Review of Business and Economic Literature
July-December (2018)
ISSN 2031-1761

Upper and Lower Bounds of Present Value Distributions of Life Insurance Contracts with Disability Related Benefits

by J. Spreeuw

Jaap Spreeuw
Cass Business School, Faculty of Actuarial Science and Statistics, London, United Kingdom

ABSTRACT

The distribution function of the present value of a cash flow can be approximated by means of a distribution function of a random variable, which is also the present value of a sequence of payments, but has a simpler structure. The corresponding random variable has the same expectation as the random variable corresponding to the original distribution function and is a stochastic upper bound of convex order. A sharper upper bound and a nontrivial lower bound can be obtained if more information about the risk is available. In this paper, it will be shown that such an approach can be adopted for some life insurance contracts under Markov assumptions, with disability related benefits. The quality of the approximation will be investigated by comparing the distribution obtained with the one derived from the algorithm presented in the paper by Hesselager and Norberg (1996).

JEL Classification code: C65.
SIBC Classification Code: IM10, IM30, IB81, IB83.
Keywords: Convex order, comonotonic joint distribution, multistate life insurance contracts, present value distributions.

I. INTRODUCTION

In life contingencies under a stochastic framework, distributions of the present value of future payments are a key component in order to derive premiums satisfying a certain criterion. The most usual principle is that of actuarial equivalence, meaning that premiums are such that the expected present value of benefits less premiums is equal to zero.

The probability distribution of present values gives an indication of the riskiness of a contract, such as the variability of the actual benefits paid out or the upper tail of the present values.

De Pril (1989) and Dhaene (1990) derive distributions of such present values in a classical life insurance framework where only the two states "Alive" and "Dead" are relevant (and where only payment by single premium is considered). As the present value depends on the outcome of only one random variable, such analyses turn out to be quite straightforward.

Deriving distributions such as those described above in general tends to be much more complicated if instead we are dealing with life contracts involving more than two states. Hesselager and Norberg (1996) show that the distribution for a multistate life insurance contract under the Markov assumption can be obtained numerically by deriving a set of integral equations. However, an explicit expression for the distribution function in terms of transition probabilities and transition intensities does not exist. The advantage of expressions is that alternative premium principles can be applied for pricing a contract. For instance, the annual premium should be determined such that the probability of a loss incurred by the insurer is smaller than 40%. Moreover, expressions could facilitate sensitivity analysis. As an example, the impact of a change in interest rates on the shape of the distribution function could be analyzed. The question is whether a method exists to derive an approximate version of the real present value distribution, which is such an expression.

A common method is to replace a random variable by a "riskier" one, i.e. a random variable which is larger with respect to some ordering relation. The probability distribution of this "riskier" random variable has a simpler structure. Goovaerts et al. (2000) consider distributions of the present value of cash flows based on stochastic interest. They conclude that the comonotonic joint distribution (the distribution that is the largest in convex order) is often a good approximation of the original distribution. The latter can usually only be derived by means of simulation.

Kaas et al. (2000) show that the convex upper bound, as derived in Goovaerts et al. (2000), can be improved if more information is known about the present values of the individual cash flows. The additional information is represented by a conditioning variable. This approach also leads to a non-trivial lower bound which can sometimes be a good approximation.

This paper focuses on life insurance contracts catering for benefits payable during any spell of disability. We will deal with generalized disability annuity treaties, which include many types of contracts which are common in practice, such as income protection policies and deferred annuity contracts with a premium reduction when disabled. For such policies, we will derive the convex upper and lower bounds as considered in Goovaerts et al. (2000) and Kaas et al. (2000). The quality of the approximations by means of the convex upper and lower bound will be analyzed by comparing it with the present value distribution derived by means of the algorithm developed by Hesselager and Norberg (1996). The latter algorithm is executed in such a way that the present value distribution thus obtained is very close to the real one.

The paper is organized as follows. A general life insurance contract is described in Section II. A summary of the algorithm developed in Hesselager and Norberg (1996), based on the general life insurance contract, is presented in Section III. We will deal with some special life insurance contracts providing disability-related benefits and other benefits. These will be explained in Section IV. If the Markov chain is hierarchical, and the number of states is limited, methods alternative to Hesselager and Norberg are available. These will be described in Section V. In Section VI the original convex upper bound, as in Goovaerts et al. (2000), is derived. In Section VII, we condition on the remaining future lifetime and hence derive an improved upper bound and a non-trivial lower bound. Section VIII gives some examples for illustrative purposes. Section IX sets out a conclusion.

II. A MULTI-STATE LIFE INSURANCE CONTRACT

The following assumptions are the same as in Hesselager and Norberg (1996) and Norberg (1995), except for the specification of the force of interest. In Hesselager and Norberg (1996) and Norberg (1995), it is state-dependent, but given the state, it is constant. In this paper, the force of interest is taken to be deterministic but it may vary over time.

Consider a set $\zeta = \{0, \ldots, J\}$ of all possible states of a general life policy, such that at any time $t \in [0, m]$, the policy is in one and only one state. The state of the policy at time t is denoted by $X(t)$. The stochastic process is taken to be right continuous with $X(0) = 0$, implying that the policy is in state 0 at time 0, being the time at issue. Let $I_j(t)$ be the indicator of the event that the contract is in state j at time t, and

$$N_{jk}(t) = \#\{\tau \in (0, t] \mid X(\tau-) = j, X(\tau) = k\} \tag{1}$$

as the total number of transitions of $\{X(t)\}_{t\geq 0}$ from state j to state k ($\neq j$) by time t. The payment function B denotes the total amount of benefits paid in the time interval $[0, t]$. It is assumed to be continuous from the right. It is specified as

$$dB(t) = \sum_{j} I_j(t-)\,\mathrm{d}B_j(t) + \sum_{j\neq k} b_{jk}(t)\,dN_{jk}(t) \tag{2}$$

where each B_j is a deterministic payment function specifying payments due during sojourns in state j (a general life annuity) and each b_{jk} is a deterministic function, specifying the payments due upon transition from state j to state k. The left limit in $I_j(t-)$ means that the state annuity is effective at time t if the policy is in state j just prior to (but not necessarily equal at) time t. Consistent with this, we define $I_0(0-) = 1$.

We assume that each B_j decomposes into an absolutely continuous part and a discrete part as

$$dB_j(t) = b_j(t)\,dt + \Delta B_j(t). \tag{3}$$

Premiums are counted as negative benefits.

It is assumed that $\{X(t)\}_{t\geq 0}$ is a Markov chain. Denote the transition probabilities by

$$p_{jk}(t, u) = P[X(u) = k \mid X(t) = j]. \tag{4}$$

The transition intensities

$$\mu_{jk}(t) = \lim_{h\downarrow 0} \frac{P_{jk}(t, t+h)}{h} \tag{5}$$

are assumed to exist for all $j, k \in \zeta, j \neq k$. The total intensity of transition out of state j is $\mu_{j\zeta}(t) = \sum_{k;k\neq j} \mu_{jk}(t)$. The probability of staying uninterruptedly in state j during the time interval from t to u is $e^{-\int_t^u \mu_{j\zeta}(s)ds}$.

We assume $\delta(t)$, the force of interest at time t, to be a deterministic and continuous function. We introduce the following discount function:

$$v(\tau) = e^{-\int_0^t \delta(s)ds}. \tag{6}$$

Our aim is to derive the distribution function of the random present value, which will be specified by S. So

$$S = \int_0^m v(\tau)\, dB(\tau). \tag{7}$$

III. THE METHOD BY HESSELAGER AND NORBERG

The formulae displayed in this section are adopted from Hesselager and Norberg (1996), bearing in mind the different specification of the interest. The difference in the formulae concerns the force of interest which in their paper is allowed to be state-dependent and fixed, conditionally given the state. In this paper, on the other hand, the force of interest is taken to be deterministic and independent of the state of the policyholder, though not necessarily constant as a function of time. Hesselager and Norberg (1996) introduce the state-dependent probability functions

$$P_j(t, u) = P\left[\int_t^m \frac{v(\tau)}{v(t)} dB(\tau) \leqslant u \middle| I_j(t) = 1\right], \; t \in [0, m],\; u \in \mathbb{R},\; j \in \zeta. \tag{8}$$

So $P_j(t, u)$ denotes the probability that at time t, given that the contract is then in state j, the present value of future payments, discounted to t, is smaller than or equal to u.

The starting point in their analysis is the recursive equation

$$P_j(t, u) = \sum_{k;k\neq j} \int_t^m e^{-\int_t^s \mu_{j\zeta}(t)dt} \mu_{jk}(s)\, ds \cdot P_k\left(s, \frac{v(t)}{v(s)}u - \int_t^s \frac{v(\tau)}{v(s)} dB_j(\tau) - b_{jk}(s)\right)$$
$$+ e^{-\int_t^m \mu_{j\zeta}(s)ds}\, I\left[\int_t^m \frac{v(\tau)}{v(t)} dB_j(t) \leqslant u\right]. \tag{9}$$

Applying the auxiliary function

$$Q_j\,(t, u) = P_j\left(t, v(t)^{-1}\left(u - \int_0^t v(t)\, dB_j(t)\right)\right) \tag{10}$$

and substituting this in (9) yields

$$e^{-\int_0^t \mu_{j\zeta}(s)ds}\, Q_j\,(t, u)$$

$$= \int_t^m e^{-\int_0^s \mu_{j\zeta}(\tau)d\tau} \sum_{k,k\neq j} \mu_{jk}(s)ds \cdot Q_k\left(s, u + \int_0^s v(\tau)d\big(B_k(\tau) - B_j(t)\big) - e^{-ds}b_{jk}(s)\right)$$

$$+\, e^{-\int_0^m \mu_{j\zeta}(s)ds}\, I\left[\int_0^m v(\tau)dB_j(\tau) \leqslant u\right]. \tag{11}$$

By differentiating with respect to t and rearranging, it follows that the functions $Q_j\,(t, u)$ in (11) are the unique solutions to the differential equations

$$dQ_j\,(t, u) = \mu_{j\zeta}\,(t)\, dt \cdot Q_j\,(t, u)$$
$$- \sum_{k;k\neq j} \mu_{jk}\,(t)dt \cdot Q_k\left(t, u + \int_0^t v(\tau)d\big(B_k(\tau) - B_j(\tau)\big) - e^{-\delta t}b_{jk}(t)\right) \tag{12}$$

subject to the constraint

$$Q_j\,(m, u) = I\left[\int_0^m v(t)dB_j(\tau) \leqslant u\right]. \tag{13}$$

The computational scheme follows by taking the finite difference version of the above equation:

$$Q_j^*\,(t-h, u) = \left(1 - \mu_{j\zeta}\,(t)h\right) \cdot Q_j^*\,(t, u)$$
$$+\, h \sum_{k;k\neq j} \cdot\, Q_k^*\left(t, u + \int_0^t v(\tau)d\big(B_k(\tau) - B_j(\tau)\big) - e^{-\delta t}b_{jk}(t)\right). \tag{14}$$

Then starting from (13) (with Q_j^* in the place of Q_j) one calculates first the functions $Q_j^*\,(m-h, \cdot)$ by (14) and continues recursively until $Q_j^*\,(0, \cdot)$ finally can be calculated. In Hesselager and Norberg (1996), the functions $Q_j^*\,(t, u)$ are defined for $t \in \{0, h, 2h, \ldots, m\}$ and $u \in \{a, a + h', a + 2h', \ldots, b\}$, where h and h' are certain step-lengths.

This method is the only general approach to tackle the problem of determining the present value distributions that has appeared in the literature until now.

IV. DESCRIPTION OF THE BENEFITS

In this paper, only three states apply, namely "Active", "Disabled" and "Dead", denoted by a, i, and d, respectively.

We distinguish between the following types of benefits

- Annuity type benefits which are payable only if the life is alive, and which depend on whether the life is Active or Disabled. In the remainder of this paper, these benefits are specified as the *disability related benefits*. Such benefits only apply during the first n years of the contract.
- Benefits which are payable while alive or on death and which make no distinction between the states "Active" and "Disabled". Such benefits are labelled as *other benefits*. These benefits comprise annuity type benefits, which are payable after the first n years of the contract, and lump sum benefits, which may be payable throughout the entire term m of the contract.

Next, we give a more thorough description of these two categories of benefits.

A. *Disability related benefits*

We assume that disability related benefits only apply during the first n years of the contract, where $n \geq 0$. Benefits are payable during sojourns in states "Active" or "Disabled". We disregard from payments due upon transition from the "Active" state to the "Disabled" state or vice versa.

Hence, the disability related benefits are mathematically specified as:

$$\begin{array}{ll} I_a(t-)\, dB_a(t) + I_i(t-)\, dB_i(t) & \text{for} \quad t \in [0, n] \\ 0 & \text{for} \quad t > n. \end{array} \tag{15}$$

Benefits related to remaining in the "Active" state are nonpositive:

$$dB_a(t) \leq 0 \quad \forall t \in [0, n]. \tag{16}$$

Benefits payable during a sojourn in the "Disabled" state are greater than or equal to the benefits related to "Active":

$$dB_i(t) \geq dB_a(t). \tag{17}$$

Recall that premiums count as negative benefits. Then inequality (17) means that a disabled life either:

- pays premiums, which are smaller than or equal to the premiums an active life would have to pay, or
- receives annuity benefits, or
- neither pays nor receives anything.

We consider the following special cases:

Case 1 *The benefits payable during a sojourn in the "Disabled" state are exactly the same as the benefits payable during a sojourn in the "Active" state:* $dB_i(t) = dB_a(t)$. *In this case, we are effectively working in a classical two state environment, comprising the states "Alive" and "Dead".*

Case 2 *The premiums payable during sojourn in the "Disabled" state are smaller than or equal to the benefits payable during sojourn in the "Active" state:* $0 \geq dB_i(t) \geq dB_a(t)$. *Then we are dealing with a premium reduction in case of disability.*

Case 3 *The benefits (or premiums) payable during sojourn in the "Disabled" state are 0:* $dB_i(t) = 0$. *This involves a premium waiver in case of disability.*

Case 4 *The benefits payable during sojourn in the "Disabled" state are nonnegative:* $dB_i(t) \geq 0$. *In such a case, we are considering a disability annuity or income protection policy.*

B. *Other benefits*

We distinguish between survival benefits and death benefits. Survival benefits are payable only after time n and are payable during a sojourn in either one of the states "Active" and "Disabled". The level of the benefits is independent of whether the individual is active or disabled. We will specify the total amount of such benefits over $[n, t]$ by $B_\ell(t)$, nondecreasing in t. Hence, $dB_\ell(t) \geq 0$, and the contribution to the payment function is $\left(I_a(t-) + I_i(t-)\right) dB_\ell(t)$.

Death benefits are comprised of lump sum benefits and annuity type benefits payable on entering the “Dead” state. The level of the lump sum benefits depends on the time of death only. We assume that it does not matter whether the life enters the state “Dead” as an active or as a disabled person. We will specify such benefits as $b_{\cdot d}(t)$, so $b_{\cdot d}(t) = b_{ad}(t) = b_{id}(t) \quad \forall t \geq 0$. The benefits payable during the stay in the state “Dead” are considered to be part of the lump sum benefit.

Hence, the total payment function reduces to:

$$\begin{aligned} dB(t) = {} & b_{\cdot d}(t)\left(dN_{ad}(t) + dN_{id}(t)\right) \\ & + I_{\{t \leq n\}}\left(I_a(t-)\, dB_a(t) + I_i(t-)\, dB_i(t)\right) \\ & + I_{\{t > n\}}\left(I_a(t-) + I_i(t-)\right) dB_\ell(t) \end{aligned} \tag{18}$$

We assume that, for $t > n$, $\mu_{ad}(t) = \mu_{id}(t)$.

Remark 5 *As after time n, no disability related benefits are payable, such an assumption is common in practice. A common additional assumption is that both $\mu_{ad}(t)$ and $\mu_{id}(t)$ are equal to $\mu(t)$, the standard force of mortality. It will turn out that the above assumption will make some equations easier to derive.*

In the next three sections, we will deal with alternative methods to those of Hesselager and Norberg (1996) of deriving the distribution of present values. First of all, we show that deriving such a distribution is a quite straightforward exercise if it is not possible to recover from sickness. Thereafter, we will allow for recovery from illness and present a way to derive, for some disability annuity contracts, an approximation to the c.d.f. of V, defined in (7).

V. THE PROBABILITY DISTRIBUTION OF THE PRESENT VALUE IF RECOVERY FROM DISABILITY IS IMPOSSIBLE

The Markov chain is hierarchical if an individual cannot recover once disabled. The situation is displayed graphically in Figure 1.

The next two examples show how to derive the distribution of the present value. The first example concerns the case of only “Disability related benefits”. The second example extends the first by incorporating “Other benefits”.

Example 6 *Consider a disability annuity with a term of n years, equal to bi per annum, payable continuously while the individual is disabled, and a benefit of ba p.a., payable continuously while the life is in the "Active" state. The contract is issued at time 0 in the "Active" state. During periods in the "Active" state, the individual pays a premium equal to c per annum, continuously. So ba = –c, and the $B_j(\tau)$ in the right hand side of (2) are equal to*

$$B_d(t) = 0 \;\; \forall t, \;\; B_a(t) = \begin{cases} 0 & t < 0 \\ ba \cdot t & 0 \leq t < n \\ ba \cdot n & t \geq n \end{cases}, \;\; B_i(t) = \begin{cases} 0 & t < 0 \\ bi \cdot t & 0 \leq t < n. \\ bi \cdot n & t \geq n \end{cases} \tag{19}$$

Furthermore, we assume $\delta(\cdot)$ to be constant and equal to δ. Hence, $v(t) = e^{-\delta t}$. For an active individual, the probability distribution of the present value at issue proves to be

$$P_a(0, u) = \begin{cases} 0 & \text{for} \quad u < G(n); \\ p_{aa}(0, n) & \text{for} \quad u = G(n); \\ \begin{aligned} & p_{aa}(0, n) \\ & + p_{aa}(0, t_1(u)) \int_{\tau=t_1(u)}^{n} p_{aa}(t_1(u), \tau)\, \mu_{ad}(\tau) d\tau \\ & + p_{aa}(0, t_2(u, n))\, p_{ai}(t_2(u, n), n) \\ & + \int_{\tau=t_1(u)}^{n} p_{aa}(0, t_2(u, \tau))\, p_{ai}(t_2(u, \tau), \tau)\, \mu_{id}(\tau) d\tau \end{aligned} & \text{for} \quad G(n) < u < 0; \\ \begin{aligned} & p_{aa}(0, n) + \int_{\tau=0}^{n} p_{aa}(0, \tau)\, \mu_{ad}(\tau) d\tau \\ & + p_{aa}(0, t_2(u, n))\, p_{ai}(t_2(u, n), n) \\ & + \int_{\tau=t_1(u)}^{n} p_{aa}(0, t_2(u, \tau))\, p_{ai}(t_2(u, \tau), \tau)\, \mu_{id}(\tau) d\tau \end{aligned} & \text{for} \quad 0 \leq u < H(n); \\ 1 & \text{for} \quad u \geq H(n) \end{cases} \tag{20}$$

with

$$G(n) = \int_{s=0}^{n} ba\, e^{-\delta s} ds; \; H(n) = \int_{s=0}^{n} bi\, e^{-\delta s}\, ds$$

$$t_1(u) = -\frac{\ln\left[1 - \frac{\delta u}{ba}\right]}{\delta}; \left(t_1(u)\text{: solution to } r \text{ of } u = \int_{s=0}^{r} ba\, e^{-\delta s} ds\right)$$

$$t_2(u, t) = -\frac{\ln\left[\frac{\delta u + bi\, e^{-\delta\tau} - ba}{bi - ba}\right]}{\delta};$$

$$t_2(u, \tau)\text{: solution to } r \text{ of } u = \int_{s=0}^{r} ba\, e^{-\delta s} ds + \int_{s=r}^{\tau} bi\, e^{-\delta s} ds. \tag{21}$$

We give a short explanation:

1. *The present value can never be lower than* $G(n)$. *This lower bound is attained if the contract remains in the state "Active" throughout the term of the contract. This happens with probability*

$$p_{aa}(0, n) = e^{-\int_{s=0}^{n} \mu_{a\zeta}(s)ds}. \tag{22}$$

2. $H(n)$ *is the maximum possible present value, corresponding to the event that the individual gets disabled immediately after issue and is still disabled at the end of the contract period;*

3. *For* $G(n) < u < H(n)$, *we distinguish between the following four mutually exclusive events:*
   (a) *the individual is still active on maturity of the contract;*
   (b) *the individual will be active until at least time* $t_1(u)$, *and die, while active, before n;*
   (c) *the individual will be active until at least time* $t_2(u, n)$, *get disabled before n, and survive to n.*
   (d) *The individual will die from "Disabled" at some time* τ *before n. Given death at time* τ, *the individual will be active until at least time* $t_2(u, \tau)$, *and get disabled before* τ.

Remark 7 *A simpler example, based on a single premium (rather than level premium) payment, can be found in Spreeuw (2000).*

We have just dealt with a relatively simple contract. Some complications could arise in case where other benefits are to be included. In such cases, it could be convenient to work with the conditioning random variable as the following example shows:

Example 8 *Consider the same contract as the one considered in Example 6, but with a constant death benefit, denoted by bd, included. Adopting the same approach as in Example 6 would be complicated. As an intermediate step, however, we could derive the conditional distribution, given the value of a certain random variable. We will choose the remaining lifetime (or time on death) of the individual as the conditioning random variable. We denote this variable by T and condition on the event that* $T = t$. *Then for* $0 \leq t \leq n$ *(death during the term of the contract), we have the following* *conditional distribution of the present value of future benefits less premiums, given T = t (denoted by P_a (0, u|T = t):*

$$P_a(0, u|T=t) = \begin{cases} 0 & \textit{for} \quad u < G(t) + bd; \\ \int_{\tau=t_1(u)}^{t} \frac{p_{aa}(0,t_1(u))\, p_{aa}(t_1(u),\tau)\, \mu_{ad}(t)dt}{p_{aa}(0,t)\, \mu_{ad}(t)dt + p_{ai}(0,t)\, \mu_{id}(t)dt} d\tau \\ + \int_{\tau=t_2(u,t)}^{t} \frac{p_{aa}(0,\tau)\, p_{ai}(\tau,t)\, \mu_{id}(t)dt}{p_{aa}(0,t)\, \mu_{ad}(t)dt + p_{ai}(0,t)\, \mu_{id}(t)dt} d\tau & \textit{for} \quad G(t) + bd \le u < bd \\ \int_{\tau=0}^{t} \frac{p_{aa}(0,t_1(u))\, p_{aa}(t_1(u),\tau)\, \mu_{ad}(t)dt}{p_{aa}(0,t)\, \mu_{ad}(t)dt + p_{ai}(0,t)\, \mu_{id}(t)dt} d\tau \\ + \int_{\tau=t_2(u,t)}^{t} \frac{p_{aa}(0,\tau)\, p_{ai}(\tau,t)\, \mu_{id}(t)dt}{p_{aa}(0,t)\, \mu_{ad}(t)dt + p_{ai}(0,t)\, \mu_{id}(t)dt} d\tau & \textit{for} \quad bd \le u < H(t) + bd; \\ 1 & \textit{for} \quad u \ge H(t) + bd. \end{cases} \tag{23}$$

Then P_a (0, u) is derived by integrating P_a (0, u|T = t) over t.

Remark 9 *In Section VII, we will apply the method of conditioning on a random variable. Just as in this example, we will choose a random variable resembling the remaining lifetime.*

The above example serves to illustrate that there are cases where one does not need to be restricted to the method by Hesselager and Norberg. The derivation of the above c.d.f. is quite straightforward because the present value depends only on two variables: the time of getting disabled and the time of death. This is due to the fact that:

1. There are only a few number of states applying, and that
2. The Markov chain is hierarchical: the state “Disabled”, once left, cannot be revisited.

Of course, deriving the present value distribution is quite complicated if there are many states involved, even if the Markov chain is hierarchical. An example is the multilife insurance contract, is given by Hesselager and Norberg (1996).

The aim of this paper, however, is to study the consequences of dropping the second assumption. If recovery is possible during the term, an insured can be disabled for more than one time interval. This makes calculations such as those above much more complicated.

Our method of dealing with problems such as those above involves replacing the random variable of present value by one being larger in convex order. If X and Y are random variables, X precedes Y in convex order (notation $X \leq_c Y$) if $\mathrm{E}[f(X)] \leqslant \mathrm{E}[f(Y)]$ for each convex function f. This method has already been applied for distributions of the present value of cash flows based on stochastic interest, see Goovaerts et al. (2000). In Kaas et al. (2000) it is shown that an improved upper bound and a nontrivial lower bound can be obtained by conditioning on a random variable. Illustrations in that paper involve the lognormal discounting process, amongst others. Vyncke et al. (2001), applying the technique in Kaas et al. (2000), centers on Vasicek and Ho-Lee models as applications.

The next section deals with the theory on the convex upper bound, as explained in Goovaerts et al. (2000), and an application to a disability annuity. The section thereafter considers the theory and applications in Kaas et al. (2000).

We will now assume that it is possible to recover from disability. Hence, Figure 2 applies.

VI. DERIVING THE ORIGINAL CONVEX UPPER BOUND

This section starts with an explanation of the theory in Goovaerts et al. (2000) in Subsection VI.A. Thereafter, in Subsection VI.B, we adapt the definition of the total present value to our needs, specifying the points of time of the several benefits. In Subsection VI.C, we show how the convex upper bound can be derived for a disability annuity.

A. *Approach by Goovaerts et al. (2000)*

Goovaerts et al. (2000) specify the present value random variable, again denoted by S, as a sum of r.v.'s, each of them involving a certain point of time:

$$\mathrm{S} = \sum_{k=1}^{M} Y_k, \tag{24}$$

with

$$Y_k = \beta_k \, e^{-X_{t_k}}, \; \beta_k \in \mathbb{R}. \tag{25}$$

where X_{t_k} represents the force of interest integrated from 0 to t_k, so is actually the stochastic discounting factor of a payment at time t_k. Furthermore, β_k is the payment itself. When deriving the expression for the probability distribution of S, one needs to realize that the X_{t_k}'s are not mutually independent. As a consequence, the variables Y_k are not independent of one another either.

The authors first derive the distribution, which, within the class of random vectors $(Y_1, \ldots, Y_M)$ with fixed marginals (such a class is called a Fréchet class) is the comonotonic one. This distribution is the largest in convex order. If $F_1, \ldots, F_M$ are the c.d.f.'s of the respective r.v.'s $Y_1, \ldots, Y_M$, this comonotonic joint distribution of $Y_1, \ldots, Y_M$ is equal to the distribution of the random vector

$$\left(F_1^{-1}(U), \ldots, F_M^{-1}(U)\right), \tag{26}$$

where U ~Uniform(0, 1) and $F_k^{-1}(u)$, $k \in \{1, \ldots, M\}$, is defined by

$$F_k^{-1}(u) = \min\left\{x \mid F_k(x) \geqslant u\right\} \tag{27}$$

The r.v. which is the sum of the components in (26) is denoted by S^c, and:

$$S \leq_c S^c. \tag{28}$$

Let $Y_k^c = F_k^{-1}(U)$, $k \in \{1, \ldots, M\}$, so $S^c = \sum_{k=1}^{M} Y_k^c$. Then, the joint c.d.f. of $Y_1^c, \ldots, Y_M^c$ is known to be

$$\Pr\left[Y_1^c \leq y_1, \ldots, Y_M^c \leq y_M\right] = \min_{k \in \{1, \ldots, M\}} F_k(y_k). \tag{29}$$

Let $F_{S^c}(\cdot)$ be the c.d.f. of S^c and $F_{S^c}^{-1}(\cdot)$ its inverse, the latter defined in the same way as in (27). The c.d.f. of S^c follows implicitly from the relationship

$$F_{S^c}^{-1}(u) = \sum_{k=1}^{M} F_k^{-1}(u), \qquad u \in [0, 1]. \tag{30}$$

This is the so called *convex upper bound*. The quality of the approximation by means of this c.d.f. can be analyzed by comparing it with the joint distribution obtained, e.g. by means of Monte Carlo simulation.

In Subsection VI.C, we will derive the convex upper bound in (30) for some disability annuity contracts. These contracts are a special case of the type of policies we have specified in Section IV, since there are only disability related benefits and no other benefits involved.

Before turning to these examples, we need to change the definition of the random variable S, as defined in (24). The reason is that Goovaerts et al. (2000) is based on the financial risk while this paper is centred on the demographic risks. We will do this in the next subsection.

B. *Specification of total present value*

In the remainder of this paper it is assumed that the interval $[0, n)$ (the interval involving the disability related benefits) can be partitioned into subsequent subintervals $[t_0, t_1)$, $[t_1, t_2)$, …, $[t_{M_1-2}, t_{M_1-1})$, $[t_{M_1-1}, t_{M_1})$, while the interval $[n, m)$ can be partitioned into subsequent subintervals $[t_{M_1}, t_{M_1+1})$, $[t_{M_1+1}, t_{M_1+2})$, …, $[t_{M-1}, t_M)$. The partitioning is such that in the period $[0, n)$, there may only be benefit payments at the M_1 different durations $t_1, t_2, \ldots, t_{M_1-1}, t_{M_1}$, while in the period $[n, m)$, there may only be benefit payments at the $M - M_1$ different durations $t_{M_1+1}, t_{M_1+2}, \ldots t_{M-1}, t_M$. If death happens in the period $[t_{k-1}, t_k)$, the benefit will be paid at time t_k.

Remark 10 *In practice, often premiums are payable upon issue, at duration* 0*. Without loss of generality, we will exclude such (negative) benefits from the calculation of present values. This means that single premiums do not count as part of the present value.*

As a consequence, we can write S in (7) as a sum of random variables:

$$S = \sum_{k=1}^{M} Y_k, \tag{31}$$

with

$$\begin{aligned} Y_k = {} & g_a(t_k)\, I_a(t_k) + g_i(t_k)\, I_i(t_k) \\ & + g_{\cdot d}(t_k)\left(N_{ad}(t_k) - N_{ad}(t_{k-1}) + N_{id}(t_k) - N_{id}(t_{k-1})\right), \end{aligned} \tag{32}$$

where

$$g_j(t_k) = v(t_k)\left(B_j(t_k) - B_j(t_{k-1})\right), \; j \in \{a, d\}, \quad k \in \{1, \ldots, M_1\} \tag{33}$$

$$g_\ell(t_k) = v(t_k)\left(B_\ell(t_k) - B_\ell(t_{k-1})\right), \quad j \in \{a, d\}, k \in \{M_1 + 1, \ldots, M\}, \tag{34}$$

and

$$g_{\cdot d}\,(t_k) = v\,(t_k)\,b_{\cdot d}\,(t_k), \quad j \in \{a, \mathrm{d}\}, k \in \{1, \ldots, M\}. \qquad (35)$$

In (33), $g_j\,(t_k)$ denotes the present value of the benefit paid at time t_k in case of remaining in state j, and $g_{\cdot d}\,(t_k)$ denotes the present value of the benefit paid at time t_k in case of death in $(t_{k-1}, t_k]$.

Remark 11 *The fully continuous case arises as a special case if we let $M \to \infty$ and besides, for each $k \in \{0, \ldots, M-1\}$, $t_{k+1} - t_k \to 0$.*

So S can be decomposed in M separate random variables where the Y_k denotes the stochastic present value of benefits due to remaining in a certain state at time t_k, $k \in 1, \ldots, M$. We have

$$\begin{aligned}
&\Pr\,[S = z] \\
&= \Pr\left[Y_1 = g_{\cdot d}\,(t_1)\right] \cdot I_{\{g_{\cdot d}(t_1)=z\}} + \sum_{\substack{j_1 \in \{a,i\} \\ g_{j_1}(t_1)+g_{\cdot d}(t_2)=z}} \Pr\left[Y_1 = g_{j_1}(t_1),\, Y_2 = g_{\cdot d}\,(t_2)\right] \\
&\quad + \ldots \\
&\quad + \sum_{\substack{(j_1,\ldots,j_{M_1-1})^t \in \{a,i\}^{M_1-1}; \\ \sum_{k=1}^{M_1-1} g_{j_k}(t_k)+g_{\cdot d}(t_{M_1+1})=z}} \Pr\left[Y_1 = g_{j_1}(t_1),\, \ldots, Y_{M_1-1} = g_{jM_1-1}\left(t_{M_1-1}\right),\, Y_{M_1} = g_{\cdot d}\left(t_{M_1}\right)\right] \\
&\quad + \sum_{\substack{(j_1,\ldots,j_{M_1})^t \in \{a,i\}^{M_1}; \\ \sum_{k=1}^{M_1} g_{j_k}(t_k)+g_{\cdot d}(t_{M_1+1})=z}} \Pr\begin{bmatrix} Y_1 = g_{j_1}(t_1),\, \ldots,\, Y_{M_1} = g_{jM_1}\left(t_{M_1}\right), \\ Y_{M_1+1} = g_{\cdot d}\left(t_{M_1+1}\right) \end{bmatrix} \\
&\quad + \sum_{\substack{(j_1,\ldots,j_{M_1})^t \in \{a,i\}^{M_1}; \\ \sum_{k=1}^{M_1} g_{j_k}(t_k)+g_{\ell}(t_{M_1+1})+g_{\cdot d}(t_{M_1+2})=z}} \Pr\begin{bmatrix} Y_1 = g_{j_1}(t_1),\, \ldots,\, Y_{M_1} = g_{jM_1}\left(t_{M_1}\right), \\ Y_{M_1+1} = g_{\ell}\left(t_{M_1+1}\right),\, Y_{M_1+2} = g_{\cdot d}\left(t_{M_1+2}\right) \end{bmatrix} \\
&\quad + \ldots \\
&\quad + \sum_{\substack{(j_1,\ldots,j_{M_1})^t \in \{a,i\}^{M_1}; \\ \sum_{k=1}^{M_1} g_{j_k}(t_k)+g_{\cdot d}(t_M) \\ +\sum_{k=1}^{M-M_1-1} g_{\ell}(t_{M_1+k})=z}} \Pr\begin{bmatrix} Y_1 = g_{j_1}(t_1),\, \ldots,\, Y_{M_1} = g_{jM_1}\left(t_{M_1}\right), \\ Y_{M_1+1} = g_{\ell}\left(t_{M_1+1}\right),\, \ldots,\, Y_{M-1} = g_{\ell}(t_{M-1}),\, Y_M = g_{\cdot d}(t_M) \end{bmatrix} \\
&\quad + \sum_{\substack{(j_1,\ldots,j_{M_1})^t \in \{a,i\}^{M_1}; \\ \sum_{k=1}^{M_1} g_{j_k}(t_k) \\ +\sum_{k=1}^{M-M_1} g_{\ell}(t_{M_1+k})=z}} \Pr\begin{bmatrix} Y_1 = g_{j_1}(t_1),\, \ldots,\, Y_{M_1} = g_{jM_1}\left(t_{M_1}\right), \\ Y_{M_1+1} = g_{\ell}\left(t_{M_1+1}\right),\, \ldots,\, Y_{M-1} = g_{\ell}(t_{M-1}),\, Y_M = g_{\ell}(t_M) \end{bmatrix}. \qquad (36)
\end{aligned}$$

Note that the above expression consists of $M+1$ terms. The first term of this equation involves the event that the insured will die between 0 and t_1, the second refers to death between t_1 and t_2, the Mth to survival to t_{M_1-1} and death between t_{M_1-1} and n ($= t_{M_1}$), the ($M_1 + 1$)th to survival to n ($= t_{M_1}$) and death between n ($= t_{M_1}$) and t_{M_1+1}, the ($M_1 + 2$)th to survival to n and death between t_{M_1+1} and t_{M_1+2}, the second last term concerns death between t_{M-1} and t_M and the last term concerns survival to $m = t_M$.

In terms of states, we get the following equality:

$$
\begin{aligned}
&\Pr[S = z] \\
&= \Pr[X(1) = d] \cdot I_{\{g_{\cdot d}(t_1)=z\}} + \sum_{\substack{j_1 \in \{a,i\} \\ g_{j_1}(t_1)+g_{\cdot d}(t_2)=z}} \Pr\left[X(1) = j_1, X(2) = d\right] \\
&+ \dots \\
&+ \sum_{\substack{(j_1,\dots,j_{M_1-1})^t \in \{a,i\}^{M_1-1}; \\ \sum_{k=1}^{M_1-1} g_{j_k}(t_k)+g_{\cdot d}(t_{M_1})=z}} \Pr\left[X(1) = j_1, \dots, X(M_1-1) = j_{M_1-1}, X(M_1) = d\right] \\
&+ \sum_{\substack{(j_1,\dots,j_{M_1})^t \in \{a,i\}^{M_1}; \\ \sum_{k=1}^{M_1} g_{j_k}(t_k)+g_{\cdot d}(t_{M_1+1})=z}} \Pr\begin{bmatrix} X(1) = j_1, \dots, X(M_1) = j_{M_1}, \\ X(M_1+1) = d \end{bmatrix} \\
&+ \sum_{\substack{(j_1,\dots,j_{M_1})^t \in \{a,i\}^{M_1}; \\ \sum_{k=1}^{M_1} g_{j_k}(t_k)+g_{\ell}(t_{M_1+1})+g_{\cdot d}(t_{M_1+2})=z}} \Pr\begin{bmatrix} X(1) = j_1, \dots, X(M_1) = j_{M_1}, \\ X(M_1+1) \in \{a, i\}, X(M_1+2) = d \end{bmatrix} \\
&+ \sum_{\substack{(j_1,\dots,j_{M_1})^t \in \{a,i\}^{M_1}; \\ \sum_{k=1}^{M} g_{j_k}(t_k)+g_{\cdot d}(t_M) \\ \sum_{k=1}^{M-M_1-1} g_{\ell}(t_{M_1+k})=z}} \Pr\begin{bmatrix} X(1) = j_1, \dots, X(M_1) = j_{M_1}, \\ X(M_1+1) \in \{a, i\}, \dots, X(M-1) \in \{a, i\}, X(M) = d \end{bmatrix} \\
&+ \dots + \sum_{\substack{(j_1,\dots,j_{M_1})^t \in \{a,i\}^{M_1}; \\ \sum_{k=1}^{M_1} g_{j_k}(t_k)+\sum_{k=1}^{M-M_1} g_{\ell}(t_{M_1+k})=z}} \Pr\begin{bmatrix} X(1) = j_1, \dots, X(M_1) = j_{M_1}, \\ X(M_1+1) \in \{a, i\}, \dots, X(M) \in \{a, i\} \end{bmatrix}.
\end{aligned}
\tag{37}
$$

We define F_k as the c.d.f. of Y_k, $k \in \{1, \dots, M\}$. Furthermore, we define, just as above, $Y_c^k = F_k^{-1}(U)$, $k \in \{1, \dots, M\}$, with $U \sim$Uniform (0, 1). So $S^c = \sum_{k=1}^{M} Y_k^c$ and

$$F_k(y_k) = \Pr[Y_k \leq y_k] = \Pr[Y_k^c \leq y_k], \quad k \in \{1, \dots, M\}. \quad (38)$$

Note that these marginals have a support consisting of a finite number of points.

C. *The convex upper bound for a disability annuity*

These contracts are a special case of the type of policies we have specified in Section IV, since there are only disability related benefits and no other benefits. Hence the term is restricted to n years, $M_1 = M$ and $t_M = n$. We will start with the simplest premium payment arrangement (payment by single premium) and then proceed to a more complicated type (level premium payment). Our aim in this subsection is to derive the expression for the convex upper bound of the present value. In the numerical examples in Section VIII, which is based on these annuities, the quality of the approximation will be judged by comparing the c.d.f. of the random variable with an accurate approximation of the c.d.f. of S. An accurate approximation of S is obtained by applying the algorithm of Hesselager and Norberg with small values for the step-lengths h and h', as defined in Section III.

In both cases, there are no payments due if the individual is dead. So we have:

$$g_d(t_k) = B_d(t_k) = 0;$$

$$g_i(t_k) = v(t_k)(B_i(t_k) - B_i(t_{k-1})) \geq 0, \quad \forall k \in \{1, \dots, M\}. \quad (39)$$

The only difference between the two policies is the method of premium payment. In Subsubsection VI.C.1 we deal with payment by single premium, while Subsubsection VI.C.2 considers the case where premiums are paid as long as the individual is active. In both subsubsections we will treat the fully continuous annuity (where all benefits are paid on a continuous basis) as a special case.

We will conclude this subsection with Subsubsection VI.C.3, containing some remarks concerning the calculation of the transition probabilities $p_{aa}(0, \cdot)$, $p_{ai}(0, \cdot)$ and $p_{ad}(0, \cdot)$.

1. Payment by single premium

Recall that, by assumption, present values do not include single premium payments, so:

$$B_a(t_k) = g_a(t_k) = 0 \quad \forall k \in \{1, \ldots, M\}. \tag{40}$$

In the given case, the marginals $F_k(y_k)$ in formula (29) are specified as:

$$F_k(y_k) = \begin{cases} 0 & \text{for } y_k < 0 \\ 1 - p_{ai}(0, t_k) & \text{for } 0 \le y_k < g_i(t_k), \\ 1 & \text{for } y_k \ge g_i(t_k) \end{cases} \quad k \in \{1, \ldots, \mathrm{M}\}. \tag{41}$$

In practice, disability annuities are often contracts valid for the period that an individual is not retired. As a consequence, for n not too large (otherwise the death rates will dominate), $p_{ai}(0, t)$ is usually an increasing function of t for $t \in [0, n]$. This leads to the following theorem:

Theorem 12 *If* $\frac{dp_{ai}(0,t)}{dt} \ge 0$ *for* $t \in (0, n]$, *the comonotonic joint distribution of the present value corresponding to a disability annuity contract, paid at single premium, with the payment scheme as specified in (39) and (40) is:*

$$\Pr[S^c \le s] = \begin{cases} 0 & \textit{for} \quad s < 0; \\ 1 - p_{ai}(0, t_w) & \textit{for} \quad \sum_{k=w+1}^{M} g_i(t_k) \le s < \sum_{k=w}^{M} g_i(t_k), \quad w \in \{1, \ldots, M\}; \\ 1 & \textit{for} \quad s \ge \sum_{k=1}^{M} g_i(t_k). \end{cases} \tag{42}$$

Proof. Note that, for any $y_w < 0$ with $w \in \{1, \ldots, M\}$:

$$F_{Y_1^c, \ldots, Y_M^c}(y_1, \ldots, y_M) = F_{Y_1^c, \ldots, Y_M^c}(0, \ldots, 0, y_w, 0, \ldots, 0) = 0. \tag{43}$$

In addition, if $p_{ai}(0, t)$ is increasing in $t \in [0, n]$ we have for $y_1, \ldots, y_{w-1} \ge 0$ and $0 \le y_w < g_i(t_w)$, $w \in \{2, \ldots, M\}$, that

$$F_{Y_1^c, \ldots, Y_M^c}(y_1, \ldots, y_M) = F_{Y_1^c, \ldots, Y_M^c}(0, \ldots, 0, y_{w+1}, \ldots, y_M). \tag{44}$$

The above equality states that if an individual is not disabled at time t_w, he cannot be disabled before that time either. This is equivalent to saying that if an individual is disabled at time t_w, he will remain so with certainty till the expiration of the contract. This implies that the present value only depends on the time at which the contract enters the state "Disabled". This proves the theorem.

Note that the Markov chain corresponding to the distribution of S^c is hierarchical. The chain is even more rigorous than the ordinary hierarchical Markov chain considered in Section V: a disabled individual can *neither recover nor die*. In other words: compared to the ordinary hierarchical Markov chain, the state "Disabled" is absorbing and not strongly transient.

On the other hand, since the marginals are fixed, the probabilities of *getting* disabled are lower and the death rates for an *active* person are *higher*.

The fully continuous version of (42) leading to a benefit payment of $dB_i\,(t)$ if the contract is in state i at time t, with $t \in (0, n]$ (so $dB_i\,(t) > 0$ on the same interval $t \in (0, n]$) is obtained by letting $M \to \infty$ and besides for all $i \in \{0, \ldots, \mathrm{M} - 1\}$, $t_{i+1} - t_i \to 0$ (cf. Remark 11). The result is:

$$\Pr\,[S^c \leq s] = \begin{cases} 0 & \text{for} \quad s < 0; \\ 1 - p_{ai}\,(0, t) & \text{for} \quad s = \int_t^n g_i(\tau)d\tau, \qquad t \in [0, n]; \\ 1 & \text{for} \quad s > \int_0^n g_i(\tau)d\tau. \end{cases} \tag{45}$$

The next equations serve to illustrate the transition intensities corresponding to the comonotonic joint distribution. These are obtained from the forward differential equations of Chapman-Kolmogorov. For notational convenience, they are accompanied by an asterisk superscript (*). The original transition intensities, corresponding to the joint distribution of S, are given between brackets.

$$\mu^*_{ia}\,(t) = 0$$
$$\left(\mu_{ia}\,(t) = \frac{1}{p_{ai}\,(0, t)} \left(\frac{dp_{aa}\,(0, t)}{dt} + p_{aa}\,(0, t)\,(\mu_{ai}\,(t) + \mu_{ad}\,(t))\right)\right); \tag{46}$$

$$\mu^*_{id}\,(t) = 0$$
$$\left(\mu_{id}\,(t) = \frac{1}{p_{ai}\,(0, t)} \left(\frac{dp_{ad}\,(0, t)}{dt} - p_{aa}\,(0, t)\,\mu_{ad}\,(t)\right)\right); \tag{47}$$

$$\mu^*_{ai}\,(\mathrm{t}) = \frac{1}{p_{aa}\,(0, t)}\,\frac{dp_{ai}\,(0, t)}{dt}$$
$$\left(\mu_{ai}\,(t) = \frac{1}{p_{aa}\,(0, t)} \left(\frac{dp_{ai}\,(0, t)}{dt} + p_{ai}\,(0, t)\,(\mu_{id}\,(t) + \mu_{ia}\,(t))\right)\right); \tag{48}$$

$$\mu^*_{ad}(t) = \frac{1}{p_{aa}(0,t)} \frac{dp_{ad}(0,t)}{dt}$$

$$\left(\mu_{ad}(t) = \frac{1}{p_{aa}(0,t)} \left(\frac{dp_{ad}(0,t)}{dt} - p_{ai}(0,t)\,\mu_{id}(t)\right)\right). \tag{49}$$

Next, we will consider annuity treaties where premiums are paid while the contract is in the "Active" state. We will treat this topic in the same way as above.

2. Premium payment in the "Active" state

In this case, the premiums discounted to time upon issue are:

$$g_a(t_k) = v(t_k)\,(B_a(t_k) - B_a(t_{k-1})) \leq 0, \quad \forall k \in \{1, \ldots, M\}. \tag{50}$$

The marginals $F_k(y_k)$ in equation (29) are, as one might expect, more complicated than in the single premium case:

$$F_k(y_k) = \begin{cases} 0 & \text{for} \quad y_k < g_a(t_k); \\ 1 - p_{ai}(0,t_k) - p_{ad}(0,t_k) & \text{for} \quad g_a(t_k) \leq y_k < 0; \\ 1 - p_{ai}(0,t_k) & \text{for} \quad 0 \leq y_k < g_i(t_k); \\ 1 & \text{for} \quad y_k \geq g_i(t_k). \end{cases}, \quad k \in \{1, \ldots, M\}. \tag{51}$$

We again assume that $p_{ai}(0,t)$ is an increasing function of t for $t \in (0, n]$, leading to the following theorem:

Theorem 13 *If* $\frac{dp_{ai}(0,t)}{dt} \geq 0$, *the comonotonic joint distribution of the present value corresponding to a disability annuity contract, with the payment scheme as specified in (39) and (50) is:*

$$\Pr[S^c \leq s] = \begin{cases} 0 & \textit{for} \quad s < \sum_{k=1}^{M} g_a(t_k); \\ 1 - p_{ai}(0,t_r) - p_{ad}(0,t_w) & \textit{for} \quad s \in [h_d(w), h_d(w-1)) \cap [h_i(r), h_i(r-1)), \\ & \qquad w \in \{1, \ldots, M\}, \quad r \in \{r_{\min}, \ldots, M\}; \\ 1 - p_{ai}(0,t_r) & \textit{for} \quad s \in [\max[h_i(r), 0], h_i(r-1)), \; r \in \{1, \ldots, r_{\min}\}; \\ 1 & \textit{for} \quad s \geq \sum_{k=1}^{M} g_i(t_h). \end{cases}$$

(52)

In the above formulae

$$h_d(w) = \sum_{k=1}^{w} g_a(t_k);$$

$$h_i(r) = \sum_{k=r+1}^{M} g_i(t_k) + \sum_{k=1}^{r} g_a(t_k); \tag{53}$$

$$r_{\min} = \max\left[r \in \{1, \ldots, M-1\} \mid \sum_{k=r+1}^{M} g_i(t_k) + \sum_{k=1}^{r} g_a(t_k) < 0 \right]$$

Proof. Note that, for any $y_w < g_a(t_w)$ with $w \in \{1, \ldots, M\}$:

$$F_{Y_1^c,\ldots,Y_M^c}(y_1, \ldots, y_M) = F_{Y_1^c,\ldots,Y_M^c}(0, \ldots, 0, y_w, 0, \ldots, 0) = 0. \tag{54}$$

Furthermore, if $p_{ai}(0, t)$ is increasing in $t \in [0, n]$ we have the following:

1. If $y_j \geq g_a(t_j)$ and $g_a(t_w) \leq y_w < 0$ for $j \in \{1, \ldots, w-1\}$ and $w \in \{2, \ldots, M\}$ then

$$F_{Y_1^c,\ldots,Y_M^c}(y_1, \ldots, y_w, \ldots, y_M) = F_{Y_1^c,\ldots,Y_M^c}(g_a(t_1), \ldots, g_a(t_w), y_{w+1}, \ldots, y_M), \tag{55}$$

implying that if an individual is active at a certain time, he is active all the time before. The consequence is that a disabled individual cannot recover.

2. For $y_1, \ldots, y_{w-1} \geq 0$ and $0 \leq y_w < g_i(t_w)$, $w \in \{2, \ldots, M\}$, that

$$F_{Y_1^c,\ldots,Y_M^c}(y_1, \ldots, y_M) = F_{Y_1^c,\ldots,Y_M^c}(0, \ldots, 0, y_{w+1}, \ldots, y_M). \tag{56}$$

The above equality states that if an individual is not disabled at time t_w, he cannot be disabled before that time either. This is equivalent to saying that if an individual is disabled at time t_w, he will remain so with certainty till the expiration of the contract. This implies that the present value only depends on the time at which the contract enters the state "Disabled" or the time at which the contract enters the state "Dead". Note that this part of the proof is equivalent to the proof of Theorem 12.

This proves the theorem.

The fully continuous version of (52) leading to a benefit payment of $dB_i(t)$ if the contract is in state i and a premium payment of $-dB_a(t)$ if the contract is in state a at time t, with $t \in (0, n]$ (so $dB_i(t) > 0$ and $dB_a(t) < 0$ on the same interval $t \in (0, n]$), is, just as in the case of single premium payment, obtained by letting $M \to \infty$ and besides for each $i \in \{0, \ldots, M-1\}$, $t_{i+1} - t_i \to 0$ (cf. Remark 11). The result is:

$$\Pr[S^c \leq s] = \begin{cases} 0 & \text{for} \quad s < \int_0^n g_a(\tau)d\tau; \\ 1 - p_{ai}(0, h_i(s)) - p_{ad}(0, h_a(s)) & \text{for} \quad \int_0^n g_a(\tau)d\tau \leq s < 0; \\ 1 - p_{ai}(0, h_i(s)) & \text{for} \quad 0 \leq s < \int_0^n g_i(\tau)d\tau; \\ 1 & \text{for} \quad s \geq \int_0^n g_i(\tau)d\tau. \end{cases} \tag{57}$$

In the above formula $h_i(s)$ and $h_a(s)$ are the solutions of t in the equalities

$$\int_{\ell}^{n} g_i(\tau)d\tau + \int_{0}^{t} g_a(\tau)d\tau = s \tag{58}$$

and

$$\int_{0}^{t} g_a(s)ds = s, \tag{59}$$

respectively.

The numerical example in Section VIII about disability annuities will deal with the continuous case, and will therefore be based on equation (57).

3. Calculation of transition probabilities

Note that up to now we assumed that the transition probabilities were known. In fact, they need to be computed first by solving a system of forward differential equations of Chapman-Kolmogorov. It turns out that the method, in order to be computationally efficient, requires that the numerical calculation of the transition probabilities does not take too much time. With the availability of modern software packages (such as Mathematica) and powerful computers, this is likely to be the case, except when the number of states is very high. One can also assume that only a limited number of transitions can take place in a certain period of time. For instance, both in the AIDS model and in the disability annuity model used in the Netherlands (cf. Alting von Geusau (1990), and Gregorius (1993) respectively) it is assumed that there can be at most one transition per year.

In one of the numerical examples of Section VIII concerned with disability annuities, it transpires that the approximation resulting from the convex upper bound is very good. In the case of a disability annuity, the convex upper bound is quite easy to derive, given some properties of the transition probabilities. However, if we add other benefits, the convex upper bound may become complicated to derive. As shown in Kaas et al. (2000), improved upper bounds and non-trivial lower bounds can be obtained by conditioning on a random variable. This is what we do in the next section.

VII. CONDITIONING ON A RANDOM VARIABLE

We will first of all, in Subsection VII.A, consider the approach in Kaas et al. (2000). Then, in Subsection VII.B we will choose our conditioning random variable.

A. *Deriving an improved upper bound and non-trivial lower bound*

In Kaas et al. (2000), it is shown that an improved upper bound an a non-trivial lower bound can be obtained if more information is known about the distribution of S, through a conditioning random variable. In what follows, we will denote this random variable by Λ, with c.d.f. $F_\Lambda(\lambda)$, which is assumed to be known. We assume that $F_\Lambda(\lambda)$ decomposes into an absolutely continuous and a discrete part, such that:

$$dF_\Lambda(\lambda) = f_\Lambda(\lambda)\, d\lambda + (\Pr[\Lambda \leq \lambda] - \Pr[\Lambda \leq \lambda-]). \qquad (60)$$

It is assumed that, for all $i \in \{1, \ldots, M\}$ and any outcome λ of Λ, the conditional distributions of Y_i, given $\Lambda = \lambda$, denoted by $F_{i|\Lambda}(\cdot)$, are known.

1. Improved upper bound

Denote $F^{-1}_{i|\Lambda}(U)$ for the random variable $f_i(\mathrm{U}, \Lambda)$, where the function f_i is defined by $f_i(\mathrm{u}, \lambda) = F^{-1}_{i|\Lambda=\lambda}(u)$.

Then the improved upper bound, denoted by S^u is known to be:

$$S^u = \sum_{i=1}^{M} F^{-1}_{i|\Lambda}(U). \qquad (61)$$

The following convex order applies:

$$S \leq_{cx} S^u \leq_{cx} S^c. \tag{62}$$

Since, given the event $\Lambda = \lambda$, the conditional random variable S^u is a sum of comonotonic random variables, the distribution function of the unconditional random variable S^u is relatively straightforward to derive:

$$F^{-1}_{S^u|\Lambda=\lambda}(p) = \sum_{i=1}^{M} F^{-1}_{i|\Lambda=\lambda}(p), \qquad p \in (0, 1). \tag{63}$$

Given $\Lambda = \lambda$, the c.d.f. of the conditional S^u follows from

$$F_{S^u|\Lambda=\lambda}(s) = \sup \left[p \in (0, 1) \,\middle|\, \sum_{i=1}^{M} F^{-1}_{i|\Lambda=\lambda}(p) \leq s \right]. \tag{64}$$

The c.d.f. of the unconditional S^u is then derived by

$$F_{S^u}(x) = \int_{\lambda=-\infty}^{\infty} F_{S^u|\Lambda=\lambda}(x)\, dF_{\Lambda}(\lambda). \tag{65}$$

2. Non-trivial lower bound

The convex lower bound, denoted by S^l, is known to be

$$S^l = \sum_{i=1}^{M} E[Y_i|\Lambda]. \tag{66}$$

We have the following order in convexity:

$$E[S] \leq_{cx} S^l \leq_{cx} S. \tag{67}$$

The c.d.f. of the lower bound can be determined as follows:

$$F_{S^l}(x) = \int_{-\infty}^{\infty} I\left[\sum_{i=1}^{M} E[Y_i|\Lambda=\lambda] \leq x \right] dF_{\Lambda}(\lambda). \tag{68}$$

Dhaene et al. (2002a) point out that, if the conditioning random variable Λ is such that all random variables $E[Y_i|\Lambda]$ are monotone functions of Λ, then the lower bound S^l is a sum of M comonotonic random variables and the c.d.f. is relatively straightforward to derive.

B. *Choice of the conditioning random variable*

As pointed out in Dhaene et al. (2002b), the choice of the optimal conditioning random variable Λ is not a trivial exercise. It is true that, the more $F_{i|\Lambda}(\cdot)$ resembles S, the better S^l approximates S. But this property usually does not hold for S^u. Moreover, there is the restriction that Λ be such that $F_{i|\Lambda}(\cdot)$ is known.

In Section VI, we concluded that the original convex upper bound can be easily derived if only disability related benefits are involved, assuming that $p_{ai}(0, t)$ is non decreasing for $t \in [0, n]$. Complications arise if other benefits are involved. The stochastic nature of these benefits is governed by the complete remaining lifetime of the life involved.

For this reason, we choose the period of death as the conditioning random variable. Denote the complete remaining lifetime by T. Then Λ is defined as follows:

$$\Lambda = \begin{cases} i & \text{for} \quad T \in [t_i, t_{i+1}) \text{ and } i \in \{0, \ldots, M-1\} \\ M & \text{for} \quad T > t_M \end{cases}. \tag{69}$$

We consider the following special cases applying for $t_M \to \infty$:

Case 14 (curtate future lifetime) *If* $t_i = i$, $i \in \{1, \ldots, M\}$, *then* Λ *is the curtate future lifetime (usually denoted by K).*

Case 15 (complete future lifetime) *If* $M \to \infty$ *and, furthermore, for each* $i \in \{0, \ldots, M-1\}$, $t_{i+1} - t_i \to 0$, *then* Λ *is the complete future lifetime and we use the notation* $\Lambda = T$.

In this way, we have chosen a random variable that captures the "Other benefits". Next, we derive the c.d.f.'s of $F_{k|\Lambda=\lambda}(\cdot)$.

For $k = \lambda + 1$ (the period of death), we have that all the mass is concentrated at $g_{\cdot d}(t_{\lambda+1})$: $F_{\lambda+1|\Lambda=\lambda}(y_\lambda) = I_{\{y_\lambda \geq g_{\cdot d}(t_{\lambda+1})\}}$. For $k > \lambda + 1$ (after the year of death), $F_{k|\Lambda=\lambda}(\cdot) \equiv 0$.

For $k < \lambda + 1$, (before the year of death) we define the c.d.f. as $F^*_{k|\Lambda=\lambda}(\cdot)$. We need to distinguish between $k \in \{1,.., M_1\}$ (the period when "Disability related benefits" are payable) and $k \in \{M_1+1,.., M\}$ (the period when only "Other benefits" are payable).

For $k \in \{1,.., M_1\}$, we have:

$$F^*_{k|\Lambda=\lambda}(y_k) = \begin{cases} 0 & \text{for} \quad y_k < g_a(t_k) \\ \Pr\left[X(t_k) = a \,\middle|\, \begin{matrix} X(0) = a, (X(t_1) = a \vee X(t_1) = i), \\ X(t_{\lambda+1}) = d \end{matrix}\right] & \text{for} \quad g_a(t_k) \leq y_k < g_i(t_k) \\ 1 & \text{for} \quad y_k \geq g_i(t_k). \end{cases} \tag{70}$$

Note that $\Pr[X(t_k) = a | X(0) = a, (X(t_\lambda) = a \vee X(t_\lambda) = i), X(t_{\lambda+1}) = d]$ is the conditional probability that the contract is in state "Active" at duration t_k, given state "Active" upon inception and death between t_λ and $t_{\lambda+1}$. It can be expressed in terms of transition probabilities in the following way:

$$\begin{aligned} &\Pr[X(t_k) = a | X(0) = a, (X(t_\lambda) = a \vee X(t_\lambda) = i), X(t_{\lambda+1}) = d] \\ &= \frac{\Pr[X(t_k) = a, (X(t_\lambda) = a \vee X(t_\lambda) = i), X(t_{\lambda+1}) = d | X(0) = a]}{\Pr[(X(t_\lambda) = a \vee X(t_\lambda) = i), X(t_{\lambda+1}) = d | X(0) = a]} \\ &= \frac{\Pr[X(t_k) = a | X(0) = a] \Pr[(X(t_\lambda) = a \vee X(t_\lambda) = i), X(t_{\lambda+1}) = d | X(t_k) = a]}{\Pr[(X(t_\lambda) = a \vee X(t_\lambda) = i), X(t_{\lambda+1}) = d | X(0) = a]} \\ &= \frac{p_{aa}(0, t_k)(p_{ad}(t_k, t_{\lambda+1}) - p_{ad}(t_k, t_\lambda))}{p_{ad}(0, t_{\lambda+1}) - p_{ad}(0, t_\lambda)} \\ &= 1 - \frac{p_{ai}(0, t_k)(p_{id}(t_k, t_{\lambda+1}) - p_{id}(t_k, t_\lambda))}{p_{ad}(0, t_{\lambda+1}) - p_{ad}(0, t_\lambda)}. \end{aligned} \tag{71}$$

For $k \in \{M_1 + 1,.., M\}$, the mass is concentrated at $g_\ell(t_k)$: $F^*_{k|\Lambda=\lambda}(y_k) = I_{\{y_k \geq g_\ell(t_k)\}}$.

1. Improved upper bound

Theorem 16 *If, for each* $\lambda \in \{1,.., M_1 - 1\}$

$$\begin{aligned} p_{ai}(0, t_1)(p_{id}(t_1, t_{\lambda+1}) - p_{id}(t_1, t_\lambda)) \leq p_{ai}(0, t_2)(p_{id}(t_2, t_{\lambda+1}) - p_{id}(t_2, t_\lambda)) \leq \ldots \\ \leq p_{ai}(0, t_\lambda) p_{id}(t_\lambda, t_{\lambda+1}) \end{aligned} \tag{72}$$

then for $\lambda \in \{1,.., M_1 - 1\}$, *the distribution of* $F_{S^u|\Lambda=\lambda}(s)$ *is given by*

$$F_{S^u|\Lambda=\lambda}(s) = \begin{cases} 0 & \text{for} \quad s < \sum_{k=1}^{\lambda} g_a(t_k) + g_{\cdot d}(t_{\lambda+1}); \\ 1 - \frac{p_{ai}(0,t_w)(p_{id}(t_w,t_{\lambda+1}) - p_{id}(t_w,t_\lambda))}{p_{ad}(0,t_{\lambda+1}) - p_{ad}(0,t_\lambda)} & \text{for} \quad \begin{matrix} s \in [h_i(w) + g_{\cdot d}(t_{\lambda+1}), h_i(w-1) + g_{\cdot d}(t_{\lambda+1})) \\ w \in \{1, \ldots, \lambda\}; \end{matrix} \\ 1 & \text{for} \quad s \geq \sum_{k=1}^{\lambda} g_i(t_k) + g_{\cdot d}(t_{\lambda+1}). \end{cases} \tag{73}$$

In the above formula:

$$h_i(w) = \sum_{k=w+1}^{\lambda} g_i(t_k) + \sum_{k=1}^{w} g_a(t_k). \tag{74}$$

Proof. Note that, for any $y_w < g_a(t_w)$ with $w \in \{1, \ldots, \lambda\}$:

$$F_{Y_l^u,\ldots,Y_M^u|\Lambda=\lambda}(y_1, \ldots, y_\lambda, 0, \ldots, 0) = 0. \tag{75}$$

Furthermore, if inequality (72) holds, we have that if $y_j \geq g_a(t_j)$ and $g_a(t_w) \leq y_w < g_i(t_w)$ for $j \in \{1, \ldots, w-1\}$ and $w \in \{2, \ldots, \lambda\}$, and $y_{\lambda+1} \geq g_{\cdot d}(t_{\lambda+1})$, then

$$\begin{aligned} &F_{Y_1^u,\ldots,Y_M^u|\Lambda=\lambda}(y_1, \ldots, y_w, \ldots, y_\lambda, y_{\lambda+1}, 0, \ldots, 0) \\ &= F_{Y_1^u,\ldots,Y_M^u|\Lambda=\lambda}(g_a(t_1), \ldots, g_a(t_w), y_{w+1}, \ldots, y_\lambda, g_{\cdot d}(t_{\lambda+1}), 0, \ldots, 0), \end{aligned} \tag{76}$$

implying that if an individual is active at a certain time, he is active all the time before. The consequence is that a disabled individual cannot recover.

This proves the theorem.

Theorem 17 *If, just as in (72),*

$$\begin{aligned} p_{ai}(0, t_1)(p_{id}(t_1, t_{\lambda+1}) - p_{id}(t_1, t_\lambda)) &\leq p_{ai}(0, t_2)(p_{id}(t_2, t_{\lambda+1}) - p_{id}(t_2, t_\lambda)) \leq \ldots \\ &\leq p_{ai}(0, t_\lambda)\, p_{id}(t_\lambda, t_{\lambda+1}) \end{aligned} \tag{77}$$

then for $\lambda \in \{M_1,.., M-1\}$, *we have that the distribution of* $F_{S^u|\Lambda=\lambda}(s)$ *is given by*

$$F_{S^u|\Lambda=\lambda}(s) = \begin{cases} 0 & \text{for} \quad s < \sum_{k=1}^{M_1} g_a(t_k) + \sum_{k=M_1+1}^{\lambda} g_\ell(t_k) + g_{\cdot d}(t_{\lambda+1}); \\ 1 - \frac{p_{ai}(0,t_w)(p_{id}(t_1,t_{\lambda+1}) - p_{id}(t_1,t_\lambda))}{p_{ad}(0,t_{\lambda+1}) - p_{id}(0,t_\lambda)} & \text{for} \quad s \in \left[\begin{matrix} h_i(w) + \sum_{k=M_1+1}^{\lambda} g_\ell(t_k) + g_{\cdot d}(t_{\lambda+1}), \\ h_i(w-1) + \sum_{k=M_1+1}^{\lambda} g_\ell(t_k) + g_{\cdot d}(t_{\lambda+1}) \end{matrix} \right), \\ & \qquad w \in \{1, \ldots, M_1\}; \\ 1 & \text{for} \quad s \geq \sum_{k=1}^{M_1} g_i(t_k) + \sum_{k=M_1+1}^{\lambda} g_\ell(t_{\lambda+1}). \end{cases} \tag{78}$$

For $\lambda = M$, we have that the distribution of $F_{S^u|\Lambda=\lambda}(s)$ is given by:

$$F_{S^u|\Lambda=\lambda}(s) = \begin{cases} 0 & \text{for} \quad s < \sum_{k=1}^{M_1} g_a(t_k) + \sum_{k=M_1+1}^{\lambda} g_\ell(t_k); \\ 1 - \frac{p_{ai}(0,t_w)(p_{id}(t_1,t_{\lambda+1}) - p_{id}(t_1,t_\lambda))}{p_{ad}(0,t_{\lambda+1}) - p_{id}(0,t_\lambda)} & \text{for} \quad s \in \left[\begin{array}{c} h_i(w) + \sum_{k=M_1+1}^{M} g_\ell(t_k), \\ h_i(w-1) + \sum_{k=M_1+1}^{M} g_\ell(t_k) \end{array} \right), \\ & \qquad w \in \{1, \ldots, M_1\}; \\ 1 & \text{for} \quad s \geq \sum_{k=1}^{M_1} g_i(t_k) + \sum_{k=M_1+1}^{M} g_\ell(t_k). \end{cases} \tag{79}$$

In the above formula:

$$h_i(w) = \sum_{k=w+1}^{M_1} g_i(t_k) + \sum_{k=1}^{w} g_a(t_k). \tag{80}$$

Proof. Consider the case $\lambda \in \{M_1,.., M-1\}$ first. Note that, for any $y_w < g_a(t_w)$ with $w \in \{1, \ldots, M_1\}$:

$$F_{Y_1^u,\ldots,Y_M^u|\Lambda=\lambda}(y_1, \ldots, y_{M_1}, g_\ell(t_{M_1+1}), g_\ell(t_{M_1+2}), \ldots, g_\ell(t_\lambda), g_{\cdot d}(t_{\lambda+1}), 0,.., 0) = 0. \tag{81}$$

Furthermore, if inequality (77) holds, we have that if $y_j \geq g_a(t_j)$ and $g_a(t_w) \leq y_w < g_i(t_w)$ for $j \in \{1, \ldots, w-1\}$ and $w \in \{2, \ldots, M_1\}$, and $y_w \geq g_\ell(t_w)$ for $w \in \{M_1 + 1,.., \lambda\}$, and $y_{\lambda+1} \geq g_{\cdot d}(t_{\lambda+1})$, then

$$F_{Y_1^u,\ldots,Y_M^u|\Lambda=\lambda}(y_1, \ldots, y_w, \ldots, y_{M_1}, y_{M_1+1}, y_{M_1+2}, \ldots, y_\lambda, y_{\lambda+1}, 0, \ldots, 0)$$
$$= F_{Y_1^u,\ldots,Y_M^u|\Lambda=\lambda}\left(\begin{array}{c} g_a(t_1), \ldots, g_a(t_w), y_{w+1}, \ldots, y_{M_1}, \\ g_\ell(t_{M_1+1}), g_\ell(t_{M_1+2}), \ldots, g_\ell(t_\lambda), g_{\cdot d}(t_{\lambda+1}), 0, \ldots, 0 \end{array} \right), \tag{82}$$

implying that if an individual is active at a certain time before n, he is active all the time before. The consequence is that a disabled individual cannot recover.

Now consider the case $\lambda = M$. Note that, for any $y_w < g_a(t_w)$ with $w \in \{1, \ldots, M_1\}$:

$$F_{Y_1^u,\ldots,Y_M^u|\Lambda=\lambda}(y_1, \ldots, y_{M_1}, g_\ell(t_{M_1+1}), g_\ell(t_{M_1+2}), \ldots, g_\ell(t_M)) = 0. \tag{83}$$

Furthermore, if inequality (77) holds, we have that if $y_j \geq g_a(t_j)$ and $g_a(t_w) \leq y_w < g_i(t_w)$ for $j \in \{1, \ldots, w-1\}$ and $w \in \{2, \ldots, M_1\}$ and $y_w \geq g_\ell(t_w)$ for $w \in \{M_1+1, \ldots, M\}$, then

$$F_{Y_1^u,\dots,Y_M^u|\Lambda=\lambda}(y_1, \dots, y_w, \dots, y_{M_1}, y_{M_1+1}, y_{M_1+2}, \dots, y_M)$$
$$= F_{Y_1^u,\dots,Y_M^u|\Lambda=\lambda}(g_a(t_1), \dots, g_a(t_w), y_{w+1}, \dots, y_{M_1}, g_\ell(t_{M_1+1}), g_\ell(t_{M_1+2}), \dots, g_\ell(t_M)), \tag{84}$$

implying that if an individual is active at a certain time before n, he is active all the time before.
The consequence is that a disabled individual cannot recover.
This proves the theorem.

Note that a sufficient condition for (72) to hold is that $p_{ai}(0, t)(p_{id}(t, t_\lambda+1) - p_{id}(t, t_\lambda))$ is increasing for $t \in [0, t_\lambda]$. Taking the derivative of the numerator of this expression with respect to t, using the Chapman-Kolmogorov backward differential equations (see e.g. Wolthuis (2003)), leads to:

$$d\frac{p_{ai}(0, t)(p_{id}(t, t_{\lambda+1}) - p_{id}(t, t_\lambda))}{dt}$$
$$= \frac{dp_{ai}(0, t)}{dt}(p_{id}(t, t_{\lambda+1}) - p_{id}(t, t_\lambda))$$
$$+ p_{ai}(0, t)\begin{pmatrix}(p_{id}(t, t_{\lambda+1}) - p_{\mathrm{id}}(t, t_\lambda))(\mu_{ia}(t) + \mu_{id}(t)) \\ -(p_{ad}(t, t_{\lambda+1}) - p_{ad}(t, t_\lambda))(\mu_{ia}(t))\end{pmatrix}. \tag{85}$$

Assuming that $p_{ai}(0, t)$ is increasing in t, which we did in Section VI, this derivative is positive if

$$p_{id}(t, t_{\lambda+1}) - p_{id}(t, t_\lambda) \geq p_{ad}(t, t_{\lambda+1}) - p_{ad}(t, t_\lambda), \tag{86}$$

(i.e. the death rates for a disabled individual are higher than those for an active person) or

$$\mu_{ia}(t) \leq \frac{\mu_{id}(t)(p_{id}(t, t_{\lambda+1}) - p_{id}(t, t_\lambda))}{(p_{ad}(t, t_{\lambda+1}) - p_{ad}(t, t_\lambda)) - (p_{id}(t, t_{\lambda+1}) - p_{id}(t, t_\lambda))}, \tag{87}$$

(i.e. the forces of recovery are "low enough").

The first condition may hold for $\lambda \leq M_1 - 1$, that is, during the period of deferment. The second condition may hold in general.

The fully continuous version is obtained by letting $M \to \infty$ and besides for each $i \in \{0, \dots, M-1\}$, $t_{i+1} - t_i \to 0$. Then for $\lambda \in [0, n)$, we get:

$$
\begin{cases}
0 & \text{for} \quad s < \int_0^\lambda g_a(s)\,ds + g_{\cdot d}(\lambda); \\
1 - \dfrac{p_{ai}(0,t)(p_{ia}(t,\lambda)\mu_{ad}(\lambda)+p_{ii}(t,\lambda)\mu_{id}(\lambda))}{p_{aa}(0,\lambda)\mu_{ad}(\lambda)+p_{ai}(0,\lambda)\mu_{id}(\lambda)} & \text{for} \quad \begin{array}{c} s = \int_0^t g_a(s)\,ds + \int_t^\lambda g_i(s)\,ds + g_{\cdot d}(\lambda); \\ 0 \leq t < n \end{array}, \\
1 & \text{for} \quad s \geq \int_0^\lambda g_i(s)\,ds + g_{\cdot d}(\lambda).
\end{cases} \tag{88}
$$

while for $\lambda \in [n, m)$, we get

$$
\begin{cases}
0 & \text{for} \quad s < \int_0^n g_a(s)\,ds + \int_n^\lambda g_\ell(s)\,ds + g_{\cdot}d(\lambda); \\
1 - \dfrac{p_{ai}(0,t)(p_{ia}(t,\lambda)\mu_{ad}(\lambda)+p_{ii}(t,\lambda)\mu_{id}(\lambda))}{p_{aa}(0,\lambda)\mu_{ad}(\lambda)+p_{ai}(0,\lambda)\mu_{id}(\lambda)} & \text{for} \quad \begin{array}{c} s = \int_0^t g_a(s)\,ds + \int_t^n g_i(s)\,ds \\ + \int_n^\lambda g_\ell(s)\,ds + g_{\cdot d}(\lambda); \quad 0 \leq t < \lambda, \end{array} \\
1 & \text{for} \quad s \geq \int_0^n g_i(s)\,ds + \int_n^\lambda g_\ell(s)\,ds + g_{\cdot d}(\lambda).
\end{cases} \tag{89}
$$

Finally, the case $\lambda = m$ yields the following result

$$
\begin{cases}
0 & \text{for} \quad s < \int_0^n g_a(s)\,ds + \int_n^m g_\ell(s)\,ds; \\
1 - \dfrac{p_{ai}(0,t)(p_{ia}(t,\lambda)\mu_{ad}(\lambda)+p_{ii}(t,\lambda)\mu_{id}(\lambda))}{p_{aa}(0,\lambda)\mu_{ad}(\lambda)+p_{ai}(0,\lambda)\mu_{id}(\lambda)} & \text{for} \quad \begin{array}{c} s = \int_0^t g_a(s)\,ds + \int_t^n g_i(s)\,ds + \int_n^m g_\ell(s)\,ds; \\ 0 \leq t < n \end{array}. \\
1 & \text{for} \quad s \geq \int_0^n g_i(s)\,ds + \int_n^m g_\ell(s)\,ds.
\end{cases} \tag{90}
$$

The numerical examples in the next section will be based on this continuous version.

As stated before, we assume that $\mu_{ad}(s) = \mu_{id}(s)$ for $s > n$. This leads to the following lemma.

Lemma 18 *Assume that, for each* $s > n$, $\mu_{ad}(s) = \mu_{id}(s)$. *Then, for* $\lambda \geq M_1$ and $t \leq n$:

$$
\frac{p_{ai}(0, t)\,(p_{id}(t, t_{\lambda+1}) - p_{id}(t, t_\lambda))}{p_{ad}(0, t_{\lambda+1}) - p_{ad}(0, t_\lambda)} = \frac{p_{ai}(0, t)\,(1 - p_{id}(t, n))}{1 - p_{ad}(0, n)}. \tag{91}
$$

Proof. For $t \leq n$ and $\lambda \geq M_1$, we have

$$\frac{p_{ai}(0, t)\,(p_{id}(t, t_{\lambda+1}) - p_{id}(t, t_\lambda))}{(p_{ad}(0, t_{\lambda+1}) - p_{ad}(0, t_\lambda))}$$

$$= \frac{p_{ai}(0, t)\,(p_{ia}(t, n)\,(p_{ad}(n, t_{\lambda+1}) - p_{ad}(n, t_\lambda)) + p_{ii}(t, n)\,(p_{id}(n, t_{\lambda+1}) - p_{id}(n, t_\lambda)))}{p_{aa}(0, n)\,(p_{ad}(n, t_{\lambda+1}) - p_{ad}(n, t_\lambda)) + p_{ai}(0, n)\,(p_{id}(n, t_{\lambda+1}) - p_{id}(n, t_\lambda))}$$

$$= \frac{p_{ai}(0, t)\,(1 - p_{id}(t, n))}{1 - p_{ad}(0, n)}, \tag{92}$$

since $p_{ad}(n, t_{\lambda+1}) - p_{ad}(n, t_\lambda) = p_{id}(n, t_{\lambda+1}) - p_{id}(n, t_\lambda)$ for $\mu_{ad}(s) = \mu_{id}(s)$, $s > n$.
The numerical examples will be based on this simplification as well.

2. Nontrivial lower bound

For $k = \lambda + 1$, we have that

$$E\,[Y_{\lambda+1} | \Lambda = \lambda] = g_{\cdot d}(t_{\lambda+1}), \tag{93}$$

while for $k > \lambda + 1$, $E\,[Y_k | \Lambda = \lambda] = 0$.

For $k \leq \lambda$, and $k \in \{1,.., M_1\}$, the following results:

$$E\,[Y_k | \Lambda = \lambda] = \frac{\begin{array}{c} p_{aa}(0, t_k)\,(p_{ad}(t_k, t_{\lambda+1}) - p_{ad}(t_k, t_\lambda))\,g_a(t_k) \\ + p_{ai}(0, t_k)\,(p_{id}(t_k, t_{\lambda+1}) - p_{id}(t_k, t_\lambda))\,g_i(t_k) \end{array}}{p_{ad}(0, t_{\lambda+1}) - p_{ad}(0, t_\lambda)}, \tag{94}$$

while for $k \leq \lambda$, and $k \in \{M_1 + 1,.., M\}$, we get

$$E\,[Y_k | \Lambda = \lambda] = g_\ell(t_k). \tag{95}$$

Hence $S^l | \Lambda = \lambda$ for $\lambda \in \{0,.., M_1 - 1\}$ is equal to:

$$\sum_{k=1}^{\lambda} \frac{p_{aa}(0, t_k)\,(p_{ad}(t_k, t_{\lambda+1}) - p_{ad}(t_k, t_\lambda))\,g_a(t_k) + p_{ai}(0, t_k)\,(p_{id}(t_k, t_{\lambda+1}) - p_{id}(t_k, t_\lambda))\,g_i(t_k)}{(p_{ad}(0, t_{\lambda+1}) - p_{ad}(0, t_\lambda))}$$

$$+ g_{\cdot d}(t_{\lambda+1}), \tag{96}$$

while for $\lambda \in \{M_1,.., M - 1\}$, we get

$$\sum_{k=1}^{M_1} \frac{p_{aa}(0, t_k)\,(p_{ad}(t_k, t_{\lambda+1}) - p_{ad}(t_k, t_\lambda))\,g_a(t_k) + p_{ai}(0, t_k)\,(p_{id}(t_k, t_{\lambda+1}) - p_{id}(t_k, t_\lambda))\,g_i(t_k)}{(p_{ad}(0, t_{\lambda+1}) - p_{ad}(0, t_\lambda))}$$

$$+ \sum_{k=M_1}^{\lambda} g_\ell(t_k) + g_{\cdot d}(t_{\lambda+1}). \tag{97}$$

Finally, for $\lambda = M$, we get:

$$\sum_{k=1}^{M_1} \frac{p_{aa}(0, t_k)\,(p_{ad}(t_k, t_{\lambda+1}) - p_{ad}(t_k, t_\lambda))\,g_a(t_k) + p_{ai}(0, t_k)\,(p_{id}(t_k, t_{\lambda+1}) - p_{id}(t_k, t_\lambda))\,g_i(t_k)}{(p_{ad}(0, t_{\lambda+1}) - p_{ad}(0, t_\lambda))}$$

$$+ \sum_{k=M_1}^{M} g_\ell(t_k) \tag{98}$$

The expectation $\mathrm{E}[S^l|\Lambda = \lambda]$ is not necessarily monotone as a function of λ.

The fully continuous version of the above is obtained by letting $M \to \infty$ and besides for each $i \in \{0, \dots, M-1\}$, $t_{i+1} - t_i \to 0$. The result is that $S^l|\Lambda = \lambda$ for $\lambda \in [0, n)$ is degenerate at

$$\int_{t=0}^{\lambda} \frac{p_{aa}(0, t)\,(p_{aa}(t, \lambda)\,\mu_{ad}(\lambda) + p_{ai}(t, \lambda)\,\mu_{id}(\lambda))}{p_{aa}(0, \lambda)\,\mu_{ad}(\lambda) + p_{ai}(0, \lambda)\,\mu_{id}(\lambda)}\,g_a(t)\,dt$$

$$+ \int_{t=0}^{\lambda} \frac{p_{ai}(0, t)\,(p_{ia}(t, \lambda)\,\mu_{ad}(\lambda) + p_{ii}(t, \lambda)\,\mu_{id}(\lambda))}{p_{aa}(0, \lambda)\,\mu_{ad}(\lambda) + p_{ai}(0, \lambda)\,\mu_{id}(\lambda)}\,g_i(t)\,dt + g_{\cdot d}(\lambda) \tag{99}$$

while for $\lambda \in [n, m)$, we get:

$$\int_{t=0}^{n} \frac{p_{aa}(0, t)\,(p_{aa}(t, \lambda)\,\mu_{ad}(\lambda) + p_{ai}(t, \lambda)\,\mu_{id}(\lambda))}{p_{aa}(0, \lambda)\,\mu_{ad}(\lambda) + p_{ai}(0, \lambda)\,\mu_{id}(\lambda)}\,g_a(t)\,dt$$

$$+ \int_{t=0}^{n} \frac{p_{ai}(0, t)\,(p_{ia}(t, \lambda)\,\mu_{ad}(\lambda) + p_{ii}(t, \lambda)\,\mu_{id}(\lambda))}{p_{aa}(0, \lambda)\,\mu_{ad}(\lambda) + p_{ai}(0, \lambda)\,\mu_{id}(\lambda)}\,g_i(t)\,dt + \int_{t=n}^{\lambda} g_\ell(t)\,dt$$

$$+ g_{\cdot d}(\lambda) \tag{100}$$

Finally, for $\lambda > m$, we get:

$$\int_{t=0}^{n} \frac{p_{aa}(0, t)\,(p_{aa}(t, \lambda)\,\mu_{ad}(\lambda) + p_{ai}(t, \lambda)\,\mu_{id}(\lambda))}{p_{aa}(0, \lambda)\,\mu_{ad}(\lambda) + p_{ai}(0, \lambda)\,\mu_{id}(\lambda)}\,g_a(t)\,dt$$

$$+ \int_{t=0}^{n} \frac{p_{ai}(0, t)\,(p_{ia}(t, \lambda)\,\mu_{ad}(\lambda) + p_{ii}(t, \lambda)\,\mu_{id}(\lambda))}{p_{aa}(0, \lambda)\,\mu_{ad}(\lambda) + p_{ai}(0, \lambda)\,\mu_{id}(\lambda)}\,g_i(t)\,dt + \int_{t=n}^{m} g_\ell(t)\,dt \tag{101}$$

In the next section, we present some numerical examples.

VIII. SOME INSURANCE CONTRACTS

This section aims to illustrate how the quality of the approximation differs with the relevant weight of the "Disability related benefits", compared with the "Other benefits". The examples will be presented in increasing order of the weight of the benefits of the latter type. We start with the disability annuity (as discussed in Spreeuw (2000)), then deal with the combined policy as discussed in Hesselager and Norberg (1996) and Norberg (1995), and finally consider a deferred annuity with a premium waiver during any spell of disability during the period of deferment. All annuity benefits are payable continuously. In all three examples, a level premium of c per annum, is payable continuously whenever the insured is in the state "Active". Hence, the premium payment function has the shape $B_a(t) = -ct; \quad t \in [0, n]$.

The following figures we use are the same as those applied in the numerical example 5.3 in Hesselager and Norberg (1996) and Table 4 of Norberg (1995):

$$\mu_{ad}(t) = \mu_{id}(t) = 0.0005 + 10^{-4.12+0.038\,(30+t)}; \quad \mu_{ia}(t) = 0.005;$$
$$\mu_{ai}(t) = 0.0004 + 10^{-5.46+0.06\,(30+t)}; \quad n = 30; \;\; \delta = \ln(1.045). \quad (102)$$

Note that, since $\mu_{ad}(t) = \mu_{id}(t)$ for all $t \geq 0$, the mortality rates for an active and disabled individual are the same. Hence inequality (86) is satisfied and so we can use the results from Theorems 16 and 17.

The value of δ corresponds to an annual level of interest of 4.5%. In all examples, the steplengths are taken to be $h = \frac{1}{1000}$ and $h' = \frac{1}{2000}$, just as in the numerical examples of Hesselager and Norberg (1996). The specification of the transition intensities results in the transition probabilities $p_{aa}(0, t)$, $p_{ai}(0, t)$ and $p_{ad}(0, t)$ as displayed graphically in Figure 3. Although it may not be completely clear from the graphics, all functions are smooth. It is clear that, for $t \in [0, n]$, $p_{ai}(0, t)$ is an increasing function of t.

A. *The disability annuity*

We take the same example as in Spreeuw (2000). In this case there are no "Other benefits" involved and the contract expires at time n. Hence $m = n$, $B_\ell(t) = 0$, and $b_{\cdot d}(t) = 0$. The contract pays an amount of 1 on a continuous basis while the insured is in the state "Disabled".

Hence $B_i(t) = t$; $t \in [0, n]$. Furthermore, the level premium satisfying the principle of equivalence is in this case equal to

$$c = \frac{\int_0^n e^{-\delta t} p_{ai}(0,t)dt}{\int_0^n e^{-\delta t} p_{aa}(0,t)dt} = 0.0175456. \tag{103}$$

The maximum possible present value is equal to

$$\int_0^{30} 1 \cdot (1.045)^t \, dt = 16.6527. \tag{104}$$

while the minimum, attained in case the individual remains in the state "Active" from the beginning till the end of the contract period, proves to be

$$-c\int_0^n e^{-\delta t} dt = a = -0.2922. \tag{105}$$

The algorithm by Hesselager and Norberg has in this case been executed for $a = -0.35$ and $b = 16.7$.

The original convex upper bound is derived as follows. The distribution of S^c, displayed in formula (57), reads as:

$$\Pr[S^c \leq s] = \begin{cases} 0 & \text{for} \quad s < -0.29; \\ 1 - p_{ai}(0, h_i(s, n)) - p_{ad}(0, h_a(s)) & \text{for} \quad -0.2922 \leq s < 0; \\ 1 - p_{ai}(0, h_i(s, n)) & \text{for} \quad 0 \leq s < 16.6527; \\ 1 & \text{for} \quad s \geq 16.6527. \end{cases} \tag{106}$$

In (106),

$$h_i(s, u) = -\frac{\ln\left(\frac{\delta s + e^{-\delta u} + c}{1+c}\right)}{\delta};$$

(solution to h_i of $s = -c\int_{w=0}^{h_i} e^{-\delta w} dw + \int_{w=h_i}^{u} e^{-\delta w} dw$)

$$h_a(s) = -\frac{\ln\left(1 + \frac{\delta s}{c}\right)}{\delta} \text{ (solution to } h_a \text{ of } s = -c\int_{w=0}^{h_a} e^{-\delta w} dw). \tag{107}$$

The improved upper bound follows by conditioning on the random variable $\Lambda = \min[T, n]$. The conditional minimum and maximum values that S^u, given $\Lambda = \lambda$ can attain are given in Figure 4. Although it may not be completely clear from the figure, the conditional minimum values are a smooth function of the time of death.

By integrating over all possible values of T, we get:

$$\Pr[S^u \leq s] = \begin{cases} 0 & \text{for } s < -0.29; \\ \int_{u=h_a(s)}^{n} p_{aa}(0, h_i(s,u)) \begin{pmatrix} p_{aa}(h_i(s,u), u)\, \mu_{ad}(u) \\ +p_{ai}(h_i(s,u), u)\, \mu_{id}(u) \end{pmatrix} du & \text{for } -0.2922 \leq s < 0; \\ \quad +p_{aa}(0, h_i(s,n))\,(1 - p_{ad}(h_i(s,n), n)) & \\ p_{ad}(0, h_d(s)) + p_{aa}(0, h_i(s,n))\,(1 - p_{ad}(h_i(s,n), n)) & \\ \quad +\int_{u=h_d(s)}^{n} p_{aa}(0, h_i(s,u)) \begin{pmatrix} p_{aa}(h_i(s,u), u)\, \mu_{ad}(u) \\ +p_{ai}(h_i(s,u), u)\, \mu_{id}(u) \end{pmatrix} du & \text{for } 0 \leq s < 16.6527; \\ 1 & \text{for } s \geq 16.6527. \end{cases} \tag{108}$$

with $h_i(s, u)$ and $h_a(s)$ as defined before. Besides,

$$h_d(s) = -\frac{\ln(1 - \delta s)}{\delta} \quad \left(\text{solution to } h_d \text{ of } s = \int_{w=0}^{h_d} e^{-\delta w} dw\right) \tag{109}$$

The conditional expectation $E[S^u | \Lambda = \lambda]$ is depicted in Figure 5. We can see that $E[S^u | \Lambda = \lambda]$ is not a monotone function of t. The minimal value is attained for $T = 15.4241$, corresponding to $E[S^u | T = 15.4241] = -0.129096$. The maximal value is attained for $T = n$ (survival to the end of term), corresponding to $E[S^u | T = 30] = 0.0143711$.

In Figure 6, the four c.d.f.'s are displayed. One can see that the convex upper bound gives a very good approximation which can hardly be distinguished from the real distribution. As expected, the improved upper bound works out even better. The lower bound, on the other hand, performs very badly.

This observation is also confirmed by comparing the variances with each other:

$$Var[S] = 1.80677; \; Var[S^c] = 1.94873;$$
$$Var[S^l] = 0.00141346; \; Var[S^u] = 1.87002. \tag{110}$$

B. *The combined policy*

We use the same values as in Hesselager and Norberg (1996) and Norberg (1995). Compared to the previous example, there are other benefits involved, namely a lump sum benefit of 1 payable immediately on death. Moreover, the disability annuity is reduced to 0.5 per annum, payable continuously. Like the previous example, the contract expires at time $n = 30$. Hence $m = n$, $B_i(t) = 0.5t$, and $b_{\cdot d}(t) = 1$; $t \in [0, n]$. Furthermore, the level premium satisfying the principle of equivalence is equal to

$$c = \frac{0.5 \int_0^n e^{-\delta t} p_{ai}(0, t)\, dt + \int_0^n e^{-\delta t} (p_{aa}(0, t)\, \mu_{ad}(t) + p_{ai}(0, t)\, \mu_{id}(t))\, dt}{\int_0^n e^{-\delta t} p_{aa}(0, t)\, dt}$$

$$= 0.0131083. \tag{111}$$

The minimum possible value is attained in case the contract survives to the end of term and stays active during an uninterrupted period of time. This value is equal to:

$$-c \int_0^n e^{-\delta t} dt = a = -0.218289. \tag{112}$$

In this case, the maximal possible value is attained if the life gets disabled immediately after issue, stays disabled until death just before the end of term. This value is equal to:

$$0.5 \int_0^n e^{-\delta t} dt + e^{-\delta n} = 8.59335. \tag{113}$$

(Note that this is not so obvious: it all depends on the levels of the interest rate, the death benefit and the disability annuity. If the volume of the disability annuity is relatively low compared to the level of the death benefit, and interest rates are high, the maximal possible value could correspond to immediate death after issue.)

In Figure 7, the conditional minima and maxima, given the time of death, are displayed. This applies if death happens before the end of term. This example is more complicated than the previous one as the minimal and maximal values in case of survival to the end of term differ from those in case of earlier death. Given that the contract survives to the end of term, the minimal and maximal values are equal to -0.218289 and 8.32635, respectively. This implies that the minimal and maximal values S can take are equal to -0.218289 and 8.59335, respectively.

The conditional expected values for the lower bound in case of earlier death are given in Figure 8. We see that the maximum expected value corresponds to immediate death. The minimum expected value attained if the contract survives to the end of the term is equal to -0.0637065.

Figure 9 displays the three different c.d.f.'s. One can see that the improved upper bound is again almost perfect. The lower bound is still poor, though relatively better than in the previous case. This is due to the fact that there are other benefits involved. The variances displayed below illustrate this case.

$$Var\,[S] = 0.486852;\; Var\,[S_l] = 0.0273882;\; Var\,[S^u] = 0.502725. \tag{114}$$

C. *Deferred annuity with premium reduction while disabled*

The "Other benefits" involved are in this case equal to a pension benefit of 1 per annum, payable continuously from duration n. A premium of c per annum, payable continuously, will be payable during the period of deferment while the life is active. During any spell of disability, a reduced premium of $\alpha \cdot c$ is payable, with $\alpha \in [0, 1]$. In this example, we have $n = 30$, $m = \infty$, $B_a(t) = -ct$, $B_i(t) = -\alpha c t$, $t \in [0, n]$, and $B_\ell(t) = t$, $t \in [n, \infty)$. The contract pays an amount of 1 on a continuous basis while the insured is in the state "Disabled". Furthermore, the level premium satisfying the principle of equivalence is equal to

$$c = \frac{\int_n^\infty e^{-\delta t}\,(1 - p_{ad}(0, t))\,dt}{\int_0^n e^{-\delta t} p_{aa}(0, t)\,dt + \alpha \int_0^n e^{-\delta t} p_{ai}(0, t)\,dt} \tag{115}$$

We will consider the case $\alpha = 0$ (premium waiver) when the difference between upper and lower bounds should be most significant. The minimum possible value is attained in case the contract survives to the end of term, stays active during an uninterrupted period of time until n and dies just after n. This value is equal to:

$$-c\int_0^n e^{-\delta t} dt = a = -2.82. \tag{116}$$

The maximal possible value is attained if the contract gets disabled immediately after issue, stays disabled until retirement and survives until the limiting age. We assume this limiting age to be equal to 115. So $m = 85$ and the maximal value is equal to:

$$\int_{30}^{115} e^{-\delta t} dt = 4.14. \tag{117}$$

In Figure 10, the conditional minima and maxima, given the time of death, are displayed. The conditional expected values, for calculation of the lower bound, are depicted in Figure 11. Obviously, the minimum expected value is attained if the contract survives to the end of the term. The maximum expected value corresponds to survival to the limiting age.

We will not display the c.d.f.'s of the three distributions, as they can hardly be distinguished from each other. This is also illustrated by comparing the three variances:

$$Var\,[S] = 2.14702;\; Var\,[S^l] = 2.08842;\; Var\,[S^u] = 2.14869. \quad (118)$$

As with the other two examples, the improved upper bound still gives the best approximation and is almost perfect. The lower bound, performs well as well. This is explained by the substantial weight of the other benefits.

For larger values of a, the two distributions of S^l and S^u will be closer to each other, coinciding completely for $a = 1$ (when there is effectively no difference between the states "Active" and "Disabled").

IX. CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

In this paper, we have applied the theory of comonotonic risks to life insurance contracts with disability related benefits. The (conditional) upper bounds are based on removing the possibility of recovery from illness. This upper bound proves to be a very good approximation in all cases. The quality of the lower bound depends on the relative weight of the disability related benefits. The lower the weight, and hence the more significant the contribution of the other benefits, the better the approximation by the lower bound. As the other benefits are governed by the chosen conditioning variable of remaining lifetime, this result is not very surprising. In Kaas et al. (2000), it was shown that the more the conditioning random variable resembles the random variable under consideration (the present value of benefits less premiums) the better the approximation by the lower bound.

This paper gives scope for more research. We intend to extend the model by allowing for financial risks. Assuming that these risks are independent of the demographic risks, convex upper and lower bounds can then be derived by applying the techniques in

Hoedemakers et al. ((2003), (2004)). Furthermore, we intend to derive present value distributions of contract with more complicated conditions in their design, such as a waiting period and a deferred period, as specified mathematically in Pitacco (1995) and Haberman and Pitacco (1999).

REFERENCES

Alting von Geusau, B.J.J., 1990, AIDS in the Eighties and Nineties, in, Liber Amicorum, Kok/Van Klinken, (Institute of Actuarial Science and Econometrics, University of Amsterdam, Amsterdam), pp.1-22.

De Pril, N., 1989, The Distribution of Actuarial Functions, *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, 173-181.

Dhaene, J., 1990, Distributions in Life Insurance, *ASTIN Bulletin*, 20, 81-92.

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.

Goovaerts, M.J., Dhaene, J., De Schepper, A., 2000, Stochastic Upper Bounds for Present Value Functions, *The Journal of Risk and Insurance* 67, 1, 1-14.

Gregorius, F.K., 1993, Disability Insurance in the Netherlands. *Insurance: Mathematics and Economics* 13, 101-116.

Haberman, S., Pitacco, E., 1999, Actuarial Models for Disability Insurance, (Chapman & Hall/CRC, Boca Raton).

Hesselager, O., Norberg, R., 1996, On Probability Distributions of Present Values in Life Insurance. *Insurance: Mathematics and Economics*, 18, 35-42.

Hoedemakers, T., Beirlant, J., Goovaerts, M.J,.Dhaene, J., 2003, Confidence Bounds for Discounted Loss Reserves. *Insurance: Mathematics and Economics* 33, 2, 297-316.

Hoedemakers, T, Darkiewicz, G., Dhaene, J., Goovaerts, M.J., 2004, On the Distribution of Life Annuities with Stochastic Interest Rates. Proceedings of the Eighth International Congress on *Insurance: Mathematics and Economics*, (Rome).

Kaas, R., Dhaene, J., Goovaerts, M.J., 2000, Upper and Lower Bounds for Sums of Random Variables. *Insurance: Mathematics and Economics* 27, 151-168.

Norberg, R., 1995, Differential Equations for Higher Order Moments of Present Values in Life Insurance, *Insurance: Mathematics and Economics* 17, 171-180.

Pitacco, E., 1995, Actuarial Models for Pricing Disability Benefits: Towards a Unifying Approach, *Insurance: Mathematics and Economics* 16, 39-62.

Spreeuw, J., 2000, Convex Order and Multistate Life Insurance Contracts, Actuarial Research Report No. 129, (Department of Actuarial Science and Statistics, City University, London).

Vyncke, D., Goovaerts, M.J., Dhaene, J., 2001, Convex Upper and Lower Bounds for Present Value Functions, *Applied Stochastic Models in Business and Industry* 17, 149-164.

Wolthuis, H., 2003, Life Insurance Mathematics (The Markovian Model), (IAE, Universiteit van Amsterdam, Amsterdam), second edition.

FIGURE 1

Disability without possible recovery (hierarchical Markov chain).

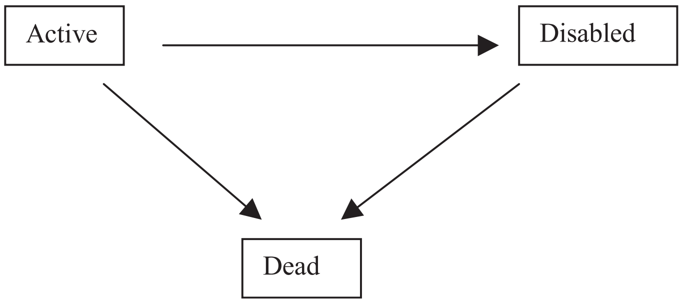

FIGURE 2

Disability with possible recovery (non-hierarchical Markov chain).

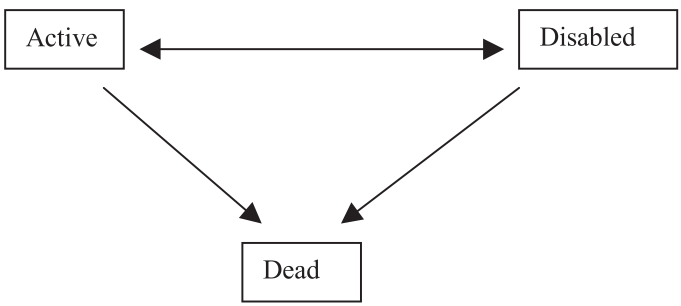

FIGURE 3

Graphical display of $p_{aa}(0, t)$ (dotted), $p_{ai}(0, t)$ (solid) *and* $p_{ad}(0, t)$ (dashed) *as a function of t.*

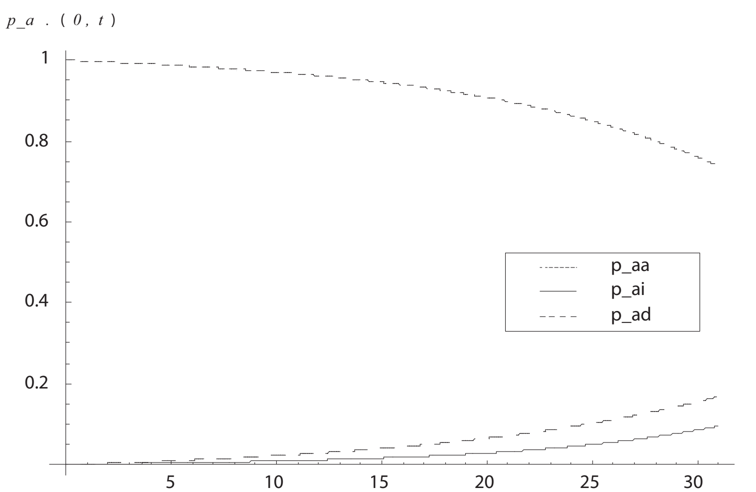

FIGURE 4

Conditional minima and maxima present values disability annuity, given time of death

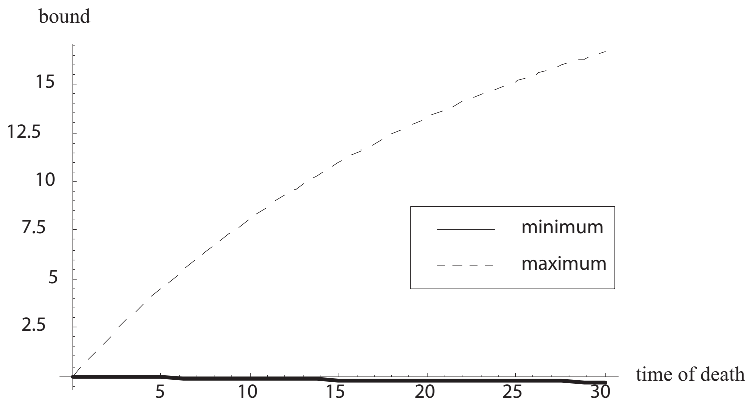

FIGURE 5

Conditional expected values disability annuity

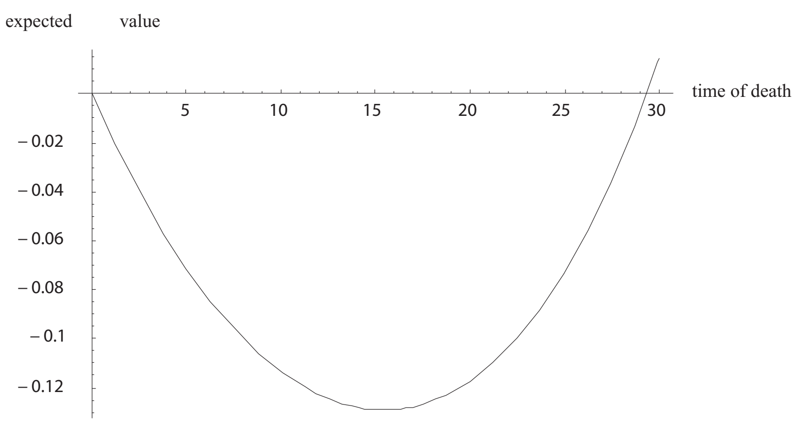

FIGURE 6

C.d.f.'s disability annuity compared: real distribution (solid)*; original convex upper bound* (dashed-dotted)*; improved upper bound* (dotted)*; lower bound* (dashed)

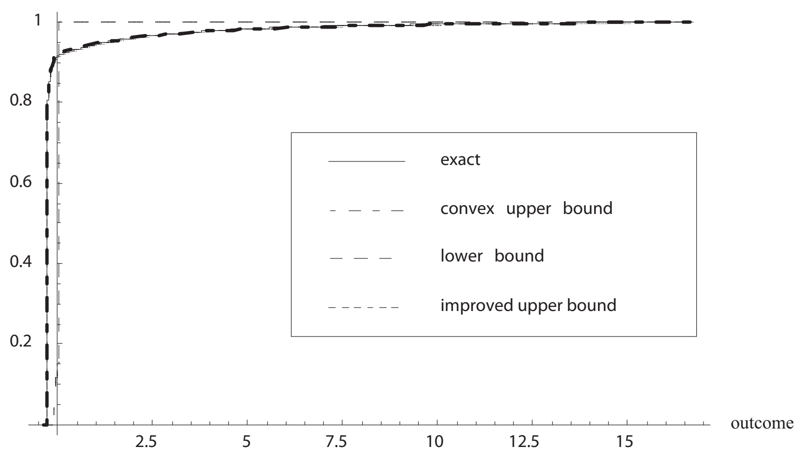

FIGURE 7

Conditional minima and maxima present values combined policy, given time of death

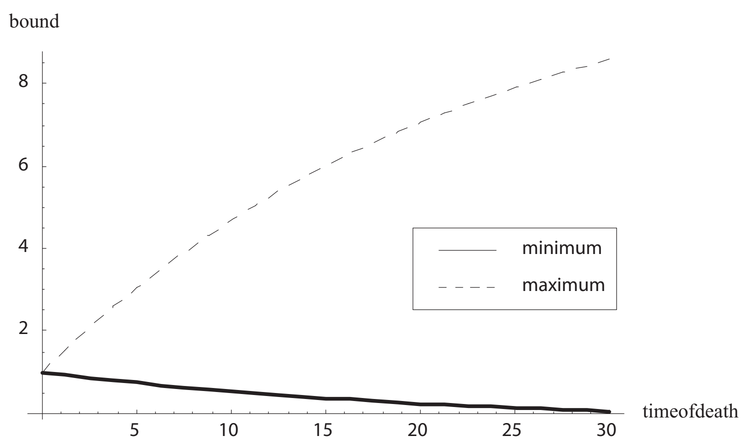

FIGURE 8

Conditional expected values in case of death: combined policy

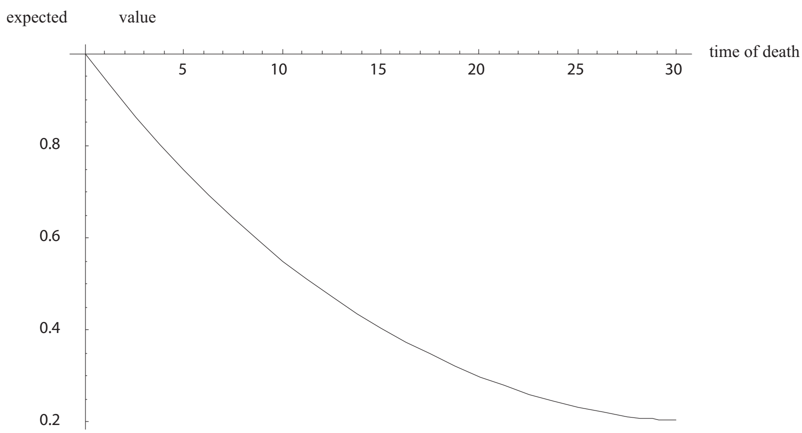

FIGURE 9

C.d.f.'s combined policy compared: real distribution (solid); *improved upper bound* (dotted); *lower bound* (dashed)

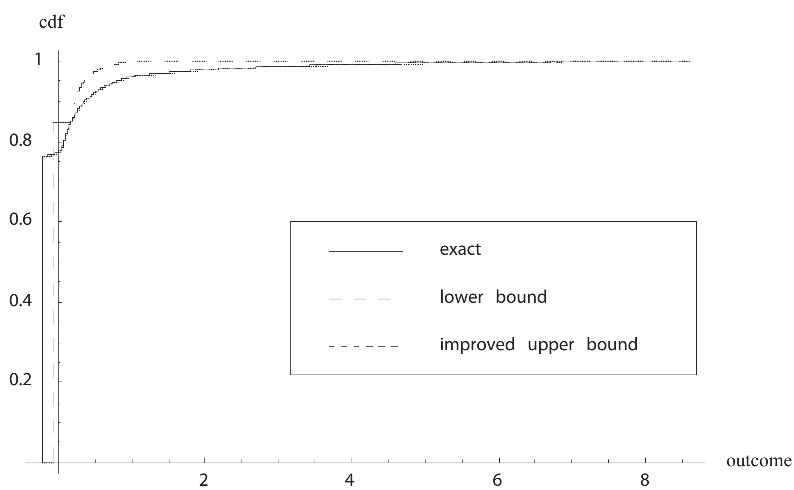

FIGURE 10

Conditional minima and maxima present values deferred annuity, given time of death

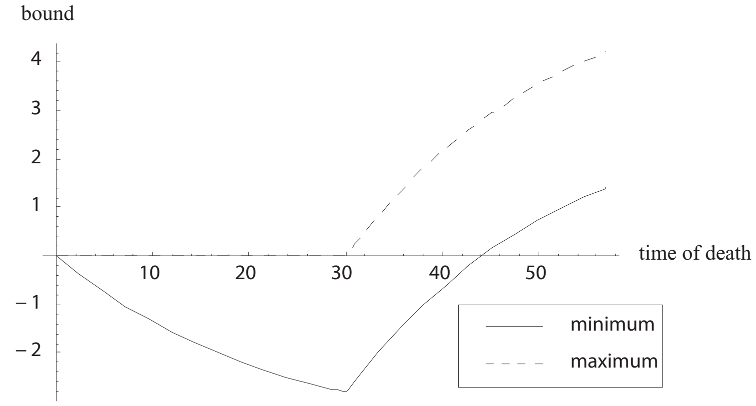

FIGURE 11

Conditional expected values deferred annuity