Understanding mortality rate deceleration and heterogeneity

David R. Steinsaltz

Kenneth W. Wachter

Department of Demography, University of California,
Berkeley, CA., USA.

September 5, 2005

David Steinsaltz was supported by Grant K12-AG00981 from the National Institute on Aging and Kenneth Wachter was supported by Grant P01-008454 from the National Institute on Aging.

Corresponding author: David Steinsaltz, Department of Demography, University of California, 2232 Piedmont Avenue, Berkeley CA 94720-2120, USA, E-mail: dstein@demog.berkeley.edu

UNDERSTANDING MORTALITY RATE DECELERATION AND HETEROGENEITY

ABSTRACT. We draw on classical mathematical results on the limiting behavior of Laplace transforms to shed light on generic relationships between heterogeneity in population frailty and flattening of aggregate population hazard functions at extreme ages. In particular, we show that the population hazard function converges to a constant precisely when the distribution of unobserved heterogeneity in initial mortalities behaves asymptotically as a polynomial near zero.

Keywords: frailty, heterogeneity, biodemography, hazard functions, Laplace Transforms, mortality plateaus.

1. INTRODUCTION

After 180 years, Benjamin Gompertz' exponential formula remains the starting point for descriptions of the age trajectory of adult mortality rates in numerous species including our own. But it is only a starting point. Discoveries emerging from the measurement of hazard rates at advanced ages in large populations of model organisms as well as humans have now directed attention onto systematic deviations from Gompertz' formula (Vaupel et al. 1998). Their theoretical importance is featured in the new collection *Lifespan* edited by Carey and Tuljapurkar (2003)

"Deceleration" in the title of this paper is the broad term for downward deviations at late ages in the exponentially accelerating hazards posited by the Gompertz model. "Flattening" and "tapering" are other terms in use. When hazards level out, it is common to speak of "mortality plateaus".

The explanations on offer to account for deviations at late ages from Gompertz models may be loosely sorted into three categories: 1) There is some more general Law of Mortality to which the Gompertz curve is only an approximation at less advanced ages; 2) There is an essential change in the aging process in the very old; 3) The Gompertz hazard function remains in force at extreme ages, but it is masked by progressive age-specific selection operating on heterogeneous frailty in the

initial population. This explanation often goes under the name of "simple culling". Like the present paper, most discussions of heterogeneity take as a point of departure the demographic frailty model of Vaupel, Manton and Stallard (1979) described in Section 3. Within this context, a lively debate is in progress about the ability of explanation (3) to account for observed patterns (Wachter 2003, p.275–277).

Much of this debate revolves around properties of one specific parametric model (the so-called "Gamma Gompertz") and its close compeers. In this paper we bring to bear classical mathematical results which allow us to separate generic properties of frailty models from relationships that depend on specific parametric forms. These results facilitate a more systematic view of the interaction between population heterogeneity and observable hazard rates.

The empirical survival curves that figure in biodemographic discussions typically serve as illustrative examples of qualitative properties calling for explanation, not as cases demanding precise fitting in their own right. A handful of high quality cohort lifetables are studied under the heading of "human curves". Needless to say, they are not intended to capture the detailed variety of human mortality patterns. Curves for medflies (Mediterranean fruit flies) or nematode worms play similar roles, exemplifying notable features but necessarily also reflecting the particularities of observational protocols. Because qualitative features are at issue, it is sensible for mathematical approaches to transcend specific parametric models and emphasize general relationships such as those treated in the present paper.

Section 2 reviews prior work and Section 3 presents the basic model for heterogeneity. Our main results on the generic properties of frailty models that lead to asymptotic limits in aggregate hazard functions are given in Section 4. Consequences for common parametric models are found in Section 5 and numerical applications and interpretive conclusions in Section 6.

2. Prior Work

Summaries of the biodemographic and evolutionary context can be found in Wachter and Finch (1997) and Carey (2003) and an overview of models in Yashin, Iachine and Begun (2000). Investigations of the potency of heterogeneity in frailty have mainly turned on the question: How much heterogeneity would be required in the population to produce observed amounts of flattening or tapering in mortality curves by age? Carey, Liedo, Orozco and Vaupel (1992) found dramatic declines in late-age mortality rates in captive Mediterranean fruit flies. While

actual declines are not characteristic of many other species, substantial data on mortality-curve flattening in flies, nematodes, humans, and even automobiles have accumulated in recent years.(Brooks et al. 1994, Wang et al. 1998, Vaupel et al. 1998, Pletcher and Curtsinger 1998, Horiuchi and Wilmoth 1998, Drapeau et al. 2000, Lynch and Brown 2001, Steinsaltz 2005) (A discussion of this research through 1996 may be found in Vaupel (1997).) A selection of comments and explanations may be found in Kowald et al. (1993) and Gavrilov et al. (1993), including suggestions that mortality at late ages is significantly affected by decreasing density of captive flies and by the increasing scale of stochastic fluctuations.

Vaupel and Carey (1993) address the heterogeneity hypothesis with a semi-parametric approach. They ask about the extent of heterogeneity that would be needed for their medflies to reconcile an underlying Gompertz hazard rate with selected mortality curves. They model the population as being composed of twelve discrete subpopulations. Each subpopulation is given a distinct value of initial mortality but the same exponential rate of progression $e^{.3x}$ by age x. The best fit to their observed curve requires their frailest group, with 41% of the population, to have a mortality rate on the order of 10^{10} times that of the superflies who begin as only .0056% of the population, but who constitute the overwhelming majority of the survivors after 100 days. A model with Weibull progression by age gives slightly less extreme answers, with five classes and a mere factor of 500 in initial mortality between *Ceratitis* hoi polloi and the one in ten thousand predisposed to centenarianity.

A broader investigation of heterogeneity and late-age hazard functions for selected invertebrate and vertebrate species has been undertaken by Horiuchi (2003). Horiuchi uses the modal age at death to establish a standardized scaling of the age axis which facilitates comparisons. By his estimates, human mortality curves reflect less unobserved heterogeneity than curves for invertebrates. He suggests reasons based on considerations of quality control. Heterogeneity is still great enough, in his view, to account for the flattening of human curves at extreme ages. Horiuchi relies on specific fully parametric models. We discuss his results in terms of our more generic approach in Section 6.

The complex trajectories of medfly hazard functions may preclude a unified explanation. Indeed, the mortality curves seem to depend in qualitative ways on details of the experiments, and frailty models are ill-suited for explaining all the ups and downs of complex curves. But what is shared among the main available examples is the phenomenon of flattening or tapering, producing a segment on graphs of mortality rates by age which resembles a plateau. It is well-known that certain

specific parametric models for frailty can generate such plateaus. Here we show how much of this behavior is a generic property of frailty models. We focus on two questions: How is the assumed frequency of extremely robust individuals, with fixed lifetime frailties far closer to zero than the average, connected to the appearance of mortality plateaus? To what extent do inferences about such plateaus depend on specifics of a chosen model?

3. Heterogeneous hazards

The basic demographic model under which we prove our results is the frailty model of Vaupel et al. (1979). It ascribes to each individual an unobserved implicit fixed lifetime frailty variable, which acts as a multiplier on a baseline hazard function shared by members of the population. The model thus combines the concept of unobserved heterogeneity with the framework of proportional hazards widely applied throughout demography.

Thus we assume that a population is composed of a large number of individuals, each of whom dies at an independent random time. The hazard rate or hazard function is the slope of the logarithm of the probability of surviving for a time t either from birth or from a time origin that marks the onset of adulthood. Under the model, the hazard rate of individual i at time t is $Z_i\mu(t)$, where the baseline hazard rate $\mu(t)$ is a fixed function, and Z_i are i.i.d. random variables with expectation 1. We denote the Laplace transform of the distribution of Z_i by $\lambda(s)$; that is,

$$\lambda(s) = \mathrm{E}\big[e^{-sZ_i}\big].$$

We assume throughout that the distribution of Z has a density supported on the non-negative real line and that Z is normalized to have unit mean. It follows that the Laplace Transform λ and the transform $\mathrm{E}\big[Z\exp(-sZ)\big]$ both have half-planes of convergence in the complex plane, as required by the theorems we shall cite. In practice we only need to evaluate the transforms for non-negative real numbers s.

The cumulative baseline hazard function is $M(t) = \int_0^t \mu(a)da$. It is monotone non-decreasing since μ is non-negative and we assume that μ is strictly positive at high ages so that $M(t)$ increases monotonically to infinity as $t \to \infty$.

The aggregate effective hazard rate function for the population (that is, for a random individual) will be denoted $\mu^*(t)$. It is the object of primary interest which can be estimated from observations.

Since the hazard function is defined to be the slope of the logarithm of the survival function, for each individual we have the conditional

probability

$$\mathrm{P}\big\{i \text{ survives until time } t \,\big|\, Z_i = Z\big\} = e^{-ZM(t)}.$$

Consequently, the population survival function is given by the unconditional probability

$$\mathrm{P}\big\{i \text{ survives until time } t\big\} = \lambda(M(t)).$$

Differentiating the logarithm of the population survival function gives the aggregate population hazard rate:

$$\mu^*(t) = -\mu(t)\frac{\lambda'(M(t))}{\lambda(M(t))}. \tag{1}$$

The rate of change of the logarithm of the aggregate hazard rate has been called the "Lifetable Aging Rate" or LAR by Horiuchi and Coale (1990) and Horiuchi and Wilmoth (1997). The LAR is constant under Gompertz mortality. Variations by age indicate deviations from the Gompertz pattern. Dividing log(2) by the LAR gives the Mortality Rate Doubling Time or MRDT. We denote the LAR by $L(t)$. It is given by

$$\begin{aligned} L(t) &:= \frac{d}{dt}\big(\log \mu^*(t)\big) \\ &= \frac{\mu'(t)}{\mu(t)} + \frac{\mu(t)}{\lambda(M(t))\lambda'(M(t))}\big[\lambda''(M(t))\lambda(M(t)) - (\lambda'(M(t)))^2\big]. \end{aligned} \tag{2}$$

Assuming, as we ordinarily do, that $M(t)$ is strictly increasing, and $\lim_{t\to\infty} M(t) = \infty$ (which will generously be satisfied if, for instance, the hazard rate μ has a nonzero infimum), we get an expression for life expectancy:

$$Life\ Expectancy = \int_0^\infty \frac{\lambda(s)}{\mu(M^{-1}(s))} ds. \tag{3}$$

4. Asymptotes

Our principal results draw on the classical mathematics of Abelian and Tauberian theorems to establish conditions for the convergence of the population-level aggregate hazard function to an asymptotic limit at high ages. An introduction to the theory is found in the textbook by Feller (1971, p. 418–423).

Intuitively, the mortality at late times t in the life cycle will depend on the individuals with low mortality who tend to survive out to those times. The asymptotic behavior of $\mu^*(t)$ as $t \to \infty$ will thus depend on the limiting behavior of the density of frailty $f(z)$ at low frailties as $z \to 0$. In fact, we show under suitable conditions that the aggregate

population hazard rate $\mu^*(t)$ which mixes individuals of all frailties converges to a positive constant if and only if $f(z)$ behaves approximately like a power of z (greater than -1) as $z \to 0$.

The conditions depend on behavior of the derivative of the logarithm of the cumulative baseline hazard $M(t)$, given by the ratio $\mu(t)/M(t)$. Suppose this ratio converges to a constant θ as $t \to \infty$. (With a Gompertz baseline, the constant is the Gompertz slope parameter.) Then $f(0) = 0$ when the limit of $\mu^*(t)$ is larger than θ, and $\lim_{z\to 0} f(z) = \infty$ when the limit is smaller than θ.

Theorem 1. *Suppose the baseline hazard function satisfies*

$$\lim_{t\to\infty} \mu(t)/M(t) = \theta \geq 0$$

and the density for the frailty distribution satisfies $f(z) \sim az^\rho$ *as* $z \to 0$ *for some* $\rho > -1$ *and* $a > 0$. *That is to say* $\lim_{z\to 0} z^{-\rho} f(z) = a$. *Then*

$$\lim_{t\to\infty} \mu^*(t) = \theta(\rho + 1). \tag{4}$$

Proof. One of the best-known Abelian theorems (Doetsch 1950, Theorem 1, p. 474) tells us that the Laplace transform λ behaves asymptotically as

$$\lambda(s) \sim a\Gamma(\rho + 1)s^{-\rho-1}$$

for s tending to infinity when $f(z) \sim az^\rho$ for z tending to zero with $\rho > -1$. Since $-\lambda'(s)$ is the Laplace transform of $zf(z)$, we see that as $s \to \infty$

$$\frac{-\lambda'(s)}{\lambda(s)} \sim \frac{a\Gamma(\rho + 2)s^{-\rho-2}}{a\Gamma(\rho + 1)s^{-\rho-1}} = (\rho + 1)s^{-1}.$$

By (1), putting $s = M(t)$,

$$\mu^*(t) \sim (\rho + 1)\mu(t)/M(t),$$

which by assumption converges to the constant $\theta(\rho+1)$ as $t \to \infty$. □

The relationship $\mu^*(t)/\mu(t) \sim (\rho + 1)/M(t)$ derived in the course of the proof tells us that the mean frailty among survivors to time t as $t \to \infty$ drops like $\rho + 1$ divided by minus the logarithm of the aggregate proportion surviving. In practice, this behavior may set in rather slowly with increasing t.

The reverse of an Abelian theorem is a Tauberian theorem. But we cannot simply reverse the argument in Theorem 1 to turn the sufficient condition for a mortality plateau into a necessary condition. Tauberian theorems, depending as they do on inverting Laplace transforms, are inevitably weaker than Abelian theorems. In addition, our starting position is weaker, since a mortality plateau is an asymptotic condition on the logarithmic derivative of the Laplace transform λ, not on λ itself.

Thus, the necessary condition has to be weaker than the sufficient condition.

Theorem 2. *Suppose that the aggregate population hazard function* $\mu^*(t)$ *satisfies* $\lim_{t\to\infty}\mu^*(t) = b$ *for some* b *when the baseline hazard function satisfies* $\lim_{t\to\infty}\mu(t)/M(t) = \theta$. *Then*

$$\lim_{z\downarrow 0}\frac{\log f(z)}{\log z} = \frac{b}{\theta} - 1. \tag{5}$$

Proof. We want to use Theorem 5 on page 476 of Doetsch (1950), which tells us that, for every $\rho > -1$,

$$\begin{aligned}\liminf_{z\downarrow 0} z^{-\rho} f(z) &\le \liminf_{s\to\infty}\frac{s^{\rho+1}}{\Gamma(\rho+1)}\lambda(s) \\ &\le \limsup_{s\to\infty}\frac{s^{\rho+1}}{\Gamma(\rho+1)}\lambda(s) \le \limsup_{z\downarrow 0} z^{-\rho} f(z).\end{aligned} \tag{6}$$

As $t\to\infty$, by assumption $\mu^*(t)\sim b$ and $\mu(t)\sim\theta M(t)$. It follows that $\mu^*(t)/\mu(t)\sim b/\mu(t)\sim(b/\theta)/M(t)$. Since $M(t)$ is monotone increasing for large t, when we set $s = M(t)$, we can express t in terms of s and write

$$\frac{d}{ds}\log(\lambda(s)) = \frac{\lambda'(s)}{\lambda(s)} \sim -\frac{b}{\theta s}.$$

Thus, for any positive ϵ, we have constants $c(\epsilon)$, $C(\epsilon)$, and $S(\epsilon)$ such that for $s > S(\epsilon)$,

$$\log c(\epsilon) + \left(-\frac{b}{\theta} - \frac{\epsilon}{2}\right)\log s \le \log\lambda(s) \le \log C(\epsilon) + \left(-\frac{b}{\theta} + \frac{\epsilon}{2}\right)\log s.$$

Thus

$$c(\epsilon)s^{-b/\theta-\epsilon/2} \le \lambda(s) \le c(\epsilon)s^{-b/\theta+\epsilon/2}$$

for $s > S(\epsilon)$. We then have

$$\begin{aligned}\lim_{s\to\infty} s^{b/\theta-\epsilon}\lambda(s) &= 0 \text{ and} \\ \lim_{s\to\infty} s^{b/\theta+\epsilon}\lambda(s) &= \infty.\end{aligned}$$

As an immediate consequence of (6) we then have

$$\forall\epsilon > 0,\ \lim_{z\downarrow 0} z^{-(b/\theta-1)+\epsilon} f(z) = 0 \text{ and } \lim_{z\downarrow 0} z^{-(b/\theta-1)-\epsilon} f(z) = \infty. \tag{7}$$

Since (7) is equivalent to (5), this completes the proof. □

The frailty distributions with power-law decay in the lower tail are the simplest distributions satisfying Equation 5. Since the frailty distribution among survivors to age t has density $f(z)\exp(-zM(t))/\lambda(M(t))$,

any power-law distribution thins down into a gamma distribution as $M(t)$ grows large and frail individuals die away. But the theorem also allows cases, for instance, with $f(z) \sim z^\rho \log(1/z)$ as $z \to 0$. For such cases, the distribution of frailty among survivors would not tend to a gamma distribution. Thus gamma distributions are common but not generic distributions for frailty among survivors.

5. Parametric cases

5.1. **Baseline hazards.** We review the main parametric forms which have been studied in the context of mortality rate deceleration. For baseline hazards, what matters for the theorems is the limiting behavior of the ratio $\mu(t)/M(t)$ at infinity.

Gompertz: $\mu(t) = \alpha e^{\beta t}$. The ratio $\mu/M = \beta/(1 - e^{-\beta t})$ converges to β.

Makeham: $\mu(t) = c + \alpha e^{\beta t}$. The ratio μ/M has the same limit as the Gompertz.

Log Quadratic : $\mu(t) = \exp(a + bt + ct^2)$. The ratio μ/M increases linearly like $2ct$ for large t and the aggregate hazard $\mu^*(t)$ does not converge to an asymptote for any frailty distribution with density conforming to a power-law near zero.

Polynomial: $\mu(t) = a_0 + a_1 t + a_2 t^2 + \ldots a_k t^k$. The ratio μ/M decreases like $(k+1)/t$ for large t, so that $\mu^*(t)$ converges to zero by Theorem 1.

The log-quadratic baseline hazard is favored by Horiuchi (2003) to take account of transient accelerations at middle ages in human cohort mortality schedules. The behavior of the ratio μ/M is obtained from its reciprocal M/μ, which can be expressed in terms of Dawson's Integral (Abramowitz and Stegun 1965, p. 298). Completing the square inside $\log(\mu)$ gives an equivalent quadratic $c(x+B)^2 - cB^2 + a$ with $B = b/(2c)$). Then

$$\frac{M(t)}{\mu(t)} = e^{-c(t+B)^2} \int_0^t e^{c(x+B)^2} dx$$

Putting $y = c^{1/2}(x+B)$, $y_0 = c^{1/2}B$ and $\tau = c^{1/2}(t+B)$ yields

$$c^{1/2}\frac{M(t)}{\mu(t)} = e^{-\tau^2} \int_0^\tau e^{y^2} dy - e^{-\tau^2} \int_0^{y_0} e^{y^2} dy$$

If we multiply the both terms on the right by 2τ, the first term is 2τ times Dawson's Integral evaluated at τ, which converges to unity as $\tau \to \infty$ (Abramowitz and Stegun 1965, p. 319) and the second term converges to zero. As a consequence, $\mu(t)/M(t) \sim b + 2ct$.

Given this behavior for the ratio, the proof of Theorem 1 shows us that $\mu^*(t)$ increases toward infinity when $f(z)$ satisfies the power-law condition of the theorem. It is not obvious whether or not there could be a frailty distribution with even thicker tails at zero which would give a finite asymptote for the aggregate hazard rate. More powerful Abelian theorems (Feller 1971, p.423) might provide an answer to this question.

For individuals with frailty z, the modal age at death, a convenient index for comparisons (Horiuchi 2003), satisfies

$$\frac{d}{dt}\log(\mu(t)) = z\mu(t) \tag{8}$$

since the density of deaths is proportional to $\mu(t)e^{-zM(t)}$. Frailty values other than unity like 1/2, 1/5 or 1/10 can be re-expressed in terms of the corresponding shift in modal age. The transformation provides a handy way of calibrating the plausibility of frailties estimated to occur in the lower tail of a distribution.

For a Gompertz baseline, the modal age is given by $(1/\beta)\log(\beta/(\alpha z)$; passing from unit frailty to frailty z shifts the modal age by $(1/\beta)log(1/z)$. For a Makeham baseline, the equation for the modal age can be solved using the quadratic formula. For numerical examples like those in Section 6, the dependence on z is close to that of the Gompertz. For a log quadratic baseline, in the absence of a closed-form solution, it is easier to solve for the frailty z which would produce any given value τ for the modal age:

$$z = (b + 2c\tau)e^{-(a+b\tau+c\tau^2)}$$

As compared to the Gompertz, it takes somewhat smaller frailties with a log quadratic baseline to produce as much of a shift in modal age.

5.2. **Frailty distributions.** For frailty distributions, what matters for the theorems is the behavior of the density $f(z)$ around $z = 0$, and in particular the exponent ρ when $f(z) \sim az^\rho$.

Gamma: $f(z) = (r^\nu/\Gamma(\nu))z^{\nu-1}e^{-rz}$. For a unit mean, $\nu = r$. The exponent ρ equals $\nu - 1$, where ν is the gamma shape parameter.

Weibull: $f(z) = rc(rz)^{c-1}\exp(rz)^c$. The exponent ρ equals $c-1$

Beta: $f(z)$ proportional to $z^p(b - z)^q$. For a unit mean, $q = p(b - 1)$. The exponent ρ equals p.

Lognormal: $f(z)$ is proportional to $z^{-1+(\log(1/z)+2\zeta)/(2\sigma^2)}$. For a unit mean, $\zeta = -\sigma^2/2$. The distribution does not conform to a power law in the lower tail.

By far the most frequently studied case is the gamma distribution, treated in more detail in Section 5.3.

5.3. **Gamma distributed initial mortalities.** The gamma distribution, adopted from the outset of frailty modeling by Vaupel et al. (1979), has the convenient property that the distribution of frailties among survivors after time t, altered by the operation of selective mortality, remains a distribution from the gamma family with the same shape parameter ν but an altered rate parameter r. For large values of the shape parameter ν the distribution is very close to normal, while for $\nu = 1$ it is an exponential. The initial frailty distribution is always normalized to have a mean of 1, leaving only one free parameter $\nu = r = 1/\operatorname{Var}(Z)$. Thus, if σ is the standard deviation of the initial mortalities, it can be fitted to a gamma distribution with $\nu = r = \mu(0)^2/\sigma^2$.

The Laplace transform of the gamma distribution is

$$\lambda(s) = \left(\frac{\nu}{\nu + s}\right)^r.$$

Thus

$$\frac{\lambda'(s)}{\lambda(s)} = -\frac{r}{\nu + s},$$

With $\nu = r$ we obtain the well-known relationship (Vaupel et al. 1979),(Yashin et al. 2000):

$$\mu^*(t) = \frac{r\mu(t)}{r + M(t)}, \tag{9}$$

The Lifetable Aging Rate is given by

$$L(t) = \frac{\mu'(t)}{\mu(t)} - \frac{\mu(t)}{r + M(t)}. \tag{10}$$

5.4. **Gompertz individuals in a gamma population.** Now assume that the baseline hazard $\mu(t)$ fits a Gompertz curve $\mu_0 e^{\beta t}$. Then $M(t) = (\mu_0/\beta)(e^{\beta t} - 1)$, and defining $C := \mu_0/r = \sigma^2/\mu_0$, we obtain the well-known expression which shows that gamma frailty with a Gompertz basline generates a logistic curve for the population hazard function:

$$\mu^*(t) = \frac{\beta\mu_0 e^{\beta t}}{\beta - C + Ce^{\beta t}}. \tag{11}$$

This logistic curve is often fitted to observed hazard functions (cf. Horiuchi and Wilmoth (1998) or Wilmoth and Robine (2003)) when deceleration of mortality at advanced ages is under examination. The

corresponding rate of logarithmic increase is

$$L(t) = \frac{\beta(\beta - C)}{\beta - C + Ce^{\beta t}}.$$

Observe that $L(0)$ is the initial rate of exponential increase of the mortality rate observed in the population, and $\beta = L(0) + C$. Thus

$$L(t) = L(0)\frac{L(0) + C}{L(0) + Ce^{(L(0)+C)t}}. \tag{12}$$

We will generally consider parameters with β much larger than C, so that the difference between β and $L(0)$ will be negligible. Note that when $\beta < C$, the mortality rate is decreasing. We also get $\mu(M^{-1}(s)) = \beta s + 1$, so that the life expectancy becomes

$$\int_0^\infty \left(1 + \frac{s}{r}\right)^{-r} (\beta s + \mu_0)^{-1} ds. \tag{13}$$

This integral cannot be written in a closed form, but may readily be approximated.

Assume now that $L(0) > 0$. If we define T_k to be the time at which this rate has fallen by a factor of k, then

$$T_k = \big(L(0) + C\big)^{-1} \log(k + (k-1)L(0)/C) \tag{14}$$

If we are given T_k and μ_0, and wish to find the necessary standard deviation that will produce this T_k, we solve for C in

$$e^{L(0)T_k} Ce^{T_k C} - kC - (k-1)L(0) = 0, \tag{15}$$

and then set $\sigma = (\mu_0 C)^{1/2}$.

5.5. **Distributions bounded away from 0.** In finite populations, frailties drawn from a distribution with a thin lower tail will have a minimum value distributed at some remove from zero. One could, alternatively, impose a minimum value for frailties. Such a step requires an extra arbitrary assumption. But it is interesting to ask what the effect would be. What would the population mortality curve look like if the initial mortality multipliers Z_i were supported on an interval with minimum value $\eta \in (0, 1)$? (We assume that η is actually in the support. What this means is that $\mathrm{P}\big\{Z_i < \eta\big\} = 0$, but $\mathrm{P}\big\{Z_i < \eta + \epsilon\big\}$ is positive for every positive ϵ.) On a coarse level, the asymptotics are obvious: the mortality rate converges to $\eta \cdot \mu(t)$. Let $f(z)$ represent the distribution of Z_i, and let $\tilde{f}(z)$ be the distribution of $Z_i - \eta$. Then

$$\lambda(s) = \int_0^\infty e^{-sz} f(z) dz = e^{-s\eta} \int_0^\infty e^{-sx} \tilde{f}(z) dz,$$

so

$$-\frac{\lambda'(s)}{\lambda(s)} = \eta + \frac{\int_0^\infty z e^{-sz}\tilde{f}(z)dz}{\int_0^\infty e^{-sz}\tilde{f}(z)dz},$$

which converges to η as $s \to \infty$.

If we follow up on the choice of gamma distributions in Section 5.3, a natural extension bounded away from 0 would be $Z_i = \eta + (1-\eta)Z_i^*$, where Z_i^* is gamma distributed with expectation 1 and parameters $r = \nu = 1/\sigma^2$. The Laplace transform will then be

$$\lambda(s) = e^{-\eta s}\big(1 + s(1-\eta)/r\big)^{-r},$$

yielding

$$-\frac{\lambda'(s)}{\lambda(s)} = \eta + \frac{r(1-\eta)}{r + s(1-\eta)}.$$

Thus

$$\mu^*(t) = \mu(t)\left[\eta + \frac{r(1-\eta)}{r + (1-\eta)M(t)}\right]. \tag{16}$$

Its logarithmic derivative is

$$L(t) = \frac{\mu'(t)}{\mu(t)} - \frac{r(1-\eta)^2\mu(t)}{\eta\big(r + (1-\eta)M(t)\big)^2 + r(1-\eta)\big(r + (1-\eta)M(t)\big)}. \tag{17}$$

If $\mu(t) = \mu_0 e^{\theta t}$, then we get

$$\mu^*(t) = \mu_0 e^{\theta t}\left[\eta + \frac{1-\eta}{1 + (1-\eta)\mu_0(r\theta)^{-1}\big[e^{\theta t} - 1\big]}\right] \tag{18}$$

and

$$L(t) = \theta - r(1-\eta)^2\theta^2\mu_0 e^{-\theta t}\Big[\eta\big((1-\eta)\mu_0\big[1 - e^{-\theta t}\big] + r\theta e^{-\theta t}\big)^2 + r\theta(1-\eta)e^{-\theta t}\big((1-\eta)\mu_0\big[1 - e^{-\theta t}\big] + r\theta e^{-\theta t}\big)\Big]^{-1}. \tag{19}$$

The initial rate of increase is then

$$L(0) = \theta - \mu_0(1-\eta)^2/r = \theta - (1-\eta)^2\alpha. \tag{20}$$

5.6. **The inverse Gompertz problem.** Suppose that the initial mortalities have a gamma distribution. What must the underlying baseline hazard function be for the mixed mortality to be exactly Gompertz? This would be the solution to the equation

$$\exp\Big\{-\frac{\mu_0}{\theta t}\big(e^{\theta t} - 1\big)\Big\} = \lambda\big(M(t)\big) = \big(1 + M(t)/r\big)^{-r}.$$

This yields a baseline hazard rate of

$$\mu_0 \exp\left\{\frac{\mu_0}{r\theta}e^{\theta t} + \theta t - \frac{\mu_0}{r\theta}\right\}. \tag{21}$$

Instead of an exponential baseline hazard rate, then, we would need a hyperexponential.

6. Applications

The practical value of Theorem 1 is the relationship it establishes between the level of an asymptotically flattening hazard rate and the exponent governing the lower tail of the frailty distribution adduced to account for it. Given observations for the limiting value of μ^*, we can estimate ρ. Using ρ, we can characterize qualitative properties of any frailty distribution that could generate such asymptotic behavior.

We go through the calculations in one leading case, and then present estimates for a variety of cases in tabular form. Assume a Gompertz-Makeham form $c + \alpha e^{\beta t}$ for the underlying baseline hazard μ. The quotient $\mu(t)/M(t)$ tends to the Gompertz slope parameter β independent of the additive Makeham term c. We take $\beta = 0.08$ and the limit of μ^* equal to 0.600, drawing on a human example described below. Theorem 1 tells us that the lower tail of any frailty distribution generating such values should have exponent $1+\rho = \lim\ \mu^*/\beta = 0.600/0.08 = 7.5$. If frailty were distributed in accordance with a full gamma distribution with unit mean, the corresponding coefficient of variation would be $(1 + \rho)^{-1/2} = 0.365$. Our estimates, however, characterize the lower tail of the frailty distribution without making any assumptions about its form for larger Z.

To interpret such an estimate for $1+\rho$, suppose that the tail behavior of the frailty distribution sets in at least below $Z = 1/2$. Proportional changes in hazard rates due to observed heterogeneity on the order of $1/2$ are commonplace among subgroups in human populations, so $Z = 1/2$ seems a reasonable benchmark for comparisons. The cumulative proportion of individuals with frailties less than some smaller Z, as a fraction of the proportion with frailties less than $1/2$ is given by $(2Z)^{1+\rho}$. Then one in ten-thousand of these robust individuals would have Z values less than $\hat{Z} = (1/2)10^{-4/(1+\rho)} = 0.146$ and one in a million would have values less than $\hat{Z} = (1/2)10^{-6/(1+\rho)} = 0.079$.

How extreme are such frailty values? With a Gompertz baseline hazard rate governing adult ages, as we have seen in Section 5, a frailty Z corresponds to a shift in the modal age of death of $(1/\beta)\log(1/Z)$, and this formula remains a good approximation with a Gompertz-Makeham baseline with parameters in the general range under consideration here.

In our example, an individual with frailty $Z = 0.146$ would have a modal age at death $(1/0.08)\log(1/0.146) = 24$ years later than the individual with average frailty.

In a cohort where members with average robustness typically live to around 73, these specially robust individuals would typically need to be living to 97 if the model is to hold good. For a strictly Gompertz baseline, in which a change in frailty corresponds to a shift in the whole hazard function, this comparison would apply at other ages as well. An 80 year old with $Z = .146$ would have to resemble a typical 56 year old with regard to risk of death. The one-in-a-million specially robust individual would experience a $(1/0.08)\log(1/0.079) = 32$ year advantage across life.

Heterogeneity on such a scale seems quite extreme, pushing the limits of the plausible. The advantage of these calculations is that they make no assumptions about the parametric form of the frailty distribution or the overall level or dispersion in initial mortality rates across the population. They depend on (1) the observation or extrapolation of an upper asymptote for the aggregate hazard rate and (2) a choice for baseline hazard.

Our examples for numerical comparison, including the one already under discussion, are based on examples in the comprehensive study by Horiuchi (2003), supplemented by empirical studies in the same volume (Carey and Tuljapurkar 2003). Higher estimates for asymptotes relax the requirements on the lower tail of the frailty distribution to some extent. Table 1 shows three alternative estimates for humans.

Table 1 here

The value $\beta = .08$ in Table 1 is chosen to agree with the maximum of the Lifetable Aging Rates in Horiuchi's Figure 2A. The figure is predicated on the cohort lifetable for Swedish men born from 1880 to 1885 available on the Human Mortality Database (Human Mortality Database January 2004). The three columns differ in the choice of asymptote. The asymptote for the first column is chosen to accord with estimated hazard rates over 105 estimated by Robine and Saito for recent Japanese females (Robine and Saito 2003). The asymptotes for the second and third columns represent the lowest and highest of a set of extrapolations obtained by Wilmoth and Robine (2003, p. 251) from fitting logistic models to two super-centenarian databases. The middle case, A2, roughly matches Horiuchi's own extrapolation. Reliance on the logistic model introduces some degree of circularity from the point of view of the comparisons undertaken here, and the asymptotes in the second and third columns may be on the high side. They exceed the

values empirically observed among the supercentenarians in the data bases. The higher asymptotes do imply thinner lower tails for frailty distributions. The one-in-ten-thousand comparisons among individuals with frailties less than $Z = 1/2$, computed as above, give increments to modal ages at death in the range of 16 to 18 years, still high but easier to imagine than 24 years.

Table 2 show similar comparisons based on four examples for invertebrates discussed by Horiuchi (2003). The empirical values are taken from Horiuchi's cases B through E in his Figures 1 to 3. (His case F gives the same table values as case E.) The data sets are drawn from Carey et al. (1995) for populations of Mediterranean fruit flies and from Vaupel et al. (1998, Figure 3) for populations of nematodes, parasitoid wasps, and *Drosophila* melanogaster. For details, see Horiuchi (2003, Appendix A).

Table 2 here

The estimates for invertebrates in Table 2 show that the inclusion of extremely robust individuals with exceptionally low frailties would be required to generate the observed asymptotes in hazard rates by selective culling alone. For medflies, with their noticeable drops in aggregate hazard functions, it has been obvious that heterogeneity in frailty alone cannot reconcile Gompertz baseline hazards with observed population curves. For the other invertebrates the increments in modal lifespans given in the table are vastly larger than the modal lifespans themselves. These increments would be experienced by one in ten-thousand of the individuals who already had frailties less than half the average. Heterogeneity on such a scale seems highly unlikely and simple culling does not seem to be a plausible explanation for the observed asymptotes. This conclusion agrees with Horiuchi's findings, but it relies on much weaker parametric assumptions than his approach.

Implicit in the human estimates in Table 1 is the use of Gompertz-Makeham hazards, including a constant along with an exponentially increasing term. Since the asymptotics of Theorem 1 are the same with a Gompertz-Makeham baseline as with an ordinary Gompertz baseline, the distinction does not affect the tabulated values here. But it does affect the goodness of fit to human cohort lifetables that can be achieved with frailty models over the middle age range. Not only the Swedish male cohort of 1880-85, but other recently completed cohorts from developed nations including France and Japan represented in the Human Mortality Database show steepening in the graph of log hazard rates before the onset of the flattening that is our chief concern.

The addition of the Makeham constant is one traditional way of accounting for the temporary steepening, seen clearly in Figure 3A of Horiuchi (2003). Horiuchi favors an alternative approach, in which the loglinear Gompertz hazard is replaced by a log-quadratic baseline hazard. The extra acceleration built into the baseline hazard by the quadratic term turns out, as mentioned in Section 5.1, to preclude any asymptotic leveling of the population hazard function no matter how strong the heterogeneity in frailty within the power-law setting considered in Theorem 1. The ratio $\mu(t)/M(t)$ and the aggregate population hazard rate μ^* ultimately grow linearly with t.

Horiuchi offers a biological rationale for the log quadratic model (Horiuchi 2003, p. 142-143) which makes an interesting comparison to the traditional rationale for the Makeham model in terms of a nonsenescent component of mortality. Flattening in μ^* with a log-quadratic baseline hazard is a temporary phenomenon balancing superexponential increases in the baseline with thin tails in the frailty distribution. Sufficiently strong heterogeneity can postpone the cessation of flattening out to ages beyond what would be observed in finite human populations, making it hard to distinguish empirically between Makeham and log-quadratic baseline specifications. The theoretical differences between the models are striking, and deserve further scrutiny. But it is also well to bear in mind that both are variants on an underlying Gompertzian theme with respect to baseline hazard functions.

When one turns to frailty distributions, the problematic feature is the arbitrariness inherent in the choice of any parametric family. The popular candidate, the gamma family, is favored largely for convenience and not for deeper reasons. An advantage of the approach taken here is the decoupling of properties determined by the lower tail of the frailty distribution from properties that involve the overall dispersion. Calculations based on an assumption of gamma-distributed frailties tie values for the coefficient of variation (CV) in initial mortality rates directly to the level of the asymptote. Thus, for example, one can predict the estimated CV's in Table 1 of Horiuchi (2003) quite closely from the asymptotes in Figure 2 of Horiuchi (2003). One may have the impression that one is obtaining information about dispersion in unobserved heterogeneity, when that information is largely an artifact of reliance on the gamma distribution.

The shape of the lower tail does put some constraints on the standard deviation, given the requirement of a unit mean, but only loose constraints. Observations of the population hazard function tell little about the upper tail of the frailty distribution. It is prudent to avoid comparisons that depend on the upper tail, such as the difference in

frailty between 5-th and 95-th percentiles. Restricting attention to the lower tail, as we have been doing, allows conclusions about the plausibility of frailty values to be inferred from observable asymptotic behavior. Concentration on the generic properties of frailty distributions rather than on parametric forms allows a franker assessment of the ability of models for heterogeneity in frailty to account for deceleration in mortality trajectories.

References

Abramowitz, M. and Stegun, I.: 1965, *Handbook of mathematical functions, with formulas, graphs, and mathematical tables*, Dover, New York.

Brooks, A., Lithgow, G. J. and Johnson, T. E.: 1994, Mortality rates in a genetically heterogeneous population of Caenorhabditis elegans, *Science* **263**(5143), 668–71.

Carey, J. R.: 2003, *Longevity: The Biology and Demography of Life Span*, Princeton University Press.

Carey, J. R. and Tuljapurkar, S.: 2003, *Life Span: Evolutionary, Ecological, and Demographic Perspectives*, Population Council. Pop. Dev. Rev. vol. 29 suppl.

Carey, J. R., Liedo, P., Orozco, D. and Vaupel, J. W.: 1992, Slowing of mortality rates at older ages in large medfly cohorts, *Science* **258**(5081), 457–61.

Carey, J. R., Liedo, P., Orozco, D., Tatar, M. and Vaupel, J. W.: 1995, A male-female longevity paradox in medfly corhorts, *Journal of Gerontology:Biological Sciences* **56**(2), 89–93.

Doetsch, G.: 1950, *Handbuch der Laplace-Transformation*, Vol. 1, Verlag Birhäuser, Basel.

Drapeau, M. D., Gass, E. K., Simison, M. D., Mueller, L. D. and Rose, M. R.: 2000, Testing the heterogeneity theory of late-life mortality plateaus by using cohorts of *drosophila melanogaster*, *Experimental Gerontology* **35**, 71–84.

Feller, W.: 1971, *An Introduction to Probability and its Applications*, Vol. 2, John Wiley & Sons, New York.

Gavrilov, L. A., Gavrilova, N. S., Olshansky, S. J., Carnes, B. A., Cassel, C. K., Nusbaum, T. J., Graves, J. L., Mueller, L. D., Rose, M. R., Vaupel, J. W., Carey, J. R. and Curtsinger, J. W.: 1993, Fruit fly aging and mortality, *Science* **260**(5114), 1565–9.

Horiuchi, S.: 2003, Interspecies differences in the lifespan distribution: Humans versus invertebrates, *in* J. R. Carey and S. Tuljapurkar (eds), *Life Span: Evolutionary, Ecological, and demographic perspectives*, Vol. 29 (Supplement) of *Population and Development Review*, The Population Council, New York, pp. 127–151.

Horiuchi, S. and Coale, A.: 1990, Age patterns of mortality for older women: An analysis using the age-specific rate of mortality change with age, *Mathematical Populatioin Studies* **2**, 245–267.

Horiuchi, S. and Wilmoth, J. R.: 1997, Age patterns of the lifetable aging rate for major causes of death in japan, 1951-1990, *Journal of Gerontology: Biological Sciences* **52**(1), B67–B77.

Horiuchi, S. and Wilmoth, J. R.: 1998, Deceleration in the age pattern of mortality at older ages, *Demography* **35**(4), 391–412.

Human Mortality Database: January 2004, *Available at www.mortality.org*, University of California, Berkeley, USA, and Max Planck Institut für Demografische Forschung, Rostock, Germany.

Kowald, A., Kirkwood, T. B. L., Robine, J.-M., Ritchie, K., Vaupel, J. W., Carey, J. R. and Curtsinger, J. W.: 1993, Explaining fruit fly mortality, *Science* **260**(5114), 1664–6.

Lynch, S. M. and Brown, J. S.: 2001, Reconsidering mortality compression and deceleration: An alternative model of mortality rates, *Demography* **38**(1), 79–95.

Pletcher, S. D. and Curtsinger, J. W.: 1998, Mortality plateaus and the evolution of senescence: Why are old-age mortality rates so low?, *Evolution* **52**(2), 454–64.

Robine, J.-M. and Saito, Y.: 2003, Survival beyond age 100: The case of japan, *in* J. R. Carey and S. Tuljapurkar (eds), *Life Span: Evolutionary, Ecological, and demographic perspectives*, Vol. 29 (Supplement) of *Population and Development Review*, The Population Council, New York, pp. 208–228.

Steinsaltz, D.: 2005, Reevaluating a test of the heterogeneity explanation for mortality plateaus, *Experimental Gerontology* **40**, 101–13.

Vaupel, J. W.: 1997, Trajectories of mortality at advanced ages, *in* K. W. Wachter and C. E. Finch (eds), *Between Zeus and the Salmon: The Biodemography of Longevity*, National Academies Press, Washington, D.C., chapter 2, pp. 17–37.

Vaupel, J. W. and Carey, J. R.: 1993, Compositional interpretations of medfly mortality, *Science* **260**(5114), 1666–7.

Vaupel, J. W., Carey, J. R., Christensen, K., Johnson, T. E., Yashin, A. I., Holm, N. V., Iachine, I. A., Kannisto, V., Khazaeli, A. A., Liedo, P., Longo, V. D., Zeng, Y., Manton, K. G. and Curtsinger, J. W.: 1998, Biodemographic trajectories of longevity, *Science* **280**(5365), 855–60.

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

Wachter, K. W.: 2003, Hazard curves and life span prospects, *in* J. R. Carey and S. Tuljapurkar (eds), *Life Span: Evolutionary, Ecological, and demographic perspectives*, Vol. 29 (Supplement) of *Population and Development Review*, The Population Council, New York, pp. 270–291.

Wachter, K. W. and Finch, C. E. (eds): 1997, *Between Zeus and the Salmon: The Biodemography of Longevity*, National Academies Press, Washington, D.C.

Wang, J.-L., Müller, H.-G. and Capra, W. B.: 1998, Analysis of oldest-old mortality: lifetables revisited, *Annals of Statistics* **26**(1), 126–63.

Wilmoth, J. R. and Robine, J.-M.: 2003, The world trend in maximum lifespan, *in* J. R. Carey and S. Tuljapurkar (eds), *Life Span: Evolutionary, Ecological, and demographic perspectives*, Vol. 29 (Supplement) of *Population and Development Review*, The Population Council, New York, pp. 239–257.

Yashin, A. I., Iachine, I. A. and Begun, A. S.: 2000, Mortality modeling: A review, *Mathematical Population Studies* **8**(4), 305–32.

TABLE 1. Asymptotes and Frailty Estimates for Humans

Species	Humans		
Case	A1	A2	A3
β per year	.08	.08	.08
Asymptote μ^*	.600	1.000	1.250
Estimated $1+\rho$	7.50	12.50	15.63
$(1+\rho)^{-1/2}$	.365	.283	.253
$\hat{Z}$ (per 10^4)	.146	.239	.277
Increment in years	24	18	16

TABLE 2. Asymptotes and Frailty Estimates for Invertebrates

Species	Medflies	Nematodes	Wasps	Drosophila mel.
Case	B	C	D	E
β per day	.04	.06	.02	.08
Asymptote μ^*	.050	.010	.001	.010
Estimated $1+\rho$	1.25	1.667	.500	1.250
$(1+\rho)^{-1/2}$	.894	.775	1.414	.894
$\hat{Z}$ (per 10^4)	$3*10^{-4}$	$2*10^{-3}$	$5*10^{-9}$	$3*10^{-4}$
Increment in days	202	104	956	101
Modal lifespan	20	15	6	45