}\)Life expectancy measures the expected lifespan of an individual facing throughout her life the age-specific mortality rates reported in vector \(\mu\).
newborn. The last term, \( LGE_\hat{a} \), is the lifespan gap expectancy relative to the age threshold:

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

\( LGE_\hat{a} \) measures the number of years that a newborn expects to lose prematurely if confronted to the mortality rates of vector \( \mu \) throughout her first years of life.\(^{13}\) The ED index therefore takes the viewpoint of a newborn and compute the expected proportion of 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 at time \( t \), while ED, by taking the viewpoint of a newborn, is actually based on an abstract, counterfactual population pyramid defined by the current mortality rates. 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 current mortality rates, 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 \( P_{GD}^\gamma \) and \( P_{ED}^\gamma \) satisfy ID Equivalence.

**Proof.** See Appendix 7.2. \( \square \)

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 a demographic equilibrium can be illustrated by graphical representations of \( LE \) and \( LGE_\hat{a} \), as shown in the left panel in Figure 2 for a population pyramid in demographic equilibrium. 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.\(^{14}\) When compared with the right panel in Figure 1, the right panel in Figure 2 reveals that ED and GD take

\(^{13}\)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).

\(^{14}\)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.
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 $a^{GD}(x, \mu)$.

Figure 2: For the pair $(x, \mu)$ in demographic equilibrium, the green area corresponds to a multiple of life-expectancy $(LE)$ and the pink area to a multiple of the lifespan gap expectancy $(LGE_a)$.

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 $\gamma$ 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).\footnote{See Appendix 7.3 for the formal definitions of these axioms.}

**Proposition 2** (Characterization of ED).

$P = P^{ED}$ 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 current population pyramid, and therefore avoids the inertia associated with the demographic evolution of population pyramids. This advantage comes at a price, as ED cannot be decomposed additively 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

\[
P(x, \mu) = \frac{\eta(x', \mu') \ast P(x', \mu') + \eta(x'', \mu'') \ast P(x'', \mu'')}{\eta(x', \mu') + \eta(x'', \mu'')},
\]

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 existing population pyramid. We further investigate the relation between our three indices by contrasting their 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 4 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 4 (\(P_{ED}^{\gamma}\) and \(P_{GD}^{\gamma}\) weigh \(\mu\) 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 \(a \in \{0, \ldots, a^* - 1\}\). Let \(\mu^*\) be the mortality vector for which \((x, \mu^*)\) is a demographic equilibrium. If \(\gamma \geq 1\), then we have \(P_{GD}^{\gamma}(x, \mu) \geq P_{ED}^{\gamma}(x, \mu)\).
if and only if

$$\sum_{a=0}^{\hat{a}-1}(\hat{a} - (a + 1)) \cdot \frac{n_a}{n_0}(\mu) \cdot \mu_a \geq \sum_{a=0}^{\hat{a}-1}(\hat{a} - (a + 1)) \cdot \frac{n_a}{n_0}(\mu^z) \cdot \mu_a$$

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

Proof. See Appendix 7.5.

This result shows that the GD index is larger than the ED index as long as the age-cohorts with a large value of $(\hat{a} - (a + 1)) \cdot \mu_a$ have a larger relative size in the current population pyramid than in the counterfactual population pyramid associated with $\mu$. 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 associated with a larger number of PYPLs.

One may object it is pointless to compare the size of two different ordinal indices, as we do in $P_{GD}^\gamma(x, \mu) \geq P_{ED}^\gamma(x, \mu)$. However, when $\gamma \geq 1$ this inequality is equivalent to comparing the relative size of their two components, i.e. equivalent to:

$$\frac{d_{GD}^\gamma(x, \mu)}{p(x) + d_{GD}^\gamma(x, \mu)} \geq \frac{LGE_a(\mu)}{LE(\mu) \cdot HC(x) + LGE_a(\mu)}.$$

Hence, inequality $P_{GD}^\gamma(x, \mu) \geq P_{ED}^\gamma(x, \mu)$ implies that GD emphasizes more lifespan deprivation than ED does. This inequality is verified in virtually all our empirical results. A simple explanation can be found in (5). 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 Appendix 8.1, 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 1 for $\gamma$ and 3 for $\hat{a}$, 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 at period zero, such that all individuals die: $\mu^0 = (1, 1, 1)$. After

\[15\]
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.\footnote{Note that the fact that the index returns to its initial value after one period is a particularity of this simple example. If instead we had $n^* = 4$, $\mu^0 = (\frac{1}{4}, 1, 1)$ and $\mu^* = (0, \frac{1}{4}, 1)$, the index would not return to its stationary value in period 1.}

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 \( \mu^* \), corresponds to the stationary population, with \( LE(\mu^*) = 3 \) and \( LGE_a(\mu^*) = 0 \), 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 \( \mu^* \) returns to its stationary value.

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

**Permanent mortality shocks**

We now investigate the consequences of a permanent mortality shock on a population in a demographic equilibrium. 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 \( \Pi_{1}^{ID} \), \( \Pi_{1}^{GD} \) and \( \Pi_{1}^{ED} \), and assume that natality is constant, and there is no alive deprivation. The age threshold is 50 and the maximal age is 100. Before the shock, the population pyramid is in equilibrium with a mortality vector such that at each age before 100, the mortality rate is equal to 1%. 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
Figure 4: Simulation of permanent mortality shocks on a stationary population

is thus $\mu = (0, 0, 1, \ldots)$. We assume that the age threshold, $\hat{a}$, is equal to 12, and $\gamma = 1$. 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 $t^*$, 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 $\mu^* = (0, 1, 0, 1, \ldots)$. Table 3 summarizes the evolution of this population after this permanent shock.

Two individuals die at the end of period $t^*$ and the GD index records 18 PYPLs. Given that four individuals lived in period $t^*$, GD is equal to $18/22$. In period $t^* + 1$, 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 demographic equilibrium 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 $t^* + 1$ but decreases in $t^* + 2$. By contrast, the ED index remains constant and equal to $10/12$ over these three periods.
Table 3: Non-monotonicity of GD indices after permanent mortality shock.

<table>
<thead>
<tr>
<th>period</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>. . .</th>
<th>11</th>
<th>GD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t &lt; t^*$</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>D</td>
<td>. . .</td>
<td>D</td>
<td>$\frac{18}{25} = 0.66$</td>
</tr>
<tr>
<td>$t^*$</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>D</td>
<td>. . .</td>
<td>D</td>
<td>$\frac{18}{25} = 0.82$</td>
</tr>
<tr>
<td>$t^* + 1$</td>
<td>NP</td>
<td>NP</td>
<td>D</td>
<td>NP</td>
<td>D</td>
<td>. . .</td>
<td>D</td>
<td>$\frac{18}{20} = 0.86$</td>
</tr>
<tr>
<td>$t^* + 2$</td>
<td>NP</td>
<td>NP</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>. . .</td>
<td>D</td>
<td>$\frac{18}{20} = 0.83$</td>
</tr>
</tbody>
</table>

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 $t^* + 1$ implies that the mortality vector $\mu^*$ does create more PYPPLs in $t^* + 1$ than in $t^* + 2$. So the GD index conveys correct information about actual deprivation. However, a fixed mortality vector $\mu^*$ 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

We first examine composite indices and show that they are typically inconsistent. To do so, we define the mortality statistic $L P_\hat{a}(\mu) = \frac{L G E_\mu(\mu)}{a}$, which measures the fraction of the normative threshold that a newborn expects not to live. $L P_\hat{a}$ 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, $L P_\hat{a}$ is unaffected by changes in mortality in old ages (above $\hat{a}$) but strictly decreases when mortality in young age is reduced. Also, $L P_\hat{a}$ decreases when the death of a lifespan deprived individual is postponed by one year. A formal proof is given in
Appendix 8.2.

At this stage, one may think that the ideal index of total deprivation aggregates alive deprivation, $HC$, and lifespan deprivation, $LP_{\alpha}$, by simply weighing these two indicators. By doing so, we obtain a composite index of total deprivation, defined as:

$$P_w(x, \mu) = w \cdot HC(x) + (1 - w) \cdot LP_{\alpha}(\mu),$$

where weight $w \in [0, 1]$.\(^{18}\)

We now show, with the help of an example, that indices based on a simple weighted sum (such as composite indices) are not consistent as they do not attribute a fixed relative weight to one PYPL compared to one PYAD. In other words, when comparing stationary pairs, they violate 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 relevant population.\(^{20}\)

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

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. We now show that, depending on the fraction of the alive population that is poor, a composite index $P_w^C$ may increase or decrease.

We illustrate this by looking at three stationary populations, presented in Table 4. The three populations have two newborns every year and individuals live for at most two years. In population 1, the two young individuals are not poor but die at the end of their birth year. There is no old individual alive. In population 2, only one young individual dies, the other survives but becomes poor in the second year of his life. In population 3, the two young individuals survive but are poor in their second year. In each year, the number of alive individuals is equal to 2 in population 1, 3 in population 2 and 4 in population 3. Across the three populations, the headcount ratio varies from 0 to 0.33 and 0.5. Assuming a normative lifespan of two years, lifespan deprivation $LP_2$ varies from 0.5 in population 1 (one year alive in two years) to 0.25 and 0 in the two other populations.

Assuming a weight $w$ equal to 0.5 for each of its two components, the composite index $P_w^C_{0.5}$ is equal to 0.25 in population 1 (0.5+0.5=1), 0.25 in population 2 (0.5+0.5=1) and 0.25 in population 3 (0.5+0.5=1). As a result, when comparing population 1 to population 2, $P_w^C_{0.5}$ considers that one PYAD is worse than one PYPL. However, when comparing population 2 to population 3, $P_w^C_{0.5}$ considers

\(^{18}\)We show in Appendix 8.2 that both the GD index and the ED index violate one of the two basic properties of a lifespan deprivation indicator, i.e. Current Mortality Focus or Current Mortality Monotonicity below $\bar{a}$.

\(^{19}\)This index satisfies the two basic properties of a lifespan deprivation indicator, i.e. Current Mortality Focus and Current Mortality Monotonicity below $\bar{a}$, defined in Appendix 8.2.

\(^{20}\)Even if index $P^C$ 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 $(x, x_{-i}), (x', x'_{-i}), (\mu), (\mu') \in O$ that are stationary, if $n(x) = n(x')$ and $b_i = b_i' > t - \bar{a}$, then $P((x, x_{-i}), \mu) \geq P((x', x'_{-i}), \mu') \iff P((x', x'_{-i}), \mu') \geq P((x, x_{-i}), \mu')$.  

20
Table 4: Composite indices are not consistent (even on stationary populations)

<table>
<thead>
<tr>
<th>Population</th>
<th>HC</th>
<th>LP</th>
<th>$P_{0.5}^C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population 1</td>
<td>NP, NP</td>
<td>D,D</td>
<td>0</td>
</tr>
<tr>
<td>Population 2</td>
<td>NP, NP</td>
<td>AP,D</td>
<td>0.33</td>
</tr>
<tr>
<td>Population 3</td>
<td>NP, NP</td>
<td>AP,AP</td>
<td>0.5</td>
</tr>
</tbody>
</table>

that one PYAD is better than one PYPL. $P_{0.5}^C$ is therefore not consistent when comparing total deprivation across these different populations.\(^{21}\) The same inconsistency also characterizes the Human Deprivation Index, a composite index which aggregates both premature mortality and alive deprivation using the headcount ratio in both dimensions (Watkins, 2006).\(^{22}\) By contrast, across the three populations, $P_{1}^{ID}$, $P_{1}^{GD}$ and $P_{1}^{ED}$ are all equal to 0.5. Our indices are equal for all populations as we consider a value of $\gamma$ equal to 1, which weighs equally one PYAD and one PYPL. A higher value of $\gamma$ involves a decreasing value of the index across the three populations.

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. Thus in population 2 of the above example, the headcount ratio is computed relative to a population of 3 alive individuals, while lifespan deprivation is based on a normative lifespan of 2 periods per individual (a total of 4 periods). 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 PYAD with the number of PYPL before normalizing by the same reference population. (Thus, in population 2, we have one person-year prematurely lost, one person-year spent in deprivation, and four person-years as the reference population.) As a result, the relative weight attributed to one PYPL over one PYAD therefore depends on the levels of alive deprivation and life expectancy.

Preference-based indicators and weight $\gamma$

We now compare our indices to the preference-based indicators used in Gary S. Becker and Soares (2005); Jones and Klenow (2016). 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 $\gamma$. By contrast, preference-based indicators are based on explicit preferences and, therefore, on a particular utility function. Clearly, measuring deprivation is less ambitious than measuring well-being, given that deprivation indices disregard the actual achievements of non-

\(^{21}\)The general problem illustrated here 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 $P_{0.5}^C$ is not consistent.

\(^{22}\)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.

\(^{23}\)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, our total deprivation indices is identical to alive deprivation, as measured by $HC$. 

21
Moreover, focusing on deprivation instead of well-being naturally leads us to use a fixed weight $\gamma$. 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 $\gamma$ (see Lemma 4 in 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:

$$P_{1 ED}(x, \mu) = \frac{LE(\mu) * HC(x) + LGE_a}{LE(\mu) + LGE_a(\mu)}.$$  

This index takes the perspective of a newborn that would be confronted throughout 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 $\mu^a$ constructed from $\mu$ by postponing all premature deaths in $\mu$ to the age threshold. 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 $t$.

---

24 It may seems odd that the ED index combines a headcount 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.

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

---

22
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. However, this argument is weak as such information is sometimes 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 separately. To make our indices 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.\(^{26}\) However, to compute such concentration-sensitive indices of total poverty, we not only need mortality rates by income groups but also information on mobility in and out alive deprivation across consecutive periods, a type of information which is typically not available.\(^{27}\)

4 Aggregate deprivation

4.1 The data

We apply our indices of total deprivation to real world data, focusing in particular on ED. In the process, we compare our results to those using more conventional

\(^{26}\)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.

\(^{27}\)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.
deprivation measures, such as the HC.\textsuperscript{28} The definition of ED requires a value for the age threshold $\hat{a}$ and the weight $\gamma$. 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 $\gamma$, 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 typically much lower than the median age at death observed in our data (which is 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) as well as the GD index, and we present these alternative results in the Online Appendix of the paper.

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.\textsuperscript{29} Details on the method used are given in the appendix of Mortality and of Death Collaborators (2016).\textsuperscript{30}

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.\textsuperscript{31} 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

\textsuperscript{28}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 country and years available in this database are very different from those for which comparable alive deprivation data is available.

\textsuperscript{29}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 ? for a critique of this approach).

\textsuperscript{30}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.

\textsuperscript{31}The website address is http://iresearch.worldbank.org/PovcalNet/povOnDemand.aspx.
in Chen and Ravallion (2013).\textsuperscript{32} 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 ED index, 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 1 for a list of those countries).

4.2 World deprivation

Figure 5 presents the evolution of world’s total deprivation, as measured by GD and ED, and of their two components, alive deprivation and lifespan deprivation. 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” lifetime 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.\textsuperscript{33} Also, the relative importance of lifespan deprivation increased over time: its share in total deprivation increased from 20\% in 1990 to 25\% 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 25\% 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.

A second point to note is that these three measures follow parallel trends. They do 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” expected lifetime in deprivation, this proportion fell down to 16\% in 2014. One should note that the alive deprivation component is systematically lower than HC: this is a mechanical effect due to the fact that HC does not take into account prematurely dead individuals. Also the difference between the alive deprivation component and HC falls over time, a consequence of both an increase in life expectancy and a decrease in lifespan gap.

\textsuperscript{32}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, it appeared that income deprivation was not always as reactive as lifespan deprivation, which may be due to the interpolated nature of the data. In the following, we therefore do not analyze these events.

\textsuperscript{33}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.
expectancy over time. 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.

**Figure 5:** Comparing world deprivation and deprivation rates

![Graph 1: Expected Deprivation at birth with \( \hat{a} = 50 \)](image1)

![Graph 2: Generated Deprivation with \( \hat{a} = 50 \)](image2)

### 4.3 Regional deprivation

Figure 6 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. 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, 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)).

---

34Remember that our measure of alive deprivation is \( \frac{LE(\mu) \times HC(x)}{LE(\mu) + LGE(\mu)} \) for ED while \( HC(x) = \frac{LE(\mu) \times HC(x)}{LE(\mu) + LGE(\mu)} \). Therefore, as \( \frac{LGE(\mu)}{LE(\mu)} \) decreases, HC and alive deprivation converge.

35In our data, the Middle East and North Africa region only includes Djibouti, Iran, Morocco and Tunisia.
In addition, note that the relative position of each region as measured with ED is similar to that measured with HC: the income-poor regions are also the most deprived ones.

**Figure 6:** Comparing continental deprivation and deprivation rates. ED with $\hat{a} = 50$

![Graphs showing continental deprivation and deprivation rates](image)

5 **Countries’ deprivation**

5.1 **Country’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 7 maps the median value 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 7:** Map of LD/ED ratio, 1990 to 2014 ($\alpha = 50$)

Individual countries are often compared and ranked according to a measure of deprivation (?). 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 8 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 32nd and Morocco 33rd in 1993. When the ranking is based on ED, Gabon was 46th and Morocco 38th. Table 5 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 $67 \times 5.9\% = 3.9$ years in alive deprivation against 3.4 in Gabon. However, as early mortality is higher in Gabon than in Morocco, a newborn in Gabon expects to lose 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 expectation to deprivation, as against 15.2% in Gabon.

More generally, our measures lead to substantial re-rankings across countries. Indeed, throughout the period, the average change in ranking across all countries is

---

36 We use the median rather than the mean to prevent extreme values to bias the general picture.
Figure 8: Examples of re-rankings: Gabon and Morocco. HC and ED with $\hat{a} = 50$

![Graphs showing re-rankings for Gabon and Morocco.]

Table 5: Decomposition of re-rankings: Gabon and Morocco in 1993.

<table>
<thead>
<tr>
<th>Country</th>
<th>HC</th>
<th>LE</th>
<th>$E(PYAD)$</th>
<th>$LGE_{50}$</th>
<th>$E(PYAD) + LGE_{50}$</th>
<th>$P_{ED}^E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morocco</td>
<td>5.9</td>
<td>67</td>
<td>3.9</td>
<td>4.5</td>
<td>8.4</td>
<td>11.8</td>
</tr>
<tr>
<td>Gabon</td>
<td>5.8</td>
<td>59</td>
<td>3.4</td>
<td>6.6</td>
<td>10</td>
<td>15.2</td>
</tr>
</tbody>
</table>

equal to 3.4 ranks. How are these changes distributed across countries and time? Figure 9 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.

Figure 9: Map of re-rankings of countries. ED with $\hat{a} = 50$ vs HC

![Map showing re-rankings across countries.]

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.

How do these re-rankings evolve over time? Figure 10 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 11 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 increase in alive deprivation. By contrast, in Botswana during the 1990s, HC decreased substantially, while ED increased dramatically due to the HIV epidemics.

Figure 11: Differences in trends between ED and HC: Comoros and Botswana. HC and ED with $\hat{a} = 50$
In Table 6, 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 expectancy 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 6: Example of evolution reversals: Comoros and Botswana**

<table>
<thead>
<tr>
<th>Year</th>
<th>HC</th>
<th>LE</th>
<th>E(YAD)</th>
<th>LGE50</th>
<th>E(YAD) + LGE50</th>
<th>P\textsuperscript{ED}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td></td>
<td>years</td>
<td>years</td>
<td>years</td>
<td>%</td>
</tr>
<tr>
<td>Comoros</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>11.2</td>
<td>57</td>
<td>6.4</td>
<td>8.5</td>
<td>14.9</td>
<td>22.7</td>
</tr>
<tr>
<td>2014</td>
<td>14.0</td>
<td>66</td>
<td>9.2</td>
<td>4.5</td>
<td>13.7</td>
<td>19.5</td>
</tr>
<tr>
<td></td>
<td>+ 25%</td>
<td>+ 12.6</td>
<td>+2.8</td>
<td>-4.0</td>
<td>-1.2</td>
<td>- 14%</td>
</tr>
<tr>
<td>Botswana</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>33.6</td>
<td>63.6</td>
<td>21.5</td>
<td>4.8</td>
<td>26.3</td>
<td>38.2</td>
</tr>
<tr>
<td>2000</td>
<td>29.5</td>
<td>45.6</td>
<td>13.5</td>
<td>11.1</td>
<td>24.6</td>
<td>43.3</td>
</tr>
<tr>
<td></td>
<td>- 12%</td>
<td>- 11</td>
<td>-8</td>
<td>+ 6.3</td>
<td>-1.7</td>
<td>+ 13%</td>
</tr>
</tbody>
</table>

How often do these opposite diagnostics arise in the last 25 years? In Figure 12, 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 \( \gamma \to \infty \).

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.
Figure 12: 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.\(^{37}\)

Our aggregation method allows placing an explicit and meaningful lower bound on the normative trade-off (the weight \( \gamma \)) 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.

References


7 Appendix 1: Proofs

7.1 Proof of Proposition 1

We first provide a formal definition of the three axioms left undefined defined 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 \( P(x) = P(x') \).

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 \( k \) sufficiently large we have \( P(x^k, z) > P(y) \) and \( P(x) > P(y^k, z) \).

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

Let \( \mathbb{Q}^+ \) denote the set of non-negative rational numbers. Consider \( \Delta \), 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 \( P(x) = P(x') \).

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. \( X^* = \{ x \in X \mid b_i > t - \hat{a} \text{ for all } i \text{ for whom } s_i = D \} \). Let mapping \( m^1 : X \to X^* \) return for any \( x \in X \) the image \( x^* = m^1(x) \) with \( n(x^*) = f(x) + p(x) + d(x) \) defined as \( x_i^* = x_j \), 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.

Let the set of status \( S = \{ NP, AP, D \} \), where \( D \) stands here for prematurely dead. Let mapping \( m^2 : X^* \to S^N \) return for any \( x \in X^* \) the image \( v = m^2(x) \) with \( n(v) = n(x) \) defined as \( v_i^* = s_i \). By construction, we have \( m^2(X^*) = S^N \). Also, 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.

Consider the set \( \mathbb{N}_0^3 \setminus \{0,0,0\} \), which contains all triplets of numbers in \( \{0, 1, 2, \ldots\} \) except the nul triplet \( (0,0,0) \). Consider the mapping \( m^3 : S^N \to \mathbb{N}_0^3 \setminus \{0,0,0\} \) that counts the number of individuals exhibiting each status. That is, mapping \( m^3 \) returns for any \( v \in S^N \) the image \( w = m^3(v) \) such that \( w'_i = \#\{ i \leq n(v) \mid v_i = NP \} \), \( w'_i = \#\{ i \leq n(v) \mid v_i = AP \} \) and \( w'_i = \#\{ i \leq n(v) \mid v_i = D \} \).38 By construction, we have \( m^3 \circ m^2(X^*) = \mathbb{N}_0^3 \setminus \{0,0,0\} \).39 Also, any two \( x, x' \in X^* \) for which

---

38 For any set \( A \), we denote the cardinality \( A \) by \( \#A \).
39 The composite mapping \( m^3 \circ m^2 : X^* \to \mathbb{N}_0^3 \setminus \{0,0,0\} \) is defined by \( m^3 \circ m^2(x) = m^3(m^2(x)) \).
$m^3 \circ m^2(x) = m^3 \circ m^2(x')$ are such that $P(x) = P(x')$ by Anonymity and Weak Independence of Birth Year.

Let mapping $m^4 : N_0^3 \setminus (0,0,0) \to \Delta$ count the fraction of individuals exhibiting each status. That is, for any $w \in N_0^3 \setminus (0,0,0)$ the image $v = m^4(w)$ is defined as

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

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 $c_1, c_2, e_1, e_2, e_3 \in \mathbb{N}$ such that $(v_1, v_2, v_3) = (c_1/e_1, c_2/e_2, e_3/e_3)$. Consider any distribution $x$ with $m(x) = e_1e_2e_3$, 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 $c_1e_2e_3 + c_2e_1e_3 + c_3e_1e_2 = e_1e_2e_3$. As a result, all individuals in $x$ who are dead are prematurely dead, hence, $x \in X^*$. By construction, 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 $w, w' \in N_0^3 \setminus (0,0,0)$ 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 $w, w' \in N_0^3 \setminus (0,0,0)$ for which $m^4(w) = m^4(w')$ are such that for some $w'' \in N_0^3 \setminus (0,0,0)$ and two $k, k' \in \mathbb{N}$ we have that $w = kw''$ and $w' = k'w''$. Then, there exist $y, y', y'' \in X^*$ with $m^3 \circ m^2 \circ m^1(y) = w$, $m^3 \circ m^2 \circ m^1(y') = w'$ and $m^3 \circ m^2 \circ m^1(y'') = w''$ such that $y$ is a $k$-replication of $y''$ and $y'$ is a $k'$-replication of $y''$. By Replication Invariance, we have that $P(y) = P(y') = P(y'')$, the desired result.

**Step 2:** Using mapping $m$ to define an ordering $\succeq$ on $\Delta$ that is equivalent to the ordering represented by $P$.

Using Step 1, we now show that the deprivation index $P$ is equivalent to some complete ordering on $\Delta$. Let $\succeq$ be an ordering on $\Delta$ 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 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

$$P(x) \leq P(x') \iff v \succeq v'.$$

**Step 3:** Identifying a particular ID index $P_y$. 

35
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 \( v \succeq \lambda v + (1 - \lambda)v' \). 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 (6), we have \( P(x) < P(y) \). By definition of \( \lambda \), there exists \( c, e \in \mathbb{N} \) such that \( \lambda = c/e \). Let \( x^e \) be a \( c \)-replication of \( x \), \( x^{(e-c)} \) be a \((e-c)\)-replication of \( x \), \( y^e \) 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 (6).

Second, we derive a value of \( \gamma > 0 \) defining the candidate ID index \( P_\gamma \). 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 (6), 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 \( v^{001} \), denoted by \( E_{001}^{100} = \{ v \in \Delta \mid v_2 = 0 \} \). As \( v^{100} \succ v^{001} \), the convexity property implies that for any \( v, v' \in E_{001}^{100} \), 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 vertices \( v^{100} \) and \( v^{001} \) such that for any \( v \in E_{001}^{100} \), if \( v_1 > v'_1 \) then \( v \succ v' \) and, if \( v_1 < v'_1 \) then \( v \prec v' \). 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 irrations. Therefore, \( v^* \) is the unique element with these properties.

We show that if \( v^* \in \Delta \), then \( v^* \sim v^{010} \). Consider the contradiction assumption that \( v^* \in \Delta \) and \( v^* \succ v^{010} \). We build a \( v' \in E_{001}^{100} \) 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 \( v \in E_{001}^{100} \) with \( v'_1 < v^*_1 \) we have \( v' \prec v^{010} \). We construct \( v' \in E_{001}^{100} \) 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 (6) 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 (6), we have \( v' \succ v^{010} \). As \( v^* \in E_{001}^{100} \), we have by construction that \( v' \in E_{001}^{100} \) and \( v'_1 < v^*_1 \), the desired result.

We take \( \gamma = \frac{1}{v^*_2} \). 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 \( \gamma > 1 \).

---

40The alternative contradiction assumption for which \( v^* \in \Delta \) and \( v^* \prec v^{010} \) also leads to an impossibility.
• Case 3: \(v^{010} \prec v^{001}\).

The construction of \(\gamma\) is similar to that proposed in Case 2. We find the unique element \(v^{**} \in \Delta^R\) 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. \(v^{010} \succ v^{001}\). The proof for Case 1 is simpler and the proof for Case 3 is very similar.

Step 4: Show that \(P_\gamma\) is ordinally equivalent to \(P\).

Let function \(F: \Delta^R \to \mathbb{R}\) be defined by \(F(v) = -(v_2 + \gamma v_3)\). By construction of mapping \(m\) and the definition of \(P_\gamma\), 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 \(\Delta\), then we get from (6) that \(P_\gamma\) is ordinally equivalent to \(P\), the desired result.

First, we show that for any \(v \in \Delta\) we have \(v \geq v^{010}\) 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 \(\Delta\) 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 \(\Delta^{001} = \{v \in \Delta \mid F(v) < -1\}\).

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 \(E_{010}\) with \(v'_1 > v'_1\). This construction is illustrated in Panel A of Figure 13. Given that any \(v' \in E_{010}\) 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 \(v' : (1 - \frac{v_3}{v_3^{101}}, 0, \frac{v_3}{v_3^{101}})\). As \(v \in \Delta\), the definition of \(v'\) is such that \(v' \in E_{010}\). Let \(v'' = \lambda v^{010} + (1 - \lambda)v'\) where \(\lambda = v_2 \in [0, 1]\) 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 \(1 - \frac{v_3}{v_3^{101}} > 1 - \frac{1}{\gamma}\), which simplifies to \(\gamma > \frac{v_3^{101}}{v_3}\). This inequality holds because, as \(v \in \Delta^{100}\), we have \(F(v) > -1\), which simplifies to the same inequality.

Figure 13: 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.

---

\[41\] We have defined \(F\) and \(\gamma\) such that \(F(v^*) = F(v^{010})\). From a geometric perspective, the set of elements \(v \in \Delta^R\) 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 \(\Delta\) is the vertex \(v^{010}\) and, therefore, \(\Delta^{010}\) degenerates to \(\Delta^{100}\). The subset \(\Delta^{100}\) contains vertex \(v^{100}\) and all elements of \(\Delta\) that are on \(v^{100}\)'s side of the segment connecting \(v^{010}\) with \(v^*\). In turn, \(\Delta^{001}\) contains vertex \(v^{001}\) and all elements of \(\Delta\) that are on \(v^{001}\)'s side of the segment connecting \(v^{010}\) with \(v^*\).
Geometrically, we have just shown that the segment connecting \( v^{010} \) to \( v^* \) is an “implicit” indifference curve of \( \succeq \).

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 13. This case is such that there exists a \( \hat{v} = (0, \hat{v}_2, 1 - \hat{v}_2) \in E_{001}^{010} \) 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 \( v'' \in E_{001}^{100} \) if \( v'' \succ \hat{v} \), if \( v'' \prec \hat{v} \) then \( v'' \prec v' \). Geometrically, we have just shown that the segment connecting \( v \) to \( \hat{v}^* \) is parallel to the segment connecting \( v^{010} \) to \( v^* \), which is \( \hat{v} = \frac{\hat{v}_1}{\hat{v}_2} \). Consider the contradiction assumption for which \( \hat{v}_2 > \frac{\hat{v}_1}{\hat{v}_2} \).

Assume that \( \hat{v}^* \in \Delta^{\mathbb{R}^+} \). Consider now \( \hat{v}' = (\frac{\hat{v}_2}{\hat{v}_1}, 0, 1 - \frac{\hat{v}_2}{\hat{v}_1}) \in E_{001}^{100} \). By the contradiction assumption, we have \( \hat{v}_1 < \hat{v}_1^* \) and, hence, \( v^{010} \succ \hat{v}' \). By construction, for the rational \( \lambda = \hat{v}_2 \) we have:

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

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 (6), 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 \( x^{\epsilon} \) be a \( c \)-replication of \( x \), \( y^{\epsilon'} \) be a \( c \)-replication of \( y \) and \( z^{(e-c)} \) be a \( (e - c) \)-replication of \( z \). By Replication Invariance, we have \( P(x^{\epsilon}) < P(y^{\epsilon'}) < P(z^{(e-c)}) \). Thus, by Subgroup Consistency, we have that \( P(x^{\epsilon}, z^{(e-c)}) < P(y^{\epsilon'}, z^{(e-c)}) \). Now, we constructed these replications such that \( \hat{v} = m((x^{\epsilon}, z^{(e-c)})) \) and \( \hat{v}^* = m((y^{\epsilon'}, z^{(e-c)})) \). By (6), we obtain \( \hat{v} \succ \hat{v}^* \), the desired contradiction.

Second, we use the previous result to show that \( v \succ v' \). Partition \( \Delta \) into three 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 \( E_{001}^{100} \) with \( v'' \prec v''^* \) 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 \( E_{001}^{100} \) with \( v'' \succ v''^* \) and, hence, such that \( v'' \succ \hat{v} \) or on the edge \( E_{010}^{100} \) and, as \( v^{100} \succ v^{010} \succ \hat{v} \), such that \( v'' \succ \hat{v} \). This implies in both cases that \( v \succ v' \).

---

44We call this indifference curve “implicit” because it is defined in \( \Delta^{\mathbb{R}^+} \) rather than in \( \Delta \).

43The alternative contradiction assumption for which \( \hat{v}_2 < \frac{\hat{v}_1}{\hat{v}_2} \) also leads to an impossibility.

42If \( \hat{v}^* \notin \Delta \), then replace \( \hat{v}^* \) by a nearby \( \hat{v}^* \in E_{001}^{010} \) such that \( \hat{v}_1^* > \hat{v}_1^* \) and \( \hat{v}_2 > \frac{\hat{v}_1^*}{\hat{v}_2^*} \). As \( \hat{v}_1 > \hat{v}_1^* \), we have \( \hat{v}^* \succ \hat{v} \).
7.2 Proof of Lemma 1

We prove in Proposition 5 a slightly more general result, which requires to refine the definition of a demographic equilibrium.

**Definition 1** (Demographic equilibrium up to \(a'\)).

The pair \((x, \mu) \in X \times M\) is a demographic equilibrium up to \(a'\) if we have for all \(a \in \{0, \ldots, a'\}\) that

\[
n_{a+1}(x) = n_a(x) \ast (1 - \mu_a). \tag{7}\]

The pair \((x, \mu)\) is a demographic equilibrium if it is a demographic equilibrium up to \(a^* - 1\), implying that (7) holds for all \(a \in \{0, \ldots, a^* - 1\}\). Observe that a demographic equilibrium does not require that past mortality rates and natality were constant.

**Proposition 5** (Equivalence between ID, GD and ED indices in equilibrium).

If the pair \((x, \mu) \in X \times M\) is a demographic equilibrium, then we have that

\[
P_{\gamma}^{ED}(x, \mu) = P_{\gamma}^{GD}(x, \mu).
\]

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

\[
P_{\gamma}^{ED}(x, \mu) = P_{\gamma}^{GD}(x, \mu) = P_{\gamma}^{ID}(x).
\]

**Proof.** We prove both equalities in turn.

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

\[
P_{\gamma}^{GD}(x, \mu) = \frac{(p(x) + f(x)) \ast HC(x)}{p(x) + f(x) + d_{GD}(x, \mu)} + \gamma \frac{d_{GD}(x, \mu)}{p(x) + f(x) + d_{GD}(x, \mu)}.
\]

Given that \((x, \mu)\) is a demographic equilibrium, we have for all \(a \in \{0, \ldots, a^*\}\) that

\[
n_a(x) = n_0(x) \ast \prod_{l=0}^{a-1} (1 - \mu_l). \tag{8}\]

Using (8), the definition of \(d_{GD}(x, \mu)\) may be rewritten

\[
d_{GD}(x, \mu) = \sum_{a=0}^{\hat{a}-1} n_a(x) \ast \mu_a \ast (\hat{a} - (a + 1)),
\]

\[
= n_0(x) \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a + 1)) \ast \mu_a \ast \prod_{l=0}^{a-1} (1 - \mu_l),
\]

\[
= n_0(x) LGE_{\hat{a}}(\mu).
\]

Similarly, given that \((x, \mu)\) is a demographic equilibrium, we may rewrite \(p(x) + f(x)\)
using (8) as

\[ p(x) + f(x) = \sum_{a=0}^{a^*} n_a(x), \]

\[ = n_0(x) \prod_{l=0}^{a^*-a-1} (1 - \mu), \]

\[ = n_0(x) LE(\mu). \]

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

\[ P^{GD}_\gamma(x, \mu) = \frac{n_0(x) LE(\mu) HC(x)}{n_0(x) LE(\mu) + n_0(x) LGE_\delta(\mu)} + \gamma \frac{n_0(x) LGE_\delta(\mu)}{LGE_\mu(\mu) + LGE_\delta(\mu)}, \]

\[ = \frac{LE(\mu) \ast HC(x)}{LE(\mu) + LGE_\delta(\mu)} + \gamma \frac{LGE_\delta(\mu)}{LE(\mu) + LGE_\delta(\mu)} = P^{ED}_\gamma(x, \mu). \]

**Second**, we show that \( P^{GD}_\gamma(x, \mu) = P^{ID}_\gamma(x) \). As the pair \((x, \mu)\) is a demographic equilibrium 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 \( P^{GD}_\gamma(x, \mu) \) becomes

\[ P^{GD}_\gamma(x, \mu) = \frac{(p(x) + f(x)) \ast HC(x)}{p(x) + f(x) + d^{GD}(x, \mu)} + \gamma \frac{d^{GD}(x, \mu)}{p(x) + f(x) + d(x)} \]

\[ = \frac{(p(x) + f(x)) \ast HC(x)}{p(x) + f(x) + d(x)} + \gamma \frac{d(x)}{p(x) + f(x) + d(x)} = P^{ID}_\gamma(x). \]

\[ \square \]

**Lemma 2** (Equivalence between \( d \) and \( d^{GD} \) in demographic equilibrium).

*If the pair \((x, \mu)\) is a demographic equilibrium 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 is a demographic equilibrium 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,

\[ d(x) = \sum_{a=1}^{\hat{a}-1} (n^* - n_a(x)). \]

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

\[ d(x) = \sum_{a=1}^{\hat{a}-1} \left( \sum_{a'=0}^{a-1} (n_{a'}(x) - n_{a'+1}(x)) \right), \]
and developing the sums, we get
\[
d(x) = (\hat{a} - 1)(n_0(x) - n_1(x)) + (\hat{a} - 2)(n_1(x) - n_2(x)) + \cdots + (\hat{a} - (\hat{a} - 1))(n_{\hat{a}-2}(x) - n_{\hat{a}-1}(x)),
\]
\[
= \sum_{a=0}^{\hat{a}-2}(n_a(x) - n_{a+1}(x))(\hat{a} - (a + 1)),
\]
and given that \(\hat{a} - ((\hat{a} - 1) + 1) = 0\), this is equivalent to
\[
d(x) = \sum_{a=0}^{\hat{a}-1}(n_a(x) - n_{a+1}(x))(\hat{a} - (a + 1)).
\]
Finally, as the pair \((x, \mu)\) is a demographic equilibrium up to \(\hat{a} - 2\) we have for all \(a \in \{0, \ldots, \hat{a} - 2\}\) that \(n_{a+1}(x) = n_a(x) - n_a(x) \ast \mu_a\) and therefore we have
\[
d(x) = \sum_{a=0}^{\hat{a}-1} n_a(x) \ast \mu_a \ast (\hat{a} - (a + 1)) = d^{GD}(x, \mu).
\]

\[\square\]

### 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 \(P((x, x_{-i}), \mu) = P((x, x_{-i}), \mu)\).

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

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

We first prove sufficiency. 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 necessity. Take any pair \((x, \mu) \in O\). We construct another pair \((x'', \mu)\) that is stationary and such that \(P(x'', \mu) = P(x, \mu)\) and \(P^{ED}(x'', \mu) = P^{ED}(x, \mu)\).

Given that \((x'', \mu)\) is stationary, we have by ID Equivalence that \(P(x'', \mu) = P^{ED}(x'', \mu)\). The characteristics of \((x'', \mu)\) then imply that \(P(x, \mu) = P^{ED}(x, \mu)\), the desired result.

We turn to the construction of the stationary pair \((x'', \mu)\). One difficulty is to ensure that the mortality rates \(\mu_a\) are feasible in a demographic equilibrium given the number of alive individuals \(n_a(x'')\), which is \(\mu_a = \frac{c_a}{n_a(x'')}\) for some \(c \in \mathbb{N}\). We first construct a \(n''\)–replication of \(x\) that has sufficiently many alive individuals to meet this constraint. By definition, mortality rates can be expressed as \(\mu_a = \frac{c_a}{n_a(x')}\) where \(c_a, e_a \in \mathbb{N}\). Let \(e = \prod_{j=0}^{a-1} e_a, n'_{a} = e \prod_{j=0}^{a-1}(1 - \frac{c_j}{n_j})\) and \(n' = \sum_{j=0}^{a} n'_j\). Let \(x'\) be a \(n''\)–replication of \(x\). Letting \(n'' = \sum_{j=0}^{a} n_j(x)\) be the number of alive individuals in distribution \(x\), we have that \(x'\) has \(n' \ast n''\) 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 \((x'', \mu)\) form a demographic equilibrium. We construct \( x'' \) with \( n(x'') = n(x') \) such that

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

One can check that \((x'', \mu)\) forms a demographic equilibrium 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 \((x''', \mu)\) is stationary. To do so, place exactly \( n_0(x'') - n_a(x'') \) dead individuals in each age group \( a \). We have \( P(x'''', \mu) = P(x''', \mu) \) by Independence of Dead*.

Together, we have that \( P(x''', \mu) = P(x, \mu) \). Finally, by construction, \( HC(x''') = HC(x) \), which implies that \( PE(x'''', \mu) = PE(x, \mu) \).

7.4 Proof of Proposition 3

We first prove sufficiency. 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

\[
d^{GD}((x', x''), \mu) = \sum_{a=0}^{\hat{a} - 1} n_a(x', x'') \ast \mu_a \ast (\hat{a} - (a + 1))
= \sum_{a=0}^{\hat{a} - 1} (n_a(x') + n_a(x'')) \ast \frac{n_a(x') \ast \mu_a' + n_a(x'') \ast \mu_a''}{n_a(x') + n_a(x'')} \ast (\hat{a} - (a + 1))
= d^{GD}(x', \mu') + d^{GD}(x'', \mu'').
\]

It is then straightforward to verify (4) by replacing \( P \) and \( \eta \) by their expressions.

We now prove necessity. 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 \( GD(x, \mu) = GD(x', \mu) \).

The proof requires to define, for each \( a \in \{0, \ldots, a^*\} \), two counterfactual pairs \((x^*_a, \mu^*_a)\) and \((x^0_a, \mu^0_a)\), which are illustrated in the center and right panels of Figure 14.

The counterfactual pair \((x^*_a, \mu^*_a)\) is stationary. The vector \( \mu^*_a \) is such that mortality rates are zero except for two cases: \( \mu^*_a = \mu_a \) and \( \mu^*_a = 1 \), which is \( \mu^*_a = \ldots \)
(0, . . . , 0, µ₀, 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 \( n_a(x^*_a) = n_a(x) * (1 - µ_a) \). At all ages \( a' \leq a \), there are no dead individuals; for ages \( a' > a \), this number is \( n_a(x) * µ_a \). 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 \( x \) who are AP, i.e. \( \#\{i \leq n(x)|s_i = AP \text{ and } b_i = t - a\} \).

The counterfactual pair \((x'_a, µ'_a)\) is not stationary and all its alive individuals are \( a \)-years old. The vector \( µ'_a = µ_a^* \) which is \( µ'_a = (0, . . . , 0, µ_a, 0, . . . , 1) \). We now turn to the construction of distribution \( x'_a \). At all ages \( a' \neq a \), there are no alive individuals (i.e. \( n_{a'}(x'_a) = 0 \)); and we have \( n_a(x'_a) = n_a(x) \). There are no dead individuals. The number of AP individuals in \( x'_a \) 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 Decomposability, we have that

\[
P(x, µ) = \frac{\sum_{j=0}^{a'} η(x'_j, µ'_j) * P(x^*_j, µ^*_j)}{\sum_{j=0}^{a'} η(x'_j, µ'_j)}.
\]

Expression (9) holds in particular for the stationary distribution \((x^*_a, µ^*_a)\):

\[
P(x^*_a, µ^*_a) = \frac{\sum_{j=0}^{a'} η((x^*_j)^0, (µ^*_j)^0) * P((x^*_j)^0, (µ^*_j)^0)}{\sum_{j=0}^{a'} η((x^*_j)^0, (µ^*_j)^0)}.
\]

where the mortality vector \((µ^*_j)^0 = µ^*_a \) for \( j = a \) and \((µ^*_j)^0 = (0, . . . , 0, 1) \) for \( j \neq a \); and for \( j = a \) we have \(((x^*_j)^0, (µ^*_j)^0) = (x'_a, µ'_a)\).

For all \( j \neq a \) we show that \( P((x^*_j)^0, (µ^*_j)^0) = 0 \). Recall that \( (µ^*_j)^0 = (0, . . . , 0, 1) \) and that \((x^*_j)^0 \) has no \( j \)-years old individuals who are AP. Consider the stationary pair \((x''', µ''')\) such that \( µ''' = (0, . . . , 0, 1) \), distribution \( x''' \) has no AP individual and no dead individual. Provided that \( n(x''') = n_j(x^*_a) * (a^* + 1) \), we have that \( P((x_j)^0, (µ_j)^0) \) appears in the decomposition (9) applied to \((x''', µ''')\). By ID Equivalence, we have that \( P(x'''', µ''') = GD(x''', µ''') = 0 \). Given that \( P \) does not yield negative images, we must have that \( P((x^*_j)^0, (µ^*_j)^0) = 0 \).

As \((x^*_a, µ^*_a)\) is stationary, we have from ID Equivalence that \( P(x^*_a, µ^*_a) = GD(x^*_a, µ^*_a) \). Given that \( P((x^*_j)^0, (µ^*_j)^0) = 0 \) for all \( j \neq a \), and \( \sum_{j=0}^{a'} η((x^*_j)^0, (µ^*_j)^0) = η(x^*_a, µ^*_a) \).

In order to be complete, there remains to show that \( η((x^*_j)^0, (µ^*_j)^0) > 0 \) when \( n_j(x^*_a) > 0 \). If not, one can derive a contradiction with the requirement that \( η(x, µ) = η(x', µ') + η(x'', µ'') \).
(10) may be rewritten as

\[ P((x^*_a)^0_a, (\mu^*_a)^0_a) = \frac{\eta(x^*_a, \mu^*_a) \ast GD(x^*_a, \mu^*_a)}{\eta(x^*_a, \mu^*_a)^0_a}. \]

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

\[ P(x^0_a, \mu^0_a) = \frac{\eta(x^*_a, \mu^*_a) \ast GD(x^*_a, \mu^*_a)}{\eta(x^0_a, \mu^0_a)}. \]

Inserting this last expression in (9), where \(\sum_{j=0}^n \eta(x^0_j, \mu^0_j) = \eta(x, \mu)\), yields

\[ P(x, \mu) = \frac{\sum_{j=0}^n \eta(x^*_j, \mu^*_j) \ast GD(x^*_j, \mu^*_j)}{\eta(x, \mu)}. \] (11)

Equation (11) holds for all pairs. If we have that function \(\eta\) is defined as \(\eta(x, \mu) = f(x) + p(x) + d^{GD}(x, \mu)\), then (11) simplifies to \(P(x, \mu) = GD(x, \mu)\) and the proof is complete. We now show that the function \(\eta\) is indeed expressed as \(\eta(x, \mu) = f(x) + p(x) + d^{GD}(x, \mu)\). Equation (11) holds in particular for any stationary pair \((x', \mu')\). Therefore, by ID Equivalence we have

\[ GD(x', \mu') = \frac{\sum_{j=0}^n \eta((x')_j, (\mu')_j) \ast GD((x')_j, (\mu')_j)}{\eta(x', \mu')}, \] (12)

which only holds if function \(\eta\) has the appropriate expression.

### 7.5 Proof of Proposition 4

The definition of \(P^{ED}_\gamma(x, \mu)\) is

\[ P^{ED}_\gamma(x, \mu) = \frac{LE(\mu) \ast HC(x)}{LE(\mu) + LGE_\gamma(\mu)} + \frac{\gamma \ast LGE_\gamma(\mu)}{LE(\mu) + LGE_\gamma(\mu)}. \]

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

\[ P^{ED}_\gamma(x, \mu) = \frac{n(x) \ast HC(x) + \gamma \ast \frac{LGE_\gamma(\mu) \ast n(x)}{LE(\mu)}}{n(x) + \frac{LGE_\gamma(\mu) \ast n(x)}{LE(\mu)}}, \]

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

\[ \frac{n(x) \ast HC(x) + \gamma \ast \frac{LGE_\gamma(\mu) \ast n(x)}{LE(\mu)}}{n(x) + \frac{LGE_\gamma(\mu) \ast n(x)}{LE(\mu)}} \leq \frac{n(x) \ast HC(x) + \gamma \ast d^{GD}(x, \mu)}{n(x) + d^{GD}(x, \mu)}. \] (13)
When $\gamma \geq 1$, each of the two fractions compared in (13) is monotonically increasing in its factor multiplying $\gamma$.\footnote{For example, $P_{GD}^{\gamma}$ is increasing in $d^{GD}$ as we have by chain derivation that}

Therefore, inequality (13) is equivalent to

\[
\frac{LGE_{a}(\mu) \ast n(x)}{LE(\mu)} \leq d^{GD}(x, \mu).
\]

Replacing $d^{GD}$ and $LGE_{a}$ in (14) by their definitions leads to

\[
\frac{n(x)}{LE(\mu)} \ast \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a + 1)) \ast \mu_{a} \ast \prod_{i=0}^{a-1} (1 - \mu_{i}) \leq \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a + 1)) \ast \mu_{a} \ast \prod_{i=0}^{a-1} (1 - \mu_{i}).
\]

As $\mu^{x}$ is the mortality vector for which $(x, \mu^{x})$ is a demographic equilibrium, we have for all $a \in \{0, \ldots, a^{*}\}$ that $n_{a}(x) = n_{0}(x) \ast \prod_{i=0}^{a-1} (1 - \mu_{i})$. Replacing $n_{a}(x)$ in last inequality leads to

\[
\frac{n(x)}{LE(\mu)} \ast \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a + 1)) \ast \mu_{a} \ast \prod_{i=0}^{a-1} (1 - \mu_{i}) \leq \sum_{a=0}^{\hat{a}-1} (\hat{a} - (a + 1)) \ast \mu_{a} \ast \prod_{i=0}^{a-1} (1 - \mu_{i}).
\]

where $n(x) = \sum_{a=0}^{a^{*}} n_{0}(x) \ast \prod_{i=0}^{a-1} (1 - \mu_{i}) = n_{0}(x) \ast LE(\mu^{x})$, yielding

\[
\frac{\sum_{a=0}^{\hat{a}-1} (\hat{a} - (a + 1)) \ast \prod_{i=0}^{a-1} (1 - \mu_{i}) \ast \mu_{a}}{LE(\mu)} \leq \frac{\sum_{a=0}^{\hat{a}-1} (\hat{a} - (a + 1)) \ast \prod_{i=0}^{a-1} (1 - \mu_{i}) \ast \mu_{a}}{LE(\mu^{x})}
\]

the desired result.

8 Appendix 2: Additional properties

8.1 Proposition 6: GD and ED count the same number of PYPLs

In the absence of migration, many aspects of distribution $x$ are mechanically related to the distribution and mortality vector of the preceding period. We say that a distribution $x'$ in period $t$ is \textit{generated} by the pair $(x, \mu)$ in period $t - 1$ if (i) the number of individuals born in each period before $t$ is the same in both distributions, (ii) all individuals in $x'$ who do not have a counterpart in $x$ are newborns, (iii) individuals that are dead in $x$ have their counterpart also dead in $x'$ and (iv) the number of $a$-year-old individuals in $x'$ is equal to the number of $(a - 1)$-year-old individuals in $x$ multiplied by the survival rate $1 - \mu_{a-1}$. Formally, that is

(i) $n_{a+1}(x') + d_{a+1}(x') = n_{a}(x) + d_{a}(x)$ for all $a \geq 0$,

(ii) $b_{j} = t$ for all $j$ present in $x'$ but not in $x$,

(iii) $(s'_{i}, b'_{i}) = (s_{i}, b_{i})$ for all $i$ present in $x$ such that $s_{i} = D$,

(iv) $n_{a+1}(x') = n_{a}(x) \ast (1 - \mu_{a})$ for all $a \geq 0$.

\footnote{For example, $P_{GD}^{\gamma}$ is increasing in $d^{GD}$ as we have by chain derivation that}

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

where $\gamma \geq HC(x)$ when $\gamma \geq 1$ as there is at least one individual who is non-poor.
Let the set of all periods up to \( t \) be denoted by \( Z_t = \{-\infty, \ldots, t-1, t\} \). When comparing total deprivation in time \( t \), objects to compare could be streams of pairs \((\tau^\tau, \mu^\tau)_{\tau \in Z_t} \in O^{Z_t}\) such that \( \tau^{\tau+1} \) is generated by \((\tau^\tau, \mu^\tau)\) for all \( \tau \in Z_t \). When evaluating \((\tau^\tau, \mu^\tau)_{\tau \in Z_t}\), ID indices only considers the information in current alive population \( \tau^\tau \) as well as past natality \( n_0(\tau^\tau) \) for \( \tau \in \{t-(\hat{a}-1), \ldots, t-1\} \). In turn, GD indices only consider the information in current alive population \( \tau^{\hat{a}} \) together with current mortality \( \mu^\tau \). ED family only considers the information in current deprivation \( HC(\tau^{\hat{a}}) \) together with current mortality \( \mu^\tau \).

Proposition 6 shows that GD and ID indices count the same number of PYPLs even outside demographic equilibrium. Consider a population in a stationary state up to period 0. Over the time-frame \( \{0, \ldots, t^*\} \), this population is hit by a series of mortality and natality shocks. From period \( t^* + 1 \) onwards, natality and mortality return to the values they had before period 0. Mechanically, the young part of the population pyramid may need up to \( \hat{a} - 1 \) periods after period \( t^* \) in order to come back to its previous stationary state. By Proposition 2, GD and ID indices are equal outside the (extended) time-frame \( \{0, \ldots, t^* + \hat{a} - 1\} \). Proposition 6 shows that GD and ID indices compute the same number of PYPLs over the (extended) time frame. Let \( n(x^\tau) = p(x^\tau) = f(x^\tau) \) be the number of alive individuals in \( x^\tau \).

**Proposition 6** (GD and ID indices count the same number of PYPLs).

Let the stream of pairs \((x^\tau, \mu^\tau)_{\tau \in Z} \in O^Z\) be such that \( x^{\tau+1} \) is generated by \((x^\tau, \mu^\tau)\) for all \( \tau \in Z \). Take any \( t^* \geq 0 \). If we have for all \( t \in Z_{\{0, \ldots, t^* + \hat{a} - 1\}} \) that

- \( n_0(x^\tau) = n^* \in \mathbb{N}_{++} \) and \( \mu^\tau = \mu^* \in M \)

then we have

\[
\sum_{t=0}^{t^*+\hat{a}-1} (n(x^t) + d(x^t)) \ast P_{\gamma}^{ID}(x^t) = \sum_{t=0}^{t^*+\hat{a}-1} (n(x^t) + d^{GD}(x^t, \mu^t)) \ast P_{\gamma}^{GD}(x^t, \mu^t),
\]

(15)

and for all \( t \in Z_{\{0, \ldots, t^* + \hat{a} - 1\}} \) we have \( P_{\gamma}^{ID}(x^t) = P_{\gamma}^{GD}(x^t, \mu^t) \).

**Proof.** We prove the two implications in turn.

- **Step 1:** For all \( t \in Z_{\{0, \ldots, t^* + \hat{a} - 1\}} \) we have \( P_{\gamma}^{ID}(x^t) = P_{\gamma}^{GD}(x^t, \mu^t) \).

  Take any \( t \in Z_{\{0, \ldots, t^* + \hat{a} - 1\}} \). By definition, we have \( P_{\gamma}^{ID}(x^t) = P_{\gamma}^{GD}(x^t, \mu^t) \) if \( d(x^t) = d^{GD}(x^t, \mu^t) \). By Proposition 2, we have \( d(x^t) = d^{GD}(x^t, \mu^t) \) if (A) the pair \((x^t, \mu^t)\) is a demographic equilibrium up to \( \hat{a} - 2 \) and (B) for all cohorts \( a \in \{0, \ldots, \hat{a} - 1\} \) we have \( n_a(x^t) + d_a(x^t) = n^* \in \mathbb{N}_{++} \).

  First, we prove (A). The pair \((x^t, \mu^t)\) is a demographic equilibrium up to \( \hat{a} - 2 \) if

  \[
  n_{a+1}(x^t) = n_a(x^t) - n_a(x^t) \ast \mu^t_a \quad \text{for all } a \in \{0, \ldots, \hat{a} - 2\}.
  \]

  As \( t \notin \{0, \ldots, t^* + \hat{a} - 1\} \), we have by the assumption that for all previous periods \( t' \in \{t-(\hat{a}-1), \ldots, t-1\} \)

  47Observe that more general indicators may consider information on past alive population \( x^\tau \) such as indicators of chronic deprivation.
- distribution \( x^{t'}+1 \) is generated by \((x^{t'}, \mu^{t'})\),
- \( n_{0}(x^{t'}) = n^* \) and \( \mu^{t'} = \mu^* \).

As a result, we have for all \( a \in \{0, \ldots, \hat{a} - 2\} \) that

\[
n_{a+1}(x^{t'}) = n^* \prod_{t=0}^{a} (1 - \mu_t^*),
\]

which implies for all \( a \in \{1, \ldots, \hat{a} - 2\} \) that \( n_{a+1}(x^{t'}) = n_{a}(x^{t'}) \ast (1 - \mu_a^*) \).

Finally, last equation also holds for \( a = 0 \) as by \( n_0(x^{t'}) \) we also have \( n_1(x^{t'}) = n^* \ast (1 - \mu_0^*) \), the desired result.

Second, we prove (B). Given that for all \( t \in \{\hat{a} - (\hat{a} - 1), \ldots, t - 1\} \) distribution \( x^{t'}+1 \) is generated by \((x^{t'}, \mu^{t'})\), we have for all \( a \in \{1, \ldots, \hat{a} - 1\} \) that \( n_{a}(x^{t'}) + d_{a}(x^{t'}) = n_{0}(x^{t' - a}) \). As for all \( t' \) \in \{\hat{a} - (\hat{a} - 1), \ldots, t\} \) we have \( n_{0}(x^{t'}) = n^* \), the previous claim implies that \( n_{a}(x^{t'}) + d_{a}(x^{t'}) = n^* \) for all \( a \in \{0, \ldots, \hat{a} - 1\} \).

**Step 2:** Equation (15) holds.

By definition of \( P^{ID}_\gamma(x^{t'}) \) and \( P^{GD\gamma}(x^{t'}, \mu^{t'}) \), equation (15) is

\[
\sum_{t=0}^{t^*+\hat{a}-1} (n(x^{t'}) + d(x^{t'})) \ast \left( \frac{n(x^{t'}) \ast HC(x^{t'})}{n(x^{t'}) + d(x^{t'})} + \gamma \frac{d(x^{t'})}{n(x^{t'}) + d(x^{t'})} \right) = \\
\sum_{t=0}^{t^*+\hat{a}-1} (n(x^{t'}) + d^{GD}(x^{t'}, \mu^{t'})) \ast \left( \frac{n(x^{t'}) \ast HC(x^{t'})}{n(x^{t'}) + d^{GD}(x^{t'}, \mu^{t'})} + \gamma \frac{d^{GD}(x^{t'}, \mu^{t'})}{n(x^{t'}) + d^{GD}(x^{t'}, \mu^{t'})} \right),
\]

which is equivalent to

\[
\sum_{t=0}^{t^*+\hat{a}-1} d(x^{t'}) = \sum_{t=0}^{t^*+\hat{a}-1} d^{GD}(x^{t'}, \mu^{t'}). \tag{16}
\]

In order to prove (16), we develop the sums ID and GD.

**First,** we develop GD. Using the short notation \( \Delta_a = \hat{a} - (a + 1) \), the definition of \( d^{GD} \) is

\[
d^{GD}(x^{t'}, \mu^{t'}) = \sum_{a=0}^{\hat{a}-1} n_{a}(x^{t'}) \ast \mu_a^* \ast \Delta_a,
\]

and as \( \Delta_{\hat{a}-1} = 0 \), we have

\[
GD = \sum_{t=0}^{t^*+\hat{a}-1} \sum_{a=0}^{\hat{a}-2} n_{a}(x^{t'}) \ast \mu_a^* \ast \Delta_a. \tag{17}
\]

Equation (17) shows that GD counts the number of person-years prematurely lost (PYPL) due to deaths occurring in the time-frame \( T = \{0, \ldots, t^* + \hat{a} - 1\} \).

All these PYPLs are lost for periods in \( \{1, \ldots, t^* + 2\hat{a} - 2\} \). Equation (18) divides these PYPLs between the PYPLs that are lost for periods in the time-frame \( T \) i.e. for periods 1 to \( t^* + \hat{a} - 1 \) and those lost for periods following
the time-frame $T$ — i.e. for periods $t^* + \hat{a}$ to $t^* + 2\hat{a} - 2$.

$$GD' = \sum_{t=0}^{t^* + \hat{a} - 2} \sum_{a=0}^{\hat{a} - 2} \sum_{\tau=t+1}^{t^* + \hat{a} - 1} n_a(x^t) \cdot \mu^t_a + \sum_{t=t^* + 1}^{t^* + \hat{a} - 2} \sum_{a=0}^{\hat{a} - 2} \sum_{\tau=t^* + 1}^{t^* + \hat{a} - 1} n_a(x^t) \cdot \mu^t_a.$$  \hspace{2cm} \text{GD' - inside } T

$$GD' = \sum_{t=0}^{t^* + \hat{a} - 2} \sum_{a=0}^{\hat{a} - 2} \sum_{\tau=t+1}^{t^* + \hat{a} - 1} n_a(x^t) \cdot \mu^t_a + \sum_{t=t^* + 1}^{t^* + \hat{a} - 2} \sum_{a=0}^{\hat{a} - 2} \sum_{\tau=t^* + 1}^{t^* + \hat{a} - 1} n_a(x^t) \cdot \mu^t_a.$$  \hspace{2cm} \text{GD' - outside } T

We show that $GD' = GD$. As term $n_a(x^t) \cdot \mu^t_a$ is independent on $\tau$ we have

$$GD' = \sum_{t=0}^{t^* + \hat{a} - 2} \sum_{a=0}^{\hat{a} - 2} n_a(x^t) \cdot \mu^t_a \cdot \sum_{\tau=t+1}^{t^* + \hat{a} - 1} 1 + \sum_{t=t^* + 1}^{t^* + \hat{a} - 2} \sum_{a=0}^{\hat{a} - 2} n_a(x^t) \cdot \mu^t_a \sum_{\tau=t^* + 1}^{t^* + \hat{a} - 1} 1.$$  \hspace{2cm} \text{(18)}

We consider the expression of $GD'$ in turn for the set of periods $\{0, \ldots, t^*\}$, then for the set of periods $\{t^* + 1, \ldots, t^* + \hat{a} - 2\}$ and finally for period $t^* + \hat{a} - 1$.

For each of these three sets of periods, we show that the expression of $GD'$ corresponds to the expression of $GD$.

- Periods $t \in \{0, \ldots, t^*\}$.

We have for all $a \in \{0, \ldots, \hat{a} - 2\}$ that

$$\min\{t + (\hat{a} - 1) - a, t^* + \hat{a} - 1\} = t + \hat{a} - 1 - a,$$

which implies

$$\sum_{\tau=t+1}^{t^* + \hat{a} - 1} 1 = \hat{a} - (a + 1) = \Delta_a,$$

and therefore

$$\sum_{t=0}^{t^*} \sum_{a=0}^{\hat{a} - 2} n_a(x^t) \cdot \mu^t_a \cdot \sum_{\tau=t+1}^{t^* + \hat{a} - 1} 1 = \sum_{t=0}^{t^*} \sum_{a=0}^{\hat{a} - 2} n_a(x^t) \cdot \mu^t_a \cdot \Delta_a.$$  \hspace{2cm} (19)

- Periods $t \in \{t^* + 1, \ldots, t^* + \hat{a} - 2\}$.

We have for all $a \in \{0, \ldots, \hat{a} - 2\}$ that

$$\sum_{t=t^* + 1}^{t^* + \hat{a} - 2} \sum_{a=0}^{\hat{a} - 2} n_a(x^t) \cdot \mu^t_a \cdot \sum_{\tau=t+1}^{t^* + \hat{a} - 1} 1 = \sum_{t=t^* + 1}^{t^* + \hat{a} - 2} \sum_{a=0}^{\hat{a} - 2} n_a(x^t) \cdot \mu^t_a \cdot \sum_{\tau=t^* + 1}^{t^* + \hat{a} - 1} 1.$$  \hspace{2cm} (20)

We consider the terms $t^* + \hat{a} - 2$ and $t^* + \hat{a} - 1$ separately.

- Period $t = t^* + \hat{a} - 1$.
We have for all $a \in \{0, \ldots, \hat{a} - 2\}$ that

$$t^* + 2(\hat{a} - 1) - a \sum_{\tau = t^* + \hat{a}} 1 = \hat{a} - (a + 1) = \Delta_a.$$ 

and therefore

$$t^* + \hat{a} - 1 \sum_{t = t^* + \hat{a} - 1} (\hat{a} - 2) \sum_{a = 0} n_a(x^t) \mu'_a \sum_{\tau = t^* + \hat{a}} 1 = \sum_{t = t^* + \hat{a} - 1} (\hat{a} - 2) \sum_{a = 0} n_a(x^t) \mu'_a \Delta_a.$$ 

Second, we develop ID. The sum ID counts the number of person-years that are prematurely lost for periods in the time-frame $T = \{0, \ldots, t^* + \hat{a} - 1\}$. As distribution $x^{t+1}$ is generated by $(x^t, \mu^t)$ for all $t \in \mathbb{Z}$, these PYPLs are lost due to deaths occurring in the set of periods $t \in \{-\hat{a} + 1, \ldots, t^* + \hat{a} - 2\}$. We express ID by counting all these PYPLs in the following way:

- run all periods $t$ at which the occurrence of a death potentially generates a PYPL for a period in $T$,
- for each such period $t$, run all age-cohorts $a$ whose death generates a PYPL for a period in $T$,
- count all periods in $T$ that are prematurely lost due to a death occurring at age $a$ in period $t$.

Using this way of counting, the sum ID is

$$ID = \sum_{t = -(\hat{a} - 1)}^{t^* + \hat{a} - 2} \sum_{a = 0}^{(\hat{a} - 2)} \sum_{\tau = \max\{0, t + 1\}}^{t + (\hat{a} - 1) - a, t^* + \hat{a} - 1} n_a(x^t) \mu'_a. \quad (19)$$

We illustrate the way in which (19) counts all the relevant PYPLs in Figure 15.

**Figure 15:** Order in which blocks of PYPLs are counted in (19). Block 1 corresponds to PYPLs due to newborns dying in period $-(\hat{a} - 1)$. Block 2 and 3 are due to newborns dying in period $-(\hat{a} - 2)$. Block 4 is due to 1-year-old dying in period $-(\hat{a} - 2)$. Block 5, 6 and 7 are due to newborns dying in period $-(\hat{a} - 3)$. Block 8 and 9 are due to 1-year-old dying in period $-(\hat{a} - 3)$. Block 10 is due to 2-year-old dying in period $-(\hat{a} - 3)$. For simplicity, this order is illustrated for pairs $(x^t, \mu^t)$ that are demographic equilibria.
Equation (20) divides these PYPLs between the PYPLs generated by deaths occurring before the time-frame \( T = \{0, \ldots, t^* + \hat{a} - 1\} \), and the PYPLs generated by deaths occurring during the time-frame \( T \). This division yields the sum \( ID' \)

\[
ID' = \sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{t+(\hat{a}-1)-a} \sum_{\tau=0}^{t+(\hat{a}-1)-a} n_a(x^\tau) \ast \mu_a^t + \sum_{t=0}^{t+(\hat{a}-1)-a} \sum_{a=0}^{t+(\hat{a}-1)-a} \sum_{\tau=t+1}^{t+(\hat{a}-1)-a} n_a(x^\tau) \ast \mu_a^t.
\]

We show that \( ID' = ID \). We consider the expression of \( ID' \) in turn for the set of periods \( \{-(\hat{a}-1), \ldots, -1\} \) and then for the set of periods \( \{0, \ldots, t^* + \hat{a} - 2\} \). For each of these two sets of periods, we show that the expression of \( ID' \) corresponds to the expression of \( ID \).

- Periods \( t \in \{-\hat{a}+1, \ldots, -1\} \).

For this case we have \( \max\{0, t + 1\} = 0. \) Furthermore, for all \( a \geq 0 \) we have \( t - a \leq 0 \leq t^* \) and thus

\[
\min\{t + (\hat{a} - 1) - a, t^* + \hat{a} - 1\} = t + (\hat{a} - 1) - a.
\]

Thus, we can rewrite the sum \( ID' \) on these periods as

\[
\sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{t+(\hat{a}-1)-a} \sum_{\tau=0}^{t+(\hat{a}-1)-a} n_a(x^\tau) \ast \mu_a^t + \sum_{t=0}^{t+(\hat{a}-1)-a} n_a(x^\tau) \ast \mu_a^t. \]

Finally, for all \( a > t + (\hat{a} - 1) \) we have \( t + (\hat{a} - 1) - a < 0 \), implying that

\[
\sum_{\tau=0}^{t+(\hat{a}-1)-a} n_a(x^\tau) \ast \mu_a^t = 0.
\]

and therefore the sum \( ID' \) is further rewritten as

\[
\sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{t+(\hat{a}-1)-a} \sum_{\tau=\max\{0, t+1\}}^{t+(\hat{a}-1)-a} n_a(x^\tau) \ast \mu_a^t.
\]

- Periods \( t \in \{0, \ldots, t^* + \hat{a} - 2\} \).

For this periods we have \( \max\{0, t + 1\} = t + 1. \) Hence, we can rewrite the sum \( ID' \) on these periods as

\[
\sum_{t=0}^{t+(\hat{a}-2)+2} \sum_{a=0}^{t+(\hat{a}-1)-a} \sum_{\tau=t+1}^{t+(\hat{a}-1)-a} n_a(x^\tau) \ast \mu_a^t = \sum_{t=0}^{t+(\hat{a}-2)+2} \sum_{a=0}^{t+(\hat{a}-1)-a} \sum_{\tau=\max\{0, t+1\}}^{t+(\hat{a}-1)-a} n_a(x^\tau) \ast \mu_a^t.
\]

Third, we show that \( GD' = ID' \). Given that \( GD' inside T \) and \( ID' inside T \) are trivially equal to each other, we only need to show that \( GD' outside T = \)

50
From last expression, the relevant $t$ and $a$ are such $t + (\hat{a} - 1) - a \geq t^* + \hat{a}$. \footnote{If $t + (\hat{a} - 1) - a < t^* + \hat{a}$, then we have $\sum_{\tau=t^*+\hat{a}}^{t+(\hat{a}-1)-a-1} 1 = 0$.}
or yet $t \geq t^* + 1 + a$, hence

$$
\sum_{t=t^*+1}^{t^*+\hat{a}-1} \sum_{a=0}^{(\hat{a}-2)} n^*_a \cdot \mu^*_a \sum_{\tau=t^*+\hat{a}}^{t^*+\hat{a}-1} 1
\leq \sum_{a=0}^{(\hat{a}-2)} n^*_a \cdot \mu^*_a \cdot (t - t^* - a),
$$

$$
= \sum_{a=0}^{(\hat{a}-2)} \sum_{t=t^*+1+a}^{t^*+\hat{a}+1+a} n^*_a \cdot \mu^*_a \cdot \left( \sum_{t=t^*+1+a}^{t^*+\hat{a}+1+a} t - (t^* + a) \sum_{t=t^*+1+a}^{t^*+\hat{a}+1+a} 1 \right),
$$

$$
= \sum_{a=0}^{(\hat{a}-2)} n^*_a \cdot \mu^*_a \cdot \left( (t^* + a) \cdot (\hat{a} - 1 - a) + \sum_{t=1}^{\hat{a} - 1 - a} 1 \right) - (t^* + a) \cdot (\hat{a} - 1 - a),
$$

$$
= \sum_{a=0}^{(\hat{a}-2)} n^*_a \cdot \mu^*_a \cdot \frac{(\hat{a} - 1 - a)(\hat{a} - a)}{2}.
$$

Consider then ID’ outside $T$. From Part 1 again, we may rewrite ID’ outside $T$ as

$$
\sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{(\hat{a}-2)} n_a(x^t) \cdot \mu^t_a = \sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{(\hat{a}-2)} n^*_a \cdot \mu^*_a \sum_{\tau=0}^{t+(\hat{a}-1)-a} 1.
$$

From last expression, the relevant $t$ and $a$ are such $t + (\hat{a} - 1) - a \geq 0$, or yet $t \geq a - (\hat{a} - 1)$. Therefore, we rewrite ID’ outside $T$ as

$$
\sum_{t=-(\hat{a}-1)}^{-1} \sum_{a=0}^{(\hat{a}-2)} n^*_a \cdot \mu^*_a \sum_{\tau=0}^{t+(\hat{a}-1)-a} 1
\leq \sum_{a=0}^{(\hat{a}-2)} \sum_{t=a-(\hat{a}-1)}^{-1} n^*_a \cdot \mu^*_a \cdot (t + \hat{a} - a)
$$

$$
= \sum_{a=0}^{(\hat{a}-2)} n^*_a \cdot \mu^*_a \cdot \left( \sum_{t=a-(\hat{a}-1)}^{-1} t + (\hat{a} - a) \sum_{t=a-(\hat{a}-1)}^{-1} 1 \right),
$$

$$
= \sum_{a=0}^{(\hat{a}-2)} n^*_a \cdot \mu^*_a \cdot \left( - \sum_{t'=1}^{\hat{a} - 1 - a} t' + (\hat{a} - a) \cdot (\hat{a} - 1 - a) \right),
$$

$$
= \sum_{a=0}^{(\hat{a}-2)} n^*_a \cdot \mu^*_a \cdot \left( - \frac{(\hat{a} - 1 - a)(\hat{a} - a)}{2} + (\hat{a} - a) \cdot (\hat{a} - 1 - a) \right),
$$

$$
= \sum_{a=0}^{(\hat{a}-2)} n^*_a \cdot \mu^*_a \cdot \frac{(\hat{a} - 1 - a)(\hat{a} - a)}{2},
$$

which shows that (21) holds.

\[ \square \]

### 8.2 LP is an index of lifespan deprivation

Formally, we say that the mortality vector $\mu'$ is obtained from $\mu$ by an increment in one year of life if there are two distributions $x'$ and $x$ such that both $(x', \mu')$

49If $t + (\hat{a} - 1) - a < 0$, then we have $\sum_{\tau=0}^{t+(\hat{a}-1)-a} 1 = 0$. 

52
and \((x, \mu)\) are demographic equilibria and for some \(a_1 \in \{1, \ldots, a^*\}\) and \(\delta \in \mathbb{N}_{++}\) we have

\[ n_{a_1}(x') = n_{a_1}(x) + \delta \]

and \(n_a(x') = n_a(x)\) for all \(a \in \{0, \ldots, a^*\} \setminus \{a_1\}\).\(^{50}\) We say that \(\mu'\) is obtained from \(\mu\) by an increment in one year of life in \textit{young age} if \(a_1 \leq \hat{a} - 1\), and in \textit{old age} otherwise.

A mortality statistic \(m : M \to \mathbb{R}\) satisfies \textit{Current Mortality Focus} if its value remains unaffected by an increment in one year of life in old age.

\textbf{Mortality axiom 1 (Current Mortality Focus).} For all \(\mu, \mu' \in M\), if the mortality vector \(\mu'\) is obtained from \(\mu\) by an increment in one year of life in \textit{old age}, then \(m(\mu) = m(\mu')\).

A mortality statistic satisfies \textit{Current Mortality Monotonicity below \(\hat{a}\)} if its value decreases after an increment in one year of life in young age.

\textbf{Mortality axiom 2 (Current Mortality Monotonicity below \(\hat{a}\)).} For all \(\mu, \mu' \in M\), if the mortality vector \(\mu'\) is obtained from \(\mu\) by an increment in one year of life in \textit{young age}, then \(m(\mu) > m(\mu')\).

\textbf{Proposition 7 \((LP_{\hat{a}}\) is an indicator of lifespan deprivation).} \(LP_{\hat{a}}\) satisfies \textit{Current Mortality Focus} and \textit{Current Mortality Monotonicity below \(\hat{a}\)}.

\textit{Proof.} It is trivial to see from its mathematical expression that \(LP_{\hat{a}}\) satisfies \textit{Current Mortality Focus}. Therefore, we only prove that \(LP_{\hat{a}}\) satisfies \textit{Current Mortality Monotonicity below \(\hat{a}\)}.

We must show that if mortality table \(\mu'\) is obtained from \(\mu\) by an increment in one year of life in young age, then we have \(LP_{\hat{a}}(\mu) > LP_{\hat{a}}(\mu')\).

By assumption, there exist two distributions \(x'\) and \(x\) such that both \((x', \mu')\) and \((x, \mu)\) are demographic equilibria, and thus we have for all \(a \in \{0, \ldots, a^* - 1\}\) that

\[ n_{a+1}(x) = n_a(x) - n_a(x) \ast \mu_a, \quad (23) \]
\[ n_{a+1}(x') = n_a(x') - n_a(x') \ast \mu'_a. \quad (24) \]

As \(\mu'\) is obtained from \(\mu\) by an increment in one year of life in young age, we also have for some \(a_1\) with \(1 \leq a_1 \leq \hat{a} - 1\) that

\[ n_{a_1}(x') = n_{a_1}(x) + \delta, \quad (25) \]
\[ n_a(x') = n_a(x) \quad \text{for all } a \in \{0, \ldots, a^*\} \setminus \{a_1\}. \quad (26) \]

Equation (26) together with (23) and (24) imply that

\[ \mu_a = \mu'_a \quad \text{for all } a \in \{0, \ldots, a^*\} \setminus \{a_1, a_2\}. \quad (27) \]

\(^{50}\)This definition implies that \(\mu'\) is any mortality vector constructed from \(\mu\) such that \(\mu'_a = \mu_a\) for all \(a \notin \{a_1 - 1, a_1\}\) and \(\mu'_{a_1 - 1}\) and \(\mu'_{a_1}\) are rational numbers such that \(\mu'_{a_1 - 1} < \mu_{a_1 - 1}\) and \((1 - \mu'_{a_1 - 1})(1 - \mu'_{a_1}) = (1 - \mu_{a_1 - 1})(1 - \mu_{a_1})\).
Equation (25) together with (23) and (24) imply that

\[ n_{a_1-1}(x) - n_{a_1-1}(x) * \mu_{a_1-1} + \delta = n_{a_1-1}(x') - n_{a_1-1}(x') * \mu'_{a_1-1} \]

and given that \( n_{a_1-1}(x') = n_{a_1-1}(x) \) by (26), we get

\[ n_{a_1-1}(x) * \mu_{a_1-1} = n_{a_1-1}(x') * \mu'_{a_1-1} + \delta \]  \( (28) \)

Similarly, as \( n_{a_1+1}(x') = n_{a_1+1}(x) \) by (26), equations (23) and (24) imply that

\[ n_{a_1}(x) - n_{a_1}(x) * \mu_{a_1} = n_{a_1}(x') - n_{a_1}(x') * \mu'_{a_1} \]

and by (25), we get

\[ n_{a_1}(x) * \mu_{a_1} = n_{a_1}(x') * \mu'_{a_1} - \delta \]  \( (29) \)

By the definition of \( LP_a \) and \( LGE_a \), we have \( LP_a(\mu) > LP_a(\mu') \) if and only if \( LGE_a(\mu) > LGE_a(\mu') \). Take any \( \nu \in M \), by definition we have

\[ LGE_a(\nu) = \sum_{a=0}^{a-1}(\hat{a} - (a + 1)) * \nu_a * \prod_{l=0}^{a-1}(1 - \nu_l). \]

Given that \( \prod_{l=0}^{a-1}(1 - \mu_l) = \frac{n_a(x)}{n_0(x)} \) and \( \prod_{l=0}^{a-1}(1 - \mu'_{l}) = \frac{n_a(x')}{n_0(x')} \) where \( n_0(x) = n_0(x') \) since \( a_1 \geq 1 \), equations (26) and (27) imply for all \( a \in \{0, \ldots, a^*\} \setminus \{a_1-1, a_1\} \) that

\[ (\hat{a} - (a + 1)) * \mu_a * \prod_{l=0}^{a-1}(1 - \mu_l) = (\hat{a} - (a + 1)) * \mu'_{a} * \prod_{l=0}^{a-1}(1 - \mu'_{l}) \]

and therefore we have \( LGE_a(\mu) > LGE_a(\mu') \) if and only if

\[ \sum_{a=a_1-1}^{a_1} (\hat{a} - (a + 1)) * \mu_a * \prod_{l=0}^{a-1}(1 - \mu_l) > \sum_{a=a_1-1}^{a_1} (\hat{a} - (a + 1)) * \mu'_{a} * \prod_{l=0}^{a-1}(1 - \mu'_{l}), \]

which is rewritten as

\[ \sum_{a=a_1-1}^{a_1} (\hat{a} - (a + 1)) * \mu_a * \frac{n_a(x)}{n_0(x)} > \sum_{a=a_1-1}^{a_1} (\hat{a} - (a + 1)) * \mu'_{a} * \frac{n_a(x')}{n_0(x')}, \]

and, when multiplying by \( n_0(x) = n_0(x') \), the sums develop as

\[ (\hat{a} - a_1) * \mu_{a_1-1} * n_{a_1-1}(x) + (\hat{a} - a_1 - 1) * \mu_{a_1} * n_{a_1}(x) > (\hat{a} - a_1) * \mu'_{a_1-1} * n_{a_1-1}(x') + (\hat{a} - a_1 - 1) * \mu'_{a_1} * n_{a_1}(x') \]

Using (28) and (29), we obtain

\[ (\hat{a} - a_1) * (n_{a_1-1}(x') * \mu'_{a_1-1} + \delta) + (\hat{a} - a_1 - 1) * (n_{a_1}(x') * \mu'_{a_1} - \delta) > (\hat{a} - a_1) * \mu'_{a_1-1} * n_{a_1-1}(x') + (\hat{a} - a_1 - 1) * \mu'_{a_1} * n_{a_1}(x') \]

and simplifying we get \( (\hat{a} - a_1) - (\hat{a} - a_1 - 1) > 0 \), the desired result. \( \square \)
Proposition 7 shows that $LP_{\tilde{t}}$ satisfies both Current Mortality Focus and Current Mortality Monotonicity below $\check{a}$. It turns out that, for a fixed distribution $x$, neither GD nor ED indices satisfy both.\footnote{It is straightforward to extend the domain of application of these two mortality axioms from mortality statistics to total deprivation indices.} ED indices satisfy Current Mortality Monotonicity below $\check{a}$ if $\gamma \geq 1$, but they violate Current Mortality Focus because life expectancy is affected by an increment in one year of life in old age.\footnote{ED indices satisfy Current Mortality Monotonicity below $\check{a}$ if $\gamma \geq 1$ because (A) when $\mu'$ is constructed from $\mu$ by an increase in one year of life in young age, we have $LE(\mu) + LGE_a(\mu) = LE(\mu') + LGE_a(\mu')$ and (B) by definition $HC < 1 \leq \gamma$.} GD indices satisfy Current Mortality Focus but violate Current Mortality Monotonicity below $\check{a}$.

The counter-example in the proof for Lemma 3 conveys the intuition. The mortality vector $\mu'$ is obtained from $\mu$ by an increment in one year of life in age one. In $\mu'$, the mortality rate of newborns is decreased and that of 1-year-old is increased, and we have $LE(\mu') > LE(\mu)$. Yet, the pair $(x, \mu)$ is not in demographic equilibrium because there are only 2 newborns and 8 1-year-old individuals in the population pyramid. So the improvement in newborns’ mortality rate only benefit one individual while the worsening in 1-year-old’s mortality rate hits 2 additional individuals. In other words, if the new mortality vector is good news from the perspective of individuals confronted to $\mu'$ throughout their lifetime, the new mortality vector increases the number of deaths in the current population pyramid and increases the number of PYPLs.

Lemma 3. GD indices violate Current Mortality Monotonicity below $\check{a}$.

Proof. We provide a pair $(x, \mu)$ and a mortality vector $\mu'$ obtained from $\mu$ by an increment in one year of life in young age such that $P^{{\gamma}_D}(x, \mu) < P^{{\gamma}_D}(x, \mu')$. Take $x$ such that $n_0(x) = 2$ and $n_1(x) = 8$. Take $\mu$ such that $\mu_0 = \mu_1 = \frac{1}{2}$ and $\mu_a = 0$ for all $a \in \{2, \ldots, \check{a} - 1\}$. Take $\mu'$ such that $\mu'_0 = 0$, $\mu'_1 = \frac{1}{2}$ and $\mu'_a = 0$ for all $a \in \{2, \ldots, \check{a} - 1\}$. Assume $\check{a} = 50$. By construction, we have $d^{GD}(x, \mu) = 241 < 288 = d^{GD}(x, \mu')$, which implies that $P^{{\gamma}_D}(x, \mu) < P^{{\gamma}_D}(x, \mu')$. \hfill $\Box$

Lemma 4. For any two pairs $(x, \mu), (x', \mu') \in O$, if we have that $HC(x) > HC(x')$ then for all $\gamma$ and $\gamma'$ such that $1 \leq \gamma \leq \gamma'$ we have

$$P^{{\gamma}_D}_{\gamma}(x) < P^{{\gamma}_D}_{\gamma}(x') \quad \Rightarrow \quad P^{{\gamma}'_D}_{\gamma'}(x) < P^{{\gamma}'_D}_{\gamma'}(x')$$

(30)

Proof. By chain derivation, we have for all pair $(x, \mu) \in O$ that

$$\frac{\partial P^{{\gamma}_D}_{\gamma}(x, \mu)}{\partial \gamma} = \frac{LGE_a(\mu)}{LE(\mu) + LGE_a(\mu)}$$

which does not depend on $\gamma$. Given that by assumption we have $P^{{\gamma}_D}_{\gamma}(x, \mu) < P^{{\gamma}_D}_{\gamma}(x', \mu')$, it is sufficient to show that $\frac{\partial P^{{\gamma}_D}_{\gamma}(x, \mu)}{\partial \gamma} < \frac{\partial P^{{\gamma}_D}_{\gamma}(x', \mu')}{\partial \gamma}$. Letting $X = \frac{\partial P^{{\gamma}_D}_{\gamma}(x, \mu)}{\partial \gamma}$ and $Y = \frac{\partial P^{{\gamma}_D}_{\gamma}(x', \mu')}{\partial \gamma}$, the precondition $P^{{\gamma}_D}_{\gamma}(x, \mu) < P^{{\gamma}_D}_{\gamma}(x', \mu')$ may be rewritten

$$HC(x) \ast (1 - X) + \gamma X < HC(x') \ast (1 - Y) + \gamma Y.$$
As $HC(x) > HC(x')$ we have that

$$HC(x') \ast (1 - X) + \gamma X < HC(x) \ast (1 - X) + \gamma X.$$ 

Last two inequalities successively imply

$$HC(x') \ast (1 - X) + \gamma X < HC(x') \ast (1 - Y) + \gamma Y,$$

$$(\gamma - HC(x')) \ast X < (\gamma - HC(x')) \ast Y,$$

which shows that $X < Y$ given that $\gamma > HC(x')$ as $\gamma \geq 1$ and $0 \leq HC(x') < 1$. $\square$