Artículo de revista:

Micheletti, Chiara & Villavicencio, Francisco (2024). "On the relationship between life expectancy, modal age at death, and the threshold age of the life table entropy". *Demographic Research*, 51(24). pp. 763-788 (ISSN 1435-9871) https://doi.org/10.4054/DemRes.2024.51.24

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

DEMOGRAPHIC RESEARCH

VOLUME 51, ARTICLE 24, PAGES 763–788
PUBLISHED 4 OCTOBER 2024
http://www.demographic-research.org/Volumes/Vol51/24/
DOI: 10.4054/DemRes.2024.51.24

Formal Relationship

On the relationship between life expectancy, modal age at death, and the threshold age of the life table entropy

Chiara Micheletti

Francisco Villavicencio

This publication is part of the Special Collection in the Memory of Professor James W Vaupel (1945–2022), founder and long-time publisher of Demographic Research. The Special Collection is edited by Jakub Bijak, Griffith Feeney, Nico Keilman, and Carl Schmertmann.

Contents

On the relationship between life expectancy, modal age at death, and the threshold age of the life table entropy

Chiara Micheletti[1]

Francisco Villavicencio[2]

Abstract

BACKGROUND
Indicators of longevity like the life expectancy at birth or the modal age at death are always positively affected by improvements in mortality. Instead, for lifespan variation it has been shown that there exists a threshold age above and below which averting deaths respectively increases or decreases such variation.

OBJECTIVE
Within a Gompertz force of mortality setting, we aim to provide approximations of the life expectancy at birth and the threshold age of the life table entropy in terms of the modal age at death, highlighting the interrelationships holding among the three.

RESULTS
In the Gompertz framework, a tight relationship exists between the life expectancy at birth, the threshold age of the life table entropy, and the modal age at death, with the former two moving together and in parallel to the latter. We apply this theoretical result to life table data from the Human Mortality Database to show how the different relationships evolve over time. We observe a remarkable association between the modal and the threshold ages, even in populations with high mortality levels.

CONTRIBUTION
We provide approximations of the life expectancy at birth and the threshold age of the life table entropy in terms of the Gompertz modal age at death. This is a mathematical demography paper that builds upon previous research by James W. Vaupel and illustrates the beauty – and oftentimes simplicity – of the mathematical relationships between demographic concepts.

[1] Centre for Demographic Studies (CED-CERCA), Autonomous University of Barcelona, Bellaterra, Spain, and Max Planck Institute for Demographic Research, Rostock, Germany.
[2] Department of Economic, Financial and Actuarial Mathematics, University of Barcelona, Barcelona, Spain. Email: villavicencio@ub.edu.

1. Introduction

The analysis of the distribution of random variables usually aims attention at three main indicators: the mean, the mode, and the variance. In mortality studies that analyse the distribution of ages at death, this translates into life expectancy (mean), modal age at death (mode), and lifespan variation (variance). Several metrics have been proposed to measure lifespan variation, such as the variance of the age-at-death distribution (Edwards and Tuljapurkar 2005; Gillespie, Trotter, and Tuljapurkar 2014), life disparity (Vaupel and Canudas-Romo 2003; Vaupel, Zhang, and van Raalte 2011), the Theil index (Smits and Monden 2009), the Gini coefficient of the life table (Hanada 1983; Shkolnikov, Andreev, and Begun 2003), or Drewnowski's index (Aburto et al. 2022).

Here we focus on the life table entropy, a dimensionless indicator of the relative variation in the length of life compared to life expectancy at birth. It was first defined by Leser (1955), and further developed and studied by Keyfitz (1977), Demetrius (1978), Goldman and Lord (1986), and Vaupel (1986), among others. Following Keyfitz (1977), the life table entropy can be defined as

$$\overline{H} = -\frac{\int_0^\infty \ell(a)\, \log \ell(a)\, da}{\int_0^\infty \ell(a)\, da}\,,$$

where $\ell(x)$ is the survival probability from birth to age x. Note that the denominator is the life expectancy at birth expressed in terms of the survival function. The life table entropy can be interpreted as a weighted average of the cumulative hazard $H(x) = -\log \ell(x)$, which justifies the notation $\overline{H}$.

Aburto et al. (2019) prove that if mortality improvements over time occur at all ages, there exists a unique threshold age that separates positive from negative contributions to $\overline{H}$: Mortality improvements below the threshold age reduce lifespan variation, whereas mortality improvements above the threshold increase lifespan inequality. This threshold is what Zhang and Vaupel (2009) identify as the age separating early from late deaths in a population.

In this paper, we study the relationship between the life expectancy at birth (e_o), the modal age at death (M), and the threshold age of the life table entropy (a^H). Assuming the risk of death increases exponentially over age following a Gompertz mortality model (Gompertz 1825), we give closed-form approximations for e_o and a^H in terms of M. To formally prove our results, we adopt a hybrid approach, combining both analytical and numerical tools. We provide an empirical application to illustrate the remarkable association between M and a^H, even in populations with high levels of infant, child, and young adult mortality or that are affected by mortality shocks.

2. Relationships

For any given age $x \geq 0$, let $\mu(x) = \beta\, \mathrm{e}^{\beta(x-M)}$ be the force of mortality (mortality hazard or risk of death) of the Gompertz model in terms of the modal age at death $M > 0$ and the rate of ageing $\beta > 0$ (Missov et al. 2015). The corresponding life expectancy at birth can be estimated as

$$e_o \approx M - \frac{\gamma}{\beta}\ , \tag{1}$$

where $\gamma \approx 0.5772157$ is the Euler-Mascheroni constant.

Let $\overline{H}$ denote the life table entropy as defied by Keyfitz (1977). Aburto et al. (2019) prove that if mortality improvements over time occur at all ages, there exists a unique threshold age a^H that separates positive from negative contributions to $\overline{H}$ resulting from those improvements. We postulate that under the Gompertz model,

$$a^H \approx M - \frac{\gamma}{\beta}\ , \tag{2a}$$

and therefore

$$a^H \approx e_o\ . \tag{2b}$$

3. Proof

Proof of (1)

Castellares et al. (2020) show that the remaining life expectancy at age x of the Gompertz mortality model with hazard $\mu(x)$ and rate of ageing $\beta > 0$ is given by

$$e(x) = \frac{1}{\beta}\, \mathrm{e}^{z}\, E_1(z)\ , \tag{3}$$

where $z = \mu(x)\,/\,\beta$ and $E_1(z) = \int_z^\infty \frac{\mathrm{e}^{-t}}{t}\, dt$ is the exponential integral. According to Abramowitz and Stegun (1964: 229), for any given $z \in \mathbb{C}$ such that $|\mathrm{Arg}(z)| < \pi$,

$$E_1(z) = -\gamma - \log(z) - \sum_{k=1}^{\infty} \frac{(-1)^k\, z^k}{k \cdot k!}\ . \tag{4}$$

Let us consider the Gompertz force of mortality in terms of the modal age at death $\mu(x) = \beta\, \mathrm{e}^{\beta(x-M)}$. Applying (3) and (4), we obtain the following approximation of the life expectancy at birth:

$$
\begin{aligned}
e_o \coloneqq e(0) &= \frac{1}{\beta}\, \mathrm{e}^{\mathrm{e}^{-\beta\, M}}\, E_1\left(\mathrm{e}^{-\beta\, M}\right) \\
&= \frac{1}{\beta}\, \mathrm{e}^{\mathrm{e}^{-\beta\, M}} \left(-\gamma + \beta\, M - \sum_{k=1}^{\infty} \frac{(-1)^k\, \mathrm{e}^{-\beta\, M\, k}}{k \cdot k!}\right) . \qquad (5)
\end{aligned}
$$

Using data from the Human Mortality Database (HMD), Missov et al. (2015) estimate values of the Gompertz parameters β and M for several populations. The lowest reported values are $\beta = 0.061$ (Russian males 2000–2009) and $M = 66.06$ (Swedish males 1800–1809). Notice that $\exp\left\{\mathrm{e}^{-0.061 \cdot 66.06}\right\} \approx 1.0179$, and that for any higher values of β and M the term $\exp\left\{\mathrm{e}^{-\beta\, M}\right\}$ in (5) gets closer to 1. On the other hand, the series $\sum_{k=1}^{\infty} \frac{(-1)^k\, z^k}{k \cdot k!}$ is convergent. Setting $z = \exp\left\{\mathrm{e}^{-0.061 \cdot 66.06}\right\}$ one can find numerically that

$$
\sum_{k=1}^{\infty} \frac{(-1)^k\, \mathrm{e}^{-0.061 \cdot 66.06 \cdot k}}{k \cdot k!} \approx -0.0177 .
$$

Similarly, the summation gets closer to 0 for higher values of β and M.[3] With these two approximations, we can infer that in human populations and under the Gompertz model

$$
e_o \approx \frac{1}{\beta}\left(-\gamma + \beta\, M\right) = M - \frac{\gamma}{\beta} ,
$$

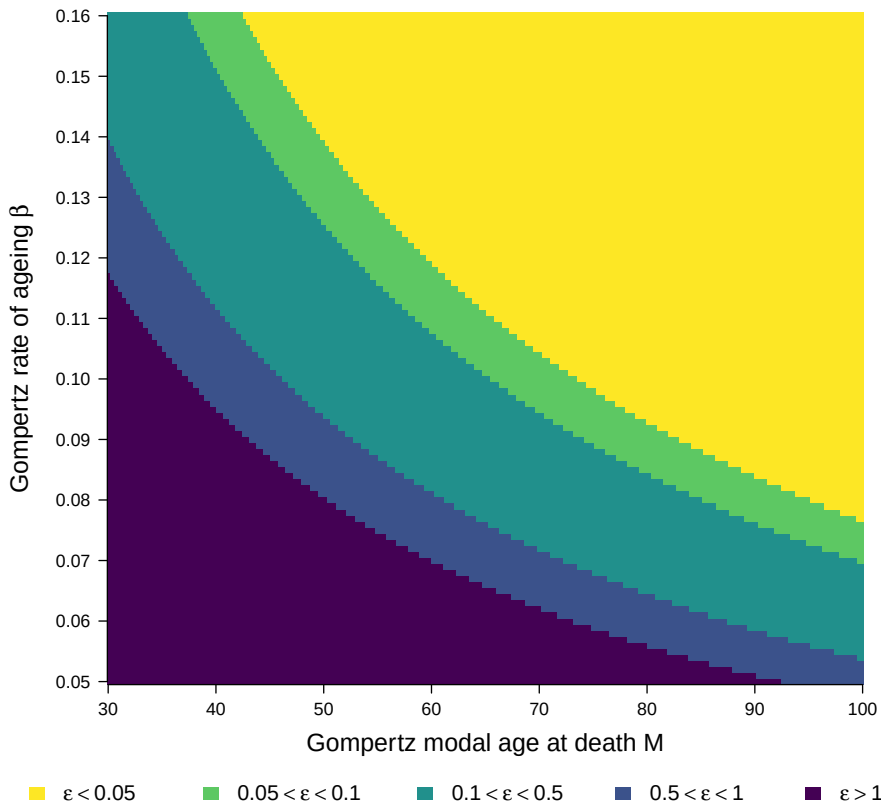
which proves (1). □

Adopting the Gompertz parameters estimated by Missov et al. (2015) as a reference, we carried out an additional sensitivity analysis by applying (5) to calculate e_o for multiple combinations of $\beta \in [0.05, 0.16]$ and $M \in [30, 100]$. Although values of $M < 60$ years are unlikely in contemporary human populations (Missov et al. 2015), we were interested in testing the validity of (1) over a wide range of values of β and M – and therefore a broad spectrum of mortality regimes. For instance, a combination of $\beta = 0.05$ and $M = 30$ corresponds to a life expectancy at birth below 30 years, whereas

[3] When estimating the summation in (5), we included terms from $k = 1$ to $k = 150$, which was tested to be sufficient to validate the convergence of the series. The same criterion was used when calculating $e(x)$ using (5) in other sections of the manuscript.

at the other end, a combination of $\beta = 0.16$ and $M = 100$ gives $e_o > 95$. The results of the sensitivity analysis are presented in Figure 1.

Figure 1: **Sensitivity analysis of** (1)**. Estimates of the error term** $\varepsilon \coloneqq |e_o - M + \frac{\gamma}{\beta}|$ **for combinations of** $\beta \in [0.05, 0.16]$ **and** $M \in [30, 100]$ **in a Gompertz setting**



We observe that the error term $\varepsilon \coloneqq |e_o - M + \frac{\gamma}{\beta}| < 0.5$ years in 70% of the cases, while $\varepsilon > 1$ about 19% of the time. The latter corresponds to combinations of β and M for which $\exp\left\{e^{-\beta M}\right\} > 1$ and the summation in (5) is not close to zero (darker area in the lower-left corner of Figure 1). Therefore, in these cases these two terms cannot be cancelled, and the approximation in (1) does not hold. For combinations of $\beta \geq 0.08$ and $M \geq 60$, $\varepsilon < 0.1$ years in 90% of the cases and $\varepsilon < 0.54$ always, so the maximum

error is just over half a year. Notice that $\beta = 0.08$ and $M = 60$ correspond to a life expectancy at birth of $e_o = 53.3$ years, and that e_o increases for higher values of β and M. The United Nations World Population Prospects (UN WPP) report very few records with such a small life expectancy at birth. Since 2010, the lowest reported value for both sexes combined is $e_o = 45.6$ years (Lesotho in 2010) and $e_o > 53.3$ in 98.3% of all country-years (United Nations 2022). Hence, we can conclude that the approximation suggested in (1) is accurate for the vast majority of contemporary human populations.

Proof of (2a) **and** (2b)

Aburto et al. (2019) show that in a Gompertz setting the threshold age a^H that separates positive from negative contributions to lifespan variation as measured by the life table entropy $\overline{H}$ is reached when

$$e(x) = \frac{e_o}{\beta e_o + 1} \,. \tag{6}$$

Let us define the function $\phi(z) := \mathrm{e}^z \, E_1(z)$. Using (3), the remaining life expectancy at age x of the Gompertz model can be expressed as $e(x) = \frac{1}{\beta}\,\phi(z)$. For all positive real numbers, $\phi(z)$ is a one-to-one function, strictly decreasing, continuous, and differentiable. Therefore, it has an inverse, here denoted by $\phi^{-1}(w)$. The threshold age occurs when

$$\begin{aligned} \frac{1}{\beta}\,\phi(z) = \frac{e_o}{\beta e_o + 1} &\iff z = \phi^{-1}\left(\frac{\beta e_o}{\beta e_o + 1}\right) \\ &\iff \mathrm{e}^{\beta(x-M)} = \phi^{-1}\left(\frac{\beta e_o}{\beta e_o + 1}\right) \\ &\iff x = M + \frac{1}{\beta}\log\left[\phi^{-1}\left(\frac{\beta e_o}{\beta e_o + 1}\right)\right] \,. \end{aligned}$$

To prove (2a), we need

$$-\log\left[\phi^{-1}\left(\frac{\beta e_o}{\beta e_o + 1}\right)\right] \approx \gamma \,. \tag{7}$$

A closed-form expression of the inverse of the exponential integral $E_1(z)$ does not exist, and it can only be approximated numerically (Pecina 1986). The same holds true for $\phi(z)$, so to compute $\phi^{-1}(w)$ we implement the Newton-Raphson method (see, for in-

stance, Quarteroni, Sacco, and Saleri 2007). For all the combinations of $\beta \in [0.05, 0.16]$ and $M \in [30, 100]$ displayed in Figure 1, we apply (5) to calculate the life expectancy at birth e_o, which leads to

$$0.586 < \frac{\beta\, e_o}{\beta\, e_o + 1} < 0.940 \; .$$

When restricting the analysis to $\beta \geq 0.08$ and $M \geq 60$, the range of this quotient reduces to $[0.810, 0.940]$.

For each fixed $w \in [0.810, 0.940]$, we aim to find z such that $\phi(z) = w$. The Newton-Raphson method starts with an initial guess z_0 and then iteratively updates the target value as

$$\phi^{-1}(w) \approx z_{n+1} = z_n - \frac{\phi(z_n) - w}{\phi'(z_n)} \; . \tag{8}$$

The procedure continues until $|\phi(z_n) - w| < 10^{-6}$ or a maximum number of predetermined iterations is reached.[4] The derivative in the denominator is obtained using the identity $\phi'(z) = \phi(z) - z^{-1}$ (Abramowitz and Stegun 1964: 230).

We apply the Newton-Raphson method to estimate $\phi^{-1}(w)$ for a large set of values of w between 0.810 and 0.940. As a result, $-\log\left[\phi^{-1}(w)\right] \in [0.475, 0.724]$ and

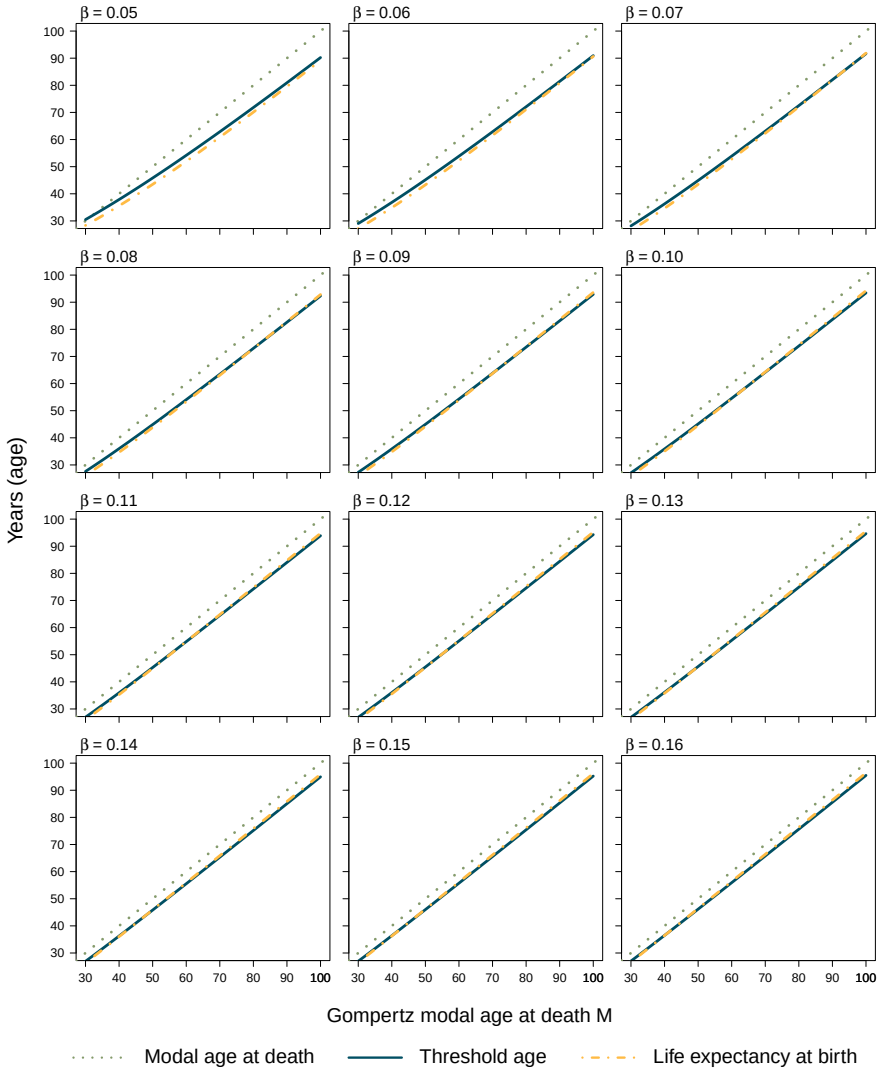
$$\max\left\{\left|\gamma + \log\left[\phi^{-1}\left(\frac{\beta e_o}{\beta e_o + 1}\right)\right]\right| : \beta \in [0.08, 0.16], M \in [60, 100]\right\} < 0.146 \; ,$$

which proves (7) and (2a). Using (1), Equation (2b) follows immediately. □

Figure 2 illustrates how close e_o and a^H are in a Gompertz setting and how they tend to evolve in parallel with M – especially for combinations of $\beta \geq 0.08$ and $M \geq 60$, the most common scenarios in contemporary human populations – which validates (1), (2a), and (2b). In addition, Table A-1 in the Appendix reports some combinations of the Gompertz parameters β and M for which (7) is a good approximation.

[4] Finding z such that $\phi(z) = w$ is analogous to searching for the roots of $g(z) := \phi(z) - w$. The Newton-Raphson method iteratively approaches a given root of $g(z)$ with the root of the tangent line to $g(z)$ at z_n, which has slope $g'(z_n) = g(z_n)/(z_n - z_{n+1})$. Hence, $z_{n+1} = z_n - g(z_n)/g'(z_n) = z_n - (\phi(z_n) - w)/\phi'(z_n)$, yielding (8).

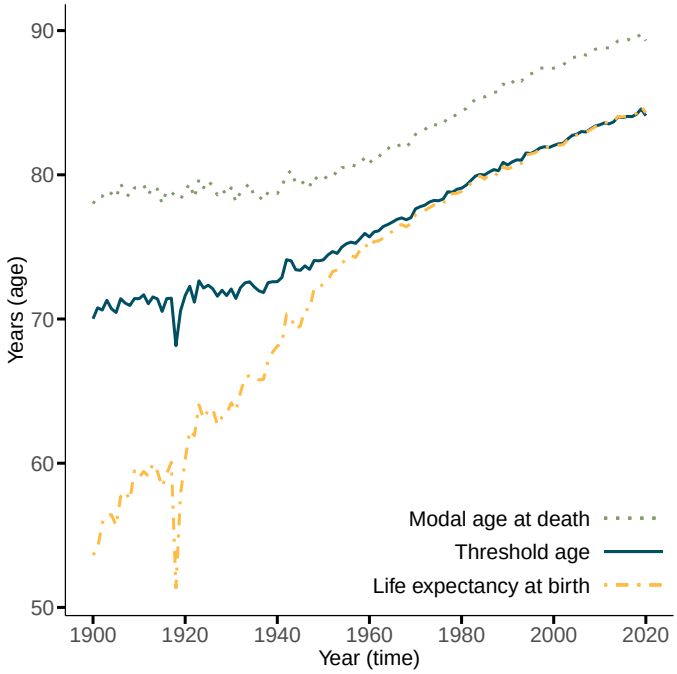
Figure 2: **Estimates of the life expectancy at birth e_o and the threshold age a^H under the Gompertz model, for a wide range of combinations of parameters $\beta \in \{0.05, \ldots, 0.16\}$ and $M \in \{30, \ldots, 100\}$**



4. History and related results

Some of the analyses by Aburto et al. (2020) on the dynamics of life expectancy and lifespan equality suggest that in low-mortality settings the life expectancy at birth e_o and the threshold age of the life table entropy a^H tend to converge. As shown in Figure 3, this is the case, for instance, of Swedish females from 1950 onward.

Figure 3: **Life expectancy at birth e_o, threshold age a^H, and modal age at death M of Swedish females, 1900–2019. This figure replicates Box 1, Figure 1, Panel A in Aburto et al. (2020: 5254) using the most recent data from the Human Mortality Database (2023)**



Values of e_o, M, and a^H in Figure 3 are estimated from life table data without assuming any parametric mortality model. The mode M is the (adult) age at which the age-at-death distribution reaches the maximum, whereas a^H is estimated numerically using the approach suggested by Aburto et al. (2019) (see Section 4.2 for details). Prior to 1950, when infant and child mortality were high, values of e_o were considerably lower than a^H. However, M and a^H evolve in parallel for the whole period 1900–2019, separated by a

nearly constant gap. We have proved in (2a) that in a Gompertz setting $a^H \approx M - \frac{\gamma}{\beta}$. In the application, we will illustrate how the Gompertz mode and rate of ageing can be used to estimate the empirical threshold age.

4.1 The Gompertz mortality model

Since its original formulation two centuries ago, the Gompertz (1825) model has been widely used by demographers, actuaries, and biologists to study mortality patterns of human and non-human populations (Colchero et al. 2016; Campos et al. 2020). The force of mortality of the Gompertz model changes exponentially over age $x \geq 0$, and it is commonly expressed in terms of the baseline mortality $\alpha > 0$ and the rate of ageing $\beta \in \mathbb{R}$ as

$$\mu(x) = \alpha \, \mathrm{e}^{\beta x} \, . \tag{9}$$

This formulation allows for a non-positive rate of ageing, something uncommon in human populations but that has been observed in some species across the tree of life (Jones et al. 2014). Aburto et al. (2022) explore the different patterns of the Gompertz model for positive, zero, and negative rates of ageing and how they relate to the threshold age of Drewnowski's index, an indicator of lifespan variation based on the Gini coefficient.

The Gompertz force of mortality as a function of the mode M and the rate of ageing β was first defined by Gumbel (1958), if not before. It is given by

$$\mu(x) = \beta \, \mathrm{e}^{\beta(x-M)} \tag{10}$$

and previous research has shown the convenience of using this formulation, provided that the modal age at death M is more informative and easy to interpret than the baseline mortality α (Canudas-Romo 2008; Horiuchi et al. 2013; Missov et al. 2015, among others).

Equations (9) and (10) are two equivalent definitions of the Gompertz force of mortality, and the latter can be derived from the former using the identity $\alpha = \beta \, \mathrm{e}^{-\beta M}$. Analogously, $M = (\log \beta - \log \alpha) / \beta$. That said, (10) is only defined for $\beta > 0$: If $\beta = 0$, then $\mu(x) = 0$ becomes meaningless; if $\beta < 0$, then $\mu(x) < 0$, which contradicts the definition of a rate. Actually, the formula by Castellares et al. (2020) in (3) to calculate the remaining life expectancy at age x is only valid for $\beta > 0$, even when applied to (9):

- If $\beta = 0$, the force of mortality $\mu(x) = \alpha$ is constant and independent of age, and the survivorship becomes $\ell(x) = \mathrm{e}^{-\alpha x}$. The remaining life expectancy at age x is

then $e(x) = \int_x^\infty \ell(a)\,da \,/\, \ell(x) = 1 \,/\, \alpha$, the inverse of $\mu(x)$. Notice the similarity with the exponential distribution with rate λ and expected value $1 \,/\, \lambda$.

- If $\beta < 0$, then $z = \mu(x) \,/\, \beta < 0$ in (3). Following Abramowitz and Stegun (1964), the exponential integral $E_1(z)$ is defined for all $z \in \mathbb{C}$ such that $|\text{Arg}(z)| < \pi$. The argument of a complex number z is the angle between the positive real axis and the segment joining the origin $(0, 0)$ and z in the complex plane. The argument of all real (not complex) negative numbers is π. Therefore, $E_1(z)$ is not defined for any $z \in \{s \in \mathbb{R} : s < 0\}$. Moreover, $\log(z)$ in (4) is not defined for negative z either.

The relationship between M and e_o in the Gompertz framework was already identified by Pollard and Valkovics (1992) and Vaupel (forthcoming) but from different approaches. Pollard and Valkovics (1992) find that $e_o = -(\log(\alpha \,/\, \beta) + \gamma) \,/\, \beta$, which is equivalent to (1) but with the parametrisation in (9). They obtain this result by calculating the first cumulant of the Gompertz distribution from the cumulant-generating function, which is the log of the moment-generating function. Vaupel (forthcoming) proves (1) using survival ages (Alvarez and Vaupel 2023) and making some mild assumptions of plausible values of β and M. Our approach is new to the extent that we use the closed-form expression of the Gompertz $e(x)$ from Castellares et al. (2020) and that we carry out a systematic sensitivity analysis to explore the combinations of the Gompertz parameters for which (1) actually holds (Figure 1).

As for the relationship between M and a^H, to our knowledge (2a) is new in the literature. Equation (2b) derives from (1) and (2a).

4.2 The life table entropy and the threshold age

Over the past two centuries, increases in life expectancy at birth have been followed by a decrease in lifespan variation, with people tending to die at more similar ages (Edwards and Tuljapurkar 2005; van Raalte, Sasson, and Martikainen 2018; Vaupel, Villavicencio, and Bergeron-Boucher 2021). Aburto et al. (2020) show that these parallel changes are not coincidental but a consequence of improvements in mortality reduction at specific ages. In particular, they find that such improvements always positively impact life expectancy, while for lifespan variation there exists an age above which averting deaths worsens inequalities. This age is referred to as the "threshold age" and has been computed for different indicators of variation in length of life, such as life disparity (Zhang and Vaupel 2009), the variance of the distribution of ages at death (Gillespie, Trotter, and Tuljapurkar 2014), or Drewnowski's index (Aburto et al. 2022). On a recent paper, Martin, Aburto, and Permanyer (2023) calculate the threshold age of the coefficient of variation and discuss the similarities among different inequality measures. They observe that, in certain populations, absolute measures of inequality, such as the variance, life

disparity, or the Gini coefficient, may not have a threshold age, implying that any improvement in mortality would increase lifespan variation. On the contrary, when using the corresponding measures of relative inequality with respect to life expectancy (i.e., coefficient of variation, life table entropy, and the relative Gini coefficient), there will always be ages for which averting deaths will reduce lifespan inequality and ages for which variation will increase.

In a general setting without assuming any underlying mortality model, Aburto et al. (2019) prove that the threshold age of the life table entropy occurs when

$$H(x) + \overline{H}(x) - 1 + \overline{H} = 0\,, \tag{11}$$

where $H(x)$ is the cumulative hazard and $\overline{H}(x)$ the life table entropy conditioned on surviving to age x. Equation (11) is a generalisation of (6), and likewise can only be solved numerically. They also suggest that under the Gompertz model the life expectancy at birth e_o and the threshold age a^H are approximately proportional. More specifically, that

$$a^H \approx e_o \cdot \delta\,, \qquad \text{where} \quad \delta = \frac{\gamma + \log(\alpha\,/\,\beta)}{\mathrm{e}^{\alpha/\beta}\,(\gamma + \log(\alpha\,/\,\beta)) - 1}\,.$$

Using the identity $\alpha = \beta\,\mathrm{e}^{-\beta M}$ yields

$$\delta = \frac{\gamma - \beta\,M}{\mathrm{e}^{\mathrm{e}^{-\beta M}}(\gamma - \beta\,M) - 1}\,. \tag{12}$$

However, (12) does not seem to work when compared to the results from the sensitivity analysis in Figure 2. For instance, when $\beta = 0.05$ and $M = 40$ one gets $\delta \approx 0.54$, which would mean that a^H is about one half of e_o. Yet from Figure 2 it is clear that $a^H > e_o$ for these values of β and M. In fact, $0 < \delta < 1$ for all combinations of $\beta \in \{0.05, \ldots, 0.16\}$ and $M \in \{30, \ldots, 100\}$ (same combinations as in Figure 2), which would imply that $a^H < e_o$. Figure 2 proves this is not always the case.

Discrepancies should be attributed to the fact that Aburto et al. (2019) do not use the correct expression of the Gompertz life expectancy (see Equation (A10) in the Appendix of Aburto et al. 2019: 101). For further discussion on the Gompertz formulae, see Castellares et al. (2020). Our sensitivity analysis in Figure 2 shows that $a^H \approx e_o$, so the constant of proportionality between the two (if any) should be close to 1 for a wide range of values of β and M.

5. Application

In this section, we assess the accuracy of the approximation of the threshold age of the life table entropy provided in Section 2. We distinguish between the empirical threshold age a^H, computed numerically using (11) following Aburto et al. (2019), and the approximated threshold age $\hat{a}^H$, obtained from (2a) in a Gompertz framework. We evaluate the approximation's precision studying the distribution of $\hat{a}^H - a^H$, while also calculating a set of errors. Furthermore, we test the behaviour of (2a) replacing the Gompertz mode M with the observed/empirical one $M_{obs} \coloneqq \max_x d(x)$, obtained by maximising the age-at-death distribution $d(x)$. The Gompertz parameters α and β in (9) are estimated via maximum likelihood, under the assumption that death counts are Poisson distributed within each age interval (Currie 2016). The mode is then obtained as $M = (\log \beta - \log \alpha) / \beta$.

All the analyses are performed with the open-source statistical software `R` (version 4.4.1) (R Core Team 2021).

5.1 The age range of the Gompertz model

Gompertz (1825) observed that human death rates start being log-linear from adult ages. Depending on countries and period, the specific age at which this log-linearity begins is not universally established: Common choices fall between 30 and 50 years. For each country-year combination, we estimate α and β across three different scenarios, assuming the Gompertz mortality model over the age ranges $[30, 90]$, $[40, 90]$, and $[50, 90]$. We use 8,022 period life tables from the Human Mortality Database (2023) for 48 countries, females and males, from 1900 (or earliest year) to the most recent year available.

Figure 4 shows the evolution of a^H, $\hat{a}^H$, and life expectancy at birth e_o for females from England and Wales, Italy, Japan, and the United States using the three different starting ages: 30, 40, and 50 years. We observe that the approximation works suitably in all cases, although depending on the country a certain initial age can produce better results. To determine the optimal age range, for each country and sex we evaluate the accuracy of our approximation by computing the mean absolute error (MAE)

$$\text{MAE}(\hat{a}^H) = \frac{1}{T} \sum_{i=1}^{T} \left| \hat{a}^H(t_i) - a^H(t_i) \right| \ ,$$

and the root mean square error (RMSE)

$$\text{RMSE}(\hat{a}^H) = \sqrt{\frac{1}{T}\sum_{i=1}^{T}(\hat{a}^H(t_i) - a^H(t_i))^2}\ ,$$

for the three Gompertz onsets over the available years $t_1, \ldots, t_T$. Here, $a^H(t_i)$ denotes the empirical threshold age of year t_i calculated via (11), and $\hat{a}^H(t_i)$ the approximated threshold age obtained with (2a). We rank each of the errors, by country and sex, from lowest (ranking 1) to highest (ranking 3): The age with the smallest sum of rankings will lead, in general, to the most accurate results.

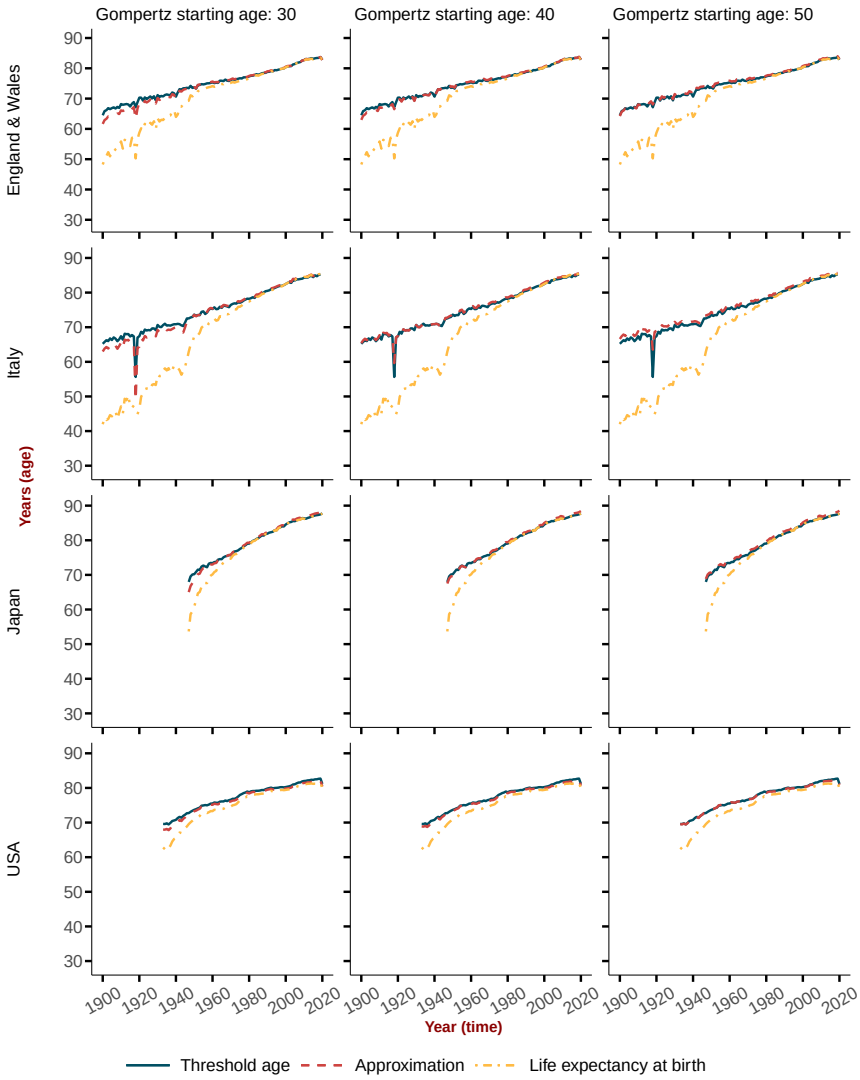
To complement this analysis and validate the outcome, we perform identical calculations using 32,904 sex-specific life tables from UN WPP, covering more than 225 countries and territories from 1950 to 2021 (United Nations 2022). As summarised in Table 1, for the two data sets both the MAE and the RMSE indicate 40 years as the preferable choice for women and 50 years for men.

Table 1: Sum of the rankings of the mean absolute error (MAE) and the root mean square error (RMSE) for the different Gompertz starting ages calculated using life tables from the Human Mortality Database (2023) and United Nations World Population Prospects (2022)

	Sum of rankings							
	Females				Males			
	HMD		UN WPP		HMD		UN WPP	
Starting age	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
30	98	101	603	615	127	122	643	643
40	77	78	369	369	95	93	418	413
50	125	121	408	396	78	85	301	306

These findings align with the sex differences caused by the "young adult mortality hump" that have long been observed in human populations (Thiele 1871; Remund, Camarda, and Riffe 2018). This phenomenon consists in higher mortality among males towards the end of adolescence and early adulthood due to increases in violent, accidental, and disease mortality (Goldstein 2011). As a result, men's Gompertz onset is delayed, and the log-linearisation of death rates is postponed to older ages.

Figure 4: **Threshold age of the life table entropy a^H (blue), its approximation $\hat{a}^H = M - \gamma / \beta$ (red), and life expectancy at birth e_o (yellow). Females of England and Wales, Italy, Japan, and USA, 1900–2021 (Human Mortality Database 2023)**



5.2 Results

Figure 5 displays the 95% central distribution of $\hat{a}^H - a^H = (M - \gamma/\beta) - a^H$, split by sex and over two periods, 1900–1950 and 1951–2000. We estimate β and M using the age range 40 to 90 years for females and 50 to 90 for males to calculate $\hat{a}^H$. We restrict our results to data from the HMD since the UN WPP covers only the period 1950–2021. Because we let the analysis begin in 1900 for countries with available data (such as Sweden, Finland, France, or Denmark), high-mortality regimes are also represented.

In all four panels of Figure 5, the mean and mode of the distributions are nearly zero. In particular, approximation (2a) works best in recent years, as for both females and males the distributions shrink. Negative values on the x-axis indicate underestimation of a^H by $M - \frac{\gamma}{\beta}$: This error is much reduced for females, while for males it slightly increases, mirroring the overestimation experienced in the 1900–1950 period. Yet men's mode shifts from 0.26 to 0.04 years, implying an almost perfect approximation of a^H by Equation (2a) for the majority of cases in the second half of the 20th century. Considering that a^H and M are usually larger than 60, the inaccuracy level is overall small.

When M is replaced by the observed mode M_{obs}, the resulting estimates of a^H worsen. This can be ascertained from the distribution of $M_{obs} - \gamma/\beta - a^H$ (Appendix Figure A-1). For all the sex-period combinations, the mode is closer to zero, but the variance is higher, and the maximum errors greater (larger x-axis values). Ultimately, the empirical mode is more affected by mortality crises than the threshold age. A striking example is Spain, heavily hit by the flu pandemic between 1918 and 1919. As shown in Appendix Figure A-2, a dramatic drop in M_{obs} is only partially reflected in a^H and $\hat{a}^H$. The Gompertz parameters can be then considered suitable candidates to estimate a^H, regardless of the underlying mortality pattern.

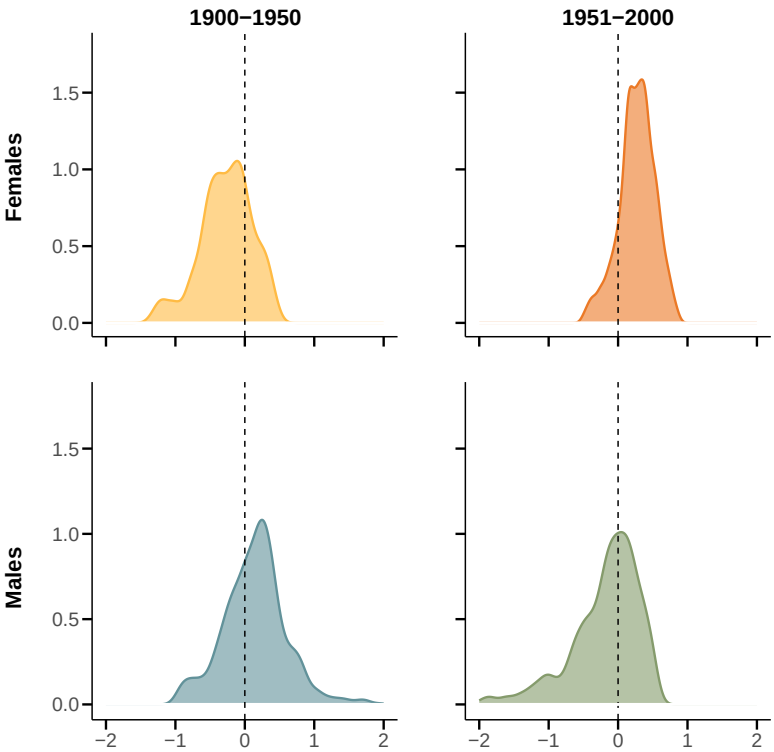
As already anticipated by Aburto et al. (2019, 2020), Figure 4 also clearly shows the convergence of the threshold age and life expectancy in recent years. This confirms relationship (2b) and proposes e_o as an alternative way to approximate a^H. Nevertheless, this is not true in high mortality settings (Figure 3). Our results instead empirically prove that (2a) works even for populations not actually following Gompertz and that may be affected by mortality crises or shocks like wars and epidemics.

6. Conclusion

The life expectancy revolution that started in the mid-19th century among the longevity vanguard countries (Oeppen and Vaupel 2002) has been accompanied by a decrease in lifespan variation: As people live longer, ages at death are becoming more similar (Edwards and Tuljapurkar 2005; Permanyer and Scholl 2019; Vaupel, Villavicencio, and Bergeron-Boucher 2021, among others). This inverse relationship, though, has been

questioned as some populations experienced growing levels of inequality in the length of life, even amidst overall improvements in life expectancy. Over the last 40 years, examples are found in different contexts, like Spain (Permanyer et al. 2018), the United States (Sasson 2016), or Denmark (Brønnum-Hansen 2017), where stagnation or increases in lifespan variation are observed among socially disadvantaged groups, notwithstanding an averagely longer life. Aburto et al. (2020) identify the reasons behind these trends in the ages at which mortality improvements are concentrated. In particular, there exists a threshold age separating early from late deaths, above and below which saving lives respectively increases or decreases lifespan inequality.

Figure 5: **The 95% central distribution of $\hat{a}^H - a^H$, by sex and period. 8,022 life tables from the Human Mortality Database (2023)**



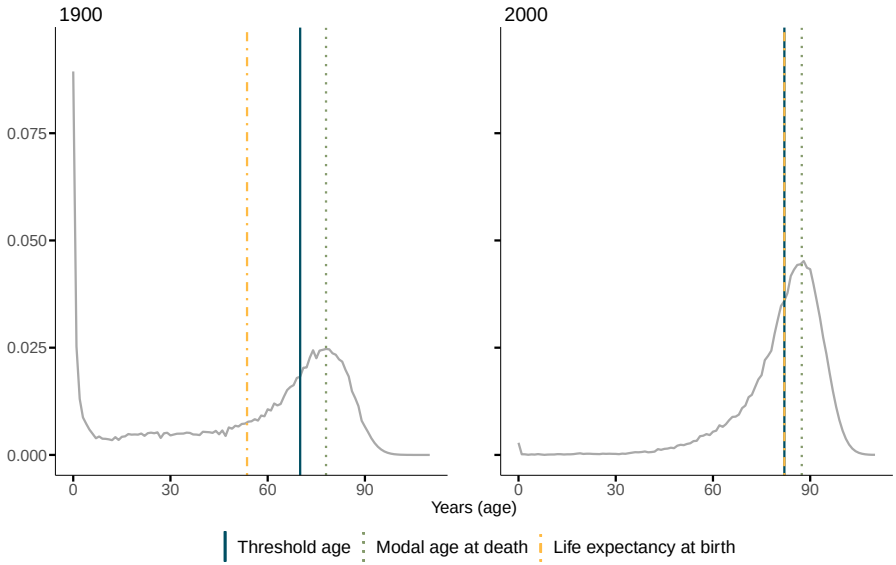
There are several indicators of lifespan variation for which the threshold age has been identified, such as life disparity (Zhang and Vaupel 2009), the variance of the ages at death (Gillespie, Trotter, and Tuljapurkar 2014), the Gini coefficient and Drewnowski's index (Aburto et al. 2022), or the coefficient of variation (Martin, Aburto, and Permanyer 2023). In this paper we focus on the threshold age a^H of the life table entropy, an indicator of the relative variation in the length of life compared to life expectancy at birth. In Equation (2b) and Figure 2 we show that the closeness between e_o and a^H is implicit in a Gompertz framework; in (1) and (2a), we prove that both indicators are a negative shift of the modal age at death M.

However, reality can be more complex than the exponential risk of death described by the Gompertz law. Figures 3 and 4 show that (1) and (2b) clearly hold in low-mortality settings, as e_o and a^H tend to converge over time. But this has not always been the case. As depicted in Figure 3 for Swedish females before 1950, e_o was notably lower than a^H, while a^H and M progressed steadily and much more gradually over the whole period 1900–2019. Our empirical results in Section 5 prove that the relationship $a^H \approx M - \gamma \,/\, \beta$ in (2a) is stronger and holds even when the Gompertz assumptions are not respected. Previous research has shown that the effects of mortality crises on the threshold age vary depending on the selected indicator of lifespan variation (Vigezzi et al. 2022). Here, we have seen that parameters β and M obtained by fitting a Gompertz model to mortality data for late-adult ages (40 to 90 years for females and 50 to 90 for males) can be used to estimate the threshold age of the life table entropy, even when $e_o << a^H$ or populations are hit by exceptional catastrophes.

While we are not able to establish a causal link between the two, similar mechanisms seem to drive the advances of the mode and the threshold age. The adult modal age at death is mostly determined by old-age mortality (Horiuchi et al. 2013) and illustrates changes from a dominance of child mortality reductions to a dominance of adult mortality reductions (Canudas-Romo 2008). Figure 6 reflects these ideas. In 1900, Swedish females enjoyed a life expectancy at birth of $e_o = 53.63$ years, while $M = 78$ and $a^H = 70.02$. A century later, e_o increased by 28.39 years, mainly thanks to remarkable and unprecedented reductions of infant and child mortality. Improvements in M and a^H were less sensational, 9.4 and 12.01 years, respectively. One could speculate that the relationship between M and a^H in (2a) is a feature of human populations, while (1) and (2b) hold only when mortality becomes roughly Gompertzian.

The success of the Gompertz model is mostly explainable by its accurate simplicity. Over the years, many have studied the interrelationships of the main parameters defining it. With this paper, we contribute to the existing literature investigating the dynamics of three lifespan indicators: the life expectancy at birth e_o, the adult modal age at death M, and the threshold age of the life table entropy a^H. Our research builds upon previous work by James W. Vaupel and illustrates the beauty of the mathematical relationships between demographic concepts.

Figure 6: **Age-at-death distribution of Swedish females in 1900 and 2000 (Human Mortality Database 2023). Comparison of the threshold age of the life table entropy (blue), the modal age at death (green), and life expectancy at birth (yellow)**



7. Acknowledgements

CM is a PhD candidate in Demography at UAB/CED-CERCA, member of the project HEALIN (H2020 ERC, grant 2019-CoG-864616), and holds a doctoral fellowship FI-AGAUR (grant 2023 FI-1 00261). FV acknowledges funding from the Spanish State Research Agency under the Ramón y Cajal programme (grant RYC2021-033979-I). The authors are grateful to two anonymous reviewers for their constructive feedback.

8. Data availability

The data and code to replicate all the results and figures presented here are publicly available on the GitHub repository https://github.com/cmicheletti/ThresholdAgeMode.

References

Abramowitz, M. and Stegun, I.A. (eds.) (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Washington, D.C.: US Department of Commerce.

Aburto, J.M., Alvarez, J.A., Villavicencio, F., and Vaupel, J.W. (2019). The threshold age of the lifetable entropy. *Demographic Research* 41(4): 83–102. doi:10.4054/DemRes.2019.41.4.

Aburto, J.M., Basellini, U., Baudisch, A., and Villavicencio, F. (2022). Drewnowski's index to measure lifespan variation: Revisiting the Gini coefficient of the life table. *Theoretical Population Biology* 148: 1–10. doi:10.1016/j.tpb.2022.08.003.

Aburto, J.M., Villavicencio, F., Basellini, U., Kjærgaard, S., and Vaupel, J.W. (2020). Dynamics of life expectancy and life span equality. *Proceedings of the National Academy of Sciences of USA* 117(10): 5250–5259. doi:10.1073/pnas.1915884117.

Alvarez, J.A. and Vaupel, J.W. (2023). Mortality as a function of survival. *Demography* 60(1): 327–342. doi:10.1215/00703370-10429097.

Brønnum-Hansen, H. (2017). Socially disparate trends in lifespan variation: A trend study on income and mortality based on nationwide Danish register data. *BMJ Open* 7(5): e014489. doi:10.1136/bmjopen-2016-014489.

Campos, F.A., Villavicencio, F., Archie, E.A., Colchero, F., and Alberts, S.C. (2020). Social bonds, social status and survival in wild baboons: A tale of two sexes. *Philosophical Transactions of the Royal Society B: Biological Sciences* 375(1811): 20190621. doi:10.1098/rstb.2019.0621.

Canudas-Romo, V. (2008). The modal age at death and the shifting mortality hypothesis. *Demographic Research* 19(30): 1179–1204. doi:10.4054/DemRes.2008.19.30.

Castellares, F., Patrício, S.C., Lemonte, A.J., and Queiroz, B.L. (2020). On closed-form expressions to Gompertz–Makeham life expectancy. *Theoretical Population Biology* 134: 53–60. doi:10.1016/j.tpb.2020.04.005.

Colchero, F., Rau, R., Jones, O.R. et al. (2016). The emergence of longevous populations. *Proceedings of the National Academy of Sciences* 113(48): E7681–E7690. doi:10.1073/pnas.1612191113.

Currie, I.D. (2016). On fitting generalized linear and non-linear models of mortality. *Scandinavian Actuarial Journal* 2016(4): 356–383. doi:10.1080/03461238.2014.928230.

Demetrius, L. (1978). Adaptive value, entropy and survivorship curves. *Nature* 275(5677): 213–214. doi:10.1038/275213a0.

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.

Gillespie, D.O.S., Trotter, M.V., and Tuljapurkar, S.D. (2014). Divergence in age patterns of mortality change drives international divergence in lifespan inequality. *Demography* 51(3): 1003–1017. doi:10.1007/s13524-014-0287-8.

Goldman, N. and Lord, G. (1986). A new look at entropy and the life table. *Demography* 23(2): 275–282. doi:10.2307/2061621.

Goldstein, J.R. (2011). A secular trend toward earlier male sexual maturity: Evidence from shifting ages of male young adult mortality. *PLoS ONE* 6(8): e14826. doi:10.1371/journal.pone.0014826.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. F.R.S. &c. *Philosophical Transactions of the Royal Society of London* 115: 513–583. doi:10.1098/rspl.1815.0271.

Gumbel, E.J. (1958). *Statistics of extremes*. New York, NY: Columbia University Press. doi:10.7312/gumb92958.

Hanada, K. (1983). A formula of Gini's concentration ratio and its application to life tables. *Journal of Japanese Statistical Society* 13(2): 95–98. doi:10.11329/jjss1970.13.95.

Horiuchi, S., Ouellette, N., Cheung, S.L.K., and Robine, J.M. (2013). Modal age at death: Lifespan indicator in the era of longevity extension. *Vienna Yearbook of Population Research* 11: 37–69. doi:10.1553/populationyearbook2013s37.

Human Mortality Database (2023). Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). http://www.mortality.org (accessed 27 March 2022).

Jones, O.R., , Scheuerlein, A., Salguero-Gómez, R., Camarda, C.G., Schaible, R., Casper, B.B., Dahlgren, J.P., Ehrlén, J., García, M.B., Menges, E.S., Quintana-Ascencio, P.F., Caswell, H., Baudisch, A., and Vaupel, J.W. (2014). Diversity of ageing across the tree of life. *Nature* 505(7482): 169–173. doi:10.1038/nature12789.

Keyfitz, N. (1977). What difference would it make if cancer were eradicated? An examination of the Taeuber paradox. *Demography* 14(4): 411–418. doi:10.2307/2060587.

Leser, C.E.V. (1955). Variations in mortality and life expectation. *Population Studies* 9(1): 67–71. doi:10.1080/00324728.1955.10405052.

Martin, J., Aburto, J.M., and Permanyer, I. (2023). Dynamics of the coefficient of variation of the age at death distribution. *Demographic Research* 49(38): 1063–1086. doi:10.4054/DemRes.2023.49.38.

Missov, T.I., Lenart, A., Nemeth, L., Canudas-Romo, V., and Vaupel, J.W. (2015). The Gompertz force of mortality in terms of the modal age at death. *Demographic Research* 32(36): 1031–1048. doi:10.4054/DemRes.2015.32.36.

Oeppen, J. and Vaupel, J.W. (2002). Broken limits to life expectancy. *Science* 296(5570): 1029–1031. doi:10.1126/science.1069675.

Pecina, P. (1986). On the function inverse to the exponential integral function. *Bulletin of the Astronomical Institute of Czechoslovakia* 37(1): 8–12.

Permanyer, I. and Scholl, N. (2019). Global trends in lifespan inequality: 1950–2015. *PLoS ONE* 14(5): e0215742. doi:10.1371/journal.pone.0215742.

Permanyer, I., Spijker, J., Blanes, A., and Renteria, E. (2018). Longevity and lifespan variation by educational attainment in Spain: 1960–2015. *Demography* 55(6): 2045–2070. doi:10.1007/s13524-018-0718-z.

Pollard, J.H. and Valkovics, E.J. (1992). The Gompertz distribution and its applications. *Genus* 48(3/4): 15–28.

Quarteroni, A., Sacco, R., and Saleri, F. (2007). *Numerical mathematics*. Berlin, Heidelberg: Springer. doi:10.1007/b98885.

R Core Team (2021). R: A language and environment for statistical computing. Vienna: The R Project for Statistical Computing. https://www.R-project.org/.

Remund, A., Camarda, C.G., and Riffe, T. (2018). A cause-of-death decomposition of young adult excess mortality. *Demography* 55(3): 957–978. doi:10.1007/s13524-018-0680-9.

Sasson, I. (2016). Trends in life expectancy and lifespan variation by educational attainment: United States, 1990–2010. *Demography* 53(2): 269–293. doi:10.1007/s13524-015-0453-7.

Shkolnikov, V.M., Andreev, E.E., 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.

Smits, J. and Monden, C. (2009). Length of life inequality around the globe. *Social Science & Medicine* 68(6): 1114–1123. doi:10.1016/j.socscimed.2008.12.034.

Thiele, T.N. (1871). On a mathematical formula to express the rate of mortality throughout the whole of life, tested by a series of observations made use of by the Danish Life Insurance Company of 1871. *Journal of the Institute of Actuaries* 16(5): 313–329. doi:10.1017/S2046167400043688.

United Nations (2022). *World Population Prospects: The 2022 revision*. United Nations, Department of Economic and Social Affairs, Population Division. https://population.un.org/wpp (acessed 11 September 2023).

van Raalte, A.A., Sasson, I., and Martikainen, P. (2018). The case for monitoring lifespan inequality. *Science* 362(6418): 1002–1004. doi:10.1126/science.aau5811.

Vaupel, J.W. (1986). How change in age-specific mortality affects life expectancy. *Population Studies* 40(1): 147–157. doi:10.1080/0032472031000141896.

Vaupel, J.W. (forthcoming). Gompertz mean, mode, and median. *Demographic Research*.

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:10.1353/dem.2003.0018.

Vaupel, J.W., Villavicencio, F., and Bergeron-Boucher, M.P. (2021). Demographic perspectives on the rise of longevity. *Proceedings of the National Academy of Sciences* 118(9): e2019536118. doi:10.1073/pnas.2019536118.

Vaupel, J.W., Zhang, Z., and van Raalte, A.A. (2011). Life expectancy and disparity: An international comparison of life table data. *BMJ Open* 1(1): e000128. doi:10.1136/bmjopen-2011-000128.

Vigezzi, S., Aburto, J.M., Permanyer, I., and Zarulli, V. (2022). Divergent trends in lifespan variation during mortality crises. *Demographic Research* 46(11): 291–336. doi:10.4054/DemRes.2022.46.11.

Zhang, Z. and Vaupel, J.W. (2009). The age separating early deaths from late deaths. *Demographic Research* 20(29): 721–730. doi:10.4054/DemRes.2009.20.29.

Appendix

Table A-1: Sample of combinations of the Gompertz parameters β and M for which Equation (7) is a good approximation ($\gamma \approx 0.5772157$)

M	β	e_o	$w = \frac{\beta e_o}{\beta e_o + 1}$	$\Phi^{-1}(w)$	$-\log\left[\Phi^{-1}(w)\right]$	$\lvert\gamma + \log\left[\Phi^{-1}(w)\right]\rvert$
40	0.16	36.46	0.854	0.570	0.561	0.016
45	0.14	40.97	0.852	0.573	0.557	0.020
45	0.15	41.21	0.861	0.563	0.575	0.002
45	0.16	41.43	0.869	0.554	0.591	0.014
50	0.13	45.64	0.856	0.568	0.566	0.011
50	0.14	45.93	0.865	0.558	0.584	0.007
50	0.15	46.18	0.874	0.549	0.601	0.024
55	0.12	50.27	0.858	0.566	0.570	0.007
55	0.13	50.61	0.868	0.555	0.589	0.012
55	0.14	50.90	0.877	0.545	0.606	0.029
60	0.11	54.84	0.858	0.566	0.570	0.007
60	0.12	55.24	0.869	0.554	0.591	0.014
65	0.10	59.33	0.856	0.568	0.566	0.011
65	0.11	59.81	0.868	0.555	0.589	0.012
70	0.09	63.72	0.852	0.573	0.557	0.020
70	0.10	64.30	0.865	0.558	0.584	0.007
70	0.11	64.79	0.877	0.545	0.606	0.029
75	0.09	68.68	0.861	0.563	0.575	0.002
75	0.10	69.27	0.874	0.549	0.601	0.024
80	0.08	72.93	0.854	0.570	0.561	0.016
80	0.09	73.65	0.869	0.554	0.591	0.014
85	0.08	77.89	0.862	0.562	0.577	0.000
85	0.09	78.63	0.876	0.546	0.605	0.028
90	0.07	81.93	0.852	0.573	0.557	0.020
90	0.08	82.86	0.869	0.554	0.591	0.014
95	0.07	86.88	0.859	0.565	0.571	0.006
95	0.08	87.84	0.875	0.547	0.604	0.027

Figure A-1: **The 95% central distribution of $M_{obs} - \gamma/\beta - a^H$, by sex and period. 8,022 life tables from the Human Mortality Database (2023)**

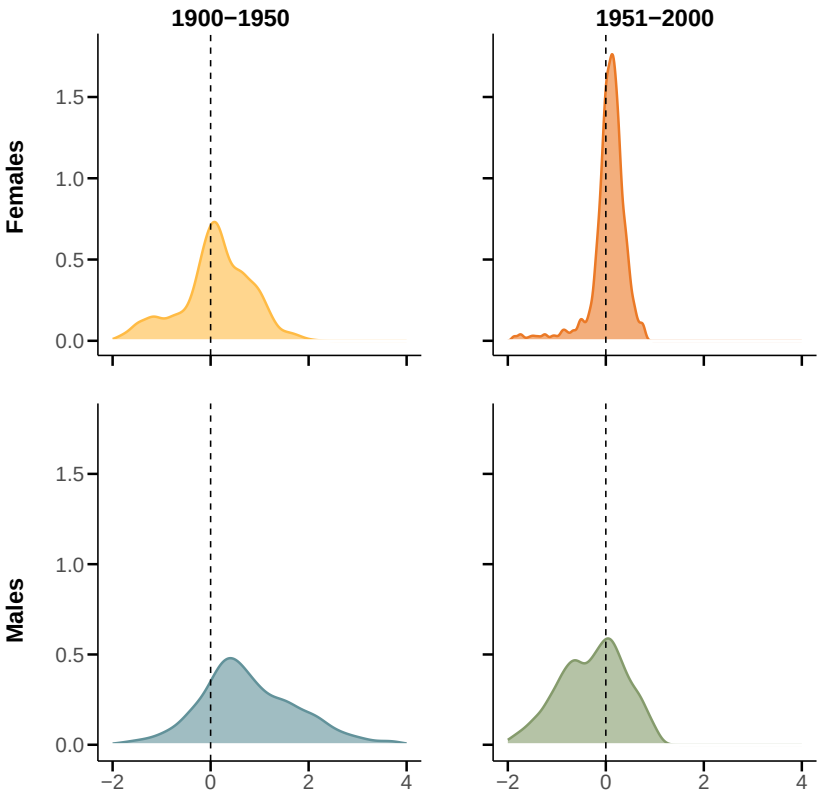


Figure A-2: **The threshold age of the life table entropy a^H (blue), its approximation $\hat{a}^H = M - \gamma \,/\, \beta$ (red), the life expectancy at birth e_o (yellow), and the observed mode M_{obs} (green). Spanish females, 1908–2021 (Human Mortality Database 2023)**