30th International Congress of Actuaries
30 March to 4 April 2014

Multi-state Microeconomic Model for Pricing and Reserving a disability insurance policy over an arbitrary period

Benjamin Schannes

Abstract

The purpose of this paper is to evaluate the cost of a disability insurance contract over an arbitrary period [0,T], thanks to a microeconomic approach inserted in the flexible and unifying framework of a Multi-State model. The key idea is that a large, complex group may be summarized by an appropriate single individual. As a matter of fact, in the context of group insurance, burdensome actuarial calculations and as a result risk assessment may be drastically streamlined by adopting a representative insured assumption. An algorithm for simulating events occurring during the period [0, T], i.e. a sequence of periods of activity and disability, potentially interrupted by the occurrence of the representative insured's death, is there constructed. This approach allows us, using a limit distribution argument, to express at $t = 0$ the theoretical value of a constant annual premium consistent with solvency constraints for the time horizon and the representative insured considered. More generally, this model can be used to price and reserve any group insurance contract for a wide range of policy conditions set by the insurer. Using a scenario-based approach, we can straightforwardly project future technical accounts determined by the specified solvency constraints and risk horizon, thus providing stress tests required for a sound risk management.

1 Introduction

We build in this paper a three-state model, 'active', 'disabled[1]' and 'dead', on the basis of estimated instantaneous transition rates depending on the representative agent selected[2]. A Poisson Model has been chosen to estimate and graduate the "active → disabled" transition rates, and a Cox Model has been implemented for the recovery, i.e. the "disabled → active" transition. French Data which have been used to estimate these models are Marsh & McLennan Companies property. Explanatory variables which have been selected to predict the transition rates are: age, gender, area of work (Large cities, Medium sized cities, Small sized cities & rural areas), and the sector of activity. Concerning the transition to the 'death' state, daily mortality rates have been computed from the French mortality tables TF00-02 and TH00-02.

The stochastic processes driving transitions have been supposed locally Homogeneous Markov, except for the disability termination which is driven by a locally Homogeneous Semi-Markov process, thanks to the Cox model which enables to take into account the time elapsed since the inception of disability. The local Homogeneity property of the different processes comes from the econometric design of rating factors: by construction, regression coefficients are constant within each age class. Daily mortality rates have been supposed age-constant besides.

2 Modeling

Let us consider the three-state model : 'active' (transient state e_1) , 'disabled' (transient state e_2), and 'death' (absorbing state e_3). On the graph 1 arrows depict possible transitions:

1. disablement: transition from the 'active' state to the disability state,
2. recovery: transition from disability to the 'active' state,
3. death of an active life: transition from the 'active' state to death,
4. death of a disabled life: transition from the 'disabled' state to death.

Time origin is considered to be the date of issuance of the written policy. We suppose the covered insured is x years old. Besides, we suppose the disability incidence process to be Markov: the instantaneous rate of disablement does not depend on the elapsed time within the 'active' state. The underlying law is exponential, which is memoryless. In this framework, a Poisson model for the disability incidence will enable us to build the transition law.

Intuitively, it seems rather relevant to get a decreasing recovery rate as a function of the disability spell duration. It requires to set up a Semi-Markov model which makes the exit rate from disablement[3] depend not only on the age, but also on the duration of disability. A Cox model is thus appropriate and we will use one which takes also into account unobserved heterogeneity. In the absence of data to assess a potential specific mortality among employees who have been selected to estimate our econometric models, we have used the mortality rates arising from the TF00-02 and TH00-02 tables. It seems however reasonable to deem that mortality rates might increase with the disability duration.

Moving to a Semi-Markov framework to model the recovery requires to take into account the length of stay in the occupied state in order to predict the future of the underlying stochastic process. In respect of this matter, let us define a bivariate process $\mathcal{XD} = \{(X_t, D_t), t \geq 0\}$, where X_t is a stochastic process with right-continuous paths providing the occupied state at time t, and D_t is the duration of stay in this state. Formally,

$$D_t = \max \left\{ s \leq t \mid X_t = X_{t-h} \text{ for any } 0 \leq h \leq s \right\}.$$

For an insured aged x at policy issue ($t = 0$), the event $\left\{X_t = e_2 \cap D_t = s\right\}$ means that at time t[4], the insured has been disabled for s units of time, i.e. s is the time spent in the state e_2 up to time t since the latest transition in that state.

[1] by an accident or sickness

[2] A model has a representative agent when agents differ, but act in such a way that the sum of their choices is mathematically equivalent to the decision of one individual or many identical individuals.

[3] due to recovery, not to mortality

[4] the variable t is often called seniority

Finally, let us recall that in the Semi-Markov framework, the stochastic process X_t is deemed Semi-Markov if the process $\{(X_t, D_t), t \geq 0\}$ is Markov. In this sense, moving to a Semi-Markov model amounts to introducing one more dimension.

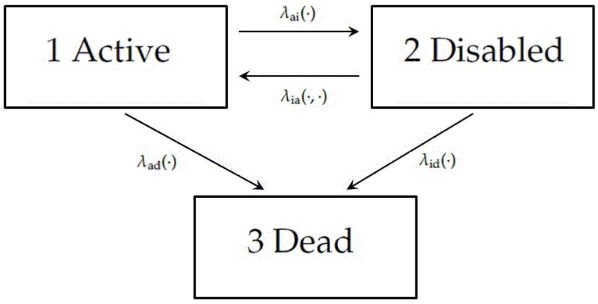

Figure 1: Three-state Model for the insurance scheme.

3 Transition intensities

3.1 First assumptions and definitions

In what follows,we will use a Hamza-like notation in which the letters 'a','i','d' denote the states 'active', 'invalid' and 'dead'. Haberman & Pitacco [8] also adopt this convention.

Since the active state is Markov, the instantaneous transition rate from this state to disability for an insured aged $x + t$ may be expressed as follows:

$$\lambda_{\text{ai}}(x+t) = \lim_{\Delta t \downarrow 0} \frac{\Pr\,[X_{t+\Delta t} = e_2 \mid X_t = e_1]}{\Delta t},$$

and the instantaneous mortality rate from the active state at age $x + t$ is given by

$$\lambda_{\text{ad}}(x+t) = \lim_{\Delta t \downarrow 0} \frac{\Pr\,[X_{t+\Delta t} = e_3 \mid X_t = e_1]}{\Delta t}.$$

For the transitions from the disability state, which have been assumed Semi-Markov, we also have a duration-in-current-state dependence. Therefore we get:

$$\lambda_{\text{ia}}(x+t,s) = \lim_{\Delta t \downarrow 0} \frac{\Pr\,[X_{t+\Delta t} = e_1 \mid X_t = e_2, D_t = s]}{\Delta t}$$

which is the recovery rate at t for an insured aged x at time 0 who has spent a time s in the disability state up to time t since the latest transition to that state. We assume the instantaneous mortality rates from the active state and from the disability state are equal:

$$\lambda_{\text{id}}(x+t) = \lambda_{\text{ad}}(x+t).$$

Finally, we define the exit rate from the active state at age $x + t$ by the sum

$$\lambda_{\text{a}\bullet}(x+t) = \lambda_{\text{ai}}(x+t) + \lambda_{\text{ad}}(x+t),$$

and we also provide the exit rate from the disability state for an insured aged $x + t$, with s the time elapsed since the latest transition into the disability state:

$$\lambda_{\text{i}\bullet}(x+t,s) = \lambda_{\text{ia}}(x+t;s) + \lambda_{\text{id}}(x+t).$$

3.2 A Transition Intensity Approach (TIA)

Transition probabilities from the active state do not depend on the information regarding the path of the process, whatever this information may be. Hence, for $u \leq t$:

$$_{s,t-u}p_{x+u}^{\mathrm{ai}} = \Pr\ [X_t = e_2 \cap D_t \leq s \mid X_u = e_1]$$

gives the probability that an active insured aged $x+u$ will be disabled at the age $x+t$ with a time spent in the disability state less than or equal to s. Similarly,

$$_{t-u}p_{x+u}^{\mathrm{aa}} = \Pr\ [X_t = e_1 \mid X_u = e_1]$$

is the probability that an insured which is active at age $x+u$ will also be active at age $x+t$. Finally,

$$_{t-u}p_{x+u}^{\mathrm{ad}} = \Pr\ [X_t = e_3 \mid X_u = e_1]$$

corresponds to the probability that an active insured aged $x+u$ will die during his lifetime until the age of $x+t$. In these expressions, we note that the path of the insured risk between ages $x+u$ and $x+t$ is not a conditioning event. For any times $u \leq t$ and $s \geq t-u$, we get the equality:

$$_{s,t-u}p_{x+u}^{\mathrm{ai}} + {}_{t-u}p_{x+u}^{\mathrm{aa}} + {}_{t-u}p_{x+u}^{\mathrm{ad}} = 1.$$

Now we give some classical assumptions about transition probabilities:

$$\lim_{\triangle t \downarrow 0} {}_{\triangle t}p_{x+t}^{\mathrm{aa}} = \lim_{\triangle t \downarrow 0} {}_{s,\triangle t}p_{x+t}^{\mathrm{ai}} = \lim_{\triangle t \downarrow 0} {}_{\triangle t}p_{x+t}^{\mathrm{ad}} = 0,$$

If a time interval has an infinitesimal length, no more than one jump (i.e. transition) can occur during this time interval.

Then, we define the transition probabilities from the disability state, for which we also need to make allowance for the time elapsed since the latest transition into this state, except for the transition to the death state. Hence, the conditioning event is more complex:

$$\begin{aligned}
_{t-u}p_{x+u,s}^{\mathrm{ia}} &= \Pr\ [X_t = e_1 \mid X_u = e_2, D_u \leq s] \\
_{t-u}p_{x+u,s}^{\mathrm{ii}} &= \Pr\ [X_t = e_2 \mid X_u = e_2, D_u \leq s] \\
_{t-u}p_{x+u}^{\mathrm{id}} &= \Pr\ [X_t = e_2 \mid X_u = e_2].
\end{aligned}$$

The first of these expressions is the probability that an insured aged $x+u$, who has been disabled since the age of $x+u-s$ will be active at the age $x+t$. As above, we assume that, for any $u \leq t$, and for any disability spell duration s:

$$_{t-u}p_{x+u,s}^{\mathrm{ia}} + {}_{t-u}p_{x+u,s}^{\mathrm{ii}} + {}_{t-u}p_{x+u}^{\mathrm{id}} = 1.$$

Thanks to an analytical calculation[5], we get the following useful approximations for the transition probabilities from the active state:

$$\begin{aligned}
_{\triangle t}p_{x+t}^{\mathrm{ai}} &\approx \lambda_{\mathrm{ai}}(x+t)\ \triangle t \\
_{\triangle t}p_{x+t}^{\mathrm{ad}} &\approx \lambda_{\mathrm{ad}}(x+t)\ \triangle t \\
_{\triangle t}p_{x+t}^{\mathrm{aa}} &\approx 1 - \lambda_{\mathrm{a}\bullet}(x+t)\ \triangle t
\end{aligned}$$

where $\triangle t$ is small enough.

Other calculations, which have been detailed in the technical annexes, also enable us to give the following approximations to the transition probabilities from the disability state:

[5] See technical annexes

$$
\begin{aligned}
{}_{\triangle t}p^{\text{ia}}_{x+t,s} &\approx \lambda_{\text{ia}}(x+t,s)\ \triangle t \\
{}_{\triangle t}p^{\text{id}}_{x+t} &\approx \lambda_{\text{id}}(x+t)\ \triangle t \\
{}_{\triangle t}p^{\text{ii}}_{x+t,s} &\approx 1-\lambda_{\text{i}\bullet}(x+t,s)\ \triangle t
\end{aligned}
$$

where $\triangle t$ is small enough.

3.3 Expressions from occupancy probabilities

In accordance with the so-called 'Hamza notation', we denote ${}_tp^{\overline{\text{aa}}}_x$ the probability that an insured aged x and active at that age continuously remains active until the age of $x+t$. From the process $\{D_t,\ t \geq 0\}$, we can write

$$
\begin{aligned}
{}_\tau p^{\overline{\text{aa}}}_{x+t} &= \Pr\ [X_{t+\varepsilon} = e_1 \text{ for any } 0 \leq \varepsilon \leq \tau \mid X_t = e_1] \\
&= \Pr\ [X_{t+\tau} = e_1 \cap D_{t+\tau} \geq \tau \mid X_t = e_1].
\end{aligned}
$$

We provide in the technical annexes the following equality for the occupancy probability for an active insured:

$$
{}_tp^{\overline{\text{aa}}}_x = \exp\left(-\int_0^t \lambda_{\text{a}\bullet}(x+\tau)\ d\tau\right) \tag{1}
$$

We define in a similar manner the probability that an insured aged $x+t$ and disabled at that age, whose age at current disability inception is $x+t-s$, continuously remains disabled until the age of $x+t+\tau$.

$$
{}_\tau p^{\overline{\text{ii}}}_{x+t,s} = \Pr\ [X_{t+\varepsilon} = e_2 \text{ for any } 0 \leq \varepsilon \leq \tau \mid X_t = e_2, D_t = s],
$$

We can get as above an explicit expression for the occupancy probability from the transition intensities. Indeed, we prove in the technical annexes the following equality:

$$
{}_\tau p^{\overline{\text{ii}}}_{x+t,s} = \exp\left(-\int_0^\tau \lambda_{\text{i}\bullet}(x+t+\varepsilon, s+\varepsilon)\ d\varepsilon\right). \tag{2}
$$

4 Calculation of Premiums and Reserves pertaining to an insurance policy

4.1 Discount factor

Considering a constant interest rate i, we denote $v = (1+i)^{-1}$ the annual discount rate and $r = \ln(1+i)$ the resulting constant force of interest. The present value at time s of an amount of € 1 paid out at time t, $s < t$, is:

$$
v(s,t) = v^{t-s} = \exp(-(t-s)\ \ln(1+i)) = \exp(-(t-s)\ r) \tag{3}
$$

If the interest rate is not constant, we denote $r(u)$ the instantaneous interest rate at time u. r can be modelled by a stochastic process. The discount factors $v(t)$ on the time interval $[0,t]$, and $v(s,t)$ on the time interval $[s,t]$ are then given by

$$v(t) = \exp\Big(-\int_0^t r(u)\,du\Big) \tag{4}$$

$$v(s,t) = \exp\Big(-\int_s^t r(u)\,du\Big) = \frac{v(t)}{v(s)} \tag{5}$$

4.2 Premium and Benefit Functions

Let n denote the term of the disability policy. Now we introduce the functions:

1. $t \mapsto \Pi_a(t)$: right-continous, non-negative and non-decreasing cumulative premium functions, which describe the commitments of the insured. Premiums are paid by the insured for an amount

$$\int_t^{t+\triangle t} d\Pi_a(\zeta) = \Pi_a(t+\triangle t) - \Pi_a(t).$$

while he is in the 'active' state in the interval $[t, t+\triangle t]$. The functions $t \mapsto \Pi_a(t)$ can be expressed as the sum of an absolutely continous function and jumps:

$$d\Pi_a(t) = \pi_a(t)dt + \Delta\Pi_a(t)$$

where

$$\Delta\Pi_a(t) = \Pi_a(t) - \Pi_a(t-).$$

The set of discontinuity points of $\Pi_a(\cdot)$ is denoted $\text{Disc}\Big(\Pi_a\Big) = \{t_{\text{a},0}, t_{\text{a},1}, \dots, t_{\text{a},q}\}$. In case of periodic premiums, we get $\Pi_a(t) = P\lceil t \rceil$, where P is a premium rate and $\lceil t \rceil$ is the smallest integer greater than or equal to t. In this case, $\int_0^t d\Pi_a(\zeta)$ will be a single sum.

2. $(t,\tau) \mapsto c_{ai}(t,\tau)$: a deterministic function specifying lump sum payments due upon transition from the 'active' state to the 'disabled' state at time t after having spent a time τ in the active state. The functions $t \mapsto c_{ai}(t, t-v)$ are left-continuous for any $v > 0$. If the 'one-shot' payment does not depend on the time spent in the occupied state (i.e. active) before the latest transition to disability, we only need to write a function $c_{ai}(t)$. For our final application, this scarcer supplementary benefit in case of transition will be omitted.

3. $(t,\tau) \mapsto B_i(t,\tau)$: non-negative cumulative benefit functions, which describe the insurer's commitments to annuities at time t when the time elapsed since the latest transition to the 'disabled' state is τ. The functions $t \mapsto B_i(t, t-v)$ are non-decreasing and right-continuous for any $t \geq v > 0$. The amount paid to the insured (disabled at time t with a disability time length equal to s) from time t to $t+\triangle t$ is

$$\int_t^{t+\triangle t} dB_i(\zeta;\zeta - s) = B_i(t+\triangle t, t+\triangle t - s) - B_i(t, t-s)$$

The functions $t \mapsto B_i(t, t-s)$ can be expressed as the sum of an absolutely continous function and jumps (pure endowments):

$$dB_i(t,t-s) = b_i(t,t-s)dt + \Delta B_i(t,t-s)$$

where

$$\Delta B_i(t,t-s) = B_i(t,t-s) - B_i(t-,(t-)-s)$$

$\Delta B_i(t,t-s)$ is the benefit paid at t to the insured who has been disabled at that time with a duration $t-s$ since the latest onset of disability. For any $s > 0$, the set of discontinuity points of $t \mapsto B_i(t,t-s)$, whose domain is the subset of real numbers t such that $t > s$, is denoted $\text{Disc}_s\Big(B_i\Big) = \{t_{\text{i},s,0}, t_{\text{i},s,1}, \dots, t_{\text{i},s,q_s}\}$. If the annuities do not depend on the time spent in the 'disabled' state, the left-continuous cumulative annuities on $[0,t]$ will be denoted $B_i(t)$. For instance, in case of periodic benefit (€ 1), we can write $B_i(t) = \lceil t \rceil$ or $B_i(t) = \lceil t \rceil - 1, t > 0$, depending on the payment mode adopted (in advance or in arrears). In the case of an annuity with a deferred period d, the benefit function is given by

$$B_i(t,\tau) = b\,\tau\,\mathbb{1}[t-\tau \geq d].$$

Prior to expressing premiums and reserves, we still have to specify a right-continuous counting process defined by $N_{\mathrm{ai}}(t) = \sharp\{\tau; X_{\tau^-} = e_1 = a, X_\tau = e_2 = i, \tau \in (0,t]\}$, providing the number of transitions from the 'active' state to the 'disabled' state during the time interval $(0,t]$. The link between this counting process and the instantaneous rate of transition to the disability state for an insured aged x at time 0 is given by the following relation:

$$\mathbb{E}\left[dN_{\mathrm{ai}}(t) \,\middle|\, (X_s, D_s)_{0\le s\le t}\right] = \lambda_{\mathrm{ai}}(x+t)\, \mathbb{1}\,[X_{t^-} = e_1]\, dt.$$

4.3 Premium calculation using the equivalence principle

In the framework defined *supra*, considering the lump sum $c_{ai}(\cdot)$ depends only on the time t elapsed since policy issue, the expected present value of benefits, i.e. the net single premium (below, SP stands for Single-Premium) is given by:

$$\begin{aligned}\mathrm{SP} &= \mathbb{E}\left[\int_0^n \mathrm{e}^{-\int_0^t r(u)\,du}\, \mathbb{1}\,[X_t = e_2]\left(dB_i(t, D_t) + c_{\mathrm{ai}}(t) dN_{\mathrm{ai}}(t)\right)\right] \\ &= \int_{t=0}^n {}_tp_x^{\mathrm{aa}}\ \lambda_{\mathrm{ai}}(x+t)\left(\int_{\varepsilon=t}^n {}_{\varepsilon-t}p_{x+t,0}^{\overline{\mathrm{ii}}}\ \mathrm{e}^{-\int_0^\varepsilon r(u)\,du}\, dB_i(\varepsilon, \varepsilon - t) + c_{\mathrm{ai}}(t)\right) dt.\end{aligned} \tag{6}$$

In case the benefits do not depend on the time spent in the 'disabled' state, (6) may be rewritten as

$$\mathrm{SP} = \int_{t=0}^n \mathrm{e}^{-\int_0^t r(u)\,du}\ {}_tp_x^{\mathrm{ai}}\left(dB_i(t) + c_{\mathrm{ai}}(t)\lambda_{\mathrm{ai}}(x+t)dt\right). \tag{7}$$

With most policies, the premiums are waived while the income benefit is being paid. By definition, the equivalence principle is fulfilled if and only if the expected present value (actuarial value) of future benefits is equal to the actuarial value of future premiums, i.e.

$$\mathrm{SP} = \int_{t=0}^n \mathrm{e}^{-\int_0^t r(u)\,du}\ {}_tp_x^{\mathrm{aa}}\ d\Pi_a(t). \tag{8}$$

When we neglect the lump sum payment $c_{ai}(\cdot)$ and assume that the premium rate and the annuity rate are constant, i.e.

$$\Pi_a(t) = \pi t \ \text{ and } \ B_i(t) = bt,$$

the premium level depends on the actuarial value of a unit-level premium paid in the 'active' state during the period $[0, n]$

$$\overline{a}^{\mathrm{aa}}_{x,\overline{n}|} = \int_0^n \mathrm{e}^{-\int_0^t r(u)\,du}\ {}_tp_x^{\mathrm{aa}}\ dt,$$

and should be linked to the actuarial value of a unit-level benefit paid in the 'disabled' state during the same coverage period

$$\overline{a}^{\mathrm{ai}}_{x,\overline{n}|} = \int_0^n \mathrm{e}^{-\int_0^t r(u)\,du}\ {}_tp_x^{\mathrm{ai}}\ dt,$$

since we obtain from the equivalence principle

$$\pi = b\,\frac{\overline{a}^{\mathrm{ai}}_{x,\overline{n}|}}{\overline{a}^{\mathrm{aa}}_{x,\overline{n}|}}.$$

Levikson, Frostig, & Bshouty [14] have built an evaluation algorithm for this constant premium rate, which does not require to use transition probabilities.

4.4 Reserves calculation

The future commitments on the part of the insurer and those of the insured are spread across time and it is impossible for the insurance company to exactly match its cash outflows with its cash inflows. Indeed, the premium paid by the insured precedes the potential benefits paid out by the insurer in virtue of the inversion of the production cycle. This intrinsic mismatch between inflows and outflows leads the insurer to report at time t his net commitment towards the insured in the form of benefit reserves, denoted $V(t)$, to fulfill his liabilities in the long run. The prospective reserve at time t is defined as the actuarial value i.e. the conditional expectation of the present value, given the path of the insured risk up to time t, of future benefits less the actuarial value of future premiums. At a future time t the story of the insured risk is summarized by the filtration generated by the stochastic processes $\{X_s,\ 0 \leq s \leq t\}$ and $\{D_s,\ 0 \leq s \leq t\}$. Any conditioning path can therefore be expressed by the 'finite information' given by the set of transition times and the set of states visited by the path. As for the premium calculation, we focus on the net reserves, in the sense that taxes, issue costs, overhead expenses such as claims management and claims administration expenses and even safety loadings have not been included in the cash flow projections. For any time t of evaluation, a specific prospective reserve is defined for each state which may be occupied by the risk at this time. When disability benefits are concerned, two prospective reserves must be defined, pertaining to the 'active' state and the 'disabled' state respectively. In the light of the above remarks, we can express the net prospective reserves at time t for an active policyholder aged x at policy issue as follows

$$\begin{aligned} V_a(t) &= \mathbb{E}\left[\int_t^{n+} \mathrm{e}^{-\int_t^z r(u)\,du}\, \mathbb{1}[X_z = e_2]\, dB_i(z, D_z) \,\middle|\, X_t = e_1\right] - \mathbb{E}\left[\int_t^{n} \mathrm{e}^{-\int_t^z r(u)\,du}\, \mathbb{1}[X_z = e_1]\, d\Pi_a(z) \,\middle|\, X_t = e_1\right] \\ &= \int_{z=t}^{n+} {}_{z-t}p_{x+t}^{\mathrm{aa}}\, \lambda_{\mathrm{ai}}(x+z) \left(\int_{\varepsilon=z}^{n} {}_{\varepsilon-z}p_{x+z,0}^{\overline{\mathrm{ii}}}\, \mathrm{e}^{-\int_t^{\varepsilon} r(u)\,du}\, dB_i(\varepsilon, \varepsilon - z)\right) dz - \int_{z=t}^{n} \mathrm{e}^{-\int_t^z r(u)\,du}\, {}_{z-t}p_{x+t}^{\mathrm{aa}}\, d\Pi_a(z). \end{aligned} \tag{9}$$

As regards the disabled prospective reserve at time t, we also need to take into account the time spent in disability:

$$\begin{aligned} V_i(t,s) &= \mathbb{E}\left[\int_t^{n+} \mathrm{e}^{-\int_t^z r(u)\,du}\, \mathbb{1}[X_z = e_2]\, dB_i(z, D_z) \,\middle|\, X_t = e_2, D_t = s\right] - \mathbb{E}\left[\int_t^{n} \mathrm{e}^{-\int_t^z r(u)\,du}\, \mathbb{1}[X_z = e_1]\, d\Pi_a(z) \,\middle|\, X_t = e_2\right] \\ &= \int_{z=t}^{n+} {}_{z-t}p_{x+t}^{\overline{\mathrm{ii}}}\, \mathrm{e}^{-\int_t^z r(u)\,du}\, dB_i(z, z-t+s) + \int_{z=t}^{n+} {}_{z-t}p_{x+t}^{\mathrm{ia}}\, \lambda_{\mathrm{ai}}(x+z) \left(\int_{\varepsilon=z}^{n} {}_{\varepsilon-z}p_{x+z,0}^{\overline{\mathrm{ii}}}\, \mathrm{e}^{-\int_t^{\varepsilon} r(u)\,du}\, dB_i(\varepsilon, \varepsilon - z)\right) dz \\ &\quad - \int_{z=t}^{n} \mathrm{e}^{-\int_t^z r(u)\,du}\, {}_{z-t}p_{x+t}^{\mathrm{aa}}\, d\Pi_a(z). \end{aligned}$$

4.5 Generalized Thiele's differential equations

Thiele's differential equation was discovered in 1875 for the case of a single-life assurance. It provides an expression of the reserve dynamics, written in differential form, which is particularly easy to interpret. This tractable differential transform can be extended and generalized to a broader disability insurance context. As mentioned above, two reserves must be analyzed by the insurance undertaking, depending on the state occupied by the insured. A Thiele's transformation can account for changes in reserves $V_a(t)$ and $V_i(t,s)$ over time.

Proposition 4.1 *Considering an insured aged x at time of policy issue, the differential of the prospective reserve pertaining to the 'active' state at time $t \notin Disc\big(\Pi_a\big)$*[6] *is given by*

$$dV_a(t) = r(t)\, V_a(t)\, dt + \pi_a(t)\, dt - \lambda_{ai}(x+t)\, dt\, \big(c_{ai}(t) + V_i(t,0) - V_a(t)\big) + \lambda_{ad}(x+t)\, dt\, V_a(t) \tag{10}$$

This differential equation is valid in any open interval $(t_{a,j}, t_{a,j+1})$, $j = 0, 1, \ldots, q-1$ and must satisfy the conditions

$$V_a(t_{a,j}) = V_a(t_{a,j}{}^{-}) + \Delta\Pi_a(t_{a,j}), \qquad j = 0, 1, \ldots, q. \tag{11}$$

(10) *and* (11) *determine the solution $V_a(\cdot)$ uniquely.*

Proof We turn back to the reserve given by (9). We are going to derive a backward differential equation through a direct backward construction. To achieve this, let us suppose the policy is in the 'active' state at time $t \notin \mathrm{Disc}\big(\Pi_a\big)$ so that any inflow $\Delta\Pi_a(t)$ vanishes. Now we can use the law of iterated expectation, conditioning on what happens in the interval $(t, t+dt]$: with probability $1 - \lambda_{\mathrm{a}\bullet}(x+t)\, dt + o(dt)$ the insured remains active and the conditional expected

[6] or, equivalently, at age $x + t$

value of the reserve is $e^{-r(t)dt}\ V_a(t+dt) - \pi_a(t)\ dt$; with probability $\lambda_{\mathrm{ai}}(x+t)\ dt + o(dt)$ the insured becomes disabled, and the conditional expected value is $c_{\mathrm{ai}}(t) + e^{-r(t)dt}\ V_i(t+dt, 0)$; with probability $\lambda_{\mathrm{ad}}(x+t)\ dt + o(dt)$ the insured dies, and the conditional expected value is 0, because the death risk is not covered. We gather these quantities as follows:

$$\begin{aligned} V_a(t) = (1 - \lambda_{a\bullet}(x+t)\ dt)\ e^{-r(t)dt}\ V_a(t+dt) - \pi_a(t)\ dt \\ + \lambda_{ai}(x+t)\ dt\left(c_{\mathrm{ai}}(t) + e^{-r(t)dt}\ V_i(t+dt, 0)\right) + o(dt) \end{aligned}$$

which may be rewritten, subtracting $V_a(t+dt)$ on both sides

$$\begin{aligned} V_a(t) - V_a(t+dt) = -(\lambda_{ai}(x+t) + \lambda_{ad}(x+t))\ dt\ e^{-r(t)dt}\ V_a(t+dt) + (e^{-r(t)dt} - 1)\ V_a(t+dt) - \pi_a(t)\ dt \\ + \lambda_{ai}(x+t)\ dt\left(c_{\mathrm{ai}}(t) + e^{-r(t)dt}\ V_i(t+dt, 0)\right) + o(dt) \end{aligned}$$

Now making $dt \to 0$ and observing that $e^{-r(t)dt} = 1 - r(t)dt + o(dt)$, one arrives at

$$dV_a(t) = r(t)\ V_a(t)\ dt + \pi_a(t)\ dt - \lambda_{ai}(x+t)\ dt\ \ (c_{\mathrm{ai}}(t) + V_i(t,0) - V_a(t)) + \lambda_{ad}(x+t)\ dt\ V_a(t)$$

which is exactly (10). ∎

Proposition 4.2 *Considering an insured aged x at time of policy issue, the differential of the prospective disabled reserve* $V_i(t,s)$ *to establish at time* $t \notin \mathrm{Disc}_s\left(B_i\right)$ *with a duration s since the disability onset may be expressed as follows:*

$$d_tV_i(t,s) = r(t)\ V_i(t,s)\ dt - d_sV_i(t,s) - b_i(t, t-s)\ ds - \lambda_{ia}(x+t, s)\ dt\left(V_a(t) - V_i(t,s)\right) + \lambda_{id}(x+t,s)\ dt\ V_i(t,s) \qquad (12)$$

This differential equation is valid in any open interval $(t_{i,s,j}, t_{i,s,j+1})$, $j = 0, 1, \ldots, q_s - 1$ *and must satisfy the conditions*

$$V_i(t_{i,s,j}) = V_i(t_{i,s,j}{}^{-}) - \Delta B_i(t_{i,s,j}, t_{i,s,j} - s), \qquad j = 0, 1, \ldots, q_s. \qquad (13)$$

(12) *and* (13) *determine the solution* $t \mapsto V_i(t,s)$ *(well-defined for* $t \geq s$*) uniquely.*

Proof Here again, we aim to derive a backward differential equation governing the evolution of the disabled reserve. To achieve this, let us suppose the policy is in the 'disabled' state at time $t \notin \mathrm{Disc}_s\left(B_i\right)$ so that any benefit jump term vanishes, with a duration s since the disability onset. Now, conditioning on what happens in a small time interval $(t, t+dt]$ we write

$$\begin{aligned} V_i(t,s) = (1 - \lambda_{i\bullet}(x+t,s)\ dt)\ e^{-r(t)dt}\ V_i(t+dt, s+dt) + b_i(t, t-s)\ dt \\ + \lambda_{ia}(x+t,s)\ dt\ e^{-r(t)dt}\ V_a(t+dt) + o(dt) \end{aligned}$$

which may be rewritten, subtracting $V_i(t+dt, s+dt)$ on both sides

$$\begin{aligned} \left(V_i(t,s) - V_i(t+dt,s)\right) + \left(V_i(t+dt,s) - V_i(t+dt,s+dt)\right) = -(\lambda_{ia}(x+t) + \lambda_{id}(x+t))\ dt\ e^{-r(t)dt}\ V_i(t+dt, s+dt) \\ + (e^{-r(t)dt} - 1)\ V_i(t+dt, s+dt) + b_i(t, t-s)\ dt \\ + \lambda_{ia}(x+t,s)\ dt\ e^{-r(t)dt}\ V_a(t+dt) + o(dt) \end{aligned}$$

Now making $dt \to 0$ and using $e^{-r(t)dt} = 1 - r(t)dt + o(dt)$, one arrives at

$$d_t V_i(t,s) + d_s V_i(t,s) = r(t)\ V_i(t,s)\ dt - b_i(t, t-s)\ dt - \lambda_{ia}(x+t,s)\ dt\ \ (V_a(t) - V_i(t,s)) + \lambda_{id}(x+t,s)\ dt\ V_i(t,s)$$

which provides (12). ■

The right-hand side of (10) exhibits positive and negative contributions determining the increment of the active reserve:

1. $V_a(t)$ bears interest at a rate $r(t)$, so that the reserve instantaneously increases with an amount equal to $r(t)V_a(t)dt$.
2. The active policyholder pays premiums for an amount of $\pi_a(t)\ dt$ from time t to time $t + dt$, which is added to the currently held reserves.
3. In the event of a transition from e_1 to e_2 at time t[7], the insurance undertaking pays the lump sum $c_{\mathrm{ai}}(t)$ and sets aside the reserve $V_i(t, 0)$, which reduces the level of the active reserve set apart at time t (i.e. when the policyholder is in the 'active' state). To supply the disabled reserve, he has at his disposal an amount of $V_a(t)$, since the insured is no longer in the 'active' state. As a result, the new conditioning information the insurer should take into account to determine and evaluate the (disabled) reserve is the fact that the insured is in e_2 at time t with a length of stay equal to zero.
4. In the event of a transition from the 'active' state to death, an amount of $V_a(t)$ becomes available to support the company's insurance business.

The relation (12) may be interpreted in a similar manner:

1. $V_i(t)$ generates interest at a rate $r(t)$, so that the reserve instantaneously increases with an amount equal to $r(t)V_i(t)dt$.
2. The insurer pays the annuity $b_i(t, t-s)\ dt$ between t and $t + dt$, the amount of which shall be deducted from the reserve.
3. The infinitesimal extension of the time spent in the 'disabled' state between t and t^+ reduces the reserve by an amount of $d_s V_i(t,s)$.
4. In the event of a transition from e_2 to e_1 (recovery) at time t with a disability duration equal to s, the insurance undertaking sets aside the reserve $V_a(t)$, which reduces the level of the disabled reserve set apart at time t (i.e. when the policyholder is in the 'disabled' state). To supply the active reserve, he has at his disposal an amount of $V_i(t,s)$, since the insured is no longer in the 'disabled' state.
5. In the event of a transition from disability to death after a time spent in this state equal to s, an amount of $V_i(t,s)$ becomes available to support the company's insurance business.

5 Procedures of simulation

5.1 Representation of the bivariate process $\mathcal{XD}$

From the process $\mathcal{XD} = \{(X_t, D_t), t \geq 0\}$, we can define the following processes:

1. a process of transition times $\{T_j,\ j = 0, 1, \dots\}$ defined by

$$T_j = \inf\{t > T_{j-1} \mid X_t \neq X_{T_{j-1}}\},\ j \geq 1,$$

   with $T_0 = 0$. As a result we get $T_0 < T_1 < T_2 < \dots$ almost surely.

2. a process of sojourn times $\{Y_j,\ j = 0, 1, \dots\}$ defined by

$$Y_j = T_j - T_{j-1},\ j \geq 1,$$

[7] or, equivalently, at age $x + t$ when we consider an insured active at t and aged x at time of policy issue

with the convention $Y_0 = 0$. Y_j as thus defined represents the time spent in the state $X_{T_{j-1}}$ between (j-1)th transition and jth transition, i.e. between T_{j-1} and T_j. Random variables $Y_j,\ j = 1, 2, \dots$ are almost surely non-negative.

We see immediately that

$$T_j = \sum_{k=1}^{j} Y_j,\ j = 1, 2, \dots$$

As a result, the information provided by the path $(X_t, D_t)(\omega)$ is equivalent to the 'finite information' given by one of the two following bivariate processes:

1. $\left\{(T_j, X_{T_j}),\ j = 0, 1, \dots\right\}$
2. $\left\{(Y_j, X_{T_j}),\ j = 0, 1, \dots\right\}$

The first bivariate process gives the set of transition times and the set of states visited by the path. The second bivariate process provides the sojourn times between each transition and the states visited. Thus, the process $\mathcal{XD}$ is equivalent to either of these two bivariate processes.

5.2 General simulation procedure

Previous expressions are helpful to simulate sample paths from a Markov stochastic bivariate process $\mathcal{XD}$, i.e. from a Semi-Markov univariate process $\mathcal{X}$. Given the initial state $X(0)$ occupied by the risk at policy issue, we just need to simulate a sequence of sojourn times and states visited. Let us suppose that at time T_n, i.e. at nth transition, the risk occupies the state e_{i_n}. Then, we can provide the distribution of the sojourn time in the state e_{i_n} given that the latest transition to e_{i_n} occurred at time T_n, which is equivalent to express the following conditional survival function, pertaining to the random variable Y_{n+1}

$$\Pr\ [Y_{n+1} > t \mid X_{T_n} = e_{i_n}, T_n = t_n], \quad i_n = 1, 2, 3, \tag{14}$$

Then we have to apply an inverse transformation method to the cumulative distribution function arising from (14) to get a realization of T_{n+1} using the standard uniform distribution on the closed interval $[0, 1]$. As usual, inversion is easier to achieve when the instantaneous transition rates are piecewise constant.

At the end of the sojourn time y_{n+1}, we have to simulate the destination state of the insured. If the insured leaves the state e_{i_n} at time t_{n+1}, he will then visit the state e_j with probability

$$\frac{\lambda_{i_n j}(t_n, t_{n+1} - t_n)}{\lambda_{i_n \bullet}(t_n, t_{n+1} - t_n)}$$

Obviously, these quantities of interest must be previously estimated in a robust and reliable manner, because otherwise the actuary could not simulate any path. Indeed, we need to partition the closed interval $[0, 1]$ into subintervals such as the length of each subinterval corresponds to the probability of one of the events that can occur. Since the sum of the probabilities of all the events is exactly equal to one, each subinterval corresponds to one, and only one, event. As a further step, we can then draw a new random sample from the standard uniform distribution on $[0, 1]$ and determine the sub-interval of $[0, 1]$ which contains this realization to get a simulation of the destination state.

Finally, we can summarize the simulation procedure as follows:

Step 1 We generate a random draw u from the standard uniform distribution on the closed interval $[0, 1]$ and determine the initial state from the vector $q = (q_1, q_2, q_3)$ whose ith component gives the probability that the risk is in the state e_i at policy issue, $0 \le i \le 3$. To achieve this, we divide the closed interval $[0, 1]$ using the partition provided by q. Let us denote $(\vartheta)_{0 \le i \le 3}$ this subdivision, where $\vartheta_0 \equiv 0$, $\vartheta_1 = q_1$, $\vartheta_2 = q_1 + q_2$, and $\vartheta_3 \equiv 1$. Now all we need to do is to determine the sub-interval of $[0, 1]$ which contains u. If $u \in [0, \vartheta_1]$ the initial state is e_1. If $u \in [\vartheta_1, \vartheta_2]$ the initial state is e_2. Otherwise, the insured is in the state e_3 at policy issue.

Step 2 We simulate the first sojourn time y_1 in the state e_{i_0} which has been previously selected. To achieve this, we apply an inversion to the c.d.f. of the underlying random duration.

Step 3 We simulate the destination state e_{i_1} of the process at the end of the first sojour time y_1. To do this, we generate a random draw from a uniform distribution and use the partition arising from the probabilities $\lambda_{i_0 j}(t_1, y_1)/\lambda_{i_0 \bullet}(t_1, y_1)$, $j \neq e_{i_0}$.

Step 4 We go back to step 2 to simulate the second sojourn time y_2. The state e_{i_1} and the sojourn time y_2 replace e_{i_0} and y_1, respectively.

⋮

Final step End of the path when death occurs or at the expiration of the contract.

6 Simulation of the three-state model

The considered time unit in the model detailed below is the day: we assume t denotes a number of days.

6.1 Formal description of the model

This model aims at getting the usual Actuarial Present Values, namely the premium and the reserves, in accordance with the characteristics of the insured (gender, age, sector of activity, type of working area). Different tariff cells have been designed:

- for the return to work from an absence due to sickness or accident: from the crossing of the mutually exclusive and exhaustive categories included in the grouped variables involved in the statistically selected Cox model with unobserved heterogeneity displayed on the table 1.
- for the transition from work to disability (work stoppage incidence): from the crossing of the categories defining grouped variables involved in the statistically selected Poisson model fitting our data. It is displayed on the table 2.
- for the transition to death from any other state: from the exhaustive crossing of the different ages and the two categories splitting the gender variable. Due to data limitations, rating factors such as the sector of activity and the type of region in which the considered employee works have not been included in the variables predicting mortality. Only age and gender - with the TH00-02 and TF00-02 tables - have an influence on mortality.

Due to our modeling, the nature of the underlying stochastic processes will vary depending on the considered transition:

- in the case of the "disabled → active" transition, the use of a Cox model and the inclusion of the age at the onset of the work stoppage in the set of explanatory variables lead to a Semi-Markov process. Indeed, the return-to-work rates depend not only on the date of disability inception but also on the time elapsed since the latest transition in the disability state. Moreover, the process is also locally homogeneous since we have designed age categories which influence the Cox output through constant coefficients.
- in the case of the "active → disabled" transition, the survival function in the 'active' state has an exponential form. Indeed, the waiting times between jumps[8] of a Poisson process follow exponential distributions. The intensity of the counting process (Poisson in our framework) is piecewise constant. In fact, in each tariff cell, defined by the crossing of an age group, a gender, a region of work and an activity sector, the transition intensity is constant. We have thus modelled a locally Time-Homogeneous Markov process. Using piecewise constant intensities is an effective way of being consistent with the required time-inhomogeneity property[9] while preserving the tractability of constant intensities.
- concerning the risk of dying from the active state or from disability, transition intensities only depend on the discrete age variable and the gender. Thus, the underlying process is again locally Homogeneous Markov.

	Mean	coef	exp(coef)	se(coef)	Pr(>|w|)
M	0.542	0	1 (ref)		
F	0.458	-0.025	0.976	0.006	0.000
Age-36-40	0.143	0	1 (ref)		
Age-16-25	0.029	0.216	1.241	0.015	0.000
Age-26-30	0.088	0.113	1.120	0.011	0.000
Age-31-35	0.120	0.041	1.042	0.010	0.000
Age-41-45	0.157	-0.048	0.953	0.010	0.000
Age-46-50	0.155	-0.113	0.893	0.010	0.000
Age-51-55	0.171	-0.218	0.804	0.011	0.000
Age-56-65	0.136	-0.294	0.745	0.012	0.000
Construction	0.097	0	1 (ref)		
Consumer goods industry ; Energy ; Building security services	0.120	-0.034	0.967	0.013	0.008
Publishing, audiovisual and media	0.036	0.036	1.037	0.017	0.035
Automotive industry	0.050	0.042	1.043	0.016	0.007
Mechanical Engineering ; Raw material	0.163	0.057	1.058	0.011	0.000
Food industry ; Finance & Insurance	0.096	0.058	1.060	0.013	0.000
Retail services ; Transport ; R&D; Consulting	0.218	-0.001	0.999	0.011	0.903
IT & Web Technologies	0.042	0.102	1.108	0.015	0.000
(Tele)Communications Equipment	0.033	0.050	1.051	0.018	0.006
Hotels & Restaurants	0.102	-0.245	0.782	0.014	0.000
Personal Services & Education	0.044	-0.071	0.931	0.017	0.000
Rural areas ; Medium-sized cities	0.651	0	1 (ref)		
Large cities	0.349	0.091	1.095	0.006	0.000

Table 1: Cox MPH (frailty gamma) model with Efron estimation.

The table 3 gives an overview of the final modellings for each possible transition. We specify that the disability incidence and severity distributions have voluntarily been estimated without splitting the disability state into different states[10] for the sake of simplicity and tractability of the final actuarial model we carry out. However, splitting the disability state into disablement by illness and disablement by accident might be strongly relevant. This approach would lead to a four-state model and applies the same logic as Cordeiro [4]. We now have to clarify the procedures which will enable us to get the different intensities and corresponding survival functions, according to the type of transition, from the aforementioned regression model outputs.

6.2 Calculation of the instantaneous transition rates from work to disability λ_{ai}

These rates have been estimated using the aforementioned Poisson model. The following relationship provides an explicit expression of these rates from the vector θ of the regression coefficients:

$$\lambda_{ai}(t \mid Z = z) = \exp(\theta' z) \tag{15}$$

where Z is the vector of regressors. Since the intensities are piecewise constant, we get the locally homogeneous Markov property. These rates depend only on age. The time spent in the active state does not impact them, in accordance with the Markov hypothesis. Disregarding the exit from the active state caused by death, we can define a "partial" survival

[8] in particular the waiting time until the first jump
[9] we have there a duration-since-initiation dependence through the age effect
[10] in particular without distinguishing between the sickness and accident causes

	Estimate	Std. Error	z value	Pr(>|z|)	Significance
(Intercept)	-1.0746	0.0110	-97.5670	0.0000	***
M	0				
F	0.3457	0.0084	40.9480	0.0000	***
Age-31-45	0				
Age-61-65	-0.2446	0.0368	-6.6497	0.0000	***
Age-16-25	0.2874	0.0207	13.8750	0.0000	***
Age-26-30	0.0313	0.0132	2.3678	0.0179	*
Age-46-50	-0.0432	0.0121	-3.5796	0.0003	***
Age-51-60	0.0449	0.0103	4.3420	0.0000	***
Construction ; Raw material	0				
Consumer goods industry	-0.0395	0.0183	-2.1613	0.0307	*
Publishing, audiovisual and media	-0.3630	0.0226	-16.0408	0.0000	***
Automotive industry ; Mechanical Engineering ; Hotels & Restaurants	0.0571	0.0125	4.5585	0.0000	***
Energy	-0.2555	0.0244	-10.4558	0.0000	***
Food industry	-0.1416	0.0286	-4.9543	0.0000	***
Retail services	-0.3032	0.0169	-17.9065	0.0000	***
Transport ; Building security services	0.1785	0.0157	11.3524	0.0000	***
Finance & Insurance	-0.2800	0.0168	-16.6795	0.0000	***
IT & Web Technologies	-0.5796	0.0302	-19.1805	0.0000	***
R&D ; Consulting	-0.2608	0.0192	-13.5925	0.0000	***
(Tele)Communications Equipment	-0.2302	0.0226	-10.1774	0.0000	***
Personal Services & Education	-0.4039	0.0209	-19.3499	0.0000	***
Rural areas	0				
Medium-sized cities	-0.0822	0.0108	-7.6368	0.0000	***
Large cities	-0.1714	0.0095	-18.0710	0.0000	***

Table 2: Poisson Model.

Transition	Modeling	Resulting Process
Disablement	Poisson	Locally Time-Homogeneous Markov renewal
Recovery	Cox MPH	Locally Time-Homogeneous Semi-Markov
Death of an active life	Mortality Tables	Locally Time-Homogeneous Markov renewal
Death of a disabled life	Mortality Tables	Locally Time-Homogeneous Markov renewal

Table 3: Overview of our multi-state model.

function in the active state, for any covariates vector z:

$$\Pr\ [T_a > t \mid Z = z] = \exp\ \Big(-\exp(\theta' z)\ t \Big).$$

These incomplete survival functions[11] are shown in the table 4 where we have restricted ourselves to the "Finance & Insurance" and "Large cities" categories to highlight the sensitivity to age group and gender. The true survival function will be provided by adding the daily mortality rates, defined below. In fact, the required mathematical operation amounts to multiplying the specific survival functions for each transition, so that the transition model from the active state can be construed as an independent competing risks[12] model, where the two competing risks are the risks of becoming disabled and dying. The term that precedes t in the exponential functions of the table 4 is the (constant) daily transition intensity,[13] specific to each set of covariates.

[11] in the sense that the mortality risk is not yet taken into consideration

[12] i.e. that produce cause-specific hazards

[13] we have divided the annual intensity given by the Poisson model - previously estimated using a one-year exposure - by 365.25 to get the desired daily intensity.

Age Group	Gender	
	F	M
16-25	$\exp(-0.001331\ t)$	$\exp(-0.000942\ t)$
26-30	$\exp(-0.001030\ t)$	$\exp(-0.000729\ t)$
31-45	$\exp(-0.000998\ t)$	$\exp(-0.000707\ t)$
46-50	$\exp(-0.000956\ t)$	$\exp(-0.000677\ t)$
51-60	$\exp(-0.001044\ t)$	$\exp(-0.000739\ t)$
61-66	$\exp(-0.000782\ t)$	$\exp(-0.000553\ t)$

Table 4: Probability (rounded to 6 digits after the point) of remaining in the active state more than t days, depending on the age category and the gender, in the "Finance and Insurance" sector and the Large cities, under the assumption that only one type of transition from the active state is possible, i.e. becoming disabled.

Age Category	Gender	
	F	M
16-25	1.404948	1.440514
26-30	1.267442	1.299527
31-35	1.179393	1.209250
36-40	1.132016	1.160673
41-45	1.078963	1.106277
46-50	1.011061	1.036656
51-55	0.910283	0.933327
56-66	0.843665	0.865022

Table 5: Values (rounded to 6 digits after the point) of the scores $\exp(z'\beta)$ emanating from the final Cox model on the recovery hazard rate, depending on gender and age, for the "Finance & Insurance" sector in Large cities. The greater the score is the faster the recovery is.

6.3 Calculation of the instantaneous recovery rates λ_{ia}

These rates have been estimated thanks to a Cox MPH (*Mixed Proportional Hazard*) model with unobserved heterogeneity, which allows us to model the instantaneous recovery rate as follows:

$$\lambda_{ia}(t \mid Z = z, v) = \lambda_{0,ia}(t)\ \exp(z'\beta)\ v,$$

where $\lambda_{0,ia}(\cdot)$ is the baseline hazard function, v is the *frailty* term, which follows a Gamma distribution with expectation equal to 1, and z is the considered covariates vector. These rates are dependent both upon the age and upon the duration of presence in the disabiliy state, in accordance with the desired Semi-Markov framework. We then show in the table 5 the values of $\exp\ (z'\beta)$, depending on gender and age group, for the "Finance & Insurance" sector in Large cities.

6.4 Calculation of the mortality intensities $\lambda_{ad} = \lambda_{id}$

In accordance with the International Association of Actuaries (IAA) standard notations, we denote ℓ_x an estimate (at time 0) of the number of individuals alive at age x ($x = 1, 2, \ldots, \omega$), out of the initial 'cohort' of ℓ_0 individuals. As ω denotes the (integer) maximum age, we have by definition $\ell_\omega > 0$ and $\ell_{\omega+1} = 0$. From ℓ_x, we easily deduce the number d_x of individuals dead at age x

$$d_x = \ell_x - \ell_{x+1},$$

and the annual death rate at age x

$$q_x = \frac{d_x}{\ell_x},$$

which provides the annual survival probability at age x, denoted p_x:

$$p_x = 1 - q_x = \frac{\ell_{x+1}}{\ell_x}.$$

We can finally introduce the instantaneous mortality rate at age x, denoted μ_x (IAA standard again), expressed in annualised form as follows:

$$\mu_x = -\frac{1}{\ell_x}\frac{d\ell_x}{dx}.$$

These instantaneous mortality rates will then enable us to write the survival probability between ages x and $x + c$ using the following relation :

$${}_c p_x = \frac{\ell_{x+c}}{\ell_x} = \exp\Big(-\int_0^c \mu_{x+s}\, ds\Big).$$

These relations can be efficiently used with regard to the simulation phase. Insofar as we have chosen to write the successive sojourn times on a daily basis, we need to provide daily hazard rates $\lambda_{id} = \lambda_{ad}$ between 0 and t so as to estimate the two targeted probabilities of survival - in the active and disability states - up to t days. The method we put forward here is easy to carry out and is premised on the assumption that the daily mortality intensities are constant between two successive ages. So now we implicitly use a daily unit time frame for our intensities and survival probabilities. We then approximate the integral $\int_0^t \lambda_{id}(x \times 365.25 + s)ds$ using the sum $\sum_{s\leq t} \hat{\lambda}_{id}(x \times 365.25 + s)$, where $\hat{\lambda}_{id}(x \times 365.25 + s)$ is the daily mortality rate on the sth day following the date of the xth birthday of the representative policyholder. We finally use the assumption of constant daily instantaneous mortality rates between two succesive ages to get the quantity of interest, i.e. an estimation of these daily intensities using the q_x. We do indeed have the following conversion to a daily unit:

$$q_x = 1 - \exp\Big(-\int_0^{365.25} \lambda_{id}(x \times 365.25 + s)\, ds\Big) \underbrace{\approx}_{\text{by assumption}} 1 - \exp\ (-365.25 \times \hat{\lambda}_{id}(x \times 365.25))$$

which provides the estimation

$$\hat{\lambda}_{id}(x \times 365.25) = -\frac{\ln\ (1 - q_x)}{365.25}. \tag{16}$$

This approach is approximately equivalent to calculating daily mortality rates, that we denote $q_{x,j}$, where j stands for the day, from the rates q_x. Indeed, using again the assumption of constant daily mortality intensities between two consecutive ages, and supposing the daily survival probabilities are independent, we get the following equivalence:

$$1 - q_x = (1 - q_{x,j})^{365.25} \Longleftrightarrow q_{x,j} = 1 - (1 - q_x)^{1/365.25}.$$

We apply the approach (16) using the certified French mortality - TF00-02 and TH00-02 - tables. As mortality estimation is beyond the scope of this paper, this process did not strike us as being particularly inappropriate, even if these tables are no longer the ideal reflection of the current French mortality.

Graph 2 depicts both annual and daily mortality rates, depending on age and gender, from the TF00-02 and TH00-02 tables and the formula (16).

6.5 Construction of the survival probabilites and formulation of the Smirnov transforms

Having estimated the intensities that are specific to each type of transition, we can easily calculate the survival probabilities given by the relations (1) and (2). However, the time unit used in (1) and (2) is the year, while our considered three-state model is built upon a daily time unit. This problem can be readily handled by a straightforward linear transformation, which enables us to transpose the aforementioned relations into a framework based on daily time unit. Indeed, we just have to replace the annual age x in (1) and (2) by $x \times 365.25$. As a consequence of this substitution $t \to 365.25 \times t$, we must appropriately transform the interval of integration.

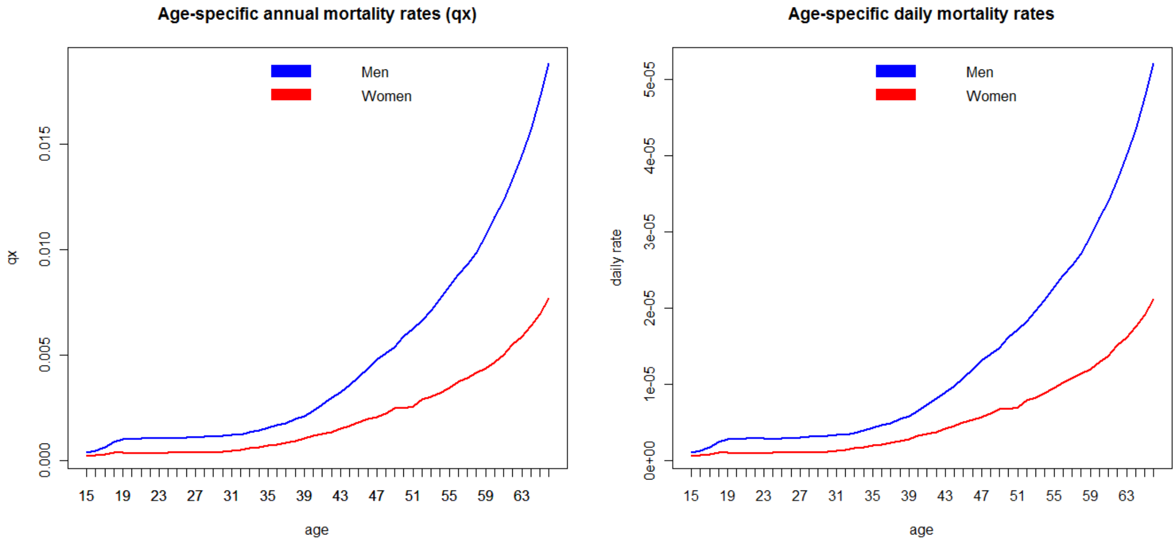

Figure 2: Annual and Daily mortality rates, depending on age and gender, and calculated from the TF00-02 and TH00-02 tables.

We can then convert the integrals (cumulative hazards) in (1) and (2) to finite sums since the all-cause exit rates (i.e. whatever the cause) from the active state $\lambda_{a\bullet}$, and from the disability state $\lambda_{i\bullet}$, are constant over daily intervals. To achieve that, we just have to divide the interval of integration of the cumulative hazard into daily intervals over which the hazard function is assumed to take a constant value. Then, using this partition and the additivity of integration on intervals (Chasles relation), we get a sum equal to the initial integrated hazard.

Compared with the relations (1) and (2), we now include the exact age[14] (as a continuous variable) at the time at which latest transition to the current state occured in the covariates vector Z. This choice enables us to alleviate notations and lies in the fact that the explanatory variables, in both the disability inception (Poisson) and the recovery (Cox) models, but also in the mortality model, take into account the real-valued age at the first day of the sojourn in the current state.

Let us now denote T^a and T^i the duration random variables related to the sojourn times in the active state and the disability state, respectively. We also denote $\hat{\lambda}_{ia}(t \mid Z)$ the estimated transition rate from the disability state to the active state for a policyholder described by a covariates vector Z at time t. The presence of the age at latest transition in the regressors Z and the sojourn duration t ensure the conditions required to set up a Semi-Markov framework. Then let $\hat{\lambda}_{ai}(t \mid Z)$ and $\hat{\lambda}_{ad}(t \mid Z) = \hat{\lambda}_{id}(t \mid Z)$ be the other estimated transition rates. These intensities are locally constant[15] and independent from the duration of presence in the active state or in the disability state. We get back to the Locally Homogeneous Markov framework. The following result holds:

Proposition 6.1 *Given a representative insured described by a vector z and the framework detailed supra, Smirnov transforms providing holding times in the respective transient states may be expressed as follows:*

$$(F_z^a)^{\leftarrow}(u) = \min \left\{ t \in \mathbb{N} \,\middle|\, \sum_{s \in \mathbb{N} \,|\, s \leq t} \exp(\hat{\theta}' z) + \hat{\lambda}_{ad}(s \mid Z = z) \geq -\ln\, u \right\}, \tag{17a}$$

$$(F_z^i)^{\leftarrow}(u) = \min \left\{ t \in \mathbb{N} \,\middle|\, \sum_{s \in \mathbb{N} \,|\, s \leq t} \hat{\lambda}_{0,ia}(s)\, \exp(z' \hat{\beta}) + \hat{\lambda}_{id}(s \mid Z = z) \geq -\ln\, u \right\}. \tag{17b}$$

Proof We can estimate the probability of survival in the active state Pr $[T^a > t \mid Z]$ by the estimator

[14] and no longer the attained age at the birthday immediately prior to the date of the latest transition

[15] constant between two consecutive ages for the mortality and constant within age categories for the disablement process. Other covariates have not been supposed time-varying

$$\exp\left(-\sum_{s\in\mathbb{N}\ |\ s\leq t}\hat{\lambda}_{ai}(s\mid Z)+\hat{\lambda}_{ad}(s\mid Z)\right)=\exp\left(-\sum_{s\in\mathbb{N}\ |\ s\leq t}\exp(\hat{\theta}'Z)+\hat{\lambda}_{ad}(s\mid Z)\right),\tag{18}$$

and use the following estimator for the survival probability in the disability state Pr $[T^i > t \mid Z]$

$$\exp\left(-\sum_{s\in\mathbb{N}\ |\ s\leq t}\hat{\lambda}_{ia}(s\mid Z)+\hat{\lambda}_{id}(s\mid Z)\right)=\exp\left(-\sum_{s\in\mathbb{N}\ |\ s\leq t}\hat{\lambda}_{0,ia}(s)\ \exp(Z'\hat{\beta})+\hat{\lambda}_{id}(s\mid Z)\right).\tag{19}$$

We can easily simulate sojourn durations from the estimators (18) and (19). We just have to write the appropriate inverse transform algorithm. To achieve this, we denote t^a_z and t^i_z realizations of the random variables whose respective distributions are the conditional probability distribution of T^a given $Z = z$ and the conditional probability distribution of T^i given the occurence of the value z of Z. We denote these distributions dF^a_z and dF^i_z respectively. Furthermore, let u be a random value drawn from the standard uniform distribution on the closed interval $[0, 1]$. Using the generalized inverse cumulative distribution function finally provides

$$(F^a_z)^{\leftarrow}(u)=\min\left\{t\in\mathbb{N}\ \middle|\ 1-\exp\left(-\sum_{s\in\mathbb{N}\ |\ s\leq t}\hat{\lambda}_{ai}(s\mid Z=z)+\hat{\lambda}_{ad}(s\mid Z=z)\right)\geq u\right\},$$

$$(F^i_z)^{\leftarrow}(u)=\min\left\{t\in\mathbb{N}\ \middle|\ 1-\exp\left(-\sum_{s\in\mathbb{N}\ |\ s\leq t}\hat{\lambda}_{0,ia}(s)\ \exp(z'\hat{\beta})+\hat{\lambda}_{id}(s\mid Z=z)\right)\geq u\right\}.$$

Thanks to a logarithmic transformation these expressions can be rewritten as follows

$$(F^a_z)^{\leftarrow}(u)=\min\left\{t\in\mathbb{N}\ \middle|\ \sum_{s\in\mathbb{N}\ |\ s\leq t}\exp(\hat{\theta}'z)+\hat{\lambda}_{ad}(s\mid Z=z)\geq-\ln\ (1-u)\right\},$$

$$(F^i_z)^{\leftarrow}(u)=\min\left\{t\in\mathbb{N}\ \middle|\ \sum_{s\in\mathbb{N}\ |\ s\leq t}\hat{\lambda}_{0,ia}(s)\ \exp(z'\hat{\beta})+\hat{\lambda}_{id}(s\mid Z=z)\geq-\ln\ (1-u)\right\}.$$

which provide the desired results, since $U \overset{d}{=} 1-U$. ■

6.6 Implementation of the trajectories

We now have to follow scrupulously the simulation procedure which was previously described. To initialize the algorithm, we use the classical and convenient assumption[16] that $q_1 = \vartheta_1 = 1$, which entails that all the trajectories will start in the active state.

Step 1 We generate a random draw from the standard uniform distribution on the closed interval $[0, 1]$ and simulate a first duration of sojourn in the active state using the generalized inversion (6.5);

Step 2 Given the first duration of presence in the active state $t_1 = t_1 - t_0 = y_1$, we update the age[17] and simulate the destination state. To achieve this, we divide the closed interval $[0, 1]$ using the probabilities $\lambda_{e_1\ j}(t_1, y_1)/\lambda_{e_1\bullet}(t_1, y_1)$, $j \neq e_1$. We then generate a new random draw from the standard uniform distribution on $[0, 1]$ and we determine the sub-interval of $[0, 1]$ given by the aforementioned partition, which contains this realization. We derive the destination state e_{i_1} in order to simulate the duration of the second stay of our trajectory (eternity if the draw is included in the sub-interval of death);

Step 3 If the insured becomes disabled at the date t_1, we generate a new random draw from the standard uniform distribution on $[0, 1]$ and simulate the second duration of stay of the trajectory (in the disability state) thanks to the inverse transform formula (6.5);

⋮

[16] an alternative choice would have been to estimate a probability to be disabled at policy issue using both the information about incidence and severity (duration) for the representative insured defined by his initial covariates vector

[17] the initial age is incremented by the time spent in the active state

Final step End of the trajectory, because of death or after the expiration of the insurance policy. Move on to the next trajectory.

6.7 Benefit function and assumptions

6.7.1 Benefit function

We remind that the time unit considered is the day. Policies normally include a deferred period so that the benefit will not begin to be payable until the work stoppage has lasted a certain minimum period[18]. As health insurance policies are usually intended to supplement the sickness benefit available from an employer or payable from the state (e.g. National Insurance in the UK, French Social Security System in France), the deferred period will tend to reflect the length of time after which these benefits reduce or cease. Obviously, the longer the deferment period is, the lower the insurance premium is.

Sometimes, cover policy conditions include a maximum benefit period. After that maximum period, other conditions may specify the benefits entitlement of the insured. For the sake of simplicity, we will suppose only one set of conditions, so that the maximum period of annuity payment depends only on the expiration date specified in the contract.

Furthermore, we introduce a ceiling value, which caps the benefit to be paid to the policyholder when he/she has been disabled for longer than the deferment period. This ceiling value (CV) is an increasing function of time. Indeed, the policy conditions usually include a clause to limit the income provided by the policy, which also reduces moral hazard.

This can also be highly desirable to introduce a maximum benefit payable from the State (BPFS). This threshold has to be deducted from the global targeted benefit by the insurer before calculating and setting the premium level. We suppose BPFS increases as time elapses.

Maybe above all, the size of the income benefit is generally linked to the insured's normal level of income. At policy issue the salary depends on the considered tariff category or covariates vector z of the representative insured (real-valued age x, in years, at policy issue, gender, sector of activity, work area) and is therefore assigned the vector $z \equiv (w, x)$ as subscript, where w summarizes the information provided by the gender, the sector of activity and work area at policy issue. Salary will also be a function of time[19]. We will denote by Salary the annual salary of the representative insured.

Now let $B_{i,w,x}(t,s)$ be the benefit a representative insured whose covariates vector at policy issue is (w, x) will receive up to time t if he has been disabled at t for $t - s$ days. $B_{i,w,x}(t,s)$ depends on the date[20] $s \le t$ of accident or sickness onset, given the real-valued age x at time 0 of policy issue, and on the duration of the disability spell[21]. Furthermore, $B_{i,w,x}(t,s)$ depends on two left-continuous functions, denoted $a_{w,x}(\cdot)$ and $b_{w,x}(\cdot)$, which are percentage levels. $a_{w,x}(\cdot)$ is the targeted replacement rate, provided that the salary does not exceed the ceiling value. $b_{w,x}(\cdot)$ is a certain percentage level consistent with a State-guaranteed minimum income in case of disability, on condition that the salary is not above the maximum benefit payable from the State. $a_{w,x}(\cdot)$ and $b_{w,x}(\cdot)$ may depend on the risk characteristics z involved and on the duration since the disability onset Finally, the value at time s of the payment stream up to time t is given by

[18] the deferred period or deferment period, usually 4, 13 or 26 weeks in the UK, on average 13 weeks in France. In the USA, the disability income policy, often called "loss of time" or "loss of income" or "long term disability", provides payments when an insured is unable to work because of injury or sickness. Policy designs are similar to those in the UK, altough a different nomenclature is used; for example "(benefit) elimination" or "waiting period" are commonly used, instead of "deferred period". It is also common to include a maximum benefit period in the event of the total disability of the insured policyholder. Benefits are normally paid monthly as fixed payments while the insured is disabled. As in the UK, this insurance is available on both an individual and a group basis. In group coverages the benefit amount is related directly to the insured's earnings, at least in part. In individual coverages the relationship is much more loose and even then only exists at time of policy issue. Benefits may be coordinated with social security, decreasing if social security benefits are received. Policies are often on a stand-alone basis and as riders to other types of long term insurance policy; however the second possibility is rare nowadays. For further information readers should consult Black & Skipper [3] and O'Grady [17]

[19] CV, BPFS and Wage will be indexed according to a fixed rate for our application.

[20] or equivalently, the real-valued age at disability inception

[21] obviously, $B_{i,w,x}(t,s)$ is equal to zero when the time spent in the disability state is lower than the deferred period of the policy, F

$$B_{i,w,x}(s,t) = \int_s^t e^{-\int_s^u r(h)\,dh} \left(a_{w,x}(u-s) \times \frac{\min\left(\text{Salary}_{(w,x)}(s), \text{CV}(u)\right)}{365.25} - b_{w,x}(u-s) \times \frac{\min\left\{\text{Salary}_{(w,x)}(s), \text{BPFS}(u)\right\}}{365.25} \right) \times G(s,u)\ \mathbb{1}\ [F \leq u-s \leq \Xi - (x * 365.25 + s)]\ du \tag{20}$$

where $G(s,u)$ is the indexing factor of the benefit for the period $[s,u]$, $a_{w,x}(\cdot)$ and $b_{w,x}(\cdot)$ are such that the expression enclosed between parentheses is non negative, F is the deferred period in days, Ξ is the maximum age (in days) at which the insured is entitled to benefits. A more intuitive form can be obtained assuming that $a_{w,x}$ and $b_{w,x}$ are constant and neglecting discounting and indexing problems, namely:

$$B_{i,w,x}(s,t) = \left(a_{w,x} \times \frac{\min\left\{\text{Salary}_{(w,x)}, \text{CV}\right\}}{365.25} - b_{w,x} \times \frac{\min\left\{\text{Salary}_{(w,x)}, \text{BPFS}\right\}}{365.25} \right) \times \min\Big((t-s-F)\,;(\Xi - (x*365.25+s) - F)\Big)\ \mathbb{1}\ [F \leq t-s]$$

Assuming now $F = 0$, we have for any $(s,t) \in \mathbb{R}^2$ such that $t - s \leq \Xi - (x * 365.25 + s)$

$$B_{i,w,x}(s,t) = \left(a_{w,x} \times \frac{\min\left\{\text{Salary}_{(w,x)}, \text{CV}\right\}}{365.25} - b_{w,x} \times \frac{\min\left\{\text{Salary}_{(w,x)}, \text{BPFS}\right\}}{365.25} \right) \times (t-s)$$

6.7.2 Assumptions about discount rate and benefits indexing

Since our model requires to evaluate future benefits cash flows, we need to model a discount-rate process $v(t)$. From a financial point of view, we have chosen to adopt a conventional approach, in the sense that a fixed, non-stochastic interest rate i is considered. As a result, both the annual discount rate $v = (1+i)^{-1}$ and the instantaneous interest rate $r = \ln(1+i)$ will be constant. Future estimated cash flows will be discounted using this constant annual interest rate.

As far as the salary is concerned, we also need to model a revaluation process. Even if our model is fully compatible with a stochastic revaluation process, we have also chosen a convenient constant annual indexation rate ζ.

Moreover, it seems reasonable to assume that the ceiling value CV and the maximum benefit payable from the State BPFS are also subject to an annual review. For the sake of simplicity we have followed the classical approach of constant annual indexation rates, denoted η and ς, respectively.

Last but not least, the benefit itself can be indexed according to an arbitrary time-varying rate. It can be relevant to introduce this type of indexing because the reference salary to be used to compute the benefit will not generally increase with the time spent in disability. For instance, the benefit provided by (20) depends only on the salary at disability inception, and nothing guarantees that the benefits will be increased annually and will *a fortiori* keep up with inflation as long as the disability continues. In fact, insofar as the maximum benefit payable from the State is upgraded, the benefit from the insurer to the insured might even decrease without such an *ad hoc* adjustment, depending on whether or not the salary at disability inception is lower than the (indexed) ceiling value. For the sake of simplicity, we will assume a constant annual rate of indexing g for the numerical application.

Thus, different types of indexation are carried out annually. We emphasize that our model is flexible enough to replace any fixed annual indexation rate by a variable rate or even a stochastic rate. Obviously, we can test any deterministic scenario of indexation. As far as the stochastic modeling is concerned, we could, for example, carry out a "one-factor model" such as the Vasicek model, or the model of Cox, Ingersoll & Ross who have built a function of the short rate to model the zero coupon yield curve. An alternative choice might be to follow Heath, Jarrow & Morton who suggest to use a similar approach to that being implemented by Black & Scholes with the risk-neutral pricing of European actions, to model the zero coupon market price. Another possibility would be to apply the Markov model of Hull & White which provides a closed-form solution to the short rate stochastic differential equation and a closed form for the zero coupon price. We could also use the Cairns stochastic model which ensures the positivity of a Markov homogeneous zero coupon price process.

7 Application

7.1 Simulations and first statistics

We now apply our simulation model to a representative man aged 25 at policy issue. We suppose he works in a large city and in the sector "Finance & Insurance". His annual salary is set to $ 65,000. We suppose the policy expires at the retirement date of the representative insured, which means our coverage period is 43 years from the aforementioned policy issue, with a retirement age of 68. For our application, the benefit function which has been selected is based on a deferment period of 13 weeks, in accordance with a usual practice for insurers. The other required values to simulate our three-state trajectories are summarized in the table 6.

Variable	Symbol	Value at policy issue
Age at policy issue	x	25
Expiry/Upper age limit for claiming disability allowance	Ξ	68
Gender	w_1	M
Sector of activity	w_2	Finance & Insurance
Type of work area	w_3	Large Cities
Annual income	$\text{Salary}_{w,x}$	$ 65,000
Annual indexation rate (Salary)	$\zeta_{w,x}$	3 %
Annual indexing rate (Benefit)	$g_{w,x}$	3 %
Deferred Period	F	91 days
Constant interest rate	i	3 %
Targeted replacement rate	$a_{w,x}$	85 %
State-guaranteed replacement rate	$b_{w,x}$	50 %
Ceiling Value	CV	$ 450,000
Maximum Benefit Payable From the State	BPFS	$ 53,000
Annual indexing rate (CV)	η	1.5 %
Annual indexing rate (BPFS)	ς	2.5 %

Table 6: Parameter values.

We generate n = 100 000 trajectories for this representative agent. Each trajectory leads to a total discounted cost C for the insurer, using the benefit function (20). We can compute the mean of these 100 000 total discounted costs to get an estimate of the expectation of the discounted future benefits and, as a result, an estimate of the net single premium defined by (6) over the considered coverage period.

Drawing on the approach of Stenberg, Manca & Silvestrov [24][22] in our three-state framework, we can estimate different central moments -up to the fourth- from our trajectories. Our n simulations lead to a sample $(C_i)_{i=1,\dots,n}$ of total discounted costs.

Figure 3 depicts the empirical distribution of the 100 000 total discounted costs we have simulated. The mass of the distribution is concentrated on the left, and the right tail is longer: the skew is positive.

Table 7 shows empirical estimates of the expectation, standard deviation, skewness, kurtosis, 5 % VaR, 5 % ES, 0.5 % VaR, 0.5 % ES, minimum and maximum from the sample of 100 000 total discounted costs and aggregate disability durations related to the 100 000 simulated trajectories. For a particular trajectory, an aggregate disability duration is defined as the sum of the durations of all the sojourns in the disability state during the representative agent working life.

We note in particular that the skewness estimated from the costs distribution is greater than the skewness estimated from the durations distribution. This discrepancy can be linked to the presence of a deferment period because benefits are paid only when the duration of the disablement spell is greater than the deferred period of the policy.

[22] The authors compute the skewness, kurtosis and high-order moments thanks to closed formula

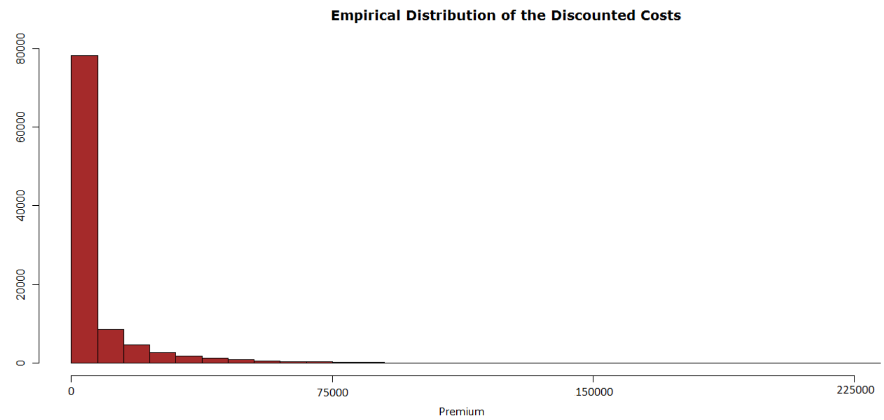

Figure 3: Empirical distribution of the 100 000 total discounted costs provided by the three-state model for a representative insured aged 25 at time of policy issue, working in a large city and in the sector "Finance & Insurance".

Statistic	Variable	
	Aggregate Duration of Disability Spells	Total Discounted Cost of DI
Mean	247.2	6,582.62
St. Dev.	205.4	14,446.32
Skewness	2.22	3.79
Kurtosis	10.37	23.2
Minimum	0	0
5 % VaR	653	35,275.21
5% ES	875.1	55,959.66
0.5 % VaR	1,204.8	84708.72
0.5% ES	1,401.2	103,636.58
Maximum	2,684.3	258,342.07

Table 7: Empirical estimates from the 100 000 total discounted costs and aggregate disability durations corresponding to the simulations for the considered representative agent. The currency unit is the dollar (we consider its value at the beginning of the 1st year of the coverage period, i.e. at policy issue) and the considered time-unit is the day.

7.2 Premium calculation using a modified standard deviation principle

Let us consider the positive random variable $S_n = \sum_{i=1}^{n} C_i$ denoting the cumulative cost of all the discounted costs C_i entailed by the claims pertaining to a representative policy, i.e. the representative insured described *supra*.

We evaluate the premium the insured has to pay to the insurer to be covered in case of sickness or accident, using a modified standard deviation principle, applied to the random variable S rather than to the variable C, i.e.:

$$\Pi(S) = \mathbb{E}\left[S\right] + \xi\sigma(S).$$

The following result will be used straightforwardly to price and reserve the representative insured policy, in a manner consistent with arbitrary solvency requirements and risk horizon:

Proposition 7.1 *The aggregate net technical surplus $S_n - \Pi(S_n)$ satisfies the following limit result:*

$$S_n - \Pi(S_n) \xrightarrow{d} \mathcal{N}(-\xi, 1)$$

Proof Since the C_i are an i.i.d. sequence, we can apply the Central Limit Theorem to the aggregate claim cost S_n. Indeed,

$$\begin{aligned}
\Pr\ [S_n - \Pi(S_n) > 0] &= \Pr\left[\sum_{i=1}^{n} C_i - \Pi(S_n) > 0\right] \\
&= \Pr\left[\sum_{i=1}^{n} C_i - n\mathbb{E}(C) > \sqrt{n}\xi\sigma(C)\right] \\
&= \Pr\left[\frac{\sum_{i=1}^{n} C_i - n\mathbb{E}(C)}{\sqrt{n}\sigma(C)} > \xi\right] \\
&\longrightarrow_{n\to\infty} 1 - \Phi(\xi),
\end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. ∎

By imposing e.g. the constraint $1 - \Phi(\xi) = 0.005$, which means that there is a 99.5% probability that the Cumulative Cost S will not exceed the cumulated premiums $\Pi(S)$, we get

$$\xi = \Phi^{-1}(0.995) \Longleftrightarrow \xi \approx 2.575.$$

We claim that the set of 100 000 simulations fulfills asymptotic conditions[23] and the initial capital of the insurer is equal to zero. To avoid the ruin for the whole coverage period with a 99.5% probability, replacing the mean and the standard deviation of the random variable C by their empirical estimates, we have to price the representative policy at the level \$ 6, 700.25, including \$ 117.63 of insurance loading. We observe the safety loading is very low compared to the premium of \$ 6, 700.25, which is due to the asymptotic approach: the greater n is the lower the loading is.

7.3 Simple scenario-based application to determine reserves and forecast technical accounts

In practice, the insurer may wish to spread the payment of the premium over the coverage period. Let us suppose the insurer wants to fix a constant annual premium $\overline{\Pi}$ to provide financial support in the event of the policyholder becoming unable to work for longer than the deferred period because of disabling illness or injury. Assuming that no waiver of premiums may be granted in the event of disability, the insurer has the following equation to solve

$$\text{Loaded Single Premium} = \sum_{k=0}^{42} \frac{\overline{\Pi}}{(1+i)^k}$$

from which we obtain

$$\overline{\Pi} = 271.25,$$

so that the constant annual premium amounts to \$ 271.25. Now suppose two years have elapsed since the policy issue. Our representative agent is 27 and earns \$ 68, 958.5 a year. The ceiling value and the maximum Benefit Payable From the State have also increased according to the indexing rules reported in the table 6. Besides, we suppose that in the past two years, he has not been disabled with duration longer than the deferred period. We always consider that the risk horizon to take into account to provide reserves is given by the policy expiration. Once again, we generate 100 000 simulations for this representative agent. The new net single premium amounts to \$ 6, 776.54 (dollar value at the beginning of the 3rd year of the coverage period that is equivalent to \$ 6, 387.54 at time of policy issue) two years after the policy issue. Using the formula (9), we express the reserves to establish at the end of the second year:

[23]i.e. the confidence interval at level α defined by $I_n = \left[\overline{C}_n - a_\alpha\sqrt{\frac{\overline{V}_n}{n}}, \overline{C}_n + a_\alpha\sqrt{\frac{\overline{V}_n}{n}}\right]$, where $\overline{C}_n$ and $\overline{V}_n$ are the empirical estimates of the mean and the variance of the simulated discounted aggregate benefit (table 7), and a_α is the two-sided α-quantile of the standard normal distribution, is considered to be small enough when α is close to 1.

$$V_a(2 \times 365.25) = 6776.54 - \sum_{k=0}^{40} \frac{\overline{\Pi}}{1.03^k}$$
$$= 235.41$$

Given the expectation of the future cash flows from the insured (anticipated premiums) and from the obligations towards the insured (anticipated benefits) over the 41 years to the come (time to expiration), the insurer shall establish a reserve for this representative policy: at the end of the second coverage year, this reserve amounts to $ 235.41. Finally, at the end of the second coverage year the insurer has collected and invested premiums, so he has got $ 567.16 (value at the beginning of the third coverage year) from the premiums received in the first two years of the policy. The insurer can therefore use this amount of cash to establish the required technical reserves. This leaves a margin of $ 331.75 at the beginning of the third year of coverage. These potential margins during the life of the contract can partly be explained by the Markov process of the disability inception process[24]. The results arising from the retained scenario for the first two years of the contract are summarized in the tables 8 and 9.

Time	At policy issue $y = 0$	At the end of the first year $y = 1^-$	Two years after issuance $y = 2^-$
Actuarial Value of Future Outflows	6,582.62	6,677.40	6,776.54
Actuarial Value of Future Inflows	6,700.25	6,621.87	6,541.13
Reserves	(117.63)	55.53	235.41

Table 8: Main Actuarial Values resulting from our simulations on the basis of the following scenario: no waiver of premiums in case of disability, no disability lasting longer than the deferred period during the first two years, risk horizon extended to contract expiry.

Time	$y = 0$	$y = 1^-$	$y = 2^-$
Assets	271.25	279.39	567.16
Reserves	(117.63)	55.53	235.41
Claims paid	0	0	0
Profit	388.88	223.86	331.75

Table 9: Projected simple technical accounts on the basis of the aforementioned scenario.

Conclusion

Admittedly, the representative-agent hypothesis is a non-standard approach to assess the pricing and reserving needs in the group insurance industry. It amounts to assuming that there is an agent whose economic behaviour is, up to normalisation, the aggregate behaviour of the insured group. But it provides an efficient bridge to connect the upstream econometric estimations to the downstream stochastic approaches. In addition, Multi-State Models define a realistic and unifying framework, strongly consistent with this assumption. An outstanding advantage of these models is their tractability and flexibility, since they are fully compatible with a specific modeling for each type of transition. This setting is also appropriate to deal with a wide range of policy conditions and long-term coverages, and to include competing risks in the risk assessment process, thus providing more consistent and accurate long-term premiums and reserves. To adopt the representative insured assumption in the multistate framework is particularly convenient to streamline risk assessment and burdensome actuarial calculations.

From the graduation of the different transition intensities, probabilities of survival within each state are expressed, and Markov chain Monte Carlo methods can be applied. Trajectories can thus be provided for the selected representative

[24] disability risk does not depend on the time spent in the active state

insured, fully described by a covariates vector at time of policy issue (e.g. gender, age, work area, sector of activity in our case). Using additional assumptions concerning the benefit function and indexing conventions, a discounted insurance cost (Net Single Premium) can be calculated for each trajectory from the sojourns in the disability state. This simulation and scenario-based approach can also be used to define an appropriate reserving policy and to design a relevant business strategy, consistent with predefined solvency requirements and a certain risk horizon, e.g. 3 to 5 years, since we are able to project future technical accounts. This microeconomic model is entirely consistent with a sound risk management in the sense that, all other things being equal, a change in a risk variable leads to a higher premium if the resulting representative agent has a higher risk of becoming disabled and/or remaining disabled.

As straightforward extensions to this paper, we might investigate the best definition of an optimal representative insured, a penalization method consistent with deviations from the rescaled limiting distribution, or the design of a more refined approach sharing the same risk philosophy, e.g. based upon heterogeneous insured models.

Technical Annexes

A Transition Intensity Approach

Starting from the equalities

$$\lambda_{\text{ai}}(x+t) = \lim_{\triangle t \downarrow 0} \frac{{}_{\triangle t}p^{\text{ai}}_{x+t} - {}_{0}p^{\text{ai}}_{x+t}}{\triangle t} = \frac{\partial}{\partial \chi}\,{}_{\chi}p^{\text{ai}}_{x+t}\Big|_{\chi=0}$$
$$\lambda_{\text{ad}}(x+t) = \lim_{\triangle t \downarrow 0} \frac{{}_{\triangle t}p^{\text{ad}}_{x+t} - {}_{0}p^{\text{ad}}_{x+t}}{\triangle t} = \frac{\partial}{\partial \chi}\,{}_{\chi}p^{\text{ad}}_{x+t}\Big|_{\chi=0}$$

we get

$${}_{\triangle t}p^{\text{ai}}_{x+t} = \lambda_{\text{ai}}(x+t)\,\triangle t + o(\triangle t) \approx \lambda_{\text{ai}}(x+t)\,\triangle t$$
$${}_{\triangle t}p^{\text{ad}}_{x+t} = \lambda_{\text{ad}}(x+t)\,\triangle t + o(\triangle t) \approx \lambda_{\text{ad}}(x+t)\,\triangle t$$

where $\triangle t$ is small enough. For any $\triangle t > 0$, the probability that at least two transitions occur in the interval $[t, t+\triangle t]$ is little-o of $\triangle t$. Finally, we have

$$\lambda_{\text{a}\bullet}(x+t) = \lim_{\triangle t \downarrow 0} \frac{{}_{\triangle t}p^{\text{ai}}_{x+t} + {}_{\triangle t}p^{\text{ad}}_{x+t}}{\triangle t} = \lim_{\triangle t \downarrow 0} \frac{1 - {}_{\triangle t}p^{\text{aa}}_{x+t}}{\triangle t} = \lim_{\triangle t \downarrow 0} \frac{{}_{0}p^{\text{aa}}_{x+t} - {}_{\triangle t}p^{\text{aa}}_{x+t}}{\triangle t} = -\frac{\partial}{\partial \chi}\,{}_{\chi}p^{\text{aa}}_{x+t}\Big|_{\chi=0}$$

Hence,

$${}_{\triangle t}p^{\text{aa}}_{x+t} = 1 - \lambda_{\text{a}\bullet}(x+t)\,\triangle t + o(\triangle t) \approx 1 - \lambda_{\text{a}\bullet}(x+t)\,\triangle t$$

for $\triangle t$ small enough.

Then, in order to express the transition probabilities from the disability state in terms of the intensities, we provide the relations

$$\lambda_{\text{ia}}(x+t,s) = \lim_{\triangle t \downarrow 0} \frac{{}_{\triangle t}p^{\text{ia}}_{x+t,s} - {}_{0}p^{\text{ia}}_{x+t,s}}{\triangle t} = \frac{\partial}{\partial \chi}\,{}_{\chi}p^{\text{ia}}_{x+t,s}\Big|_{\chi=0}$$
$$\lambda_{\text{id}}(x+t) = \lim_{\triangle t \downarrow 0} \frac{{}_{\triangle t}p^{\text{id}}_{x+t} - {}_{0}p^{\text{id}}_{x+t}}{\triangle t} = \frac{\partial}{\partial \chi}\,{}_{\chi}p^{\text{id}}_{x+t}\Big|_{\chi=0}$$

which lead to the approximations

$${}_{\triangle t}p^{\text{ia}}_{x+t,s} = \lambda_{\text{ia}}(x+t,s)\,\triangle t + o(\triangle t) \approx \lambda_{\text{ia}}(x+t,s)\,\triangle t$$
$${}_{\triangle t}p^{\text{id}}_{x+t} = \lambda_{\text{id}}(x+t)\,\triangle t + o(\triangle t) \approx \lambda_{\text{id}}(x+t)\,\triangle t.$$

where $\triangle t$ is small enough.

We finally obtain

$$\lambda_{\text{i}\bullet}(x+t,s) = \lim_{\triangle t \downarrow 0} \frac{{}_{\triangle t}p^{\text{ia}}_{x+t,s} + {}_{\triangle t}p^{\text{id}}_{x+t}}{\triangle t} = \lim_{\triangle t \downarrow 0} \frac{1 - {}_{\triangle t}p^{\text{ii}}_{x+t,s}}{\triangle t} = \lim_{\triangle t \downarrow 0} \frac{{}_{0}p^{\text{ii}}_{x+t,s} - {}_{\triangle t}p^{\text{ii}}_{x+t,s}}{\triangle t} = -\frac{\partial}{\partial \chi}\,{}_{\chi}p^{\text{ii}}_{x+t,s}\Big|_{\chi=0}$$

from which we deduce, for $\triangle t$ small enough:

$${}_{\triangle t}p^{\text{ii}}_{x+t,s} = 1 - \lambda_{\text{i}\bullet}(x+t,s)\,\triangle t + o(\triangle t) \approx 1 - \lambda_{\text{i}\bullet}(x+t,s)\,\triangle t.$$

B Expressions from occupancy probabilities

We recall that

$$ {}_{\tau}p_{x+t}^{\overline{\mathrm{aa}}} = \Pr\ [X_{t+\tau} = e_1 \cap D_{t+\tau} \geq \tau \mid X_t = e_1]. $$

Using Markov property, we get the equality

$$ {}_{t+\triangle t}p_x^{\overline{\mathrm{aa}}} = {}_t p_x^{\overline{\mathrm{aa}}}\, {}_{\triangle t}p_{x+t}^{\overline{\mathrm{aa}}} = {}_t p_x^{\overline{\mathrm{aa}}} \left(1 - \lambda_{\mathrm{a}\bullet}(x+t)\, \triangle t + o(\triangle t)\right) $$

Rearranging this expression, we find

$$ \frac{{}_{t+\triangle t}p_x^{\overline{\mathrm{aa}}} - {}_t p_x^{\overline{\mathrm{aa}}}}{\triangle t} = -\lambda_{\mathrm{a}\bullet}(x+t)\, {}_t p_x^{\overline{\mathrm{aa}}} + \frac{o(\triangle t)}{\triangle t}. $$

If $\triangle t$ vers 0, we get the differential equation

$$ \frac{\partial}{\partial t}\, {}_t p_x^{\overline{\mathrm{aa}}} = -\lambda_{\mathrm{a}\bullet}(x+t)\, {}_t p_x^{\overline{\mathrm{aa}}} $$

whose solution must satisfy the initial condition ${}_0 p_x^{\overline{\mathrm{aa}}} = 1$. So we have

$$ {}_t p_x^{\overline{\mathrm{aa}}} = \exp\left(-\int_0^t \lambda_{\mathrm{a}\bullet}(x+\tau)\, d\tau\right) $$

Then, recalling that

$$ {}_{\tau}p_{x+t,s}^{\overline{\mathrm{ii}}} = \Pr\ [X_{t+\varepsilon} = e_2 \text{ for any } 0 \leq \varepsilon \leq \tau \mid X_t = e_2, D_t = s], $$

we can, as previously, obtain an expression of the probability of survival in the disability state from the instantaneous exit rate. To achieve this, we note that

$$ \begin{aligned} {}_{\tau+\triangle\tau}p_{x+t,s}^{\overline{\mathrm{ii}}} &= {}_{\tau}p_{x+t,s}^{\overline{\mathrm{ii}}}\ {}_{\triangle\tau}p_{x+t+\tau,s+\tau}^{\overline{\mathrm{ii}}} \\ &= {}_{\tau}p_{x+t,s}^{\overline{\mathrm{ii}}} \left(1 - \lambda_{\mathrm{i}\bullet}(x+t+\tau, s+\tau)\, \triangle\tau + o(\triangle\tau)\right) \\ \iff \quad \frac{{}_{\tau+\triangle\tau}p_{x+t,s}^{\overline{\mathrm{ii}}} - {}_{\tau}p_{x+t,s}^{\overline{\mathrm{ii}}}}{\triangle\tau} &= -\lambda_{\mathrm{i}\bullet}(x+t+\tau, s+\tau)\, {}_{\tau}p_{x+t,s}^{\overline{\mathrm{ii}}} + \frac{o(\triangle\tau)}{\triangle\tau}, \end{aligned} $$

which provides - taking the limit - the differential equation

$$ \frac{\partial}{\partial\tau}\, {}_{\tau}p_{x+t,s}^{\overline{\mathrm{ii}}} = -\lambda_{\mathrm{i}\bullet}(x+t+\tau, s+\tau)\, {}_{\tau}p_{x+t,s}^{\overline{\mathrm{ii}}} $$

whose solution is given by

$$ {}_{\tau}p_{x+t,s}^{\overline{\mathrm{ii}}} = \exp\left(-\int_0^{\tau} \lambda_{\mathrm{i}\bullet}(x+t+\varepsilon, s+\varepsilon)\, d\varepsilon\right). $$

C Integro-differential equations for transition probabilities from the Markov 'active' state

We can also express the transition probabilities ${}_t p_x^{\mathrm{aa}}$, ${}_t p_x^{\mathrm{ad}}$ and ${}_{s,t}p_x^{\mathrm{ai}}$ as solutions of a system of integro-differential equations. Indeed, the following equality

$$ {}_{t+\triangle t}p_x^{\mathrm{aa}} = {}_t p_x^{\mathrm{aa}} \left(1 - \lambda_{\mathrm{a}\bullet}(x+t)\,\triangle t\right) + \int_0^t {}_\tau p_x^{\mathrm{aa}}\, \lambda_{\mathrm{ai}}(x+\tau)\, {}_{t-\tau}p_{x+\tau,0}^{\overline{\mathrm{ii}}}\, \lambda_{\mathrm{ia}}(x+t,\tau-t)\,\triangle t\, d\tau + o(\triangle t) $$

provides, rearranging this expression and making $\triangle t \downarrow 0$

$$ \frac{\partial}{\partial t} {}_t p_x^{\mathrm{aa}} = -{}_t p_x^{\mathrm{aa}}\, \lambda_{\mathrm{a}\bullet}(x+t) + \int_0^t {}_\tau p_x^{\mathrm{aa}}\, \lambda_{\mathrm{ai}}(x+\tau)\, {}_{t-\tau}p_{x+\tau,0}^{\overline{\mathrm{ii}}}\, \lambda_{\mathrm{ia}}(x+t,\tau-t)\, d\tau. $$

Then we express the probability ${}_{t+\triangle t}p_x^{\mathrm{ad}} - {}_t p_x^{\mathrm{ad}}$ that an insured who is active at age x will die between ages $x+t$ and $x+t+\triangle t$, as follows:

$$ {}_{t+\triangle t}p_x^{\mathrm{ad}} - {}_t p_x^{\mathrm{ad}} = {}_t p_x^{\mathrm{aa}}\, \lambda_{\mathrm{ad}}(x+t)\,\triangle t + \int_0^t {}_\tau p_x^{\mathrm{aa}}\, \lambda_{\mathrm{ai}}(x+\tau)\, {}_{t-\tau}p_{x+\tau,0}^{\overline{\mathrm{ii}}}\, \lambda_{\mathrm{id}}(x+t,\tau-t)\,\triangle t\, d\tau + o(\triangle t) $$

This provides, rearranging and making $\triangle t \downarrow 0$

$$ \frac{\partial}{\partial t} {}_t p_x^{\mathrm{ad}} = {}_t p_x^{\mathrm{aa}}\, \lambda_{\mathrm{ad}}(x+t) + \int_0^t {}_\tau p_x^{\mathrm{aa}}\, \lambda_{\mathrm{ai}}(x+\tau)\, {}_{t-\tau}p_{x+\tau,0}^{\overline{\mathrm{ii}}}\, \lambda_{\mathrm{id}}(x+t)\, d\tau $$

in accordance with the Markov property characterizing the transition process from disability to death [i.e. $\forall \nu > t > 0,\ \lambda_{\mathrm{id}}(x+t,\nu-t) = \lambda_{\mathrm{id}}(x+t)$]. Calculations required to solve the system of integro-differential equations directly may be heavy and time-consuming. In practice, approximations have been proposed. We refer the reader to Waters [29] for a more detailed description.

Now the probability ${}_{s,t}p_x^{\mathrm{ai}}$ may be expressed in the same way as the previous probability by noticing that for $s > t$

$$ {}_{s,t}p_x^{\mathrm{ai}} = {}_{t,t}p_x^{\mathrm{ai}} \quad\Longrightarrow\quad \frac{\partial}{\partial s} {}_{s,t}p_x^{\mathrm{ai}} = 0. $$

Then, for $s \le t$, we can analyze the difference ${}_{s+\triangle s,t}p_x^{\mathrm{ai}} - {}_{s,t}p_x^{\mathrm{ai}}$, which gives the probability of an active insured aged x being disabled at age $x+t$ with a disability duration between s and $s+\triangle s$:

$$ {}_{s+\triangle s,t}p_x^{\mathrm{ai}} - {}_{s,t}p_x^{\mathrm{ai}} = {}_{t-s-\triangle s}p_x^{\mathrm{aa}}\, \lambda_{\mathrm{ai}}(x+t-s-\triangle s)\, {}_s p_{x+t-s,0}^{\overline{\mathrm{ii}}}\, \triangle s + o(\triangle s), $$

The latter expression provides, rearranging terms and letting $\triangle t$ going to 0,

$$ \frac{\partial}{\partial s} {}_{s,t}p_x^{\mathrm{ai}} = {}_{t-s}p_x^{\mathrm{aa}}\, \lambda_{\mathrm{ai}}(x+t-s)\, {}_s p_{x+t-s,0}^{\overline{\mathrm{ii}}} \quad \text{for } s \le t. $$

References

[1] Aalen, O.O., & Johansen, S., An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *Scandinavian Journal of Statistics* **5**, pp. 141-150, 1978.

[2] Billingsley, P., *Statistical inference for Markov processes*, Chicago University Press, 1961.

[3] Black, K., & Skipper, H.D., Life Insurance, *Prentice Hall, Englewood Cliffs*, New Jersey, USA, 1987.

[4] Cordeiro, I.M.F., A multiple state model for the analysis of permanent health insurance claims by cause of disability. *Insurance: Mathematics & Economics* **30**, pp.167-186, Elsevier, 2002.

[5] Cox, D.R., Regression models and life-tables. *Journal of the Royal Statistical Society* B **34**, pp. 187-220, 1972.

[6] De Dominicis, R., & Janssen, J., An algorithmic approach to non-homogeneous semi-Markov processes. *Insurance: Mathematics & Economics* **3**, pp.157-165, Elsevier, 1984.

[7] Denuit, M., & Robert, C., *Actuariat des assurances de personnes, Modélisation, tarification et provisionnement*, Economica, 2007.

[8] Haberman, S., & Pitacco, E., *Actuarial Models for Disability Insurance*, Chapman & Hall/CRC, 1999.

[9] Hoem, J.M., Markov chains models in life insurance. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik* **9**, pp. 91-107, 1969.

[10] Hoem, J.M., A Markov chain models of working life tables. *Scandinavian Actuarial Journal* **1**, pp. 1-20, 1977.

[11] Iosifescu Manu, A., Non homogeneous semi-Markov processes. *Studii si Cercetuari Matematice* **24**, pp. 529-533, 1972.

[12] Janssen, J., Application des processus semi-Markoviens à un problème d'invalidité. *Bulletin de l'Association Royale des Actuaires Belges* **63**, 2006.

[13] Janssen, J., & Manca, R., *Applied Semi-Markov Processes*, Springer, 2006.

[14] Levikson, B., Frostig, E., & Bshouty, D., An algorithm evaluating generalized life insurance programs. *Methodology and Computing in Applied Probability* **3**, pp.329-340, 2001.

[15] Lévy, P., Processus semi-Markoviens. *Proceeding of International Congress Mathematics*, Amsterdam, 1990.

[16] Limnios, N., & Ouhbi, B., Non-parametric estimation for semi-Markov processes based on K-sample paths with application to reliability. *Applied Stochastic Models and Data Analysis - Conference Proceedings*, pp. 1061-1068, 2005.

[17] O'Grady, F.T., Individual Health Insurance, *Society of Actuaries, Schaumburg*, Illinois, USA, 1988.

[18] Ouhbi, B., & Limnios, N., Non-parametric estimation for semi-Markov processes based on its hazard rate functions. *Statistical Inference for Stochastic Processes* **2**, pp. 151-173, 1999.

[19] Ouhbi, B., & Limnios, N., Non-parametric reliability estimation of semi-Markov processes. *Journal of Statistical Planning and Inference* **109**, pp. 155-165, 2000.

[20] Pitacco, E., Actuarial models for pricing disability benefits: Towards a unifying approach. *Insurance: Mathematics & Economics* **16**, pp.39-62, Elsevier, 1994.

[21] Planchet, F., & Thérond, P., *Modèles de durée, applications actuarielles*, Economica, 2006.

[22] Renshaw, A., & Haberman, S., On the graduation associated with a mutiple state model for permanent health insurance. *Insurance: Mathematics & Economics*, **17**, pp.1-17, Elsevier, 1995.

[23] Renshaw, A., & Haberman, S., Modelling the recent time trends in UK permanent health insurance recovery, mortality and claim inception transition intensities. *Insurance: Mathematics & Economics*, **27**, pp.365-396, Elsevier, 2000.

[24] Stenberg, F., Manca, R., & Silvestrov, D., An Algorithmic Approach to Discrete Time Non-homogeneous Backward Semi-Markov Reward Processes with an Application to Disability Insurance. *Methodol Comput Appl Probab* **9**, pp. 497-519, 2007.

[25] Waters, H.R., An approach to the study of multiple state models. *Journal of the Institute of Actuaries* **111**, pp. 363-374, 1984.

[26] Waters, H.R., Some aspects of the modelling of permanent health insurance. *Journal of the Institute of Actuaries* **116**, pp. 611-624, 1989.

[27] Waters, H.R., The recursive calculation of the moments of the profit on a sickness insurance policy. *Insurance: Mathematics & Economics* **9**, pp.101-113, Elsevier, 1990.

[28] Waters, H.R., A multiple state model for permanent health insurance. *CMIR* **12**, pp. 5-20, 1991.

[29] Waters, H.R., Part D: Computational procedures for the model. *CMIR* **12**, pp. 79-96, 1991.

[30] Wolthuis, H., & Hoem, J.M., The retrospective premium reserves. *Insurance: Mathematics & Economics*, **5**, pp.217-254, Elsevier, 1986.