Collegio Carlo Alberto

Cross-generational comparison of stochastic mortality of coupled lives

Elisa Luciano
Jaap Spreeuw
Elena Vigna

No. 187
December 2010

Carlo Alberto Notebooks
www.carloalberto.org/working_papers

Cross-generational comparison of stochastic mortality of coupled lives

Elisa Luciano* Jaap Spreeuw† Elena Vigna‡

26th November 2010

Abstract

This paper studies the evolution across generations of dependence among individuals of a couple. We consider a well-known data set of couples of individuals provided by a large Canadian insurer, and select three different generations of couples. For each of them, we model the marginal survival functions with a doubly stochastic approach and perform a best-fit selection of the copula. Since the effect of censoring on the dependence structure of the couple varies across different generations, in the best-fit copula test it is necessary to restrict the attention to the subset of complete data. Despite the small sample available, the remarkable result is that the Kendall's tau varies between 30% for the young generation and 45% for the old one. As a consequence, for every candidate copula the dependence parameter decreases when younger generations are taken into account. This result is intuitive and in accordance with the observed increase in the rate of divorces, the creation of enlarged families and so on. The best fit copula is not invariant across generations, but different Archimedean copulas perform similarly. An actuarial application to pricing and reserving of joint life products and reversionary annuities shows that not only insurance companies should dismiss the simplifying independence assumption, but they should also select different dependence parameters for different generations.

1 Introduction

The ageing population phenomenon in industrialized countries creates a natural link between financial/pricing and actuarial/pension problems. Indeed, due to ageing, public (first pillar) and private (second pillar) pension systems will play a crucial role in financing the needs of a large part of individuals. Both in the

*Università di Torino, ICER and Collegio Carlo Alberto, Italy. Email: luciano at econ.unito.it

†Faculty of Actuarial Science, Cass Business School, City University, London. Email: j.spreeuw at city.ac.uk

‡Università di Torino, CeRP and Collegio Carlo Alberto, Italy. Email: elena.vigna at econ.unito.it

first and in the second pillar the offer of insurance products includes reversionary annuities. Given the expected increase of reversionary annuities in the insurance market, the problem of their correct and accurate pricing will become more and more relevant. Assessing the correct dependency between two members of a couple is the first step in this direction. Investigating whether and how this dependency evolves over time, i.e. considering different generations, is a second important aim. Whereas the first issue has been addressed in the literature, to the best of our knowledge the second has not been tackled. However, we believe that knowledge of the trend of dependency through generations could help in the long-time horizon planning of a life office. Therefore, we address both issues in this paper.

Regarding the dependency between members of a couple, the actuarial practice has been sticking to the independence assumption. This is due to a number of reasons, including lack of data, paucity of models and computational convenience. Expectedly, the existing actuarial literature rejects the independence assumption and measures the extent of mispricing through the comparison between premiums of insurance products on two lives with and without independence. In this field, the seminal paper is [9], which introduces a dataset of couples of individuals provided by a large Canadian insurer. Their paper has been followed by a few others, including [3], [5], [21], [19], [14].

The present paper, that is a follow-up of [14], addresses the issue of trend of dependency. We consider the same well-known dataset, and select three different generations of couples. For each generation we calibrate the marginal survival functions of male and female with the doubly stochastic approach. We then apply the Genest and Rivest method (see [12]) in order to find the best fit copula among a sample of Archimedean. The remarkable result is that for each copula tested, the dependence parameter decreases when younger generations are taken into account. This result is intuitive and in accordance with the observed increase in the rate of divorces, the creation of enlarged families and so on.

When we apply dependency and take its evolution into account in pricing and reserving of joint life products - as we do in the last part of the paper - we do get different results across generations. Then,the key point of our study is that not only insurance companies should dismiss the simplifying independence assumption, but they should also select different dependence parameters for different generations.

The paper is organized as follows. In section 2, we briefly review the existing literature on mortality of coupled lives. In section 3, we present the methodology used in this paper, underlining the differences with the methods used in [14], due to the selection of different generations. In section 4, we present the results of the calibration. In section 5, we present an actuarial application, by comparing the premiums of last survivor products and reversionary annuities under both the independence assumption and the generation-based dependency. Section 6 concludes.

2 Mortality of coupled lives

In this section, we review the existing actuarial literature on mortality of couples. The first paper in this field is [9], who use for the first time a large dataset of couples of individuals provided by a Canadian insurer to assess the impact of dependence on the pricing of life insurance products on two lives. They describe the marginal survival functions using the Gompertz model, the parameters of which are estimated using the Kaplan Meier empirical survival functions on the whole dataset. In order to obtain the joint survival function, they use the copula approach, and choose the Frank copula. They take into consideration the fact that data are censored when providing the maximum likelihood (ML) estimate of the parameters of the marginal survival function and the copula. They apply the calibrated joint survival probabilities in the pricing of a joint and last survivor contract. They measure the impact of the independence assumption on the premium by studying the ratio of the annuity price with dependence to that with independence. In addition, they make a sensitivity analysis of such ratio with respect to different ages of the members of the couple and different interest rates. To conclude, they make a robustness check of the choice of marginals and copula model by extending the annuity exam to Weibull margins and common shock bivariate models.

A richer description of competing models for joint survivorship is contained in [5]. They build a different dataset starting from the NIS (National Institute of Statistics) Belgian marginals for men and women. In order to describe joint mortality, they need data related to couples rather than individuals, which they obtain by sampling data from cemeteries. To this data set they first apply the Makeham model for the marginals with Markov switching intensities, then a copula model. For the dependence structure, they provide a best fit analysis of competing Archimedean copulas, using the Genest and Rivest method. The two-step ML procedure selects the Gumbel Hougaard copula, which is then used for pricing joint life products and comparing them with the case of independence. Since the dependence in their data set is low, the Gumbell Hougaard premiums are very close to the independence ones.

A rigorous marginal and copula best fit analysis is performed by [3]. He uses the same Canadian dataset introduced in [9] and focuses on the subset of complete data. The best marginal model turns out to be the Gompertz. The best fit copula is the linear-mixing frailty model.

[21] use the same Canadian dataset to describe the joint survival function of couples, but focus on the age difference between the spouses. Like [9], they select the Gompertz and the Weibull laws for the marginals, and two different copulas for the description of dependence, namely the Frank and the Gumbell Hougaard. Similarly to [9], they also study the impact of the dependence on the price of insurance products on two lives, i.e. the joint and last survivor policy. In a companion paper, [19] focus on the Weibull marginal law and on the Gumbel Hougaard copula to study the impact of dependence on the joint and last survivor premia. They still classify results depending on the age difference of the members of the couple, but they use Bayesian methods for the estimate

of the parameters.

Differently from the papers mentioned, [14] use the doubly stochastic approach to model the marginal survival functions of the spouses. In particular, they select a Feller model for the stochastic intensity of mortality, based on previous investigation of goodness of fit performed by [15] and [16]. At bivariate level, they provide a best fit copula among different Archimedean ones. The main innovation with respect to the previous literature is that the marginal survival functions in their paper are generation-based, due to the fact that the doubly stochastic approach is generation-based. As a relevant consequence, the joint survival function is based on the generations selected too. Moreover, the marginals are described by a stochastic force of mortality, that turns out to be a natural extension of the Gompertz model.

This paper extends [14] through a comparison among different generations and an application on the pricing of insurance products on two lives.

3 Methodology

This paper models mortality of couples, using a copula approach: the joint survival probability is written in terms of the marginal survival probabilities and a function - namely, the copula - which represents dependence. The calibration procedure is two-step as usual in the copula field, in that the best fit parameters of the margins are chosen separately from the best-fit parameter of the copula. This section briefly describes the modelling and calibration choices in the two steps.

Let the heads of the generation selected be (x) and (y), belonging respectively to the gender m (males) and f (females). They have remaining lifetimes T_x^m and T_y^f, which are assumed to have continuous distributions. As usual in actuarial notation, in the following (x) and (y) will refer to the initial ages of male and female, respectively. Denote by S_x^m and S_y^f the corresponding marginal survival functions:

$$S_x^m(t) = \Pr\left[T_x^m > t\right], \qquad \forall t \geq 0$$
$$S_y^f(t) = \Pr\left[T_y^f > t\right], \qquad \forall t \geq 0$$

Denote as $S_{xy}(s,t)$ the joint survival function of the couple (x,y), i.e.

$$S_{xy}(s,t) = \Pr\left[T_x^m > s, T_y^f > t\right] \qquad \forall s,t \geq 0$$

As is known, Sklar's theorem states the existence (and uniqueness over the range of the margins) of a function $C_{xy}^S : [0,1] \times [0,1] \to [0,1]$ such that, for all $(s,t) \in [0,\infty] \times [0,\infty]$, S can be represented in terms of S_x^m, S_y^f:

$$S_{xy}(s,t) = C_{xy}^S(S_x^m(s), S_y^f(t)).$$

Such copula is the so-called survival copula, obtained from the corresponding copula C via the relationship

$$C^S(v,z) = v + z \square 1 + C(1 \square v, 1 \square z) \tag{1}$$

It follows that

$$S_{xy}(s,t) = S_x^m(s) + S_y^f(t) - 1 + C_{xy}(1 - S_x^m(s), 1 - S_y^f(t)). \tag{2}$$

3.1 Marginal life distributions

At marginal level, for each generation, we find the survival functions of male and female with the doubly stochastic approach. This approach is well established in the actuarial literature, see [4], [2], [18]. The mathematical framework is the one used in the credit risk theory for the description of the random time to default, see [1], [13], [8] and [6]. Within this approach, the random time of death τ of the individual is modelled as the first jump time of a doubly stochastic process, i.e. a counting process, the intensity of which is itself a stochastic process (rather than being a constant or a deterministic function of time as in standard Poisson processes). Let the intensity of the counting process at time s be represented by a (nonnegative, measurable) function Λ of a vector of state variables $X(s)$. Under some technical properties, this construction permits to write the marginal survival probabilities as

$$S_i^j(t) = \Pr(\tau_i^j > t) = \mathbb{E}\left[\exp\left(-\int_0^t \Lambda_i^j(X(s))ds\right)\right], \tag{3}$$

where $i = x, y, j = m, f$, and the expectation is taken over the filtration generated by the state variable processes in X.

We focus on the case in which the state variables - and consequently the intensity - evolve according to a diffusion process. In order to make the model amenable to computation and calibration, we assume such process to be affine:

$$dX(t) = f(X(t))dt + g(X(t))d\tilde{W}(t),$$

where $\tilde{W}$ is a standard Brownian motion and both the drift $f(X(t))$ and the square of the diffusion coefficient $g^2(X(t))$ have affine dependence on $X(t)$.

This permits to write the marginal survival probability in (3) in terms of the intensity evaluated at time 0 - the evaluation date - and two functions of time, denoted as $\alpha(\cdot)$ and $\beta(\cdot)$. Using the results in [7] one can indeed show that

$$S_i^j(t) = \exp\left[\alpha_i^j(t) + \beta_i^j(t)X(0)\right], \tag{4}$$

where the functions $\alpha_i^j(\cdot)$ and $\beta_i^j(\cdot)$ satisfy generalized Riccati ODEs.

The modelling restrictions described up to now have then an evident advantage in terms of representation of the survival probabilities in closed form. A further simplification can be obtained by considering the intensity itself as the only relevant state variable for each gender and generation. In this case we have $\Lambda_i^j(X(s)) = \Lambda_i^j(s)$ This can be written formally as:

$$S_i^j(t) = \exp\left[\alpha_i^j(t) + \beta_i^j(t)\Lambda_i^j(0)\right]. \tag{5}$$

Practically, in order to write down the survival function for each gender and generation, it is sufficient to determine the functions $\alpha(\cdot)$ and $\beta(\cdot)$, together with the initial (observed) value of the mortality intensity. The functions will solve ODEs which depend on the type process assumed for the intensity.

Previous papers motivated the appropriateness, in the affine family, of nonnegative intensities without mean reversion. In [15] and [16], using the evidence provided by a comparison of competing models over the UK population, we focused on the following intensity:

$$d\Lambda_i^j(s) = a_i^j \Lambda_i^j(s) ds + \sigma_i^j \sqrt{\Lambda_i^j(s)} dW_i^j(s),$$

where W_i^j is a one-dimensional Wiener process. For it to be well defined, we assume that the percentage drift is positive - $a_i^j(\cdot) > 0$ - and the diffusion coefficient is nonnegative: $\sigma_i^j \geq 0$. The above process belongs to the Feller family. It stays nonnegative and shows no mean-reversion.

From the modelling point of view, this is a parsimonious choice, involving only two parameters (a_i^j, σ_i^j) for each generation i and gender j. From the empirical point of view, it proved to fit quite accurately a number of different generation mortality tables. These are the reasons that motivate its adoption in this paper too. We know that for such process $\alpha_i^j(t) = 0$, while

$$\beta_i^j(t) = \frac{1 - \exp\left(b_i^j t\right)}{c_i^j + d_i^j \exp(b_i^j t)} \tag{6}$$

where

$$\begin{cases} b_i^j = -\sqrt{\left(a_i^j\right)^2 + 2\left(\sigma_i^j\right)^2} \\ c_i^j = \frac{b_i^j + a_i^j}{2} \\ d_i^j = c_i^j - a_i^j \end{cases}$$

The survival function given by (5) with (6) is biologically reasonable (i.e. it is decreasing over time) if and only if the following condition holds (sub- and superscripts are omitted for easier notation)[1]:

$$e^{bt}(\sigma^2 + 2d^2) > \sigma^2 - 2dc. \tag{7}$$

In order to provide an estimate of the marginal parameters for each generation, $(\hat{a}_i^j, \hat{\sigma}_i^j)$, we first identify the generations. A generation could be identified with males or females born in a specific calendar year, or in several of them. The choice of the definition of generation strongly depends on the availability of data. We consider the large Canadian dataset introduced by [9]. In this dataset thousands of couples of individuals are observed in a timeframe of five years, from 29 December 1988 till 31 December 1993. Given the scarcity of data for each single year of birth of the dataset, and observing that persons with years

[1] This condition always holds in the calibrations of this paper.

of birth close to each other can be considered to belong to the same generation, we define a generation as the set of all individuals born in a fourteen-years time-interval (as in [14]). Due to strong censoring of the dataset, in selecting the three generations, we make the years of birth partially overlap. As in [14], in this work we keep the three-years age difference between male and female of the same couple, as this is the average age-difference between spouses in the whole dataset. The generations we consider are: 1900-1913, 1907-20, 1914-27 for males, 1903-1916, 1910-23, 1917-30 for females[2]. From now on, we will refer to these generations as "old", "middle" and "young". Notice that the members of each of these generations may enter the observation period in nineteen different years. For instance, the male members of the old generation may start to be observed at every age between 75 and 94. For notational convenience and according to [14], we will consider as initial age of each generation the smallest possible entry age, namely, $x = 75$ for the old male, $x = 68$ for the middle male, $x = 61$ for the young male, $y = 72$ for the old female, $y = 65$ for the middle female, $y = 58$ for the young female.

Then, we extract from the raw data the Kaplan-Meier (KM) distribution for each generation and each gender. Details of the procedure adopted in this step are in [14].

The last step consists in using the KM data to calibrate the parameters of the intensity of each gender and generation, $(\hat{a}_i^j, \hat{\sigma}_i^j)$. This is done by minimizing the mean square error between empirical and theoretical probabilities, the former being the KMs, the latter being obtained by replacing the appropriate function $\beta(\cdot)$ - i.e., (6) - in the survival function (5):

$$S_i^j(t) = \exp\left[\frac{1 - \exp\left(b_i^j t\right)}{c_i^j + d_i^j \exp(b_i^j t)} \Lambda_i^j(0)\right]. \tag{8}$$

3.2 Copula

We follow a quite established tradition in survival modelling of couples, by restricting our attention to Archimedean copulas. These copulas share a number of mathematical properties. Each one of them is obtained from a (different) generator, which is a continuous, decreasing, convex function $\phi : [0, 1] \to [0, +\infty]$, such that $\phi(1) = 0$. If one defines the inverse of the generator, ϕ^{-1}, as the function such that $\phi^{-1}(\phi(v)) = v$, the Archimedean copula corresponding to ϕ can be obtained as follows:

$$C(v, z) = \phi^{-1}(\phi(v) + \phi(z)) \tag{9}$$

One can indeed check that the resulting function has the right properties for being a copula. Usually the generator - and consequently the copula - contains one parameter, which we denote as θ. We will write indifferently ϕ or ϕ_θ.

[2] To be more precise, the males of the older generation were born between 1.1.1900 and 31.12.1913, while the corresponding females between 1.1.1903 and 31.12.1916, and so on.

On top of their mathematical similarities, Archimedean copulas share the possibility of being calibrated using the so-called Kendall's process. For the sake of simplicity, we omit the dependence on generation and gender while describing such calibration. Let K be the distribution function of the random number

$$Z = S\left(T_x^m, T_y^f\right),$$

namely

$$K(z) = \Pr(Z \leq z).$$

Genest and Rivest in [12] proved that K is linked to the generator - and consequently to the parameter which is contained in it, θ - by the relationship

$$K(z) = z - \frac{\phi_\theta(z)}{\phi_\theta^{'}(z)}, \quad 0 < z \leq 1 \tag{10}$$

where the prime denotes differentiation.

Given a complete dataset of observed death ages $\left\{\left(x_{(i)}, y_{(i)}\right), i = 1, ..., n\right\}$, of size n, Genest and Rivest suggested to choose the best-fit copula in a set of candidates through the following steps:

1) start from an estimate $\hat{\tau}$ of the Kendall's tau coefficient between the raw data $\left(x_{(i)}, y_{(i)}\right)$;

2) determine - for each candidate copula - the parameter value $\hat{\theta}$ which corresponds to $\hat{\tau}$, by working the parameter out of the relationship

$$\hat{\tau} = 4\int_0^1 [v - K(v)]\, dv + 1 \tag{11}$$

$$= 4\int_0^1 \left[\frac{\widehat{\phi}_\theta(v)}{\widehat{\phi}_\theta^{'}(v)}\right] dv + 1 \tag{12}$$

3) build - again for each copula - a theoretical K, $K_{\phi_{\hat{\theta}}}$, by substituting in (10) the estimate $\hat{\theta}$;

4) compare the theoretical K of each candidate Archimedean copula with the so-called empirical K, or Kendall's process, denoted by $\widehat{K}_n$:

$$\widehat{K}_n(z) = \frac{1}{n}\#\left\{i \,|z_i \leq z\right\}, \tag{13}$$

where

$$z_i = \frac{1}{n-1}\#\left\{\left(x_{(j)}, y_{(j)}\right)\left|x_{(j)} < x_{(i)}, y_{(j)} < y_{(i)}\right.\right\}$$

and $\#\{.\}$ indicates the cardinality of the set $\{.\}$.

5) select as best fit copula the one whose theoretical K is the least distant from the empirical one, $\widehat{K}_n$. The comparison can be performed using any distance between K and $\widehat{K}_n$. Using the L^2-norm, which is the typical one, the best fit copula is the one which minimizes the following error

$$\int_0^1 \left(K_{\phi_\theta}(v) - \widehat{K}_n(v)\right)^2 dv. \tag{14}$$

The couple data we start from are - per se - censored, since the observation period is five years

For censored data the procedure has to be adapted. The adapted procedure was first described in [20]. The modification involves selecting a starting point for K, namely ξ, as follows

$$\xi = \min_{v \geq 0} \left\{ v : \widehat{K}(v) > 0 \right\} = \min_{v \geq 0} \left\{ v : \Pr \left(\widehat{S} \, \square \, T_x^m, T_y^f \right) > v \right) < 1 \right\}, \qquad (15)$$

where $\widehat{S}$ is a non parametric estimator of the empirical joint survival function that has to take censoring into account.

The adapted procedure has been used, on the same dataset, in [14], where a single generation was considered. In the Canadian dataset the starting point ξ differs across generations. This can be explained by inspection of the meaning of ξ given by (15). In fact, ξ is the minimum value for which the empirical joint survival function is positive. Due to the definition of generation and the five years observation period, the empirical survival function takes the minimum value in $\widehat{S}(19, 19)$. Thus, we have $\xi = \widehat{S}(19, 19)$. It is evident that the value of the joint survival probability for 19 years depends on the age of the members of the generation at the beginning of the observation period. In particular, an older generation has a lower ξ than a younger generation. The ensuing cut of the domain of the $\widehat{K}$ function – that is unavoidable in the case of censored data – has an impact also on the estimate of Kendall's tau, which for censored data, is obtained after - and not before - a best fit copula has been obtained. In fact, with censored data, the calculation of $\widehat{\tau}$ is the last step of the whole procedure of [20]. The first step consists in choosing as parameter value $\hat{\theta}$ for each copula the one which minimizes the distance between the corresponding theoretical $K_{\phi_{\widehat{\theta}}(xy)}$ and the empirical $\widehat{K}_{(xy)}$. The second step consists in selecting as best fit copula the one which minimizes such a distance. The third step consists in getting an estimate of Kendall's tau from the parameter value of the best fit copula. In particular, $\widehat{\tau}$ is calculated from $\widehat{\theta}$ using the relationship:

$$\hat{\tau} = 4 \int_{\xi}^{1} \left[v \, \square \, K_{\phi_{\widehat{\theta}}(xy)}(v) \right] dv + 1. \qquad (16)$$

that involves the cutting point ξ.

Clearly, the higher is $\xi > 0$ the more overestimated is the Kendall's tau – considering that the integral in (16) provides a negative value. Therefore, the effect of ξ on $\widehat{\tau}$ strongly depends on the age of the generation chosen[3]. For instance, in the case of a very old generation such that $\xi \simeq 0$ the effect is very small, but in the case of a very young generation with $\xi \simeq 1$ the effect is such that the integral in (16) becomes negligible, leading to a distorted value of tau equal to or very close to 1. Thus, the younger the generation, the higher the overestimation of $\widehat{\tau}$. This phenomenon is not acceptable in our context, given that the focus of this paper is on the comparison of dependence among

[3] Notice that we can speak of "age of a generation", since the members of the couple have a 3-years age difference by definition.

different generations. It is then clear that the straight use of the Wang and Wells procedure, which was appropriate in the case of one generation, is not admissible in the presence of different generations. If we want to compare the association within a couple across different generations we are bound to select smaller subsets of complete data for all the generations selected. Evidently, the price to pay in order to be able to do consistent comparisons across generations is a remarkable reduction of the size of the sample, which in our case becomes 66, 102 and 66 for the old, middle and young generations respectively.

4 Calibration results

In this section, we report the results of the calibrations done on the three different generations selected.

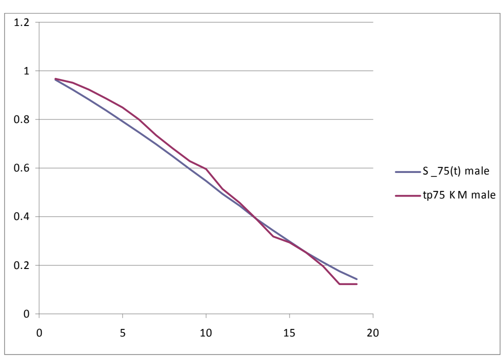
4.1 Marginal calibrations

The parameters of their marginal survival functions, for females and males, are presented (in basis points) in the following table, where OG, MG and YG stay for old generation, middle generation and young generation, respectively:

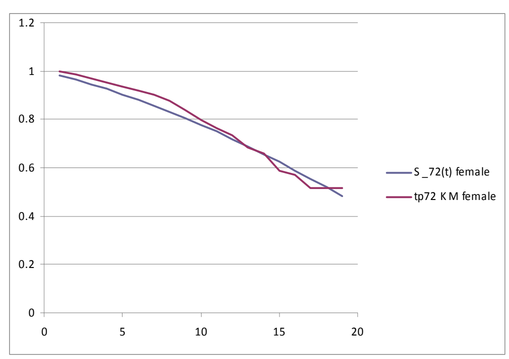
	OG Male	OG Female	MG Male	MG Female	YG Male	YG Female
a	961.045	790.232	810.051	1249.792	528.581	619.733
σ	0.007	0.057	2.426	0.021	0.019	0.5

Parameters of the marginal survival functions (in basis points).

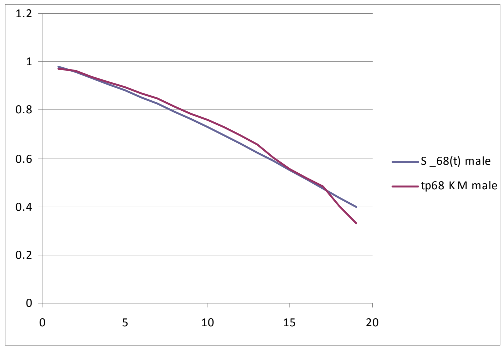
The following six figures report the plot of the survival probabilities, grouped by generation and gender. Each figure reports the analytical survival function $S_z(t)$ for initial age z, and the empirical survival function obtained with Kaplan Meier methodology.



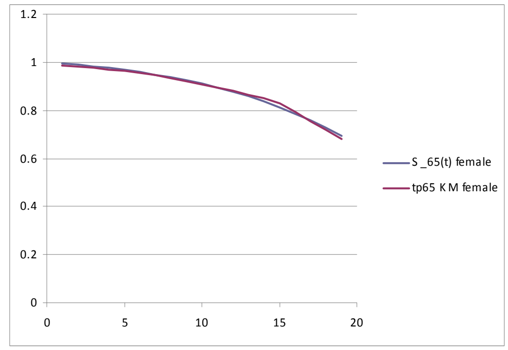
Survival function of the male of the old generation.



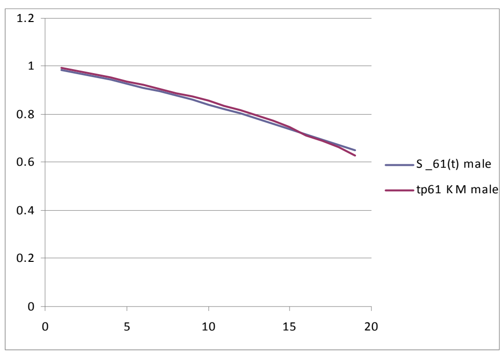
Survival function of the female of the old generation.



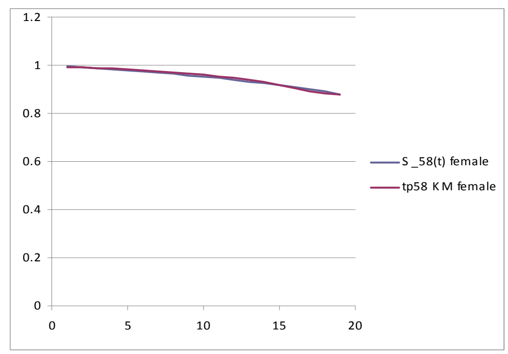
Survival function of the male of the middle generation.



Survival function of the female of the middle generation.



Survival function of the male of the young generation.



Survival function of the female of the young generation.

4.2 Joint calibration

Following Genest and Rivest, we first compute an estimate $\widehat{\tau}$ of the Kendall's tau coefficient for each generation. The empirical estimate $\widehat{\tau}$ for the Kendall's tau is decreasing over generations: it is 0.4396 for the old generation, 0.3826 for the middle one, 0.2792 for the younger one. As intuition would suggest, dependence decreases as we consider younger generations.

In order to describe this (decreasing) dependence, we examine - through a best-fit procedure - the following copulas:

Name	$C(u,v)$	**Kendall's** τ
Clayton	$\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{1}{\theta}}$	$\frac{\theta}{\theta+2}$
Gumbel-Hougaard	$\exp\left[-\left((-\ln u)^{\theta}+(-\ln v)^{\theta}\right)^{\frac{1}{\theta}}\right]$	$1-\frac{1}{\theta}$
Frank	$-\frac{1}{\theta}\ln\left[1+\frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{e^{-\theta}-1}\right]$	$1+\frac{4}{\theta}\left(\int_{t=0}^{\theta}\frac{t}{\theta(e^{t}-1)}dt-1\right)$
Nelsen 4.2.20	$\left[\ln\left(\exp\left(u^{-\theta}\right)+\exp\left(v^{-\theta}\right)-e\right)\right]^{-\frac{1}{\theta}}$	$1-\frac{4}{\theta}\frac{1}{\theta+2}-\int_{t=0}^{1}\frac{t^{\theta+1}}{\theta}\exp\left[1-t^{-\theta}\right]$
Special	$\left(\frac{-W+\sqrt{4+W^{2}}}{2}\right)^{\frac{1}{\theta}}$; with $W=\frac{1}{u^{\theta}}-u^{\theta}+\frac{1}{v^{\theta}}-v^{\theta}$	Complicated form

For each copula we estimate the parameter $\hat{\theta}$ - from $\widehat{\tau}$ - by inverting the relationship reported in the table above. We then calculate the theoretical K function, $K_{\phi_{\widehat{\theta}}}$, using (10). From the original joint data, we obtain the Kendall's process $\widehat{K_n}$ of each generation ($n = 66, 102, 66$ for the old, middle and young generations respectively), using (14). For each copula, we compute the distance between the corresponding $K_{\phi_{\widehat{\theta}}}$ function and the empirical $\widehat{K_n}$ function, according to three different norms: a) the quadratic distance (Q.d), b) the Cramer and Mises distance (CM.d), and c) the Kolomogorov-Smirnov distance (KS.d).

To test the goodness-of-fit of a copula C_θ, [10] have developed a parametric bootstrap procedure which consists of the following steps:

1. Estimate θ by a consistent estimator θ_n (like the aforementioned "inversion of Kendall's tau" approach) and calculate the distance S_n

2. Generate a large number N random samples of size n of C_{θ_n} (the copula under the null hypothesis), and for each of these samples, estimate θ by the same method as under 1., and determine the distances, given by S_j^*.

3. If $S_{1:N}^* \leq ... \leq S_{N:N}^*$ are the ordered values of the distances calculated in step 2, an estimate of the critical value of the test at level α based on S_n is given by $S^*_{\lfloor(1-\alpha)N\rfloor:N}$, while

$$\frac{1}{N}\#\left\{j : S_j^* \geq S_n\right\}$$

yields an estimate of the p-value relating to the observed value S_n (here $\lfloor x\rfloor$ denotes the integer value of x).

[10] demonstrate that this method works for any copula that satisfies a weak convergence of Kendall's process. It is particularly worthwhile for Archimedean copulas, since for this class the K function derived from a copula a) is available in closed form, and b) has a one-to-one correspondence with that copula.

The distances, according to different distance criteria, are reported in the following tables, together with the p-values. Exploring a dataset, [10] have shown that cases exist, where copulas, that are selected on the basis of minimum distance, should be rejected on the basis of the p-value. This explains the importance of the formal goodness-of-fittest as outlined above. The best fit copula – the one that provides the highest p-value – is in bold in each table. This is also the copula providing the minimum distance between $K_{\phi_{\widehat{\theta}}}$ and $\widehat{K}_n$, which is reassuring.

Old Generation (1900-13).

Copula	Q d.	p-value	CM d.	p-value	KS d.	p-value
Clayton	0.138	0.011	0.16	0.026	0.821	0.137
Gumbel Hougaard	0.048	0.395	0.06	0.486	0.689	0.351
Frank	**0.032**	**0.725**	**0.044**	**0.789**	**0.53**	**0.816**
4.2.20 Nelsen	0.204	0.004	0.248	0.017	0.998	0.051
Special	0.21	0	0.236	0.003	0.937	0.039

Middle Generation (1907-20).

Copula	Q d.	p-value	CM d.	p-value	KS d.	p-value
Clayton	**0.069**	**0.184**	**0.098**	**0.17**	**0.722**	**0.306**
Gumbel Hougaard	0.246	0	0.268	0	1.068	0.007
Frank	0.117	0.028	0.147	0.029	0.829	0.096
4.2.20 Nelsen	0.097	0.071	0.14	0.068	0.892	0.091
Special	0.078	0.123	0.111	0.12	0.868	0.101

Young Generation (1914-27).

Copula	Q d.	p-value	CM d.	p-value	KS d.	p-value
Clayton	0.052	0.377	0.096	0.271	0.832	0.2
Gumbel Hougaard	0.197	0.001	0.285	0.001	1.149	0.007
Frank	0.124	0.023	0.198	0.015	0.988	0.034
4.2.20 Nelsen	0.05	0.411	0.094	0.309	0.816	0.249
Special	**0.034**	**0.669**	**0.064**	**0.572**	**0.66**	**0.564**

Then, we see that the Frank copula performs best for the old generation, the Clayton for the middle generation and the Special for the young generation. The best fit copula is not invariant across different generations. We think that this may be due to the small size of the restricted dataset for each generation (namely, 66 couples for the old generation, 102 for the middle one, 66 for the young one). [11] observe that for a dataset of size 50 it can be hard to compare different copulas.

We also observe that different Archimedean copulas describe the dependency between two random variables in a similar way. In fact, they share some important characteristics in describing the kind of dependence (symmetry, associativity, boundedness of the diagonal section, as described in theorem 4.1.5 of [17]). Passing from one generation to the other these properties are preserved, despite the change in best fit copula.

A closer look at the table reveals that Frank and Gumbel Hougaard exhibit quite similar performance. For the old generation, these families are the only two that can be accepted at 5% significance (disregarding the 5.1% p-value for the KS-statistic reported for 4.2.20 Nelsen). On the other hand, for both the middle and young generation, only these two families would be rejected at a level of 5%. One might be able to explain these observations by considering some properties of these copulas:

a) While Clayton, 4.2.20 Nelsen and Special survival copulae all feature lower tail dependence, this is not the case for Gumbel Hougaard (upper tail dependence) and Frank (no tail dependence). (In fact, these tail dependence properties are reverted for distribution copulae).

b) It is only for Gumbel Hougaard and Frank that dependence decreases over time. For the 4.2.20 Nelsen and Special copula, dependence increases over time, while for Clayton it is constant.

In our view, the most important finding is given by the following table, that collects the parameters of all copulas for each generation (the Kendall's tau is also reported, for the sake of completeness).

GENERATION	1900-13	1907-20	1914-27
Kendall's tau	0.439627	0.382644	0.279254
Clayton	1.569	1.239	0.774
Gumbel Hougaard	1.784	1.619	1.387
Frank	4.734	3.926	2.686
4.2.20 Nelsen	0.597	0.492	0.33
Special	2.068	1.72	1.213

For each copula selected the dependence parameter is decreasing when passing from older to younger generations. This is consistent with the decrease of the Kendall's tau across generations. This result is not surprising and is in accordance with the observed increase in the rate of divorces, the creation of enlarged families, the increased independence of the woman in the family and so on. We would like to remark that this phenomenon, that is intuitive and can be considered common wisdom, in this context has been exactly quantified and measured. Furthermore, given that this result has been found with a significant small size of dataset, we would expect this decreasing trend of dependency to appear also in larger parts of the populations. Thus, we believe that an insurance company, that has at disposal large dataset of insured people, could well perform a test of dependence among members of couples of different generations to see how this dependency evolves over time. In fact, we believe the changing dependency factor across different generations has an impact on the pricing of insurance policies on two lives that should not be overlooked. Indeed, in the next section, we will show the effect of mispricing induced by the wrong formulation of the model, both with respect to the copula selection and with respect to the dependence factor.

5 Pricing bivariate contracts

In this section, we present two different actuarial applications related to the pricing of policies on two lives. Namely, we will consider the last survivor policy and the reversionary annuity. In both cases, we will measure the mispricing induced by the independence assumption.

5.1 Last survivor contract

The last survivor insurance policy pays a fixed amount in every period of time as long as at least one member of the couple is alive. In what follows, we consider the case of annual unitary payments, made at the end of the year.

If the interest rate used in the actuarial evaluation is constant at the level i over the maturity of the contract[4], the last survivor contract has the following fair premium:

$$\sum_{t=1}^{\omega} v^t \cdot {}_t p_{\overline{xy}}$$

where ${}_tp_{\overline{xy}}$ is the probability that at least one member of the couple survives t years from now, v is the discount factor, $v = (1+i)^{-1}$ and ω is, as usual, the maximal age. The needed probability is computed from the marginal and joint survival ones as

$${}_tp_{\overline{xy}} = {}_t p_x^m + {}_t p_y^f - {}_t p_{xy} = S_x^m(t) + S_y^f(t) - S_{xy}(t,t) \tag{17}$$

where the joint survival probability up to time t is obtained from the corresponding copula as in (2):

$$S_{xy}(s,t) = S_x^m(s) + S_y^f(t) - 1 + C_{xy}(1 - S_x^m(s), 1 - S_y^f(t)). \tag{18}$$

Putting the two relationships together, we have

$${}_tp_{\overline{xy}} = 1 - C_{xy}(1 - S_x^m(t), 1 - S_y^f(t)).$$

Therefore, the annuity price is equal to

$$\sum_{t=1}^{\omega} v^t \left[1 - C_{xy}(1 - S_x^m(s), 1 - S_y^f(t))\right]$$

which reduces to

$$\sum_{t=1}^{\omega} v^t \left[1 - (1 - S_x^m(s))(1 - S_y^f(t))\right]$$

in case of independency between the two lives.

We have implemented numerically the previous pricing formulas for the three generations in correspondence to different interest rates, namely $i = 1\%, 2\%, 3\%, 4\%, 5\%$. We collect the results in the following three tables, which report, for each rate, the price in correspondence to the best fit copula, the price which one would obtain by assuming independent lives, and the ratio of the two.

[4] The extension to interest rates changing deterministically over time is straightforward.

i	Frank	Independence	Ratio
1%	16.654	17.793	0.936
2%	14.924	15.926	0.937
3%	13.465	14.351	0.938
4%	12.226	13.011	0.94
5%	11.164	11.864	0.941

Old Generation.

i	Clayton	Independence	Ratio
1%	20.43	21.68	0.942
2%	18.039	19.104	0.944
3%	16.052	16.964	0.946
4%	14.387	15.173	0.948
5%	12.981	13.662	0.95

Middle Generation.

i	Special	Independence	Ratio
1%	31.329	32.752	0.957
2%	25.896	27.034	0.958
3%	21.784	22.704	0.959
4%	18.619	19.372	0.961
5%	16.143	16.766	0.963

Young Generation.

The reader can notice that the ratio is always less than 1. This result is expected and intuitive. In fact, the last survivor policy pays until the last spouse dies. The positive dependence between spouses is reflected by the fact that the joint survival probability $S_{xy}(t,t)$ is higher than in the independence case. Thus, looking at (17), a direct consequence is that the probability ${}_tp_{\overline{xy}}$ that at least one member of the couple survives t years is lower than in the independence case, implying a lower fair premium. Clearly, in the case of joint life annuity the effect of dependence is opposite and we would expect to find ratios higher than 1. This is indeed the case, as we will show in the next section on reversionary annuities. The practical consequence is that insurance companies, by assuming independence when pricing the last survivor policy, are overestimating the premium. This can be interpreted as a prudential manouvre from the point of view of insurers, and the previous tables give a measure of the extent of prudentiality so obtained. They also indicate that, for the generations under scrutiny, prudentiality decreases when the interest rate decreases or when a younger generation is selected. This trend may be particularly interesting in an era of decreasing interest rates.

5.2 Reversionary annuity

The annuity described in the previous section is just an example of a more general contract, named *reversionary annuity*, which pays 1 as long as both members are alive and a fraction R of it (R stays for "reduction factor") when only one member of the couple is alive. In fact, in this scheme, the last survivor product corresponds to $R = 1$.

More in general, the fair price of the reversionary annuity with reduction factor $R \in [0, 1]$ is

$$\sum_{t=1}^{\omega} v^t \left[R({}_tp_x^m \square_t\, p_{xy}) + R({}_tp_y^f \square_t\, p_{xy}) +_t p_{xy}\right] \tag{19}$$

since $({}_tp_x^m \square_t\, p_{xy})$ is the probability that the benefit R is paid only to the male, $({}_tp_y^f \square_t p_{xy})$ is the probability that the benefit R is paid only to the female, and ${}_tp_{xy}$ is the probability that the benefit 1 is paid when both are alive.

In practice, such contracts are quite common (see [9]) for $R = 1/2, 2/3$. However, we want to stress that $R = 0$ is nothing but the joint life annuity and $R = 1$ is nothing but the last survivor contract described before. Therefore, we have implemented the pricing formulas when R takes the values $0, 1/4, 1/3, 1/2, 2/3, 3/4, 1$.

The results are presented in the following three tables, one for each generation. As in the interest case, we devote a column to the best–fit copula price, one to the independence assumption and one to the ratio of the two prices.

R	Frank	Independence	Ratio
0	8.722	7.72	1.13
$\frac{1}{4}$	10.273	9.772	1.051
$\frac{1}{3}$	10.79	10.456	1.032
$\frac{1}{2}$	11.823	11.823	1
$\frac{2}{3}$	12.857	13.191	0.975
$\frac{3}{4}$	13.374	13.875	0.964
1	14.924	15.926	0.937

Old Generation.

R	Clayton	Independence	Ratio
0	12.326	11.261	1.095
$\frac{1}{4}$	13.754	13.222	1.04
$\frac{1}{3}$	14.23	13.875	1.026
$\frac{1}{2}$	15.183	15.183	1
$\frac{2}{3}$	16.135	16.49	0.978
$\frac{3}{4}$	16.611	17.143	0.969
1	18.039	19.104	0.944

Middle Generation.

R	Special	Independence	Ratio
0	17.547	16.41	1.069
$\frac{1}{4}$	19.635	19.066	1.03
$\frac{1}{3}$	20.33	19.951	1.019
$\frac{1}{2}$	21.722	21.722	1
$\frac{2}{3}$	23.113	23.492	0.984
$\frac{3}{4}$	23.809	24.378	0.977
1	25.896	27.034	0.958

Young Generation.

The reader can notice the following.

- Each single annuity has a value increasing in R, as expected, both under dependency and independency. This is obvious and due to the fact that a higher reduction factor implies a higher actuarial value of the benefits to be paid.

- The ratio is decreasing when R increases. This can be explained too. Let us observe that for $R = 0$ we have the joint life annuity (as nothing is paid to the last survivor) and for $R = 1$ we have the last survivor policy (where the benefit paid remains constant also after the first death). Then, R measures the weight given to the last survivor part of the reversionary annuity, with respect to the joint life annuity part. When $R = 0$ the positive dependence implies that the joint survival probability is higher than in the independence case, leading to a ratio greater than 1. At the opposite, when $R = 1$ we have the last survivor, for which we have already observed that positive dependence implies ratios lower than 1. The values $0 < R < 1$ give all the intermediate situations between these two extremes. In particular, for $R \in (0, 1/2)$ we still have ratios greater than 1, for $R \in (1/2, 1)$ we have ratios lower than 1. For $R = 1/2$, the ratio is exactly 1, in that the annuity price is shown to be not affected by the level of dependence. In fact, due to (19), the joint survival probability does not enter the premium that turns out to be:

$$\sum_{t=1}^{\omega} v_t^t \quad \frac{p_x^m +_t p_y^f}{2} \Bigg) .$$

  In this case, the weight given to the last survivor benefit is equal to that given to the joint life annuity, and the two opposite effects of overestimation and underestimation of the premium perfectly offset.

- By considering younger generations (both with and without independence) the value of the annuity increases, for any given reduction factor. This is expected, because all survival probabilities increase when considering younger generations.

6 Conclusions and extensions

This paper analyzes - first theoretically, then empirically, on the largest data set publicly available - the evolution of dependence between couples survivorship. The evolution has been studied via comparison between different generations belonging to the well-known dataset of insured couples introduced by [9]. We model the marginals of the two spouses with the doubly stochastic setup, while their dependence is captured with the copula approach. At marginal level we adopt a time-homogeneous non mean-reverting affne process for the intensity of mortality, while at bivariate level we perform a best fit copula test among different Archimedean copulas.

At theoretical level, we find that the Wang and Wells methodology for censored data should not be adopted when the focus is on comparison among different generations. This depends on the fact that its straight adoption induces an intrinsec overestimation of the degree of dependence and this overestimation varies remarkably across generations. This result implies that the dataset to be considered for the calibration should be complete rather than censored.

At empirical level, the main conclusion is that dependence does matter in pricing - and consequently in reserving - reversionary annuities, including joint and last survivor ones. In particular, the effect of dependence on the premium of insurance products on two lives varies considerably across generations and should not be overlooked. The effect of mispricing is exactly quantified. Thus, we believe that not only insurance companies should dismiss the simplifying independence assumption: they should also measure properly the degree of dependence of the particular generation the two policyholders belong to.

The natural extension of the current work, which includes stochastic mortality but deterministic interest rates, is to stochastic financial rates too. Given that it seems natural to assume independency between the evolution of interest rates and mortality over calendar time - and across generations - we do not expect this extension to provide substantial modification to our conclusions. In fact, we would expect changes at most in the magnitude of the effect, but not in its direction.

References

[1] Artzner, P., and Delbaen, F., (1992). Credit risk and prepayment option, *ASTIN Bulletin*, 22, 81-96.

[2] Biffis, E. (2005). Affine processes for dynamic mortality and actuarial valuations, *Insurance: Mathematics and Economics*, 37, 443–468.

[3] Carriere, J.F. (2000). Bivariate survival models for coupled lives. *Scandinavian Actuarial Journal*, 17-31.

[4] Dahl, M. (2004) Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts, *Insurance: Mathematics and Economics*, 35, 113–136.

[5] Denuit, M., Dhaene, J., Le Bailly de Tilleghem, C. and Teghem, S. (2001). Measuring the impact of dependence among insured lifelengths. *Belgian Actuarial Bulletin*, 1 (1), 18-39.

[6] Duffie D., Dynamic Asset Pricing Theory, Third Edition, Princeton University Press.

[7] Duffie, D., J. Pan and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions *Econometrica, 68,* 1343-1376.

[8] Duffie, D., and Singleton, K. (2003). *Credit risk*, Princeton University Press.

[9] Frees, E.W., Carriere, J.F. and Valdez, E.A. (1996). Annuity valuation with dependent mortality. *Journal of Risk and Insurance*, 63 (2), 229-261.

[10] Genest, P., Quessy, J.F., and Rémillard, B. (2006). Goodness of fit procedures for copula models based on the probability integral transformation, *Scandinavian Journal of Statistics,*1-30.

[11] Genest, C., Rémillard, B. and D. Beaudoin, (2009), Goodness-of-fit tests for copulas: A review and a power study, *Insurance: Mathematics and Economics,* 44 (2), 199-213.

[12] Genest, C. and Rivest, L.P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association* 88, 1034-1043.

[13] Lando, D. (1998). On Cox processes and credit risky securities, *Review of Derivatives Research*, 2, 99–120.

[14] Luciano, E., Spreeuw, J. and E. Vigna, (2008), Modelling stochastic mortality for dependent lives, *Insurance: Mathematics and Econonomics,* 43, 234-244.

[15] Luciano, E. and Vigna, E. (2005), A note on stochastic survival probabilities and their calibration, ICER Working Paper 1/05.

[16] Luciano, E. and Vigna, E., (2008), Non mean reverting affine processes for stochastic mortality, *Belgian Actuarial Bulletin,* 8 (1), 5-16.

[17] Nelsen, R.B. (2006). *An Introduction to Copulas*, Springer Series, Second edition.

[18] Schrager, D. F. (2006). Affine stochastic mortality. *Insurance: Mathematics and Economics* 40, 81-97.

[19] Shemyakin, A., and Youn, H., (2006), Copula models oj joint last survivor analysis, *Applied Stochastic Models in Business and Industry,* 22, 211-224.

[20] Wang, W. and Wells, M.T. (2000). Model selection and semiparametric inference for bivariate failure-time data. *Journal of the American Statistical Association* 95, 62-72.

[21] Youn, H. and Shemyakin, A. (2001). Pricing practices for joint last survivor insurance. *Actuarial Research Clearing House*, 2001.1.