

Article

The Contributions of Social Factors to Lifespan Disparity: Multi-Factor Analysis and a Four-Way Case Study of U.S. Lifespan

Hal Caswell 

Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA; hcaswell@whoi.edu

Abstract

Lifespan varies among individuals, and this variation arises both from heterogeneity (e.g., in sex or race) and from unavoidable individual stochasticity. Treating a heterogeneous population as a mixture of groups, each experiencing different rates, makes it possible to partition lifespan variance into a between-group component due to heterogeneity and a within-group component due to chance. Until now, such studies could only treat factors one at a time. This paper presents an analysis that applies to multi-factor studies and quantifies the contribution of multiple factors and their interactions to variance in lifespan. As a case study it is applied to life table data from a four-factor (sex, marital status, education, and race) study of longevity in the U.S. population, 2015–2019. Even accounting for four main factors, six two-way interactions, four three-way interactions, and one four-way interaction, between-group heterogeneity accounts for only 7% to 10% of lifespan variance. Education and its interactions make the largest contribution. Main effects make the largest contributions; those of two-way, three-way, and four-way interactions are orders of magnitude smaller.

Keywords: lifespan inequality; variance partition; factorial study designs; education and lifespan

1. Introduction

In a homogeneous population, all individuals are subject to the same age schedules of mortality and fertility. Individual lifespan in such a population is a random variable whose probability distribution is given by a column, often denoted d_x , of the life table. It is computed from the probability of death, often denoted q_x and the survivorship probability, often denoted by ℓ_x . The variability among individuals in this distribution of lifespan is called individual stochasticity [1] or, with some savings of syllables, luck [2].

A heterogeneous population, in contrast, is a mixture of groups, each subject to a different set of demographic rates. Groups may be defined on the basis of measured characteristics (e.g., sex, race, education) or on the basis of latent unmeasured properties (e.g., frailty). Lifespan in such a population follows a mixture distribution in which variability arises from two sources: the differences among the groups and the individual stochasticity that operates within each group.

This variability in lifespan (or any demographic outcome) is will be referred to here as disparity. It is often referred to as an ‘inequality,’ a term that has become politically loaded in some places. Disparity in lifespan has been called the most fundamental of all inequalities because “every other type of inequality is conditional upon being alive.” [3].



Academic Editors: Jia Wang, Wei Xu, Shiro Furuya and David Lopez-Carr

Received: 12 July 2026

Revised: 29 September 2026

Accepted: 5 October 2026

Published: 12 October 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Inequality can be quantified in many ways; e.g., [4] (Chap. 2), including the variance of the distribution. When measured by the variance, disparity can be partitioned uniquely into two components, one due to differences between groups and one due to variation within groups. These variance partitioning calculations now have a considerable history in studies of longevity and of lifetime reproduction; see Caswell [5] for a review.

Caswell [5] surveys the contributions of heterogeneity (among groups) and individual stochasticity (within groups) to (1) disparity in human longevity among latent frailty classes, income groups, education groups, and neighborhood deprivation groups; (2) disparity in adult longevity among insects differing in early life nutrition; disparity in longevity and lifetime reproduction in a wild seabird population; (3) disparity in longevity and lifetime reproduction in laboratory populations of a rotifer, in which heterogeneity is defined by maternal age; (4) disparity in healthy longevity within and among European countries; and (5) disparity in longevity and lifetime reproduction within and among species of plants and animals. The conclusion from this survey, and of many studies dating back at least to Vaupel [6], is that the contribution of heterogeneity to variance in lifespan is much smaller than that of stochasticity.

However, those studies are limited because they examine only a single source of heterogeneity at a time, although longevity is clearly influenced by many social and biological factors. In this paper, I present an analysis that can incorporate multiple factors and identify the simultaneous contributions of those factors to the variance in any demographic outcome. I apply the analysis to a study of lifespan disparity in the United States from 2015 to 2019, in which heterogeneity in four interacting factors has the opportunity to contribute to the lifespan inequality.

In the next section, I present the calculations for multi-factor partitioning of disparity, as a natural extension of the now-classical single factor partitioning. Then the methodology is applied to the data of Bergeron-Boucher et al. [7] on lifespan inequality in the United States. A final discussion section considers at some length some of the issues and questions that arise in partitioning of disparity and suggests directions for future developments.

2. Variance Partitioning

Consider a heterogeneous population composed of a mixture of groups, with the representation of the groups given by a probability distribution (the mixing distribution) π . Let ξ denote some demographic outcome (here we will be concerned with longevity); the distribution of this outcome is a mixture distribution. The variance in ξ is

$$V(\xi) = E_{\pi} \left[V(\xi | \text{group}) \right] + V_{\pi} \left[E(\xi | \text{group}) \right] \quad (1)$$

$$= V_{\text{within}}(\xi) + V_{\text{between}}(\xi), \quad (2)$$

where E denotes the expected value; for example, see [8] or [9] (Section 5.6). The within-group variance V_{within} is the mean of the variances within each group, weighted by the mixing distribution π . The between-group variance V_{between} is the variance of the group means, again weighted by the distribution π . The variance ratio

$$\mathcal{K} = \frac{V_{\text{between}}}{V_{\text{between}} + V_{\text{within}}} \quad (3)$$

quantifies the proportion of the variance that is due to between-group differences.

The partition (2) has now been applied in many studies, and the general conclusion is that for many kinds of heterogeneity, the variance ratio is on the order of a few percent, even in the face of heterogeneity in important social factors.

2.1. From One Factor to Many

Studies of lifespan variation have focused on contributions of one factor at a time, using Equation (2) as a tool. For example, Luy et al. [10] reported life tables for men and women as functions of income, education, and occupation, but each factor was treated separately, for men and for women. Caswell [5] used these life tables to partition the variance in longevity into within-group and between-group components and found that $\mathcal{K}$ ranged from 0.01 to 0.05 for each of the factors separately. In a more extensive analysis, van Raalte et al. [11] analyzed lifespan inequality for males and females in eleven European countries and four educational groups, measuring disparity by both Thiel's entropy-based index and the variance in age at death. Inequality due to educational differences within each of the 22 sex-country combinations was analyzed singly and partitioned into within-group and between-group contributions. The between-group differences accounted for 0.6% to 11% of the variation in lifespan.

Treating each factor singly rules makes it impossible to see how factors combine in their effects and rules out the possibility of calculating the contributions to variance of interactions between them. The approach presented here, extending that of Caswell and van Daalen [12] makes it possible to partition the variance into contributions from the each factor and from all of their interactions.

2.2. Notation

The following notation is used throughout this paper. Matrices and arrays are denoted by upper-case bold characters (e.g., $\mathbf{U}$) and vectors are denoted by lower-case bold characters (e.g., $\mathbf{a}$). Vectors are column vectors by default; $\mathbf{x}^T$ is the transpose of $\mathbf{x}$. The vector $\mathbf{1}$ is a vector of ones, and the matrix $\mathbf{I}$ is the identity matrix. When necessary, subscripts are used to denote the size of a vector or matrix; e.g., $\mathbf{I}_\omega$ is an identity matrix of size $\omega \times \omega$. The notation $\|\mathbf{x}\|$ denotes the 1-norm of $\mathbf{x}$ (i.e., the sum of the absolute values of the entries). The symbol $\otimes$ denotes the Kronecker product. The `vec` operator stacks the columns of a $m \times n$ matrix into a $mn \times 1$ column vector. When applied to an array with more than two dimensions, the `vec` operator stacks columns from all dimensions. We will make use of a `reshape` operator (using MATLAB notation), that is, the inverse of the `vec` operator, changing the vector back into an array with specified dimensions. (In the R programming language, the `as.vector` and `as.matrix` commands implement some of this functionality).

Two element-by-element operations are important in what follows. The symbol $\circ$ denotes the Hadamard, or element-by-element product (implemented by `.*` in MATLAB and by `*` in R). The symbol $\oslash$ is used to denote the less commonly encountered element-by-element quotient. For two matrices $\mathbf{A}$ and $\mathbf{B}$,

$$\mathbf{A} \circ \mathbf{B} = \begin{pmatrix} a_{ij} b_{ij} \end{pmatrix} \quad \mathbf{A} \oslash \mathbf{B} = \begin{pmatrix} \frac{a_{ij}}{b_{ij}} \end{pmatrix} \quad (4)$$

with the obvious restrictions that $\mathbf{A}$ and $\mathbf{B}$ must be of the same size and that for the quotient, none of the entries of $\mathbf{B}$ can be zero.

2.3. Factorial Study Designs

Consider a study examining the effects of k factors, in all possible combinations, on some demographic outcome. In the example to follow, $k = 4$ and the outcome is lifespan. The study includes n_1 levels of factor 1, n_2 levels of factor 2, and so on up to n_k levels of factor k . What follows is the general calculation for any number of factors. For clarity, a simple 2×2 example (two factors, each at only two levels) is presented in Appendix A.

The means and variances of lifespan, and the mixing distribution, are contained in k -dimensional arrays, each of size $n_1 \times n_2 \cdots \times n_k$:

$$\mathbf{M} = \text{array of means} \quad (5)$$

$$\mathbf{V} = \text{array of variances} \quad (6)$$

$$\mathbf{\Pi} = \text{mixing distribution array.} \quad (7)$$

2.3.1. Marginal Mean and Mixing Arrays

The partition of variance is based on the weighted marginal distributions from $\mathbf{M}$ and $\mathbf{\Pi}$. For notational convenience, let us denote the names of the factors as A, B, C , and so on. The entry $\mathbf{M}(a, b, \dots)$ of $\mathbf{M}$ is the mean longevity in the group defined by level a of factor A , level b of factor B , and so on. Similarly for the variances $\mathbf{V}$ and the mixing distribution $\mathbf{\Pi}$. We require that

$$\sum_{a,b,\dots} \mathbf{\Pi} = 1. \quad (8)$$

The variance components are calculated from the weighted marginal distributions obtained from the arrays $\mathbf{M}$, $\mathbf{V}$, and $\mathbf{\Pi}$. Caswell and van Daalen [12] presented formulas for these marginal distributions for several special cases. To write a general formula, define a set F containing all the factors in the study. For example, suppose there are four factors, labeled A, B, C , and D . Then

$$F = \{A, B, C, D\} \quad (9)$$

Each possible combination of factors is a partition of the set F into a set S of factors included in the combination and a complementary set S^c of factors not included. For example the two-way interaction of A and C is defined by

$$S = \{A, C\} \quad \text{and} \quad S^c = \{B, D\}. \quad (10)$$

Using these sets, the marginal mean array and the marginal mixing distribution array for the interaction among the factors in the set S are

$$\mathbf{m}_S = \left(\sum_{S^c} (\mathbf{M} \circ \mathbf{\Pi}) \right) \oslash \left(\sum_{S^c} \mathbf{\Pi} \right) \quad (11)$$

$$\boldsymbol{\pi}_S = \sum_{S^c} \mathbf{\Pi}. \quad (12)$$

where $\oslash$ is the Hadamard, or element-by-element quotient. For example, the marginal AC mean and mixing distribution, with S and S^c given by (10) are

$$\mathbf{m}_{AC} = \left(\sum_b \sum_d (\mathbf{M} \circ \mathbf{\Pi}) \right) \oslash \left(\sum_b \sum_d \mathbf{\Pi} \right) \quad (13)$$

$$\boldsymbol{\pi}_{AC} = \sum_b \sum_d \mathbf{\Pi}. \quad (14)$$

2.3.2. Within-Group and Between-Group Variances

The study design contains $n_1 n_2 \cdots n_k$ factor combinations. By treating each of these combinations as a group, the analysis is reduced to a one-factor design. Then, as in Equation (2), the within-group variance is given by the expectation, over the mixing distribution, of the variances in the groups:

$$V_{\text{within}} = \left(\text{vec } \mathbf{\Pi} \right)^T \text{vec } \mathbf{V}. \quad (15)$$

The between-group variance is the variance, over the mixing distribution, of the group means. In general, we can define the variance of the entries of a vector $\mathbf{x}$ over a distribution $\boldsymbol{\pi}$ as the function

$$\mathbb{V}(\mathbf{x}, \boldsymbol{\pi}) \equiv \boldsymbol{\pi}^\top (\mathbf{x} \circ \mathbf{x}) - (\boldsymbol{\pi}^\top \mathbf{x})^2. \quad (16)$$

The first term in $\mathbb{V}(\cdot, \cdot)$ is the second moment (the expectation of the squares of the entries of $\mathbf{x}$). The second term is the square of the first moment. In terms of this function, the overall between-group variance is

$$V_{\text{between}} = \mathbb{V}(\text{vec } \mathbf{M}, \text{vec } \boldsymbol{\Pi}) \quad (17)$$

$$= (\text{vec } \boldsymbol{\Pi})^\top (\text{vec } \mathbf{M} \circ \text{vec } \mathbf{M}) - \left((\text{vec } \boldsymbol{\Pi})^\top \text{vec } \mathbf{M} \right)^2. \quad (18)$$

2.3.3. Partitioning Variance: Main Effects and Interactions

In a study including k factors, the between-group variance V_{between} in Equation (17) can be partitioned into k single-factor contributions, $\binom{k}{2}$ two-way interactions, $\binom{k}{3}$ three-way interactions, and so on up to a single k -way interaction.

The variance contribution of each of these interactions is defined in terms of the set S as

$$V_S = \mathbb{V}(\text{vec } \mathbf{m}_S, \text{vec } \boldsymbol{\pi}_S) - \sum \text{lower order interactions in } S \quad (19)$$

where $\mathbf{m}_S$ and $\boldsymbol{\pi}_S$ are the marginal mean array and marginal mixing array corresponding to the set S , calculated according to (11) and (12).

The case study to be examined here contains four factors, sex, marital status, education, and race (see Section 3), labeled A, B, C, and D. The between-group variance in lifespan is partitioned into components due to four single-factor effects (V_A, V_B, V_C, V_D), six two-way interactions ($V_{AB}, V_{AC}, V_{AD}, V_{BC}, V_{BD}, V_{CD}$) between pairs of factors, four three-way interactions ($V_{ABC}, V_{ABD}, V_{ACD}, V_{BCD}$) and one four-way interaction (V_{ABCD}).

Examining a single example of each order of interaction reveals the general pattern. The explicit formulas for all the variances are given in Appendix B.

- Single-factor variances (e.g., V_A)

$$V_A = \mathbb{V}(\mathbf{m}_A, \boldsymbol{\pi}_A) \quad (20)$$

- Two-way interactions (e.g., V_{AB})

$$\begin{aligned} V_{AB} &= \mathbb{V}(\text{vec } \mathbf{m}_{AB}, \text{vec } \boldsymbol{\pi}_{AB}) \\ &\quad - V_A - V_B \end{aligned} \quad (21)$$

- Three-way interactions (e.g., V_{ABC})

$$\begin{aligned} V_{ABC} &= \mathbb{V}(\text{vec } \mathbf{m}_{ABC}, \text{vec } \boldsymbol{\pi}_{ABC}) \\ &\quad - V_A - V_B - V_C \\ &\quad - V_{AB} - V_{AC} - V_{BC} \end{aligned} \quad (22)$$

- Four-way interaction (e.g., V_{ABCD})

$$\begin{aligned} V_{ABCD} &= \mathbb{V}(\text{vec } \mathbf{m}_{ABCD}, \text{vec } \boldsymbol{\pi}_{ABCD}) \\ &\quad - V_A - V_B - V_C - V_D \\ &\quad - V_{AB} - V_{AC} - V_{AD} - V_{BC} - V_{BD} \\ &\quad - V_{ABC} - V_{ABD} - V_{ACD} - V_{BCD} \end{aligned} \quad (23)$$

In each case the variance calculated from the marginal distributions by the function $\mathbb{V}(\cdot, \cdot)$ includes the effects of all lower-order interactions, and so those must be subtracted from $\mathbb{V}(\cdot, \cdot)$ to obtain the variance V_S .

2.4. Choosing the Mixing Distribution

As discussed in Caswell and van Daalen [12] (and as familiar from experimental design texts), it is only possible to partition variance into its components if the mixing distribution is either uniform (or “flat” in the usage here) or a rank-one, proportional distribution. A flat mixing distribution gives equal weight to every factor combination. In a rank-one distribution, each column of Π is proportional to every other column and each row is proportional to every other row. Thus the flat distribution is a special case of a rank-one distribution.

A rank-one mixing distribution is assembled from its marginals as follows. Consider the case with four factors, and let π_A , π_B , π_C , and π_D be the marginal distributions over levels of A , B , C , and D respectively. The rank one mixing distribution array with these marginals is

$$\Pi = \text{reshape}\left((\pi_D \otimes \pi_C \otimes \pi_B \otimes \pi_A), n_A, n_B, n_C, n_D\right). \quad (24)$$

The Kronecker product term creates a vector of dimension $n_A n_B n_C n_D \times 1$. The reshape operator assembles the entries of this vector into a four-dimensional array of dimension $n_A \times n_B \times n_C \times n_D$. Note the order of the terms.

Flat and rank-one mixing distributions ask, and answer, different questions. Neither is right or wrong. For more discussion, see Section 4.3.

3. A Case Study: Lifespan Inequality in the United States

Bergeron-Boucher et al. [7] presented an analysis of lifespan and mortality risk in the United States from 2015 to 2019. Their study is unique because they reported life tables for all 54 combinations of four factors; see Table 1.

Table 1. The definitions and levels of the factors in the study of Bergeron-Boucher et al. [7].

Factor	Levels
A: sex	female, male
B: marital status	married, previously married, never married
C: education	high school degree or less, some college, college degree
D: race	Black, Hispanic White, non-Hispanic White

The life tables were constructed starting at age 30 and were truncated at age 90. They were based on data from the Multiple Causes of Death Dataset of the National Center for Health Statistics and the American Community Survey. Data were pooled from 2015 to 2019; years from 2020 onward were excluded to avoid effects of the COVID-19 pandemic. For further details see Bergeron-Boucher et al. [7]. They analyzed partial life expectancy between ages 30 and 90, seeking to identify factors influencing that index, and related it to causes of death. They report the population sizes of each of the 54 groups, which we have used to construct one of the weighting distributions.

These four factors are of undoubted social and demographic significance. Their combinations make a sizable impact on life expectancy: Bergeron-Boucher et al. [7] report an 18-year difference between the factor combination yielding the lowest partial life expectancy and the combination yielding the highest partial life expectancy, a result that is replicated in our calculations.

These data provide an ideal opportunity to explore lifespan disparity by partitioning the variance in longevity into contributions from each factor and from the six two-way, four

three-way, and one four-way interaction among the factors. For the analysis here, means and variances in remaining longevity were computed from each life table using Markov chain methods. These means and variances were used in constructing the arrays **M** and **V** in Equations (5) and (6).

The variance components for all main factor effects and all interactions were computed, as in Section 2.3.3, using both the flat and a population-weighted mixing distribution. In the flat distribution $\Pi(a, b, c, d) = 1/54$. The population-weighted distribution was assembled as a rank-one distribution from the marginal population sizes over each factor, using Equation (24). Figure 1 displays these mixing distributions graphically, emphasizing the weight given to the various factor combinations, from the largest group to the smallest.

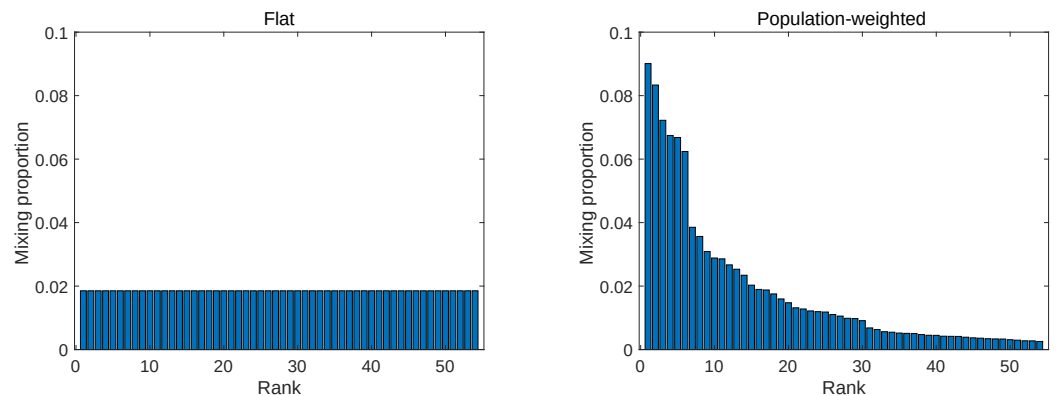


Figure 1. The representation of the 54 factor combinations, ranked from highest to lowest, in the flat and the population-weighted mixing distributions.

3.1. Components of Variance

The variance components can be presented in various ways. I present several of these here because they highlight different aspects of the same results. A complete tabulation of the variance contributions from all main effects and interactions, for both mixing distributions, is given in Table 2. The between-group variance V_{between} contributes only 11% (flat) or 8% (pop-weighted) of the total variance. The remainder is due to individual stochasticity.

Education makes the largest contribution, accounting for 4.9% (flat) or 3.2% (pop-weighted) of the total variance. Sex and marital status contribute similar amounts, with race a distant third. The contributions of interactions are smaller than the contributions of the main effects.

Interactions are of particular interest because they draw attention to the ways in which the response to one factor depends on the value of a different factor. Table 3 displays the contributions of the main effects and interactions, grouped by the order of the effects. The contributions of the two-way interactions are an order of magnitude smaller than those of the main effects. The three-way interaction contributions are an order of magnitude smaller yet.

The contributions of all factors and interactions are smaller under the population-weighted mixing distribution than under the flat mixing distribution. This reflects the greater concentration of the population-weighted mixing distribution into the most abundant of the 54 groups (Figure 1).

Even when, as here, heterogeneity creates large differences in survival and life expectancy, the contribution of that heterogeneity to lifespan disparity is small, being overwhelmed by individual stochasticity. It may be useful to focus on the importance of the factors as measured by their contributions to the between-group variance V_{between} rather than the total variance. Table 4 shows these contributions. The main effect of education con-

tributes 46% (flat) or 40% (pop-weighted) of the between-group variance. Sex and marital status each contribute about 20% and race contributes 3.7% (flat) or 5.9% (pop-weighted).

Table 2. Components of variance in remaining longevity due to sex, marital status, education, and race, and the percent of the total variance contributed by each. Starting age = 30. Results are shown for flat and for population-weighted mixing distributions.

Component	Flat Mix		Pop. Mix	
	Variance	%	Variance	%
A = sex	3.94	2.2	2.64	1.8
B = marital status	3.96	2.2	2.95	2.0
C = education	8.81	4.9	4.78	3.2
D = race	0.71	0.4	0.70	0.5
AB = sex $\times$ marital status	0.12	0.07	0.09	0.06
AC = sex $\times$ education	0.31	0.17	0.20	0.13
AD = sex $\times$ race	0.03	0.02	0.01	0.01
BC = marital status $\times$ education	0.34	0.18	0.12	0.08
BD = marital status $\times$ race	0.37	0.20	0.13	0.09
CD = education $\times$ race	0.35	0.20	0.12	0.08
ABC = sex $\times$ marital status $\times$ education	0.02	0.01	0.01	0.01
ABD = sex $\times$ marital status $\times$ race	0.04	0.02	0.04	0.02
ACD = sex $\times$ education $\times$ race	0.03	0.02	0.01	0.01
BCD = marital status $\times$ education $\times$ race	0.10	0.05	0.06	0.04
ABCD = sex $\times$ marital status $\times$ education $\times$ race	0.01	0.01	0.00	0.00
V_{between}	19.15		11.86	
V_{within} (stochasticity)	161.25		138.45	
Total	180.39		150.72	
$\mathcal{K}$	0.11		0.08	

Table 3. Components of variance in remaining longevity due to sex, marital status, education, and race. Main effects and interactions, grouped by two-way, three-way, and four-way interactions. Starting age = 30 years. Results are shown for flat and for population-weighted mixing distributions.

Component	Flat Mix Variance	Pop. Mix Variance
Main effects	17.42	11.07
Two-way interactions	1.52	0.67
Three-way interactions	0.19	0.11
Four-way interaction	0.01	0.00
V_{between}	19.15	11.86
V_{within} (stochasticity)	161.24	138.45
Total	180.39	150.72
$\mathcal{K}$	0.11	0.08

Table 4. Components of variance in longevity due to sex, marital status, education, and race. Variance components expressed as percentages of V_{between} . Results are shown for flat and for population-weighted mixing distributions.

Component	% of V_{between}	
	Flat Mix	Pop. Mix
A = sex	20.57	22.37
B = marital status	20.68	24.84
C = education	46.00	40.29
D = race	3.73	5.87
AB = sex $\times$ marital status	0.64	0.80
AC = sex $\times$ education	1.60	1.67
AD = sex $\times$ race	0.15	0.12
BC = marital status $\times$ education	1.76	1.00
BD = marital status $\times$ race	1.93	1.10
CD = education $\times$ race	1.86	0.97
ABC = sex $\times$ marital status $\times$ education	0.12	0.09
ABD = sex $\times$ marital status $\times$ race	0.22	0.30
ACD = sex $\times$ education $\times$ race	0.15	0.07
BCD = marital status $\times$ education $\times$ race	0.51	0.48
ABCD = sex $\times$ marital status $\times$ education $\times$ race	0.06	0.03

3.2. Importance of Factors: Sobol' Global Sensitivity

The importance of education as a factor can be formalized using the Sobol' global sensitivity index [13,14]. If some quantity is a function of a set of variables $x_1, x_2, \dots$, Sobol' argued for using the components of variance as a measure of the importance of the various factors [15]. The Sobol' index S of a factor is obtained by summing the variances due to the main effect of the factor and all the interactions that include that factor. In our case

$$S_A = V_A + V_{AB} + V_{AC} + V_{AD} + V_{ABC} + V_{ABD} + V_{ACD} + V_{ABCD} \quad (25)$$

$$S_B = V_B + V_{AB} + V_{BC} + V_{BD} + V_{ABC} + V_{ABD} + V_{BCD} + V_{ABCD} \quad (26)$$

$$S_C = V_C + V_{AC} + V_{BC} + V_{CD} + V_{ABC} + V_{ACD} + V_{BCD} + V_{ABCD} \quad (27)$$

$$S_D = V_D + V_{AD} + V_{BD} + V_{CD} + V_{ABD} + V_{ACD} + V_{BCD} + V_{ABCD} \quad (28)$$

Table 5 shows the Sobol' indices for each of the four factors. Under either mixing distribution, the Sobol' index for education is the largest. This supports other claims about the importance of education in determining longevity, e.g., [11,16,17]. The Sobol' indices for sex and marital status are smaller and similar to each other, showing the two factors as similarly important in determining longevity.

3.3. Inequality as a Function of Age

The variance in remaining lifetime can be calculated conditional on any starting age. Figure 2 shows the within- and between-group contributions, V_{within} and V_{between} , as a function of age. Under either mixing distribution, the total variance and the between-group variance decline with age. The between-group component never becomes larger than at age 30, which is the first age reported in the data.

Table 5. Sobol’ indices for the four factors. The index gives the contribution to variance of the factor, pooling main effects and all interactions involving that factor. Initial age = 30.

Factor	Mixing	
	Flat	Proportional
Sex	4.50	3.02
Marital status	4.96	3.39
Education	9.97	5.29
Race	1.65	1.06

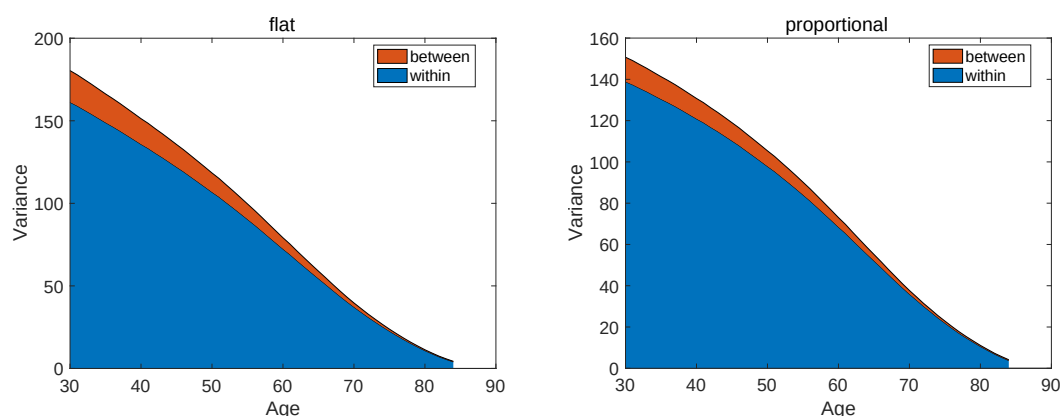


Figure 2. Within-group V_{within} and between-group V_{between} variances in remaining longevity, as a function of age, for flat and population-weighted mixing distributions.

The variance ratio K declines with the starting age of the calculation up to about 80 years, after which it increases (Figure 3). The final increase may reflect the truncation of the life tables at age 90 because, at ages approaching this maximum, V_{within} must decline towards zero.

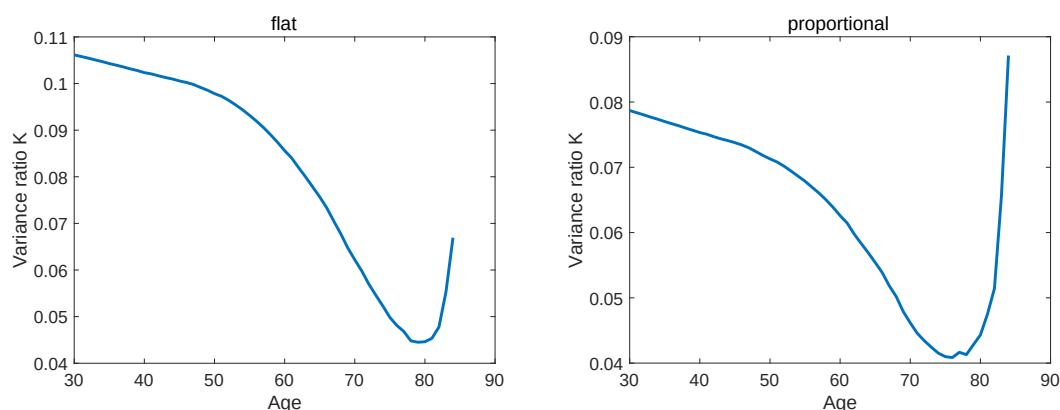


Figure 3. The variance ratio K for remaining longevity as a function of starting age, for flat and proportional mixing distributions.

4. Discussion

The ability to partition variance in multi-factor studies opens new perspectives on the contributions of social and biological factors to lifespan disparity, and how that is related to inequalities in those factors. As shown in Section 2.3.1, the components of variance due to each factor and to all of their interactions can be written in terms of the partitions of the set of factors and a function giving the variance of the entries of a vector. The terms of this

partition account for the simultaneous operation of all the factors, rather than treating each one separately.

In the present four-way factorial case study, the contributions of heterogeneity to lifespan variance are typical of those found in single-factor studies. Note that this is a novel empirical finding; it need not have been true. Even with four factors and 11 interactions, the variance ratio $\mathcal{K}$ is still small, on the order of 8–10%. About 90% or more of the variance in remaining longevity is due to stochasticity. The contributions to inequality of the factor interactions are even smaller. This is also a novel empirical result. It is not difficult to imagine interactions among sex, marital status, race, and education that might have made a large contribution to variance, but this did not happen.

Within-group variance due to stochasticity is large and, because lifespan is the result of a repeated random outcome of survival, unavoidable. This suggests that it may be interesting to evaluate factors not by their contribution to total variance, but by their contribution to the between-group variance, thus removing stochasticity from the picture. Such a comparison (Table 4) reveals the importance of education as a source of variance.

4.1. Education as a Factor

Education and interactions involving education make the biggest contributions to between-group variance (52% under a flat mixing distribution and 44% under a population-weighted mixing distribution). The Sobol' index for education is larger than for any of the other factors, under either mixing distribution (Table 5).

Education has long been identified as a factor with major influences on lifespan, disability, and health; e.g., [17–19] as part of a general interest in human capital and its development [20]. Indeed, Lutz and his colleagues have argued that “education drives demography” and that educational attainment is not merely one of many social factors, but “is the single most important source of empirically observable population heterogeneity next to age and sex” [16] (Chap. 2). The Vienna Institute of Demography now maintains a large database, the Wittgenstein Centre Human Capital Data & Graphic Explorer (<https://iiasa.ac.at/models-tools-data/wcde>), that provides mortality by age, sex, and education for 200 countries. These data have been used to produce a large set of population projections for all countries that differ from other such efforts by projecting joint education/age structures. The authors conclude that incorporating education as a factor significantly changes predictions of future global population.

Because education has obvious associations with other factors, including wealth, a number of studies have controlled for such factors in examining the role of education and concluded that even after such controls, education has the larger effect [21]. It has been argued that some of the effects of education are due to its influence on neurological development during childhood and adolescence [22,23].

Although the present paper is not concerned with fertility, it is worth noting parenthetically the well-known effects of education on fertility; e.g., [24–26].

The contribution of educational differences to lifespan inequality, while simultaneously accounting for several other important factors suggests that further multi-factor studies exploring education in the context of lifespan inequality will be particularly valuable.

4.2. Factors over Age and Time

It is worthwhile to emphasize the nature of the factors, including education. The calculations of Bergeron-Boucher et al. [7] produce period life tables. Thus, any statistics calculated from them apply to a synthetic cohort of individuals living their lives under those conditions, including the factors that determine those rates. Sex and race are, to a good approximation, fixed properties of an individual. Education is not, and

Bergeron-Boucher et al. [7] follow the practice, standard in studies of education and human capital, of beginning the life tables at an age (30 years in this case) at which education is mostly completed; e.g., [11,16,18,19]. Marital status is a more dynamic characteristic. It is operationally defined as marital status at the time of measurement, which treats it as an individual property, regardless of previous history. To quote Bergeron-Boucher and colleagues, the four factors are treated “as a snapshot of a series of profiles and [we] assume that these profiles are what an individual experiences for their whole life” [7] (p. 9). A fully dynamic analysis of marital status would require a multistate model coupling marital status to mortality. That is not available in this study. It would be awkward, and I believe unnecessary, to change the term ‘marital status’ to something more pedantic such as ‘marital status at the point of the measurement of age-specific mortality’.

On a longer time scale or a broader cultural scale, it is of course important not to reify any social factors. The differences between men and women, and the implications of those differences, have changed dramatically over historical time and across different cultures. The same is true for racial differences, for degrees of education, and for different marital statuses.

The same is true of any attempt to relate demographic results to factors of any kind. This study, and the study on which it is based [7], is based on data obtained at one time (2015–2019) in one country (the U.S.). It has nothing to say about how the variance components contributed by these factors would change at different times or in different places. (Note that the same issue arises in measurements of heritability. Heritability is defined by a partition of phenotypic variance, and is specific to the time and place of measurement [27], not an invariant property of the trait or the groups). Needless to say, applying the method introduced here in a comparative study across times and places would be a valuable contribution, but is not addressed by this study.

4.3. *Mixing Distributions*

Mixture distributions are a powerful tool for statistical analysis of heterogeneous populations; see Everitt and Hand [28] or Frühwirth-Schnatter [8] for surveys. The mixing distribution π is an indispensable component of the description of a heterogeneous population. Although mixing distributions are, in general, enormously flexible, the goal of partitioning variance in multi-factor studies imposes the restriction to flat or rank-one distributions.

The idea that one can choose among multiple possibilities for a mixing distribution deserves clarification. In a flat mixing distribution, each factor combination receives the same weight in the calculation of means and variances. This maximizes the information obtained about the effects of the factors involved. Imagine that you wanted to know how sex, marital status, education, and race, as factors, contribute to inequality in longevity. Imagine that you collect 54,000 individuals and randomly assign 1000 to each of the 54 combinations of those factors (an obvious impossibility, but a useful thought experiment). This flat mixture over the factor combinations would maximize the information about the contributions of each factor provided by the study.

Alternatively, you might sample 54,000 individuals from a population and assemble the mixing distribution from the relative abundance of the factor combinations in that sample to assemble the mixing distribution. In this calculation, some groups would have more influence than others because they are more common in the population. In the case study analyzed here, the largest group (married white males with a university degree) is 82 times larger than the smallest group (previously married Hispanic males with a university degree). If your goal is to draw conclusions about the sources of inequality in a population that has this composition, then the population-weighted mixing distribution is your preferred tool.

Branches of science in which experiments like the imaginary one described above are regular occurrences have long had to confront this issue. Kendall and Stuart [29] (pp. 25–26) describe the choices particularly clearly.

In many experimental contexts there is no question of the observations being a random sample from some population. The $r \times c$ cross-classification is deliberately set up to throw light on the variable (y) being studied. It may even be meaningless to consider any set of weights as the ‘right’ ones, in the sense of reflecting an underlying population distribution; for example, if we have a 2×3 cross-classification to study the effects of two different doses of Fertilizer A and three different doses of Fertilizer B on the yield of a crop (y), one may be simply interested in the effects and interactions as such, and not as representing any population at all.

There is a crucial distinction here between the “experimental” and the “survey” approach to data . . . In experimental investigations, therefore, it is common (for lack of any known appropriate system of weights) to use equal weights throughout. The distinction between these two weighting systems will become acute, if only because, in this most general case, we may have a very small number of very large frequencies which tend to dominate the frequency weighting system, and perhaps distort its interpretation.

The example presented by Kendall and Stuart [29] of two types of fertilizer in an agricultural experiment corresponds to our position if we view sex, marital status, education, and race as factors and want to evaluate their effects on lifespan variation. This leads naturally to the flat mixing distribution. Evaluating these effects is a different question from evaluating the role of the factors conditional on the composition of the U.S. population. Treating the results as a sample from that population leads naturally to a mixing distribution weighted by population size. The distortion due to a small number of large frequencies mentioned by Kendall and Stuart [29] corresponds to the distortion due to the 82-fold difference between the most and least abundant groups in this analysis.

It is possible to create other rank-one mixing distributions, assembled from marginal distributions as in Equation (24), based on criteria other than population size. The use of such distributions is an open research question.

In the end, the important message is that in analyzing contributions of heterogeneity and stochasticity, the mixing distribution is an analytical tool that we can deploy to answer more than a single question.

4.4. Stochasticity Is Unavoidable; We Should Embrace It

*My apologies to chance for calling it necessity,
My apologies to necessity if I'm mistaken, after all.*

Wisława Szymborska, *Under One Small Star*

The life table is a stochastic model for individual mortality. The variance in longevity within any homogeneous population is due to the individual stochasticity inherent in the life table and is calculated as part of basic survival analysis. It is tempting to imagine that some invisible hidden heterogeneity could contribute to this within-group variance and explain it as due to something other than chance. But this is impossible until that source is included as a factor, for any analysis based on calculation of means and variances of longevity from a life table or a Markov chain.

The logic is different for empirical measurement, rather than calculation, of the distribution of longevity [5]. The differences among individuals in an observed frequency distribution of age at death arise both from stochasticity and from whatever heterogeneity

exists among those individuals. Estimation of latent heterogeneity (e.g., frailty; see [30]) make it possible to partition some of that unobserved heterogeneity [31].

Stochasticity produces inequality of outcome. Heterogeneity produces inequality of opportunity because individuals in different groups experience lives governed by different sets of rates. How to respond to these types of inequalities, especially those due to chance, is a non-trivial problem [5,32]. Frank [33] surveys arguments about the relative importance of effort and talent on one hand, and luck on the other, in determining individual economic success. Conservative positions emphasize the former, liberal positions the latter. Behavioral experiments show that there is a tendency to underestimate the role of chance in success. The large contributions of stochasticity to inequality of lifespan does not mean that social, behavioral, and biological factors are not worthy of attention.

4.5. What if We Have Individual Measurements?

Although the partition of variance results in Table 2 is reminiscent of the results of statistical Analysis of Variance (ANOVA), the method used to get those results is fundamentally different. As is well known, ANOVA begins with a linear model,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon}, \quad (29)$$

where $\mathbf{y}$ is a vector of observations, $\boldsymbol{\theta}$ is an unknown vector of parameters $\mathbf{X}$ is a known design matrix, and $\boldsymbol{\epsilon}$ is a vector of random deviations. Because there are more observations than t parameters to be estimated, there is no solution to the equation, and an estimate $\hat{\boldsymbol{\theta}}$ is found by minimizing the sum of squared deviations between observations and estimates. The vector $\mathbf{y}$ contains individual observations. In our context these would be observations of individual lifespans, but the life tables that form the basis of much of demography contain no such individual measurements. Instead, life tables are computed from counts of deaths and of individuals at risk of death at each age. Treating the moments of longevity as describing a mixture distribution permits the variance decomposition without individual observations.

However, with considerable research effort it is possible to extract individual longevity data from published records; this has been done in the CenSoc project at the University of California at Berkeley [34]. That project links U.S. Census and Social Security records for the cohort of men born in 1910 and tracks them until their deaths between ages 65 and 95. The Census data provide information on a variety of social and environmental variables on these individuals. Breen and Seltzer [35] use these data to attempt to predict individual lifespan on the basis of a set of 32 variables (including detailed information on education, race, occupation, household structure, and movements). Rather than a simple linear model, they deployed an ensemble of machine learning algorithms to find a function predicting individual mortality from these variables (full details in their paper). They found statistically significant relations between the variables and individual lifespan but that relationship explained only 1.3% of the variance. As with the case study reported here, they found education-related factors to be the most important, but those explained only 0.8% of the variance in age at death.

Although detailed individual data sets (census records, registers) will yield increasingly detailed predictions of individual mortality, it seems unlikely that simply knowing more and more about everybody will eliminate the importance of individual stochasticity.

4.6. Measures of Inequality

Economists have employed a great variety of measures of economic inequality [4]. These have been applied in various demographic contexts; when applied to a single population, the results are highly correlated; e.g., [36–38].

But comparing two or more groups in a heterogeneous population requires a measure that can be partitioned into contributions that reflect that heterogeneity, being sensitive to inequality both within and between the groups. Of the indices in common use, the variance and entropy are the only indices that permit this partition. In one study that partitioned lifespan disparity using both measures, the results were very similar [11]. The standard deviation is sometimes used on the grounds that it is measured in the same units as the outcome (years, in the case of longevity). However, it cannot be partitioned. The standardized variance, defined as the ratio of the variance to the square of the mean, has been used, and it can be partitioned [39], making it preferable to the coefficient of variation, which is its square root. Its use in multi-factor study designs is still to be explored. The Gini coefficient, which is based on the mean absolute difference among individuals, is popular for studies in a single population. However, it cannot be partitioned into within-group and between-group components.

4.7. Possible Extensions

Variance partitioning can be applied to any demographic outcome for which means and variances can be calculated. In addition to longevity, other obvious candidates include healthy longevity [40], state occupancy times in multistate models [41], lifetime reproductive output [2,42,43], and the abundance of, and lifetime overlap with, various types of kin [44,45].

Extending the mixing distribution beyond the flat and the population-weighted cases provides a novel opportunity to address longevity from other perspectives. The case study here uses a weighting distribution proportional to the population sizes reported by Bergeron-Boucher et al. [7] for each of the 54 groups. Other quantities could be used. Population size of a particular age classes an obvious choice. Mixing distributions might also be defined based on the number or proportion of individuals with particular religious or political affiliations, or levels of wealth, income, or debt. Exploring these is an open research area.

Factorial study designs are those that include all possible combinations of the factors studied. Other multi-factor designs are possible, in particular hierarchical designs in which groups are composed of sub-groups, each of which is composed of sub-sub-groups, and so on. The largest demographic data sets in the world (e.g., the U.N. World Population Prospects) have such a hierarchical structure, in which geographical regions are composed of subregions which in turn are composed of countries. Variance partitioning in such studies are also possible and will be presented elsewhere (Caswell, unpublished results).

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/populations2040021/s1>. This includes the means and variances of longevity in all factor combinations, in spreadsheet form.

Funding: This research was begun under support from the European Research Council, through the European Union's Horizon 2020 research and innovation program, under ERC Advanced Grant 788195 (FORMKIN).

Institutional Review Board Statement: This study uses only anonymous statistical summary data which include no individually identifiable elements. Hence, ethical review board approval is not needed.

Informed Consent Statement: This study uses only anonymous statistical summary data that contain no individually identifiable information. Therefore, informed consent is not required.

Data Availability Statement: The original contributions presented in this study are included in the article and Supplementary Material. Further inquiries can be directed to the corresponding author.

Acknowledgments: I am happy to thank the late John L. Gill of Michigan State University. I also thank Marie-Pier Bergeron-Boucher for generously providing the life table data used in these calculations.

Conflicts of Interest: The author declares no conflicts of interest.

Appendix A. A Simple 2×2 Example

It may be helpful to see a simple worked example of the variance partitioning, with two factors and two levels of each factor. This section presents such an example. Note that it is intended only for illustration and not as a substantive exploration of the factors.

The example uses life tables for the U.S. in 2023 [46]. The design here has two levels of factor A, sex (1 = female, 2 = male) and two levels of factor B, race (1 = white, 2 = black). The arrays of means and variances of longevity at birth are

$$\mathbf{M} = \begin{pmatrix} 81.4 & 78.0 \\ 76.5 & 70.8 \end{pmatrix} \quad \mathbf{V} = \begin{pmatrix} 234.0 & 338.0 \\ 296.2 & 406.0 \end{pmatrix} \quad (\text{A1})$$

Appendix A.1. Flat Mixing Distribution

The flat mixing distribution applies equal weight to all factor combinations,

$$\mathbf{\Pi} = \begin{pmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{pmatrix} \quad (\text{A2})$$

The within-group variance is the mean, over the mixing distribution, of the variances in the four groups,

$$V_{\text{within}} = \text{vec}(\mathbf{\Pi})^T \text{vec} \mathbf{V} \quad (\text{A3})$$

$$= \begin{pmatrix} 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix} \begin{pmatrix} 234.0 \\ 338.0 \\ 296.2 \\ 406.0 \end{pmatrix} \quad (\text{A4})$$

$$= 318.6 \quad (\text{A5})$$

The between-group variance is the variance, over the mixing distribution, of the entries of $\mathbf{M}$,

$$V_{\text{between}} = \mathbb{V}(\text{vec} \mathbf{M}, \text{vec} \mathbf{\Pi}) \quad (\text{A6})$$

$$= \begin{pmatrix} 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix} \left[\begin{pmatrix} 81.4 \\ 78.0 \\ 76.5 \\ 70.8 \end{pmatrix} \circ \begin{pmatrix} 81.4 \\ 78.0 \\ 76.5 \\ 70.8 \end{pmatrix} \right] \quad (\text{A7})$$

$$= \left[\begin{pmatrix} 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix} \begin{pmatrix} 81.4 \\ 78.0 \\ 76.5 \\ 70.8 \end{pmatrix} \right]^2 \quad (\text{A8})$$

$$= 14.7 \quad (\text{A9})$$

where $\circ$ denotes the element-by-element product of the two vectors. The variance ratio is

$$\mathcal{K} = \frac{14.7}{318.6 + 14.7} = 0.044. \quad (\text{A10})$$

That is, 4.4% of the variance in longevity is accounted for by heterogeneity between sexes and races.

To partition the V_{between} into its components, we need the marginal means and marginal mixing distributions for sex and race, calculated using Equations (11) and (12),

$$\mathbf{m}_A = \begin{pmatrix} 79.8 \\ 73.6 \end{pmatrix} \quad \mathbf{m}_B = \begin{pmatrix} 79.0 \\ 74.4 \end{pmatrix} \quad (\text{A11})$$

$$\boldsymbol{\pi}_A = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \quad \boldsymbol{\pi}_B = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \quad (\text{A12})$$

main effect variances

$$V_A = \mathbb{V}(\mathbf{m}_A, \boldsymbol{\pi}_A) = 9.24 \quad (\text{A13})$$

$$V_B = \mathbb{V}(\mathbf{m}_B, \boldsymbol{\pi}_B) = 5.16 \quad (\text{A14})$$

$$V_{AB} = \mathbb{V}(\text{vec } \mathbf{M}, \text{vec } \boldsymbol{\Pi}) - V_A - V_B \quad (\text{A15})$$

$$= 14.7 - 9.24 - 5.16 \quad (\text{A16})$$

$$= 0.35 \quad (\text{A17})$$

In this simple example, sex contributes almost twice as much to lifespan inequality as race, and the sex $\times$ race interaction is an order of magnitude smaller.

Appendix A.2. Population-Weighted Mixing Distribution

Now we consider the variance components accounting for the different sizes of the female and male, white and black populations. We obtain an array of 2023 population sizes from U.S. Census Bureau data (<https://www.census.gov/data/datasets/time-series/demo/popest/2020s-national-detail.html>, accessed on 12 July 2026). Since only relative sizes matter for the mixing distribution, this array is normalized to sum to 1.

$$P = \begin{pmatrix} 124,310,110 & 23,712,829 \\ 121,353,289 & 21,644,013 \end{pmatrix} \longrightarrow \begin{pmatrix} 0.43 & 0.08 \\ 0.42 & 0.07 \end{pmatrix} \quad (\text{A18})$$

This array is not of rank-one, so we compute a rank-one approximation from its marginals,

$$\boldsymbol{\pi}_A = \begin{pmatrix} 0.51 \\ 0.49 \end{pmatrix} \quad \boldsymbol{\pi}_B = \begin{pmatrix} 0.84 \\ 0.16 \end{pmatrix} \quad (\text{A19})$$

The rank-one mixing distribution from these marginals is

$$\boldsymbol{\Pi} = \text{reshape}((\boldsymbol{\pi}_B \otimes \boldsymbol{\pi}_A), 2, 2) \quad (\text{A20})$$

$$= \begin{pmatrix} 0.429 & 0.079 \\ 0.415 & 0.077 \end{pmatrix} \quad (\text{A21})$$

The rows and columns of this distribution are proportional; it has marginals

$$\boldsymbol{\pi}_A = \begin{pmatrix} 0.51 \\ 0.49 \end{pmatrix} \quad \boldsymbol{\pi}_B = \begin{pmatrix} 0.84 \\ 0.16 \end{pmatrix} \quad (\text{A22})$$

That is, about 51% female, 49% male; 84% white, 16% black.

The choice of mixing distribution does not affect the arrays $\mathbf{M}$ and $\mathbf{V}$ of means and variances. The within-group and between-group variances are thus

$$V_{\text{within}} = \text{vec}(\mathbf{\Pi})^\top \text{vec} \mathbf{V} \quad (\text{A23})$$

$$= \begin{pmatrix} 0.43 & 0.44 & 0.08 & 0.08 \end{pmatrix} \begin{pmatrix} 234.0 \\ 338.0 \\ 296.2 \\ 406.0 \end{pmatrix} \quad (\text{A24})$$

$$= 281.3 \quad (\text{A25})$$

and

$$V_{\text{between}} = \mathbb{V}(\text{vec} \mathbf{M}, \text{vec} \mathbf{\Pi}) \quad (\text{A26})$$

$$= \begin{pmatrix} 0.43 & 0.44 & 0.08 & 0.08 \end{pmatrix} \left[\begin{pmatrix} 81.4 \\ 78.0 \\ 76.5 \\ 70.8 \end{pmatrix} \circ \begin{pmatrix} 81.4 \\ 78.0 \\ 76.5 \\ 70.8 \end{pmatrix} \right] \quad (\text{A27})$$

$$- \left[\begin{pmatrix} 0.43 & 0.44 & 0.08 & 0.08 \end{pmatrix} \begin{pmatrix} 81.4 \\ 78.0 \\ 76.5 \\ 70.8 \end{pmatrix} \right]^2 \quad (\text{A28})$$

$$= 9.80 \quad (\text{A29})$$

The variance ratio is

$$\mathcal{K} = \frac{9.80}{281.3 + 9.80} = 0.034. \quad (\text{A30})$$

That is, 3.4% of the variance in longevity is accounted for by differences between sexes and races.

To partition the between-group variance, we apply the population-weighted mixing distribution to the marginal means.

$$V_A = \mathbb{V}(\mathbf{m}_A, \boldsymbol{\pi}_A) = 6.93 \quad (\text{A31})$$

$$V_B = \mathbb{V}(\mathbf{m}_B, \boldsymbol{\pi}_B) = 2.69 \quad (\text{A32})$$

$$V_{AB} = \mathbb{V}(\text{vec} \mathbf{M}, \text{vec} \mathbf{\Pi}) - V_A - V_B \quad (\text{A33})$$

$$= 9.80 - 6.93 - 2.69 \quad (\text{A34})$$

$$= 0.35 \quad (\text{A35})$$

Overall, variance is reduced when using the population-weighted mixing distribution, because more of the distribution is concentrated on white females and males, and less on black females and males. Once again, race accounts for less of the variance than sex, and the sex $\times$ race interaction is an order of magnitude smaller.

Appendix B. Complete Formulas for Four-Way Variance Components

The four factors are labelled A, B, C , and D . The between-group variance is partitioned into four single factor main effects, six two-way interactions, four three-way interactions, and one four-way interaction.

- Single factor contributions:

$$V_A = \mathbb{V}(\mathbf{m}_A, \boldsymbol{\pi}_A) \quad (\text{A36})$$

$$V_B = \mathbb{V}(\mathbf{m}_B, \boldsymbol{\pi}_B) \quad (\text{A37})$$

$$V_C = \mathbb{V}(\mathbf{m}_C, \boldsymbol{\pi}_C) \quad (\text{A38})$$

$$V_D = \mathbb{V}(\mathbf{m}_D, \boldsymbol{\pi}_D) \quad (\text{A39})$$

- Two-way interactions:

$$V_{AB} = \mathbb{V}(\text{vec } \mathbf{m}_{AB}, \text{vec } \boldsymbol{\pi}_{AB}) - V_A - V_B \quad (\text{A40})$$

$$V_{AC} = \mathbb{V}(\text{vec } \mathbf{m}_{AC}, \text{vec } \boldsymbol{\pi}_{AC}) - V_A - V_C \quad (\text{A41})$$

$$V_{AD} = \mathbb{V}(\text{vec } \mathbf{m}_{AD}, \text{vec } \boldsymbol{\pi}_{AD}) - V_A - V_D \quad (\text{A42})$$

$$V_{BC} = \mathbb{V}(\text{vec } \mathbf{m}_{BC}, \text{vec } \boldsymbol{\pi}_{BC}) - V_B - V_C \quad (\text{A43})$$

$$V_{BD} = \mathbb{V}(\text{vec } \mathbf{m}_{BD}, \text{vec } \boldsymbol{\pi}_{BD}) - V_B - V_D \quad (\text{A44})$$

$$V_{CD} = \mathbb{V}(\text{vec } \mathbf{m}_{CD}, \text{vec } \boldsymbol{\pi}_{CD}) - V_C - V_D \quad (\text{A45})$$

- Three-way interactions:

$$V_{ABC} = \mathbb{V}(\text{vec } \mathbf{m}_{ABC}, \text{vec } \boldsymbol{\pi}_{ABC}) - V_A - V_B - V_C - V_{AB} - V_{AC} - V_{BC} \quad (\text{A46})$$

$$V_{ABD} = \mathbb{V}(\text{vec } \mathbf{m}_{ABD}, \text{vec } \boldsymbol{\pi}_{ABD}) - V_A - V_B - V_D - V_{AB} - V_{AD} - V_{BD} \quad (\text{A47})$$

$$V_{ACD} = \mathbb{V}(\text{vec } \mathbf{m}_{ACD}, \text{vec } \boldsymbol{\pi}_{ACD}) - V_A - V_C - V_D - V_{AC} - V_{AD} - V_{CD} \quad (\text{A48})$$

$$V_{BCD} = \mathbb{V}(\text{vec } \mathbf{m}_{BCD}, \text{vec } \boldsymbol{\pi}_{BCD}) - V_B - V_C - V_D - V_{BC} - V_{BD} - V_{CD} \quad (\text{A49})$$

- The four-way interaction:

$$V_{ABCD} = \mathbb{V}(\text{vec } \mathbf{m}_{ABCD}, \text{vec } \boldsymbol{\pi}_{ABCD}) - \text{all lower-order interactions} \quad (\text{A50})$$

References

1. Caswell, H. Stage, age and individual stochasticity in demography. *Oikos* **2009**, *118*, 1763–1782. [\[CrossRef\]](#)
2. Snyder, R.E.; Ellner, S.P. Pluck or luck: Does trait variation or chance drive variation in lifetime reproductive success? *Am. Nat.* **2018**, *191*, E90–E107. [\[CrossRef\]](#) [\[PubMed\]](#)
3. van Raalte, A.A.; Sasson, I.; Martikainen, P. The case for monitoring life-span inequality. *Science* **2018**, *362*, 1002–1004. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Sen, A. *On Economic Inequality*, 2nd ed.; Oxford University Press: Oxford, UK, 1997.
5. Caswell, H. The contributions of stochasticity and social inequality to lifespan variability. *Demogr. Res.* **2023**, *49*, 309–354. [\[CrossRef\]](#)
6. Vaupel, J.W. Inherited frailty and longevity. *Demography* **1988**, *25*, 277–287. [\[CrossRef\]](#)
7. Bergeron-Boucher, M.P.; Calloway, J.; Strozza, C.; Oeppen, J. Inequalities in lifespan and mortality risk in the US: A cross-sectional analysis of subpopulations by social determinants of health. *BMJ Open* **2024**, *14*, e079534. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Frühwirth-Schnatter, S. *Finite Mixture and Markov Switching Models*; Springer: New York, NY, USA, 2006.
9. Rényi, A. *Probability Theory*; North-Holland: Amsterdam, The Netherlands, 1970.
10. Luy, M.; Wegner-Siegmundt, C.; Wiedemann, A.; Spijker, J. Life expectancy by education, income and occupation in Germany: estimations using the longitudinal survival method. *Comp. Popul. Stud.* **2015**, *40*, 339–436. [\[CrossRef\]](#)
11. van Raalte, A.A.; Kunst, A.E.; Lundberg, O.; Leinsalu, M.; Martikainen, P.; Artnik, B.; Deboosere, P.; Stirbu, I.; Wojtyniak, B.; Mackenbach, J.P. The contribution of educational inequalities to lifespan variation. *Popul. Health Metrics* **2012**, *10*, 1–10. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Caswell, H.; van Daalen, S.F. Inequality, heterogeneity, and chance: Multiple factors and their interactions. *Vienna Yearb. Popul. Res.* **2025**, *23*, 1–27. [\[CrossRef\]](#)

13. Sobol', I.M. Sensitivity estimates for nonlinear mathematical models (Russian). English translation: 1993, Mathematical Modeling and Computational Experiments 1:407–414. *Matem Model.* **1990**, *2*, 112–118.
14. Sobol', I.M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* **2001**, *55*, 271–280. [[CrossRef](#)]
15. Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. *Global Sensitivity Analysis: The Primer*; John Wiley & Sons: New York, NY, USA, 2008.
16. Lutz, W.; Butz, W.P.; KC, S. *World Population and Human Capital in the Twenty-First Century*; Oxford University Press: Oxford, UK, 2014.
17. Lutz, W. *Advanced Introduction to Demography*; Edward Elgar Publishing: Cheltenham, UK, 2021.
18. Lutz, W.; Cuaresma, J.C.; Sanderson, W. The demography of educational attainment and economic growth. *Science* **2008**, *319*, 1047–1048. [[CrossRef](#)] [[PubMed](#)]
19. KC, S.; Lentzner, H. The effect of education on adult mortality and disability: A global perspective. *Vienna Yearb. Popul. Res.* **2010**, *8*, 201–235. [[CrossRef](#)]
20. Lutz, W.; KC, S. Global human capital: Integrating education and population. *Science* **2011**, *333*, 587–592. [[CrossRef](#)] [[PubMed](#)]
21. Lutz, W.; Kebede, E. Education and health: Redrawing the Preston curve. *Popul. Dev. Rev.* **2018**, *44*, 343. [[CrossRef](#)] [[PubMed](#)]
22. Baker, D.P.; Leon, J.; Smith Greenaway, E.G.; Collins, J.; Movit, M. The education effect on population health: A reassessment. *Popul. Dev. Rev.* **2011**, *37*, 307–332. [[CrossRef](#)] [[PubMed](#)]
23. Baker, D.P.; Salinas, D.; Eslinger, P.J. An envisioned bridge: Schooling as a neurocognitive developmental institution. *Dev. Cogn. Neurosci.* **2012**, *2*, S6–S17. [[CrossRef](#)] [[PubMed](#)]
24. Durowaa-Boateng, A.; Yildiz, D.; Goujon, A. A Bayesian model for the reconstruction of education-and age-specific fertility rates. *Demogr. Res.* **2023**, *49*, 809–848. [[CrossRef](#)]
25. Bora, J.K.; Saikia, N.; Kebede, E.B.; Lutz, W. Revisiting the causes of fertility decline in Bangladesh: The relative importance of female education and family planning programs. *Asian Popul. Stud.* **2023**, *19*, 81–104. [[CrossRef](#)]
26. Skirbekk, V. *Decline and Prosper!: Changing Global Birth Rates and the Advantages of Fewer Children*; Palgrave Macmillan, Springer Nature: Cham, Switzerland, 2022.
27. Lewontin, R.C. The analysis of variance and the analysis of causes. *Am. J. Hum. Genet.* **1974**, *26*, 400–411. [[PubMed](#)]
28. Everitt, B.; Hand, D. *Finite Mixture Distributions*; Chapman and Hall: London, UK, 1981.
29. Kendall, M.G.; Stuart, A. *The Advanced Theory of Statistics, Volume 3. Design and Analysis, and Time Series*, 3rd ed.; Charles Griffin: London, UK, 1976.
30. Vaupel, J.W.; Missov, T.I. Unobserved population heterogeneity: A review of formal relationships. *Demogr. Res.* **2014**, *31*, 659–688.
31. Hartemink, N.; Missov, T.I.; Caswell, H. Stochasticity, heterogeneity, and variance in longevity in human populations. *Theor. Popul. Biol.* **2017**, *114*, 107–117. [[CrossRef](#)] [[PubMed](#)]
32. Anderson, E.S. What is the point of equality? *Ethics* **1999**, *109*, 287–337. [[CrossRef](#)]
33. Frank, R.H. *Success and Luck: Good Fortune and the Myth of Meritocracy*; Princeton University Press: Princeton, NJ, USA, 2016.
34. Breen, C.F.; Osborne, M.; Goldstein, J.R. CenSoc: Public linked administrative mortality records for individual-level research. *Sci. Data* **2023**, *10*, 802. [[CrossRef](#)] [[PubMed](#)]
35. Breen, C.F.; Seltzer, N. Structured inequality, uncertain lifespans: Demographic perspectives on predicting individual-level longevity. *Popul. Dev. Rev.* **2026**, *52*, 299–318. [[CrossRef](#)] [[PubMed](#)]
36. Wilmoth, J.R.; Horiuchi, S. Rectangularization revisited: Variability of age at death within human populations. *Demography* **1999**, *36*, 475–495. [[CrossRef](#)]
37. Anand, S.; Nanthikesan, S. *A Compilation of Length-of-Life Distribution Measures for Complete Life Tables*; Working Paper Volume 10 7; Harvard Center for Population and Development Studies, Harvard School of Public Health: Boston, MA, USA, 2000.
38. Shkolnikov, V.M.; Andreev, E.E.; Begun, A.Z. Gini coefficient as a life table function: Computation from discrete data, decomposition of differences and empirical examples. *Demogr. Res.* **2003**, *8*, 305–358.
39. Rosenbluth, G. Note on Mr. Schutz's measure of income inequality. *Am. Econ. Rev.* **1951**, *41*, 935–937.
40. Zarulli, V.; Caswell, H. Longer healthy life, but for how many? A stochastic analysis of healthy lifespan inequality. *Ann. Oper. Res.* **2024**, *363*, 1161–1178. [[CrossRef](#)] [[PubMed](#)]
41. Caswell, H.; van Daalen, S.F. Healthy longevity from incidence-based models: More kinds of health than stars in the sky. *Demogr. Res.* **2021**, *45*, 397–452. [[CrossRef](#)]
42. van Daalen, S.F.; Caswell, H. Lifetime reproductive output: Individual stochasticity, variance, and sensitivity analysis. *Theor. Ecol.* **2017**, *10*, 355–374. [[CrossRef](#)] [[PubMed](#)]
43. Tuljapurkar, S.; Zuo, W.; Coulson, T.; Horvitz, C.; Gaillard, J.M. Skewed distributions of lifetime reproductive success: Beyond mean and variance. *Ecol. Lett.* **2020**, *23*, 748–756. [[CrossRef](#)] [[PubMed](#)]

44. Caswell, H. The formal demography of kinship VI: Demographic stochasticity and variance in the kinship network. *Demogr. Res.* **2024**, *51*, 1201–1256. [[CrossRef](#)]
45. Caswell, H.; de Vries, C. The formal demography of kinship. VII. Lifetime kin overlap within and across generations. *Demogr. Res.* **2025**, *53*, 123–174. [[CrossRef](#)]
46. Arias, E.; Xu, J.; Kochanek, M.A. United States Life Tables, 2023. *Natl. Vital Stat. Rep.* **2025**, *74*, 1–63. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.