A parametric family of hazard trajectories for infant mortality

Jonas Schöley*

Abstract

Background While there is a consensus that the risk of death follows a Gompertz law over much of the adult age span, no such agreement exists about the parametric form of mortality at the very beginning of life with most literature on the topic suggesting either an exponential or various power-law expressions.

Objective I aim to identify a parsimonious, interpretable, and well-fitting model for the infant mortality age trajectory as observed in recent U.S. birth cohorts across a range of social and medical strata.

Results Age-specific infant mortality in the U.S. displays both power-law and exponential behavior and is better described by a product of those functions: a "power law with an exponential tail." Across all infant populations under consideration, the age trajectory of mortality following birth is initially dominated by a power-law regime. Throughout infancy, it eventually approaches an exponential decline of less than a one percent reduction per additional day of age. The hazard of infant death varies in a highly non-proportional fashion by prematurity and the infant's health upon birth.

Contribution The power-exponential hazard is a novel tool for the study of infant mortality being more parsimonious than smoothing-splines while providing interpretable parameters and an excellent fit over a range of diverse populations. The transition from a power-law dominated neonatal mortality schedule to an exponential post-neonatal hazard has not been noted before and suggests an underlying shock-recovery or mortality selection process.

*Interdisciplinary Centre on Population Dynamics, University of Southern Denmark. Correspondence: jschoeley@health.sdu.dk. During the writing of this article the author was a guest at the Max-Planck Institute for Demographic Research and funded by a grant from AXA Insurance.

1. Background

Since Benjamin Gompertz published his eponymous law of mortality for the adult ages, many suggestions were made for an expression of similar generality describing the age pattern of infant or childhood mortality. Perhaps the earliest attempt was published by Oppermann (1870; as stated by Thiele 1871) who proposed[1] $h(x) = ax^{-1/2} + b + cx^{1/2}$ for mortality prior to age 20. With a power law as the first term, this formula has the curious feature of predicting an infinite risk of death at the moment of birth, a property that reportedly was important to Oppermann though the reasons remain unclear (Steffensen 1930). One year later, Thiele (1871), motivated by the search for a mortality law covering the whole human life span, assumed $h(x) = ae^{-bx}$, the hazard of a negative-Gompertz distribution, for the changing risk of death before maturity. *These two expressions, a power-law and an exponential, constitute the functional basis for most parametric models of infant/childhood mortality employed thereafter.* Different power-law hazards have been used to describe the age pattern of infant mortality by Brillinger (1961), Choe (1981), de Beer and Janssen (2016), and Berrut et al. (2016); Gompertz-like exponential functions appeared as infant mortality terms in Siler (1979), Mode and Busby (1982), Rogers and Little (1994); and compositions of power- and exponential functions have been suggested by Wittstein and Bumsted (1883) and Heligman and Pollard (1980).

The mortality models above can be interpreted as describing an unspecified risk to which an infant adapts over time, resulting in a continuous decline in mortality as the child grows (as explicitly stated by Siler 1979; Heligman and Pollard 1980).[2] Levitis (2011) called this the "acquired robustness hypothesis." A "competing-risks" explanation was proposed by Bourgeois-Pichat (1951), who hypothesized that the observed age pattern of mortality over the first year of life is the result of two separate processes: intrinsic mortality due to congenital disorders and extrinsic mortality due to accidents and maltreatment of the child. Bourgeois-Pichat noticed that the cumulative distribution of deaths during the post-neonatal period (> 1 month of age) is closely matched by a linear function of $\log^3_{10}$(days since birth + 1) while neonatal deaths follow a different age trajectory. He then proposed that intrinsic and extrinsic mortality over age do not share the same functional form and that extrinsic mortality only starts to dominate over its intrinsic counterpart after the first month of life.

Some attempts have been made to express the parametric form of the infant hazard of death as a *frailty model* (Vaupel and Yashin 1983; Hougaard 1984; Vaupel and Yashin 1985). A valuable insight from these models is that in a population where individuals differ significantly in their risk of death the average mortality over age is determined not only by individual-level age effects but also by *mortality selection*, i.e., the changing composition of a cohort towards individuals with low frailty (Vaupel et al. 1979). It has been demonstrated that a *declining* infant mortality age trajectory may result from *constant* individual level hazards of different magnitudes (Vaupel and Yashin 1983, 1985). Hougaard (1984)

[1] Throughout the paper $h(x)$ denotes the force of mortality at age x.

[2] An exception is Oppermann (1870), who essentially proposed a competing risks model though the meaning Oppermann gave to the three components of his formula is unknown.

suggested that high mortality right after birth and the fast subsequent decline may be the result of substantial heterogeneity in individual frailties in the population of newborns.

In the presence of the ubiquitous semi-parametric Cox proportional hazards model (Cox 1972) and smoothing methods for count data such as penalized splines (Eilers and Marx 1996; Camarda et al. 2016) it may seem like an exercise in nostalgia to consider parametric hazard expressions as a basis for the analysis of the infant mortality age pattern. Nevertheless, a parametric treatment of deaths during the first year of life comes with unique advantages:

1) Successes in combating infant mortality over the past century made infant death an increasingly rare event in large parts of the world (World Health Organization 2006, 2015). The statistician thus is challenged to develop a methodology suited to inference from rare events, especially if the object of study is infant mortality over age, on the regional level, by season, in nations with a small population, by socio-economic group, by cause-of-death, or by any combination of those criteria. If correctly specified, parametric survival models alleviate the statistical challenge of small event counts. By imposing a tight structure on the distribution of life-times, one can gain insight from comparatively little data. Furthermore, a parametric specification of infant mortality expressed in terms of observable quantities (e.g., mortality at the day of birth, rate of post-neonatal mortality decline) facilitates the use of Bayesian methodology – a useful tool for working with sparse data – because informative prior distributions on the parameters are easily specified if the parameters are well understood.

2) Parametric models may allow insight into the mechanisms that determine the age distribution of infant deaths. This advantage is especially apparent with regard to the frailty model (Vaupel et al. 1979), which expresses the age-pattern of mortality on the population level as emerging from heterogeneous individual level hazards via selective mortality.

3) Information on the ages at death during infancy may only be available in broad age groups (e.g., birth to one day, one day to one week, one week to one month, months one to twelve). If the parametric form of the ages at death during the first year of life is known, this information can be used to graduate the grouped data. Demographers interested in a precise estimation of the "average age at death during infancy," a relevant statistic for the construction of life tables, can make use of such graduation.

The advantages above depend on a parsimonious yet well-fitting expression of the age distribution of infant deaths. Validating such an expression requires a) data on the precise timing of each infant death in a birth cohort (ideally on a day-to-day basis); b) a cohort of infants large enough for a clear trend in the age trajectory of infant mortality rates to dominate over random variation; and c) data on multiple birth-cohorts, and subpopulations so that it is possible to check the generality of the parametric model. The individual-level data on millions of births and infant deaths provided by the National Center for Health

Statistics (2016) fit these requirements, and for this article, I confine the search for the shape of the hazard of infant death to the U.S. population.

After a brief description of the data, the methods, and the current day-to-day age pattern of infant mortality in the United States, I will present evidence for both power-law *and* exponential behavior of the age trajectory of infant death. Both types of models are then synthesized into a *power-exponential hazard*, a generalization of both the power law and negative-Gompertz hazards, whose fit is evaluated on life tables for various cohorts and subpopulations of U.S. infants. A discussion of the generative mechanisms which could give rise to a power-exponential hazard concludes the paper.

2. Data

The "NCHS Cohort Linked Birth – Infant Death Data Files" (National Center for Health Statistics 2016) contain a complete census of births and infant deaths on the territory of the United States[3] and feature most fields present on the birth and infant death certificates. The size and detail of the data allow for the calculation of day-to-day infant life tables for various subpopulations.

I confine the analysis to the quinquennial birth cohorts 1995-1999 and 2005-2009 stratified by sex, prematurity, five-minute Apgar score (Apgar 1953), and social background of the mother. Those particular variables constitute major sources of heterogeneity in infant mortality along biological, medical, and social dimensions and are among the most reliably reported fields. In total, I analyze data on 277,004 infant deaths out of 40,727,055 births over the birth cohorts 1995–1999 and 2005–2009.

Female and male infant life tables are calculated for both birth cohorts, each life table stratified by either the five minute Apgar score (an ordinal measure of an infant's vitality shortly after delivery discretized into three levels), prematurity (defined via the age of gestation upon delivery and discretized into four levels), mother's education (three levels) or mother's ethnicity (four levels). Those 56 life tables thus cover heterogeneity in the age trajectory of infant mortality across period, medical and social strata.

[3] Available data on the overseas territories has been excluded due to compatibility issues.

3. Methods

To estimate the functional shape of infant mortality over the first year of life, I fit the power-exponential family of hazards to the day-to-day death counts and exposures for various populations of U.S. infants. Let $h(x)$ be the force of mortality (the hazard) at non-negative age x. The power-exponential family is given by the expression

$$h_{\mathrm{PE}}(x) = a(x + c)^{-p} e^{-bx},$$

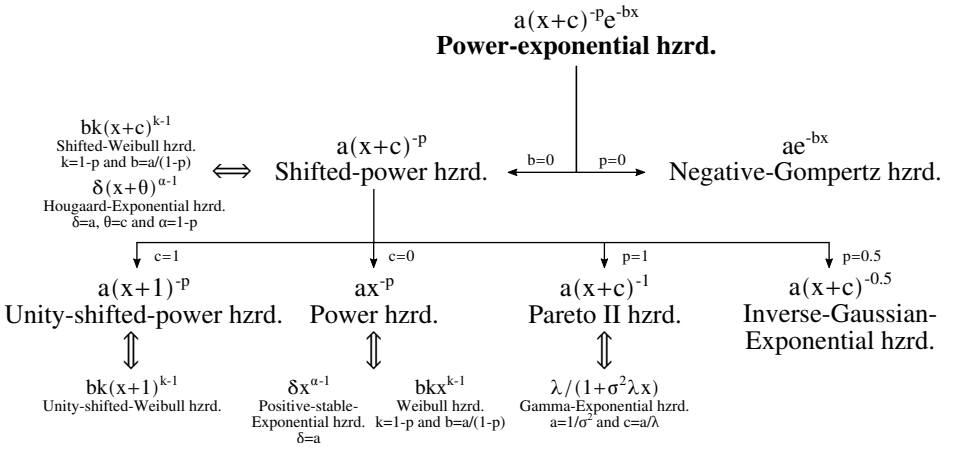
with $a \geq 0$, $p \geq 0$, $b \geq 0$, $c > 0$, and survivorship function

$$S_{\mathrm{PE}}(x) = e^{ab^{p-1}e^{bc}(\Gamma[1-p,b(c+x)]-\Gamma[1-p,bc])},$$

where Γ is the incomplete Gamma function (Abramowitz and Stegun 1964, 6.5.3).

Many of the models proposed in the literature on infant mortality age trajectories are contained within the power-exponential family as special cases, such as the hazard functions of the negative-Gompertz (Lomax 1954; Marshall and Olkin 2007, pp. 368) and the Pareto II distribution (Lomax 1954; Marshall and Olkin 2007, pp. 400), the exponential-Gamma frailty parametrization of the Pareto II distribution (Clayton and Cuzick 1985; Wienke 2011, p. 78), (shifted)-power hazards and their (shifted)-Weibull parametrizations (Lehman 1963; Marshall and Olkin 2007, pp. 321), and the Hougaard-Exponential frailty model (Hougaard 1984). Having all these models nested in a single family (Figure 1 for a family tree) facilitates goodness-of-fit comparisons: A simpler model may be rejected in favor of a more complicated one on the grounds of significance test on the parameters, and the whole family of models can be rejected if the full expression fails to describe the data accurately.

Figure 1: The power-exponential family of hazards considered for the age trajectory of infant mortality. A range of distributions previously applied to the survival analysis of infant death appear as special cases.



Prior to fitting the power-exponential model to the data, I will discuss the respective properties and parameter interpretation of the nested power and negative-Gompertz hazards and demonstrate their fit (or lack thereof) to the observed death rates of the complete cohort of infants born in the U.S. 2005 to 2009. This procedure builds evidence that it takes both a power- and an exponential component to describe mortality over infancy adequately, a subtlety that, to my best knowledge, has remained unmentioned in the literature.

In a second analysis, I test the fit of the power-exponential hazard by birth cohort, sex, Apgar score, gestation at birth, maternal origin, and education. Confronting a mortality model with such a varied collection of life tables acts as a validity test and allows for the quantification of heterogeneous mortality trajectories via comparison of model coefficients across population strata. Particular attention will be paid to non-proportional variations of the hazard, the hazard's behavior during the neonatal period (the first month of life), and the post-neonatal period (months one to twelve of age).

All the models in this paper are fitted as generalized linear models (GLMs). Such an approach guarantees concave likelihood surfaces and consequently stable estimation of the model parameters while also capitalizing on the wide availability of software to fit and evaluate GLMs. By using suitable link functions and transformations of the age variable, a range of parametric hazard functions can be linearized and thus fitted to observed age-specific death counts and exposures via a Poisson GLM (Aitkin and Clayton 1980; Clayton 1983; Currie 2016).

The power-exponential family of hazards, and therefore the distributions shown in Figure 1, may be fit, after deciding on the value for c, as Poisson GLMs with a log-link. This is possible due to the log-linearity of $h_{\text{PE}}(x)$ given c:

$$\log h_{\text{PE}}(x) = \log(a) - p\log(x + c) - bx = \beta_0 + \beta_1 \log(x + c) + \beta_2 x,$$

with $a = \exp(\beta_0)$, $p = -\beta_1$ and $b = -\beta_2$. The non-linear c parameter is estimated by maximizing the *profile likelihood* (McCullagh and Nelder 1989, chapter 7.2.4) of c over a range of GLM fits[4]. Model evaluation is performed via the model deviances, LOESS-smoothed Pearson-residuals[5], and McFadden's R^2 criterion (McFadden 1974). Confidence intervals around the parameter estimates of the power-exponential hazard are derived from 1,000 repeated fits of the model on 1,000 different parametric-bootstrap replicates of each life table. Unlike the asymptotic standard errors retrieved from the GLM fit, the bootstrap approach incorporates the uncertainty associated with the estimation of the non-linear c parameter.

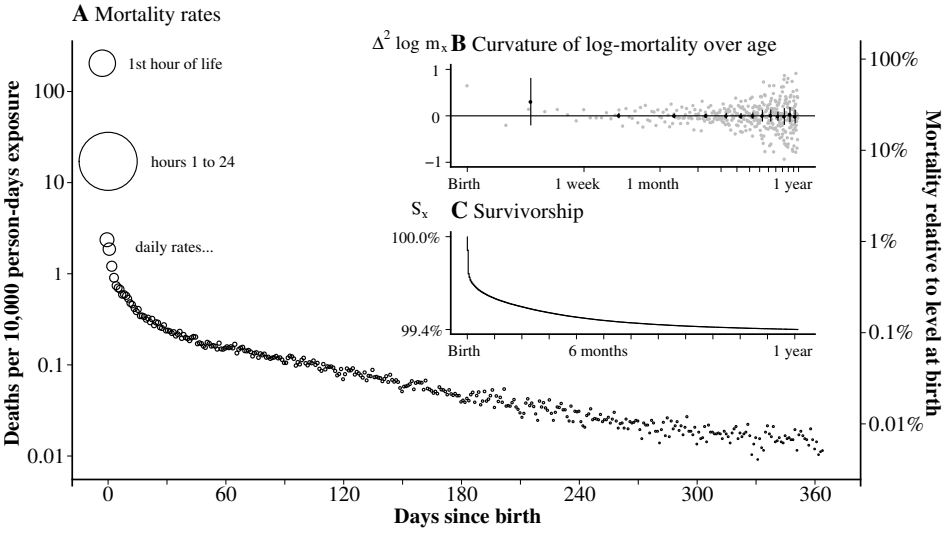
[4] Let $L(\boldsymbol{\beta}, c)$ be the full likelihood of a GLM fit with coefficients $\boldsymbol{\beta}$ and age-offset c. The profile likelihood of c is given by $L_p(c) = \max_{\boldsymbol{\beta}} L(\boldsymbol{\beta}, c)$ and c is estimated as $\hat{c} = \arg\max_c [\max_{\boldsymbol{\beta}} L(\boldsymbol{\beta}, c)]$. I found $L_p(c)$ to be concave over c and the maximization was quick and stable. See also Rinne (2009), chapter 11.3.2.2 for the closely related strategy of using the profile likelihood to determine the location parameter of a three parameter Weibull distribution.

[5] In order to smooth Pearson residuals over age x I use the `loess()` function implemented in R (R Core Team 2020) specified with quadratic polynomials, a tri-cubic weight function, Gaussian errors and a span of 0.4.

4. The age trajectory of U.S. infant mortality

The age trajectory of infant mortality, as observed in the U.S. birth cohort 2005-2009, is characterized by a peak at birth with a rapid decline thereafter (Figure 2A). While the mortality rate over the first hour following birth is around 200 deaths per 10,000 person-days of exposure, 24 hours later, the rate is at only 1.1 percent of its initial value and then again drops by a factor of ten over the next 29 days. Around one month after birth, mortality rates start to decline with a near-constant relative rate, approaching an exponential behavior. This log-linear trend is further illustrated by the second-differences of log-mortality rates over day-of-age, which have a mean close to zero for every month but the first (Figure 2C).

Figure 2: Infant life table measures for the U.S. birth cohort 2005-2009. A) Mortality rates drop rapidly during the neonatal period, after which the decline becomes exponential. Note that the area of the circles is proportional to the observed death counts. B) The exponential decline in mortality is further illustrated by the second differences in log-mortality over age clustering around zero. 95% CI via t-test of second differences aggregated over week one, weeks two to four and each of the following 11 months. C) As indicated by the survival curve, more than half of the infant deaths occur in the first few days following birth.

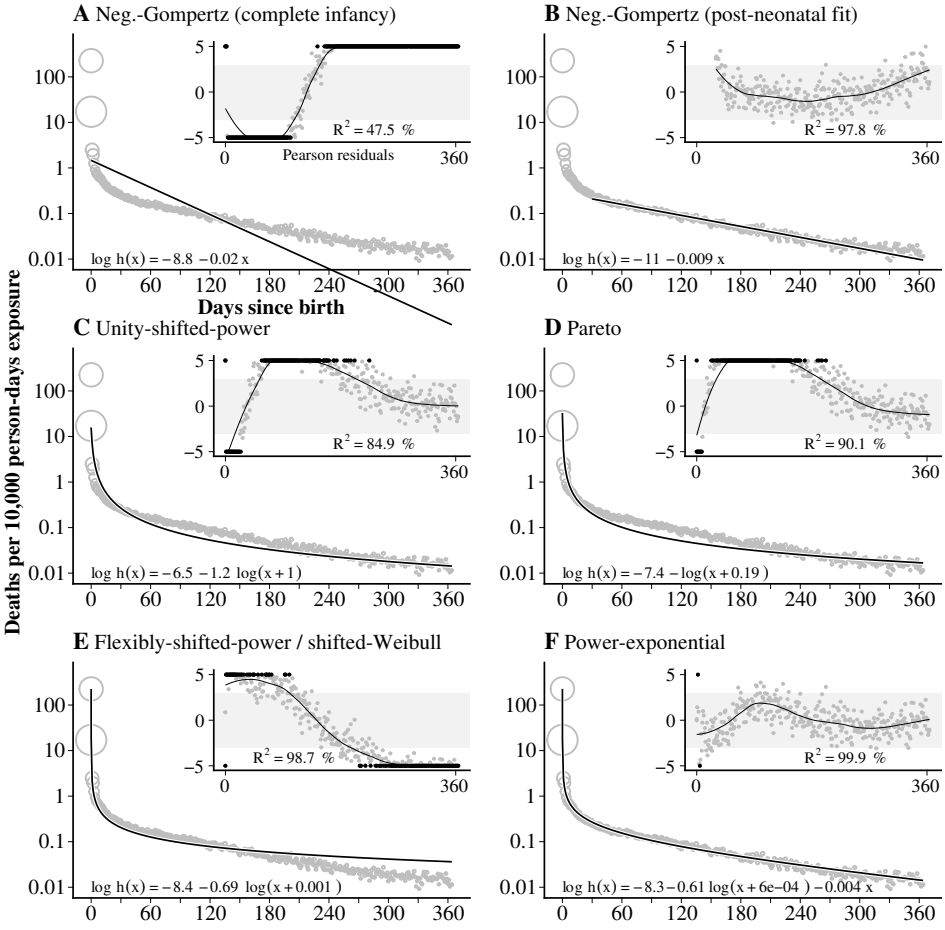


The pattern of a "super-exponential" decline in mortality during the neonatal period followed by an exponential tail over the remainder of the first year of life holds for girls and boys, irrespective of the social background of their mothers, the Apgar score upon birth, the gestation at birth or birth cohort (Figures 4 and 5).

Out of 162,541 infant deaths, 23,683 (14.5 percent) die during the first hour following birth and 81,506 (50.1 percent) before the age of 5 days. The overall probability of survival is 99.4 percent (Figure 2B).

5. Evidence for a power-exponential hazard

Figure 3: Fitted versus observed mortality over the first year of life for the U.S. birth cohort 2005-2009. Combining a power law (C, D, E) and a negative-Gompertz hazard (A, B) into the power-exponential model (F) accounts for the different behavior of the hazard in the neonatal and post-neonatal periods and achieves a close fit over the entire first year of life.



5.1 Negative-Gompertz hazard

In an early attempt to model mortality across the whole human lifespan, Thiele (1871) proposed to express the risk of death before maturity as an exponential function of age,

$$h_{\text{GP}}(x) = ae^{-bx},$$

with $a \geq 0$ and $a \geq 0$. This is the hazard function of a Gompertz distribution with a negative b parameter and therefore mirrors the shape of the hazard commonly assumed for adult humans. The simplicity of an exponential term for the hazard facilitates analytic treatment and gives the parameters a direct interpretation as important quantities in the study of infant mortality, $h_{\text{GP}}(0) = a$ being the hazard of death at the moment of birth and $-\frac{h'_{\text{GP}}(x)}{h_{\text{GP}}(x)} = b$ the instantaneous relative rate of mortality decline over age. The fact that the risk drops with a constant relative rate is the defining feature of the negative-Gompertz hazard. Related quantities are independent of age as well: the time t it takes for the hazard to drop by a factor k is the solution to equation $\frac{h_{\text{GP}}(x+t)}{h_{\text{GP}}(x)} = \frac{1}{k}$ given by $t_{1/k} = \frac{\log(k)}{b}$ and for each unit increase in age the hazard changes by a factor $\frac{h_{\text{GP}}(x+1)}{h_{\text{GP}}(x)} = \exp(-b)$.

The negative-Gompertz hazard can be expressed as a log-linear model with Poisson distributed age-specific death counts D_x as the outcome and age-specific exposure times E_x as a fixed offset[6]

$$\log \text{E}[D_x] = \beta_0 + \beta_1 x + \log E_x, \text{with } D \sim \text{Poisson}.$$

From this fit the parameters of h_{GP} are recovered as $a = \exp(\beta_0)$ and $b = -\beta_1$.

The negative-Gompertz model of infant mortality is especially popular in biology due to the influential paper by Siler (1979), who included it as part of his "Competing-Risk Model for Animal Mortality." In this model, a Gompertz term with negative b parameter is used "to account for the hazard due to immaturity" (Siler 1979). For life tables of human infants and children, however, the negative-Gompertz law has long been found to provide an insufficient fit (Thiele 1871, p. 326; Choe 1981; Gage and Dyke 1986).

Figure 3A clearly shows that the age trajectory of infant mortality for the 2005-2009 U.S. birth cohort deviates from a pure log-linear (i.e., exponential) form, ruling out the Gompertz hazard as a suitable model for the entire infant age range. *Yet*, a log-linear decline constitutes a near-perfect description of the mortality trajectory over the post-neonatal period (Figure 3B). The rate of death at day 30 after birth is around 2.03 per 100,000 person-days of exposure, with a subsequent decline of about 0.93 percent per additional day of age. Remarkably these two numbers explain 97.8 percent of the total deviance in the data during the post-neonatal period. Further investigating this result, I fit negative-Gompertz and power-law hazards to the post-neonatal life tables for each of the 60 monthly U.S. birth cohorts from January 2005 to December 2009 separate by sex. In most cases, the negative-Gompertz hazard provided a closer fit to the observed

[6] Assuming that the width of each age group $j = 1, \ldots, 366$ is small enough as not to introduce substantial aggregation bias. To simplify the notation I write D_x and E_x for total deaths and exposures in age groups with starting age x, where $x \in \{x_{j=1}, \ldots, x_{j=366}\}$.

Table 1: Number of months over the five year period 2005-2009 a negative-Gompertz fit achieved a lower deviance on the post-neonatal cohort life tables compared to a given power-law fit.

Gompertz vs.	Female	Male
Pareto II	44/60 73.3%	54/60 90%
Unity-shifted power	59/60 98.3%	60/60 100%
Flexibly-shifted power	23/60 38.3%	41/60 68.3%

post-neonatal mortality trajectory than either of the alternative two or three-parameter power-law hazards (Table 1).

5.2 Power-law hazard

Various power-laws have been specified to characterize the age-specific hazard of infant death (Oppermann 1870; Brillinger 1961; Choe 1981; de Beer and Janssen 2016; Berrut et al. 2016) all of which are variations on the basic expression

$$h_{\mathrm{PW}}(x) = ax^{-p},$$

with $a \geq 0$ and $p \geq 0$. Just like a negative-Gompertz hazard, a power-law hazard monotonically approaches 0 as $x \to \infty$; however, the relative rate of decline of a power hazard varies with age according to $-\frac{h'_{\mathrm{PW}}(x)}{h_{\mathrm{PW}}(x)} = \frac{p}{x}$, i.e., it is greatest right after birth and approaches 0 as age increases. It is this behavior that allows a power-law to initially decline faster than a Gompertz hazard, yet become slower for larger x, rendering it potentially useful for the description of the day-to-day age pattern of infant mortality.

Similar to the Gompertz case, the parameters of the power hazard can be interpreted in terms of mortality level and rate of change: The risk of death at age $x = 1$ is given by $h_{\mathrm{PW}}(1) = a$ and the proportional change in mortality for change in age by proportion w is $\frac{h_{\mathrm{PW}}(wx)}{h_{\mathrm{PW}}(x)} = w^{-p}$. The proportional increase in age w needed for the hazard to drop by factor k is the solution to equation $\frac{h_{\mathrm{PW}}(wx)}{h_{\mathrm{PW}}(x)} = 1/k$ given by $t_{1/k} = k^{\frac{1}{p}}$. The power parameter p by itself represents the elasticity of the hazard, $\frac{\mathrm{d}\log h_{\mathrm{PW}}(x)}{\mathrm{d}\log x} = -p$, i.e. for every infinitesimal proportional increase in age the hazard drops by proportion p. A convenient property of power-laws is that the elasticity is invariant to any rescaling of age in the form $x' = wx$. In practice this means that the units used for the age column of the infant life tables (e.g. hours, days, weeks) have no effect on the estimation of the exponent p.

Note that h_{PW} can be reparametrized into the hazard of the Weibull distribution $h_{\mathrm{UW}}(x) = bkx^{k-1}$ by substituting $p = 1 - k$ and $a = bk$.

Berrut et al. (2016) fit h_{PW} to death rates starting shortly after birth for a range of human and non-human populations and identify power-law behavior in most populations[7]. Specifically, they identify a segmented power-law relationship between age and death rate

[7] Note that while occurrence-exposure rates are used in this article, Berrut et al. (2016) use the total number of births in the denominator of their death-rates.

for a cohort of Swiss infants. The first segment starts one hour after birth and lasts for eight hours, while a different power coefficient is identified for the remainder of the first month of life.

Unity-shifted-power hazard

As x approaches 0, hazard h_{PW} approaches infinity. Due to this behavior, a pure power-law hazard can not be used to describe the age pattern of infant mortality starting from the moment of birth. Instead one can either choose to exclude the moment/hour/day of birth from the study period (as did Choe 1981; Berrut et al. 2016) or add a positive location parameter to h_{PW}, i.e. $h_{\mathrm{PW}}(x + c)$. The offset can either be estimated from the data – allowing for additional model flexibility – or set to some constant. The latter option has the advantage that the resulting hazard can be written as a fully linear function of log-mortality. For mathematical convenience one may use a unity offset resulting in the *unity-shifted-power* expression

$$h_{\mathrm{UP}}(x) = a(x + 1)^{-p},$$

with $a \geq 0$ and $p \geq 0$. Due to the unity offset, the risk of death at the moment of birth is $h_{\mathrm{UP}}(0) = a$. The elasticity of $h_{\mathrm{UP}}(x)$ approaches $-p$ as $x \to \infty$. In practice I found $-p$ to be a very close approximation to the true elasticity of $h_{\mathrm{UP}}(x)$ for at least the post-neonatal period; thus one may safely interpret p as the approximate proportional drop in the hazard of infant death for an infinitesimal proportional change in age starting at day 30 after birth.

The unity-shifted-power hazard can be fitted as the log-linear Poisson regression

$$\log \mathrm{E}[D_x] = \beta_0 + \beta_1 \log(x + 1) + \log E_x,$$

with $a = \exp(\beta_0)$ and $p = -\beta_1$.

Note that h_{UP} can be translated into the unity-shifted Weibull hazard $h_{\mathrm{UW}}(x) = bk(x+1)^{k-1}$ by substituting $k = 1 - p$ and $b = a/(1 - p)$.

While the unity-shifted-power law fits the 2005-2009 U.S. infant life table better than the Gompertz hazard, the Pearson residuals in Figure 3C still indicate systematic misspecification. Due to its very nature, the unity-shifted-power law can not capture the exponential portion of the hazard, and it also fails to adequately describe the swift drop in the risk of death over the first hours following birth.

Pareto hazard

In their model for the age pattern of human mortality de Beer and Janssen (2016) express the hazard of death during infancy and childhood as $h(x) = \frac{a}{c+x}$. This is the hazard function of a Pareto type II distribution (Lomax 1954; Marshall and Olkin 2007, pp. 400).

Rewriting the hazard reveals a shifted-power law with exponent -1, scaling factor a and location offset c,

$$h_{\mathrm{PT}}(x) = a(x + c)^{-1},$$

where $a \geq 0$ and $c > 0$. While for the unity-shifted-power-law, the location offset is fixed, and the power exponent varies, the situation is reversed for the Pareto II hazard with a fixed power and a variable location offset. As $h_{\mathrm{PT}}(x)$ is finite for $x > -c$ one may interpret the c parameter as the duration h_{PT} extends into the prenatal period. For such an interpretation to be justified the intrapartum death-rates (death during labor, i.e., at $x < 0$) have to follow the same functional form as the infant death rates and, given that $h(x + c)$ is declining with age, have to be higher than the mortality rates after birth. Of course the mortality at birth is $h_{\mathrm{PT}}(0) = a/c$ with a being the hazard level at $x = 1 - c$. The elasticity of the Pareto hazard is $\frac{\mathrm{d}\log h_{\mathrm{PT}}(x)}{\mathrm{d}\log x} = -\frac{x}{x+c}$ which approaches -1 as $x \to \infty$. Here a central characteristic of the Pareto hazard is implied: in the limit any proportional increase in age by factor w results in a proportional change of the hazard by $1/w$, e.g., doubling age halves the hazard. In practice, the estimated values for c are small enough for this limiting behavior to set in shortly after birth.

The Pareto hazard can be fitted via the Poisson regression

$$\log \mathrm{E}[D_x] = \beta_0 - \log(x + c) + \log E_x,$$

an intercept-only model with additional offset $-\log(x + c)$ and $a = \exp(\beta_0)$. As described earlier c is estimated by maximizing its profile likelihood.

Vaupel and Yashin (1983) propose to describe the infant and childhood component of human mortality by a Gamma-Exponential frailty model with population hazard function $h_{\mathrm{GE}}(x) = \lambda/(1 + \sigma^2\lambda x)$, i.e., a mixture of constant baseline hazards (corresponding to an exponential baseline distribution of deaths) with Gamma distributed rate parameter λ. As noted by e.g. Wienke (2011) this model is equivalent to h_{PT} when $a = 1/\sigma^2$, i.e., the inverse of the variance parameter of the Gamma-Exponential frailty model, and $c = a/\lambda$. Therefore h_{PT} may be interpreted as the population hazard resulting from mortality selection among individuals with constant hazards of Gamma varying magnitudes.

The fit of the Pareto hazard to the 2005-2009 U.S. infant life tables is inadequate and similar to that of the shifted-power-law (Figure 3D).

Flexibly-shifted-power hazard

Adding a positive location offset c to h_{PW} results in the flexibly-shifted-power hazard

$$h_{\mathrm{FP}}(x) = a(x + c)^{-p},$$

where $a \geq 0$, $b \geq 0$ and $c > 0$. This hazard shape contains the Pareto II hazard and the unity-shifted-power hazard as special cases. Note that h_{FP} can be interpreted as the

hazard function of the shifted Weibull distribution $h_{\mathrm{SW}}(x) = bk(x+c)^{k-1}$ with $k = 1-p$ and $b = a/(1-p)$.

The elasticity of the flexibly-shifted-power hazard is $\frac{\mathrm{d}\log h_{\mathrm{FP}}(x)}{\mathrm{d}\log x} = -\frac{px}{x+c}$ which approaches $-p$ as $x \to \infty$. Due to small estimates for c, this limiting behavior sets in shortly after birth, allowing an interpretation of the p parameter as the proportional drop in mortality for an infinitesimal proportional increase in age for all but the very first moments after birth. As for the Pareto II hazard, a is the hazard level at age $1-c$, and c may be interpreted as the time that h_{FP} extends into the prenatal period.

The flexibly-shifted-power hazard can be fitted as a Poisson GLM of the form

$$\log \mathrm{E}[D_x] = \beta_0 + \beta_1 \log(x+c) + \log E_x, \text{ with,}$$

with the original parameters recovered as $a = \exp(\beta_0)$ and $p = -\beta_1$. Again, c is estimated by maximizing its profile likelihood.

Substituting $a = \delta$, $c = \theta$ and $p = 1-\alpha$ recovers the hazard of the Hougaard-Exponential distribution $h_{\mathrm{HE}}(x) = \delta(x+\theta)^{\alpha-1}$, a continuous mixture of exponential distributions employed by Hougaard (1986) to describe the hazard in the days following myocardial infarction. The mixing distribution is flexible and contains the Gamma ($\alpha = 0$), the Inverse-Gaussian ($\alpha = 0.5$), and the positive-stable distribution ($\theta = 0$) as special cases. The Hougaard-Exponential hazard is a frailty model, albeit in a non-standard parametrization, with average frailty upon birth equal to $\delta\theta^{\alpha-1}$ and a baseline hazard fixed at unity (cp. Hougaard 1986, eq. 5.1). Because of the constant individual level hazard, all of the declining mortality observed on the population level is assumed to be due to mortality selection.

Allowing for both a free power parameter p and a free location offset c greatly improves the fit to the U.S. 2005-2009 infant life tables compared to the more restricted unity-shifted-power and Pareto hazards (Figure 3E). The estimate for c is 0.0012 (i.e. a location shift by ≈ 1.7 minutes of age) and for p equals 0.69, corresponding to an approximate drop in mortality by $(1-2^{-0.69}) \times 100 = 38$ percent for every doubling of age. While fitting better than the hazards discussed above, the Pearson residuals in Figure 3E still exhibit a clear trend signifying a systematic lack of fit.

5.3 Power-exponential hazard

Multiplying the flexibly-shifted-power law hazard h_{FP} with an exponential term results in the power-exponential expression

$$h_{\mathrm{PE}}(x) = a(x+c)^{-p}e^{-bx},$$

with $a \geq 0$, $p \geq 0$, $b \geq 0$, and $c > 0$. This hazard contains all of the above models as special cases (Figure 1).

The power-exponential hazard smoothly transitions from power-law to exponential behavior: With increasing age h_{PE} approaches a constant relative rate of change and thus increasingly resembles a negative-Gompertz hazard, formally $\lim_{x\to\infty} \frac{h'_{\text{PE}}(x)}{h_{\text{PE}}(x)} = -b$.
The Poisson GLM form of the power-exponential hazard is

$$\log \mathrm{E}[D_x] = \beta_0 + \beta_1 \log(x + c) + \beta_2 x + \log E_x,$$

where $a = \exp(\beta_0)$, $p = -\beta_1$, $b = -\beta_2$ and the non-linear coefficient c is estimated via profile likelihood maximization.
Adding the exponential term to h_{FP} significantly ($p < 0.001$ via deviance-ratio test, Table 2) improves the fit of the Poisson GLM to the 2005-2009 U.S. birth cohort and flattens the trend in the Pearson residuals over age (Figure 3F) albeit a hump-shaped curvature remains with a peak around 90 days of age. A likely source for this residual pattern is Sudden Infant Death, which exhibits an incidence-hump with a peak at around two to four months of age (Kinney and Thach 2009).

6. Evaluation of the power-exponential hazard

The power-exponential hazard achieves an excellent fit irrespective of cohort, sex, five minute Apgar score, gestational age at birth, maternal origin or education. The percentage of deviance explained by the model ranges from 94.5 to 99.6 percent in all life tables under consideration. For every single population, the inclusion of an exponential term in addition to the shifted-power term significantly ($p < 0.001$ via deviance-ratio test, Table 2) improves the fit, reducing the residual deviance by more than 50% on 43 out of the 56 populations.
The shape of the hazard function over age varies most strongly by Apgar score (Figure 4A) with the exponential behavior most pronounced among infants born with a score of nine or ten (the majority). Within two weeks after delivery, the initial mortality spike at birth transitions into a log-linear decline in mortality risk of around 0.7 percent per day. Conversely, the hazard trajectory for infants with a "low" Apgar score (indicating health problems of upon delivery), while still exponential in the tail, features a more gradual transition into log-linear behavior. The most considerable difference in the behavior of the hazard among the Apgar groups can be seen over the first week of life. While for Apgar group zero to five, the hazard drops by a factor of 1000 from the day of birth to an age of seven days, it only drops by a factor of nine or less for the other groups (Table 3). Thus hazards vary by Apgar score in a highly *non-proportional* fashion as can also be inferred from the varying power (p), exponential-rate (b), and offset (c) parameters of the fitted models. Non-proportional behavior, albeit less pronounced, is also evident when comparing hazards by prematurity, with higher values for the power parameter $p = -\beta_1$ for lower ages of gestation at birth (Figure 4B), a result also observed by Berrut et al. (2016) for Swiss and Norwegian infants after fitting a simple power-law to age-specific mortality rates.

Table 2: Percent reduction in residual deviance after multiplying the flexibly-shifted-power-law hazard with an exponential term. All results significant at $p < 0.001$ (via deviance-ratio tests).

		1995-1999		2005-2009	
		Female	**Male**	**Female**	**Male**
APGAR	Very low [0,5)	50.2	45.6	61.3	60.3
	Low [5,9)	20.6	27.9	49.5	49.7
	Regular 9+	86.8	86.4	89.8	92.0
Origin	Non-Hispanic Black	56.0	60.1	65.1	63.9
	Non-Hispanic White	76.3	74.2	81.9	79.6
	Hispanic	55.4	63.6	60.2	69.4
	Other	20.3	29.4	24.7	33.5
Prematurity	Extremely preterm <28w	67.4	67.1	74.1	75.7
	Very preterm [28,32)w	54.7	56.1	60.9	63.0
	Moderate to late preterm [32,37)w	58.2	60.0	71.0	70.2
	Term 37w+	79.8	78.1	83.7	80.9
Education	Elementary or less	38.6	41.9	30.5	40.9
	High school	76.9	74.7	79.9	77.3
	College or university	67.3	66.0	76.6	76.8

The relative rate of mortality decline during the post-neonatal period, as approximated by b, is higher for premature infants and infants with a low Apgar score compared to infants born at term or with a regular Apgar score, implying that the relative difference in mortality between very frail and less frail infants diminishes over age.

Hazards are mostly proportional by ethnicity of the mother (Figure 5B) with the risk of death during the later stages of infancy declining exponentially with a rate of 0.3 to 0.5 percent per day of age and the power parameter ranging from -0.51 to -0.55 across the life table strata. Notably, during the post-neonatal period mortality over age declines slower for children of African-American mothers compared to infants of white mothers[8]; taken together with the observation that the intercept of the hazard curve is highest among infants of African-American mothers, this puts the group at a double disadvantage. Along education strata, the hazards are proportional (Figure 5B).

There are no systematic differences in the age trajectory of infant mortality between the sexes, apart from a proportionally higher hazard for males than females. Similarly, the difference between the two birth cohorts is mostly proportional, with the younger cohort having a lower intercept β_0 (Figures 4 and 5).

[8] Significant at $p < 0.05$ for female and male infants of either birth cohort (Table 5).

Figure 4: Age specific hazard of death as predicted by the power-exponential hazard contrasted with life table mortality rates by sex, birth cohort, and condition of the infant upon birth.

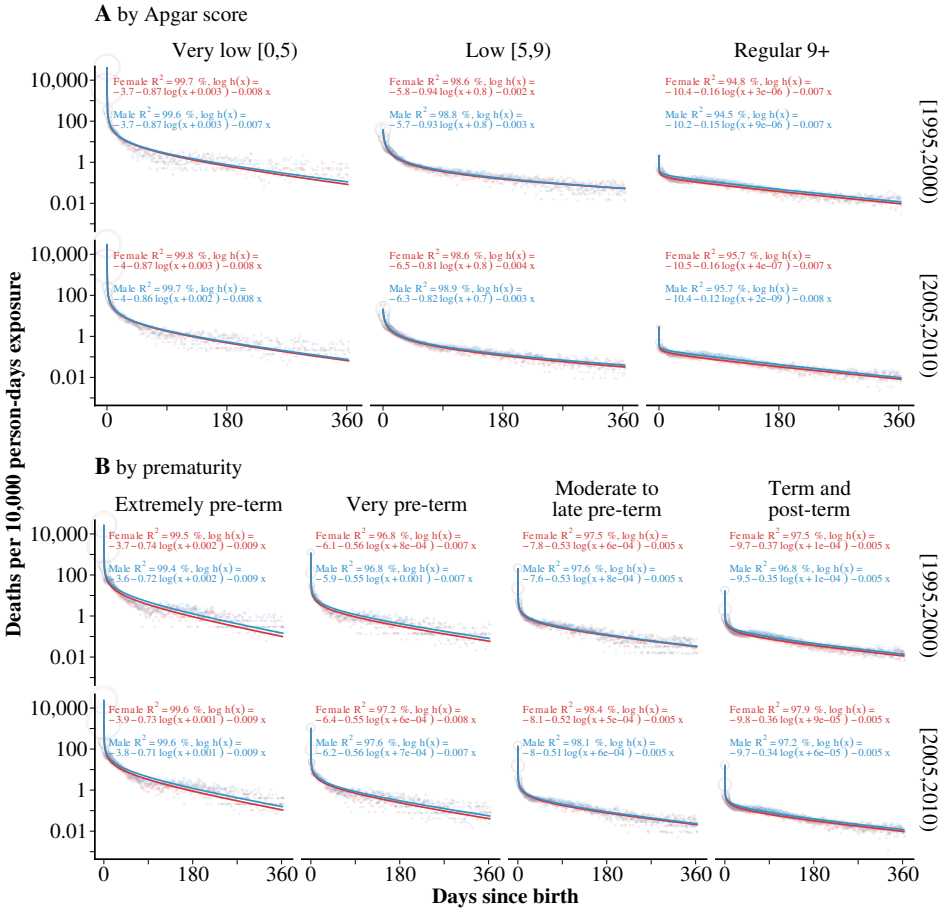
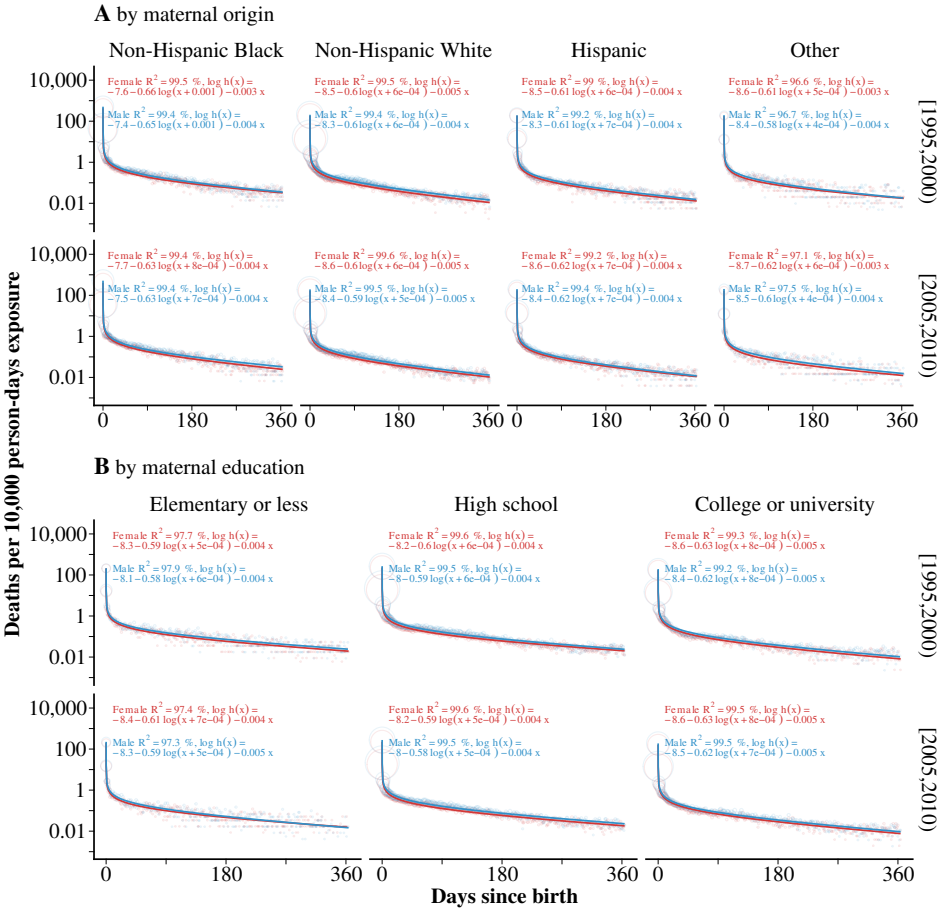


Figure 5: Age specific hazard of death as predicted by the power-exponential hazard contrasted with life table mortality rates by sex, birth cohort, and maternal background.



7. Interpretation of the exponential-power hazard

7.1 A shock-recovery process

The power-exponential product in the population hazard h_{PE} can be interpreted as a *non-homogeneous split Poisson process*, where *shocks* to an infant's health arrive with rate $\lambda(x)$ per unit person-time, each shock resulting in infant death with probability $p(x)$. Shock models in the context of human mortality have, for example, been studied by Strehler and Mildvan (1960), Finkelstein (2005), Cha and Finkelstein (2016). A similar explanation for the age trajectory of preadolescent mortality across species has been put forward by Levitis (2011) under the name "transitional timing hypothesis," stating that *"transcriptional, developmental and environmental transitions are dangerous, and these are concentrated early in life."*

Let $N(x)$ be the number of infants alive at age x and let $\mathrm{E}[Y]$ be the expected value of a Poisson distributed random variable Y with rate parameter $\int_x^{x+n} \lambda(s)N(s)\,\mathrm{d}s$ representing the total number of *health-shocks* the population of infants is expected to experience over age interval $[x, x+n)$. If each shock leads to death with probability $p(x)$ then the number of deaths D over age interval $[x, x+n)$ follows a Poisson distribution with expected value

$$\mathrm{E}[{}_nD_x] = \int_x^{x+n} \lambda(s)p(s)N(s)\,\mathrm{d}s,$$

see Prékopa (1958) for a proof. The hazard of death experienced by survivors N at time x is $h(x) = \lambda(x)p(x)$. If the rate of shocks $\lambda(x)$ varies over time according to a flexibly-shifted-power hazard and if the probability of a shock leading to death $p(x)$ is exponentially declining the power-exponential hazard is recovered. Note that neither the rate of shocks nor the probability of death following a shock are identified as this would require inferring the model $h(x) = a_1(x+c)^{-p} \times a_2e^{-bx}$ from the fit $h(x) = a(x+c)^{-p}e^{-bx}$ – a problem with infinitely many solutions. However, the power and the exponential rate parameters p and b are completely identified. They can be interpreted as follows: For the extremely preterm female births of the U.S. birth cohort 1995-1999 (Figure 3F) the rate of shocks, $\lambda(x) \propto (x+0.002)^{-0.74}$, declined rapidly in the vicinity of birth by approximately $(1-2^{-0.74}) \times 100 = 40$ percent for every doubling of age whereas the probability of a shock leading to death, $p(x) \propto e^{-bx}$, declined by about 0.9 percent per additional day of age.

Why use the power law for the rate of shocks and the exponential term for the shock lethality instead of the other way around – after all, both options lead to the same expression for $h(x)$? I find it likely that the rapid neonatal mortality decline is predominantly a result of the waning stresses of birth – a sudden transition which may inflict a series of "shocks" to the infant – and not the result of a fast decline in the mortality risk associated with each shock. Furthermore, as $\lim_{x\to\infty} \frac{h'_{\text{FP}}(x)}{h_{\text{FP}}(x)} = 0$, the power-law approaches a constant hazard as age increases, allowing it to capture the age-independent rate of accidents which may dominate the rate of shocks later in infancy.

A particular shock-recovery process is described by the Strehler-Mildvan model of senescent mortality. Strehler and Mildvan (1960) propose a model where individuals experience

shocks to their health at a constant rate α. The magnitude of each shock, ranging from mild to severe, is a value drawn from an exponential random variable with rate β. The shocks are counteracted by the vitality of an individual, which is a positive and declining function of age. When the magnitude of a shock at age x exceeds the individual's vitality at that age, death occurs. Strehler and Mildvan (1960) show that this process leads to a Gompertz distribution of life-times with corresponding hazard function $h(x) = \alpha e^{\beta x}$.

The power-exponential hazard of infant mortality can be derived from such a Strehler-Mildvan process by changing two assumptions: 1) Where Strehler and Mildvan (1960) assume a decrease of vitality with time due to aging, one assumes an increase due to growth, and 2) instead of that shocks arrive with a constant rate α, one assumes that the risk of experiencing complications is highest at birth and falls over time (according to a shifted Weibull distribution).

7.2 A mortality selection process

In the multiplicative frailty model (Vaupel et al. 1979), it is assumed that all individuals in a population share the same age-specific "baseline hazard" of death $h_0(x)$ but on different "frailty" levels z which act multiplicatively on the baseline. Consequently, the age-specific hazard conditioned on frailty is given by

$$h(x|z) = zh_0(x).$$

Frailty is treated as a random variable Z with density $f(z|x)$ at age x. Integrating out the frailty yields the expression for the hazard observed at the population level, which is a mixture of the individual level hazards weighted by the age-specific distribution of frailties

$$\overline{h}(x) = \int_0^\infty h(x|z)f(z|x)\,\mathrm{d}z,$$

which may be re-expressed as

$$\overline{h}(x) = h_0(x)\int_0^\infty zf(z|x)\,\mathrm{d}z = h_0(x)\mathrm{E}[Z|x],$$

showing that the population hazard is a product of the baseline hazard and the average frailty among the population at age x. By choosing $h_0(x) = ae^{-bx}$ and $\mathrm{E}[Z|x] = (x+c)^{-p}$ we express the power-exponential hazard h_{PE} as a multiplicative frailty model.

Frailty models can be interpreted in terms of *mortality selection*: Because individuals with high frailty on average die earlier than those with lower values for z, the average frailty among a cohort of individuals declines over time. Above, we modeled this decline as a shifted-power function of age, and consequently, p can be interpreted as the approximate elasticity of average frailty w.r.t. age, 0.69 for the U.S. birth cohort 2005-1999, corresponding to an approximate drop in average frailty by 38 percent for every doubling of age – a substantial mortality selection shortly after birth.

While differing in frailty, all individuals of a cohort are assumed to share the same baseline hazard. By modeling this *individual level* risk as an exponential function of age, we assume that each additional unit of age decreases an infants risk of death by a factor $\exp(-b)$, around 0.4 percent per additional day of age for the 2005-1999 U.S. birth cohort (Figure 3F). This relatively slow rate of decline may be attributed to *acquired robustness* (Levitis 2011) due to the infant's continued growth and development.

Hougaard (1986) describes a class of multiplicative frailty models that exhibit a shifted-power law decline in average frailty over age. Frailty is assumed to be distributed according to a three-parameter extension of the stable distributions $f_Z(z;\alpha,\delta,\theta)$, a density without closed-form representation, but with closed-form Laplace transform

$$\mathcal{L}\{f_Z\}(s) = \mathrm{E}\left[e^{-sZ}\right] = \exp\left[-\frac{\delta}{\alpha}\left[(\theta+s)^\alpha - \theta^\alpha\right]\right].$$

Substituting the cumulative baseline hazard $H_0(x) = \int_0^x h_0(s)\,\mathrm{d}s$ for argument s one recovers the population survival function

$$\overline{S}(x) = \mathrm{E}\left[e^{-H_0(x)Z}\right] = \exp\left[-\frac{\delta}{\alpha}\left[(\theta+H_0(x))^\alpha - \theta^\alpha\right]\right]$$

and corresponding population hazard

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}x}\log\overline{S}(x) &= \overline{h}(x) \\ &= \mathrm{E}[Z|x]h_0(x) \\ &= \delta(\theta+H_0(x))^{\alpha-1}h_0(x). \end{aligned}$$

Assuming a negative-Gompertz baseline hazard $h_0 = \exp(-bx)$ and substituting $a = \delta$, $c = \theta$ and $p = -\alpha + 1$ into $\overline{h}(x)$ yields the Hougaard-Negative-Gompertz hazard

$$h_{\mathrm{HG}}(x) = a\left(\frac{1-e^{-bx}}{b} + c\right)^{-p} e^{-bx}.$$

This hazard has the same qualitative behavior as the power-exponential expression, namely a power-law decline shortly after birth and a gradual transition into an exponential tail. Both models achieve a virtually identical fit to the U.S. data. Under a frailty interpretation, both hazards lead to the same inference: mortality selection explains most of the mortality decline over the first month of life. The estimated $c = \theta$ parameter in the neighborhood of 0 approaches a positive-stable distribution of frailties, which has extreme positive skewness and exhibits infinite mean and variance.

Of course one could, to the same effect, also choose $h_0(x) = a(x+c)^{-p}$ and $\mathrm{E}[Z|x] = e^{-bx}$. This model implies that the power-law behavior of the population hazard (mostly seen during the first month of life) arises from individual-level processes, while above, we have assumed that the rapid mortality decline after birth is mostly due to changing average frailty, i.e., mortality selection.

Mortality selection can also be modeled via finite mixtures (discrete frailty models) of the form $F_{\mathrm{FM}}(x) = p_1 F_1(x) + p_2 F_2(x)$, where $F_{\mathrm{FM}}(x)$ is the probability of death until age x, p_1 and $p_2 = 1 - p_1$ are the relative proportions of population sub-groups 1 and 2 at birth and F_1, F_2 denote the corresponding stratum-specific cumulated probabilities of death. Choosing the form of a negative-Gompertz distribution for F_1 and F_2, yet allowing the parameters to differ between both populations, can produce hazards that smoothly transition into an exponential tail similar to the power-exponential expression (e.g., Marshall and Olkin 2007, cp. 3)[9]. Different choices for the baseline hazard and the distribution of frailties can result in qualitatively or mathematically equivalent expression for the population hazard – an identifiability problem well-known in the frailty-model literature (e.g., Hougaard 1995).

8. Discussion

Using highly detailed individual-level data on the timing of death over the first year of life, I found strong evidence for an exponentially modulated power-law behavior of the hazard of infant death in two recent U.S. birth cohorts. The shift from a power-law regime to an exponential decline has not previously been noted in the literature and invites speculation regarding the mechanisms giving rise to such a pattern. Observed hazards may be the result of mechanisms that have been discussed at length in the context of senescent risk of death: mortality selection due to heterogeneous frailties and a shock-recovery process.

The frailty hypothesis may be tested via a decomposition analysis. Vaupel and Zhang (2010) prove that, if the population hazard $\overline{h}(x)$ and the stratum specific hazards $h(x|z)$ are known for a cohort, then the rate of change over age of $\overline{h}(x)$ can be decomposed into a "direct change" component and change due to the process of mortality selection. Using this decomposition one could, for example, calculate how much of the decline in the risk of death over the first week of life as shown in Figure 2 is due to the changing population composition by prematurity status resulting from the early death of extremely premature infants, i.e., mortality selection.

A test of the shock-recovery model is less straightforward. It requires both a clear definition of what constitutes a shock and data on the timing of such shocks to see if their rate of occurrence corresponds to either the exponential or the power-law component of the power-exponential hazard.

Note that the power-exponential hazard discussed in this paper does not permit an interpretation as a competing-risks model. Such a model was implied by Bourgeois-Pichat (1951) for the age trajectory of infant mortality, which he partitioned into deaths due to "intrinsic" and "extrinsic" causes. An alternative competing-risks model featuring a power-law and an exponential term could take the form $h_{\mathrm{CR}}(x) = a_1(x + c)^{-p} + a_2 \exp(-bx)$, which may be re-written as $h_{\mathrm{CR}}(x) = \exp[\log a_1 - p \log(x + c)] + \exp[\log a_2 - bx]$. Sums-of-exponentials can be fit in the GLM framework by using a composite-link-function

[9] Adopting this strategy and increasing the number of subpopulations one can approximate the hazard trajectory of human mortality over the entire lifespan, as demonstrated by Avraam et al. (2013), Avraam et al. (2014).

(Thompson and Baker 1981; Camarda et al. 2016). Assuming that deaths due to intrinsic causes follow a power-law behavior, whereas extrinsic deaths feature the hazard of a negative-Gompertz distribution, this competing-risks formulation of the age-specific hazard of infant death over time permits a shape-based cause-of-death decomposition.

The results of this paper indicate that the proportional hazards assumption does not hold across Apgar score and prematurity strata. The association of those variables with the baseline hazard of death over the first year of life is highly non-proportional with much of the "effect" focused on the first weeks of life. Indeed differences in the rate of post-neonatal mortality decline imply that the relative difference in mortality between very frail and less frail infants diminishes over age to some degree.

It must be noted that without further study, the results of this paper can not be generalized to populations other than present-day U.S. infants. While similar results for countries with similar overall levels of infant mortality (implying a similar level of development) are to be expected, it would be foolish to assume the same age pattern of infant mortality in pre-20th century populations or present populations suffering under a crisis: Historically the age-pattern of infant mortality was very much shaped by the interaction of seasonality effects, improper substitutes for breastfeeding, and deadly infectious diseases (Knodel and Kintner 1977; Huck 1995), factors which have lost relevance in the U.S. over the course of the 20th century. Likewise, an analysis of 21st century U.S. infant mortality yields little insight into the characteristics of infant death in parts of the world suffering from humanitarian crises and violent conflicts.

Insofar as the power-exponential hazard *does* generalize to populations other than the present-day U.S., it can become a valuable tool for working with low-quality data on the timing of infant deaths. Based on h_{PE}, one can calculate life table a_0 from coarse data; design model infant-life tables; smooth over data artifacts such as age-heaping; estimate seasonal effects on mortality separately from age effects, or include realistic infant mortality schedules into a simulation model.

A. Parameter estimates and confidence intervals for the power-exponential hazard

Table 3: The factor of mortality reduction over the first 7 days of life. Parameter estimates and 95% CIs calculated from the power-exponential hazard GLM fits for various U.S. infant cohorts.

		Female	**Male**
1995-1999			
APGAR	Very low [0,5)	1011 (826, 1242)	962 (810, 1157)
	Low [5,9)	8 (6, 10)	8 (6, 10)
	Regular 9+	10 (7, 14)	8 (5, 10)
Origin	Non-Hispanic Black	328 (282, 385)	315 (277, 357)
	Non-Hispanic White	303 (272, 340)	275 (249, 304)
	Hispanic	320 (262, 391)	277 (229, 331)
	Other	329 (228, 498)	303 (219, 444)
Prematurity	Extremely preterm <28w	524 (449, 604)	448 (399, 508)
	Very preterm [28,32)w	181 (135, 246)	145 (113, 188)
	Moderate to late preterm [32,37)w	147 (116, 185)	127 (103, 158)
	Term 37w+	58 (49, 67)	49 (43, 57)
Education	Elementary or less	281 (203, 385)	236 (178, 316)
	High school	280 (254, 312)	256 (232, 281)
	College or university	330 (286, 377)	303 (268, 345)
2000-2005			
APGAR	Very low [0,5)	1039 (866, 1241)	1008 (869, 1193)
	Low [5,9)	6 (5, 8)	7 (5, 8)
	Regular 9+	14 (11, 18)	13 (10, 16)
Origin	Non-Hispanic Black	318 (272, 373)	340 (299, 391)
	Non-Hispanic White	301 (268, 338)	277 (250, 304)
	Hispanic	326 (278, 389)	310 (268, 359)
	Other	348 (243, 508)	347 (258, 472)
Prematurity	Extremely preterm <28w	538 (471, 613)	477 (427, 531)
	Very preterm [28,32)w	194 (152, 249)	178 (141, 225)
	Moderate to late preterm [32,37)w	142 (115, 175)	119 (98, 144)
	Term 37w+	60 (51, 69)	58 (50, 66)
Education	Elementary or less	310 (221, 436)	290 (218, 397)
	High school	280 (248, 312)	272 (248, 300)
	College or university	320 (285, 366)	315 (282, 351)

Table 4: Estimated a parameters of power-exponential hazard GLM fits for various U.S. infant cohorts. Mean estimates and 95% confidence intervals are based on 1000 parametric bootstrap replications.

		Female	**Male**
1995-1999			
APGAR	Very low [0,5)	2.5e-2 (2.5e-2, 2.6e-2)	2.6e-2 (2.5e-2, 2.6e-2)
	Low [5,9)	2.9e-3 (2.7e-3, 3.2e-3)	3.3e-3 (3.1e-3, 3.6e-3)
	Regular 9+	3.0e-5 (2.9e-5, 3.2e-5)	3.8e-5 (3.6e-5, 3.9e-5)
Origin	Non-Hispanic Black	4.8e-4 (4.7e-4, 4.9e-4)	5.8e-4 (5.7e-4, 5.9e-4)
	Non-Hispanic White	2.0e-4 (1.9e-4, 2.0e-4)	2.5e-4 (2.4e-4, 2.5e-4)
	Hispanic	2.0e-4 (1.9e-4, 2.0e-4)	2.4e-4 (2.3e-4, 2.4e-4)
	Other	1.8e-4 (1.8e-4, 1.9e-4)	2.1e-4 (2.1e-4, 2.2e-4)
Prematurity	Extremely preterm <28w	2.4e-2 (2.3e-2, 2.4e-2)	2.8e-2 (2.7e-2, 2.8e-2)
	Very preterm [28,32)w	2.2e-3 (2.1e-3, 2.3e-3)	2.7e-3 (2.6e-3, 2.8e-3)
	Moderate to late preterm [32,37)w	4.3e-4 (4.1e-4, 4.4e-4)	4.9e-4 (4.8e-4, 5.0e-4)
	Term 37w+	6.3e-5 (6.2e-5, 6.5e-5)	7.5e-5 (7.3e-5, 7.6e-5)
Education	Elementary or less	2.5e-4 (2.4e-4, 2.6e-4)	2.9e-4 (2.8e-4, 3.1e-4)
	High school	2.7e-4 (2.7e-4, 2.8e-4)	3.4e-4 (3.4e-4, 3.4e-4)
	College or university	1.9e-4 (1.8e-4, 1.9e-4)	2.2e-4 (2.2e-4, 2.3e-4)
2000-2005			
APGAR	Very low [0,5)	1.8e-2 (1.8e-2, 1.8e-2)	1.8e-2 (1.7e-2, 1.8e-2)
	Low [5,9)	1.6e-3 (1.4e-3, 1.7e-3)	1.8e-3 (1.7e-3, 2.0e-3)
	Regular 9+	2.7e-5 (2.6e-5, 2.8e-5)	3.1e-5 (3.0e-5, 3.2e-5)
Origin	Non-Hispanic Black	4.5e-4 (4.4e-4, 4.6e-4)	5.4e-4 (5.3e-4, 5.5e-4)
	Non-Hispanic White	1.9e-4 (1.8e-4, 1.9e-4)	2.3e-4 (2.3e-4, 2.3e-4)
	Hispanic	1.8e-4 (1.8e-4, 1.9e-4)	2.2e-4 (2.2e-4, 2.3e-4)
	Other	1.7e-4 (1.6e-4, 1.8e-4)	2.0e-4 (1.9e-4, 2.1e-4)
Prematurity	Extremely preterm <28w	2.0e-2 (1.9e-2, 2.0e-2)	2.3e-2 (2.3e-2, 2.3e-2)
	Very preterm [28,32)w	1.7e-3 (1.6e-3, 1.8e-3)	2.0e-3 (1.9e-3, 2.1e-3)
	Moderate to late preterm [32,37)w	3.0e-4 (2.9e-4, 3.0e-4)	3.2e-4 (3.2e-4, 3.3e-4)
	Term 37w+	5.3e-5 (5.2e-5, 5.5e-5)	6.1e-5 (6.0e-5, 6.3e-5)
Education	Elementary or less	2.2e-4 (2.1e-4, 2.3e-4)	2.6e-4 (2.4e-4, 2.7e-4)
	High school	2.6e-4 (2.6e-4, 2.7e-4)	3.3e-4 (3.2e-4, 3.3e-4)
	College or university	1.8e-4 (1.7e-4, 1.8e-4)	2.1e-4 (2.1e-4, 2.1e-4)

Table 5: Estimated b parameters of power-exponential hazard GLM fits for various U.S. infant cohorts. Mean estimates and 95% confidence intervals are based on 1000 parametric bootstrap replications.

		Female	**Male**
1995-1999			
APGAR	Very low [0,5)	8.0e-3 (7.1e-3, 8.8e-3)	7.3e-3 (6.7e-3, 8.1e-3)
	Low [5,9)	2.2e-3 (1.7e-3, 2.7e-3)	2.5e-3 (2.1e-3, 3.0e-3)
	Regular 9+	7.0e-3 (6.7e-3, 7.2e-3)	7.2e-3 (6.9e-3, 7.4e-3)
Origin	Non-Hispanic Black	3.1e-3 (2.8e-3, 3.4e-3)	3.6e-3 (3.3e-3, 3.8e-3)
	Non-Hispanic White	4.5e-3 (4.3e-3, 4.8e-3)	4.5e-3 (4.3e-3, 4.7e-3)
	Hispanic	3.9e-3 (3.5e-3, 4.3e-3)	4.1e-3 (3.8e-3, 4.4e-3)
	Other	3.0e-3 (2.3e-3, 3.7e-3)	3.7e-3 (3.1e-3, 4.3e-3)
Prematurity	Extremely preterm <28w	9.4e-3 (8.9e-3, 9.9e-3)	9.2e-3 (8.8e-3, 9.6e-3)
	Very preterm [28,32)w	7.2e-3 (6.5e-3, 7.8e-3)	7.0e-3 (6.5e-3, 7.6e-3)
	Moderate to late preterm [32,37)w	5.0e-3 (4.6e-3, 5.4e-3)	5.1e-3 (4.8e-3, 5.6e-3)
	Term 37w+	5.1e-3 (4.9e-3, 5.3e-3)	5.3e-3 (5.1e-3, 5.5e-3)
Education	Elementary or less	3.8e-3 (3.2e-3, 4.3e-3)	3.8e-3 (3.3e-3, 4.3e-3)
	High school	3.7e-3 (3.5e-3, 3.9e-3)	4.0e-3 (3.8e-3, 4.2e-3)
	College or university	4.7e-3 (4.4e-3, 4.9e-3)	4.7e-3 (4.4e-3, 5.0e-3)
2000-2005			
APGAR	Very low [0,5)	7.7e-3 (7.0e-3, 8.4e-3)	7.6e-3 (7.1e-3, 8.1e-3)
	Low [5,9)	3.9e-3 (3.4e-3, 4.4e-3)	3.5e-3 (3.1e-3, 3.9e-3)
	Regular 9+	7.1e-3 (6.8e-3, 7.3e-3)	7.7e-3 (7.5e-3, 7.9e-3)
Origin	Non-Hispanic Black	4.1e-3 (3.8e-3, 4.3e-3)	3.8e-3 (3.5e-3, 4.0e-3)
	Non-Hispanic White	4.5e-3 (4.3e-3, 4.7e-3)	4.7e-3 (4.5e-3, 4.9e-3)
	Hispanic	4.0e-3 (3.6e-3, 4.3e-3)	4.4e-3 (4.1e-3, 4.7e-3)
	Other	3.4e-3 (2.7e-3, 4.1e-3)	3.7e-3 (3.1e-3, 4.3e-3)
Prematurity	Extremely preterm <28w	8.8e-3 (8.3e-3, 9.3e-3)	8.6e-3 (8.3e-3, 9.0e-3)
	Very preterm [28,32)w	7.7e-3 (7.1e-3, 8.3e-3)	7.2e-3 (6.7e-3, 7.8e-3)
	Moderate to late preterm [32,37)w	5.3e-3 (4.9e-3, 5.7e-3)	5.4e-3 (5.1e-3, 5.7e-3)
	Term 37w+	5.2e-3 (5.0e-3, 5.4e-3)	5.3e-3 (5.1e-3, 5.5e-3)
Education	Elementary or less	3.7e-3 (3.1e-3, 4.3e-3)	4.7e-3 (4.1e-3, 5.3e-3)
	High school	4.1e-3 (3.9e-3, 4.3e-3)	4.2e-3 (4.0e-3, 4.3e-3)
	College or university	4.8e-3 (4.5e-3, 5.0e-3)	4.8e-3 (4.6e-3, 5.0e-3)

Table 6: Estimated c parameters of power-exponential hazard GLM fits for various U.S. infant cohorts. Mean estimates and 95% confidence intervals are based on 1000 parametric bootstrap replications.

		Female	**Male**
1995-1999			
APGAR	Very low [0,5)	2.7e-3 (2.6e-3, 2.9e-3)	2.7e-3 (2.6e-3, 2.8e-3)
	Low [5,9)	8.1e-1 (7.1e-1, 9.2e-1)	7.9e-1 (7.1e-1, 8.8e-1)
	Regular 9+	5.7e-6 (4.2e-7, 2.6e-5)	1.6e-5 (1.2e-6, 7.2e-5)
Origin	Non-Hispanic Black	1.1e-3 (1.0e-3, 1.2e-3)	1.0e-3 (9.1e-4, 1.1e-3)
	Non-Hispanic White	5.6e-4 (5.2e-4, 6.1e-4)	5.9e-4 (5.5e-4, 6.4e-4)
	Hispanic	6.1e-4 (5.2e-4, 7.1e-4)	6.9e-4 (6.0e-4, 8.0e-4)
	Other	5.4e-4 (3.9e-4, 7.2e-4)	4.3e-4 (3.1e-4, 5.8e-4)
Prematurity	Extremely preterm <28w	1.7e-3 (1.6e-3, 1.8e-3)	1.5e-3 (1.4e-3, 1.6e-3)
	Very preterm [28,32)w	7.8e-4 (5.8e-4, 1.0e-3)	1.0e-3 (7.8e-4, 1.3e-3)
	Moderate to late preterm [32,37)w	5.7e-4 (4.5e-4, 7.2e-4)	8.2e-4 (6.5e-4, 1.0e-3)
	Term 37w+	1.3e-4 (9.2e-5, 1.8e-4)	1.3e-4 (9.2e-5, 1.8e-4)
Education	Elementary or less	5.3e-4 (3.9e-4, 6.7e-4)	6.5e-4 (5.1e-4, 8.3e-4)
	High school	6.3e-4 (5.8e-4, 6.8e-4)	6.4e-4 (6.0e-4, 6.9e-4)
	College or university	7.9e-4 (7.2e-4, 8.6e-4)	7.8e-4 (7.0e-4, 8.5e-4)
2000-2005			
APGAR	Very low [0,5)	2.6e-3 (2.4e-3, 2.7e-3)	2.3e-3 (2.2e-3, 2.4e-3)
	Low [5,9)	8.0e-1 (6.9e-1, 9.3e-1)	7.4e-1 (6.4e-1, 8.3e-1)
	Regular 9+	6.6e-7 (4.6e-8, 2.9e-6)	6.3e-9 (1.0e-10, 3.5e-8)
Origin	Non-Hispanic Black	8.2e-4 (7.3e-4, 9.1e-4)	7.4e-4 (6.7e-4, 8.1e-4)
	Non-Hispanic White	5.5e-4 (5.1e-4, 6.1e-4)	5.1e-4 (4.7e-4, 5.5e-4)
	Hispanic	7.0e-4 (6.2e-4, 7.9e-4)	7.0e-4 (6.2e-4, 7.9e-4)
	Other	6.3e-4 (4.8e-4, 8.0e-4)	4.2e-4 (3.2e-4, 5.3e-4)
Prematurity	Extremely preterm <28w	1.5e-3 (1.4e-3, 1.6e-3)	1.3e-3 (1.2e-3, 1.4e-3)
	Very preterm [28,32)w	5.6e-4 (4.3e-4, 7.2e-4)	7.1e-4 (5.8e-4, 8.7e-4)
	Moderate to late preterm [32,37)w	5.2e-4 (4.1e-4, 6.4e-4)	6.2e-4 (5.0e-4, 7.6e-4)
	Term 37w+	9.5e-5 (6.5e-5, 1.4e-4)	6.1e-5 (4.3e-5, 8.4e-5)
Education	Elementary or less	6.6e-4 (5.1e-4, 8.4e-4)	4.9e-4 (3.8e-4, 6.3e-4)
	High school	5.5e-4 (5.0e-4, 6.0e-4)	5.0e-4 (4.6e-4, 5.4e-4)
	College or university	7.7e-4 (7.0e-4, 8.5e-4)	7.1e-4 (6.6e-4, 7.8e-4)

Table 7: Estimated p parameters of power-exponential hazard GLM fits for various U.S. infant cohorts. Mean estimates and 95% confidence intervals are based on 1000 parametric bootstrap replications.

		Female	**Male**
1995-1999			
APGAR	Very low [0,5)	8.7e-1 (8.6e-1, 8.8e-1)	8.7e-1 (8.6e-1, 8.8e-1)
	Low [5,9)	9.3e-1 (9.0e-1, 9.7e-1)	9.3e-1 (9.1e-1, 9.6e-1)
	Regular 9+	1.6e-1 (1.5e-1, 1.7e-1)	1.5e-1 (1.4e-1, 1.6e-1)
Origin	Non-Hispanic Black	6.6e-1 (6.5e-1, 6.6e-1)	6.5e-1 (6.4e-1, 6.5e-1)
	Non-Hispanic White	6.0e-1 (6.0e-1, 6.1e-1)	6.0e-1 (5.9e-1, 6.0e-1)
	Hispanic	6.1e-1 (6.0e-1, 6.2e-1)	6.1e-1 (6.0e-1, 6.1e-1)
	Other	6.1e-1 (5.9e-1, 6.3e-1)	5.8e-1 (5.7e-1, 6.0e-1)
Prematurity	Extremely preterm <28w	7.4e-1 (7.4e-1, 7.5e-1)	7.2e-1 (7.1e-1, 7.2e-1)
	Very preterm [28,32)w	5.6e-1 (5.5e-1, 5.8e-1)	5.5e-1 (5.4e-1, 5.7e-1)
	Moderate to late preterm [32,37)w	5.3e-1 (5.1e-1, 5.4e-1)	5.3e-1 (5.2e-1, 5.4e-1)
	Term 37w+	3.7e-1 (3.6e-1, 3.8e-1)	3.5e-1 (3.5e-1, 3.6e-1)
Education	Elementary or less	5.9e-1 (5.7e-1, 6.0e-1)	5.8e-1 (5.7e-1, 6.0e-1)
	High school	6.0e-1 (6.0e-1, 6.1e-1)	5.9e-1 (5.9e-1, 6.0e-1)
	College or university	6.3e-1 (6.3e-1, 6.4e-1)	6.2e-1 (6.2e-1, 6.3e-1)
2000-2005			
APGAR	Very low [0,5)	8.7e-1 (8.6e-1, 8.8e-1)	8.6e-1 (8.5e-1, 8.6e-1)
	Low [5,9)	8.1e-1 (7.8e-1, 8.4e-1)	8.2e-1 (8.0e-1, 8.5e-1)
	Regular 9+	1.6e-1 (1.4e-1, 1.7e-1)	1.2e-1 (1.0e-1, 1.3e-1)
Origin	Non-Hispanic Black	6.3e-1 (6.3e-1, 6.4e-1)	6.3e-1 (6.3e-1, 6.4e-1)
	Non-Hispanic White	6.0e-1 (6.0e-1, 6.1e-1)	5.9e-1 (5.8e-1, 5.9e-1)
	Hispanic	6.2e-1 (6.2e-1, 6.3e-1)	6.2e-1 (6.1e-1, 6.3e-1)
	Other	6.2e-1 (6.0e-1, 6.4e-1)	6.0e-1 (5.8e-1, 6.1e-1)
Prematurity	Extremely preterm <28w	7.3e-1 (7.3e-1, 7.4e-1)	7.1e-1 (7.0e-1, 7.1e-1)
	Very preterm [28,32)w	5.5e-1 (5.4e-1, 5.6e-1)	5.6e-1 (5.5e-1, 5.7e-1)
	Moderate to late preterm [32,37)w	5.2e-1 (5.1e-1, 5.3e-1)	5.1e-1 (5.0e-1, 5.2e-1)
	Term 37w+	3.6e-1 (3.5e-1, 3.7e-1)	3.4e-1 (3.4e-1, 3.5e-1)
Education	Elementary or less	6.1e-1 (6.0e-1, 6.3e-1)	5.9e-1 (5.7e-1, 6.0e-1)
	High school	5.9e-1 (5.9e-1, 6.0e-1)	5.8e-1 (5.8e-1, 5.9e-1)
	College or university	6.3e-1 (6.2e-1, 6.4e-1)	6.2e-1 (6.2e-1, 6.3e-1)

References

M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. National Bureau of Standards, Washington, D. C., 10th edition, 1964. doi:10.2307/2004284.

M. Aitkin and D. Clayton. The fitting of exponential, Weibull and extreme value distributions to complex censored survival data using GLIM. *Journal of the Royal Statistical Society C*, 29(2):156–163, 1980. doi:10.2307/2986301.

V. Apgar. A proposal for a new method of evaluation of the newborn infant. *Current Researches in Anesthesia and Analgesia*, 32(4):260–267, 1953. doi:10.1213/ANE.0b013e31829bdc5c.

D. Avraam, J. P. de Magalhaes, and B. Vasiev. A mathematical model of mortality dynamics across the lifespan combining heterogeneity and stochastic effects. *Experimental Gerontology*, 48(8):801–811, 2013. doi:10.1016/j.exger.2013.05.054 .

D. Avraam, S. Arnold, D. Jones, and B. Vasiev. Time-evolution of age-dependent mortality patterns in mathematical model of heterogeneous human population. *Experimental Gerontology*, 60:18–30, 2014. doi:10.1016/j.exger.2014.09.006.

S. Berrut, V. Pouillard, P. Richmond, and B. M. Roehner. Deciphering infant mortality. *Physica A: Statistical Mechanics and its Applications*, 463:400–426, 2016. doi:10.1016/j.physa.2016.07.031.

J. Bourgeois-Pichat. La mesure de la mortalité infantile. II. les causes de décès. *Population*, 6(3):459–480, 1951. doi:10.2307/1523958.

D. R. Brillinger. A justification of some common laws of mortality. In *Transactions of the Society of Actuaries*, volume XIII, pages 115–119, 1961.

C. G. Camarda, P. H. Eilers, and J. Gampe. Sums of smooth exponentials to decompose complex series of counts. *Statistical Modelling*, 16(4):279–296, 2016. doi:10.1177/1471082X16641796.

J. H. Cha and M. Finkelstein. Justifying the Gompertz curve of mortality via the generalized Polya process of shocks. *Theoretical Population Biology*, 109:54–62, 2016. doi:10.1016/j.tpb.2016.03.001.

M. K. Choe. Fitting the age pattern of infant and child mortality with the Weibull survival distribution. In *Asian and Pacific census forum*, volume 7, pages 10–13, 1981.

D. Clayton and J. Cuzick. Multivariate generalizations of the proportional hazards model. *Journal of the Royal Statistical Society A*, 148(2):82–108, 1985. doi:10.2307/2981943.

D. G. Clayton. Fitting a general family of failure-time distributions using GLIM. *Applied Statistics*, 32(2):102–109, 1983. doi:10.2307/2347288.

D. R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society B*, 34(2):187–220, 1972. doi:10.1007/978-1-4612-4380-9_37.

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

J. de Beer and F. Janssen. A new parametric model to assess delay and compression of mortality. *Population Health Metrics*, 14(1):46, 2016. doi:10.1186/s12963-016-0113-1.

P. H. C. Eilers and B. D. Marx. Flexible smoothing with B-splines and penalties. *Statistical Science*, 11(2):89–102, 1996. doi:10.1214/ss/1038425655.

M. S. Finkelstein. Lifesaving explains mortality decline with time. *Mathematical Biosciences*, 196(2):187–197, 2005. doi:10.1016/j.mbs.2005.04.004.

T. B. Gage and B. Dyke. Parameterizing abridged mortality tables: The Siler three-component hazard model. *Human Biology*, 58(2):275–291, 1986.

L. Heligman and J. H. Pollard. The age pattern of mortality. *Journal of the Institute of Actuaries*, 107(1):49–80, 1980. doi:10.1017/s0020268100040257.

P. Hougaard. Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika*, 71(1):75–83, 1984. doi:10.1093/biomet/71.1.75.

P. Hougaard. Survival models for heterogeneous populations derived from stable distributions. *Biometrika*, 73(2):387–396, 1986. doi:10.1093/biomet/73.2.387.

P. Hougaard. Frailty models for survival data. *Lifetime Data Analysis*, 1(3):255–273, 1995. doi:10.1007/bf00985760.

P. Huck. Infant mortality and living standards of English workers during the industrial revolution. *The Journal of Economic History*, 55(3):528–550, 1995. doi:10.1017/s0022050700041620.

H. C. Kinney and B. T. Thach. The sudden infant death syndrome. *New England Journal of Medicine*, 361(8):795–805, 2009. doi:10.1056/nejmra0803836.

J. Knodel and H. Kintner. The impact of breast feeding patterns on the biometric analysis of infant mortality. *Demography*, 14(4):391–409, 1977. doi:10.2307/2060586.

E. H. Lehman. Shapes, moments and estimators of the Weibull distribution. *IEEE Transactions on Reliability*, 12(3):32–38, 1963. doi:10.1109/tr.1963.5218214.

D. A. Levitis. Before senescence: the evolutionary demography of ontogenesis. *Proceedings of the Royal Society B*, 278(1707):801–809, 2011. doi:10.1098/rspb.2010.2190.

K. S. Lomax. Business failures: Another example of the analysis of failure data. *Journal of the American Statistical Association*, 49(268):847–852, 1954. doi:10.1080/01621459.1954.10501239.

A. W. Marshall and I. Olkin. *Life Distributions*. Springer, New York, 2007. ISBN 9780387203331. doi:10.1007/978-0-387-68477-2.

P. McCullagh and J. A. Nelder. *Generalized Linear Models*. Monographs on Statistics and Applied Probability. CRC Press, New York, 2nd edition, 1989. ISBN 9780412317606.

D. McFadden. Conditional logit analysis of qualitative choice behavior. In P. Zarembka, editor, *Frontiers in econometrics*, pages 105–142. Academic Press, New York, 1974.

C. J. Mode and R. C. Busby. An eight-parameter model of human mortality—the single decrement case. *Bulletin of Mathematical Biology*, 44(5):647–659, 1982. doi:10.1007/bf02462273.

National Center for Health Statistics. Birth cohort linked birth-infant death data files (U.S. data), 2016. URL ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/DVS/cohortlinkedus/.

Oppermann. On the graduation of life tables, with special application to the rate of mortality in infancy and childhood. *Insurance Record*, page 42, Feb. 1870. Minutes from a meeting in the Institute of Actuaries.

A. Prékopa. On secondary processes generated by a random point distribution of Poisson type. *Annales Univ. Sci. Budapest de Eötvös Nom. Sectio Math*, 1:153–170, 1958.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL https://www.R-project.org/.

H. Rinne. *The Weibull Distribution*. CRC Press, New York, 1st edition, 2009. ISBN 978-1-4200-8743-7. doi:10.1201/9781420087444.

A. Rogers and J. S. Little. Parameterizing age patterns of demographic rates with the multiexponential model schedule. *Mathematical Population Studies*, 4(3):175–195, 1994. doi:10.1080/08898489409525372.

W. Siler. A competing-risk model for animal mortality. *Ecology*, 60(4):750–757, 1979. doi:10.2307/1936612.

J. F. Steffensen. Infantile mortality from an actuarial point of view. *Scandinavian Actuarial Journal*, 1930(2):272–286, 1930. doi:10.1080/03461238.1930.10416902.

B. L. Strehler and A. S. Mildvan. General theory of mortality and aging. *Science*, 132 (3418):14–21, 1960. ISSN 0036-8075. doi:10.1126/science.132.3418.14.

T. N. Thiele. 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, 1871. doi:10.1017/s2046167400043688.

R. Thompson and R. J. Baker. Composite link functions in Generalized Linear Models. *Journal of the Royal Statistical Society C*, 30(2):125–131, 1981. doi:10.2307/2346381.

J. W. Vaupel and A. I. Yashin. The deviant dynamics of death in heterogeneous populations. Technical report, Laxenburg, Austria, 1983. URL http://user.demogr.mpg.de/jwv/pdf/IIASA-83-001.pdf.

J. W. Vaupel and A. I. Yashin. Heterogeneity's ruses: Some surprising effects of selection on population dynamics. *The American Statistician*, 39(3):176–185, 1985. doi:10.2307/2683925.

J. W. Vaupel and Z. Zhang. Attrition in heterogeneous cohorts. *Demographic Research*, 23(26):737–748, 2010. doi:10.4054/DemRes.2010.23.26.

J. W. Vaupel, K. G. Manton, and E. Stallard. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3):439–54, 1979. doi:10.2307/2061224.

A. Wienke. *Frailty Models in Survival Analysis*. Biostatistics Series. Chapman and Hall, Boca Raton, 2011. ISBN 978-1-4200-7388-1. doi:10.1111/j.1541-0420.2012.01769.x.

T. Wittstein and D. A. Bumsted. The mathematical law of mortality. *Journal of the Institute of Actuaries and Assurance Magazine*, 24(3):153–173, 1883. doi:10.1017/s0020268100006260.

World Health Organization. *Neonatal and Perinatal Mortality: Country, Regional and Global Estimates*. World Health Organization, 2006. ISBN 9241563206.

World Health Organization. *Health in 2015. From Millennium Development Goals to Sustainable Development Goals*. World Health Organization, 2015. ISBN 9789241565110. doi:10.1111/padr.12006.