Modeling Stochastic Mortality for Joint Lives through Subordinators

Authors

ABSTRACT

Mortality model for joint lives is critical to institutions which offer joint and last survivor financial products. In this paper, we propose a new model in which we use the time-changed Brownian motion with dependent subordinators to describe the mortality of joint lives. We then employ this model to estimate the mortality rate of joint lives in a well-known Canadian insurance dataset, and a dataset collected from National Health Interview Survey (NHIS). Specifically, we first depict an individual's death time as the stopping time when the value of the hazard rate process first reaches one, and then introduce the dependence through dependent subordinators. Compared with existing mortality models, our model better interprets the correlation of death between joint lives, and allows more flexibility of the evolution of the hazard rate process. Empirical results show that our model yields highly accurate estimations of mortality, compared to the baseline non-parametric (Dabrowska) estimation, and the most widely used Copula model. Besides, our model also has high accuracy when modeling the joint mortality for couples in advanced old age.

JEL code: C6, C1, I1, J1

Keywords: Mortality rate; Joint lives; Survival; Stochastic process; Subordinator

1 Introduction

In practice in the current life contingencies markets, most issuers assume independence of mortality rates when pricing and issuing multiple life contingent products such as Joint and Survivor Annuities. This, however, neglects the possible dependency between joint lives, although there is evidence of, for example, dependence in the life trajectory of married couples. Indeed, several empirical studies suggest that the survival times of joint couples are not independent events. Some earlier findings on dependence of joint lives include Jagger and Sutton (1991), which show that an individual's risk of mortality increases after the partner dies. This phenomenon is even more significant when the husband dies first. Additionally, Houggard et al. (1992) study the longevity impact of the death of one partner in a relationship by analyzing the joint survival of Danish twins.[1] Assuming independence of lifetimes may cause inaccurate pricing and introduce moral hazard when selling joint life products to interrelated individuals.

We study a joint life mortality model in this paper without imposing the assumption of independence of survival times of the individuals. Previous work in parametrically modeling dependent joint life distributions includes Brockett (1984) who uses a functional equation method to find the totality permissible classes of bivariate Gompertz distributions, and then introduces additional dependence in the risk of simultaneous accidental death through the Marshal Olkin bivariate exponential distribution so as to obtain a general dependent bivariate Makeham distribution. These bivariate joint life models have Gompertz (or Makeham) marginal distributions for the individual lives' survival distributions while accommodating dependence in their joint distribution. The Gompertz and Makeham distributions fit the univariate (marginal) mortality of humans very well until age 100, as discussed by Gavrilov and Gavrilova (2011).

An alternative empirical approach to model dependence in joint life mortality is to start with marginal distributions, and then introduce dependence using a copula model for the bivariate mortality structure. Articles using this approach include Frees et al. (1996) who use a Frank copula function to model the dependence between the survival times of couples. Carriere (2000) argues that a mixed frailty copula model and generalized Frank copula can describe joint lives well. Spreeuw (2006) derives the conditional law of mortality of a person given the partner's mortality status by using a copula model. Luciano et al. (2008) also describe the dependency through an Archimedean copula.

[1]Besides the above earlier studies, there are some other related papers that find the same conclusion. For example, Frees et al. (1996) prove and model the joint survival through a copula model using a large Canadian insurance company data set. Manor and Eisenbach (2003) examine the effect of spousal death on mortality and how this effect is driven by the duration of bereavement and other demographic factors such as age, sex, education, and household size. Seifter et al. (2014) investigate the impact of bereavement and the mortality of the surviving spouse among Amish couples.

In this paper, we introduce a novel stochastic mortality model to describe the dependency of joint lives. We model the hazard rate process (also known as the force of mortality rate or the failure rate process) of an individual by utilizing a time changed Brownian motion, and introduce the dependence into the joint life process through the use of dependent subordinator processes. This results in a stochastic mortality rate process exhibiting dependence between lives.

The assumption of a stochastic mortality process has been widely used for modeling human mortality. For example, Lorenzo et al. (2006) consider stochastic volatility in the force of mortality (hazard rate) process, where the mortality rate is modeled by a Cox Ingersoll Ross (CIR) stochastic process. Luciano et al. (2008) model the marginal stochastic mortality via a Cox process which allows "jumps" in the arrival of death.

Besides the widely used copula model in modeling dependence between stochastic processes (e.g. Luciano et al., 2008), another good method of introducing dependence into correlated stochastic processes is through the use of what is known as subordinator (also stochastic processes). For example, Semeraro (2006) introduces a multivariate subordinator having gamma marginals, and Luciano and Semeraro (2010) build a theoretical framework for multivariate subordination of Brownian motions. In our paper, we construct the subordinators as a combination of an idiosyncratic component and a common one. The idiosyncratic subordinators express the individual component of an individual's mortality within a couple's joint mortality structure, while the common subordinator expresses the lifestyle connecting dependence between the mortality rate processes of the couple (e.g., eating similar foods, having similar exercise patterns, living in similar environments, driving in automobiles together, etc.). We propose a stochastic mortality model which involves a subordinator in the hazard rate process, and we model the death time as a stopping time of this process.

As mentioned previously, the subordinator of each individual's hazard rate process contains both a common process that is shared with his/her partner, as well as an unique one which is determined by his/her age, gender, genome, etc. This formulation respects the fact that the evolution of the probability of death is determined both separately and connectedly with the partner, because people living together share similar external factors. For example, Nielsen et al. (2018) prove that an individual who lives together with a spouse that has diabetes also experiences a higher risk of having it, even after adjusting for obesity. Besides, because there is a tradeoff of internal and external factors in determining mortality propensity, we also include a weighting parameter that determines how much the exhibited individual mortality relies on these internal and external processes, and we allow this weight to change over time. These can be seen as the change of the level of impact coming from marriage. For example, Johnson et al. (2000) prove that, by getting married, people in younger age group reduce more risk of death than their counterparts in older age group. Therefore, our model not only models the dependence between joint lives, but also allows this dependence to change with time.

Compared with statistical copula models and nonparametric models, our model has three major strengths. First, as mentioned in Frees et al. (1996), copula models focus on modeling and measuring the effects of dependence, but do not attempt to describe the mechanism behind joint mortality. Our model, however, provides a more straightforward way to interpret how joint lives impact each other. Second, we model how the dependence evolves as people age. We model the fact that, as people age, their internal characteristics could play a more important role in determining their probability of death, compared to when they are in younger ages. This coincides with previous observations and analyses (e.g., Austad, 2006). We find that this phenomenon is more significant among males. Specifically, as the male ages, his probability of death becomes less dependent on external factors, while this trend is not similarly significant among females. Third, our model does not constrain the hazard process to be an increasing function, and allows it to fluctuate within certain ranges. This feature reflects the reality that the evolution of individuals' probability of death can be impacted by various life events, and can both speed up and slow down after negative or positive shocks related to changes of lifestyles, habits, or life events (job loss or promotion; illness; recovery from severe diseases; etc.).

The rest of this paper is organized as follows. Section 2 introduces the detailed setting of our model and its basic properties. Section 3 describes the dataset in our empirical analysis. In Section 4 and 5, we apply our model to datasets and test its performance. We conclude our findings and discuss further research in Section 6.

2 Modeling Dependent Mortality through Time Change Brownian Motions with Dependent Subordinators

2.1 Modeling Individual Death Time through Time Change Brownian Motion

Assume a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ which supports two independent stochastic processes B_t and G_t. Here, B_t is assumed to be a standard Brownian motion, and G_t is assumed to start from zero and be an increasing Càdlàg process (i.e., has sample paths that are right continuous with left limits). Construct a Brownian motion with drift $X_t = X_0 + \sigma B_t + \beta\sigma^2 t$, assuming $\sigma > 0$. Then, the time-changed process can be built as X_{G_t}. Intuitively, the "time" scale t for X is moving according to the process G_t. Here, G_t is called a *subordinator* of X.

We model the log of the hazard rate process h_t as the time changed Brownian motion $log(h_t) = X_{G_t}$. Following Hurd (2009), we define the death time of an individual as the stopping time $\tau = \inf\{t | G_t \geq t^*\}$. Then

$$P(\tau > t \mid \mathscr{F}_t) = e^{\int_0^t -e^{X_{G_s}} ds} \tag{1}$$

Following Hurd (2009), we define the death time of an individual as the stopping time $\tau = inf\{\ t \mid G_t \geq t^*\ \}$, with $t^* = inf\{\ X_t \geq 0\ \}$, i.e., the time when the hazard becomes one, or death occurs. Accordingly, the cumulative distribution function the death time can be calculated (Hurd 2009) as

$$P(\tau \leq t, x) = \int_0^{\text{inf}} [N(\frac{-x - \beta\sigma^2 y}{\sigma\sqrt{y}}) + e^{-2\beta x} N(\frac{-x + \beta\sigma^2 y}{\sigma\sqrt{y}})]\rho_t(y) dy \tag{2}$$

with ρ_t being the density of G_t.

As is shown by Hurd (2009), the death time, or the stopping time, defined in the above way, exhibits the *reduced form property*, a concept borrowed from financial mathematics. By Jarrow and Protter (2004), the reduced form property proven by Hurd for this process allows the stopping time to be inaccessible, and also allows for the observer to have only incomplete knowledge of the evolution of the time changed Brownian motion. These properties are indeed desirable in our mortality modeling, since (a) death could happen unexpectedly, and (b) with the biological knowledge that humans currently have, we have very limited knowledge of how the individual hazard rate process of a person is precisely determined.

We model the log of the hazard rate process h_t as the time changed Brownian motion $\log(h_t) = X_{G_t}$, modeling the death time of an individual as the stopping time for a hazard function that corresponds to a time changed Brownian motion, i.e., for $\log(h_t) = X_{G_t}$. This model provides a new way to understand the increasing probability of death as an individual is aging. The base process with the regular time t describes a generic pattern of how the probability of death would evolve over time for people with certain characteristics such as genetics and place of residence, and the stochastic process G_t assigns a unique and unobservable "internal clock" to each individual that speeds up or slows down this initial clock t. This internal clock mechanism G_t could be either continuous or discontinuous with jumps, but should not flow backward at any time. Therefore, in addition to the regular time, the internal clock impacts how an individual ages based on his characteristics such as age, health condition, etc. This clock could also depend on some common factors such as air quality and macroeconomic condition.

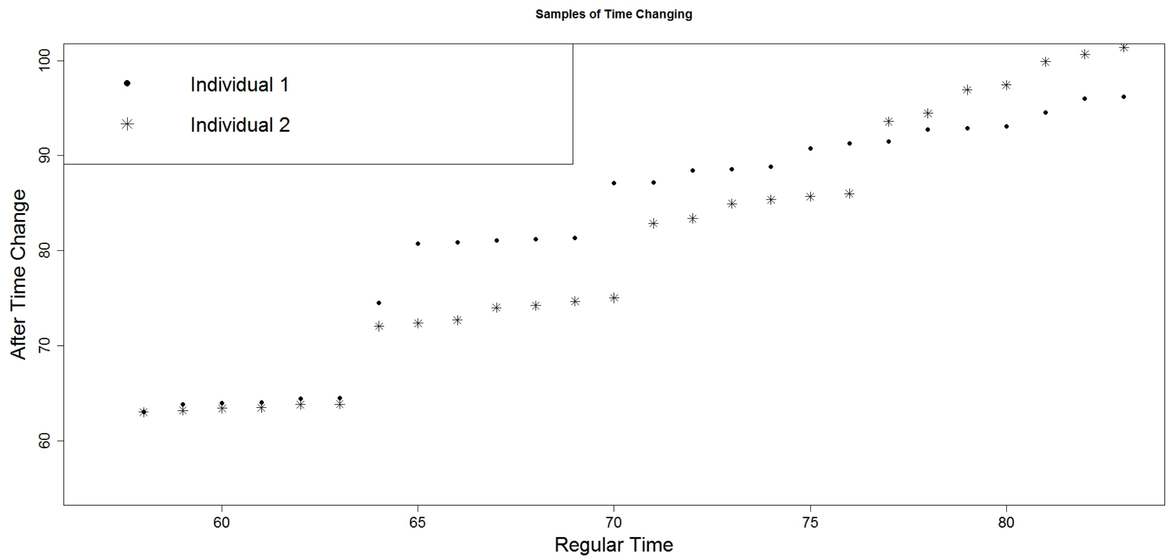
Figure 1 shows two samples of moving paths of subordinator G_t, with G_t assumed to be Inverse Gaussian Process with the same parameters. We can observe from the graph that individual 2 has a relatively fast increasing speed, with big jumps at age 64, 71 and 77, while individual 1 yields a relatively slow increasing speed after 65 years old. Individual 2's jumps could describe the impacts caused by unexpected life events or change of health conditions at age 64, 71 and 77, and a more smooth path of evolution is exhibited in individual 1's probability of death evolution.

2.2 modeling Mortality Dependency through Dependent Subordinators

In our model, we model the dependence of joint lives by allowing the two individuals of the couple have correlated subordinators in their log-hazard functions.

Formally, consider a couple in which the male partner is denoted as M, and the female partner is denoted as F. As per the above discussion, we assume that both of them have a base log hazard

Figure 1. Samples of Time Changing Path
(G_t's being Inverse Gaussian Processes)



rate (force of mortality) process X^M and X^F, with

$$X_t^M = X_0^M + \sigma^M B_t^M + \beta^M {\sigma^M}^2 t, \tag{3}$$

and

$$X_t^F = X_0^F + \sigma^F B_t^F + \beta^F {\sigma^F}^2 t \tag{4}$$

respectively. Here, B_t^M (and similarly, B_t^F) is a standard Brownian Motion, where X_0^M (X_0^F) is the starting value of the hazard rate process for X_t^M (X_t^F), and σ^M (σ^F) and β^M (β^F) are two parameters that determine their volatility and drift. The subordinators are noted as G_t^M and G_t^F, with X^M, X^F, G_t^M and G_t^F being mutually independent stochastic processes. Assume that the subordinators G_t^M and G_t^F are constructed dependently, in the following way.

$$G_t^M = \alpha^M(t) G_t + (1-\alpha^M(t)) G_t^{M0} \quad ,0 \le \alpha^M(t) \le 1 \tag{5}$$

$$G_t^F = \alpha^F(t) G_t + (1-\alpha^F(t)) G_t^{F0} \quad ,0 \le \alpha^F(t) \le 1 \tag{6}$$

Here, G_t, G_t^{M0} and G_t^{F0} are three increasing Càdlàg processes which are mutually independent, allowing jumps, and $\alpha^M(t)$ and $\alpha^F(t)$ are real valued deterministic function of t, ranging between 0 and 1. Note that both G_t^M and G_t^F involve two processes — a common process G_t which causes the dependence between the male and the female, and individual processes G_t^{M0} and G_t^{F0} which reflect the unique characteristics of each person as they apply to their trajectory along the time axis. Hence, the time change of each person is determined by these two processes and their weights (relative importance of each idiosyncratic and common effect). Clearly, the higher the $\alpha^M(t)$ (or $\alpha^F(t)$), the heavier the G_t weights at time t; thus the higher the impact the common process has on the time changing process, and hence the more synchronized the male and the female. When $\alpha^M(t)$ (or $\alpha^F(t)$) takes the minimum value of 0, both the subordinators of the male and the female individual, G_t^M and G_t^F, and the time changing process X^M and X^F are mutually independent. Similarly, a value of 1 for $\alpha^M(t)$ (or $\alpha^F(t)$) implies that the subordinator of the male (female) follows the common temporal factor entirely and the lives are perfectly correlated. In our analysis we allow $\alpha^M(t)$ (or $\alpha^F(t)$) to change over time, which reflects the possibility of changing dependency over time.

Under this model, the conditional cumulative density function of the individual i dies before time t, given that individual j dies after time t, can be expressed as follows (Hurd, 2009).

$$P(\tau^i \leq t; \tau^j \geq t) = \int_0 (1 - F^i(x^i, y)) F^j(x^j, y) \rho_t(y) dy\,, \quad i \neq j\,,\ i, j \in \{M, F\} \tag{7}$$

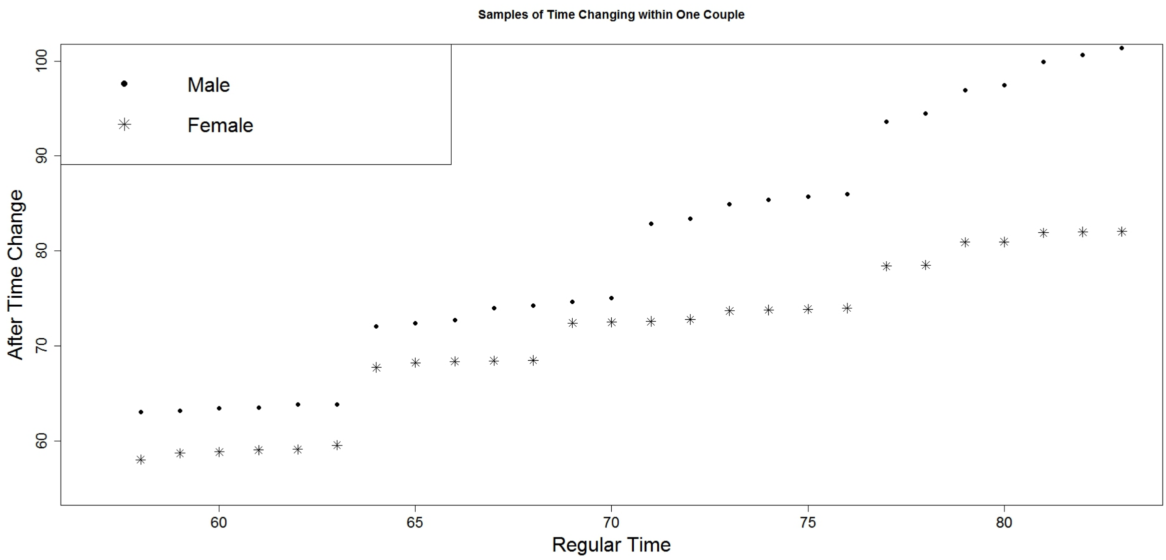
where

$$F^m(x,y) = \frac{e^{-\beta_m x}}{\pi} \int_{-} \frac{zsin(zx)}{z^2 + \beta_m^2} e^{-\alpha_m \sigma_m^2 (z^2+\beta_m^2) y/2 - \Phi^{G^{m0}}((1-\alpha_m)\sigma_m^2(z^2+\beta_m^2)/2, y)} dz\,, \quad m = M, F \tag{8}$$

.

Our model describes how an individual's probability of death evolves. On one hand, an individual's probability of death can be impacted by external factors (air quality, macroeconomic condition, etc.). Individuals living together share the same external characteristics, hence it is reasonable to assume that a factor that causes a jump in one individual's probability of death

Figure 2. Samples of Subordinators of One Couple



process also affects on the other individual (resulting in dependent life trajectories). The shared process G_t captures such common factors. On the other hand, the probability of death also depend heavily on individuals' idiosyncratic characteristics such as genome and age, and these idiosyncratic factors are not associated with marital status. The processes G_t^{M0} and G_t^{F0} capture these individual-level characteristics.

We also allow the weight $\alpha^M(t)$ (or $\alpha^F(t)$) to be a time dependent function in our model. Previous mortality models, such as Frees et al. (1996) and Luciano et al. (2008), use the copula function to model the association between couples' mortality, which is based on an assumption that the correlation between the male and the female should be consistent through their entire lives. However, it has been shown by previous studies (e.g., Austad, 2006) that the dependent level between couples could change as people age. Our model respects the reality that the internal characteristics and external factors could have different impacts throughout life course.

Figure 2 shows a sample of the subordinators within one couple. We can see from the figure that on one hand, the male and the female yield different time changing path, while on the other hand, they share a very similar pattern of increasing and jumping. Both the male and the female have big jumps at age 64 and 76, which are caused by jumps of the common process G_t. Additionally, the female has a unique jump at age 69 which is caused by her unique process G_t^{F0}, while a similar unique jump happens to the male at age 71.

3 Data

3.1 Data Description

We apply the methodology to a well-known Canadian insurance data set.[2]. This dataset contains $14,947$ insurance contracts each of which represents a pair of couple observed between December 29^{th}, 1988, to December 31^{st}, 1993. Each observation includes the birth year of both the male and female members, the time when they begin their joint life contract, if any of the members is observed to have died during the observation period, and if yes, the year of death. We select samples with the male and female both born between 1910 and 1925 and whose age differences are not greater than five years, as we take (a) changing trends in mortality among different generations, and (b) potential impacts of age differences into consideration in the data analysis. This reduces the set for analysis to a subset of $7,270$ pairs of observations. Due to the short observing window and the limit of the size of the observed population, we do not consider cohort effects, and we also do not include same-sex marriage. However, our model can be easily applied to all types of joint-life contexts in a similar manner. Table 3 in the Appendix shows a summary of the birth year of our data set.

3.2 Dependence between Joint Lives and Kaplan-Meier Estimation of Marginal Survival Probability

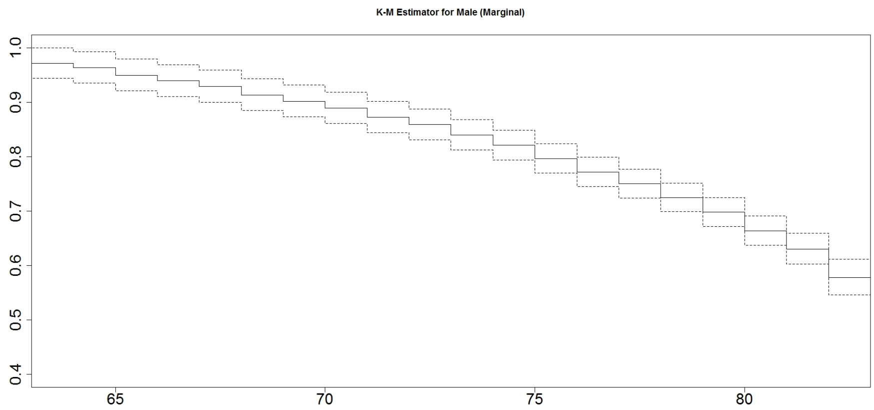
Before applying any bivariate estimation model to our data set, we need to examine the dependence between the joint lives. We calculate Kendall's tau coefficient to measure the association between the male and female' life length. Kendall's tau coefficient is expected to be zero if the male and female are independent. However, in our data set, Kendall's tau coefficient is 0.58, with a p-value of 2.22×10^{-16}. This indicates a strong association between the joint lives. This assessment agrees with that of Frees et al. (1996) that also show the correlation of the mortality rate processes between joint lives.

[2]We acknowledge the Society of Actuaries who originally furnished the data to Edward (Jed) Frees and Emiliano Valdez. Frees and Valdez provided the data to us, and we gratefully acknowledge this as well

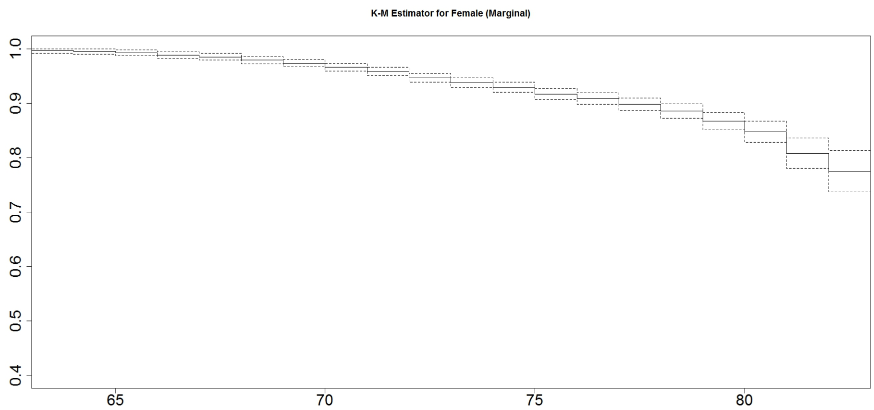
We also apply the Kaplan-Meier nonparametric maximum likelihood estimation method to estimate the marginal survival probability for the male and female, to show the trend of the evolution of the probability of death (which are shown in Figure 3 below, and Table 4 in the Appendix). The marginal survival probability we estimate is the conditional probability $P(\tau^m > t \mid \tau^m \geq 63)$, $m = \{M\ (Male), F\ (Female)\}$, as the minimum age of our data set is 63. It is significant that the female experience a lower probability of death in the period that we examine. We have rounded down any age in between integers (i.e., used the age last birthday value) in our analysis.

Figure 3. Kaplan-Meier Estimation of Marginal Survival Probability

(a) Male



(b) Female



4 Estimating and Evaluating the Model

We apply both a classical nonparametric method, Dabrowska (1988), and our time changed hazard rate process model to estimate the conditional probability $P(\tau^M > t_1, \tau^F > t_2 \mid \tau^M \geq 63, \tau^F \geq 63)$. In this section, we will show that our method obtains good results.

4.1 Nonparametric Estimation of the Bivariate Survival Function

Dabrowska (1988) proposes a nonparametric method of estimating the bivariate survival function $F(t_1,t_2) = P(\tau^M > t_1, \tau^F > t_2)$. We employ this estimation as the baseline estimation of the probability. Here, τ^M is the death age of the male individual, and τ^F is the death age of the female individual. We estimate the conditional probability because the starting age of our observation is 63 for both genders. This estimation takes the form of

$$\begin{aligned}\hat{P}(\tau^1 > t_1, \tau^2 > t_2 \mid \tau^M \geq 63, \tau^F \geq 63) = \hat{P}(\tau^1 > t_1 \mid \tau^M \geq 63) \times \hat{P}(\tau^2 > t_2 \mid \tau^F \geq 63) \\ \times M(t_1,t_2 \mid \tau^M \geq 63, \tau^F \geq 63)\ ,\end{aligned} \tag{9}$$

with $\hat{P}(\tau^1 > t_1 \mid \tau^M \geq 63)$ and $\hat{P}(\tau^2 > t_2 \mid \tau^F \geq 63)$ being the univariate Kaplan-Meier estimators of each individual, and

$$M(t_1,t_2 \mid \tau^M \geq 63, \tau^F \geq 63) \triangleq \prod_{0<s_1\leq t_1, 0<s_2\leq t_2} (1 - \hat{L}(\Delta s_1, \Delta s_2 \mid \tau^M \geq 63, \tau^F \geq 63))$$

being a multiplier for each pair of (t_1,t_2). Here, Δs corresponds to the time period between $s-1$ and s. If the probabilities of death of the male and female are independent, then $M(t_1,t_2 \mid \tau^M \geq 63, \tau^F \geq 63)$ should be 1 at each pair of age. However, we find multipliers at all ages are greater than 1, which indicates that the evolution of the probability of death of male and female are indeed dependent.

4.2 Estimating Bivariate Survival Function through Time-Changed Hazard Rate Process

We apply our model to the $7,270$ observations in our data set. We first estimate the initial value of X_0^m, $m = \{M, f\}$, as the log of the hazard rate estimated by the Kaplan-Meier estimator at age

63, i.e.,

$$X_0^M = log(-(0.968-1)/1) = -3.450\ , \qquad X_0^F = log(-(0.998-1)/1) = -6.125\ .$$

In our model, we select our subordinators to be the IG (Inverse Gaussian) process since it has a good goodness-to-fit, and is easy to simulate. It is also widely used in other mortality models such as Mitchell et al. (2013) and Wang et al. (2011). The hazard rate process h_t, i.e., the time changed Brownian motion X_{G_t}, thus follows the NIG (Normal Inverse Gaussian) process. The NIG process requires the subordinators take the form of

$$\begin{aligned} G_t &= IG(\ t\ ,\ b\ ) \\ G_t^{M0} &= IG(\ \frac{1-\sqrt{\alpha^M}}{\sqrt{1-\alpha^M}} t\ ,\ \frac{b \times \sqrt{1-\alpha^M}}{\sqrt{\alpha^M}}\ ) \\ G_t^{F0} &= IG(\ \frac{1-\sqrt{\alpha^F}}{\sqrt{1-\alpha^F}} t\ ,\ \frac{b \times \sqrt{1-\alpha^F}}{\sqrt{\alpha^F}}\ )\ , \end{aligned} \tag{10}$$

and the Brownian motion takes the form of

$$\begin{aligned} \beta^M &= \sqrt{\alpha^{M2} - b^2/(\alpha^M \sigma^{M2})} \\ \beta^F &= \sqrt{\alpha^{F2} - b^2/(\alpha^F \sigma^{F2})}\ , \end{aligned} \tag{11}$$

as discussed in Luciano and Semeraro (2010).

Therefore, we have the set of parameters $\{\alpha^M, \alpha^F, b, \sigma^M, \sigma^F\}$ to be estimated. We simulate $10,000$ pairs of (X_t^M, G_t^F) and (X_t^F, G_t^F), and calculate the stopping time defined in Section 3.1. We further calculate the joint survival probabilities $\hat{P}(\tau^1 > t_1, \tau_2 > t_2 \mid \tau^M \geq 63, \tau^F \geq 63)$ for $63 \leq t_1 \leq 83$ and $63 \leq t_2 \leq 83$.

We fit the parameters so as to yield the minimal sum of the L^1 distance between the estimated probability and the nonparametric estimation by the Dabrowska's estimator. In order to show the possible change of the level of association and the level of impact from internal and external factors, we first assume α^M and α^F are constants and calculate the mean, median and standard deviation of the L^1 distance, and then set them to be functions of regular time, calculating the same statistical measures. We find that results appear to be better in terms of fitness when allowing $\alpha^M(t)$ and $\alpha^F(t)$ to be functions of time. To be more specific, we let $\alpha^M(t)$ and $\alpha^F(t)$ be in the simple form of $\alpha^M(t) = \alpha^M \times {C_M}^t$, and $\alpha^F(t) = \alpha^F \times {C_F}^t$.

Table 1-a shows the estimated value of the parameters when α^M and α^F are fixed constants, and Table 1-b shows the estimated value of the parameters when α^M and α^F take the form of $\alpha^M(t) = \alpha^M \times {C_M}^t$ and $\alpha^F(t) = \alpha^F \times {C_F}^t$. We can see from the table that both males and females show decreasing degrees of dependence on the external factors, and males generally have more significant decreasing trends than females.

Table 1. Estimated Value of Parameters

(a) Fixed α^M and α^F

	α^M	α^F	b	σ^M	σ^F
Estimated	0.673	0.800	0.193	0.660	0.95

(b) $\alpha^M(t)$ and $\alpha^F(t)$ as functions of time

	α^M	α^F	b	σ^M	σ^F	C_M	C_F
Estimated	0.62	0.62	0.0108	1.55	1.16	0.955	0.955

Table 2 shows the mean, median and the standard deviation of the L^1 distance between our estimation and the Dabrowska's estimator for both estimations. These values are the averages of 10,000 repeated simulations.

We can see from Table 2 that our model yields highly accurate estimations. Specifically, when assuming $\alpha^M(t)$ and $\alpha^F(t)$ as functions of time, we find that the mean and median of the L^1 distance between our estimation and Dabrowska's estimation are only 0.008 and 0.0064, respectively, which are less than the most widely used Copula model. These results show that our model is reliable, and has useful applications in the risk and insurance practice.

5 Model Capability for More Advanced Age Group

In Section 4, we have shown that our model performs well on people in moderate old age. However, as pointed out by previous studies, e.g. Gavrilov et al. (2017), human mortality may

Table 2. L^1 Distance

(a) Fixed α^M and α^F

	Mean of L^1 Distance	Median of L^1 Distance	Std. of L^1 Distance
Estimated	0.015	0.012	0.019

(b) $\alpha^M(t)$ and $\alpha^F(t)$ as functions of time

	Mean of L^1 Distance	Median of L^1 Distance	Std. of L^1 Distance
Estimated	0.0080	0.0064	0.0074

(c) L^1 Distance (Frees et al. (1996))

	Mean of L^1 Distance	Median of L^1 Distance	Std. of L^1 Distance
Estimated	0.0100	0.0090	0.0071

show different patterns when people are in advanced age and the Gompertz-Makeham law may not work there, although some other studies argue that human mortality follow very similar pattern until extreme old age. In this section, we test our model on couples in a more advanced age group, and show that our model still perform well. More specifically, our model outperform the Copula model significantly.

5.1 Data Description

We collect data from the National Health Interview Survey (NHIS), which is a national survey with information on the health, and health behaviors of the U.S. population, conducted by the U.S. Census Bureau. In this survey, each year, samples of households across the entire US. are randomly selected to be national representative samples. U.S. Census Bureau links each household members to death certificate data (by CDC) where couples' death years can be obtained. However, death records used in this dataset only covers people who were reported to be died during 1986 to 2011. Therefore, people in the dataset were either observed died before 2011, or right censored. In our study, we select couples born between 1901 to 1910, in consideration of the generation effect and the limited observation. It narrows down the dataset to a total of $1,979$ pairs of observations. We do not apply further constraints to our dataset because of the limit of the sample size.

5.2 Estimating and Evaluating the Model

We apply our model to the $1,979$ pairs of observations, following the same estimation methods as in Section 4. Conditional probabilities estimated here is the probability $P(\tau^M > t_1, \tau^F > t_2 \mid \tau^M \geq 75, \tau^F \geq 75)$. We can see from Table 4 that our model still obtain high accuracy and outperforms the Copula model with Gompertz Marginal significantly.

Table 3. Estimated Value of Parameters

	α^M	α^F	b	σ^M	σ^F	C_M	C_F
Estimated	0.550	0.550	0.0090	3.500	2.550	0.800	0.800

Table 4. L^1 Distance

(a) $\alpha^M(t)$ and $\alpha^F(t)$ as functions of time

	Mean of L^1 Distance	Median of L^1 Distance	Std. of L^1 Distance
Estimated	0.0100	0.0077	0.0090

(b) L^1 Distance (Gompertz & Copula)

	Mean of L^1 Distance	Median of L^1 Distance	Std. of L^1 Distance
Estimated	0.0167	0.0139	0.0118

6 Discussion and Conclusion

In this paper, we propose a new mortality model for dependent lives which uses a time changed Brownian motion to describe the hazard rate process of an individual, and model the death time of the individual as the stopping time when the value of the subordinator first becomes no less than the regular time when the base process first reaches zero. Based on this, we model the dependence of the probability of death between joint lives through dependent subordinators. The subordinator of each individual contains both a common process that is shared with his/her partner, as well as a unique factor process which is determined by his/her own age, gender, genome, etc.

In general, our modeling strategies provide a flexible framework for describing the evolution of the probability of death for joint lives, compared with prior research that uses either copula functions (e.g., Frees et al, 1996) or non-parametric estimators (e.g., Dabrowska, 1988). Specifically, our model allows the association level between joint lives to be changing over time, which captures the fact that individuals' internal characteristics could play an increasingly more important role in determining the probability of death as they age (e.g., Austad, 2006). In addition, our model allows the non-monotonicity of the hazard rate process, which allows for the possibility that an individuals' health condition can be improved and the evolution of the probability of death can be slowed down.

In the empirical analysis, we exploit a Canadian insurance data set (Luciano et al, 2008; Frees et al, 1996), and first run the baseline estimation using a non-parametric method (Dabrowska, 1988). We then apply our model to this data set. Results show that our model yields estimates that are close to the baseline estimates. Specifically, the mean and median L^1 distance between two estimates is only 0.008 and 0.0064, respectively. Our model implementation is also straightforward, and is computationally easy. We also examine our model performance on more advance age groups, and it also shows high estimation accuracy.

Our paper has potential of applications in the risk and insurance practice. First, our method can be easily applied to data sets on other types of relationships (e.g., the pair of the owner and pet) in insurance practice, even including non-health related distributional relationships (e.g. auto and house). This further helps insurance companies to set appropriate pricing strategies. Second, as our model points out trends in joint lives' mortality, it can be used to guide household financial management and retirement planning.

The idea of this paper leads to several future research paths. First, a natural continuation of our current modeling strategies is to take the cohort effect and the age difference between couples into account, and explore how these factors change modeling results. Second, we can further extend our model to describe dependence among multiple household members. A more complicated extension of our model is to determine $\alpha^m(t)$ in a non-parametric manner, which leads to a semi-parametric structure in our current model. In the empirical analysis, we need to use a more rich data set on joint lives to conduct the cohort study and estimate effects brought by the age difference between joint lives.

Appendix

Table 5. Summary of Birth Years (Female by Male)

		Year of Birth (F)															
		1910	1911	1912	1913	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923	1924	1925
Year of Birth (M)	1910	13	11	14	9	20	15	8	0	0	0	0	0	0	0	0	0
	1911	12	19	25	20	26	24	14	11	0	0	0	0	0	0	0	0
	1912	4	16	18	34	26	40	23	23	10	0	0	0	0	0	0	0
	1913	10	12	23	27	36	56	37	56	26	13	0	0	0	0	0	0
	1914	1	15	6	23	45	48	52	51	59	56	22	0	0	0	0	0
	1915	1	6	19	19	40	66	84	60	64	67	74	27	0	0	0	0
	1916	0	2	14	10	44	47	71	51	76	74	68	56	42	0	0	0
	1917	0	0	0	14	15	25	44	72	76	86	83	76	47	30	0	0
	1918	0	0	0	1	10	16	39	57	68	77	112	104	71	61	38	0
	1919	0	0	0	0	5	18	28	31	36	64	84	116	76	95	71	26
	1920	0	0	0	0	0	6	17	29	51	85	118	136	105	96	101	83
	1921	0	0	0	0	0	0	10	15	26	35	83	114	128	89	119	101
	1922	0	0	0	0	0	0	0	7	15	28	55	78	110	129	87	99
	1923	0	0	0	0	0	0	0	0	7	14	34	50	49	98	105	107
	1924	0	0	0	0	0	0	0	0	0	11	23	36	27	60	91	96
	1925	0	0	0	0	0	0	0	0	0	0	9	24	22	39	48	102

Table 6. Kaplan-Meier Estimation of Marginal Survival Probability

	Male
Age	Survival Probability
63	0.968
64	0.960
65	0.946
66	0.936
67	0.926
68	0.910
69	0.898
70	0.886
71	0.870
72	0.856
73	0.837
74	0.817
75	0.792
76	0.766
77	0.742
78	0.718
79	0.690
80	0.650
81	0.618
82	0.558
83	0.492

	Female
Age	Survival Probability
63	0.998
64	0.996
65	0.994
66	0.989
67	0.986
68	0.980
69	0.975
70	0.967
71	0.959
72	0.946
73	0.938
74	0.930
75	0.917
76	0.908
77	0.898
78	0.884
79	0.864
80	0.846
81	0.806
82	0.791
83	0.767

Table 7. Estimated Probability of Death (Hazard Rate Process as Time Changed Brownian Motion)

		Year of Birth (F)																				
		63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83
	63	1.000	0.999	0.997	0.995	0.993	0.989	0.985	0.980	0.974	0.966	0.957	0.946	0.934	0.920	0.905	0.888	0.869	0.848	0.826	0.803	0.778
	64	0.996	0.996	0.994	0.992	0.990	0.986	0.982	0.977	0.971	0.963	0.954	0.944	0.932	0.918	0.903	0.886	0.867	0.847	0.825	0.801	0.776
	65	0.990	0.990	0.989	0.988	0.985	0.982	0.978	0.973	0.967	0.959	0.951	0.940	0.929	0.915	0.900	0.883	0.864	0.844	0.822	0.799	0.774
	66	0.983	0.983	0.982	0.981	0.979	0.976	0.972	0.967	0.961	0.954	0.945	0.935	0.924	0.910	0.895	0.878	0.860	0.840	0.818	0.795	0.770
	67	0.973	0.973	0.972	0.972	0.971	0.968	0.964	0.959	0.954	0.947	0.938	0.929	0.917	0.904	0.889	0.873	0.854	0.834	0.813	0.790	0.766
Year	68	0.961	0.961	0.961	0.960	0.959	0.957	0.954	0.949	0.944	0.937	0.929	0.920	0.909	0.896	0.881	0.865	0.847	0.827	0.806	0.783	0.759
of	69	0.947	0.947	0.946	0.945	0.944	0.943	0.941	0.937	0.932	0.926	0.918	0.909	0.898	0.885	0.871	0.855	0.838	0.818	0.798	0.775	0.752
Birth	70	0.930	0.930	0.929	0.928	0.927	0.926	0.924	0.922	0.917	0.911	0.904	0.895	0.885	0.873	0.859	0.844	0.826	0.807	0.787	0.765	0.742
(M)	71	0.910	0.910	0.909	0.909	0.908	0.907	0.905	0.903	0.900	0.894	0.887	0.879	0.869	0.857	0.844	0.829	0.813	0.794	0.774	0.753	0.730
	72	0.888	0.888	0.888	0.887	0.886	0.885	0.883	0.881	0.878	0.875	0.868	0.860	0.851	0.840	0.827	0.813	0.797	0.779	0.760	0.739	0.717
	73	0.864	0.864	0.863	0.863	0.862	0.861	0.859	0.857	0.854	0.851	0.847	0.840	0.831	0.820	0.808	0.795	0.779	0.762	0.743	0.723	0.701
	74	0.838	0.838	0.837	0.837	0.836	0.835	0.833	0.831	0.829	0.826	0.822	0.817	0.809	0.799	0.787	0.774	0.759	0.743	0.725	0.705	0.685
	75	0.811	0.810	0.810	0.809	0.809	0.807	0.806	0.804	0.802	0.799	0.795	0.790	0.785	0.775	0.765	0.752	0.738	0.722	0.705	0.686	0.666
	76	0.782	0.782	0.782	0.781	0.780	0.779	0.778	0.776	0.774	0.771	0.767	0.763	0.757	0.751	0.741	0.729	0.716	0.701	0.684	0.666	0.647
	77	0.753	0.753	0.752	0.752	0.751	0.750	0.749	0.747	0.745	0.742	0.738	0.734	0.729	0.723	0.716	0.705	0.692	0.678	0.662	0.645	0.626
	78	0.723	0.723	0.723	0.722	0.721	0.720	0.719	0.717	0.715	0.713	0.709	0.705	0.700	0.695	0.688	0.680	0.668	0.654	0.639	0.623	0.605
	79	0.694	0.693	0.693	0.693	0.692	0.691	0.690	0.688	0.686	0.683	0.680	0.676	0.672	0.666	0.660	0.652	0.643	0.630	0.616	0.600	0.584
	80	0.664	0.664	0.664	0.663	0.663	0.662	0.660	0.659	0.657	0.654	0.651	0.648	0.643	0.638	0.631	0.624	0.616	0.606	0.593	0.578	0.562
	81	0.635	0.635	0.635	0.634	0.634	0.633	0.632	0.630	0.628	0.626	0.623	0.619	0.615	0.610	0.604	0.597	0.589	0.580	0.569	0.555	0.540
	82	0.607	0.607	0.607	0.606	0.605	0.605	0.603	0.602	0.600	0.598	0.595	0.592	0.587	0.582	0.577	0.570	0.562	0.554	0.544	0.533	0.519
	83	0.580	0.580	0.579	0.579	0.578	0.577	0.576	0.575	0.573	0.571	0.568	0.565	0.561	0.556	0.550	0.544	0.537	0.528	0.519	0.509	0.497

Reference

Austad, S.N. (2006). Why women live longer than men: sex differences in longevity. *Gender Medicine*, 3(2), 79-92.

Brockett, P.L. (1984). General bivariate Makeham laws. *Scandinavian Actuarial Journal*, 1984(3), 150-156, DOI: 10.1080/03461238.1984.10413763.

Carriere, J.F. (2000). Bivariate Survival Models for Coupled Lives. *Scandinavian Actuarial Journal* , 2000(1), 17-32, DOI: 10.1080/034612300750066700.

Dabrowska, B.M. (1988). Kaplan-Meier estimate on the plane. *The Annals of Statistics*, 16(4), 1475-1489.

Frees, E.W., Carriere, J., & Valdez, E. (1996). Annuity valuation with dependent mortality. *The Journal of Risk and Insurance*, 63(2), 229-261.

Gavrilov, L.A., & Gavrilova, N.S. (1996). Mortality measurement at advanced ages: a study of the social security administration death master file. *North American Actuarial Journal*, 63(2), 229-261.

Gavrilova, N.S., & Gavrilov., L.A. (2011). Aging and longevity: Mortality laws and mortality forecasts for aging population [In Czech: Stárnutí a dlouhov kost: Zákony a prognózy úmrtnosti pro stárnoucí populace]. *Demografie*, 53(2), 109-128.

Gavrilov., L.A., & Gavrilova, N.S., Krutko, V.N.& (2017). Historical Evolution of Old-Age Mortality and New Approaches to Mortality Forecasting. *Living 100 Monogr.*, available in PMC 2017 November 21.

Hougaard, P., Harvald, B., & Holm, N.V. (1992). Measuring the similarities between the lifetimes of adult Danish twins born between 1881-1930. *Journal of the American Statistical Association* , 87(417), 17-24.

Hurd, T.R. (2009). Credit Risk modeling using time-changed Brownian motion. *International Journal of Theoretical and Applied Finance*, 12, 1213–1230.

Jagger, C., & Sutton, C.J. (1991). Death after marital bereavement - is the risk increased? *Statistics in Medicine*, 10, 395-404.

Jarrow, R.A., & Protter, P. (2015). Structural versus Reduced-Form Models: A New Information-Based Perspective . *The Credit Market Handbook: Advanced Modeling Issues*, https://doi.org/10.1002/9781119201892.ch6.

Lorenzo, E.D., Sibillo, M., & Tessitore, G. (2006). A Stochastic Proportional Hazard Model for the Force of Mortality. *Journal of Forcasting*, 25(7), 529-536.

Luciano, E., & Schoutens, W. (2006). A multivariate jump-driven financial asset. *Quantitative Finance*, 29, www.carloalberto.org.

Luciano, E., Spreeuw, J., & Vigna, E. (2008). modeling stochastic mortality for dependent lives. *Insurance Mathematics and Economics,*, 43(2), 234-244 DOI: 10.1016/j.insmatheco.2008.06.005.

Luciano, E., & Semeraro, P. (2010). Multivariate time changes for Lévy asset models: Characterization and calibration. *Journal of Computational and Applied Mathematics*, 233, 1937-1953.

Manora, O., & Eisenbach, Z. (2003). Mortality after spousal loss: are there socio-demographic differences? *Social Science Medicine* , 56, 405–413.

Mitchell, D., Brockett, P.L., Mendoza-Arriaga, R., & Muthuraman, K. (2013). Credit risk modeling using time-changed Brownian motion. *Insurance: Mathematics and Economics*, 52, 275–285.

Nielsen J.J., Hulman A., & Witte DR. (2018). Spousal cardiometabolic risk factors and incidence of type 2 diabetes: a prospective analysis from the English Longitudinal Study of Ageing. *Diabetologia*, 61(7), 1572-1580. doi: 10.1007/s00125-018-4587-1. Epub 2018 Mar 8..

Seifter, A., Singh, S., McArdle, P.F., Ryan, K.A., Shuldiner, A.R., Mitchell, B.D., & Schäffer, A.A. (2014). Analysis of the bereavement effect after the death of a spouse in the Amish: a population-based retrospective cohort study. *BMJ Open* 4:(e003670). doi:10.1136/bmjopen-2013-003670.

Semeraro, P. (2008). A multivariate variance gamma model for financial applications. *International Journal of Theoretical and Applied Finance*, 11(1), 1–18.

Spreeuw, J. (2006). Types of dependence and time-dependent association between two lifetimes in single parameter copula models . *Scandinavian Actuarial Journal*, 2006(5), 286-309, DOI: 10.1080/03461230600952880.

Wang, C.W., Huang, H.C., Liu, I.C. (1988). A Quantitative Comparison of the Lee-Carter Model under Different Types of Non-Gaussian Innovations. *The Annals of Statistics* , 16(4), 1475-1489.