INSTITUT DE STATISTIQUE
BIOSTATISTIQUE ET
SCIENCES ACTUARIELLES
(ISBA)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

DISCUSSION
PAPER

2014/05

Efficient approximations for numbers of survivors in the Lee-Carter model

GBARI, S. and M. DENUIT

EFFICIENT APPROXIMATIONS FOR NUMBERS OF SURVIVORS IN THE LEE-CARTER MODEL

Samuel Gbari & Michel Denuit
Institut de statistique, biostatistique et sciences actuarielles - ISBA
Université Catholique de Louvain
B-1348 Louvain-la-Neuve, Belgium
samuel.gbari@uclouvain.be
michel.denuit@uclouvain.be

February 21, 2014

Abstract

In portfolios of life annuity contracts, the payments made by an annuity provider (an insurance company or a pension fund) are driven by the random number of survivors. This paper aims to provide accurate approximations for the present value of the payments made by the annuity provider. These approximations account not only for systematic longevity risk but also for the diversifiable fluctuations around the unknown life table. They provide the practitioner with a useful tool avoiding the problem of simulations within simulations in, for instance, Solvency 2 calculations, valid whatever the size of the portfolio.

Key words and phrases: Life annuity, mortality projection, Lee-Carter model, comonotonicity, supermodular order, increasing directionally convex order, risk measures.

1 Introduction and motivation

In this paper, we consider the present value of life annuity payments accounting for the stochastic nature of decrements. Precisely, the systematic longevity risk coming from the unknown underlying life table as well as the theoretically diversifiable risk of random fluctuations around this life table are both taken into account. Thus, the size of the portfolio now enters the calculations and this dimension is very important for small to medium-sized portfolios (see, e.g., Donnelly, 2011).

In the literature, the case of life annuity policies has been treated quite extensively but only in the limiting case, for homogeneous portfolios comprising infinitely many (conditionally) independent contracts. The applicability of these limiting results may be questioned in insurance practice as life annuity portfolios do not always contain enough policies to reach full diversification. For these reasons, Hoedemakers et al. (2005) proposed to approximate the distribution of the number of survivors using the Normal Power formula. In this paper, we pursue this idea and we allow for unknown future mortality improvements, the death probabilities prevailing in the future being difficult to assess.

After Lee and Carter (1992), we assume that the death rate at age x in calendar year t is of the form $\exp(\alpha_x+\beta_x\kappa_t)$. Here, the time index κ_t reflects the general level of mortality and the age-specific component β_x represents how rapidly or slowly mortality at each age varies when the general level of mortality changes. The dynamics of the time index is usually described by ARIMA models. Conditional survival probabilities, given the time index future trajectory, are complicated functions of the κ_ts. As there is no analytical expression available for their distribution function, Denuit and Dhaene (2007) used comonotonicity to approximate the distribution of the sums of strongly correlated LogNormal random variables playing a central role in the Lee-Carter framework. Expanding on this approach, Denuit (2008) derived analytic approximations for the quantiles of the life annuity conditional expected present value given the κ_ts. This is made by supplementing the comonotonic approximations for the conditional survival probabilities worked out in Denuit and Dhaene (2007) with a second approximation of the same type for the life annuity conditional expected present value, given the κ_ts. Denuit, Haberman and Renshaw (2010) further studied the quality of these approximations, allowing for general ARIMA models instead of the simple random walk with drift adopted in the majority of papers using Lee-Carter methodology.

In this paper, our aim is to develop accurate approximations for the present value of the payments made in favor of a group of n annuitants. The size n of the group now explicitly enters the computations so that our results apply also to small portfolios. Deriving the exact distribution for the present value of life annuity payments requires extensive simulations or numerical evaluations. The approximations derived in this paper after Denuit and Dhaene (2007) and Denuit (2008) avoid the requirement to conduct simulations within simulations in, for instance, Solvency 2 reserving calculations. Numerical illustrations show that the comonotonic approximations perform well, which suggests that they can be used in practice to evaluate the consequences of the uncertainty in future death rates.

To derive an effective comonotonic approximation, it is essential to identify in the problem under consideration random variables that are as much positively correlated as possible. Partial sums are often good candidates in that respect, as demonstrated in Denuit and Dhaene (2007). For portfolios of life annuities, the numbers of survivors up to times $1, 2, 3, \ldots$ form a strongly positively dependent sequence for which the comonotonic approximation is expected to work well. This is precisely the intuitive idea exploited in the present paper, which turns out to provide accurate approximations. Notice that making the lifetimes comonotonic in an homogeneous portfolio means that all policyholders die at the same time, which is very crude. Hence, it is important to select appropriately the random variables which will be replaced by their comonotonic versions.

The paper is organized as follows. In Section 2, we briefly recall the comonotonic approximations for the conditional survival probabilities derived by Denuit and Dhaene (2007) and for the conditional expectation of annuity payments present value derived by Denuit (2008). We supplement previous results with increasing directionally convex stochastic inequalities between the Lee-Carter conditional survival probabilities and their approximations. Section 3 proposes new approximations for the consecutive numbers of survivors. It is established there that the approximate numbers of survivors dominate the Lee-Carter ones in the increasing directionally convex order, which allows us to derive stop-loss order stochastic inequalities for the present value of life annuity payments. Numerical illustrations are discussed in Section 4. Section 5 briefly concludes.

2 Comonotonic approximations

2.1 Conditional survival probabilities

In this paper, we assume that the force of mortality at age x and time t, denoted as $\mu_x(t)$, is constant within bands of age and time in the Lexis diagram, but allowed to vary from one band to the next. Specifically, given any integer age x and calendar year t, it is supposed that

$$\mu_{x+\xi}(t+\tau) = \mu_x(t) \text{ for } 0 \leq \xi, \tau < 1. \tag{2.1}$$

Furthermore, the force of mortality is of the form

$$\ln \mu_x(t) = \alpha_x + \beta_x \kappa_t. \tag{2.2}$$

Henceforth, we will assume that the values $\kappa_1, \ldots, \kappa_{t_0}$ are known but that $\kappa_{t_0+1}, \kappa_{t_0+2}, \ldots$ are unknown and have to be projected from some appropriate time series model. The future trajectory $\kappa_{t_0+1}, \kappa_{t_0+2}, \ldots$ is henceforth denoted as $\boldsymbol{\kappa}$. Therefore, the force of mortality $\mu_x(t)$ given in (2.2) is not constant but develops over time following a stochatic process.

Consider an individual aged x_0 in calendar year t_0, with remaining lifetime T subject to (2.1)-(2.2). Define $\delta_j = \exp(\alpha_{x_0+j}) > 0$, $Z_j = \beta_{x_0+j}\kappa_{t_0+j}$ and

$$S_d \;=\; \sum_{j=0}^{d-1} \exp\left(\alpha_{x_0+j} + \beta_{x_0+j}\kappa_{t_0+j}\right) = \sum_{j=0}^{d-1} \delta_j \exp(Z_j).$$

In the applications, the time index is generally modelled by means of ARIMA time series models. Hence, we assume that $\boldsymbol{\kappa}$ is multivariate Normal so that we have $Z_j \sim \mathcal{N}or(\mu_j, \sigma_j^2)$. Then, the conditional survival probability over the next d years, given the future trajectory $\boldsymbol{\kappa}$ of the time index is given by

$$\Pr[T > d|\boldsymbol{\kappa}] = \exp(-S_d) = {}_dP_{x_0}(t_0), \;\; d = 1, 2, \ldots$$

Denuit and Dhaene (2007) proposed comonotonic approximations for the conditional survival probabilities ${}_dP_{x_0}(t_0)$. Specifically, these conditional probabilities are expected to be closely dependent for increasing values of d since they can be viewed as the exponential of the sum of death rates from age x_0 to age x_0+d-1. So, it may be reasonable to approximate the random vector of conditional survival probabilities with its comonotonic version.

Recall that a random vector $(X_1,\ldots,X_d)$ is said to be comonotonic if, and only if, there exist a random variable Z and non-decreasing functions $g_1,\ldots,g_d$, such that $(X_1,\ldots,X_d)$ is distributed as $\big(g_1(Z),\ldots,g_d(Z)\big)$. Equivalently, $(X_1,\ldots,X_d)$ is comonotonic if it is distributed as $\big(g_1(Z),\ldots,g_d(Z)\big)$ with $g_1,\ldots,g_d$ non-increasing. In particular, we may choose Z to be uniformly distributed over the unit interval $[0,1]$ and g_i to be the quantile function of X_i, i.e. the left-continuous inverse of the distribution function of X_i. A detailed account of comonotonicity can be found in Dhaene et al. (2002a,b) and Denuit et al. (2005).

In order to determine whether the approximations derived in this paper are conservative, we can use the following stochastic order relations. For more details, the readers are referred, e.g., to Denuit et al. (2005). Considering two random variables X and Y, X is said to be smaller than Y in the increasing convex order, or stop-loss order, henceforth denoted as $X \preceq_{\text{icx}} Y$, if the inequality $\mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)]$ holds true for all the non-decreasing and convex functions g for which the expectations exist. A usual strengthening of the stop-loss order is obtained by requiring in addition that the means of the random variables to be compared are equal. More precisely, X is said to be smaller than Y in the convex order, henceforth denoted by $X \preceq_{\text{cx}} Y$, if $\mathbb{E}[X]=\mathbb{E}[Y]$ and $X \preceq_{\text{icx}} Y$ simultaneously hold. The term "convex" is used since $X \preceq_{\text{cx}} Y \Leftrightarrow \mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)]$ for all the convex functions g for which the expectations exist.

Stochastic orderings $\preceq_{\text{cx}}$ and $\preceq_{\text{icx}}$ aim to mathematically express the intuitive ideas of "being less variable than" and "being smaller and less variable than" for random variables. Dealing with random vectors, $\preceq_{\text{cx}}$ and $\preceq_{\text{icx}}$ may apply marginally to each component but we also need multivariate stochastic order relations that translate the fact that the components of one of these vectors are "more positively dependent" than those of the other random vector. The supermodular order translates this idea in mathematical terms. Precisely, recall that a function $g:\mathbb{R}^d \to \mathbb{R}$ is said to be supermodular if the inequality

$$
\begin{aligned}
g(x_1,\ldots,x_i+\epsilon,\ldots,x_j+\delta,\ldots,x_d) - g(x_1,\ldots,x_i+\epsilon,\ldots,x_j,\ldots,x_d)\\
\geq g(x_1,\ldots,x_i,\ldots,x_j+\delta,\ldots,x_d) - g(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_d)
\end{aligned}
$$

holds for all $\boldsymbol{x}\in\mathbb{R}^d$, $1\leq i<j\leq d$ and all $\epsilon,\delta>0$. If the function is regular enough then supermodularity corresponds to $\frac{\partial^2}{\partial x_i \partial x_j} g \geq 0$ for every $i \neq j$. Now, consider two d-dimensional random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ such that $\mathbb{E}[g(\boldsymbol{X})] \leq \mathbb{E}[g(\boldsymbol{Y})]$ for all supermodular functions $g:\mathbb{R}^d\to\mathbb{R}$, provided the expectations exist. Then $\boldsymbol{X}$ is said to be smaller than $\boldsymbol{Y}$ in the supermodular order, which is denoted by $\boldsymbol{X}\preceq_{\text{sm}}\boldsymbol{Y}$. In words, $\boldsymbol{X}\preceq_{\text{sm}}\boldsymbol{Y}$ means that $X_1,\ldots,X_d$ are less positively related than $Y_1,\ldots,Y_d$. Notice that $\boldsymbol{X}\preceq_{\text{sm}}\boldsymbol{Y} \Rightarrow X_i$ and Y_i are identically distributed for each i so that $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same univariate marginals. Also, $\boldsymbol{X}\preceq_{\text{sm}}\boldsymbol{Y} \Rightarrow \mathbb{C}\text{ov}[X_i,X_j] \leq \mathbb{C}\text{ov}[Y_i,Y_j]$ for all $i\neq j$ as $g(\boldsymbol{x})=x_ix_j$ is obviously supermodular, which shows that $\preceq_{\text{sm}}$ indeed has the intuitive meaning stated above.

If $X_i=g_i(Z_i)$, where $g_1,\ldots,g_d$ are all decreasing or all increasing and where $Z_1,\ldots,Z_d$ are identically distributed but strongly correlated, then we can approximate $\sum_{i=1}^d X_i$ by the comonotonic sum $\sum_{i=1}^{d} g_i(Z)$, where Z is distributed as Z_i. This strategy often appears to be conservative as the stochastic inequalities

$$\big(g_1(Z_1),\ldots,g_d(Z_d)\big) \preceq_{\text{sm}} \big(g_1(Z),\ldots,g_d(Z)\big) \text{ and } \sum_{i=1}^{d} g_i(Z_i) \preceq_{\text{cx}} \sum_{i=1}^{d} g_i(Z) \tag{2.3}$$

both hold true, the latter being a consequence of the former.

Thus, we approximate S_d by a sum of perfectly dependent random variables, with the same marginal distributions, that is, by

$$S_d^u = \sum_{j=0}^{d-1} \delta_j \exp(\mu_j + \sigma_j Z), \text{ with } Z \sim \mathcal{N}or(0,1). \tag{2.4}$$

Let $S_1^u, S_2^u, \ldots$ be defined from (2.4) for $d = 1, 2, \ldots$ with the same random variable Z (so that they are comonotonic), and define

$$_dP_{x_0}^u(t_0) = \exp(-S_d^u), \;\; d = 1, 2, \ldots$$

The distribution of $_dP_{x_0}^u(t_0)$ is easily obtained from the quantile additivity for sums S_d^u of comonotonic random variables. For more details, we refer the reader to Denuit and Dhaene (2007).

The corresponding one-year survival probabilities are given by

$$P_{x_0+k}(t_0+k) = \exp(-\delta_k \exp(Z_k)) \text{ and } P_{x_0+k}^u(t_0+k) = \exp(-\delta_k \exp(\mu_k+\sigma_k Z)) \text{ for } k = 1, 2, \ldots$$

Clearly,

$$_dP_{x_0}(t_0) = \prod_{k=0}^{d-1} P_{x_0+k}(t_0+k) \text{ and } {_dP_{x_0}^u}(t_0) = \prod_{k=0}^{d-1} P_{x_0+k}^u(t_0+k).$$

Marginally, i.e. for fixed k, the one-year survival probabilities $P_{x_0+k}(t_0+k)$ and $P_{x_0+k}^u(t_0+k)$ are identically distributed. This is not the case for the d-year survival probabilities $_dP_{x_0}(t_0)$ and $_dP_{x_0}^u(t_0)$. As pointed out by Denuit and Dhaene (2007), the stochastic inequality $_dP_{x_0}(t_0) \preceq_{\text{icx}} {_dP_{x_0}^u}(t_0)$ holds true for any integer $d \geq 2$. The next result extends these stochastic inequalities to random vectors of conditional survival probabilities up to different time horizons. To this end, we need the increasing directionally convex order. Recall that the function g is said to be directionally convex if it is supermodular and coordinatewise convex. If g is twice differentiable then it is directionally convex if, and only if, $\frac{\partial^2}{\partial x_i \partial x_j} g \geq 0$ for all $i, j \in \{1, \ldots, d\}$. Now, the d-dimensional random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ are said to be ordered in the increasing directionally convex order, which is denoted by $\boldsymbol{X} \preceq_{\text{idir-cx}} \boldsymbol{Y}$, if $\mathbb{E}[g(\boldsymbol{X})] \leq \mathbb{E}[g(\boldsymbol{Y})]$ for all non-decreasing functions $g : \mathbb{R}^d \to \mathbb{R}$ that are directionally convex, provided the expectations exist. Notice that $\boldsymbol{X} \preceq_{\text{idir-cx}} \boldsymbol{Y} \Rightarrow X_i \preceq_{\text{icx}} Y_i$ for each $i = 1, \ldots, d$. The increasing directionally convex order is closely related to the supermodular order with main difference that supermodular order compares only dependence structures of random vectors with fixed marginals, whereas the directionally convex order additionally takes into account the variability of the marginals, which may then be different, as it is the case here for d-year conditional survival probabilities $_dP_{x_0}(t_0)$ and $_dP_{x_0}^u(t_0)$ that are known to be ordered in the $\preceq_{\text{icx}}$-sense.

Proposition 2.1. *For every integer $d \geq 2$, we have*

(i) $\Big(P_{x_0}(t_0), P_{x_0+1}(t_0+1), \ldots, P_{x_0+d}(t_0+d)\Big) \preceq_{sm} \Big(P^u_{x_0}(t_0), P^u_{x_0+1}(t_0+1), \ldots, P^u_{x_0+d}(t_0+d)\Big)$.

(ii) $\Big(P_{x_0}(t_0), {}_2P_{x_0}(t_0), \ldots, {}_dP_{x_0}(t_0)\Big) \preceq_{idir\text{-}cx} \Big(P^u_{x_0}(t_0), {}_2P^u_{x_0}(t_0), \ldots, {}_dP^u_{x_0}(t_0)\Big)$.

Proof. The stochastic inequality stated under (i) directly follows from (2.3). Considering (ii), we know from (2.3) that

$$\Big(\exp(-S_1), \exp(-S_2), \ldots, \exp(-S_d)\Big)$$

$$\preceq_{\text{sm}} \Big(\exp\big(-F^{-1}_{S_1}(U)\big), \exp\big(-F^{-1}_{S_2}(U)\big), \ldots, \exp\big(-F^{-1}_{S_d}(U)\big)\Big)$$

where U is uniformly distributed over the unit interval (0,1). Now, both random vectors $\Big(\exp\big(-F^{-1}_{S_1}(U)\big), \ldots, \exp\big(-F^{-1}_{S_d}(U)\big)\Big)$ and $\Big(\exp(-S^u_1), \ldots, \exp(-S^u_d)\Big)$ are comonotonic and marginally $\exp\big(-F^{-1}_{S_j}(U)\big) \preceq_{\text{icx}} \exp(-S^u_j)$ for $j = 1, 2, \ldots, d$. They are then ordered in the increasing directionally convex sense by Theorem 2.4 in Balakrishnan et al. (2012):

$$\Big(\exp\big(-F^{-1}_{S_1}(U)\big), \exp\big(-F^{-1}_{S_2}(U)\big), \ldots, \exp\big(-F^{-1}_{S_d}(U)\big)\Big)$$

$$\preceq_{\text{idir-cx}} \Big(\exp(-S^u_1), \exp(-S^u_2), \ldots, \exp(-S^u_d)\Big).$$

Combining the two stochastic inequalities provides the announced result. □

2.2 Systematic risk

Let us consider a basic life annuity contract paying € 1 at the end of each year, as long as the annuitant survives. Let $v(s,t)$ be the present value at time s of a unit payment made at time t, $s \leq t$. The random variable

$$a_{\overline{T}|} = \sum_{d=1}^{\lfloor T \rfloor} v(0,d)$$

corresponds to the present value of the payments made to an annuitant aged x_0 in calendar year t_0 whose remaining lifetime is T (with the convention that the empty sum is zero). Here, the discount factors $v(0,d)$ can be deduced from an appropriate yield curve and are thus treated as known, deterministic values. The conditional expectation of the payments made to this annuitant, given the time index, is

$$a_{x_0}(t_0|\boldsymbol{\kappa}) \quad = \quad \mathbb{E}[a_{\overline{T}|}|\boldsymbol{\kappa}] = \sum_{d\geq 1} {}_dP_{x_0}(t_0)v(0,d).$$

Proposition 2.1 allows us to get the following result derived by Denuit (2008) which provides an upper bound on $a_{x_0}(t_0|\boldsymbol{\kappa})$ in the $\preceq_{\text{icx}}$-sense.

Corollary 2.2. $a_{x_0}(t_0|\boldsymbol{\kappa}) \preceq_{icx} a^u_{x_0}(t_0|Z) = \sum_{d\geq 1} {}_dP^u_{x_0}(t_0)v(0,d)$.

Proof. Recall that any positive linear combinations of the components of random vectors ordered in the $\preceq_{\text{idir-cx}}$-sense are ordered in the $\preceq_{\text{icx}}$-sense; see e.g. Proposition 3.4.67 in Denuit et al. (2005). The announced result then follows from Proposition 2.1(ii). □

The random variable $a_{x_0}(t_0|\boldsymbol{\kappa})$ can be seen as the residual risk per annuity contract in an infinitely large portfolio where only systematic longevity risk remains. Corollary 2.2 provides, thus, a conservative approximation for the risk borne by the annuity provider when the portfolio is sufficiently large.

3 Present value of life annuity payments

Consider an insurance portfolio made of n life insurance policies covering individuals aged x_0 in calendar year t_0 with identically distributed remaining lifetimes $T_1, T_2, \ldots, T_n$. Precisely, given the future trajectory $\boldsymbol{\kappa}$ of the time index, the random variables $T_1, T_2, \ldots, T_n$ are independent and subject to the common death rate (2.2). Generating a realization of each T_i gives a very detailed picture of the cash-flows but requires n simulations. If the portfolio is homogeneous with respect to survival probabilities and sum insured, it is often enough to simulate the number L_k of survivors to age $x_0 + k$ for $k = 1, 2, \ldots$, starting from $L_0 = n$. This considerably speeds the simulation process.

Let $\mathbb{I}[A]$ be the indicator variable of the event A, equal to 1 if A is realized and to 0 otherwise. We can decompose L_k into

$$\begin{aligned} L_k &= \sum_{i=1}^{n} \mathbb{I}[T_i > k] \\ &= L_{k+1} + D_k = \sum_{j\geq k} D_j \end{aligned}$$

where D_k is the number of deaths recorded during the time interval $(k, k+1)$. Therefore, given the conditional survival probabilities, we can expect that replacing the random variables $L_1, L_2, \ldots$ with comonotonic ones provides an accurate approximation. This approach considerably reduces the computational burden (as a single realization gives the whole trajectory $L_1, L_2, \ldots$ needed to determine the future cash-flows).

The present value of the payments made to this homogeneous group of n annuitants is

$$V = \sum_{i=1}^{n} \sum_{k\geq 1} \mathbb{I}[T_i > k] v(0,k) = \sum_{k\geq 1} L_k v(0,k).$$

Clearly, if the future mortality is unknown and predicted by means of the Lee-Carter model, the random variables L_k depend on the future trajectory $\boldsymbol{\kappa}$ of the time index. Denuit (2008) studied the residual risk per policy $a_{x_0}(t_0|\boldsymbol{\kappa})$ in such a portfolio with n tending to $+\infty$: considering the identity $\mathbb{E}[V|\boldsymbol{\kappa}] = n a_{x_0}(t_0|\boldsymbol{\kappa})$, the large portfolio approximation $V \approx n a_{x_0}(t_0|\boldsymbol{\kappa})$ is expected to provide accurate results provided n is large enough. The present paper aims to study the random variable V and to develop efficient approximations for its quantiles, not only in large portfolios but also in smaller ones.

To ease the computations, we replace the original L_k with its approximation L_k^u built as follows. Given independent random variables U_{ik}, $i = 1, 2, \ldots$, $k = 1, 2, \ldots$ uniformly distributed over the unit interval $(0, 1)$, we can represent L_k, $k \geq 1$, as

$$\begin{aligned} L_k &=_d \sum_{i=1}^{L_{k-1}} \mathbb{I}[U_{ik} < P_{x_0+k-1}(t_0 + k - 1)] \\ &=_d \sum_{i=1}^{n} \prod_{j=1}^{k} \mathbb{I}[U_{ij} < P_{x_0+j-1}(t_0 + j - 1)] \end{aligned}$$

starting from $L_0 = n$. The first step consists in replacing L_k with $\widetilde{L}_k$ given by

$$\begin{aligned} \widetilde{L}_k &=_d \sum_{i=1}^{L_{k-1}} \mathbb{I}[U_{ik} < P^u_{x_0+k-1}(t_0 + k - 1)] \\ &=_d \sum_{i=1}^{n} \prod_{j=1}^{k} \mathbb{I}[U_{ij} < P^u_{x_0+j-1}(t_0 + j - 1)] \end{aligned}$$

starting from $\widetilde{L}_0 = n$.

The family of Binomial distributions with exponent n and success probability p is known to be stochastically increasing convex in its parameter p. Therefore, Property 3.4.38 in Denuit et al. (2005) ensures that ${}_kP_{x_0}(t_0) \preceq_{\text{icx}} {}_kP^u_{x_0}(t_0) \Rightarrow L_k \preceq_{\text{icx}} \widetilde{L}_k$. The next result extends these marginal comparisons to random vectors $(L_1, L_2, \ldots, L_d)$ by means of the increasing directionally convex order.

Proposition 3.1. *For every integer $d \geq 2$, we have $(L_1, \ldots, L_d) \preceq_{\text{idir-cx}} (\widetilde{L}_1, \ldots, \widetilde{L}_d)$.*

Proof. Given a function $g : \{0, 1, \ldots, n\}^d \to \mathbb{R}$, define the auxiliary function $g^\star : [0, 1]^d \to \mathbb{R}$ as

$$\begin{aligned} g^\star(p_{x_0}, \ldots, p_{x_0+d-1}) &= E[g(L_1, \ldots, L_d) | P_{x_0+k-1}(t_0) = p_{x_0+k-1},\ k = 0, 1 \ldots, d] \\ &= \sum_{l_1=0}^{n} \sum_{l_2=0}^{l_1} \cdots \sum_{l_d=0}^{l_{d-1}} g(l_1, l_2, \ldots, l_d) \\ &\quad \binom{n}{l_1} p_{x_0}^{l_1} q_{x_0}^{n-l_1} \binom{l_1}{l_2} p_{x_0+1}^{l_2} q_{x_0+1}^{l_1-l_2} \cdots \binom{l_{d-1}}{l_d} p_{x_0+d-1}^{l_d} q_{x_0+d-1}^{l_{d-1}-l_d}. \end{aligned}$$

Now, we need to prove that the auxiliary function $g^\star$ is supermodular provided g is increasing and directionally convex. If this is the case, then

$$\begin{aligned} E[g(L_1, \ldots, L_d)] &= E[g^\star(P_{x_0}(t_0), \ldots, P_{x_0+d-1}(t_0))] \\ &\leq E[g^\star(P^u_{x_0}(t_0), \ldots, P^u_{x_0+d-1}(t_0))] \text{ by Proposition 2.1(i)} \\ &= E[g(\widetilde{L}_1, \ldots, \widetilde{L}_d)]. \end{aligned}$$

Precisely, define the first-order and the second-order differences of g as

$$\begin{aligned}\Delta_j g(l_1,l_2,\ldots,l_d) &= g(l_1,l_2,\ldots,l_j+1,\ldots,l_d)-g(l_1,l_2,\ldots,l_j,\ldots,l_d)\\ \Delta^2_{j_1,j_2} g(l_1,l_2,\ldots,l_d) &= \Delta_{j_2} g(l_1,l_2,\ldots,l_{j_1}+1,\ldots,l_d)-\Delta_{j_2} g(l_1,l_2,\ldots,l_{j_1},\ldots,l_d).\end{aligned}$$

We then have to show that

$$\frac{\partial^2}{\partial p_{x_0+j_1}\partial p_{x_0+j_2}} g^\star(p_{x_0},\ldots,p_{x_0+d-1}) \geq 0 \text{ for every } j_1<j_2$$

provided the first-order and the second-order differences of g are all non-negative. The proof is inspired from Property A.1 in Denuit and Mesfioui (2013) who established the following useful result about the family $\{X_{n,p}, n\in\mathbb{N}, p\in(0,1)\}$ of Binomially distributed random variables with mean np and variance $np(1-p)$: for any $g:\{0,\ldots,n\}\to\mathbb{R}$, we have

$$\frac{\partial}{\partial p}\mathbb{E}[g(X_{n,p})] = n\mathbb{E}\left[\Delta g(X_{n-1,p})\right]. \tag{3.1}$$

As, given $L_{j_1}=l$, L_{j_1+1} is Binomially distributed with parameters l and $p_{x_0+j_1}$, formula (3.1) allows us to write

$$\begin{aligned}\frac{\partial g^\star}{\partial p_{x_0+j_1}}(p_{x_0},\ldots,p_{x_0+d-1}) &= \sum_{l_1=0}^{n}\sum_{l_2=0}^{l_1}\cdots\sum_{l_{j_1+1}=0}^{l_{j_1}-1}\cdots\sum_{l_d=0}^{l_{d-1}}\Delta_{j_1+1}g(l_1,l_2,\ldots,l_d)\\ &\quad \binom{n}{l_1}p_{x_0}^{l_1}q_{x_0}^{n-l_1}\binom{l_1}{l_2}p_{x_0+1}^{l_2}q_{x_0+1}^{l_1-l_2}\ldots l_{j_1}\binom{l_{j_1}-1}{l_{j_1+1}}p_{x_0+j_1}^{l_{j_1+1}}q_{x_0+j_1}^{l_{j_1}-l_{j_1+1}-1}\ldots\\ &\quad \binom{l_{d-1}}{l_d}p_{x_0+d-1}^{l_d}q_{x_0+d-1}^{l_{d-1}-l_d},\end{aligned}$$

with the convention that the empty sum is zero. A similar reasoning shows that for every $j_1<j_2$

$$\begin{aligned}&\frac{\partial^2}{\partial p_{x_0+j_1}\partial p_{x_0+j_2}} g^\star(p_{x_0},\ldots,p_{x_0+d-1})\\ &= \sum_{l_1=0}^{n}\sum_{l_2=0}^{l_1}\cdots\sum_{l_{j_1+1}=0}^{l_{j_1}-1}\cdots\sum_{l_{j_2+1}=0}^{l_{j_2}-1}\cdots\sum_{l_d=0}^{l_{d-1}}\Delta^2_{j_1+1,j_2+1}g(l_1,l_2,\ldots,l_d)\\ &\quad \binom{n}{l_1}p_{x_0}^{l_1}q_{x_0}^{n-l_1}\binom{l_1}{l_2}p_{x_0+1}^{l_2}q_{x_0+1}^{l_1-l_2}\ldots l_{j_1}\binom{l_{j_1}-1}{l_{j_1+1}}p_{x_0+j_1}^{l_{j_1+1}}q_{x_0+j_1}^{l_{j_1}-l_{j_1+1}-1}\ldots\\ &\quad \ldots l_{j_2}\binom{l_{j_2}-1}{l_{j_2+1}}p_{x_0+j_2}^{l_{j_2+1}}q_{x_0+j_2}^{l_{j_2}-l_{j_2+1}-1}\ldots\binom{l_{d-1}}{l_d}p_{x_0+d-1}^{l_d}q_{x_0+d-1}^{l_{d-1}-l_d},\end{aligned}$$

which ends the proof. □

Given $\widetilde{L}_{k-1}$, $\widetilde{L}_k$ is Binomially distributed with exponent $\widetilde{L}_{k-1}$ and success probability $P^u_{x_0+k-1}(t_0+k-1)$. Unconditionally, $\widetilde{L}_k$ is Binomially distributed with exponent n and success probability ${}_kP^u_{x_0}(t_0)$. However, given ${}_kP^u_{x_0}(t_0)$, $k=1,2,\ldots$, the random variables $\widetilde{L}_k$

are not comonotonic. The second step now consists in making them perfectly conditionally dependent, switching to

$$L_k^u = \sum_{i=0}^{n-1} \mathbb{I}\left[U > \sum_{j=0}^{i} \binom{n}{j} \big({}_kP_{x_0}^u(t_0)\big)^j \big(1 - {}_kP_{x_0}^u(t_0)\big)^{n-j}\right] \tag{3.2}$$

where U is a unit uniform random variable independent of Z involded in the ${}_kP_{x_0}^u(t_0)$. Note that here, the same random variables U and Z are used to define all the L_k^u. However, the L_k^u are not comonotonic as they depend on two random variables U and Z.

Proposition 3.2. *For every integer $d \geq 2$, we have $(L_1, \ldots, L_d) \preceq_{idir\text{-}cx} (L_1^u, \ldots, L_d^u)$.*

Proof. Considering Proposition 3.1 and the transitivity of $\preceq_{\text{idir-cx}}$, we need to establish that the stochastic inequality $(\widetilde{L}_1, \ldots, \widetilde{L}_d) \preceq_{\text{idir-cx}} (L_1^u, \ldots, L_d^u)$ is valid. Given a supermodular function g,

$$\begin{aligned} \mathbb{E}[g(\widetilde{L}_1, \ldots, \widetilde{L}_d)] &= \mathbb{E}\Big[\mathbb{E}[g(\widetilde{L}_1, \ldots, \widetilde{L}_d)|Z]\Big] \\ &\leq \mathbb{E}\Big[\mathbb{E}[g(L_1^u, \ldots, L_d^u)|Z]\Big] \\ &\leq \mathbb{E}[g(L_1^u, \ldots, L_d^u)] \end{aligned}$$

so that $(\widetilde{L}_1, \ldots, \widetilde{L}_d) \preceq_{\text{sm}} (L_1^u, \ldots, L_d^u)$ holds, which ends the proof. □

The present value of life annuity payments V can then be approximated by

$$V^u = \sum_{k \geq 1} L_k^u v(0, k).$$

Simulating V^u only requires two independent realizations of unit uniform U and standard Normal Z random variables. Proposition 3.2 allows us to compare V with V^u in the $\preceq_{\text{icx}}$-sense, as stated next.

Corollary 3.3. *$V \preceq_{icx} V^u$.*

Proof. It suffices to proceed as for Corollary 2.2. □

Corollary 3.3 shows that replacing V with V^u is a conservative strategy in the increasing convex sense. This result is intuitive since it corresponds to the strategy where we increase the conditional survival probabilities in the $\preceq_{\text{icx}}$-sense and simultaneously worsen the dependence structure of the numbers of survivors into conditional comonotonicity.

4 Numerical illustrations

4.1 Assumptions

Here, we use the same α_x and β_x as in Denuit (2008), under the random walk with drift model for the time index, that is,

$$\kappa_t = \kappa_{t-1} + \theta + \xi_t \tag{4.1}$$

with independent errors $\xi_t \sim \mathcal{N}or(0, \sigma^2)$. The point estimate of the time index at time $t_0 + k$ with all data available up to t_0 is thus $\kappa_{t_0} + k\theta$, which follows a straight line as a function of the forecast horizon k, with slope θ. The conditional variance of the forecast is $k\sigma^2$. Therefore, the conditional standard errors for the forecast increase with the square root of the distance to the forecast horizon k. Here, $\hat{\theta} = -0.4169175$, $\hat{\sigma}^2 = 0.3333644$, and $\kappa_{t_0} = -7.79652$. The constant technical interest rate is 3%. We consider portfolios of size $n = 100$ (small), 1000 (medium) and 10000 (large) to illustrate the trade-off between diversifiable risk and systematic risk. Finally, we consider $x_0 = 65$, the usual retirement age.

4.2 Another approximation

In addition to the upper bounds for the survival probabilities and for the numbers of survivors proposed in the preceding sections, we also consider an approximation proposed by Denuit and Dhaene (2007) which does not dominate its exact counterpart in the $\preceq_{\text{icx}}$-sense. Precisely, considering the first-order approximation $\Lambda_d = \sum_{j=0}^{d-1} \delta_j \exp(\mu_j) Z_j$ of S_d, define

$$\begin{aligned} S_d^l &= \mathbb{E}[S_d | \Lambda_d] \\ &= \sum_{j=0}^{d-1} \delta_j \exp\left(\mu_j + r_j(d)\sigma_j Z + \frac{1}{2}(1 - (r_j(d))^2)\sigma_j^2\right) \end{aligned}$$

where $r_i(d)$, $i = 0, \ldots, d-1$, is the correlation coefficient between Λ_d and Z_i, that is,

$$r_i(d) = \frac{\mathbb{C}\text{ov}[Z_i, \Lambda_d]}{\sigma_i \sigma_{\Lambda_d}} = \frac{\sum_{j=0}^{d-1} \delta_j \exp(\mu_j) \mathbb{C}\text{ov}[Z_i, Z_j]}{\sigma_i \sqrt{\sum_{j=0}^{d-1} \sum_{k=0}^{d-1} \delta_j \delta_k \exp(\mu_j + \mu_k) \beta_{x+j} \beta_{x+k} \min\{j, k\} \sigma^2}}$$

where $\mathbb{C}\text{ov}[Z_i, Z_j] = \beta_{x_0+i} \beta_{x_0+j} \min\{i, j\} \sigma^2$, with the convention $r_0(d) = 0$. We then have $S_d^l \preceq_{\text{cx}} S_d$.

Now, define

$${}_dP_{x_0}^l(t_0) = \exp(-S_d^l), \;\; d = 1, 2, \ldots$$

Note that, again, the same random variable Z is used for all the values of d. As pointed out by Denuit and Dhaene (2007), we have ${}_dP_{x_0}^l(t_0) \preceq_{\text{icx}} {}_dP_{x_0}(t_0)$. However, these comparisons do not extend to random vectors of conditional survival probabilities nor to linear combinations of ${}_dP_{x_0}^l(t_0)$ and ${}_dP_{x_0}(t_0)$. Indeed, as shown in Denuit (2008), defining

$$a_x^l(t_0|Z) = \sum_{d \geq 1} {}_dP_{x_0}^l(t_0) v(0, d),$$

we only have

$$\mathbb{E}[a_{x_0}(t_0|\boldsymbol{\kappa})] \geq \mathbb{E}[a_{x_0}^l(t_0|Z)]. \tag{4.2}$$

Considering the single crossing visible on Figure 4.2 in Denuit (2008), no $\preceq_{\text{icx}}$-comparison generally holds between $a_{x_0}^l(t_0|Z)$ and $a_{x_0}(t_0|\boldsymbol{\kappa})$. It can be seen there that the quantiles of $a_{x_0}^l(t_0|Z)$ are larger than the corresponding quantiles of $a_{x_0}(t_0|\boldsymbol{\kappa})$ after the unique crossing point but the respective expected values contradict a possible $\preceq_{\text{icx}}$-ranking as the difference in the stop-loss transforms must change sign at least once.

We also consider V^l defined by means of survivors L_k^l obtained from (3.2) with ${}_dP_{x_0}^l(t_0)$ replacing ${}_dP_{x_0}^u(t_0)$. The same reasoning based on Figures 4.1-4.3 below shows that, in general, V and V^l cannot be ordered in the increasing convex sense as $\mathbb{E}[V] \geq \mathbb{E}[V^l]$ holds.

4.3 Results

We made 1 000 000 simulations of the future trajectory $\boldsymbol{\kappa}$ of the time index and of pairs (U, Z) producing V^u. Figures 4.1-4.3 display the distribution functions of V, V^u and V^l, together with those of the large-portfolio approximations $n \times a_x(t_0|\boldsymbol{\kappa})$, $n \times a_x^u(t_0|Z)$ and $n \times a_x^l(t_0|Z)$ for $n = 100$, 1000 and 10000, respectively.

We clearly see that the distribution function of V single crosses the distribution function of V^u and dominates it after the crossing point, which confirms the increasing convex inequality established in Corollary 3.3 as $\mathbb{E}[V^u] = 1215.363 > \mathbb{E}[V] = 1215.232$ for $n = 100$. Despite the single crossing between the distribution functions of V and V^l, no $\preceq_{\text{icx}}$-ranking holds as $\mathbb{E}[V^l] = 1215.219 < \mathbb{E}[V] = 1215.232$. The same comments apply for larger n, multiplying the expected values by 10 or by 100.

The unique crossing is also clearly visible between the distribution function of V and the large portfolio approximations $na_{x_0}(t_0|\boldsymbol{\kappa})$, $na_{x_0}^u(t_0|Z)$ and $na_{x_0}^l(t_0|Z)$. The single crossing between the distribution functions of V and of $na_{x_0}(t_0|\boldsymbol{\kappa})$ confirms the stochastic inequality $na_{x_0}(t_0|\boldsymbol{\kappa}) \preceq_{\text{cx}} V$. However, as

$$n\mathbb{E}[a_{x_0}^l(t_0|Z)] \leq \mathbb{E}[V] \leq n\mathbb{E}[a_{x_0}^u(t_0|Z)],$$

Figures 4.1-4.3 suggest different $\preceq_{\text{icx}}$-rankings between V and either $na_{x_0}^u(t_0|Z)$ or $na_{x_0}^l(t_0|Z)$, depending on the size of the portfolio. For $n = 100$, the unique crossing visible on Figure 4.1 together with the inequality between the respective expected values shows that $na_{x_0}^l(t_0|Z) \preceq_{\text{icx}} V$ holds, whereas for $n = 10000$, we see from Figure 4.3 that $V \preceq_{\text{icx}} na_{x_0}^u(t_0|Z)$ is valid in the numerical example.

When the portfolio size n increases, we see that the distribution functions get closer. Also, for $n = 10000$, the distribution functions of V and $na_{x_0}(t_0|\boldsymbol{\kappa})$ are nearly identical but still exhibit the single crossing supporting the stochastic inequality $na_{x_0}(t_0|\boldsymbol{\kappa}) \preceq_{\text{cx}} V$. However, we see from Figure 4.3 that the distribution functions of $na_{x_0}^u(t_0|Z)$ and $na_{x_0}^l(t_0|Z)$ are now both smaller than the distribution of V after the crossing point, suggesting that these approximations are conservative in the right tail.

In addition to distribution functions, we also evaluate the accuracy of the approximation by means of the quantile functions. Specifically, the relative differences between the VaRs of V and V^l or V^u are depicted in Figures 4.4, 4.5, and 4.6. The quantiles of the approximation V^u and V^l are quite close to the exact quantiles of V since their relative differences varies between -5% and +5% when $n = 100$, -2% and +2% when $n = 1000$, and -1.5% and +1.5% when $n = 10000$ for usual probability levels.

5 Discussion

This paper proposes approximations for the number of survivors when future mortality is predicted by means of the Lee-Carter model. Accurate approximations for the present value

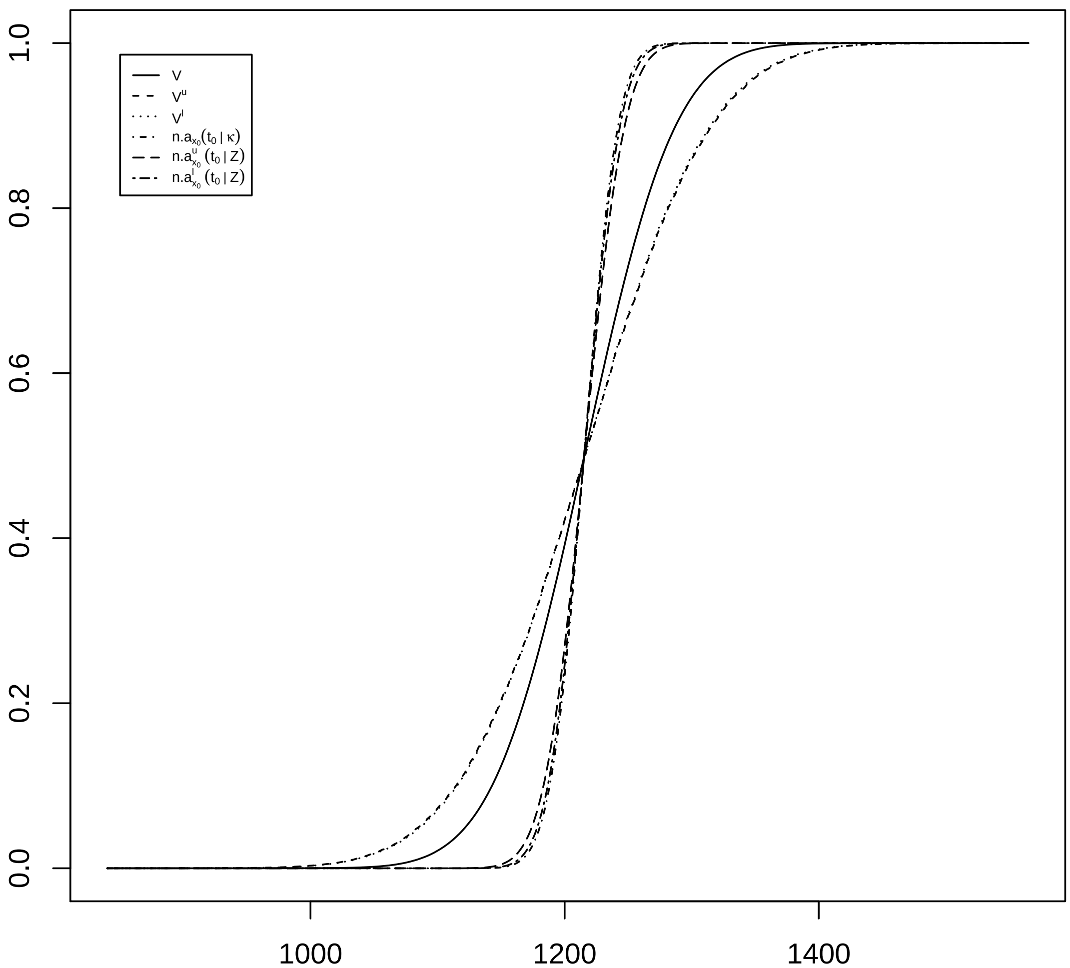

Figure 4.1: Distribution function of V (—), of V^u (- - -), of V^l $(\cdots)$ and of the large-portfolio approximations $n \times a_x(t_0)$ $(\cdot - \cdot - \cdot)$, $n \times a^u_x(t_0)$ $(---)$ and $n \times a^l_x(t_0)$ $(\cdot - \cdot - \cdot)$ for $n = 100$.

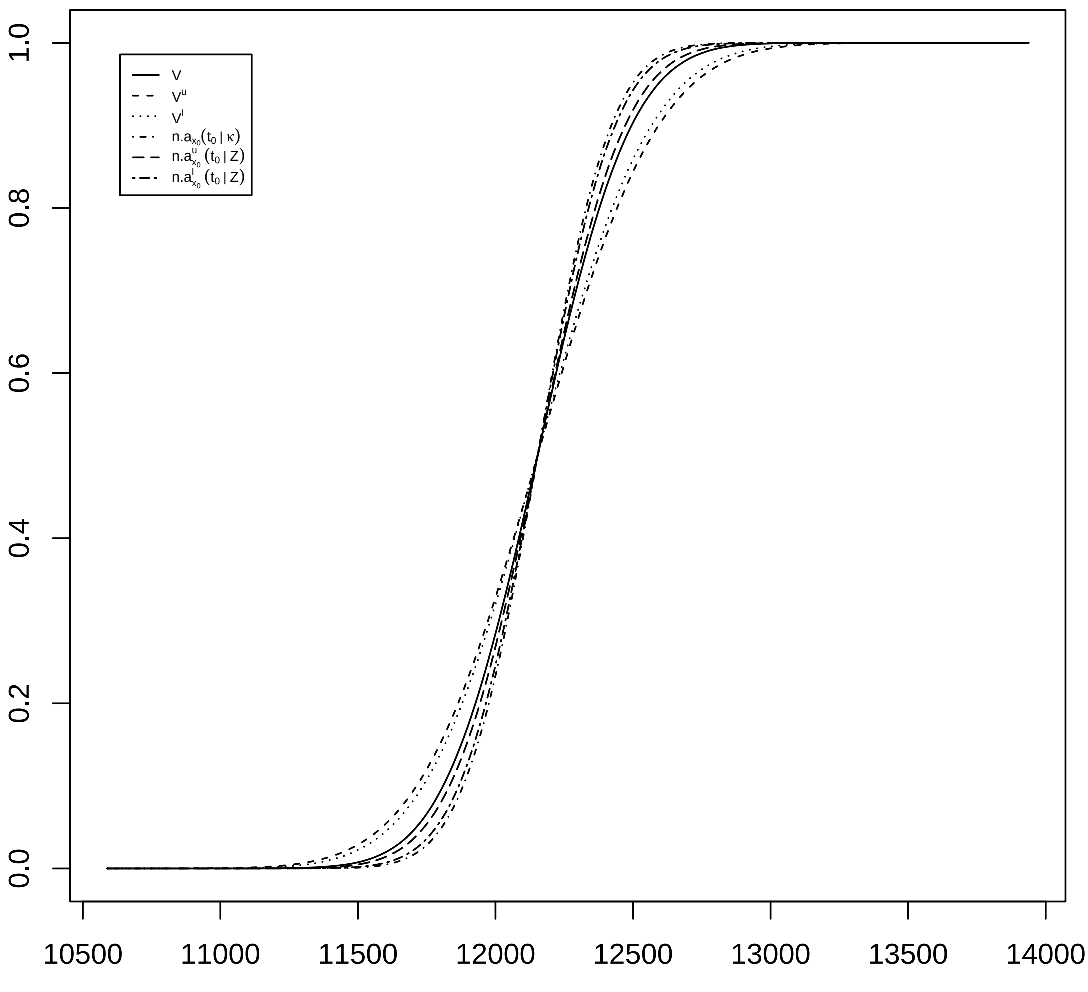

Figure 4.2: Distribution function of V (—), of V^u (- - -), of V^l $(\cdots)$ and of the large-portfolio approximations $n \times a_x(t_0)$ $(\cdot - \cdot - \cdot)$, $n \times a_x^u(t_0)$ $(---)$ and $n \times a_x^l(t_0)$ $(\cdot - \cdot - \cdot)$ for $n = 1000$.

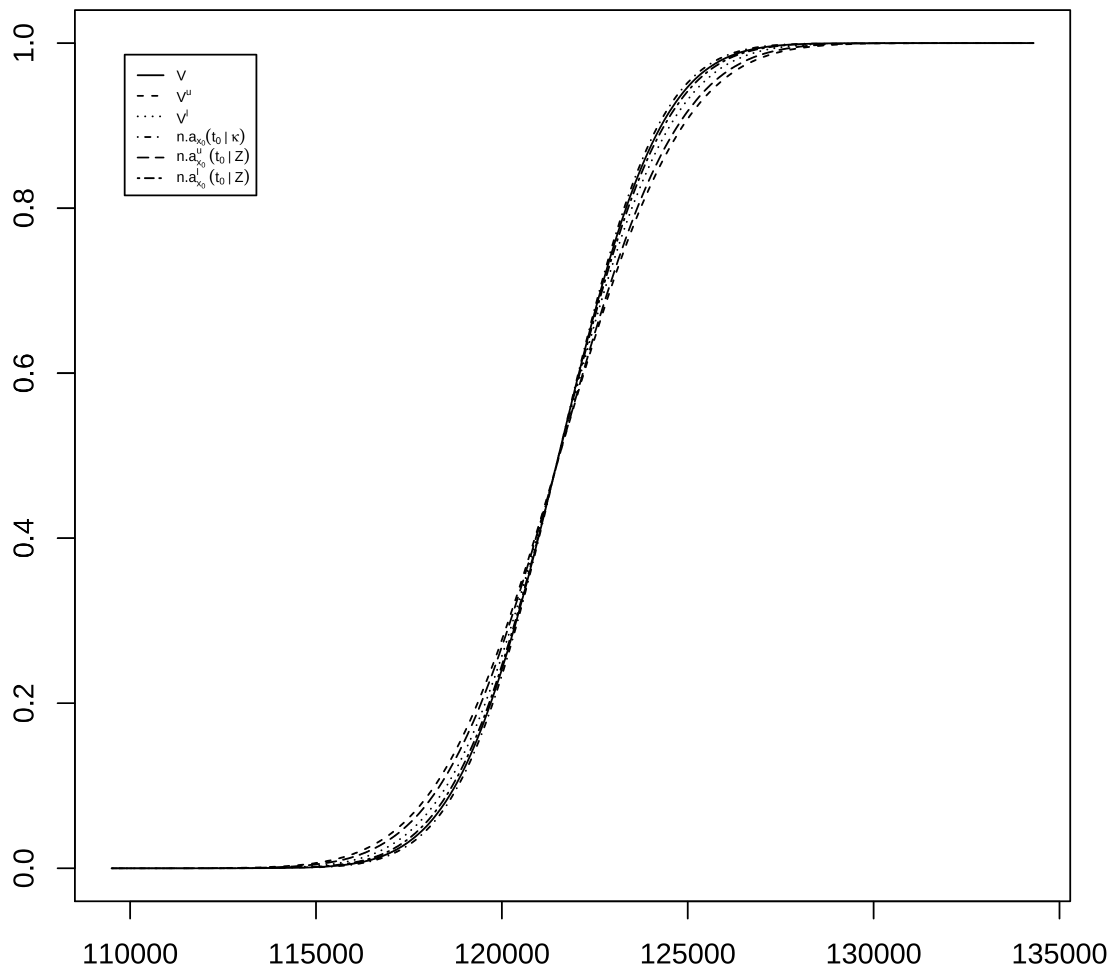

Figure 4.3: Distribution function of V (—), of V^u (- - -), of V^l $(\cdots)$ and of the large-portfolio approximations $n \times a_x(t_0)$ $(\cdot - \cdot - \cdot)$, $n \times a_x^u(t_0)$ $(---)$ and $n \times a_x^l(t_0)$ $(\cdot - \cdot - \cdot)$ for $n = 10000$.

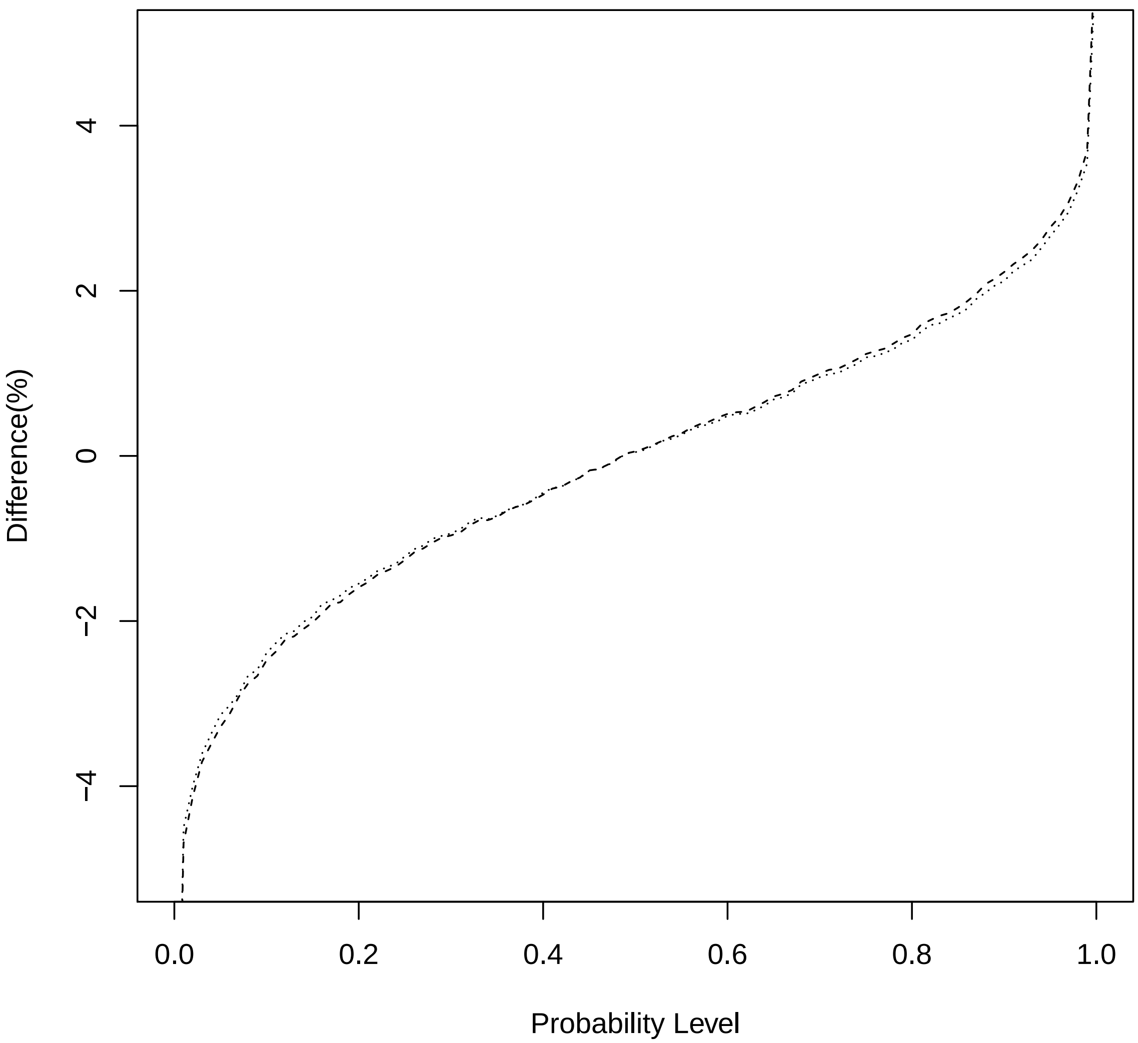

Figure 4.4: Relative differences between the quantiles of the approximations V^u (- - -) , V^l $(\cdots)$ and the quantiles of V for $n = 100$.

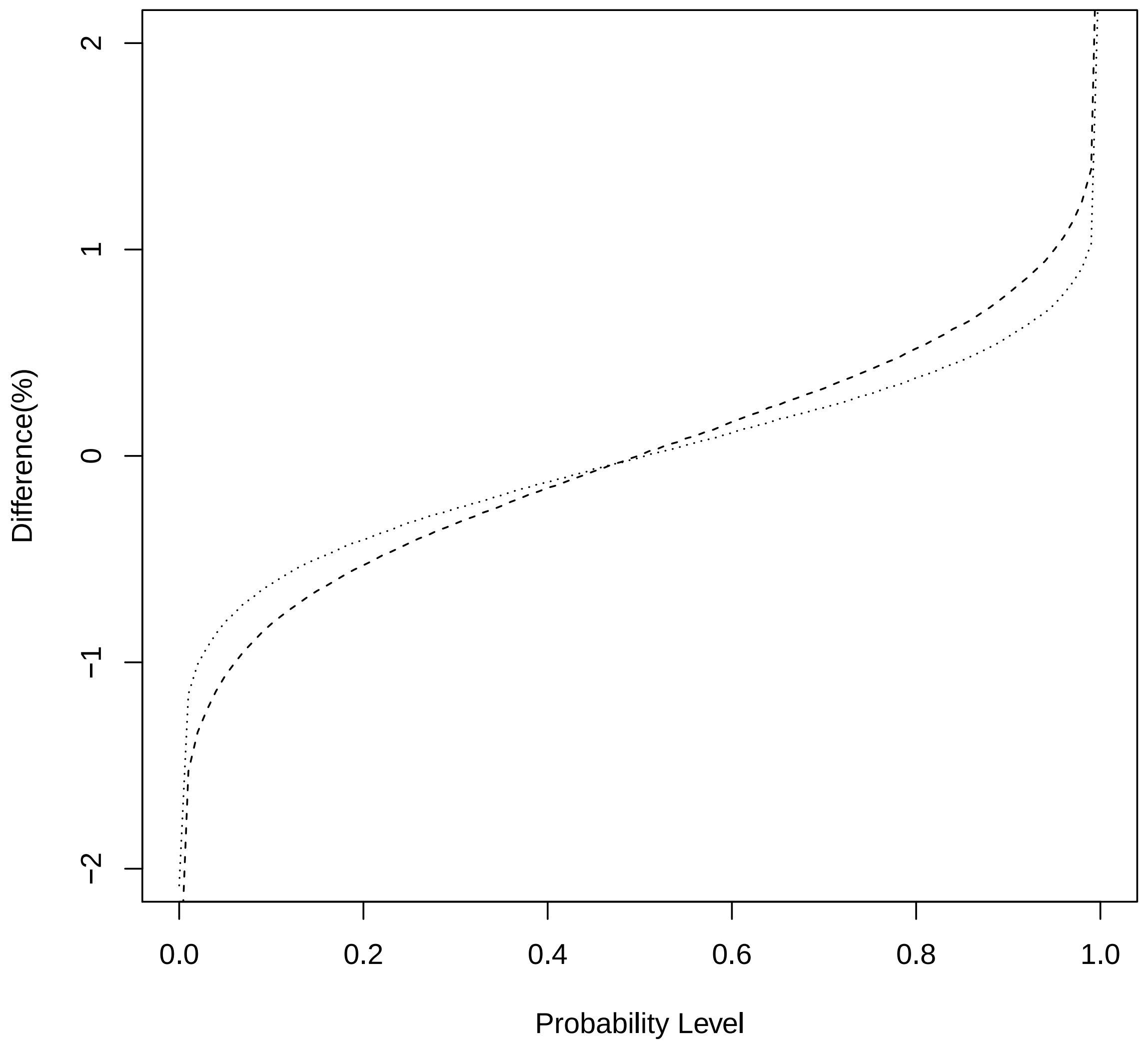

Figure 4.5: Relative differences between the quantiles of the approximations V^u (- - -) , V^l $(\cdots)$ and the quantiles of V for $n = 1000$.

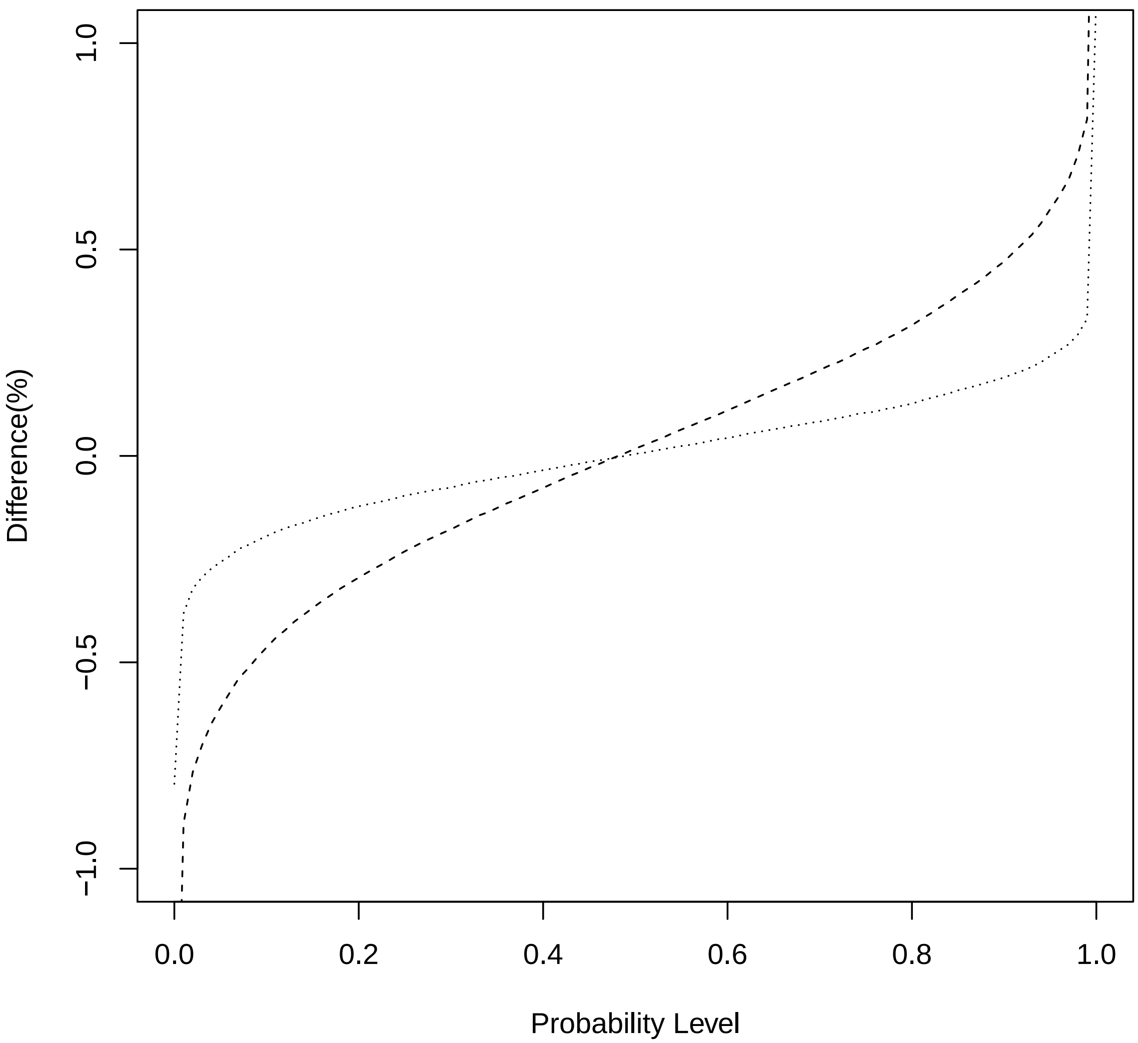

Figure 4.6: Relative differences between the quantiles of the approximations V^u (- - -) , V^l $(\cdots)$ and the quantiles of V for $n = 10000$.

of future benefits in a life annuity portfolio are then derived. This random variable plays a key role in the pricing and reserving process so that the approximations developed here can help insurers to speed their simulation routines in the Solvency 2 calculations.

Dickson and Waters (1999) and Sundt (1999, 2000) have derived algorithms for the calculation of the joint distribution of aggregate claims from a life insurance portfolio over several time periods. The approach developed in the present paper can be seen as an alternative to these recursive algorithms. By means of appropriate comonotonic approximations, we replace the vector of annual aggregate claim amounts with a random vector with a much simpler structure, whose joint distribution is easily derived.

Acknowledgements

The authors acknowledge the financial support from the contract "Projet d'Actions de Recherche Concertées" No 12/17-045 of the "Communauté française de Belgique", granted by the "Académie universitaire Louvain". Michel Denuit gratefully thanks the UNIL "Chaire Pensions et Longévité" financed by Retraites Populaires, directed by Professor François Dufresne.

References

Balakrishnan, N., Belzunce, F., Sordo, M.A., Suarez-Llorens, A. (2012). Increasing directionally convex orderings of random vectors having the same copula, and their use in comparing ordered data. Journal of Multivariate Analysis 105, 45-54.

Denuit, M. (2008). Comonotonic approximations to quantiles of life annuity conditional expected present values. Insurance: Mathematics and Economics 42, 831-838.

Denuit, M., Dhaene, J. (2007). Comonotonic bounds on the survival probabilities in the Lee-Carter model for mortality projections. Computational and Applied Mathematics 203, 169-176.

Denuit, M., Dhaene, J., Goovaerts, M.J., Kaas, R. (2005). Actuarial Theory for Dependent Risks: Measures, Orders and Models. Wiley, New York.

Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R., Vyncke, D. (2002a). The concept of comonotonicity in actuarial science and finance: Theory. Insurance: Mathematics and Economics 31, 3-33.

Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R., Vyncke, D. (2002b). The concept of comonotonicity in actuarial science and finance: Applications. Insurance: Mathematics and Economics 31, 133-161.

Denuit, M., Haberman, S., Renshaw, A. (2010). Comonotonic approximations to quantiles of life annuity conditional expected present values: extensions to general ARIMA models and comparison with the bootstrap. ASTIN Bulletin 40, 331-349.

Denuit, M., Mesfioui, M. (2013). Multivariate higher-degree stochastic increasing convexity. ISBA Discussion Paper - 2013/16.

Dickson, D.C.M, Waters, H.R. (1999). Multi-period aggregate loss distributions for a life portfolio. ASTIN Bulletin 29, 295-309.

Donnelly, C. (2011). Quantifying mortality risk in small defined-benefit pension schemes. Scandinavian Actuarial Journal 2011, 1-17.

Hoedemakers, T., Darkiewicz, G., Goovaerts, M.J. (2005). Approximations for life annuity contracts in a stochastic financial environment. Insurance: Mathematics and Economics 37, 239-269.

Lee, R.D., Carter, L. (1992). Modelling and forecasting the time series of US mortality. Journal of the American Statistical Association 87, 659-671.

Sundt, B. (1999). Discussion on D.C.M Dickson and H.R. Waters. ASTIN Bulletin 29, 311-314.

Sundt, B. (2000). On error bounds for approximations to multivariate distributions. Insurance: Mathematics and Economics 27, 137-144.