ON EXPECTATION OF LIFE IN BIVARIATE SURVIVAL ANALYSIS

Lorenzo Moreno

Mathematica Policy Research, Inc.
Princeton, NJ 08540-2393

June 1992

Short title: On conditional life expectancy

Presented at: The Population Association of America, April 1992, Denver, Colorado.

ABSTRACT

Survival analysis of groups, particularly pairs, of related individuals has been receiving growing attention by biostatisticians, demographers, econometricians and other social scientists. In this article, an interpretation of the properties of bivariate life table functions for individuals who have common frailty is provided. In particular, expressions for the conditional life expectancy — assuming that pairs of linked individuals face the same baseline hazard functions — for three parametric duration models commonly used in demographic analysis, i.e., exponential, Weibull and Gompertz distributions are derived. The implications of assuming that frailty follows a gamma or an inverse Gaussian distributions are discussed. This work illustrates that the pace at which improvements in the survivorship of one of the members of the pair translates into a higher mean duration of life is directly related to the variance of frailty and, more importantly, to the assumptions about its distribution.

Key words: *Bivariate survival analysis, random effects, kindred frailty, expectation of life.*

1. INTRODUCTION

Survival analysis of related or linked individuals has been receiving growing attention by demographers, economists and sociologists. For instance, recent studies of the determinants of infant and child mortality have explicitly accounted for the fact that the offspring of the same woman share common unobserved or unobservable environmental or genetic factors that may affect their chances of individual survival (Curtis, McDonald and Diamond, 1991; Guo and Rodríguez, 1991; Zenger, 1991). Also in the mortality setting, Vaupel (1988) has studied the role that inherited frailty, i.e., a set of susceptibles and risk factors that alter the chances of death of an individual at all ages, plays in explaining the correlation between life spans of parents and their children. In addition, he has proposed to analyze the genetic and environmental components of the longevity of monozygotic twins (Vaupel, 1990a and 1990b). Other recent examples of studies that have modelled the dependence between the lifetimes of twins are the work of Hougaard, Harvald and Holm (1992) using Danish data, and the analysis by Guo and Grummer-Strawn (1992) on the determinants of infant mortality among twins using survey data from low-income countries. Mare and Palloni (1988) studied couple survival data from the Panel Study of Income Dynamics to assess the socioeconomic effects on the mortality of older individuals. In a different context, Larsen and Vaupel (1989) analyzed the pattern of effective fecundity over age for Hutterite couples. Their analysis, however, is based on related events rather than individuals. Anderton et al. (1987) studied the intergenerational relationship of marriage and other fertility-related events utilizing proportional hazards models, arguing that "fertility behaviors transmitted through the family are behavioral propensities relative to prevailing social behavior". Another example is the work of Haurin and Mott (1990) on adolescent sexual activity, who examined the influence of an older sibling's age at first sexual intercourse on the sexual initiation of a younger sibling.

Based on the extensive work done by statisticians in constructing multivariate generalizations of the proportional hazards model (Clayton, 1978; Clayton and Cuzick, 1985; Cox and Oakes, 1984, Chapter 10; Holt and Prentice, 1974; Hougaard, 1986b; Oakes, 1982 and 1989), most of the studies of bivariate (or more generally, multivariate) survival analysis mentioned before have concentrated on testing and estimating an association parameter from a right-censored sample of pairs of life times, as well as the extent to which the estimates of the coefficients of observed covariates included in the model vary with the introduction of frailty effects. However, there is little work on the interpretation of the demographic properties of multivariate life-table functions induced by common dependence on an unobserved random effect.

In this article, I attempt partly to fill this gap by exploring bivariate life table functions, in particular those conditional on the survival experience of one of the members of the pair. The central objective is to discuss the demographic implications of the choice of a functional form for the distribution of the dependence of the members of a pair on an unobserved (or not adequately observed) random effect, as well as of the trajectory of the hazard function. To the best of my knowledge, there is little discussion of the demographic implication of these selections, with the exception of the work of Vaupel and Yashin (1985a and 1985b), Trussell and Richards (1985), Montgomery and Trussell (1986) and Trussell, Rodríguez and Vaughan (in press) in the case of univariate survival analysis. In addition, general expressions are provided for the conditional hazard and survival functions, as well as for the conditional life expectancy, under the assumption that the heterogeneity in the susceptibility to experiencing an event — *frailty* — follows a gamma or an inverse Gaussian distributions. These frailty models have been shown to be mathematically convenient, mainly because they are characterized by having the *closure* property, i.e., that the conditional distribution of frailty, given survival until a specific age or duration, is also gamma or inverse Gaussian but with a different scale parameter (Hougaard, 1984 and 1986a). This is an attractive attribute that I will utilize in the derivation of the results. In particular, I obtain expressions for the

conditional life expectancy — assuming that pairs of linked individuals face the same baseline hazard functions — for three parametric duration models commonly used in demographic analysis, namely: exponential, Weibull and Gompertz distributions. Examples derived from published estimates for human populations will be used to illustrate the results. Finally, I explore the implications of these findings in demographic research.

2. THE ASSOCIATION MODEL

2.1 Preliminaries

Let T_1 and T_2 denote the lifetimes of individuals in a pair. Central to this work is the assumption that these individuals share a common set of unobservable characteristics, z, or *frailty* (Vaupel, Manton and Stallard, 1979) which is distributed over non negative values of z with density function $f(z)$ at the beginning of exposure. Under the strong, but frequently used, assumptions that this common factor is lifetime-invariant and has no time-dependent effects, frailty is equivalent to relative risk in a proportional-hazard model (Ibid.). Hence,

$$\lambda_i(t_i \mid z) = \lambda_{0i}(t_i)\, z \tag{1}$$

denotes the hazard function for the i-th member of the pair. Although the baseline hazard $\lambda_{0i}(t_i)$ can be allowed to have different specifications for the components of the duo, as suggested by Clayton (1978), I will assume it equal.

Conditional on the common frailty z, the T_i are mutually independent, and hence imply that the (conditional) bivariate survival function takes the form

$$S(t_1,t_2 \mid z) = [S_0(t_1)]^z\, [S_0(t_2)]^z \tag{2}$$

It follows that the unconditional joint bivariate survival function is given by

$$S(t_1,t_2)=\int_0^{\infty} e^{-z[\Lambda_0(t_1)+\Lambda_0(t_2)]} f(z)dz \qquad (3)$$

where $\Lambda_0(t_i)=\int_0^{t_i}\lambda_0(u)du$ denotes the integral or cumulative baseline hazard for the i-th member of the couple.

2.2 The distribution of frailty

The specification of the distribution of the common factor z is crucial to any analysis of bivariate life tables because, as shown by Oakes (1989), it uniquely determines the functional form of the association between the observed lifetimes of the individuals in the pair. In other words, the specification of the distribution of the unobservable across the population of pairs (or mixing function), uniquely *induces* the degree of dependence between T_1 and T_2.

The utilization of a non-parametric, discrete probability distribution as a mixing function of the underlying hazards is closely identified with the work of Heckman and Singer (1982 and 1984) on univariate survival analysis with unobservables. However, Vaupel used a two-point distribution in his analysis of the correlation of the lifetimes of fathers and sons (1988). Vaupel also employed a non-parametric distribution to determine when, in a clinical trial of a new therapy, the experiment must stop if there is evidence of adverse consequences on the group receiving the treatment (1990a). More common, however, has been the utilization of continuous functions. In particular, the gamma

distribution — popularized by Vaupel et al. (1979) in demographic applications — has been favored because of the flexibility of shapes offered, its parsimony, as well as its mathematical tractability [Clayton (1978), Oakes (1982), Clayton and Cuzick (1985), and others]. The inverse Gaussian distribution has also been advocated in the univariate analysis of survival times with an unobserved covariate by Aalen (1988), Hougaard (1984 and 1986a) and, in the demographic context, by Vaupel and Yashin (1985a). Another continuous distribution frequently used in demographic applications, e.g., the study of fecundability, is the beta distribution (Sheps and Menken, 1973, pp. 73-74; Keyfitz, 1985, pp. 395-399; Heckman and Walker, 1990).

In this document, I will concentrate on the implications of assuming that the heterogeneity in the shared unobserved characteristics follows a gamma or inverse Gaussian distributions among pairs of individuals. My interest on these distributions arises from the now common utilization of these distributions in modelling random effects in survival models.

2.2.1. The gamma distribution case

Assume that frailty z has a gamma distribution with density

$$f(z)=z^{\eta-1}e^{-\eta z}\eta^{\eta}/\Gamma(\eta) \tag{4}$$

so the mean is unity and the variance is $\sigma^2 = \eta^{-1}$. As long as the baseline hazard in (1) includes a constant term, the assumption that mean frailty is one involves no loss of generality.

It follows immediately from (3) and (4) that the joint bivariate survival function is given by

$$S(t_1,t_2)=[1+\sigma^2(\Lambda_0(t_1)+\Lambda_0(t_2))]^{-\eta} \tag{5}$$

From the conditional probability formula [i.e., $Pr(A|B) = Pr(A \cap B) / Pr(B)$], and the expression for the marginal survival function

$$S(t_i)=S(0,t_i)=S(t_i,0)=[1+\sigma^2\Lambda_0(t_i)]^{-\eta} \tag{6}$$

(Oakes, 1989), it follows that the probability of one member surviving to t_2 given that the other member of the pair has survived to t_1, is

$$S(t_2|t_1)=[1+\sigma^2\Lambda_0(t_1)/(1+\sigma^2(\Lambda_0(t_1)+\Lambda_0(t_2)))]^{\eta} \quad . \tag{7}$$

In addition, the conditional hazard function takes the form

$$\lambda(t_2|t_1)=\lambda_0(t_2)/[1+\sigma^2(\Lambda_0(t_1)+\Lambda_0(t_2))] \quad . \tag{8}$$

Note that, unless there is no heterogeneity across the population (i.e., the variance of frailty is zero at the beginning of exposure), the conditional hazard decreases with respect to t_1. For example, in the context of familial survivorship, this implies that the age-specific rates for offsprings decline with increased survival time of their predecessors. This is an interesting result because, as will be shown below, gains in the expectation of life of, say, the father, might translate into even greater improvements in life expectancy for the son.

As shown by Clayton (1978), and later confirmed by Oakes (1979), the selection of a gamma distribution of frailty uniquely implies that the association between survival times of the members of a pair is constant for all durations. In other words, the measure of the degree of dependence, i.e., the ratio of the hazard rate of the conditional distribution of T_2, given $T_1 = t_1$, to that of T_2 given $T_1 > t_1$,

$$\theta(t_1,t_2) = \lambda(t_2 \mid T_1 = t_1) \,/\, \lambda(t_2 \mid T_1 > t_1) \tag{9}$$

is constant (say, θ) across time.

Moreover, Clayton showed that

$$\theta = 1 + \eta^{-1} = 1 + \sigma^2 \tag{10}$$

so that the variance of frailty can be interpreted as a "component of variance" (p.146) of survival times of the pairs of individuals. Moreover, when there is no heterogeneity in the population, i.e., $\sigma^2 = 0$, the association parameter θ is unity, which implies that the lifetimes within pairs are independent. As argued by Vaupel (1990b), *kindred-frailty models* or models for related individuals, provide the "conceptual basis for dividing the heterogeneity among individuals into a frailty component and a residual component given by the distribution of lifetimes of individuals with the same frailty" (p. 171).

Finally, an interesting result derived by Oakes (1989) relates the measure of dependence, $\theta(t_1, t_2)$ to Kendall's coefficient of concordance τ (Kendall, 1938). For instance, in the case of the gamma distribution,

$$\theta = (1 + \tau) \,/\, (1 - \tau) \tag{11}$$

and, using the delta method or method of statistical differentials (Namboodiri and Suchindran, 1987; pp. 115-116), its variance is given by

$$\text{var}(\theta) = 4 \,\text{var}(\tau) \,/\, (1 - \tau)^4. \tag{12}$$

These results provide an inferential procedure to assess the degree of association between pairs of individuals, even from censored data[1], directly from an easily interpretable measure, such as Kendall's coefficient of concordance. In other words, it is easy to test for independence between lifetimes using expressions (11) and (12). This simple test is achieved, of course, at the cost of assuming that frailty is gamma-distributed among population pairs. If the null hypothesis is rejected,

the issue is then the estimation of the degree of dependence between observations, which I briefly discuss below (see Section 3).

2.2.2. The inverse Gaussian distribution case

Now assume that frailty z has an inverse Gaussian distribution with density

$$f(z)=(\kappa/\pi)^{1/2}e^{2\kappa}z^{-3/2}e^{-\kappa(z+z^{-1})} \tag{13}$$

so that, without loss of generality, the mean is unity and the variance is $\sigma^2 = (2\kappa)^{-1}$.

From (3) and (13), the joint bivariate survival distribution is given by

$$S(t_1,t_2)=e^{-2\kappa[(1+2\sigma^2(\Lambda_0(t_1)+\Lambda_0(t_2)))^{1/2}-1]} \tag{14}$$

and based on similar arguments as in the case of the gamma distribution, it can be shown that the conditional survivorship function is given by

$$S(t_2|t_1)=e^{-2\kappa[(1+2\sigma^2(\Lambda_0(t_1)+\Lambda_0(t_2)))^{1/2}-(1-2\sigma^2\Lambda_0(t_1))^{1/2}]} \tag{15}$$

and the conditional hazard now takes the form

$$\lambda(t_2|t_1)=\lambda_0(t_2)/[1+2\sigma^2(\Lambda_0(t_1)+\Lambda_0(t_2))]^{1/2} \tag{16}$$

Again, as in the case of the gamma distribution, the conditional duration-specific rate for one member of the pair declines with increased survivorship of other individual in the duo. If the lifetimes are independent, i.e., $\sigma^2 = 0$, then the conditional hazard is identical to the baseline hazard.

As stated before, the specification of an inverse Gaussian distribution of frailty uniquely determines the functional representation of the dependence between lifetimes. Hence, in this case, $\theta(t_1,t_2)$ is inversely related to the bivariate survivorship function, as shown by Oakes (1989). For instance, after some algebraic manipulation, the measure of association takes the form

$$\theta(t_1,t_2)=1+\sigma^2[1+2\sigma^2(\Lambda_0(t_1)+\Lambda_0(t_2))]^{-1/2} \quad . \quad (17)$$

This expression indicates, according to Oakes's formulation (1989), that knowledge that an individual died at t_1 increases the hazard for the other member of the pair by a time-varying percentage to what it would be if the other member had not failed at duration t_1.

An alternative interpretation of $\theta(t_1,t_2)$ is possible by noting that the last term in equation (17) is equivalent to the expectation of frailty among survivors at t_2 given that the other member of the pair survived to T_1, i.e., $E[\, z \mid T_2 > t_2;\, T_1 > t_1 \,]$. This is a generalization to the bivariate case of one of the better-known results in unobserved heterogeneity for the univariate case (Sheps and Menken, 1973; Vaupel, Manton and Stallard, 1979; and Hougaard, 1984). Hence

$$\theta(t_1,t_2) = 1 + \sigma^2 \;\; E[\, z \mid T_2 > t_2;\, T_1 > t_1 \,]. \quad (18)$$

Note that this measure decreases from $1 + \sigma^2$ when $t_1 = t_2 = 0$ — an identical expression as in the case of the gamma distribution [*cf.* (10)] — to unity when both lifetimes approach infinity. Moreover, as $\sigma^2 \to 0$, $\theta(t_1,t_2) \to 1$, which implies independence between the lifetimes.

To illustrate the discussion of the expression derived above, as well as other results to be presented below, I have selected three parametric distributions — exponential, Weibull and Gompertz, which are commonly used in demographic research (Vaupel, 1990a; p. 160). The hazard and integrated hazard functions for these three parametric representations of the baseline lifetime distribution for each of the members of a pair are presented in Table I. The exponential distribution has been extensively used in studies of human fecundability (Sheps and Menken, 1973; p. 73; Aalen, 1987) and IUD expulsion (Aalen, 1987), whereas the Weibull and Gompertz distributions have been traditionally used to study mortality: cause-specific and total, respectively (Horiuchi and Coale, 1990; Manton et al., 1981 and 1986). All these parametric functions correspond to extreme-value distributions (Kalbfleisch and Prentice, 1980; pp. 21-30), and their statistical properties are studied elsewhere (e.g. Sheps and Menken, 1973).

2.3 Illustration of the measure of association $\theta(t_1, t_2)$ between survival times

A graphical representation of $\theta(t_1, t_2)$ is rendered in Figure 1 when the Gompertz trajectory (left-hand panel) and the Weibull distribution (right-hand panel) represent the underlying hazard function $\lambda_0(t)$. The baseline hazard is supposed to be equal for both individuals in the pair. That is, the members of the duo share the same *underlying risk*, as well as the same level of frailty (cf. Section 2.1). This assumption, although restrictive and possibly unrealistic, is based on the interest to illustrate, *ceteris paribus*, how $\theta(t_1, t_2)$, as well as other results related to the conditional hazard and expectation of life (shown below), strongly depend on the assumption of the distribution of frailty among pairs.

The values of the parameters for these distributions were selected to resemble those reported by Manton et al. (1986) in their analysis of overall and cause-specific mortality in the United States using Medicare and National Center for Health Statistics (NCHS) data. For the Gompertz, α is equal to $1x10^{-3}$ for both members of the pair, and β takes the value of 0.05 for both individuals. The latter parameter represents the annual increase in the force of mortality. For the Weibull distribution, which is commonly fit to lung cancer mortality data, α takes values of $1.25x10^{-5}$ for both individuals, and β is set equal to 2.45. The latter parameter represents the slope of the logarithm of the cumulative hazard rate as a function of the logarithm of age. The parameters selected for both distributions represent approximately the mortality conditions of the male and female cohorts born in high-income countries with an expectation of life above age 70. The parameter σ^2 is assumed to take values between 0 — homogeneity — and 0.5 — high heterogeneity — with increments of 0.05, values which researchers have assumed to encompass most situations to be encountered in human populations. Most empirical studies of mortality suggest that a σ^2 value of 0.25 fits human mortality data reasonably well (Manton et al., 1986; Mare and Palloni, 1988), although a recent study of the mortality of monozygotic and dyzogotic twins reveals that the variance of frailty might be as high as 0.75 in human populations (Hougaard, Harvald and Holm, 1992; Table 5).

The bottom diagonal depicts the situation where the lifetimes of both individuals are independent, i.e., $\sigma^2 = 0$. When $\sigma^2 > 0$, $\theta(t_1, t_2)$ remains fairly constant for the first 50 years of life, thereafter rapidly declining towards unity in both distributions. This is the case when heterogeneity is high, so the rapid removal of frail individuals at early durations causes the association parameter to approach unity very rapidly after age 60, as illustrated for the Gompertz distribution. Indeed, for both representations of the underlying hazard, the graphs indicate that the effect of assuming an inverse Gaussian distribution rather than a gamma distribution will be reflected mainly among the older persons. As shown before [cf. Eq.(10)], the association parameter remains constant at a value of $1 + \sigma^2$ for all ages when frailty is gamma-distributed. The graph depicts clearly the magnitude of

the departure from this constant value when frailty follows an inverse Gaussian distribution.

Unfortunately, there is no simple relation between the coefficient of concordance τ and $\theta(t_1,t_2)$ — the estimation of τ involves the evaluation of an exponential integral function, $E_i(x)$ (Gradshteyn and Ryzhik, 1980) — although Oakes (1989) proposed the ratio $[\theta(t_1,t_2) - 1] / [\theta(t_1,t_2) + 1]$ as a conditional (on t_1 and t_2) version of Kendall's τ.

Note that the measure of association $\theta(t_1, t_2)$ is independent of the assumption about the baseline survival distributions when frailty is assumed to be gamma distributed, whereas this parameter depends on the functional representation of $\Lambda_0(t_i)$ (i=1,2) when the common characteristic to the members of the pair follows an inverse Gaussian distribution.

3. RESULTS ON CONDITIONAL HAZARD AND LIFE EXPECTANCY AT BIRTH

I next discuss some of the demographic implications of selecting a distribution of frailty in bivariate life table models. In particular, the interest is in studying the behavior of the conditional hazard and life expectancy functions.

3.1 The conditional hazard function

The behavior of $\lambda(t_2 \mid t_1)$ can directly be ascertained from equations (8) or (16), given that the common frailty follows either a gamma or an inverse Gaussian distribution. As mentioned before, the hazard rate for one member of the pair declines with increasing survivorship of the other member. However, this rate of change depends on the assumption made about the distribution of frailty. For

instance, by taking derivatives of the conditional hazard function with respect to t_1 in (8) and (16), the following expressions are obtained

$$d/dt_1[\lambda(t_2|t_1)] = -\sigma^2\lambda_0(t_1)\lambda_0(t_2)/[1+\sigma^2(\Lambda_0(t_1)+\Lambda_0(t_2))]^2 \quad (19)$$

and

$$d/dt_1[\lambda(t_2|t_1)] = -2\sigma^2\lambda_0(t_1)\lambda_0(t_2)/[1+2\sigma^2(\Lambda_0(t_1)+\Lambda_0(t_2))] \quad (20)$$

for the gamma and inverse Gaussian distributions, respectively. Note that the pace of decline is determined by the assumption about the distribution of frailty. The specific form of the derivatives depends on the selection of the functional form of the underlying hazard function.

In Figure 2, I present the conditional hazard $\lambda(t_2|t_1)$ and its derivative with respect to t_1 under the assumption that the underlying hazard follows a Gompertz distribution, and frailty is distributed either as a gamma or as an inverse Gaussian distribution (left- and right-hand panels, respectively). The values of the parameters for the Gompertz distribution are the same as those used in Figure 1. The parameter σ^2 also takes values between 0 and 0.5 with increments of 0.05, for both distributions of frailty. This assumption implies that η takes values between 2 and ∞ for the gamma distribution, and κ ranges between 1 and ∞ for the inverse Gaussian distribution.

The situation when lifetimes are independent — so that the conditional hazard is equal to the underlying hazard $\lambda_0(t_2)$ — is depicted by the first line from the left. When the variance of frailty becomes positive, the conditional hazard $\lambda(t_2 \mid t_1)$ is always lower than $\lambda_0(t_2)$ for all values of t_1. However, the departure is evident only at older ages. This can be seen from the derivative of the conditional hazard with respect to t_1, shown in the lower panels of Figure 2. As stated before, this

function is always negative when $\sigma^2 > 0$ and increases (in absolute terms) with increases in heterogeneity. The largest impact of changes in t_1 on the conditional hazard occurs for ages above 60 years. Note also that the assumption that frailty follows a gamma distribution implies a slower pace of decline than when this distribution is an inverse Gaussian curve. The same situation arises when the underlying hazard is assumed to follow a Weibull distribution (not shown). Hence, because of the common influence shared by the members of the pair, improvements in the survivorship of one induces a smaller reduction in the hazard of the other member under the assumption that frailty follows a gamma distribution rather than an inverse Gaussian distribution.

Overall, the previous analysis implies that improvements in the mortality of one of the members of the pair may translate into higher expectations of life at birth for the other associate. This is discussed below.

3.2 The conditional expectation of life at birth

The expectation of life at birth is obtained by integrating the conditional survival function [*cf.* equations (7) or (15)],

$$e_o(T_2|t_1) = \int_0^\omega S(u|t_1)du \quad . \tag{21}$$

The solution to this integral can be greatly simplified if the results are presented in terms of the three parametric distributions discussed before. The conditional expectation of life at birth for the two distributions of frailty (i.e., gamma and inverse Gaussian distributions) and the distributions of the underlying lifetimes in the pair (i.e., exponential, Weibull and Gompertz functions) are reported

in Table II. As is clear from the results shown, the formulae for estimating the conditional expectation of life become quite cumbersome, in particular when heterogeneity is represented by the inverse-Gaussian distribution. The evaluation of several of the integrals reported would require the utilization of numerical methods.

However, it is still possible to interpret these expressions. For instance, use of the closure property of the distributions of frailty,[2] Bayes's theorem and the conditional probability formula reveal that *the conditional expectation of life at birth has the same form as does the (cohort) expectation of life in the univariate case*[3] but with a different scale parameter. For instance, this takes the form $\eta^* = \eta + \Lambda_0(t_1))$ for the gamma distribution; and $\kappa^* = \kappa + \Lambda_0(t_1)$ for the inverse Gaussian distribution (see Table II). These parameters are a function of the survival time of the other member of the pair, more specifically, the cumulative (baseline) hazard function. Moreover, the mean is different from unity for both distributions.[4] These results imply that the conditional expectation of life is equivalent to mixing a baseline life expectancy at birth (i.e., exponential, Weibull or Gompertz distributions) among those with frailty z^*, i.e., $e_o(z^*)$, with a distribution of heterogeneity which is either gamma with parameters η and η^*, or inverse Gaussian with parameters κ and κ^*. In general,

$$e_o(T_2|t_1)=\int_0^\omega e_o(z^*)g(z^*|t_1)dz^* \tag{22}$$

where $g(z^*|t_1)$ is the density function of frailty given that the other member of the pair survived to t_1.

Instead of attempting to offer a general interpretation of the behavior of the conditional expectation of life as the lifespan of one of the members of the pair varies, I will illustrate the discussion with a specific example — the case when the underlying hazard follows a Weibull

distribution and frailty is gamma or inverse Gaussian distributed (see Figure 3). The parameters for the Weibull distribution are $\alpha = 1.25 \times 10^{-5}$ and $\beta = 2.45$, as in the example presented in Figure 1.

As anticipated, improvements in the survival of one of the members of the pair translates, in most instances, into larger increases in life expectancy for the other associate. However, the magnitude of the impact of prolonged survivorship depends on both the level of heterogeneity shared by the members of the pair and, prominently, by the distribution of frailty. For instance, if the common (unobserved) characteristic is assumed to follow a gamma distribution among the members of the population (upper panel), only when **both** heterogeneity assumes atypical values (i.e., $\sigma^2 \geq 0.3$) **and** one of the members of the pair has survived to age 80 or higher the conditional life expectancy increases as the other individual survives to older ages. If the distribution of frailty follows an inverse Gaussian distribution, however, the conditional life expectancy rapidly increases as the other individual survives to ages 60 years or older. For instance, when $\sigma^2 = 0.25$ and the lifespan of the first individual reaches the age of 60, the second member of the group is expected to survive to about age 80 years. This is not necessarily the case when the distribution of heterogeneity has a lower variance (i.e., below 0.10); in this case a longer lifespan for one of the individuals does not necessarily translate into a higher expectation of life at birth for the other member of the duo. Still, as shown in Figure 4, the difference between the conditional life expectancy when frailty follows an inverse Gaussian distribution and when frailty is gamma distributed — for a given level of heterogeneity — is quite startling and increases the longer is the lifespan of the first individual in the pair. Finally, note that the rate of improvement in life expectancy also depends on the parameters of the underlying hazard function.

The explanation of these patterns lies in the fact that the (conditional) probability distribution of the lifespan of one of the members of the pair is the weighted average of the distributions across frailty levels with weights given by the probability density functions selected, i.e., that of the gamma

or inverse Gaussian distributions. Because the inverse Gaussian distribution has a longer tail than the gamma distribution, surviving individuals are more homogeneous (with regard to their own frailty) when the former distribution is used to model the common characteristics of the individuals. This implies that individuals with higher longevity are more frequently found when the inverse Gaussian distribution is adopted than when the gamma distribution is assumed. Overall, for a given level of heterogeneity, the choice of the inverse Gaussian distribution to represent the dispersion of frailty instead of a gamma distribution may imply substantial gains in life expectancy for one member of the pair given the survival time of the other member of the duo.

3. ESTIMATION

Although estimation of the parameters of the models presented above goes beyond the objective of this work, it is important to mention the estimation methodologies that have been proposed.

Clayton (1978) provided an expression to estimate θ — the measure of association between survival times — by maximum likelihood when frailty is gamma-distributed and when a parametric representation is assumed for the baseline hazard. He also discussed a method to estimate θ when the analyst is unwilling to make parametric assumptions concerning the baseline function. Oakes (1982) examined the inference of the association measure after reparametrizing Clayton's model, provided expressions for the log-likelihood function and the information matrix, reviewed Clayton's method when the baseline hazards are completely unknown, and proposed as an alternative a non-parametric estimator based on Kendall's coefficient of concordance. Hougaard (1986b) used maximum likelihood methods to estimate the bivariate exponential and Weibull models (as particular cases of a more general multivariate model) assuming a special case of the inverse Gaussian distribution.[5] In all three instances, inference about the association parameter θ is used to test the *hypothesis of independence*

of the lifetimes of the members of the pair. Finally, Vaupel (1990a) proposed a general expression for the likelihood function of related or kindred lifetimes obtained by integrating the frailty distribution. This formulation is very elegant because the entire history of survival data on grouped individuals (combining death times and censoring times) can be summarized by three statistics based on the hazard function and the number of deaths: namely, the total log-hazard at observed death times, the total cumulative hazard, and the number of deaths. Vaupel's method can be used with a parametric representation for the baseline hazard and a wide range of frailty distributions, including the gamma and inverse Gaussian cases. However, Vaupel does not offer an interpretation of the parameters of the model in terms of the association function.

Considerable work has been done when the assumption about the parametric representation of the baseline hazard is relaxed to adopt Cox's semi-parametric representation of the baseline hazard (Cox, 1972), as well as to estimate the parameters of observable covariates. Holt and Prentice (1974) discussed the cases when uncensored and censored observations were available, but they did not make explicit assumptions about the distribution of frailty nor about the properties of the measure of association θ. They also compared their results with those for the exponential and Weibull distributions. Clayton and Cuzick (1985) developed a method for testing and estimating the association parameter from right-censored sample pairs using only the rank-order information under the assumption that frailty follows a gamma distribution. Their expressions for the likelihood function seem to be quite cumbersome. More recently, Guo and Rodríguez (1991) proposed the utilization of the EM algorithm (Dempster, Rubin and Laird, 1977) to estimate a multivariate hazard model also under the assumption that the heterogeneity among the individuals follows a gamma distribution, as well as when frailty is represented by a two-point non parametric distribution. They argue that use of the EM algorithm greatly simplifies the expressions that result from the direct extension of Clayton's method for the bivariate hazard model to the multivariate case.

4. IMPLICATIONS FOR DEMOGRAPHIC RESEARCH

Although it is certainly encouraging that recent demographic analyses of related individuals or events have modelled their shared dependence via an unobserved (or not adequately observed) random effect, it is judicious to ponder the implications of adopting an analytical framework that Vaupel has designated as *frailty modelling* (1990b). As several demographers have highlighted, theories and empirical findings of, say, biologists (in the context of mortality) are critical in the construction of frailty models, an exercise that consists of determining the functional forms to be used for both frailty distributions and hazards functions (Trussell and Rodríguez, 1990; Vaupel, 1990b; Weiss, 1990). Very often, it is mathematical convenience more than biological, social or economic reality that dictates the assumptions made. For instance, the postulate that frailty is gamma distributed is now a basic staple of multivariate hazard and frailty models due to the flexibility and mathematical tractability of this distribution.

In this paper I have exhibited one aspect of the consequences of imposing a specific trajectory to the distribution of the (unobserved) characteristic shared by individuals in a group, that has not yet received enough attention. This work illustrates how the selection of specific distributions of frailty *induces* a specific functional form to a measure of association between lifetimes, an important result due to Oakes (1989). For instance, the choice of a gamma distribution to represent unobserved heterogeneity implies that the ratio of the conditional hazard for one member of the pair at t_2 given that the other member died at t_1 to the conditional hazard given that the other member of the survived at least to age t_1 is constant at all durations. In contrast, the selection of an inverse Gaussian distribution, implies that the measure of association is a function of the bivariate survivorship function and, consequently, of the parametric representation assumed for the underlying force of mortality. Additional work may be needed on elucidating the form of $\theta(t_1, t_2)$ when discrete distributions of frailty, e.g. Poisson, binomial or N-point distributions, are selected. For some of

these distributions, which have a finite probability that a pair's frailty is zero, the model ensures infinite survival, a feature desirable in applications where events do not affect everyone (Oakes, 1989; Vaupel, 1990a). The discrete distributions are now regularly in use by econometricians in analyses of (univariate) survival data (e.g. Berhman, Sickles and Taubman, 1990) and, as properly noted by Vaupel (1990a), can be applied to situations "where individuals can be divided into discrete groups (...) and estimates are needed of the relative risk of each group".

I have also discussed the properties of bivariate life table functions conditional on the survival experience of one of the members of the pair. In particular, under the assumption that pairs of individuals share lifetime invariant frailty and have the same underlying risk of experiencing an event, the survival model implies that improvements in life expectancy for one of the members of the pair frequently translate into larger gains for the other group member. The pace of improvement depends, primarily, on the assumption about the distribution of unobserved heterogeneity across the population and, secondarily, on the chosen trajectory for the underlying hazard function. The implications of this result should be carefully pondered when applying frailty models to the study of bio-demographic or socio-economic issues. For instance, in the context of teenage pregnancy, if intergenerational transmission of fertility-determining behavior occurs within the family (Anderton et al., 1987; Kahn and Anderson, 1992), a random-effects bivariate survival model implies that if a group of mothers had a teenage pregnancy, then their daughters would be more likely to become pregnant during their adolescence than the daughters whose mother or sibling delayed their first pregnancy until the early thirties, ages when the population is very heterogeneous in regard to their propensity to have a teenage pregnancy. This is an important caveat for those researchers who utilize frailty models to study socio-economic processes. An equivalent argument can be made about the transmission of longevity or a genetic trait between parents and offspring, between twins, or between related events such as unemployment spells.

Although the knowledge of a person's life span provides fairly weak information about the person's susceptibility to experiencing death (Vaupel, 1988), the lifespan for an individual is of considerable importance for the other member of the pair, as shown before. The results, however, should be expected to be highly sensitive to the assumptions made on how to model the unobserved or unobservable characteristics shared by the members of the pair. Even when the variance of frailty is low, the assessment of the impact of prolonged lifespans on life expectancy is heavily dependent on the distribution of heterogeneity across the population.

Whereas the modelling of the distribution of frailty seems to render a firmer research basis to those interested in studying bivariate survival distributions, the results in this paper make it difficult to see how strong confidence can be placed on any *a priori* choice of the distribution of frailty without a sound theoretical basis. If the issue is then to acquire additional information on this prior, we may well have to go beyond the realm of survival analysis to support the selection.

ACKNOWLEDGMENTS

I am grateful to Noreen Goldman, James Trussell and Ian Diamond for their valuable comments to this paper. The Office of Population Research, Princeton University, supported part of this study. At Mathematica Policy Research, Rebecca Maynard facilitated my work on this document.

Table I. Hazard and integrated hazard functions for baseline failure time models.

Model	Hazard $\lambda_0(t)$	Integrated Hazard $\Lambda_0(t) = \int_0^t \lambda_0(y)\, dy$	
Exponential	α	αt	$\alpha > 0$
Weibull	$\alpha t^{\beta-1}$	$\alpha t^{\beta}/\beta$	$\alpha,\beta > 0$
Gompertz	$\alpha e^{\beta t}$	$\alpha e^{\beta t}/\beta$	$\alpha,\beta > 0$

Table II. Conditional expectation of life at birth, $e_o(T_2|t_1)$, for two distributions of frailty and three parametric models of failure.

Model	*Gamma Distribution*[a]
Exponential	$\eta^*/[\alpha(\eta-1)]$
Weibull	$\beta^{\gamma-1}(\eta^*/\alpha)^{\gamma}(\eta-\gamma)^{-1}\Gamma(\gamma)$ where $\gamma = \beta^{-1}$
Gompertz	$\beta^{\eta-1}\eta^{-1}[\eta^*/(\alpha-\beta\eta^*)]^{\eta}.{}_2F_1[1,\eta;\eta+1;-\eta^*((\alpha/\beta)-\eta^*))^{-1}]$
	$=\beta^{\eta-1}\eta^{-1}[\eta^*/(\alpha-\beta\eta^*)]^{\eta}.[1/B(\eta,1)\int_0^1 y^{\eta-1}[\eta^*\alpha/(\eta^*-\alpha-(\alpha\eta^*)y)]^{\eta}dy]$
	where ${}_2F_1$ is a Gauss Hypergeometric function and B(a,b) is a Beta function.
	Inverse Gaussian Distribution[b]
Exponential	$\alpha^{-1}e^{(\kappa\kappa^*)^{1/2}}(\kappa/\pi)^{1/2}[2(\kappa/\kappa^*)^{-3/4}K_{-3/2}(2(\kappa\kappa^*)^{1/2})]$ where K_υ is a Bessel function.
Weibull	$\Gamma(\gamma)(\beta^{1-\gamma}\alpha^{\gamma})^{-1}(\kappa/\pi)^{1/2}e^{2(\kappa\kappa^*)^{1/2}}[2(\kappa/\kappa^*)^{\delta/2}K_{\delta}(2(\kappa\kappa^*)^{1/2})]$ where $\gamma = \beta^{-1}$ and $\delta = -(1/2+\gamma)$ and K_δ is a Bessel function.
Gompertz	$-\beta^{-1}(\kappa/\pi)^{1/2}e^{2(\kappa\kappa^*)^{1/2}}\int_0^{\infty}E_i(-\alpha\beta^{-1}z)z^{-3/2}e^{(-\kappa^*-\beta z^{-1})}dz$ where E_i is an Exponential Integral function.

[a] $\eta^* = \eta + \Lambda_0(t_1)$. The density function is $f(z) = z^{\eta-1}\, e^{-\eta z}\, \eta^{\eta}\, /\Gamma(\eta)$ with mean 1 and variance $\sigma^2 = \eta^{-1}$.

[b] $\kappa^* = \kappa + \Lambda_0(t_1)$. The density function is $h(z) = (\kappa/\pi)^{1/2}\, e^{2\kappa}\, z^{-3/2}\, e^{-\kappa(z+1/z)}$ with mean 1 and variance $\sigma^2 = (2\kappa)^{-1}$.

REFERENCES

Aalen, O.O. (1988). Heterogeneity in Survival Analysis. *Statistics in Medicine* 7: 1121-1137.

Aalen, O.O. (1987). Two Examples of Modelling Heterogeneity in Survival Analysis. *Scandinavian Journal of Statistics* 14: 19-25.

Anderton, D.L., N.O. Tsuya, L.L. Bean and G.P. Mineau (1987). Intergenerational transmission of relative fertility and life course patterns. *Demography* 24: 467-480.

Behrman, J.R., R.C. Sickles, P. Taubman (1990). Age-specific death rates with tobacco smoking and occupational activity: sensitivity to sample length, functional form, and unobserved frailty. *Demography* 27: 233-250.

Clayton, D. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 65: 141-151.

Clayton, D. and J. Cuzick (1985). Multivariate generalizations of the proportional hazards model (with discussion). *Journal of the Royal Statistical Society*, Series A, 148: 82-117.

Cox, D.R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society* B 34: 187-220.

Cox, D.R. and D. Oakes (1984). *Analysis of Survival Data.* Chapman and Hall: London.

Curtis, S., J.W. McDonald and I. Diamond (1991). Random-effects models for birth intervals effects on infant mortality in Brazil. Paper presented at the 1991 PAA Meeting, Washington, DC.

Dempster, A.P., N.M. Laird and D.B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society* B 39: 1-38.

Gradshteyn, I.S. and I.M. Ryzhik (1980). *Tables of Integrals, Series and Products.* Academic Press: San Francisco.

Guo, G. and G. Rodríguez (1991). Estimating a Multivariate Proportional Hazards Model for Correlated Risks via the EM Algorithm, with an application to Child Survival in Guatemala. *Office of Population Research Working Paper* 91-12. Forthcoming in the *Journal of the American Statistical Association.*

Guo, G. and L. Grummer-Strawn (1992). Child Mortality among Twins from DHS Countries. Paper presented at the PAA Meeting, Denver, CO.

Haurin, R.J. and F.L. Mott (1990). Adolescent Sexual Activity in the Family Context: The Impact of Older Siblings. *Demography* 27: 537-557.

Heckman, J. and B. Singer (1982). Population heterogeneity in demographic models. In *Multidimensional Mathematical Demography*, edited by K. Land and A. Rogers. Academic Press: New York.

Heckman, J. and B. Singer (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* **52**: 231-241.

Heckman, J. and J. Walker (1990). Estimating Fecundability from Data on Working Times to First Conception. *Journal of the American Statistical Association* **85**, pp. 283-294.

Holt, J. and R.L. Prentice (1974). Survival analysis in twin studies and matched pair experiments. *Biometrika* **61**: 17-30.

Horiuchi, S. and A.J. Coale (1990). Age patterns of mortality for older women: an analysis using the age-specific rate of mortality change with age. *Mathematical Population Studies* **2**: 245-267.

Hougaard, P. (1984). Life table methods for heterogenous populations: distributions describing the heterogeneity. *Biometrika* **71**: 75-83.

Hougaard, P. (1986a). Survival models for heterogenous populations derived from stable distributions. *Biometrika* **73**: 387-396.

Hougaard, P. (1986b). A class of multivariate failure time distributions. *Biometrika* **73**: 671-678.

Hougaard, P., B. Harvald, and N.V. Holm (1992). Measuring the Similarities Between the Lifetimes of Adult Danish Twins Born Between 1881-1930. *Journal of the American Statistical Association* **87**: 17-24.

Kahn, J. and K.E. Anderson (1992). Intergenerational Patterns of Teenage Fertility. *Demography* **29**: 39-57.

Kalbfleisch, J.D. and R.L. Prentice (1980). *The Statistical Analysis of Failure Time Data.* John Wiley: New York.

Kendall, M.G. (1938). A New Measure of Rank Correlation. *Biometrika* **30**: 81-93.

Keyfitz, N. (1985). *Applied Mathematical Demography. Second Edition.* New York: Springer-Verlag.

Larsen, U.M. and J. Vaupel (1989). Frailty models of Hutterite fertility. Working paper **WP-89-09-1**, Center for Population Analysis and Policy, University of Minnesota.

Manton, K.G., E. Stallard, and J. Vaupel (1981). Methods for Comparing the Mortality Experience of Heterogeneous Population. *Demography* **18**: 389-410.

Manton, K.G., E. Stallard, and J. Vaupel (1986). Alternative models for the heterogeneity of mortality risks among the aged. *Journal of the American Statistical Association* **81**: 635-644.

Mare, R.D. and A. Palloni (1988). Couple models for socioeconomic effects on the mortality of older persons. Paper presented at the 1987 Annual Meeting of the Population Association of America, Chicago. Also in Center for Demography and Ecology **88-7**. University of Wisconsin-Madison.

Montgomery, M. and J. Trussell (1986). Models of marital status and childbearing, pp. 205-270 in *Handbook of Labor Economics*, Volume 1, edited by O. Ashenfelter and R. Layard. Elsevier Science Publishers: New York.

Namboodiri, K. and C.M. Suchindran (1987). *Life Table Techniques and Their Applications.* Academic Press: Orlando.

Oakes, D. (1982). A model for association in bivariate survival data. *Journal of the Royal Statistical Society*, Series B, 44: 414-422.

Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association* 84: 487-493.

Sheps, M.C. and J.M. Menken (1973). *Mathematical Models of Conception and Birth.* University of Chicago Press: Chicago.

Trussell, J. and T. Richards (1985). Correcting for unmeasured heterogeneity in hazard models using the Heckman-Singer procedure, pp. 242-276 in *Sociological Methodology 1985*, edited by N. Tuma. Jossey-Bass: San Francisco.

Trussell, J. and G. Rodríguez (1990). Heterogeneity in demographic research. In *Convergent Issues in Genetics and Demography*, edited by J. Adams, D. Lam, A. Hermalin, and P. Smouse. New York: Oxford University Press: London.

Trussell, J., G. Rodríguez and B. Vaughan. Union Dissolution in Sweden, In *Demographic Applications of Event-History Analysis*, edited by J. Trussell, R. Hankinson and J. Tilton. Clarendon Press: Oxford (in press).

Vaupel, J. (1988). Inherited Frailty and Longevity. *Demography* 25: 277-287.

Vaupel, J. (1990a). Kindred Lifetimes: Frailty models in Population Genetics. In *Convergent Issues in Genetics and Demography*, edited by J.Adams, A. Hermalin, D. Lam and P. Smouse. Oxford University Press: New York.

Vaupel, J. (1990b). Relatives' Risks: Frailty Models of Life History Data. *Theoretical Population Biology* 37: 220-234.

Vaupel, J., K.G. Manton, and E. Stallard (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16: 439-454.

Vaupel, J. and A. Yashin (1985a). The Deviants Dynamics of Death in Heterogeneous Populations. In *Sociological Methodology 1985*, edited by N. Tuma. Jossey-Bass: San Francisco.

Vaupel, J. and A. Yashin (1985b). Heterogeneity's ruses: some surprising effects on selection on population dynamics. *American Statistician* 39: 176-185.

Weiss, K.M. (1990). The Biodemography of Variations in Human Fertility. *Demography* 27: 185-206.

Zenger, E. (1991). Heterogeneity among Families in Neonatal Mortality and Birth Interval in Matlab, Bangladesh. Paper presented at the PAA Meeting, Washington, DC.

NOTES

1. In the case of censored data, Cox and Oakes (1984, p. 161) suggest to modify the Kendall's coefficient of concordance by counting only definite concordances and definite discordances.

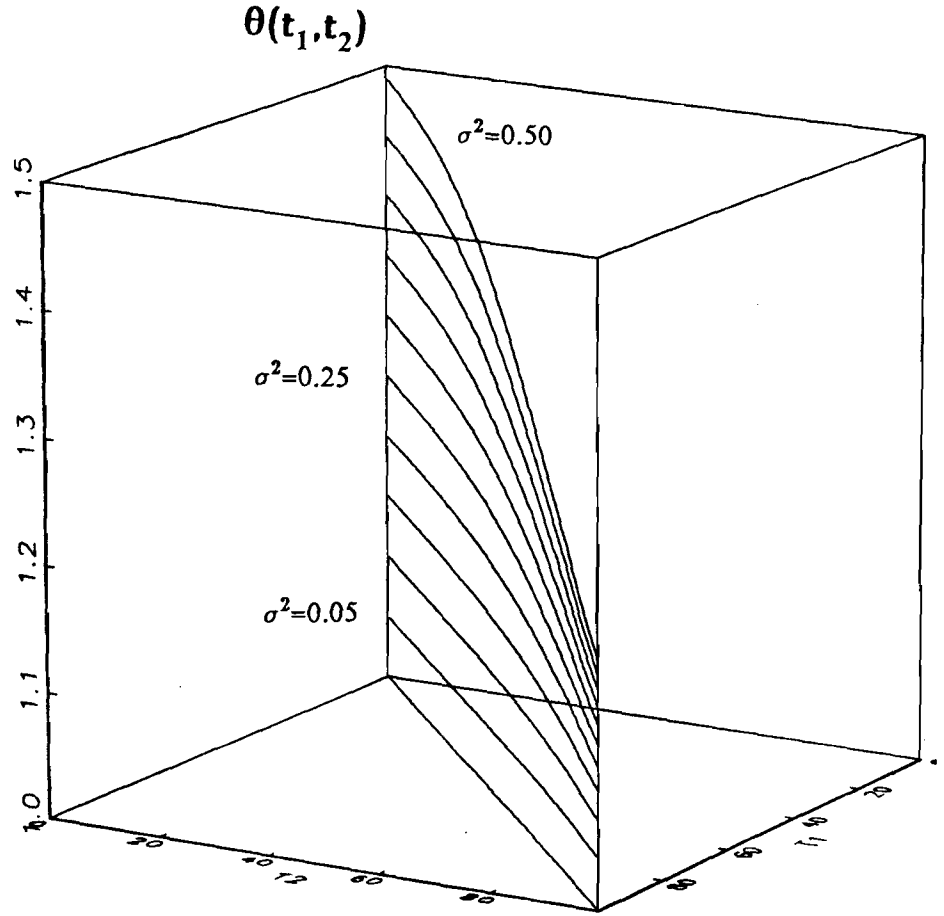
2. Hougaard (1984) showed that the *closure* property applies to all distributions of the exponential family, to which both the gamma and inverse Gaussian distributions belong.

3. Also derived under the assumption that heterogeneity is gamma or inverse Gaussian distributed (Vaupel et al., 1979; Hougaard, 1984, 1986; Aalen, 1988; Vaupel, 1988).

4. For a gamma distribution with parameters α and β, $\Gamma(\alpha,\beta)$, the mean is equal to α/β and the variance is given by α/β^2 (Vaupel et al., 1979). Whereas for the inverse Gaussian with parameters γ and δ, $N^{-}(\gamma,\delta)$, the mean is given by $(\delta/\gamma)^{1/2}$ and the variance by $1/2\ \delta^{1/2}\ \gamma^{-3/2}$ (Hougaard, 1984).

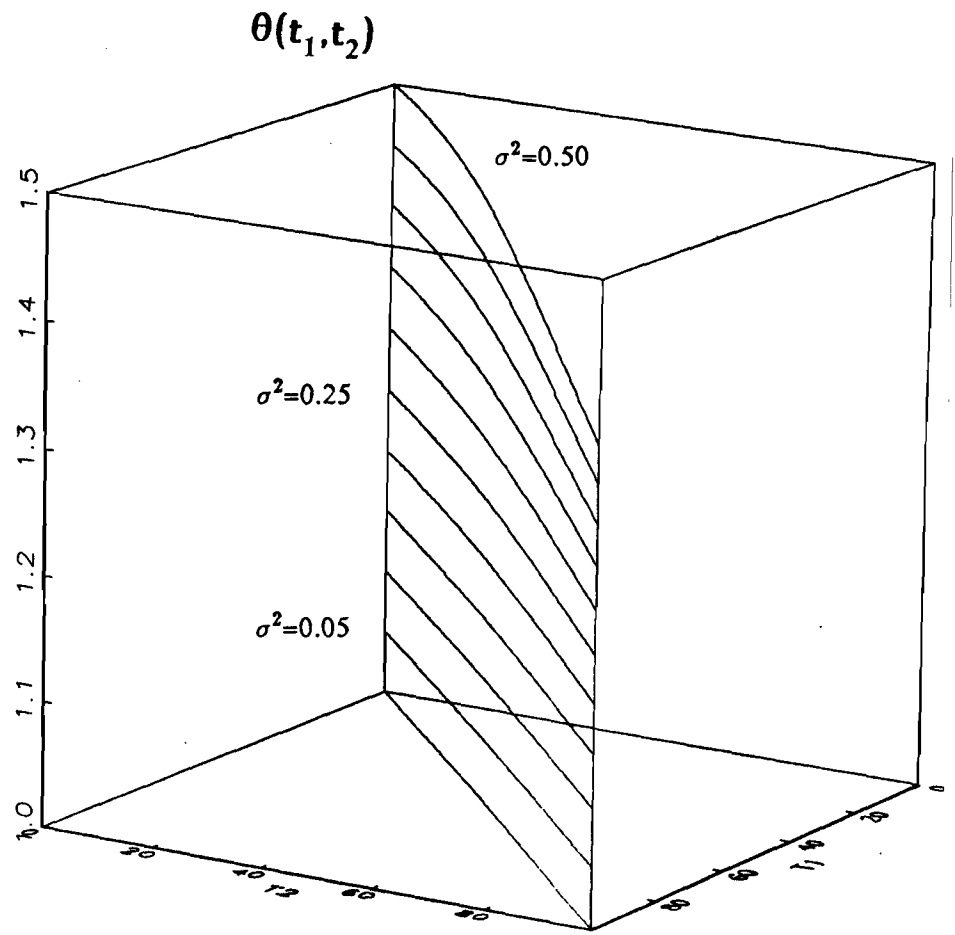
5. Specifically, when $\gamma = 0$ and $\delta = 1/4$, which has mean equal to ∞ (Hougaard, 1984; p. 78).

Figure 1. Association parameter $\theta(t_1,t_2)$ when frailty follows an Inverse Gaussian distribution.



GOMPERTZ

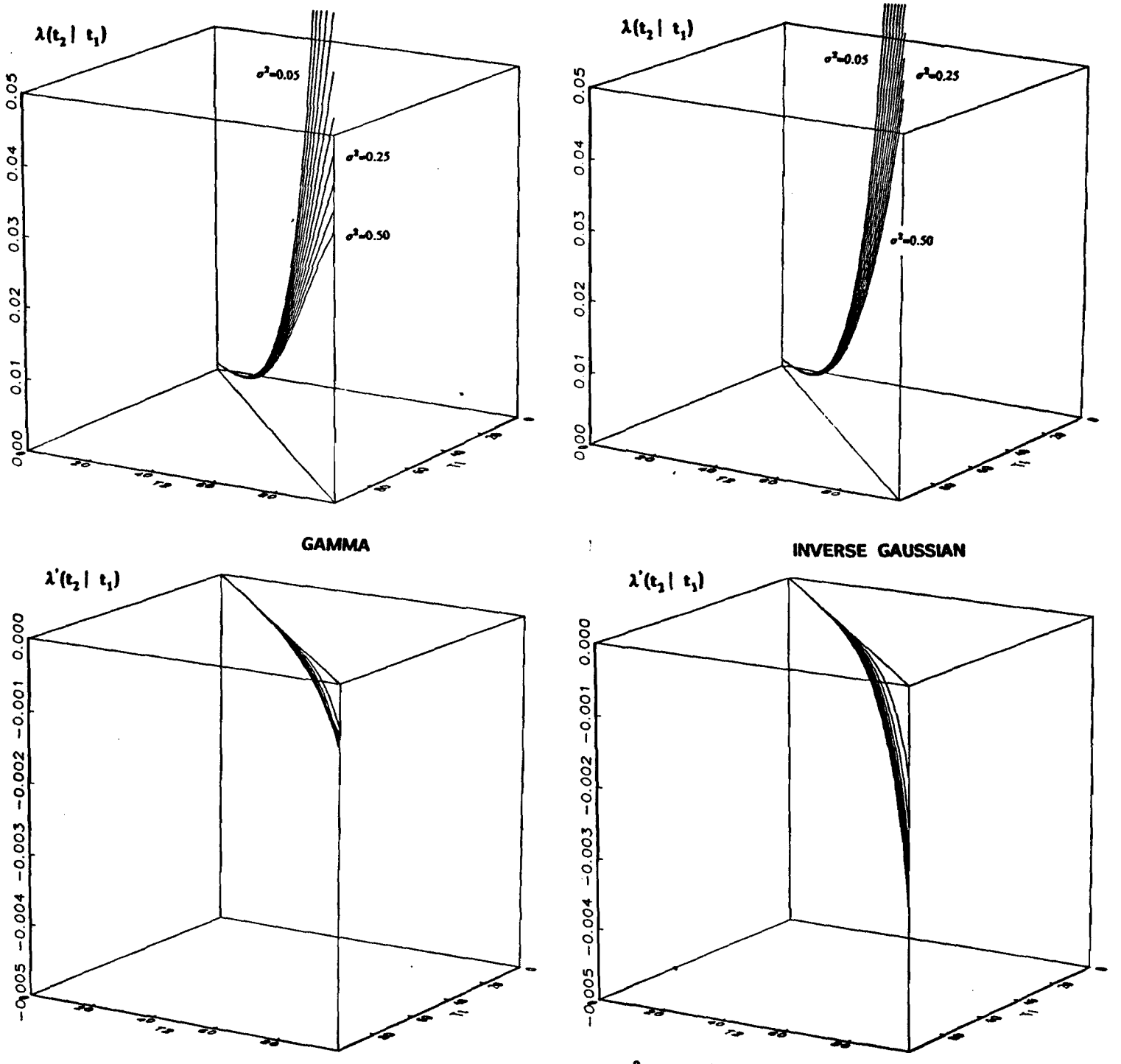
Parameter values: $\alpha = 1x10^{-3}$; $\beta = 0.05$



WEIBULL

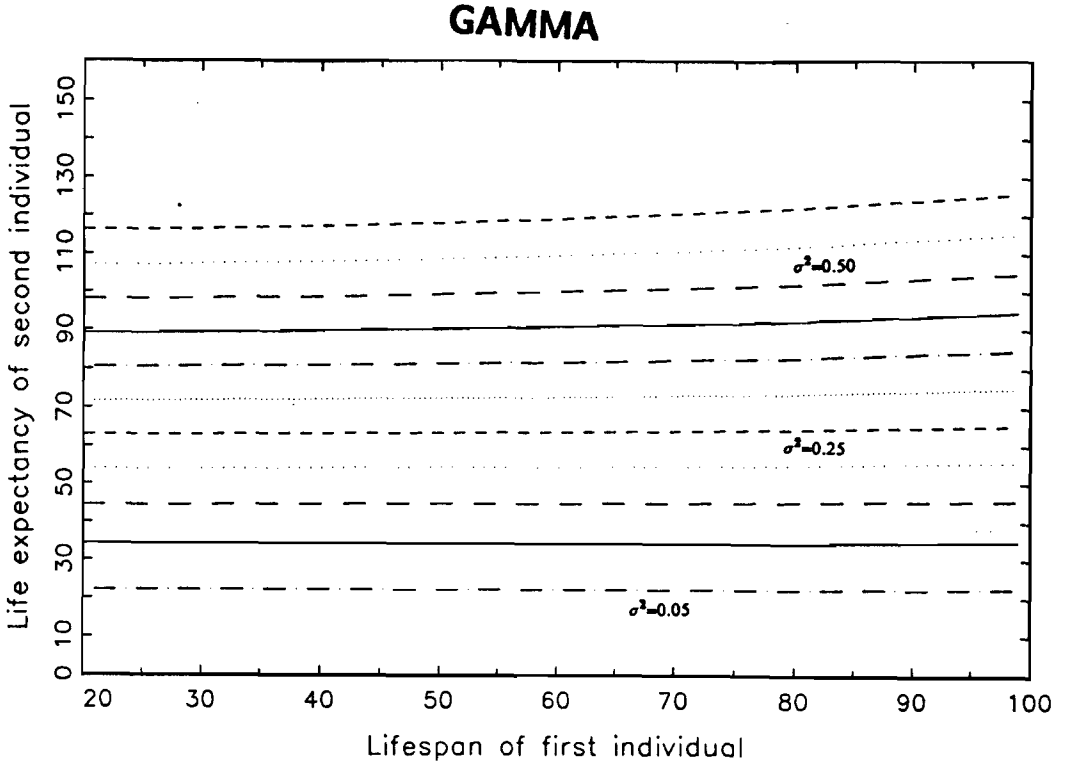
Parameter values: $\alpha = 1.25x10^{-5}$; $\beta = 2.45$

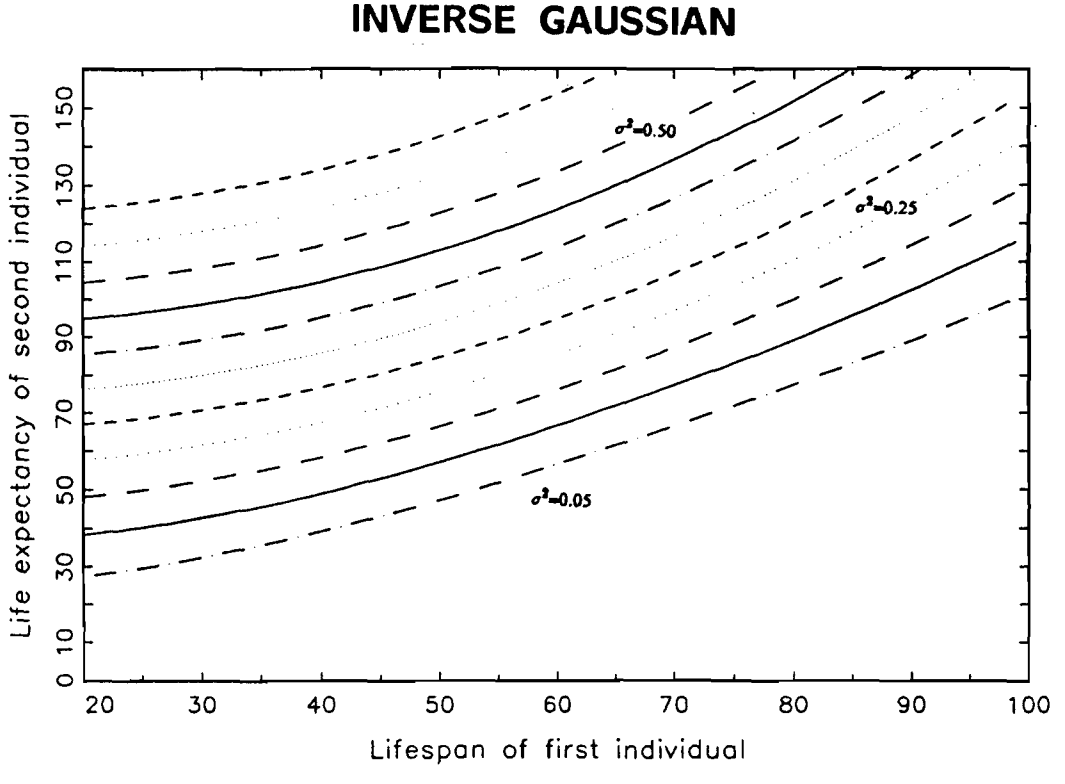
Figure 2. Conditional hazard $\lambda(t_2 \mid t_1)$ and derivative with respect to t_1 when $\lambda_0(t)$ follows a Gompertz distribution.



Parameter values: $\alpha = 1 \times 10^{-3}$; $\beta = 0.05$

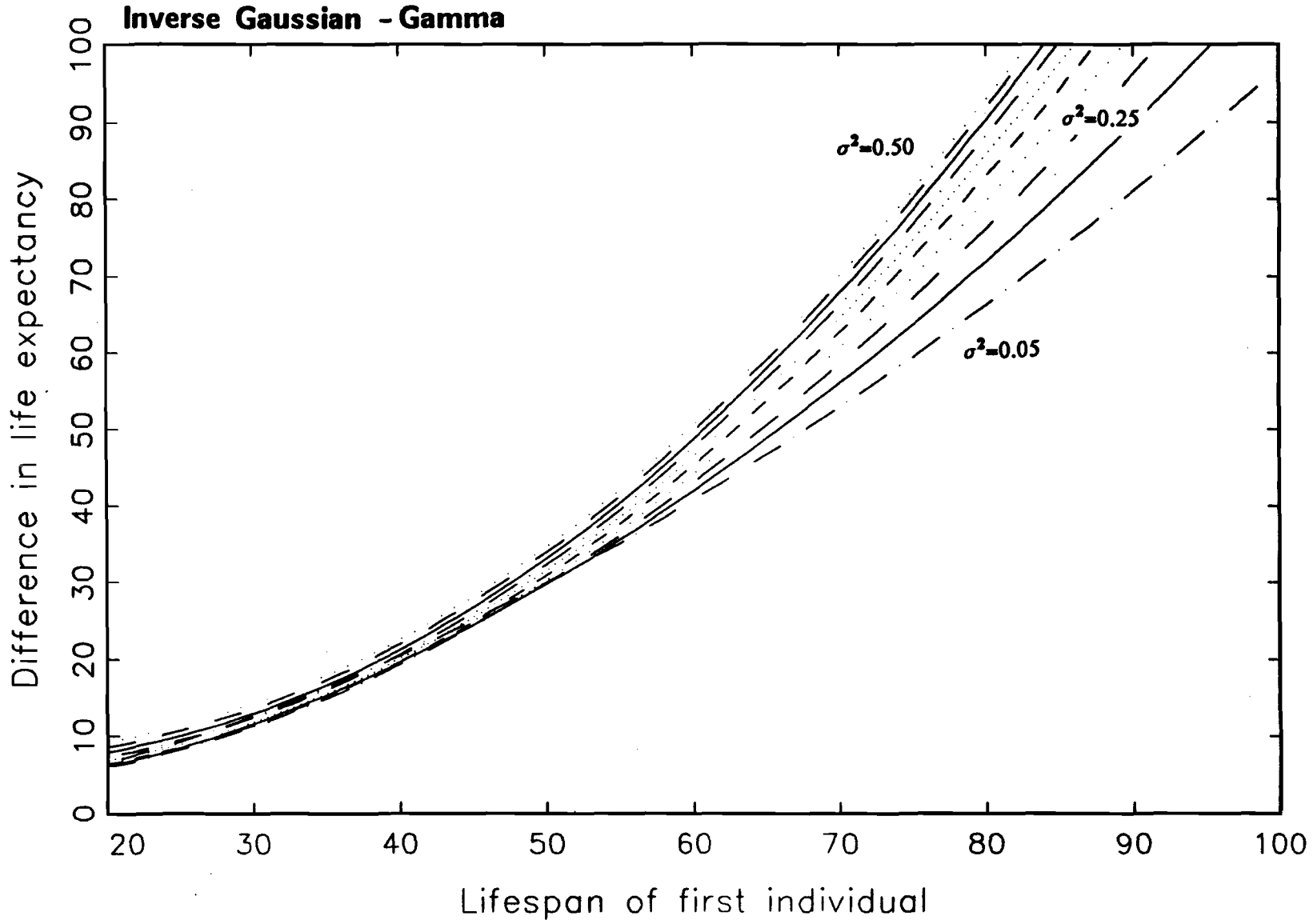
Figure 3. Conditional expectation of life at birth $e_0(T_2|t_1)$ when $\lambda_0(t)$ follows a Weibull distribution and frailty is gamma or inverse Gaussian distributed.





Parameter values: $\alpha = 1.25 \times 10^{-5}$; $\beta = 2.45$

Figure 4. Difference in conditional expectation of life $e_0(T_2|t_1)$ when $\lambda_0(t)$ follows a Weibull distribution and frailty is inverse Gaussian or gamma distributed



Parameter values: $\alpha = 1.25 \times 10^{-5}$; $\beta = 2.45$