DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 54, ARTICLE 14, PAGES 441–470
PUBLISHED 6 MARCH 2026
http://www.demographic-research.org/Volumes/Vol54/14/
DOI: 10.4054/DemRes.2026.54.14

Formal Relationship

The groupwise decomposition: Estimating group-specific contributions to differences between demographic measures

Eugenio Paglino

Contents

Demographic Research: Volume 54, Article 14
Formal Relationship

The groupwise decomposition: Estimating group-specific contributions to differences between demographic measures

Eugenio Paglino[1]

Abstract

BACKGROUND
Several general decomposition algorithms have been developed to decompose the difference between the values of demographic measures in two populations. However, fewer techniques are available to researchers wishing to understand how different subpopulations are contributing to the difference in the value of a demographic measure between the total population and a reference subpopulation.

METHODS
This paper proposes a general decomposition to answer these 'groupwise' questions. As for other general decomposition techniques, the method assumes that the aggregate measure is computed from matrices of discrete demographic data but is agnostic to the specific function used to compute the measure of interest. After proving some of the decomposition's properties, its relationships with other methods are discussed, and extensions are considered.

RESULTS
The proposed method is used to investigate the contribution of United States counties by metropolitan/nonmetropolitan code to the life expectancy gap between large metropolitan areas and the US as a whole. Results from this new approach are compared to those obtained by adapting other general decomposition algorithms, and their computational efficiency is evaluated.

CONTRIBUTION
This paper provides a new decomposition method to investigate how different subpopulations contribute to the difference between a demographic measure for the total population and for a reference population. This setup is common in demography, and the decomposition method developed in this paper adds one more tool to improve our understanding of how different groups contribute to the demographic profile of a population.

---

[1] Helsinki Institute for Demography and Population Health, University of Helsinki, Finland, and Max Planck – University of Helsinki Center for Social Inequalities in Population Health. Email: eugenio.paglino@helsinki.fi.

1. Introduction

The decomposition of demographic measures is a fundamental instrument in a demographer's toolkit. Through the decomposition of demographic measures, researchers can gain crucial insights into how different mechanisms produce demographic change. Kitagawa's classic method (Kitagawa 1955) decomposes the difference between the crude mortality rates of two populations into differences in age-specific mortality rates and differences in the age composition. Das Gupta generalizes this approach to more than two dimensions (Das Gupta 1978), with the restriction that the aggregate demographic quantity had to be a linear function of the covariates of interest. In the 1980s, the focus shifted to nonlinear functions, and methods to decompose differences in life expectancy into age-specific components were independently developed by at least four scholars (Andreev 1982; Arriaga 1984; Pollard 1982; Pressat 1985), although similar methods had been developed earlier (e.g., Chandrasekaran 1949, see Ponnapalli 2005 for an historical account). The next development involved designing general decomposition methods that could decompose the difference between two values of an arbitrary demographic index, with the requirement that it could be derived from a matrix of demographic quantities into cell-specific components. Three such methods have been developed so far, two based on models for continuous change (Caswell 1989; Horiuchi, Wilmoth, and Pletcher 2008) and one based on a model for discrete, stepwise change (Andreev, Shkolnikov, and Begun 2002). Additional methods exist to solve decomposition problems in which changes in a demographic metric over time are to be attributed to (1) changes in group-specific mortality rates, and (2) changes in the share of the population represented by each group (Shkolnikov, Andreev, and Begun 2003; Torres, Canudas-Romo, and Oeppen 2019; Vaupel and Canudas-Romo 2002; Canudas Romo 2003; Riffe, Tursun-Zade, and Trias-Llimós 2024). While these techniques are most often used to decompose differences in life expectancy (Enroth et al. 2022; Luy et al. 2019; Shkolnikov et al. 2006; Su et al. 2024a,b), they can be applied to more complex indicators such as lifespan inequality (Edwards and Tuljapurkar 2005), variances of life expectancy within and between groups, and inequality measures defined on length-of-life distribution (Shkolnikov, Andreev, and Begun 2003; Su et al. 2024a; Timonin et al. 2016). With one exception (Riffe, Tursun-Zade, and Trias-Llimós 2024), these methods are either based on stepwise replacement (Andreev, Shkolnikov, and Begun 2002) or on formulas for decomposing continuous change in a demographic metric over time into direct and compositional change (Canudas Romo 2003; Preston, Himes, and Eggers 1989; Vaupel and Canudas Romo 2003; Vaupel and Canudas-Romo 2002).

To the best of my knowledge, none of these methods, nevertheless, is immediately suitable for tackling 'groupwise' decomposition problems. Consider a demographic measure for the total population $f(T)$, such as US life expectancy, and the same demographic measure for a reference population $f(R)$, such as life expectancy for the US-born population. In a groupwise problem we wish to understand (1) how much individuals who are not in the reference population but are part of the total population, such as foreign-born population, contribute to the difference $f(T) - f(R)$, and (2) how their total contribution can be split into group-specific components, such as by race and Hispanic origin, educational attainment, or sex. Although not previously noted in the literature, in the empirical part of the paper I demonstrate that the general decomposition techniques based on a continuous change model (Caswell 1989; Horiuchi, Wilmoth, and Pletcher 2008) can be adapted to solve this decomposition problem, with appropriate sets of inputs, but they quickly become computationally inefficient even when few groups are involved. This paper thus develops a new decomposition technique which overcomes these limitations, discusses one practical application, and shows how the technique can be further extended.

2. Relationship

Let T be the total population, R a reference subpopulation, and $G_1, G_2, \ldots, G_I$ a partition of the remaining population into I mutually exclusive subgroups such that

$$D_x^R + \sum_{i=1}^{I} D_x^{G_i} = D_x^T$$

$$E_x^R + \sum_{i=1}^{I} E_x^{G_i} = E_x^T,$$

where D_x^T, D_x^R, and $D_x^{G_i}$ are death counts for age x for the total population T, the reference subpopulation R, and subpopulation $G_i (i \in 1,\ 2,\ \ldots,\ I)$, respectively. E_x^T, E_x^R, and $E_x^{G_i}$ are the corresponding exposures in person-years. Then, we wish to identify an expression for the contribution C^i of each subgroup $G_i (i \in 1,\ 2,\ \ldots,\ I)$ to the difference between the demographic measure $f(\cdot)$ for the total population T and the reference population R that satisfies

$$\sum_{i=1}^{I} C^i = f(T) - f(R). \tag{1}$$

3. Proof

3.1 Strategy to develop a proof

The goal is to find a closed-form expression for the C^i terms that satisfies Equation (1). I start by considering the simplest possible case, the one in which the population (T) is partitioned into two subgroups, a reference population (R) and a single contributing group (G_1). We can think of T, R, and G_1 as vectors of age-specific death counts and exposures stacked onto each other. This is not essential for the proof but allows us to write expressions such as $T = R + G_1$ with standard notation for operations between vectors. In the simplest case, it is trivial to show that the difference in the value that any life table function $f(\cdot)$ takes for the total and the reference population is equal to the contribution of the G_1 subgroup or $C^1 = f(T) - f(R)$. Starting from this basic expression, I will then move to the case of two contributing groups G_1 and G_2 and show that the expression for C^1 involves computing and averaging the differences between the value that the life table function takes for each combination of R, G_1, and G_2 that includes the target group G_1 and the one that excludes it. Since the notation quickly becomes complicated, I will first introduce a more complex notation for addressing the case of three contributing groups $I = 3$, before finally moving to the general case of any I. In each case, the proof consists in showing that there exists a symmetric expression for each C^i term such that they collectively satisfy Equation (1). I demonstrate that the expression for each of the C^i involves computing a set of differences between the value taken by $f(\cdot)$ for combinations of R and all $G_j s$ that include and exclude G_i, always asking what is the marginal contribution of adding the G_i group, and averaging them with appropriate weights.

3.2 Proof for *I* = 1, 2

Let us consider a population (T) that is composed of a reference population (R) and two mutually exclusive contributing groups (G_1, G_2). Let us also denote the population combining (G_1, G_2) as G, so $G = G_1 + G_2$. We can think of T, R, and each of the $G_i s$ as vectors stacking age-specific death counts and exposures. In standard demographic notation, we would write

$$T = \begin{matrix} D_0 \\ D_1 \\ D_3 \\ \vdots \\ D_{110} \\ E_0 \\ E_1 \\ E_3 \\ \vdots \\ E_{110} \end{matrix},$$

where D_x is the number of deaths to individuals aged x and E_x is the number of person-years lived by individuals aged x. We can then write $G = T - R$, which becomes a standard element-wise difference operation between vectors. At this point, I will be as general as possible and leave the demographic metric of interest denoted with $f(\cdot)$ unspecified. As in other general decomposition techniques, we will see that computationally all that matters is that we can write the demographic measure as a function of a vector or matrix storing exposures and deaths counts or alternatively rates and age- and group-specific population shares (Andreev, Shkolnikov, and Begun 2002; Caswell 1989; Horiuchi, Wilmoth, and Pletcher 2008; Riffe 2019). We can write the total contribution of the two groups as $C^G = f(T) - f(R)$. Because both the exposures and the death counts involved in the computation of life tables combine additively, we have that $f(R) = f(T - G)$. Equivalently, we could write $f(T) = f(R + G)$. Together, these two equations allow us to write

$$C^G = f(T) - f(R) = f(R + G) - f(R) = f(T) - f(T - G). \tag{2}$$

Looking at Equation (2) it is apparent that the total contribution of population G can equivalently be understood as an addition process starting from the reference population, $C^G = f(R + G) - f(R)$, or as a subtraction process starting from the total population, C$^G = f(T) - f(T - G)$. It now becomes useful to introduce another layer of notation. We define $C^{G+} \equiv f(R+G) - f(R)$ and call it the addition contribution of G. We instead define $C^{G-} \equiv f(T) - f(T - G)$ and call it the subtraction contribution of G. When a single contributing subpopulation is considered, $C^{G+} = C^{G-}$ as Equation (2) showed. However, we will see that even in the case of two contributing subpopulations (G_1, G_2), $C^{1+} \neq C^{1-}$ and $C^{2+} \neq C^{2-}$ so that the notation I just introduced becomes useful.

To see why these two inequalities arise, let us consider $C^{1+} = f(R + G_1) - f(R)$ and $C^{1-} = f(T) - f(T - G_1)$. Suppose for the moment that both G_1 and G_2 contribute positively to the life table function we are considering, and more precisely that

$C^{1+}, C^{1-}, C^{2+}, C^{2-} > 0$. When we compute C^{1-} we are then starting from a population T which already encapsulates the positive effect of population G_2. On the contrary, when we compute C^{1+} we are starting from the reference population R that has not yet received the positive effect of population G_2. Furthermore, population R from which we start our calculations for C^{1+} is smaller than population $R + G_2$ from which we start our calculations for C^{1-}. Consequently, the effect of G_1 will be larger when computed as C^{1+} than when computed as C^{1-} and $C^{1+} > C^{1-}$. A specular argument can be made for C^{2+} and C^{2-}. While the inequality $C^{1+} > C^{1-}$ depends on the assumption that both populations have a positive effect, it illustrates the general principle that leads to the inequality $C^{1+} \neq C^{1-}$. More generally, only when $f(\cdot)$ is linear we have that

$$C^{1+} = f(R + G_1) - f(R) = f(R) + f(G_1) - f(R) = f(G_1)$$

and

$$C^{1-} = f(T) - f(T - G_1) = f(T) - f(T) + f(G_1) = f(G_1)$$

so that

$$C^{1+} = C^{1-}.$$

When instead $f(\cdot)$ is nonlinear, such as life expectancy, we have

$$\begin{aligned} C^{1-} &= f(T) - f(T - G_1) \\ &= f(R + G_1 + G_2) - f(R + G_2) \\ &\neq f(R + G_1) + f(G_2) - f(R) - f(G_2) \end{aligned}$$

because of nonlinearity of $f(\cdot)$

$$= f(R + G_1) - f(R) = C^{1+}.$$

The inequality $C^{1+} \neq C^{1-}$means that we now have two different measures of the contribution of each group. One way of solving this inconsistency is to introduce a third type of contribution $C^1 \equiv \frac{1}{2}(C^{1+} + C^{1-})$, which we call the average contribution. It turns out that the average contributions have a nice property that both the addition and the subtraction contributions generally lack. In particular, we have

$$\begin{aligned}
C^1 + C^2 =& \frac{1}{2}(C^{1+} + C^{1-}) + \frac{1}{2}(C^{2+} + C^{2-}) \\
=& \frac{1}{2}[(f(R+G_1) - f(R)) + (f(T) - f(T-G_1)) + (f(R+G_2) - f(R)) \\
& \qquad + (f(T) - f(T-G_2))] \\
=& \frac{1}{2}[2(f(T) - f(R)) + (f(R+G_1) - f(T-G_2)) \\
& \qquad + (f(R+G_2) - f(T-G_1))] \\
=& f(T) - f(R) + \frac{1}{2}[(f(R+G_1) - f(T-G_2)) + (f(R+G_2) - f(T-G_1))] \\
=& f(T) - f(R) + \frac{1}{2}[(f(R+G_1) - f(R+G_1)) + (f(R+G_2) - f(R+G_2))]
\end{aligned}$$

where the four terms within square brackets in the last line cancel out and leave

$$C^1 + C^2 = f(T) - f(R). \tag{3}$$

Equation (3) proves that the two average contributions sum to $f(T) - f(R)$, satisfying Equation (1), which establishes C^1, C^2 as a legitimate decomposition of the difference $f(T) - f(R)$. The passage from the third-to-last to the penultimate row is possible because with only two G_i subgroups, $T - G_2 = R + G_1$ and $T - G_1 = R + G_2$. Notice that no special properties of the life table function were used in the proof, which establishes that this approach is valid for any function that can be written as a function of vectors of death or event counts and exposures, or equivalently rates and population shares. Also note that although expressions like $T - G_1 = R + G_2$ only strictly make sense if T, R, G_1, and G_2 are vectors of additive quantities, such as death counts and exposures, if we think of $R + G_2$ as the rates for the population rather than a combination of all individuals in the R and G_2 populations, we can still write $f(T - G_1) = f(R + G_2)$, which is all we need for the proof above and its more general version in the following sections. The only restriction imposed by the use of rates and population shares vis-à-vis event counts and exposures is that the former are not suitable to compute metrics $f(\cdot)$ that depend on total population size. However, most commonly used demographic metrics – including all life table measures – are a function of rates only, which means they can be computed for any population combining R and any number of G_i groups starting from group-specific rates and population shares.

3.3 Proof for *I* = 3

The case of three groups (G_1, G_2, G_3) introduces a few additional complications which illuminate some properties of this decomposition and will allow me to develop a general expression for any number of groups. The main difference with the $(I = 2)$ case is that our notation C^{i+}, C^{i-} for $i = 1, 2, 3$ is no longer sufficient to describe all possible ways of computing the contribution of group G_i to the difference $f(T) - f(R)$. Indeed, we now have four ways of computing the contribution of G_1:

1. $f(R + G_1) - f(R)$
2. $f(R + G_2 + G_1) - f(R + G_2)$
3. $f(R + G_3 + G_1) - f(R + G_3)$
4. $f(R + G_3 + G_2 + G_1) - f(R + G_3 + G_2) = f(T) - f(R + G_3 + G_2)$.

In the first expression, the contribution of G_1 is computed as the difference between the life table function $f(\cdot)$ for the reference population and G_1 combined $(R{+}G_1)$ and for the reference population alone (R). This is equivalent to C^{1+}. In the second expression, a new contribution is computed by comparing the value of $f(\cdot)$ for the reference population combined with G_1 and G_2 $(R + G_1 + G_2)$ and for the reference population combined with $G_2(R + G_2)$. In the third expression, G_2 is replaced by G_3. Finally, in the fourth expression, we compare the life table function for the total population $(T = R + G_1 + G_2 + G_3)$ with the one for the total population minus $G_1, (T - G_1 = G_2 + G_3)$, which makes it equivalent to C^{1-}.

In the next section, I will introduce a more general notation to capture these cases. However, for the moment, let us focus on how to define the average contribution in this case, with the idea of finding a set of terms for C^1, C^2, and C^3 that sum to $f(T) - f(R)$ and satisfy Equation (1). It turns out that to do so we need to construct a weighted average of the four different contribution types with weights inversely proportional to the number of groups $I = 3$ and to the number of contributions involving the same number of G_i subpopulations. Notice that contribution 1 contains a single subpopulation (G_1), contribution 4 involves 3 subpopulations (G_1, G_2, and G_3), while contributions 2 and 3 each involve 2 subpopulations, G_1 and G_2 and G_1 and G_3, respectively. As such, in this weighting scheme contributions 1 and 4 have weight $\frac{1}{3}\frac{1}{1} = \frac{1}{3}$, while contributions 2 and 3 have weight $\frac{1}{3}\frac{1}{2} = \frac{1}{3}\frac{1}{2} = \frac{1}{6}$. The use of these weights is equivalent to first averaging within contributions involving the same number of subpopulations and then averaging between contributions involving different numbers of subpopulations. It is important to note that contributions involving the total population T always involve all the G_i subgroups. While this weighting scheme might seem unintuitive, I will now show that these weights lead to a set of average contributions that sum to the total contribution, as in the $I = 2$ case.

$$\begin{aligned}
C^1 + C^2 + C^3 =& \frac{1}{3}(f(R+G_1) - f(R)) + \frac{1}{6}(f(R+G_2+G_1) - f(R+G_2)) \\
&+ \frac{1}{6}(f(R+G_3+G_1) - f(R+G_3)) + \frac{1}{3}(f(T) - f(R+G_2+G_3)) \\
&+ \frac{1}{3}(f(R+G_2) - f(R)) + \frac{1}{6}(f(R+G_1+G_2) - f(R+G_1)) \\
&+ \frac{1}{6}(f(R+G_3+G_2) - f(R+G_3)) + \frac{1}{3}(f(T) - f(R+G_1+G_3)) \\
&+ \frac{1}{3}(f(R+G_3) - f(R)) + \frac{1}{6}(f(R+G_1+G_3) - f(R+G_1)) \\
&+ \frac{1}{6}(f(R+G_2+G_3) - f(R+G_2)) + \frac{1}{3}(f(T) - f(R+G_1+G_2)).
\end{aligned}$$

It is easy to see that one can the take target value $f(T) - f(R)$ out of this complicated expression. We can then reorder the remaining terms and write

$$\begin{aligned}
C^1 + C^2 + C^3 =& (f(T) - f(R)) \\
&+ \frac{1}{3}(f(R+G_1) + f(R+G_2) + f(R+G_3)) \\
&- \frac{1}{3}(f(R+G_1+G_2) + f(R+G_1+G_3) + f(R+G_2+G_3)) \\
&- \frac{1}{6}(2 \cdot f(R+G_1) + 2 \cdot f(R+G_2) + 2 \cdot f(R+G_3)) \\
&+ \frac{1}{6}(2 \cdot f(R+G_1+G_2) + 2 \cdot f(R+G_1+G_3) + 2 \cdot f(R+G_2+G_3)).
\end{aligned}$$

The second and fourth lines cancel out, as do the third and fifth lines, leaving just

$$C^1 + C^2 + C^3 = f(T) - f(R). \tag{4}$$

Equation (4) proves that (C^1, C^2, C^3) satisfy Equation (1) and establishes them as a legitimate decomposition of $f(T) - f(R)$.

3.4 Proof for $I \geq 4$

To study the general case of I contributing groups, it is useful to introduce a slightly more general notation. We define

$$C^{\,i,m,k} = f(R + G_k^m + G_i) - f(R + G_k^m), \tag{5}$$

with $i = \{1, 2, \ldots, I\}$, $m = \{0, 1, \ldots, I-1\}$, $k = 1, 2, \ldots, \binom{I-1}{m}$ and where G_k^m is the k^{th} set of subpopulations G_j with $j \neq i$ such that $\#G_k^m = m$. In other words, $C^{i,m,k}$ is the contribution of subgroup G_i calculated as the difference between the value of $f(\cdot)$ for the population obtained by combining the reference population, the k^{th} subset of G_j's not including G_i, and G_i, $f(R + G_k^m + G_i)$, and the value of $f(\cdot)$ for the population obtained by combining the reference population and the k^{th} possible subset of G_j's not including G_i, $f(R + G_k^m)$. This definition is a generalization of the types of contributions we have seen in the cases $I = 2, 3$. Indeed, for $I = 2$, C^{1+} is simply $C^{1,0,1}$ while C^{1-} is $C^{1,1,1}$. The logic is that we are always comparing the value a demographic measure takes for a population with and without the group whose contribution we want to estimate. Back to the $I = 3$ case, we can now write

$$C^{1,0,1} = f(R + G_1) - f(R) = C^{1+} \tag{6}$$

for the simplest case with no other groups except G_1 involved in the expression. We can also write

$$C^{1,1,1} = f(R + G_1^1 + G_1) - f(R + G_1^1) = f(R + G_2 + G_1) - f(R + G_2) \tag{7}$$

$$C^{1,1,2} = f(R + G_2^1 + G_1) - f(R + G_2^1) = f(R + G_3 + G_1) - f(R + G_3) \tag{8}$$

for the two intermediate cases with another group involved $(m = 1)$, and

$$\begin{aligned} C^{1,2,1} &= f(R + G_1^2 + G_1) - f(R + G_1^2) \\ &= f(R + G_2 + G_3 + G_1) - f(R + G_2 + G_3) \\ &= f(T) - f(T - G_1) \\ &= C^{1-} \end{aligned} \tag{9}$$

for the case in which both remaining groups, G_2 and G_3, are included $(m = 2)$. With these expressions, we can see why m should be bounded between 0 and $I - 1$ and k

between 1 and $\binom{I-1}{m}$. With $m = 0$, we recover C^{1+}. In this case there is no need for an index k, which we can conventionally set to 1. For $m = 1$, G_k^m contains only one element which we can choose out of the $I - 1$ groups that are not G_i. The first of these groups will be denoted as G_1^1, the second as G_2^1, and so on. The order in which we select the groups does not matter because addition is commutative. Notice that, for a general value m there will be $\binom{I-1}{m}$ of these groups, which explains the limits imposed on k. Finally, when we reach $m = I - 1$, we obtain C^{1-} because even in the general case,

$$f(R + G_1^{I-1} + G_1) - f(R + G_1^{I-1}) = f(T) - f(T - G_1) = C^{1-}. \tag{10}$$

Clearly, there is only one way of choosing $I - 1$ elements from a set of $I - 1$ elements.

As for the $I = 3$ case, we now need to find appropriate weights for each contribution $C^{i,m,k}$ to construct C^i terms that sum to $f(T) - f(R)$. The general principle is still the same, we need weights that are inversely proportional to the number of subgroups I and inversely proportional to the number of contributions of type m. This consideration leads to the simple expression

$$w^m = \frac{1}{I\binom{I-1}{m}} \tag{11}$$

which is just a generalization of the weights we derived for the $I = 3$ case. The expression for w^m in Equation (11) denotes the weight for all terms $C^{i,m,k}$ as defined in Equation (5). These weights are equal for all m-type contributions and ensure that each set of m contributions is collectively assigned the same weight $\frac{1}{I}$ while each of its members is also assigned the same weight $\frac{1}{I\binom{I-1}{m}}$. With our new notation and having defined appropriate weights, we can define the average contribution of group G_i as:

$$C^i = \sum_{m=0}^{I-1} \frac{1}{I} \sum_{k=1}^{\binom{I-1}{m}} \frac{1}{\binom{I-1}{m}} C^{i,m,k}. \tag{12}$$

To establish that the average contributions defined according to Equation (12) form a valid decomposition of $f(T) - f(R)$ into group-specific contributions, we need to prove that they satisfy Equation (1):

$$\sum_{i=1}^{I} C^i = f(T) - f(R).$$

We already saw that for $I = 2, 3$, the proof involves showing that all terms except $f(T)$ and $f(R)$ have weights summing to 0. This is harder to do directly now that we have many terms. However, we can start by recognizing that the life table function computed for a given population composed by a set of subgroups plus the reference population can appear in just two ways in Equation (1):

1. As a term of the form $f(R + G_k^m + G_i)$ in a contribution of the type $C^{i,m,k}$
2. As a term of the form $f(R + G_k^{m+1})$ in a contribution of the type $C^{j,m+1,k}$ where $j \neq i$.

In the first case, the life table function will have weight w^m and in the second case it will have weight $-w^{m+1}$. To understand what the final weight will be for each term $f(R + G_k^m + G_i)$, we just need to know how many times it will appear with weight w^m and how many times with weight w^{m+1}. For terms of the first type, we have $m + 1$ ways of choosing i while keeping the subgroups involved the same. On the other hand, terms of the second type will only appear in contributions involving the $I - (m + 1)$ excluded groups. With this information, we are now able to compute the weight associated with the demographic measure computed for each population:

$$\begin{aligned} w(f(R + G_k^m + G_i)) = (m+1)w^m - (I-m-1)w^{m+1} \\ &= (m+1)\frac{1}{I\binom{I-1}{m}} - (I-m-1)\frac{1}{I\binom{I-1}{m+1}} \\ &= (m+1)\frac{m!(I-1-m)!}{I\cdot(I-1)!} \\ &\qquad - (I-m-1)\frac{(m+1)!(I-1-m-1)!}{I\cdot(I-1)!} \\ &= \frac{(m+1)!(I-1-m)!}{I\cdot(I-1)!} - \frac{(m+1)!(I-1-m)!}{I\cdot(I-1)!} \end{aligned}$$

The two terms cancel out and we are left with

$$w(f(R + G_k^m + G_i)) = 0. \tag{13}$$

Equation (13) proves that life table functions for all populations of the type $R + G_k^m + G_i$ cancel out in Equation (1). The only two exceptions are $f(R)$, which appears I times with weight $-\frac{1}{I}$, and $f(T)$, which appears I times with weight $\frac{1}{I}$. Thus, summing the average contributions as defined in Equation (12) for our I groups we obtain

$$\sum_{i=1}^{I} C^i = f(T) - f(R)$$

as desired. This result shows that the contributions in Equation (12) satisfy Equation (1) and provide a decomposition of $f(T) - f(R)$ even in the case of I subgroups.

4. Origins of the decomposition and notes on its interpretation

The idea that the contribution of a group to total life expectancy could be expressed as the difference between total life expectancy and life expectancy of the population excluding the group of interest was originally proposed to investigate the contribution of the foreign-born population to total life expectancy at age 65 across high income countries (Dupre, Gu, and Vaupel 2012). Similar analyses were carried out for Australia (Page et al. 2007) and for New York City (Preston and Elo 2014), but without explicitly characterizing the difference between total and native-born life expectancy as a contribution of the foreign-born population to total life expectancy. The same logic has since been applied to investigate the foreign-born contribution to total life expectancy in the United States (Hendi and Ho 2021; Paglino and Elo 2025) and in Nordic countries (Wallace et al. 2022; Wallace and Drefahl 2024). In these applications, a total population T including foreign-born and native-born residents, a reference population R including only native-born residents, and a single contributing group G including all foreign-born residents are considered. Dupre, Gu, and Vaupel (2012) proposes to compute the contribution of the foreign-born population to total life expectancy at age 65 as

$$C^{FB} = e_{65}^{T} - e_{65}^{NB} = e_{65}^{T} - e_{65}^{R}, \tag{14}$$

which equals C^1 in the notation I have been using. In Equation (14), e_{65}^{T} is the life expectancy at age 65 for the total population and $e_{65}^{NB} = e_{65}^{R}$ is the life expectancy at age 65 for native-born residents (the reference population). This particular case mirrors the general expression in Equation (1), where the generic demographic measure $f(\cdot)$ is replaced with life expectancy at age 65. The logic behind this formula is clear: The difference between the life expectancy of all residents and native-born residents must be explained by the contribution of the foreign-born population. However, it is important to note that previous works have sometimes conflated the contribution of the nonreference population to the life expectancy gap between the total and the reference population $f(T) - f(R)$

with the contribution of the nonreference population to total life expectancy $f(T)$. This interpretation is problematic and leads to inconsistencies even in the simple case with a single nonreference population. In the context of native-born vis-à-vis foreign-born mortality, if M^{NB} is the vector of age-specific mortality rates for native-born residents, M^{FB} is the vector of age-specific mortality rates for the foreign-born residents, and M^{T} is the vector of age-specific mortality rates for the total population, and assuming $e_x^{FB} \geq e_x^{NB}$, there is no guarantee that life expectancy at age x for the total population e_x^{T} will be bounded by the corresponding life expectancies of the native-born e_x^{NB} and foreign-born e_x^{FB} populations:

$$e_x^{NB} \leq e_x(M^{FB}w^{FB} + M^{NB}w^{NB}) = e_x^{T} \leq e_x^{FB},$$

where w^{FB} and w^{NB} are vectors of age-specific weights proportional to the share of the foreign-born and native-born in the population, respectively, and such that $w_x^{FB} + w_x^{NB} = 1 \; \forall x$. The two inequalities fail to hold because a priori,

$$\max\left(e_x^{T}(M^{FB}w^{FB} + M^{NB}w^{NB})\right) \; w.r.t. \; (w^{FB}, w^{NB}) = e_x(\min\left(M^{FB}, M^{NB}\right)),$$

where $\min\left(M^{FB}, M^{NB}\right)$ is the set of elementwise (age-specific) minimums, and

$$\min\left(e_x^{T}(M^{FB}w^{FB} + M^{NB}w^{NB})\right) \; w.r.t. \; (w^{FB}, w^{NB}) = e_x(\max\left(M^{FB}, M^{NB}\right)),$$

where $\max\left(M^{FB}, M^{NB}\right)$ is the set of elementwise (age-specific) maximums. Equivalently, as long as the two sets of mortality rates M^{FB}, M^{NB} cross at some point (i.e., $\exists i, j \; s.t. \; M_i^{FB} > M_i^{NB} \; \bigwedge \; M_j^{FB} < M_j^{NB}$), there will be a set of weights such that $e_x^{T} > \max(e_x^{FB}, e_x^{NB})$ and also a set of weights such that $e_x^{T} < \min(e_x^{FB}, e_x^{NB})$. The failure of e_x^{T} to be bounded by (e_x^{FB}, e_x^{NB}) and in the more general notation of e_x^{T} to be bounded by (e_x^{G}, e_x^{R}) has two consequences that affect how we interpret the contribution of a group G computed as $f(T) - f(R)$:

1. It is possible that $C^{G} = e_x^{T} - e_x^{R} > e_x^{G} - e_x^{R}$. In other words, the contribution of group G could be larger than the difference between its life expectancy and that of the reference population.
2. Suppose we exchange the role of the contributing group G and that of the reference population R. We would then define $C^{R} = e_x^{T} - e_x^{G}$. It is possible that both $C^{R} > 0$ and $C^{G} > 0$ or that $C^{R} < 0$ and $C^{G} < 0$ even though R and G, by definition, form the total population T.

In addition, contrary to what the interpretation of C^G as the contribution of G to e_x^T would suggest, $C^R + C^G \neq e_x^T$. In fact, $C^R + C^G = e_x^T - e_x^R + e_x^T - e_x^G = 2e_x^T - e_x^R - e_x^G$, which only equals e_x^T in the unlikely event that $e_x^R + e_x^G = e_x^T$. Given these observations, it is important to understand that the interpretation of $f(T) - f(R)$ as the contribution of group G to $f(T)$ given by Dupre, Gu, and Vaupel (2012) and subsequent studies is incorrect. In contrast to what this interpretation suggests, the contribution of a group G, as well as the contributions of subgroups G_1, G_2, $\ldots$, G_I to $f(T) - f(R)$, depend on the reference population. Changing the reference will lead to a different decomposition problem and thus to different contributions. Keeping this in mind, researchers should be explicit in formulating their research question when asking about the contribution of certain subgroups to the difference in a demographic indicator between the total and the reference population to avoid any ambiguity.

5. Use cases

The decomposition developed in this paper can be used to investigate the contribution of any set of mutually exclusive groups to the difference in the values of a demographic measure for the total and reference populations. As mentioned before, Dupre, Gu, and Vaupel (2012) study how the foreign-born population contributes to the gap between national life expectancies and life expectancies of native-born populations. The groupwise decomposition would allow to extend this analysis to examine the contributions of subgroups of the foreign-born population, by country of origin, race/ethnicity (as in Paglino and Elo 2025), or educational attainment. Other interesting applications include the investigation of how different regions contribute to the gap between national life expectancy and the life expectancy of a reference region, how different countries contribute to the gap between the regional life expectancy and the life expectancy of a reference country, or how different educational groups contribute to the difference between the life expectancy of the total population and that of the group with the lowest educational attainment.

Like other general decompositions, while I have so far considered examples related only to mortality, applications to fertility, migration, or even projections are possible. Examples include computing the contribution of the foreign-born population to the difference between the national total fertility rate and the total fertility rate for the native-born population and decomposing it into education-specific terms. The same exercise could be carried out for the net reproduction rate, combining fertility and mortality rates. In principle, even questions about future quantities can be addressed with this method. For example, one could use a Lee–Carter model (Lee and Carter 1992) to forecast mortality rates for the United States as a whole and for large central metros in the United States and ask how much counties in other metropolitan codes (fringe metros, medium metros, etc.) contribute to the difference between the two forecasts.

6. Application

For this application, I use death counts and population data from the National Center for Health Statistics (NCHS) obtained through CDC WONDER. Deaths and population are classified by year (2017–2019), sex, five-year age groups (<1, $1-4$, $5-9$, $\ldots$, $85+$), and one of four metropolitan–nonmetropolitan codes (large central metro, large fringe metro, medium and small metro, and nonmetro). Details on the classification criteria are available in the NCHS online documentation (NCHS 2023).

At the national level, large fringe metro counties have the highest life expectancy for males, followed by large central metro counties, medium and small metro counties, and nonmetro counties (Figure 1A). The ordering is essentially the same for females, the only difference being that large central metro counties have the highest life expectancy, with large fringe metro counties coming second (Figure 1A). Large central metros and medium and small metros each account for about 30% of the national population for both sexes (Figure 1B). The remaining 40% is split between large fringe metros (25%) and nonmetropolitan counties (15%). Figure 2 shows the group-specific contributions to the gap between national life expectancy and the life expectancy of large central metros using the groupwise decomposition. As we would have expected based on the life expectancy ranking in Figure 1A, all groups but large fringe metro counties contribute negatively. Nonmetropolitan counties have the largest absolute contribution, followed by medium and small metro counties, and large fringe metro counties. The size of each contribution is determined by the interaction of three factors: the population size of counties belonging to the specific group, the population's age distribution, and its age-specific mortality rates (which life expectancy conveniently summarizes).

Figure 1: Life expectancy at birth (panel A) and mid-year population (panel B) by year, sex, and metropolitan–nonmetropolitan category

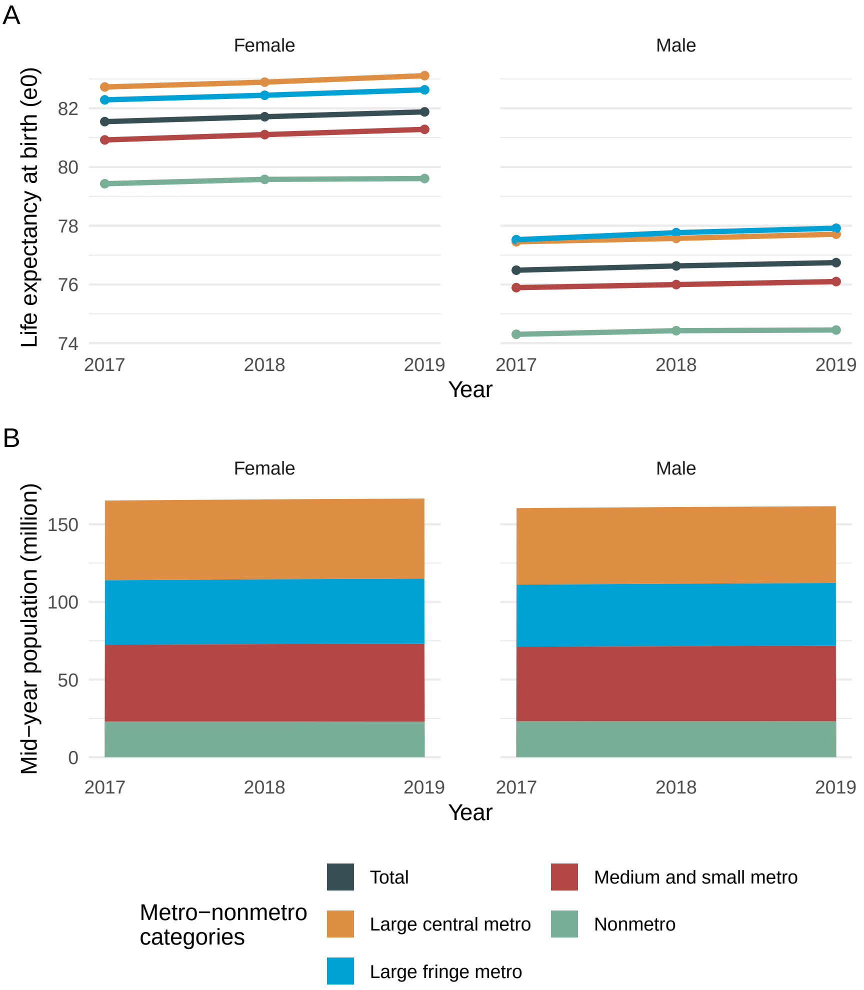
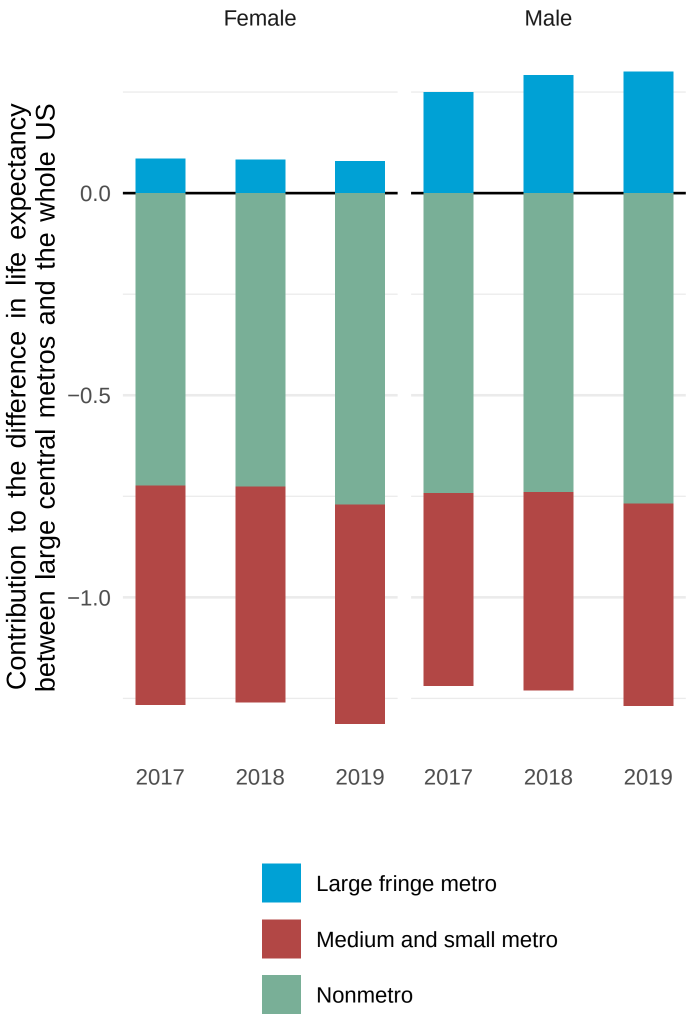


Figure 2: **Decomposing the contribution of counties by metropolitan–nonmetropolitan category to the gap between national life expectancy and life expectancy in large central metro counties**



As a further test of the validity of the groupwise decomposition, I adapted the three general decomposition methods (Andreev, Shkolnikov, and Begun 2002; Caswell 1989; Horiuchi, Wilmoth, and Pletcher 2008) as implemented in the DemoDecomp package

(Riffe 2019) to obtain a decomposition by group similar to the one developed in this paper. The relationship between the three general decomposition methods and the groupwise decomposition is not immediately obvious. The three general decomposition methods ask how element-specific changes in the matrix of input values are related to changes in a demographic metric of interest. The groupwise decomposition asks instead how different subpopulations, with their death counts and exposures, contribute to the difference in a demographic metric of interest between the total population and a reference subpopulation. Additionally, in the case of the line-integral decomposition, the authors explicitly stated that this technique is not suitable to study compositional effects. However, we can, for a moment, ignore the original framing of the existing general methods and focus on the practical requirement of being able to reframe the groupwise decomposition problem as a question about how element-specific changes between two input matrices contribute to changes in a demographic indicator. This is not trivial because it is not obvious what the two input matrices leading to two different values of the demographic metric of interest should be. It turns out that what we need is a matrix which contains stacked death counts by age, followed by stacked exposures by age, with one column for each population subgroup, including the reference population. The exact structure of the matrix, such as whether exposures or death counts come first, whether they are organized by rows or columns, or the order of the subgroups, is not crucial and changes only how the matrix elements are reorganized to compute the demographic metric of interest. To get the target demographic metric for the reference population, death counts and exposures should be filled only for the reference population and set to zero for the remaining subgroups. Conversely, to get the target demographic metric for the total population, the matrix should contain actual death counts and exposures for all subgroups, including the reference population. Equipped with these two matrices, we are left with the task of writing a function that takes such matrices as inputs and outputs the demographic metric of interest. This is the same generic function $f(\cdot)$ we used in the proof of the decomposition approach developed in this paper. Each of the general decomposition methods will then give us a matrix of element-specific contributions to the difference between the two values of the demographic indicator. To summarize these contributions and obtain quantities comparable to those produced by the groupwise approach, we need to sum over rows to obtain a single contribution for each subgroup.

The comparison between the results obtained with the groupwise decomposition and those obtained with each of the three methods following the procedure detailed above are presented in Figure 3 and Figure 4 (in which the stepwise decomposition is removed). They show that the line-integral method (Horiuchi, Wilmoth, and Pletcher 2008) and the life table response experiment (Caswell 1989) produce results almost identical to those of the groupwise decomposition. The fact that the results are practically equivalent establishes that the two continuous change general decomposition approaches are also useful to tackle groupwise decomposition problems. The stepwise decomposition fails to produce reasonable results when simply averaging the contributions obtained through element-wise replacement starting (1) from the top-left element and moving down across rows and rightward across columns, and (2) from the bottom-right element and moving up across rows and leftward across columns, as recommended by the authors (Andreev, Shkolnikov, and Begun 2002) and in line with similar methods specific to life expectancy (Andreev 1982; Pressat 1985). The failure of the basic stepwise decomposition is likely related to the strong dependence of the results on the order in which elements are replaced. In principle, averaging across contributions obtained with all possible permutations should give results close to those of the line-integral and life table response experiment methods, but even in the simple case with two subgroups and 19 age groups (< 1, $1-4$, $5-9, \dots, 85+$), it would involve $3 \cdot 19 \cdot 2 = 114$ elements, and thus the computation of $2^{114} > 10^{34}$ life expectancies (or some other metric). Although it might be possible to identify a more efficient approximation – for example, one that averages across a sample of random replacement orders – this was beyond the scope of the present paper.

A legitimate question is why one should use the groupwise decomposition if both the line-integral and the life table response experiment methods provide similar results as I showed. One answer is that the groupwise decomposition is computationally very efficient compared with the alternatives. For the example presented above, where a very simple life table function is used (LTabr from DemoDecomp), the average execution time measured with the microbenchmark package (Mersmann et al. 2023) was 0.05 seconds for the groupwise decomposition, 42.80 seconds for the line-integral method, and 187.22 seconds for the life table experiment, both of the latter as implemented in the DemoDecomp package. In this relatively simple application, the groupwise decomposition is thus more than 850 times faster than the line-integral method and about 3,700 times faster than the life table experiment method. While an execution time of 43 seconds might still appear fast, the difference between an execution time of 0.05 vis-à-vis 43 seconds for a single computation of the decomposition becomes more relevant when wishing to use bootstrapping or other simulation methods to produce uncertainty intervals around the decomposition results. With 1,000 replications, the total execution time for the line-integral decomposition method would increase to approximately 12 hours vis-à-vis 50 seconds for the groupwise approach.

Figure 3: Comparing decomposition results from the groupwise decomposition with those obtained with other general decomposition methods: Horiuchi, Life table response experiment, and stepwise

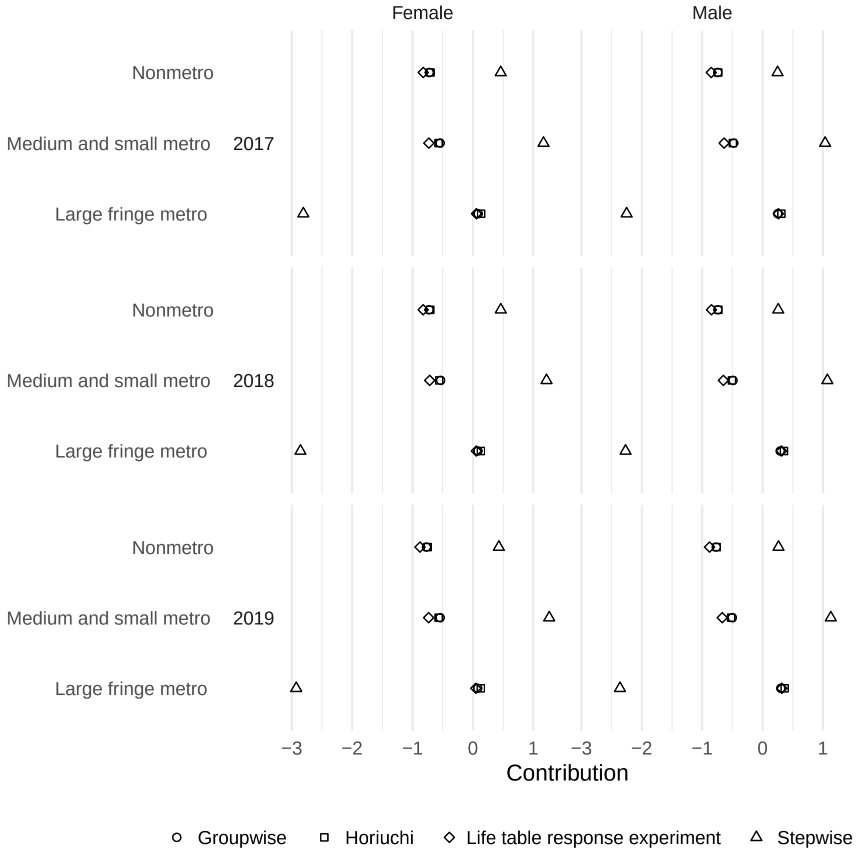
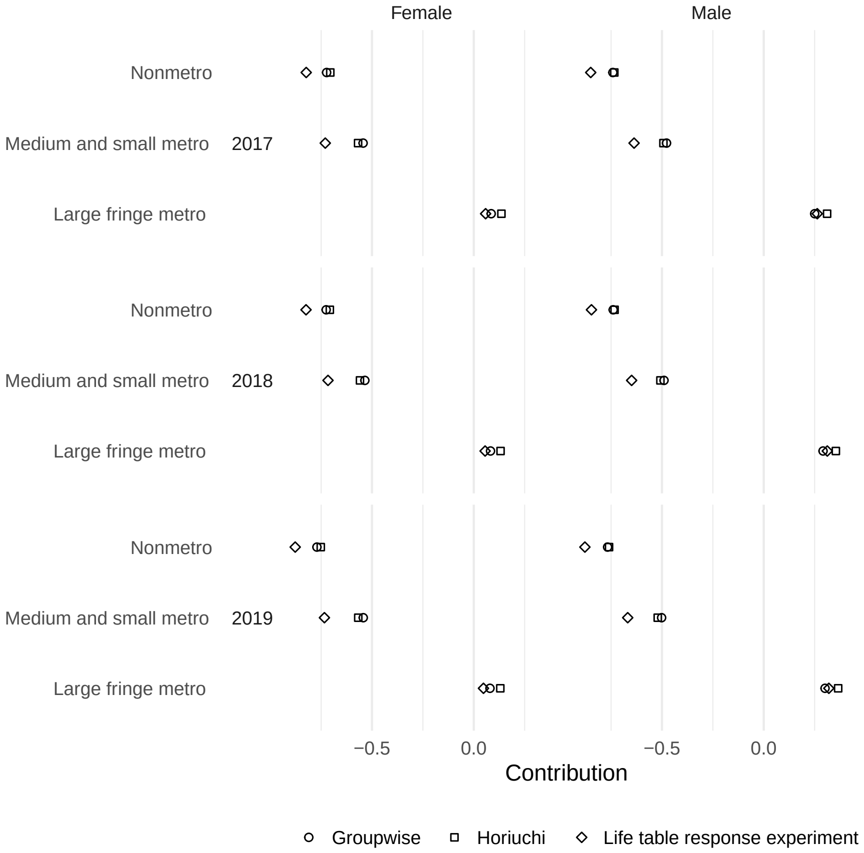


Figure 4: Comparing decomposition results from the groupwise decomposition with those obtained with the Horiuchi and life table response experiment general decomposition methods



7. A simple extension: Combining decomposition techniques to compute age- and group-specific contributions

Once we have seen how the decomposition works, we can quickly recognize two facts. First, while we considered a function mapping vectors to scalars, nothing in the proof requires it. We can thus extend the decomposition to functions that take as input a vector of death counts and exposures and outputs a vector of demographic metrics. Second, general decomposition techniques are themselves a useful class of vector-to-vector functions that satisfy this property. Recognizing these two facts, we can easily combine other decomposition techniques with the groupwise decomposition to further decompose group-specific contributions by other dimensions of interest. The only difference is that now each of the $C^{i,m,k} = f\left(R + G_k^m + G_i\right) - f\left(R + G_k^m\right)$ contributions in Equation (5) is a vector of contributions – for example, by age and cause of death in applications involving mortality, or by parity or age of mother in applications involving fertility – obtained by applying one of the general decomposition methods (or another analytical method when possible) to the difference $f\left(R + G_k^m + G_i\right) - f\left(R + G_k^m\right)$. This is the same trick proposed by Riffe, Tursun-Zade, and Trias-Llimós (2024), who combine the Kitagawa (Kitagawa 1955) and Arriaga (Arriaga 1984) decompositions to express differences in life expectancy – but the technique can be extended to other demographic metrics – between two time points or populations as a function of group- or age-specific mortality rates and compositional change.

Figure 5 shows an example of this extension, using age as the additional dimension. Figure 6 replicates the decomposition using the line-integral method and once again shows that the results are nearly identical. In this more complex application, the computational advantages of the groupwise decomposition are smaller because the function being decomposed now outputs vectors rather than scalars, while the computations for the line-integral decomposition remain the same, only the aggregation of contributions changes. Again, using the microbenchmark package, I timed the mean execution time for the groupwise decomposition at 0.3 seconds compared with 46.8 seconds for the line-integral method as implemented in the DemoDecomp package (with the groupwise decomposition 160 times faster).

Figure 5: **Groupwise decomposition of contributions to the gap between national life expectancy and life expectancy of large central metro counties by metro type and age (2019)**

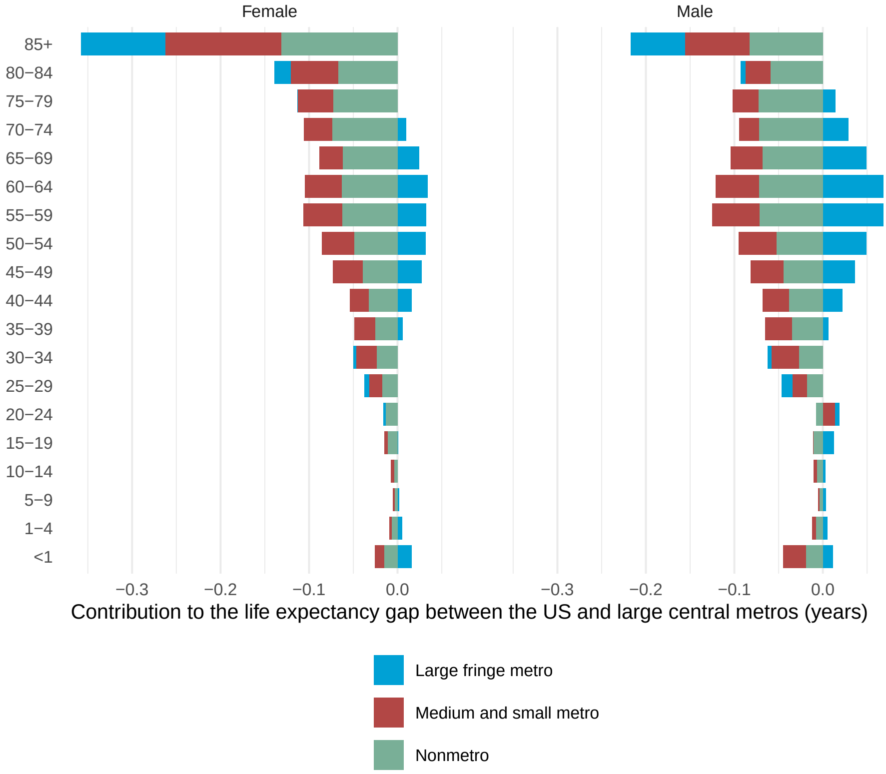
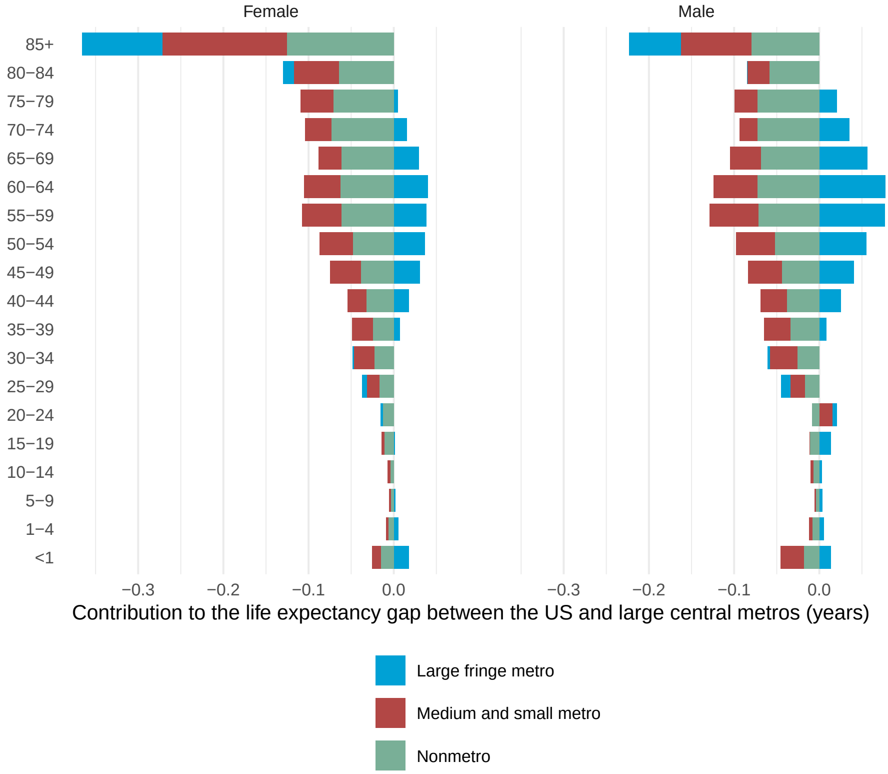


Figure 6: Line-integral decomposition (Horiuchi, Wilmoth, and Pletcher 2008) of contributions to the gap between national life expectancy and life expectancy of large central metro counties by metro type and age (2019)



8. Conclusion

This paper provides a new decomposition method to investigate how different subpopulations contribute to the difference between a demographic measure for the total population and for a reference population. This setup is not infrequent in demography and arises with research questions such as "How much do counties belonging to different groups in the metropolitan–nonmetropolitan continuum contribute to the large central metro mortality advantage compared with national rates?" or "How much do different educational groups contribute to the life expectancy gap between the population with a professional or graduate degree and the total population?" The decomposition method developed in this paper will enable researchers to answer these and other questions, expanding our understanding of how different groups contribute to the demographic profile of a population.

9. Acknowledgements

I wish to thank the four anonymous reviewers and the editor for providing feedback on this manuscript. The paper has truly benefited from their constructive comments and has greatly improved through the peer-review process. I also wish to thank Samuel Preston, Elizabeth Wrigley-Field, Irma Elo and José Monteiro da Silva for reading previous drafts of this paper and for giving me great advice and encouragement. Finally, I wish to thank the attendees of the Innovations in Demographic Methods and Formal Demography session at the 2024 PAA Annual Meeting. This research was supported by the Max Planck – University of Helsinki Center for Social Inequalities in Population Health, LIFECON project (#345219), and the International Max Planck Research School for Population, Health and Data Science (IMPRS-PHDS).

References

Andreev, E.M. (1982). Metod komponent v analize prodoljitelnosty zjizni. [The method of components in the analysis of length of life]. *Vestnik Statistiki* 9: 42–47.

Andreev, E.M., Shkolnikov, V.M., and Begun, A.Z. (2002). Algorithm for decomposition of differences between aggregate demographic measures and its application to life expectancies, healthy life expectancies, parity-progression ratios and total fertility rates. *Demographic Research* 7(14): 499–522. doi:10.4054/DemRes.2002.7.14.

Arriaga, E.E. (1984). Measuring and explaining the change in life expectancies. *Demography* 21(1): 83–96. doi:10.2307/2061029.

Canudas Romo, V. (2003). *Decomposition methods in demography*. (Thesis fully internal (DIV), University of Groningen). Amsterdam: Rozenberg, University of Groningen.

Caswell, H. (1989). Analysis of life table response experiments I. Decomposition of effects on population growth rate. *Ecological Modelling* 46(3): 221–237. doi:10.1016/0304-3800(89)90019-7.

Chandrasekaran, C. (1949). The effect of the change in mortality conditions in an age group on the expectation of life at birth. *Human Biology* 21(1): 35–46.

Das Gupta, P. (1978). A general method of decomposing a difference between two rates into several components. *Demography* 15(1): 99–112. doi:10.2307/2060493.

Dupre, M.E., Gu, D., and Vaupel, J.W. (2012). Survival differences among native-born and foreign-born older adults in the United States. *PLOS ONE* 7(5): e37177. doi:10.1371/journal.pone.0037177.

Edwards, R.D. and Tuljapurkar, S. (2005). Inequality in life spans and a new perspective on mortality convergence across industrialized countries. *Population and Development Review* 31(4): 645–674. doi:10.1111/j.1728-4457.2005.00092.x.

Enroth, L., Jasilionis, D., Németh, L., Strand, B.H., Tanjung, I., Sundberg, L., Fors, S., Jylhä, M., and Brønnum-Hansen, H. (2022). Changes in socioeconomic differentials in old age life expectancy in four Nordic countries: The impact of educational expansion and education-specific mortality. *European Journal of Ageing* 19(2): 161–173. doi:10.1007/s10433-022-00698-y.

Hendi, A.S. and Ho, J.Y. (2021). Immigration and improvements in American life expectancy. *SSM – Population Health* 15: 100914. doi:10.1016/j.ssmph.2021.100914.

Horiuchi, S., Wilmoth, J.R., and Pletcher, S.D. (2008). A decomposition method based on a model of continuous change. *Demography* 45(4): 785–801. doi:10.1353/dem.0.0033.

Kitagawa, E.M. (1955). Components of a difference between two rates. *Journal of the American Statistical Association* 50(272): 1168–1194. doi:10.1080/01621459.1955.10501299.

Lee, R.D. and Carter, L.R. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* 87(419): 659–671. doi:10.1080/01621459.1992.10475265.

Luy, M., Zannella, M., Wegner-Siegmundt, C., Minagawa, Y., Lutz, W., and Caselli, G. (2019). The impact of increasing education levels on rising life expectancy: A decomposition analysis for Italy, Denmark, and the USA. *Genus* 75(11): 1–21. doi:10.1186/s41118-019-0055-0.

Mersmann, O., Beleites, C., Hurling, R., Friedman, A., and Ulrich, J.M. (2023). Microbenchmark: Accurate timing functions. R package. https://github.com/joshuaulrich/microbenchmark.

NCHS (2023). NCHS urban rural classification scheme for counties. Hyattsville, MD: NCHS. https://www.cdc.gov/nchs/data-analysis-tools/urban-rural.html.

Page, A., Begg, S., Taylor, R., and Lopez, A.D. (2007). Global comparative assessments of life expectancy: The impact of migration with reference to Australia. *Bulletin of the World Health Organization* 85(6): 474–481. doi:10.2471/BLT.06.036202.

Paglino, E. and Elo, I.T. (2025). US-born and foreign-born life expectancy by race and Hispanic origin before and during the COVID-19 pandemic in the United States. *Social Science and Medicine* 380: 118191. doi:10.1016/j.socscimed.2025.118191.

Pollard, J.H. (1982). The expectation of life and its relationship to mortality. *Journal of the Institute of Actuaries* 109(2): 225–240. doi:10.1017/S0020268100036258.

Ponnapalli, K.M. (2005). A comparison of different methods for decomposition of changes in expectation of life at birth and differentials in life expectancy at birth. *Demographic Research* 12(7): 141–172. doi:10.4054/DemRes.2005.12.7.

Pressat, R. (1985). Contribution des écarts de mortalité par âge à la différence des vies moyennes. *Population (French Edition)* 40(4/5): 766–770. doi:10.2307/1532986.

Preston, S.H. and Elo, I.T. (2014). Anatomy of a municipal triumph: New York City's upsurge in life expectancy. *Population and Development Review* 40(1): 1–29. doi:10.1111/j.1728-4457.2014.00648.x.

Preston, S.H., Himes, C., and Eggers, M. (1989). Demographic conditions responsible for population aging. *Demography* 26(4): 691–704. doi:10.2307/2061266.

Riffe, T. (2019). DemoDecomp: General demographic decomposition methods. R package. https://github.com/timriffe/DemoDecomp.

Riffe, T., Tursun-Zade, R., and Trias-Llimós, S. (2024). Arriaga meets Kitagawa. Life expectancy decomposition with population subgroups. (MPIDR Working Paper WP-2024-029). Rostock: MPIDR. doi:10.4054/MPIDR-WP-2024-029.

Shkolnikov, V.M., Andreev, E.M., and Begun, A.Z. (2003). Gini coefficient as a life table function: Computation from discrete data, decomposition of differences and empirical examples. *Demographic Research* 8(11): 305–358. doi:10.4054/DemRes.2003.8.11.

Shkolnikov, V.M., Andreev, E.M., Jasilionis, D., Leinsalu, M., Antonova, O.I., and McKee, M. (2006). The changing relation between education and life expectancy in Central and Eastern Europe in the 1990s. *Journal of Epidemiology and Community Health* 60(10): 875–881. doi:10.1136/jech.2005.044719.

Su, W., Raalte, A., Aburto, J.M., and Canudas-Romo, V. (2024a). Subnational contribution to life expectancy and life span variation changes: Evidence from the United States. *Demographic Research* 50(54): 583–624. doi:10.4054/DemRes.2024.50.22.

Su, W., Welsh, J., Korda, R.J., and Canudas-Romo, V. (2024b). Educational composition effect on the sex gap in life expectancy: A research note based on evidence from Australia. *Population Studies* 78(2): 361–369. doi:10.1080/00324728.2023.2273466.

Timonin, S., Shkolnikov, V.M., Jasilionis, D., Grigoriev, P., Jdanov, D.A., and Leon, D.A. (2016). Disparities in length of life across developed countries: Measuring and decomposing changes over time within and between country groups. *Population Health Metrics* 14(29). doi:10.1186/s12963-016-0094-0.

Torres, C., Canudas-Romo, V., and Oeppen, J. (2019). The contribution of urbanization to changes in life expectancy in Scotland, 1861–1910. *Population Studies* 73(3): 387–404. doi:10.1080/00324728.2018.1549746.

Vaupel, J.W. and Canudas-Romo, V. (2002). Decomposing demographic change into direct vs. compositional components. *Demographic Research* 7(1): 1–14. doi:10.4054/DemRes.2002.7.1.

Vaupel, J.W. and Canudas Romo, V. (2003). Decomposing change in life expectancy: A bouquet of formulas in honor of Nathan Keyfitz's 90th Birthday. *Demography* 40(2): 201–216. doi:10.2307/3180798.

Wallace, M. and Drefahl, S. (2024). Against the grain: International migrants, the children of migrants and national life expectancy in Sweden, 1990–2019. *SSM – Population Health* 28: 101726. doi:10.1016/j.ssmph.2024.101726.

Wallace, M., Thomas, M.J., Aburto, J.M., Pallesen, A.V.J., Mortensen, L.H., Syse, A., and Drefahl, S. (2022). Immigration, mortality, and national life expectancy in the Nordic region, 1990–2019. *SSM – Population Health* 19: 101177. doi:0.1016/j.ssmph.2022.101177.