Cass Business School
CITY UNIVERSITY LONDON

Faculty of Actuarial
Science and Insurance

Actuarial Research Paper No. 199

Evolution of coupled lives' dependency across generations and pricing impact

Elisa Luciano, Jaap Spreeuw
& Elena Vigna

May 2012

Cass Business School
106 Bunhill Row
London EC1Y 8TZ
Tel +44 (0)20 7040 8470
www.cass.city.ac.uk

ISBN 978-1-905752-33-1

Evolution of coupled lives' dependency across generations and pricing impact

Elisa Luciano* Jaap Spreeuw† Elena Vigna‡

25th May 2012

Abstract

This paper studies the dependence between coupled lives - both within and across generations - and its effects on prices of reversionary annuities in the presence of longevity risk. Longevity risk is represented via a stochastic mortality intensity. Dependence is modelled through copula functions. We consider Archimedean single and multi-parameter copulas. We find that dependence decreases when passing from older generations to younger generations. Not only the level of dependence but also its features - as measured by the copula - change across generations: the best-fit Archimedean copula is not the same across generations. Moreover, for all the generations under exam the single-parameter copula is dominated by the two-parameter one. The independence assumption produces quantifiable mispricing of reversionary annuities. The misspecification of the copula produces different mispricing effects on different generations. The research is conducted using a well-known dataset of double life contracts.

JEL classification: C12, C18, G22, J12.

Keywords: copula, goodness-of-fit, significance test, stochastic mortality, generation effect, reversionary annuity.

1 Introduction

Longevity risk, i.e. the risk that individuals live longer than expected, has become an issue. The ageing population in industrialized countries creates a natural link between financial and actuarial problems, i.e. annuity pricing for pension schemes. Due to ageing population, public and private pension systems will play a crucial role in financing the needs of most 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, Collegio Carlo Alberto and CeRP, Italy. Email: elena.vigna at econ.unito.it

first and in the second pillar the offer of insurance products increasingly includes last survivor and reversionary annuities. Given these features, the problem of their correct and accurate pricing becomes 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. The first issue has been addressed in the literature, at least to a certain extent. To the best of our knowledge, the second has not been tackled, since it requires a treatment of longevity risk too. We believe that knowledge of the nuances of dependence and its evolution through generations could help in the long-time horizon planning of a life office. Therefore, we address both issues in this paper.

The existing actuarial literature rejects the independence assumption and measures both the level of dependence and the extent of mispricing through the comparison between premia of insurance products on two lives with and without independence. In this field, the seminal paper is Frees et al. (1996), which introduces a dataset of couples of individuals provided by a large Canadian insurer. Their paper has been followed by a few others, including Carriere (2000), Denuit et al. (2001), Youn and Shemyakin (2001), Shemyakin and Youn (2006), and Luciano et al. (2008). Some of these papers were written before longevity risk became an issue. Others already include stochastic mortality. Consciousness of longevity risk indeed arose at the individual level, leading to the adoption of stochastic mortality - or stochastic intensity - models, in which the actual mortality rate can be different from the forecasted one. When couples are considered, one has to model both marginal stochastic mortality - taking the due age and generation effects into account - and dependence. The latter aim is achieved by coupling marginal longevity risks with a copula function. Youn and Shemyakin (2001) and Shemyakin and Youn (2006) address the existence of dependence between the members of a couple, without performing a best-fit copula analysis. Carriere (2000), Denuit et al. (2001) and Luciano et al. (2008) investigate which copula encapsulates dependence better, i.e. how dependence "looks like".

The academic literature on the static features of dependence is therefore existing, even though young. Up to our knowledge studies on the evolution of dependency across generations do not exist yet.

The present paper aims at deepening the study of dependence between members of a couple in two directions. Firstly, by addressing its evolutions over time, i.e. across generations. Secondly, by extending the features of dependence addressed and the copulas used, mainly from one parameter to multi-parameter ones. By so doing, we extend Luciano et al. (2008) from a statistical (or best-fit) point of view, and also in terms of pricing consequences.

We consider a set of Archimedean single-parameter copulas and their two- and three-parameter extensions. We implement both the marginal and joint modelling on a well-known dataset of double life contracts, provided by a large Canadian insurer. A preview of our results is the following. We find that dependence decreases when passing from older to younger 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 women in the family and so on. Not only the level of dependence but also its features - as measured by the copula - change across generations. Indeed, the best-fit Archimedean copula for a single generation is not the same across generations. However, for all the generations considered, goodness-of-fit and significance tests indicate that two-parameters copulas are significantly more suitable to describe dependence than one-parameter copulas. Less can be said on the comparison between two- and three-parameters copulas: the tests do not give clear answers regarding the dominance of one model on the other. The analysis of prices of reversionary annuities confirms that dependence matters on pricing, for the independence assumption produces a quantifiable over- and under-pricing. Mispricing is heavier for older generations than for younger ones. Surprisingly, we find that even when the copula approach is adopted, the misspecification of the copula has different and opposite effects on the mispricing for different generations.

The paper is organized as follows. In Section 2 we present the methodology. In Sections 3 and 4 we present respectively the calibration method - as required by the specific data-set - and its results. In Section 5 we discuss the effect of different models on premia of reversionary annuities, comparing them with the independence assumption. Section 6 concludes.

2 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 survival copula - which represents dependence. The calibration procedure is two-step, as usual in the copula field. The best fit parameters of the margins are chosen separately from the best-fit parameters of the copula. This section briefly describes the modelling and calibration choices in the two steps.

Let the lives 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) 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 survival copula $C : [0,1] \times [0,1] \to [0,1]$ 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:

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

2.1 Marginal survival functions

For each generation we model the marginal survival functions of males and females with the stochastic-intensity or doubly-stochastic approach. This approach is well established in the actuarial literature, see Dahl (2004), Biffis (2005) and Schrager (2006). Within this approach, the random time of death T 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 nonnegative, measurable stochastic process $\Lambda(s)$. Under some technical properties, this construction permits to write the marginal survival probabilities as

$$S_i^j(t) = \Pr(T_i^j > t) = \mathbb{E}\left[\exp\left(-\int_0^t \Lambda_i^j(s)ds\right)\right], \tag{1}$$

where $i = x, y, j = m, f$.

As usual, we focus on the case in which the intensity is an affine diffusion. This permits to write the marginal survival probability in (1) in terms of the intensity evaluated at time 0 and two functions of time, denoted as $\alpha(\cdot)$ and $\beta(\cdot)$. One can indeed show that

$$S_i^j(t) = \exp\left[\alpha_i^j(t) + \beta_i^j(t)\Lambda_i^j(0)\right], \tag{2}$$

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

Previous papers motivate the appropriateness, among generation-based affine intensities, of processes without mean reversion. Luciano and Vigna (2005, 2008), using the evidence provided by a comparison of competing models over the UK population, focused on the following intensity:

$$d\Lambda_i^j(s) = a_i^j\Lambda_i^j(s)ds + \sigma_i^j\sqrt{\Lambda_i^j(s)}dW_i^j(s), \tag{3}$$

where W_i^j is a one-dimensional Wiener process, with $a_i^j(\cdot) > 0$ and $\sigma_i^j \geq 0$. Notice that the process (3), that belongs to the Feller family, is a natural stochastic extension of the Gompertz model.[1] This is a desirable property, given that in general the Gompertz model is appropriate for ages greater than 35. In addition, Carriere (2000) finds that the Gompertz model outperforms competing mortality models on the same dataset used in this paper.

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 a number of different datasets quite accurately. These are the reasons that

[1] In fact, $\Lambda_i^j(s)$ is an exponential force of mortality if $\sigma_i^j = 0$.

motivate its adoption in this paper too. For such process we have

$$\begin{cases} \alpha_i^j(t) = 0 \\ \beta_i^j(t) = \frac{1-\exp\left(b_i^j t\right)}{c_i^j + d_i^j \exp(b_i^j t)} \end{cases} \tag{4}$$

where

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

and therefore[2]

$$S_i^j(t) = \exp\left[\frac{1-\exp\left(b_i^j t\right)}{c_i^j + d_i^j \exp(b_i^j t)} \Lambda_i^j(0)\right]. \tag{5}$$

2.2 Copulas

We follow a quite established tradition in survival modelling of couples, by restricting our attention to Archimedean copulas. We start from one-parameter copulas. Each of these copulas is obtained from a continuous, decreasing, convex function $\phi : [0,1] \to [0,+\infty]$, the generator, such that $\phi(1) = 0$. Using ϕ and its generalized inverse ϕ^{-1}, the copula is defined as follows:

$$C(v,z) = \phi^{-1}\ \left(\phi(v) + \phi(z)\right). \tag{6}$$

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. As in Luciano et al. (2008), we consider the following Archimedean copulas:

Name	$C(u,v)$
Clayton	$\left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}}$
Gumbel-Hougaard	$\exp\left[-\left((-\ln u)^{\theta} + (-\ln v)^{\theta}\right)^{\frac{1}{\theta}}\right]$
Frank	$-\frac{1}{\theta}\ln\left[1 + \frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{e^{-\theta}-1}\right]$
Nelsen 4.2.20	$\left[\ln\left(\exp\left(u^{-\theta}\right) + \exp\left(v^{-\theta}\right) - e\right)\right]^{-\frac{1}{\theta}}$
Special	$\left(\frac{-W+\sqrt{4+W^2}}{2}\right)^{\frac{1}{\theta}}$; with $W = \frac{1}{u^{\theta}} - u^{\theta} + \frac{1}{v^{\theta}} - v^{\theta}$

Table 1. Archimedean single-parameter copulas considered.

[2] The survival function given by (5) is biologically reasonable (i.e. it is decreasing over time) if and only if the following condition holds: $e^{b_i^j t}\left[\left(\sigma_i^j\right)^2 + 2\left(d_i^j\right)^2\right] > \left(\sigma_i^j\right)^2 - 2d_i^j c_i^j$. In our calibrations this is always true.

While the first three copulae are common in the literature, the last two are not. They have been included because, differently from the first three copulae, their association – as measured by the cross-ratio function – is increasing over time, which is what one would expect (see Spreeuw, 2006). In order to improve the fit of the copula, we then extend our exam to two-parameter families, obtained by combining the Archimedean copulas with the product copula. A combination which still produces a copula is the following:

$$C_{\alpha,\theta}(u,v) = u^{1-\alpha}v^{1-\alpha}C_\theta(u^\alpha, v^\alpha) \tag{7}$$

where C_θ is any Archimedean copula included in Table 1 and $\alpha \in [0,1]$. Last, we extend our consideration to three-parameter families, built in a similar way:

$$C_{\alpha,\beta,\theta}(u,v) = u^{1-\alpha}v^{1-\beta}C_\theta(u^\alpha, v^\beta) \tag{8}$$

where $(\alpha, \beta) \in [0,1] \times [0,1]$. Both the Archimedean and the two–parameter Archimedean copulas of type (7) are symmetric, in the sense that

$$C_\theta(u,v) = C_\theta(v,u)$$

and

$$C_{\alpha,\theta}(u,v) = C_{\alpha,\theta}(v,u)$$

for any $(u,v) \in [0,1] \times [0,1]$.[3]

The three-parameter Archimedean copulas in (8) are not symmetric in the sense that

$$C_{\alpha,\beta,\theta}(u,v) \neq C_{\alpha,\beta,\theta}(v,u),$$

Symmetry implies that conditional survival probabilities are equal across genders, while asymmetry allows for survivorship conditional on death of the spouse that are different when we condition on the male or on the female's death. Indeed, the conditional probabilities that the male (female) dies after $t_1(t_2^*)$, given that the female (male) dies at $t_2(t_1^*)$ are:

$$\begin{aligned} \Pr(T_x^m &> t_1 \mid T_y^f = t_2) = \left(\frac{\partial C(u,v)}{\partial v}\right)_{(u,v)=(S_x^m(t_1),S_y^f(t_2))} \\ \Pr(T_y^f &> t_2^* \mid T_x^m = t_1^*) = \left(\frac{\partial C(v,u)}{\partial v}\right)_{(v,u)=(S_x^m(t_1^*),S_y^f(t_2^*))}, \end{aligned}$$

defining t_1^* and t_2^* such that $S_x^m(t_1^*) = S_y^f(t_2)$ and $S_y^f(t_2^*) = S_x^m(t_1)$.

These probabilities are not necessarily equal in reality. As an example one could consider the case where $t_1 = t_2$. In the majority of cases, we would have $t_1^* < t_2^*$, due to the higher mortality of males, compared to females. This is the main reason why we include asymmetric copulas in the current study. Some people would argue that the first probability may be lower than the second due

[3] It can be proved (see e.g. Nelsen, 2006), that two random variables are exchangeable if and only if they are identically distributed and have a symmetric copula.

to bereavement effects which usually have a more severe impact on males than on females.

By testing whether a two-parameter family does capture dependence in a more significant way than a single parameter copula, we are interested in exploring the trade-off between a less parsimonious model and its fit. By testing the three-parameter model, not only are we interested in pushing the trade-off between parsimony and fit one step further, but we are also interested in capturing the symmetry or asymmetry between dependence of males on females and vice-versa. The three parameter model nests the two-parameter one,[4] when $\alpha = \beta$. Obviously, the last one nests all the Archimedean copulas, including the maximum, minimum and product, when $\alpha = \beta = 1$, as well as the product copula alone, when $\alpha = 0$ or $\beta = 0$.

For the sake of simplicity, in the remainder of the paper we classify in different ways the copulas considered, according to their number of parameters. In particular, we have:

1. Class 1P: the class of one-parameter copulas included in Table 1;
2. Class 2P: the class of two-parameters copulas, extensions of copulas in Class 1P via (7);
3. Class 3P: the class of three-parameters copulas, extensions of copulas in Class 1P via (8).

3 Calibration methods

3.1 Marginal survival functions

We consider the large Canadian dataset introduced by Frees et al. (1996). In this dataset thousands of couples of individuals are observed in a timeframe of five years, from 29 December 1988 till 31 December 1993. 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. The definition of generation strongly depends on the availability of data. Given the scarcity of data for each single year of birth of the dataset, and observing that persons with ages of birth close to each other can intuitively belong to the same generation, we define a generation as the set of all individuals born in a fourteen-years time-interval, as in Luciano et al. (2008). We also 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. We select two generations: 1900-1913, 1914-27 for males, 1903-1916, 1917-30 for females.[5] From now on, we refer to these generations as "old" and

[4] The asymmetric three-parameter model was introduced in the doctoral dissertation of Khoudraji (1995), under the supervision of Genest and Rivest. It is discussed in Genest et al. (1998). The latter paper also contains the proof that the three parameter function defined in (8) is a copula, and a way to simulate from it.

[5] 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.

"young". Notice that the members of each generation may enter the observation period in nineteen different years of age. 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 Luciano et al. (2008), we will consider as initial age of each generation the smallest possible entry age, namely, $x = 75$ for the old male, $x = 61$ for the young male, $y = 72$ for the old female, $y = 58$ for the young female.

Then, we extract from the raw data the Kaplan-Meier (KM) empirical distribution for each generation and each gender.

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 squared error between empirical and theoretical probabilities, the former being the KMs, the latter being obtained by replacing the appropriate function $\beta(\cdot)$ - i.e., (4) - in the survival function (5).

3.2 Copulas

The calibration of the copulas belonging to classes 1P, 2P and 3P proceeds through the pseudo-maximumlikelihood (PML) approach as in Genest et al. (1995), a brief outline of which will follow below. In order to use the methodology we need complete data. This is a relevant restriction for the dataset available. In fact, the five-years observation window of the dataset implies that the majority of couples under observation are censored data. Restricting our attention only to complete data makes the number of couples suitable for the investigation drop remarkably.

The restriction to complete data is unavoidable even in the absence of multiparameter copulas, whenever the accent is on the evolution of dependence across generations. Indeed, suppose that the focus of the investigation is only on the comparison of dependence of different generations and the copulas considered are only one-parameter. Ideally it is possible to keep the vast number of censored data and perform the best-fit copula test within each generation with the Wang and Wells procedure, see Wang and Wells (2000), that is suitable for censored data.[6] The Wang and Wells procedure involves selecting a starting point $\xi \in [0, 1]$ in the calculation of the empirical Kendall's tau $\widehat{\tau}$. In turn, the selection of ξ induces an overestimation of $\widehat{\tau}$, on which the whole procedure is based, and the higher ξ the higher the overestimation of $\widehat{\tau}$. It can be shown that the value of ξ strongly depends on the generation considered: the older the generation, the lower the cutting point ξ and vice versa. Therefore, the overestimation of $\widehat{\tau}$ strongly depends on the generation chosen, and is bigger with younger generations. This phenomenon is not acceptable if the focus of the paper is on the comparison of dependence among different generations. If the investigator wants to compare in a consistent way the association within a couple across different generations, she is bound to select smaller subsets of complete data.

[6] This was indeed the method used in Luciano et al. (2008), where the authors used censored data on the same dataset.

This has disadvantages and advantages. On the first side, the price to pay in order to be able to make consistent comparisons across generations is a remarkable reduction of the size of the sample, which in our case becomes $n = 66$ for each generation. In spite of this, as Carriere (2000) notices, this approach is inefficient due to the scarcity of couples, but informative. Indeed Carriere, when testing the null hypothesis of independence of the data, considered only the complete observations of the same data. As in Frees et al. (1996), he pooled all the generations together, rather than considering different generations. Also in his case the number of couples drops to a surprisingly low 229. On the second side, the advantage of working with complete data only is that it allows to employ relatively straightforward goodness-of-fit tests that appeared in the literature in recent years. We greatly benefit of this possibility in performing significance tests of goodness-of-fit of copulas with different numbers of parameters.

We now illustrate the PML approach. For survival copulas, the rank-based loglikelihood to be maximized with respect to the parameters has the form:

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log \left\{ c_{\boldsymbol{\theta}} \left(1 - \frac{R_i^{(1)}}{n+1}, 1 - \frac{R_i^{(2)}}{n+1} \right) \right\}, \tag{9}$$

where $\ell(\boldsymbol{\theta})$ is the resulting log-likelihood with parameter $\boldsymbol{\theta}$ (which can be real or vector valued), n is the number of data, $c_{\boldsymbol{\theta}}(u,v) = \frac{\partial^2 C_{\boldsymbol{\theta}}(u,v)}{\partial u \partial v}$ is the density of the copula $C_{\boldsymbol{\theta}}$ and $R_i^{(j)}$ is the rank of part j of observation $i, j \in \{1,2\}$, $i \in \{1,..,n\}$. The motivation for using this pseudolikelihood and its intuitive meaning are illustrated for instance in Cherubini et al. (2004).

3.3 Best fit copula among classes 1P, 2P and 3P

Once for each copula in each class (1P, 2P, 3P) the parameters have been selected using (9), we can find the best-fit copula among copulas with one, two and three parameters. In fact, there is no guarantee (and, indeed, it is not the case) that the best-fit copula remains the same in the three classes. Therefore we perform a best-fit-copula test among copulas belonging to the same class (in terms of number of parameters). This is done in a straightforward way, by comparing the loglikelihoods.

3.4 One vs multi-parameter copulas: significance test

Once the best-fit copula test has been performed within each of the three classes of copulas (1P, 2P, 3P), it is important to compare the three best copulas obtained. Clearly, the higher the number of parameters, the better the fit. There are several ways to compare the performance of models with different numbers of parameters in order to achieve the right balance between goodness-of fit and parsimony. One way to do this is to calculate the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) for each model. These criteria adjust the likelihood by penalizing for the number of parameters. A second way, for nested models, is to test the hypothesis that a model with a reduced number of parameters performs better than a full model with more parameters, within which it is nested. Both methods will be described briefly below.

3.4.1 AIC and BIC values

Let $\ell(\boldsymbol{\theta})$ be defined as in (9). Then, the Akaike Information Criterion (AIC) is defined as

$$\mathbf{AIC} = -\frac{2}{n}\left(\ell(\boldsymbol{\theta}) - p\right),$$

where p is defined as the number of parameters of $c_{\boldsymbol{\theta}}$. On the other hand, the Bayesian Information Criterion (BIC) is defined as

$$\mathbf{BIC} = -\frac{2}{n}\left(\ell(\boldsymbol{\theta}) - \frac{\log n}{2}p\right).$$

In the vast majority of applications **BIC** penalizes more than **AIC** does, since **BIC** $>$ **AIC** for $n > e^2 \approx 7.39$. Both criteria have also found their way in the actuarial literature. **AIC** has been applied by Frees and Valdez (1998) in comparing non-nested copula models fitted to a dataset consisting of losses and ALAE's (Allocated Loss Adjustment Expenses), while Cairns et al. (2009) use **BIC** in judging on the performance of several mortality forecasting models.

According to these methods, the lower the **AIC** (or **BIC**) value, the more suitable the model.

3.4.2 Statistical tests for nested models

When copula models are nested, it is possible to use the pseudolikelihood ratio test for nested models, introduced by Chen and Fan (2005). In this case, the following hypotheses have been tested:

1. (Model 1P versus Model 2P) $H_0 : \alpha = 1$
2. (Model 2P versus Model 3P) $H_0 : \alpha = \beta$.

For instance, consider Hypothesis 1 with parameter estimates for Models 1P and 2P denoted by $\overline{\theta}$ and $\left(\widehat{\theta}, \widehat{\alpha}\right)$ respectively. The test serves to reject the null hypothesis if twice the difference between the loglikelihoods, that is

$$2\sum_{i=1}^{n} \log\left[\frac{c_{\overline{\theta}}\left(1-\frac{R_i^{(1)}}{n+1}, 1-\frac{R_i^{(2)}}{n+1}\right)}{c_{\left(\widehat{\theta},\widehat{\alpha}\right)}\left(1-\frac{R_i^{(1)}}{n+1}, 1-\frac{R_i^{(2)}}{n+1}\right)}\right],$$

is sufficiently small. For details of the nonparametric bootstrap procedure, consult Chen and Fan (2005) or Genest and Favre (2007). A similar procedure applies to testing Hypothesis 2 in order to compare Models 2P and 3P. Recall that Model 2P features symmetry while Model 3P does not.

4 Calibration results

4.1 Marginal survival functions

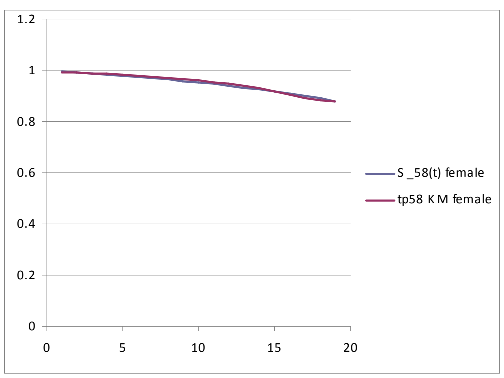
The parameters of the marginal survival functions, for both the OG (old generation) and the YG (young generation) are presented (in basis points) in the following table:

	OG male	OG female	YG Male	YG Female
a	961.045	790.232	528.581	619.733
σ	0.007	0.057	0.019	0.5

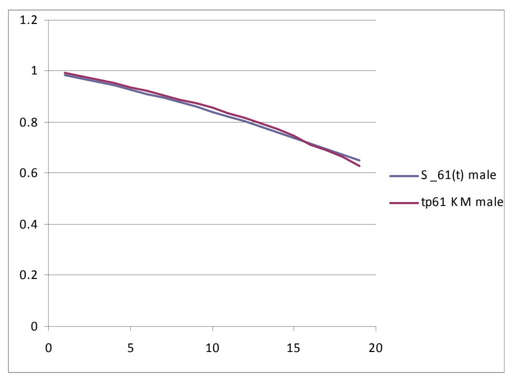
Table 2. Parameters of the marginal survival functions.

Their initial values for the two generations are $\lambda_x(0) = 0.036097$ and. $\lambda_y(0) = 0.016453$.

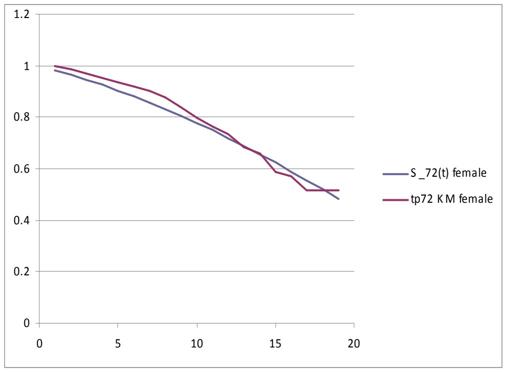
The following figures report the plot of the survival probabilities, grouped by generation and gender. Each figure reports the analytical survival function $S_z(t)$ for initial age z, and the empirical survival function obtained with the Kaplan Meier methodology.



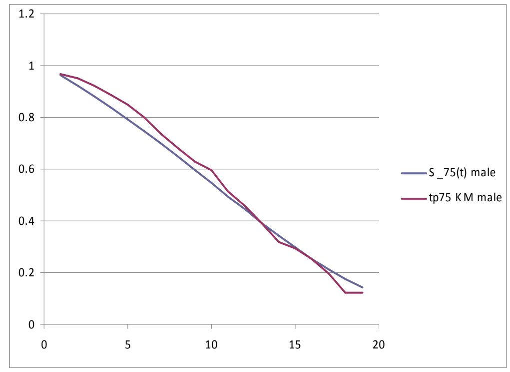
Female YG: analytical and empirical survival functions.



Male YG: analytical and empirical survival functions.



Female OG: analytical and empirical survival functions.



Male OG: analytical and empirical survival functions.

4.2 Joint calibration: decreasing dependence from OG to YG

Following Genest and Rivest (1993), 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 given by Table 3.

	OG	YG
Kendall's τ	0.440	0.279

Table 3. Kendall's τ across generations.

Dependence decreases as we consider younger generations. Consistently with the decrease of the Kendall's tau across generations, for each copula selected the dependence parameter is decreasing when passing from older to younger generations.

This interesting 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 women in the family and so on. We call this effect "cohort effect".

We acknowledge that this result can be subject to skepticism. One might wonder whether the decrease in dependency from older to younger generations is really due to cohort effect, or rather to the following effect. The two generations enter the observation window (from December 1988 to December 1993) at different ages. In particular, the OG enters at ages 75-72 (male-female) while the YG enters at ages 61-58 (male-female). Thus, the higher dependence of the old generation could be partially explained by higher dependence of a couple at older ages: one could argue that also for the same generation dependence increases with age or with the duration of marriage or coupling. We call this effect "age effect". Because of the age effect, one would expect a higher dependence coefficient for the older generation even *without* a cohort effect.

The dataset at hand does not allow us to separate the age and generation effect. In order to do so, one would need either to observe the same cohort over two different time windows, i.e. at different ages, or different generations over different windows but when they have the same initial age. No one of these possibilities is open to us, since we have a unique time window. In order to cope with the problem, we proceed as follows:

- We create artificially two observation windows out of the unique one, by distinguishing the period 29 December 1988– 30 June 1991 from the period 1 July 1991– 31 December 1993.

- For each cohort we compute the Kendall's tau in the two sub-windows, in which the same individuals have different initial ages.

Implementing this procedure we are observing the same generation with different initial ages. We obtain the following results:

τ	Dec 88- Jun 91	Jul 91- Dec 93
OG	0.636	0.502
YG	0.543	0.411

Table 4. Kendall's τ across two close observations windows.

Table 4 shows the unexpected result that for each generation τ decreases when time passes, meaning that the dependence seems to decrease when the two members of the same cohort become older. This may be due to the small number of couples at disposal in the dataset, and to the fact that in order to implement the procedure we had to further reduce the number of couples in each sub-window. Had we found the opposite results (i.e. an increasing τ) we would have been puzzled by the issue whether the decreasing Kendall's tau of Table 3 was due to cohort effects or to age effects. The likely answer would have been "by both", and it would have been impossible (with this dataset) to measure the extent of age effect and cohort effect. This would have remained an open issue.

However, the evidence produced by Table 4 indicates that – for the two generations under scrutiny – increasing age does not imply higher dependence. Transferring this conclusion to Table 3 (which is the best that we can do with this dataset), there seems to be no age effect on the decreasing Kendall's tau from OG to YG. In other words, in this dataset dependence decreases when passing from older to younger generations and this seems to be due only to cohort effects and not also to age effects.

We do not intend to claim that in general dependency decreases when considering younger generations. This would require extensive investigations that we cannot perform. However, an insurance company endowed with a complete series of data on coupled lives would be able to perform this investigation and separate age from generation effect on the Kendall's tau. This would permit to further support our results that dependence is affected by generations effect and decreases when considering younger cohorts. A complete answer to this open issue would have important implications. 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.

4.3 Joint calibration: Archimedean copulas

Following the method illustrated in Section 3, for each copula of Table 1 we calibrate the single-parameter and the two- and three-parameter versions. We then perform a best fit copula test among all copulas belonging to the same class in terms of number of parameters. The results are provided by the following two tables (the first one reports data for the old generation, the second one those for the young generation):[7]

OG	LH-1P	AIC-1P	LH-2P	AIC-2P	LH-3P	AIC-3P
Frank	11.806	−0.327	32.097	−0.912	34.834	−0.965
Clayton	6.04	−0.153	31.583	−0.896	36.035	−1.001
Gumbel-Hougaard	**14.396**	**−0.406**	**41.015**	**−1.182**	**41.212**	**−1.158**
Nelsen 4.2.20	4.919	−0.12	18.754	−0.508	31.959	−0.877
Special	2.216	−0.037	6.624	−0.14	18.283	−0.463

Table 5. Likelihood and AIC values for the Old Generation.

YG	LH-1P	AIC-1P	LH-2P	AIC-2P	LH-3P	AIC-3P
Frank	4.736	−0.113	9.31	−0.221	12.016	−0.273
Clayton	6.696	−0.173	9.555	−0.229	**13.941**	**−0.331**
Gumbel-Hougaard	2.39	−0.042	9.057	−0.214	13.386	−0.315
Nelsen 4.2.20	6.444	−0.165	8.915	−0.209	13.749	−0.326
Special	**6.832**	**−0.177**	**10.049**	**−0.244**	10.951	−0.241

Table 6. Likelihood and AIC values for the Young Generation.

The results can be summarized as follows.

1. Class 1P: The one-parameter Archimedean family that performs best according to Tables 5 and 6 is the Gumbel-Hougaard for the old generation and the Special for the young generation.[8]

[7] In the tables, LH stays for loglikelihood. In the tables we report results only for AIC, the BIC values providing the same conclusions.

[8] One sees that, for the old generation, the loglikelihood values of Clayton, Nelsen 4.2.20 and Special are low. Same observation for the young generation regarding Clayton and Gumbel-Hougaard. This is consistent with the results – not displayed here – of the parametric bootstrap procedure due to Genest et al. (2006), indicating that these 1P models are no suitable candidates. The other 1P models could not be dismissed at 5% significance.

2. Class 2P: The two-parameters Archimedean family that performs best is the Gumbel-Hougaard for the old generation and the Special for the young generation.

3. Class 3P: The three-parameters Archimedean family that performs best is the Gumbel-Hougaard for the old generation and the Clayton for the young generation.

We see that for the old generation the winner copulas are nested, but they are not for the young generation. The following two tables report the value of the parameters for the winner copulas (the first one reports data for the old generation, the second one those for the young generation):

OG	Gumb-Houg-1P	Gumb-Houg-2P	Gumb-Houg-3P
θ	1.758	13.331	12.773
α	–	0.653	0.670
β	–	–	0.657

Table 7. Value of parameters for the best-fit copulas for the Old Generation.

YG	Special-1P	Special-2P	Clayton-3P
θ	1.116	2.899	46.366
α	–	0.786	0.396
β	–	–	0.526

Table 8. Value of parameters for the best-fit copulas for the Young Generation.

We then apply the significance tests described in Section 3.4 to compare:

(i) for the old generation: 1P-GH versus 2P-GH, then 2P-GH versus 3P-GH;

(ii) for the young generation: 1P-Special versus 2P-Special, then 2P-Special versus 3P-Clayton.

For the first three out of four comparison (nested models) we apply both AIC values and statistical tests, for the fourth comparison (non-nested models) we only use AIC values.

4.3.1 Significance tests

For the Old Generation the p-values of the statistical test described in Section 3.4.2 are summarized in the following table:

OG	2P-vs-1P	3P-vs-2P
P-value	0.001	0.019

Table 9. Significance test with p-value for the Old Generation

From Table 9 we find that for the test 2P-vs-1P the Null Hypothesis ($H_0 : \alpha = 1$) should be rejected at the significance level of 5%, hence the two-parameters copula performs significantly better than the single-parameter one. This is in accordance with the comparison between AIC values (and also BIC values, here not displayed), since AIC-2P<AIC-1P. Moreover, an additional investigation (the results of which are not displayed here) shows that the dominance of the 2P-copula on the corresponding nested 1P-one occurs for every copula of Table 1. This allows us to conclude that for the Old Generation of this dataset two-parameters copulas are significantly more suitable to describe dependence than one-parameter ones.

Regarding the test 3P-vs-2P the Null Hypothesis ($H_0 : \alpha = \beta$) should be rejected at the significance level of 5%, hence the three-parameters copula performs significantly better than the two-parameter one. However, this is not in accordance with the comparison between AIC values (and also BIC values, here not displayed), since this time AIC-2P<AIC-3P. So it is not clear which model is to be preferred.

As for the Young Generation, the test 2P-vs-1P gives a p-value of 0.0224, so on that basis we can reject the Null Hypothesis that 1P is a better model than 2P. Like the Old Generation, this is in accordance with the observation of the AIC (and BIC) values of both models.

Testing the Null Hypothesis that the 2P-model is significantly better than the 3P-model is problematic. The Special and Clayton families are not nested, so it is not possible to conduct the aforementioned Chen and Fan tests. Chen and Fan (2005) also discuss a test for non-nested models, which however in this case leads to a high p-value (about 0.4, possibly due to the small dataset), and is therefore not informative. Relying just on AIC (and BIC) values would lead to the conclusion that the 3P-model (Clayton) is best for the given generation.

Since for both generations the comparison 3P-vs-2P is either problematic or gives contradictory answers, and recalling that 2P copulas display symmetry while 3P ones do not, we have conducted a generic test on symmetry of bivariate copulas, due to Genest et al. (2012). Unlike the Chen and Fan test, this test does not require *a priori* specification of the parametric copula, but is based on a consistent nonparametric estimate of the copula:

$$\widehat{C}_n(v,z) = \frac{1}{n}\sum_{i=1}^{n} I_{\left(\widehat{V}_i \le v, \widehat{Z}_i \le z\right)},$$

where $\widehat{V}_i$ and $\widehat{Z}_i$ are such that $n\widehat{V}_i = R_i^{(1)}$ and $n\widehat{Z}_i = R_i^{(2)}$. The null hypothesis H_0 simply states that the underlying copula is symmetric. The test statistic centers on a distance between $\widehat{C}(v,z)$ and $\widehat{C}(z,v)$ of the Cramér-Von Mises type:

$$S_n = \int_0^1\int_0^1 \left\{\widehat{C}_n(v,z) - \widehat{C}_n(z,v)\right\}^2 d\widehat{C}_n(v,z),$$

which, according to Genest et al. (2012) was found to be superior to alternative tests proposed. For our data set, the p-values are 0.247 (OG) and 0.162

(YG).[9] So on this basis, we cannot reject the null hypothesis of symmetry at 5% significance.

To sum up, for each generation two-parameters copulas are statistically significantly better than their one-parameter versions to describe dependence among the couples of this dataset. The comparison between two-parameters and three-parameters copulas does not provide clear answers. Whether the 3P-copula is dominant on the 2P-one depends on different factors, such as the generation and the criterion chosen.

5 Effects of dependence on pricing

In this section we present an actuarial application related to the pricing of policies on two lives. We consider a *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 this scheme, the last survivor product corresponds to $R = 1$ and the joint life annuity corresponds to $R = 0$.

If the interest rate used in the actuarial evaluation is constant at the level i over the maturity of the contract,[10] the fair price of the reversionary annuity with reduction factor $R \in [0, 1]$ is

$$\sum_{t=1}^{+\infty} v^t \left[R({}_tp_x^m -_t p_{xy}) + R({}_tp_y^f -_t p_{xy}) +_t p_{xy} \right], \tag{10}$$

where $v = (1 + i)^{-1}$ is the discount factor, $({}_tp_x^m -_t p_{xy})$ is the probability that the benefit R is paid only to the male, $({}_tp_y^f -_t p_{xy})$ 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. Connecting the survival probabilities needed to the marginal and the joint survival functions we have

$${}_tp_x^m -_t p_{xy} = S_x^m(t) - S_{xy}(t,t), \qquad \text{and} \qquad {}_tp_x^f -_t p_{xy} = S_y^f(t) - S_{xy}(t,t),$$

and also

$${}_tp_{xy} = S_{xy}(t,t) = C(S_x^m(t), S_y^f(t)).$$

Therefore, the price of the reversionary annuity is equal to

$$\begin{aligned}&\sum_{t=1}^{+\infty} v^t R \left(S_x^m(t) + S_y^f(t) - 2C(S_x^m(t), S_y^f(t)\right) + \\ &+ \sum_{t=1}^{+\infty} v^t C(S_x^m(t), S_y^f(t)).\end{aligned}$$

[9] The function exchTest of the copula package "copula" in R (R Development Core Team, 2011) was used to compute the results. For more details about the copula package in R, see Kojadinovic and Yan (2010).

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

In practice, such contracts are quite common (see Frees et al., 1996) for $R = 1/2, 2/3$. Therefore, we have implemented the pricing formulas when R takes the values $0, 1/4, 1/3, 1/2, 2/3, 3/4, 1$. Results are in the next section.

5.1 Prices of reversionary annuities

Tables 10 and 11 report the single premia of $R-$reversionary annuities for the Old Generation and the Young Generation, respectively. The interest rate used is $i = 2\%$. While the first column reports the value of R, the second reports the price of the annuity under the independence assumption. For each model specification (1P, 2P, 3P) there is a column reporting the best-fit copula price and another one reporting the ratio of the cum-dependence price to the independence one. In Table 10 "GH" stays for "Gumbel Hougaard", while in Table 11 "Spec." stays for "Special" and "Clay" stays for "Clayton".

R	Indep.	GH-1P	Ratio	GH-2P	Ratio	GH-3P	Ratio
0	7.72	8.786	1.138	8.665	1.122	8.672	1.123
1/4	9.772	10.305	1.054	10.244	1.048	10.247	1.049
1/3	10.456	10.811	1.032	10.771	1.030	10.773	1.030
1/2	11.823	11.823	1	11.823	1	11.823	1
2/3	13.191	12.835	0.975	12.876	0.976	12.874	0.976
3/4	13.875	13.342	0.964	13.402	0.966	13.399	0.966
1	15.926	14.860	0.937	14.981	0.941	14.975	0.940

Table 10. Reversionary annuity price for OG under independence and using best-fit 1P–, 2P–, 3P–copulas.

R	Indep.	Spec.-1P	Ratio	Spec.-2P	Ratio	Clay.-3P	Ratio
0	16.421	17.056	1.039	17.25	1.050	17.330	1.055
1/4	19.271	19.589	1.016	19.686	1.021	19.726	1.024
1/3	20.221	20.433	1.010	20.498	1.014	20.524	1.015
1/2	22.121	22.121	1	22.121	1	22.121	1
2/3	24.021	23.810	0.991	23.745	0.989	23.718	0.987
3/4	24.971	24.654	0.987	24.557	0.983	24.517	0.982
1	27.822	27.187	0.977	26.993	0.970	26.912	0.967

Table 11. Reversionary annuity price for YG under independence and using best-fit 1P–, 2P–, 3P–copulas.

The reader can notice the following.

1. For each R and each model specification (1P, 2P, 3P) the annuity price of the YG are higher than those of the OG. This is expected, because all survival probabilities are higher for younger insureds.

2. For each R and each model specification (1P, 2P, 3P) the YG shows ratios of cum-dependence to independence price that are closer to 1 than those of the OG. This is a clear consequence of the decreasing τ from OG to YG: the milder dependence of the YG generates prices that deviate less from the independence prices than the OG prices.

3. In both tables and for each model specification (1P, 2P, 3P) each annuity has a value increasing in R, as expected, both under dependency and independency. This is obvious: a higher reduction factor implies a higher actuarial value of the benefits to be paid.

4. In both tables and for each model specification (1P, 2P, 3P) the ratio cum-dependence/independence is decreasing when R increases. This can be explained too. Let us recall 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 part. When $R = 0$ 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 positive dependence implies lower survivorship after the spouse's death, implying 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 unaffected by the level of dependence. In fact, due to (10), the joint survival probability does not enter the premium that turns out to be:

$$\sum_{t=1}^{+\infty} v^t \left(\frac{{}_tp_x^m +_t p_y^f}{2} \right).$$

   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 each other.

5. The practical consequence of having ratios greater than 1 as long as $R < 1/2$ and ratios smaller than 1 when $R > 1/2$ is that insurance companies, by assuming independence when pricing the former, are under-pricing contracts, while they are over-pricing - or prudentially pricing - the latter. Consider the joint life case. Insurance companies which assume independence are not "on the safe side". The previous tables give a measure of the lack of safety so obtained. Consistently with item 2 above, the lack of safety is greater for the older generation. Consider now the last survivor policy, for which insurance companies which assume independency are overpricing the contract. This can be interpreted as a prudential manoeuvre from the point of view of insurers, and the previous tables give a

measure of the extent of prudence so obtained. Consistently with item 2 above, prudence decreases when the younger generation is selected.

6. In both tables the prices and the ratios of 2P and 3P copulas are almost identical. This is reassuring, given that the significance tests displayed in Section 4.3.1 do not assess the dominance of either model on the other. For this reason, in the following items we will focus only on the 2P class and neglect the 3P class.

7. For the OG the impact on prices and on the ratio cum-dependence/independence is smaller for the 2P copulas than for the 1P copula; the opposite happens for the young generation. As a consequence, for the OG, the width of the range of prices and ratios when R changes is smaller for the 2P copulas than for the 1P copula; the opposite happens for the YG.

8. Misspecification of the copula produces opposite mispricing effects on the two generations. Indeed, given the assessed superiority of the 2P model with respect to the 1P one (see Section 4.3.1), a comparison of annuity prices implies that for the OG when the dependence is described with a 1P copula rather than with a 2P one, the insurer over-prices annuities with $R < 1/2$, and under-prices annuities with $R > 1/2$. Opposite results apply to the YG: when the dependence is described with a 1P copula rather than with a 2P one, the insurer under-prices annuities with $R < 1/2$, and over-prices annuities with $R > 1/2$.

Once more, not only dependence matters, but also its evolution across generations - and the impact on pricing - matters and is not uniform.

6 Conclusions

This paper analyzes - first from a statistical, then from a pricing point of view - dependence between coupled lives of insureds and its evolution. We model the margins of the two spouses with the doubly stochastic setup, and their dependence with the copula approach. Our main aim is to perfect the existing research in single generation dependence models, when longevity risk is accounted for. We develop the statistical analysis in two directions.

The first direction is to study the evolution of dependency across generations. We find that in the largest dataset on couples publicly available, dependence decreases when passing from older to younger generations. We find that this decrease in dependence is due to the cohort effect only and not to the age effect. We provide a methodology that, whenever sufficiently rich datasets were available, would allow to separate further age and cohort effect on the evolution of dependence.

The second direction is to consider not only a class of single-parameter Archimedean copulas, but also their two- and three-parameters extensions. In each class of copulas we perform a best-fit copula test, and then compare the best copulas with significance tests that penalize for the number of parameters. We find that two-parameters copulas are significantly more suitable to describe dependence than single-parameter ones. This seems to be a new result in the literature on dependency among coupled lives. On the other hand, there are not enough elements to assess the dominance of the three-parameters class with respect to the two-parameters class, or vice versa.

After having performed statistical tests within and across generations, we study the effect of dependence on pricing insurance products. Dependence does matter in pricing reversionary annuities, including joint-life and last-survivor ones. Indeed, when assuming independence the insurer under/over-prices reversionary annuities with a reduction factor lower/greater than a half. As for the comparison between generations, surprisingly we find that the misspecification of dependence affects in a different way each cohort. In fact, when the insurer misspecifies the copula (taking one-parameter rather than two-parameters) the sign in the over- and underestimation is reversed depending on the generation under scrutiny and the reduction factor. This shows that mispricing caused by inaccurate modelling of dependence is not uniform across cohorts.

References

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

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

[3] Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A., and Balevich, I. (2009). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal* 13 (1), 1-35.

[4] Chen, X. and Y. Fan (2005). Pseudo-likelihood ratio tests for semiparametric multivariate copula selection. *The Canadian Journal of Statistics* 33 (3), 389-414.

[5] Cherubini, U., Luciano, E. and Vecchiato W. (2004). *Copula methods for finance,* John Wiley and Sons.

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

[7] 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.

[8] 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.

[9] Frees, E.W., and Valdez, E.A. (1998). Understanding relationships using copulas. *North American Actuarial Journal* 2 (1), 1-25.

[10] Genest, C and A.-C. Favre (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* 12 (4), 347-368.

[11] Genest, C., Ghoudi, K. and L. Rivest (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* 82 (3), 543-552.

[12] Genest, C., Ghoudi, K. and L. Rivest (1998). Commentary on "Understanding relationships using copulas", by Freez, E.W. and Valdez, E.A., *North American Actuarial Journal* 2, 143-149.

[13] Genest, C., Nešlehová, J., and Quessy, J.-F. (2012). Tests of symmetry for bivariate copulas. *Annals of the Institute of Statistical Mathematics*, forthcoming.

[14] Genest, C., 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.

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

[16] Khoudraji, A. (1995). *Contributions à I'étude des copules et à la modélisation des valeurs extrêmes bivariées.* Ph.D. thesis, Université Laval, Quebec, Canada.

[17] Kojadinovic, I. and Yan, J. (2010). Modeling distributions with continuous margins using the copula R package. *Journal of Statistical Software* 34 (9), 1-20.

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

[19] Luciano, E. and Vigna, E. (2005). Non mean reverting affine processes for stochastic mortality", Carlo Alberto Notebook 30/2006. *ICER Working Paper* 1/05.

[20] Luciano, E. and Vigna, E. (2008). Mortality risk via affine stochastic intensities: calibration and empirical relevance, *Belgian Actuarial Bulletin,* 8 (1), 5-16.

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

[22] R Development Core Team (2011). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

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

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

[25] Spreeuw, J. (2006). Types of dependence and time-dependent association between two lifetimes in single parameter copula models, *Scandinavian Actuarial Journal* 5, 286-309.

[26] 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.

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

Cass Business School
CITY UNIVERSITY LONDON

FACULTY OF ACTUARIAL SCIENCE AND INSURANCE

Actuarial Research Papers since 2009

Report Number	Date	Publication Title	Author
188.	January 2009	The Market Potential for Privately Financed Long Term Care Products in the UK. ISBN 978-1-905752-19-5	Leslie Mayhew
189.	June 2009	Whither Human Survival and Longevity or the Shape of things to Come. ISBN 978-1-905752-21-8	Leslie Mayhew David Smith
190	October 2009	ilc: A Collection of R Functions for Fitting a Class of Lee Carter Mortality Models using Iterative fitting Algorithms* ISBN 978-1-905752-22-5	Zoltan Butt Steven Haberman
191.	October 2009	Decomposition of Disease and Disability Life Expectancies in England, 1992-2004. ISBN 978-1-905752-23-2	Domenica Rasulo Leslie Mayhew Ben Rickayzen
192.	October 2009	Exploration of a Novel Bootstrap Technique for Estimating the Distribution of Outstanding Claims Reserves in General Insurance. ISBN 978-1-905752-24-9	Robert Cowell
193.	January 2010	Surplus Analysis for Variable Annuities with a GMDB Option. ISBN 978-1-905752-25-6	Steven Haberman Gabriella Piscopo
194.	January 2010	UK State Pension Reform in a Public Choice Framework. ISBN 978-1-905752-26-3	Philip Booth
195.	June 2010	Stochastic processes induced by Dirichlet (B-) splines: modelling multivariate asset price dynamics. ISBN 978-1-905752-28-7	Vladimir Kaishev
196.	August 2010	A comparative Study of Parametric Mortality Projection Models. ISBN 978-1-905752-29-4	Zoltan Butt Steven Haberman
197.	April 2012	Are the Dimensions of Private Information More Multiple than Expected? Information Asymmetries in the Market of Supplementary Private Health Insurance in England* ISBN 978-1-905752-31-7	Martin Karlsson Florian Klohn Ben Rickayzen
198.	May 2012	Re-thinking households – Using administrative data to count and classify households with some applications. ISBN 978-1-905752-32-4	Gillian Harper Les Mayhew
199.	May 2012	Evolution of coupled lives' dependency across generations and pricing impact. ISBN 978-1-905752-33-1	Elisa Luciano, Jaap Spreeuw Elena Vigna

Statistical Research Papers

1.	December 1995.	Some Results on the Derivatives of Matrix Functions. ISBN 1 874 770 83 2	P. Sebastiani
2.	March 1996	Coherent Criteria for Optimal Experimental Design. ISBN 1 874 770 86 7	A.P. Dawid P. Sebastiani
3.	March 1996	Maximum Entropy Sampling and Optimal Bayesian Experimental Design. ISBN 1 874 770 87 5	P. Sebastiani H.P. Wynn
4.	May 1996	A Note on D-optimal Designs for a Logistic Regression Model. ISBN 1 874 770 92 1	P. Sebastiani R. Settimi
5.	August 1996	First-order Optimal Designs for Non Linear Models. ISBN 1 874 770 95 6	P. Sebastiani R. Settimi
6.	September 1996	A Business Process Approach to Maintenance: Measurement, Decision and Control. ISBN 1 874 770 96 4	Martin J. Newby
7.	September 1996.	Moments and Generating Functions for the Absorption Distribution and its Negative Binomial Analogue. ISBN 1 874 770 97 2	Martin J. Newby
8.	November 1996.	Mixture Reduction via Predictive Scores. ISBN 1 874 770 98 0	Robert G. Cowell.
9.	March 1997.	Robust Parameter Learning in Bayesian Networks with Missing Data. ISBN 1 901615 00 6	P.Sebastiani M. Ramoni
10.	March 1997.	Guidelines for Corrective Replacement Based on Low Stochastic Structure Assumptions. ISBN 1 901615 01 4.	M.J. Newby F.P.A. Coolen
11.	March 1997	Approximations for the Absorption Distribution and its Negative Binomial Analogue. ISBN 1 901615 02 2	Martin J. Newby
12.	June 1997	The Use of Exogenous Knowledge to Learn Bayesian Networks from Incomplete Databases. ISBN 1 901615 10 3	M. Ramoni P. Sebastiani
13.	June 1997	Learning Bayesian Networks from Incomplete Databases. ISBN 1 901615 11 1	M. Ramoni P.Sebastiani
14.	June 1997	Risk Based Optimal Designs. ISBN 1 901615 13 8	P.Sebastiani H.P. Wynn
15.	June 1997.	Sampling without Replacement in Junction Trees. ISBN 1 901615 14 6	Robert G. Cowell
16.	July 1997	Optimal Overhaul Intervals with Imperfect Inspection and Repair. ISBN 1 901615 15 4	Richard A. Dagg Martin J. Newby
17.	October 1997	Bayesian Experimental Design and Shannon Information. ISBN 1 901615 17 0	P. Sebastiani. H.P. Wynn
18.	November 1997.	A Characterisation of Phase Type Distributions. ISBN 1 901615 18 9	Linda C. Wolstenholme

19.	December 1997	A Comparison of Models for Probability of Detection (POD) Curves. ISBN 1 901615 21 9	Wolstenholme L.C
20.	February 1999.	Parameter Learning from Incomplete Data Using Maximum Entropy I: Principles. ISBN 1 901615 37 5	Robert G. Cowell
21.	November 1999	Parameter Learning from Incomplete Data Using Maximum Entropy II: Application to Bayesian Networks. ISBN 1 901615 40 5	Robert G. Cowell
22.	March 2001	FINEX : Forensic Identification by Network Expert Systems. ISBN 1 901615 60X	Robert G.Cowell
23.	March 2001.	Wren Learning Bayesian Networks from Data, using Conditional Independence Tests is Equivalant to a Scoring Metric ISBN 1 901615 61 8	Robert G Cowell
24.	August 2004	Automatic, Computer Aided Geometric Design of Free-Knot, Regression Splines. ISBN 1-901615-81-2	Vladimir K Kaishev, Dimitrina S.Dimitrova, Steven Haberman Richard J. Verrall
25.	December 2004	Identification and Separation of DNA Mixtures Using Peak Area Information. ISBN 1-901615-82-0	R.G.Cowell S.L.Lauritzen J Mortera,
26.	November 2005.	The Quest for a Donor : Probability Based Methods Offer Help. ISBN 1-90161592-8	P.F.Mostad T. Egeland., R.G. Cowell V. Bosnes Ø. Braaten
27.	February 2006	Identification and Separation of DNA Mixtures Using Peak Area Information. (Updated Version of Research Report Number 25). ISBN 1-901615-94-4	R.G.Cowell S.L.Lauritzen J Mortera,
28.	October 2006	Geometrically Designed, Variable Knot Regression Splines : Asymptotics and Inference. ISBN 1-905752-02-4	Vladimir K Kaishev Dimitrina S.Dimitrova Steven Haberman Richard J. Verrall
29.	October 2006	Geometrically Designed, Variable Knot Regression Splines : Variation Diminishing Optimality of Knots. ISBN 1-905752-03-2	Vladimir K Kaishev Dimitrina S.Dimitrova Steven Haberman Richard J. Verrall
30.	November 2008	Scheduling Reentrant Jobs on Parallel Machines with a Remote Server. ISBN 978-1-905752-18-8	Konstantin Chakhlevitch Celia Glass
31.	November 2009	Probabilistic Expert Systems for Handling Artifacts in Complex DNA Mixtures. ISBN 978-1-905752-27-0	R.G.Cowell S.L. lauritzen J. Mortera
32.	April 2012	Detecting a Family Trio in a DNA Mixture. ISBN 978-1-905752-30-0	Robert G. Cowell

Papers can be downloaded from

http://www.cass.city.ac.uk/facact/research/publications.html

Faculty of Actuarial Science and Insurance

Cass Business School

Cass Business School
106 Bunhill Row,
London EC1Y 8TZ.

ISBN 978-1-905752-33-1