CONTINUOUS-TIME SEMI-MARKOV INFERENCE OF BIOMETRIC LAWS ASSOCIATED WITH A LONG-TERM CARE INSURANCE PORTFOLIO

BY

GUILLAUME BIESSY

ABSTRACT

Unlike the mortality risk on which actuaries have been working for more than a century, the long-term care (LTC) risk is relatively new and as of today hardly mastered. Semi-Markov processes have been identified as an adequate tool to study this risk. Nevertheless, access to data is limited and the associated literature still scarce. Insurers mainly use discrete time methods directly inspired from the study of mortality in order to build experience tables. Those methods however are not perfectly suited for the study of competing risk situations. This article provides a theoretical framework to estimate biometric laws associated with a LTC insurance portfolio. The presented method relies on a continuous-time semi-Markov model with three states: autonomy, disability and death. The process describing the state of disability is defined through its transition intensities. We provide a formula to infer the mortality of autonomous people from the mortality of the whole portfolio, on which we have more reliable knowledge. We then propose a parametric expression for the remaining intensities of the model. In particular, incidence in LTC is described by a logistic formula. Under the assumption that the disabled population is a mixture of two latent populations with respect to the category of pathology that caused LTC, we show that the resulting intensity of mortality in LTC takes a very peculiar form and depends on time spent in the LTC state. Estimation of parameters relies on the maximum likelihood method. Our parametric approach, while inducing model uncertainty, eliminates issues related to segmentation in age categories, smoothing or extrapolation at higher ages and thus proves very convenient for the practitioner. Finally, we provide an application using data from a real LTC insurance portfolio.

KEYWORDS

Long-term care insurance, continuous-time semi-Markov process, competing risks, maximum likelihood, mixture model, parametric model.

Astin Bulletin 47(2), 527-561. doi: 10.1017/asb.2016.41

1. Introduction

Disability among elderly people can be defined as a permanent state of inability to autonomously perform activities of daily living. It is mostly caused by diseases linked to ageing, such as dementia, neurological diseases, cardiovascular diseases and cancer. Disabled elderly people require regular care whose frequency increases with the severity of their status. While some people can rely at least partially on their family or their friends for help, others have to hire professional caregivers or join a nursing home, whose average cost exceeds 3,000€ per month. Despite public aids, this cost proves overwhelming for most pensioners. Therefore, to long-term care (LTC) is associated a financial risk to which most people are exposed. In France, part of this risk is transferred through private insurance contracts.

The LTC risk is complex. Its study requires to take into account incidence in LTC as well as probabilities of death for both autonomous and disabled people, which are very different from another. This risk is directly related to ageing through pathologies, and longevity gains in the second half of the 20th century made it paramount. The very first LTC insurance products appeared in the United States during the 1980s and shortly after in France. Average age at subscribing for those products is close to 60, whereas the average age at which LTC occurs is near to 85. Therefore, even on older portfolios, the number of claims remains limited. Moreover, in France, insurers and public aids use different definitions to assess the level of required care. Those definitions as well as insurers underwriting and claims policies often change over time. All those elements make data aggregation from several sources very difficult, which may explain the difficulty of getting a better knowledge of the risk.

Markov processes are such that their transition probabilities only depend on the current state of the process. A semi-Markov process is a generalization for which transition probabilities depend on both the current state and the time spent in the current state. One can find more details about those processes in Cinlar (1969). Multi-state models based on Markov and semi-Markov processes have led to many applications in the field of epidemiology. As the LTC state is mainly caused by pathologies, those processes appear as natural candidates to study the LTC risk. This framework has already been described for example in Haberman and Pitacco (1998) or Christiansen (2012). Several studies based on U.S. national data have also been performed. One can refer to Robinson (1996), Pritchard (2006) or more recently Fong *et al.* (2015). On the other hand, studies based on portfolio data (Guibert and Planchet, 2014) are very rare. Practitioners nevertheless played a key role in the knowledge of the LTC risk. One of the very first models on the French market was presented by SCOR (1995). Relying on a parametric approach, it highlights the exponential increase in the probabilities of incidence in LTC, and defines mortality in LTC (resp. autonomous mortality) as a linear function of the general population mortality, computed via an exogenous mortality table. With only five parameters required to model the whole process, it is remarkably simple. It is however based on the Markov assumption that mortality in LTC only depends on the age of the disabled life, and not on the time since the entry in LTC. The Markov assumption is still used today by many insurers as well as in recent academic papers like Pitacco (2015) or Fong *et al.* (2015), because it allows for simpler models. However, it does not reflect the reality of the LTC process, for which mortality is much higher during the first year in LTC than for the subsequent year. For an insurance company, ignoring this feature of the risk can be very damaging. Indeed, it leads to greatly overestimating mortality in LTC based on the first-year mortality experience and therefore underestimating the required amount of reserve, which results in heavy losses in the future.

Semi-Markov processes have already been used for disability insurance, especially through the illness-death model as described in Pitacco (2014). However, one has to keep in mind that on one hand disability insurance only lasts until retirement age with a limited period for benefits. On the other hand, individual LTC insurance relies on lifetime annuities with no expiry date. Therefore, although a similar model may be used for both risks, issues related to extrapolation of biometric laws at higher ages and higher duration in the disabled state arise in the study of the LTC risk. For the same reason, non-parametric methods based on Nelson–Aalen estimator (Klein, 1991) that have also been used to study the LTC risk, for example in Guibert and Planchet (2015) still need to be associated with parametric methods for the extrapolation step.

In this article, we present a parametric approach relying on a continuous-time semi-Markov process, which is defined using its transition intensities. Compared to a discrete-time approach, it allows to get a more straightforward modelling of the process, while correctly taking into account the competing risks (disability and death). Section 2 introduces the model and derives an equation to express the autonomous mortality using general mortality and other intensities of the model. Benefits to use general mortality instead of autonomous mortality are discussed with more details. We then introduce the intensity for general mortality of the portfolio using a simple relational model as in Brass (1971). We propose a parametric expression for the intensity of incidence in LTC, based on the logistic form introduced by Perks (1932) for the study of mortality. We use a complex parametric model for the intensity of mortality in LTC, corresponding to a latent mixture model where we consider two homogeneous populations of disabled people, with two different levels of mortality. Estimation of parameters relies on the maximum likelihood method. We also introduce formulas for pricing and reserving based directly on the transition intensities. Section 3 provides an application of the model based on data from a real insurance portfolio. For each transition intensity, several models of increasing complexity are compared using the Bayesian Information Criterion (BIC). Comparison with empirical transition rates is also provided. Robustness of estimation is then assessed using a non-parametric quantile bootstrap method. Finally, Section 4 summarizes the results obtained and discusses limits and potential improvements of the model.

2. Model

2.1. Notations

For $x_0 \geq 0$, let us consider a continuous-time process $(Z_x)_{x \geq x_0}$ with values in the three-state set $E = \{A, I, D\}$ of autonomy, LTC (or "illness"), death, respectively. Let us assume that Z is *càd-làg* and that $Z_{x_0} = A$. The index variable of the process Z is called age of the process. Therefore, x_0 is an initial age where all individuals are assumed to be autonomous. For $x \geq x_0$ let us denote by A_x (resp. I_x, D_x) the probability for the process to be in the state of autonomy (resp. LTC, death) at age x or more formally

$$
\begin{aligned}
A_x &= P(Z_x = A | Z_{x_0} = A), \\
I_x &= P(Z_x = I | Z_{x_0} = A), \\
D_x &= P(Z_x = D | Z_{x_0} = A).
\end{aligned}
$$

Hence, $A_{x_0} = 1$ and for all $x \geq x_0$, $A_x + I_x + D_x = 1$.

We now assume that $(Z_x)_{x \geq x_0}$ is a non-homogeneous semi-Markov process and introduce the transition intensities, also called instantaneous transition probabilities. Transition intensities allow us to fully describe the behaviour of the process

$$
\begin{aligned}
\mu_a(x) &= \lim_{h \to 0} \frac{1}{h} P\left(Z_{x+h} = D | Z_x = A\right), \\
\lambda(x) &= \lim_{h \to 0} \frac{1}{h} P\left(Z_{x+h} = I | Z_x = A\right), \\
\mu_i(x, t) &= \lim_{h \to 0} \frac{1}{h} P\left(Z_{x+t+h} = D | Z_{x^-} = A, Z_x = I, Z_{x+t} = I\right).
\end{aligned}
$$

Those intensities are called respectively intensity of entry in LTC, intensity of autonomous mortality and intensity of mortality in LTC, with the latter intensity depending on both the age at onset of LTC and time spent in LTC. We consider that death is an absorbing state and that there is no transition allowed from LTC to autonomy. To understand this last assumption, one has to keep in mind that on the French LTC insurance market, the LTC benefit is only granted when the disabled state is expected to be permanent. Therefore, cases of return to the autonomy state are quite rare, compared to other markets where this is not the case. Furthermore, once the benefit is granted, the annuitant is not required to provide any proof that they are still disabled. Hence, ignoring cases of return to the autonomy state does not introduce any inconsistency with the way the insurance products are priced, and it allows for simpler models. Given the limited amount of available data this proves very convenient. A representation of the model can be found in Figure 1.

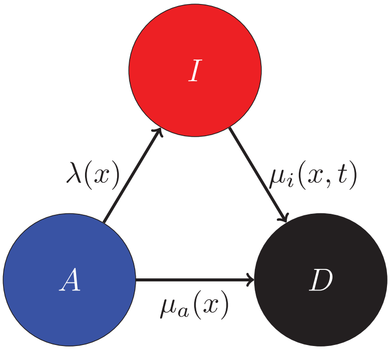

FIGURE 1: The three states continuous-time semi-Markov model and the associated transition intensities. (Color online)

Lemma 1. *Let $x \geq x_0$. The probability A_x (resp. I_x) to be in the autonomous (resp. disabled) state at age x may be expressed directly from the transition intensities of the model and we have*

$$A_x = \exp\left(-\int_{x_0}^{x} [\lambda(u) + \mu_a(u)]\, du\right), \tag{1}$$

$$I_x = \int_{x_0}^{x} \lambda(u)\, A_u \exp\left(-\int_{u}^{x} \mu_i(u, v-u) dv\right) du. \tag{2}$$

Proof. For $x \geq x_0$, $h \geq 0$, we have

$$P(Z_{x+h} = A) = [1 - P(Z_{x+h} = I | Z_x = A) - P(Z_{x+h} = D | Z_x = A)] \\ \times P(Z_x = A),$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{d}x} P(Z_x = A) = -[\mu_a(x) + \lambda(x)]\, P(Z_x = A).$$

As $A_{x_0} = 1$, this equation has a unique solution

$$A_x = \exp\left(-\int_{x_0}^{x} [\lambda(u) + \mu_a(u)]\, \mathrm{d}u\right). \tag{3}$$

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

For $x \geq x_0$, $t, h \geq 0$, we can write

$$P(Z_{x+t+h} = I | Z_{x^-} = A, Z_x = I) = P(Z_{x+t+h} = I | Z_{x^-} = A, Z_x = I, Z_{x+t} = I) \times P(Z_{x+t} = I | Z_{x^-} = A, Z_x = I),$$

which gives us

$$\frac{\mathrm{d}}{\mathrm{d}t} P(Z_{x+t} = I | Z_{x^-} = A, Z_x = I) = -\mu_i(x, t) P(Z_{x+t} = I | Z_{x^-} = A, Z_x = I).$$

As

$$P(Z_x = I | Z_{x^-} = A, Z_x = I) = 1,$$

we obtain

$$P(Z_{x+t} = I | Z_{x^-} = A, Z_x = I) = \exp\left(-\int_0^t \mu_i(x, u)\mathrm{d}u\right).$$

Then, as we have the following decomposition:

$$I_x = \int_{x_0}^{x} P(Z_u = A) P(Z_u = I | Z_{u^-} = A) P(Z_x = I | Z_{u^-} = A, Z_u = I),$$

we get an expression of the probability to be disabled at age $x \geq x_0$

$$I_x = \int_{x_0}^{x} \lambda(u) A_u \exp\left(-\int_u^x \mu_i(u, v - u)\mathrm{d}v\right). \tag{4}$$

■

2.2. Link with general mortality

Let us consider the intensity of mortality for the aggregated population of autonomous and disabled (hereafter general mortality) defined by

$$\mu_g(x) = \lim_{h \to 0} \frac{1}{h} P\left(Z_{x+h} = D | Z_x \in \{A, I\}\right).$$

Figure 2 represents the fourth transition in our model, a transition between life and death for the general population.

Lemma 2. *For $x \geq x_0$ and $t \geq 0$, let us denote by $\Delta(x, t)$ the difference between the intensity of mortality in LTC and the intensity of autonomous mortality for*

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

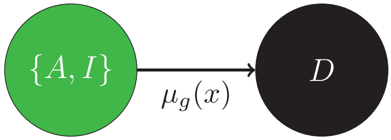

FIGURE 2: Intensity of mortality for the general population. (Color online)

the same current age, so that $\Delta(x, t) = \mu_i(x, t) - \mu_a(x + t)$. Then the intensity of autonomous mortality is solution of the following equation:

$$\mu_a(x) = \mu_g(x) - \frac{\int_{x_0}^{x} \lambda(u) \Delta(u, x-u) \exp\left(-\int_u^x [\Delta(u, v-u) - \lambda(v)]\, dv\right) du}{1 + \int_{x_0}^{x} \lambda(u) \exp\left(-\int_u^x [\Delta(u, v-u) - \lambda(v)]\, dv\right) du}. \quad (5)$$

Proof. Differentiating (1) and (2) gives us Equations (6) and (7) below which describe the evolution of the probabilities A_x and I_x. Similarly, from the definition of μ_g, we get Equation (8). We obtain a system of three differential equations:

$$\frac{\mathrm{d}}{\mathrm{d}x} A_x = -[\lambda(x) + \mu_a(x)] A_x, \quad (6)$$

$$\frac{\mathrm{d}}{\mathrm{d}x} I_x = \lambda(x) A_x - \int_{x_0}^{x} \lambda(u) A_u \exp\left(-\int_u^x \mu_i(u, v-u)\mathrm{d}v\right) \mu_i(u, x-u)\mathrm{d}u, \quad (7)$$

$$\frac{\mathrm{d}}{\mathrm{d}x}(A_x + I_x) = -\mu_g(x)(A_x + I_x). \quad (8)$$

Summing the evolution Equations (6) and (7), and then identifying with Equation (8) yields

$$\mu_g(x)(A_x + I_x) = \mu_a(x) A_x + \int_{x_0}^{x} \lambda(u) A_u \exp\left(-\int_u^x \mu_i(u, v-u)\mathrm{d}v\right) \mu_i(u, x-u)\mathrm{d}u.$$

With simple algebra we get

$$\mu_a(x) = \mu_g(x)\left(1 + \frac{I_x}{A_x}\right) - \int_{x_0}^{x} \lambda(u) \frac{A_u}{A_x} \exp\left(-\int_u^x \mu_i(u, v-u)\mathrm{d}v\right) \mu_i(u, x-u)\mathrm{d}u.$$

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

Using (2) and (1), we obtain after a few simplifications

$$\mu_a(x) = \mu_g(x) - \int_{x_0}^{x} \lambda(u) \exp\left(-\int_{u}^{x} [\mu_i(u, v-u) - \lambda(v) - \mu_a(v)]\, \mathrm{d}v \right) \times \left[\mu_i(u, x-u) - \mu_g(x)\right] \mathrm{d}u.$$

Now, we replace the intensity of mortality in LTC using the formula:

$$\mu_i(x, t) = \mu_a(x+t) + \Delta(x, t),$$

which gives us

$$\mu_a(x) = \mu_g(x) - \int_{x_0}^{x} \lambda(u) \exp\left(-\int_{u}^{x} [\Delta(u, v-u) - \lambda(v)]\, \mathrm{d}v \right) \times \left[\mu_a(x) - \mu_g(x) + \Delta(u, x-u)\right] \mathrm{d}u.$$

This finally leads to the result. ■

Equation (5) allows us to use the general mortality instead of the autonomous mortality in the model, at the cost of the introduction of Δ, the difference between autonomous mortality and mortality in LTC. On one hand, mortality of autonomous people is complex to predict, because people can leave the state of autonomy either by becoming disabled or dying. Furthermore, the scope of autonomous people depends directly on the definition used for LTC. Therefore, predicting the autonomous mortality requires to have intensive knowledge of the LTC process beforehand. On the other hand, general mortality has been studied for a long time by actuaries, demographers, biologists and is very well documented. One can therefore rely on reference mortality tables for ages where no portfolio data is available.

The formula does not give an analytic expression for the intensity of autonomous mortality in the most general case. Numerical methods can however be used to compute it. As will be seen in Section 2.4.3, choosing an *ad hoc* model, the inner integrals in the formula take an analytic expression and numerical approximation is only required for the outer integrals. Intensity of general mortality appears directly in the equation, which is very convenient if we want to use an external reference for this intensity.

2.3. Data structure

Data issued from insurance portfolios generally consists of two databases. The first database gathers information on contributors and the second on annuitants. We also define the database of insured lives obtained by merging the two previous bases which will be used for the estimation of general mortality. From

TABLE 1

EXAMPLE OF A DATABASE OF CONTRIBUTORS.

DoB	DoS	DoE	CoE
12/23/1941	11/10/1992	09/27/2006	2
06/14/1926	03/28/1997	12/31/2014	0
04/17/1937	04/27/1995	04/08/2003	1

one portfolio to another, data quality and available information may vary a lot. In what follows, we assume both databases contain at least the variables of Table 1, listed as follows:

- DoB: date of birth of the individual,
- DoS: date of start. For contributors, it is the date of subscribing. For annuitants, the date of entry in LTC,
- DoE and CoE: the date of end and cause of end for the individuals, respectively. In the case where the observation ends because of death, we use code 1 for the cause, in the case of exit because of entry in LTC, we use code 2. For individuals still autonomous when the observation stops, trajectories are right-censored. We use code 0 and the date of exit is the date at which observation ends.

Other covariates such as gender, type of residence (home or facility), marital status, cause of disability, amount of annuity bought or premium for substandard risk may be available and bring additional information. In what follows, we assume that only gender is available and we estimate a separate model for male and female.

The observation period must often be limited in some way:

- By removing the last year of individual exposure. With each database is associated a date of extraction, which is the date of the latest entry in the database. In practice, most claims are reported up to 1 year after their occurrence, which may result in some missing information during the last year of observation. It may therefore be a good idea to set a date for the end of the observation 1 year prior to the date of extraction, in order not to underestimate the number of events. For events that occurred during the last year of observation, the associated code must then be set to 0.
- By removing the first 3 years of individual exposure. On the French individual LTC insurance market, there is usually an elimination period of 3 years for dementia and neurological pathologies that results in fewer claims during those 3 years. In order not to underestimate the incidence rate, we therefore remove the exposure for the first 3 years of observation for each individual.
- By shortening the observation period. When we study the behaviour of a population for a specific risk, it may change over time. There are several factors involved, such as changes in the definition of LTC, underwriting and

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

claim management policy in addition to underlying changes of biometric laws. Shortening the observation period is therefore required in order to minimize those effects and set a good compromise between large volume and stability of the underlying risk over the period.

Once the data has been processed, we may easily compute quantities of interest which will be used in the estimation procedure:

- The age of entry $x = \text{DoS} - \text{DoB}$,
- The age of exit $y = \text{DoE} - \text{DoB}$,
- The cause of exit $c = \text{CoE}$.

2.4. Parametric modelling of the intensities

In this section, we propose to rely on a parametric expression for each of the transition intensities in the model.

2.4.1. *Intensity of general mortality.* We want to assess the general mortality of our portfolio, which is seen as a specific population inside the French population. To do so, we rely on the database of insured lives as well as on an external mortality reference, using the Brass relational model as described in Brass (1971, 1974) or Hannerz (2001).

Let F be the cumulative distribution function associated with an intensity of mortality μ such that

$$F(x) = 1 - \exp\left(-\int_{x_0}^{x} \mu(u)\mathrm{d}u\right).$$

Then, we define the cumulative distribution odds (CDO) by

$$\text{CDO}(x) = \frac{F(x)}{1 - F(x)}.$$

In his model, Brass relies on the assumption that the logarithm of the CDO associated with the mortality of a reference population and the mortality of a specific population are parallel curves. We denote by μ_g and F_g (resp. μ_g^{ref} and F_g^{ref}) the intensity of mortality and cumulative distribution function associated with the mortality of the specific (resp. reference) population. Under this assumption, the Brass estimator for the intensity of mortality of the specific population is

$$\widehat{\mu_g}(x) = \frac{\widehat{\beta}\mu_g^{\text{ref}}(x)}{1 - (1 - \widehat{\beta})F_g^{\text{ref}}(x)},$$

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

where $\widehat{\beta}$ is the solution of the equation

$$\sum_x D_x = \sum_x D_x^{\text{ref}} \frac{\widehat{\beta} N_x}{N_x^{\text{ref}}(1-(1-\widehat{\beta})F_g^{\text{ref}}(x))},$$

and D_x, N_x (resp. D_x^{ref}, N_x^{ref}) are the total number of deaths observed and the number of years lived between ages x and $x+1$ by the specific population (resp. by the reference population). The Brass model only requires the estimation of a single parameter $\widehat{\beta}$. It gives an estimator for the intensity of mortality that converges smoothly towards the mortality reference at higher ages while predicting the same number of deaths as in the empirical data, given the empirical exposure.

2.4.2. *Intensity of incidence in LTC.* For the intensity of incidence in LTC, we consider the logistic model introduced in Beard (1959, 1971) and Perks (1932)

$$\lambda(x) = \frac{\exp(a_\lambda x + b_\lambda)}{1+\exp(a_\lambda x + c_\lambda)} + d_\lambda, \tag{9}$$

with $a_\lambda > 0$, $b_\lambda, c_\lambda \in \mathbb{R} \cup \{-\infty\}$ and $d_\lambda \geq 0$.

Experience from insurers shows that the intensity of incidence in LTC increases exponentially with respect to age (SCOR, 1995). At higher ages, data becomes scarcer. As LTC is linked to ageing, it is reasonable to expect that the behaviours of mortality and morbidity are quite similar and that an exponential or logistic form is suited to model incidence in LTC. The logistic model has already been used to this extent, e.g., in Rickayzen and Walsh (2002). Let us notice that the exponential models introduced in Gompertz (1825) and Makeham (1867) may be seen as limit cases of the logistic model, for which $c_\lambda = -\infty$.

For an individual p defined by their age of entry in the portfolio $x_p \geq 0$, their age of exit $y_p > x_p$ and the associated exit cause $c_p \in \{0, 1, 2\}$ the log-likelihood has the following expression:

$$\begin{aligned} l_p(\lambda) &= \log\left[\exp\left(-\int_{x_p}^{y_p} \lambda(u)\mathrm{d}u\right)\lambda(y_p)^{\delta^2_{c_p}}\right] \\ &= \delta^2_{c_p}\log(\lambda(y_p)) - \int_{x_p}^{y_p}\lambda(u)\mathrm{d}u \\ &= \delta^2_{c_p}\log\left[\frac{\exp(a_\lambda y_p + b_\lambda)}{1+\exp(a_\lambda y_p + c_\lambda)} + d_\lambda\right] \\ &\quad - \frac{\exp(b_\lambda - c_\lambda)}{a_\lambda}\log\left[\frac{1+\exp(a_\lambda y_p + c_\lambda)}{1+\exp(a_\lambda x_p + c_\lambda)}\right] - d_\lambda(y_p - x_p), \end{aligned}$$

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

where for $k, l \in \mathbb{N}$, $\delta_k^l = \begin{cases} 1 & \text{if } k = l, \\ 0 & \text{otherwise.} \end{cases}$

2.4.3. *Intensity of mortality in LTC.* Disability may be caused by a wide range of underlying pathologies. Unfortunately most of the time those pathologies are not available in the data. This results in heterogeneity among disabled people. In this section, we provide a simple parametric model which accounts for the heterogeneity caused by pathologies. In order to do this, we must rely on several strong assumptions. First of all, we assume that underlying pathologies can be divided into two main groups. On the one hand, we have pathologies associated with very high mortality such as terminal cancer mainly or more rarely respiratory diseases. For such diseases, life expectancy at the onset of LTC is within a few months. On the other hand, dementia, neurological or cardiovascular diseases have an associated life expectancy which is closer to 5 years. We further assume that among each group the population can be considered as homogeneous. We could consider three or more groups of pathologies but then inference of parameters would prove extremely difficult and this would be at the expense of robustness in the estimation procedure.

We then consider an additive model for the intensity of mortality in LTC, so that the mortality within each group is the sum of a common term (which may be for example the autonomous mortality at the current age) plus a term that only depends on the pathology group and the age at onset of LTC. The underlying assumption to this additive model of mortality is that people who become disabled have increased mortality from the pathology that caused disability but are still exposed to other causes of death. Also, considering that the additional mortality depends on the age at onset of LTC rather than on the current age may seem very restrictive and the model may not be accurate for very high duration in the disabled state. However, due to the very high mortality in the disabled state, cases of exceptional longevity should remain rare enough and the resulting impact quite limited. Under those assumptions, the resulting intensity of mortality in LTC takes a very peculiar form as we show in the following lemma.

Lemma 3. *Let us consider a model with two distinct states of disability I_1 and I_2, such that the respective transition intensities from autonomy to those states are λ_1 and λ_2 respectively and no transition is allowed between those states or back to autonomy (see Figure 3). Let us further assume that the intensity of mortality in state I_k can be written as*

$$\mu_{i,k}(x, t) = \mu_0(x, t) + \Delta_k(x),$$

with $x \geq x_0$, $t \geq 0$ where $\mu_0(x, t)$ is a common mortality term and $\Delta_k(x)$ a state-specific mortality term, for $k \in \{1, 2\}$.

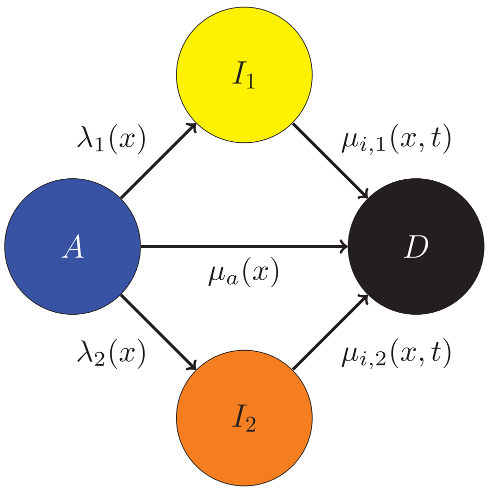

FIGURE 3: Model with two completely separate states of LTC. (Color online)

Then, to ensure the embedding of the 3-states model in this model, the following relations must be satisfied for all $x > x_0$, $t \geq 0$:

$$\lambda(x) = \lambda_1(x) + \lambda_2(x)$$

$$\mu_i(x,t) = \begin{cases} \mu_0(x,t) + \Delta_1(x) + \frac{\Delta_2(x)-\Delta_1(x)}{1+\frac{\lambda_1(x)}{\lambda_2(x)}\exp([\Delta_2(x)-\Delta_1(x)]t)} & \textit{if } \lambda_2(x) \neq 0 \\ \mu_0(x,t) + \Delta_1(x) & \textit{otherwise.} \end{cases}$$

Proof. The first relation on the incidences in LTC is obvious, as well as the case where $\lambda_2(x) = 0$. For the second relation, let us denote by $\eta_k(x,t)$ the proportion of disabled people in state I_k among the population of people who became disabled at age $x \geq x_0$ and then survived for a time $t \geq 0$.

Let us define for $x \geq x_0$, $t \geq 0$, $k \in \{1,2\}$ and $h > 0$

$$\eta_k(x,t) = \frac{P(Z_{x+t} = I_k | Z_{x^-} = A, Z_x \in \{I_1, I_2\})}{P(Z_{x+t} \in \{I_1, I_2\} | Z_{x^-} = A, Z_x \in \{I_1, I_2\})},$$

and

$$\eta_k(x,t,h) = \frac{P(Z_{x+t} = I_k | Z_{x-h} = A, Z_x \in \{I_1, I_2\})}{P(Z_{x+t} \in \{I_1, I_2\} | Z_{x-h} = A, Z_x \in \{I_1, I_2\})}.$$

On the one hand, we have

$$\eta_k(x,t,h) \underset{h \to 0}{\to} \eta_k(x,t),$$

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

and on the other hand

$$\eta_k(x,t,h) = \frac{P(Z_{x+t} = I_k, Z_{x-h} = A, Z_x \in \{I_1, I_2\})}{P(Z_{x+t} \in \{I_1, I_2\}, Z_{x-h} = A, Z_x \in \{I_1, I_2\})}$$
$$= \frac{P(Z_{x+t} = I_k, Z_x = I_k, Z_{x-h} = A)}{\sum_{l=1}^{2} P(Z_{x+t} = I_l, Z_x = I_l, Z_{x-h} = A)}$$

$$= \frac{P(Z_{x-h} = A)\, P(Z_x = I_k | Z_{x-h} = A)\, P(Z_{x+t} = I_k | Z_x = I_k, Z_{x-h} = A)}{\sum_{l=1}^{2} P(Z_{x-h} = A)\, P(Z_x = I_l | Z_{x-h} = A)\, P(Z_{x+t} = I_l | Z_x = I_l, Z_{x-h} = A)}$$
$$= \frac{P(Z_x = I_k | Z_{x-h} = A)\, P(Z_{x+t} = I_k | Z_x = I_k, Z_{x-h} = A)}{\sum_{l=1}^{2} P(Z_x = I_l | Z_{x-h} = A)\, P(Z_{x+t} = I_l | Z_x = I_l, Z_{x-h} = A)}$$
$$\underset{h\to 0}{\rightarrow} \frac{\lambda_k(x) \exp\left(-\int_0^t \mu_{i,k}(x,u)\mathrm{d}u\right)}{\sum_{l=1}^{2} \lambda_l(x) \exp\left(-\int_0^t \mu_{i,l}(x,u)\mathrm{d}u\right)}$$
$$= \frac{\lambda_k(x) \exp\left(-\int_0^t \mu_0(x,u)\mathrm{d}u\right) \times \exp\left[-\Delta_k(x) \times t\right]}{\sum_{l=1}^{2} \lambda_l(x) \exp\left(-\int_0^t \mu_0(x,u)\mathrm{d}u\right) \times \exp\left[-\Delta_l(x) \times t\right]}$$
$$= \frac{\lambda_k(x) \exp\left[-\Delta_k(x) \times t\right]}{\sum_{l=1}^{2} \lambda_l(x) \exp\left[-\Delta_l(x) \times t\right]}.$$

By uniqueness of the limit we obtain

$$\eta_k(x,t) = \frac{\lambda_k(x) \exp\left[-\Delta_k(x) \times t\right]}{\sum_{l=1}^{2} \lambda_l(x) \exp\left[-\Delta_l(x) \times t\right]}.$$

Now, the intensity of mortality for the population of disabled people is

$$\begin{aligned}
\mu_i(x,t) &= \sum_{k=1}^{2} \eta_k(x,t)\mu_{i,k}(x,t) \\
&= \sum_{k=1}^{2} \frac{\lambda_k(x)\exp(-\Delta_k(x)t)}{\sum_{l=0}^{2}\lambda_l(x)\exp(-\Delta_l(x)t)}\mu_{i,k}(x,t)
\end{aligned}$$

$$\begin{aligned}
&= \mu_0(x,t) + \sum_{k=1}^{2} \frac{\lambda_k(x)\exp(-\Delta_k(x)t)}{\sum_{l=0}^{2}\lambda_l(x)\exp(-\Delta_l(x)t)}\Delta_k(x) \\
&= \mu_0(x,t) + \Delta_1(x) \\
&\quad + \frac{\lambda_2(x)\exp(-\Delta_2(x)t)}{\lambda_1(x)\exp(-\Delta_1(x)t) + \lambda_2(x)\exp(-\Delta_2(x)t)}[\Delta_2(x) - \Delta_1(x)] \\
&= \mu_0(x,t) + \Delta_1(x) + \frac{\Delta_2(x) - \Delta_1(x)}{1 + \frac{\lambda_1(x)}{\lambda_2(x)}\exp\{[\Delta_2(x) - \Delta_1(x)]\,t\}},
\end{aligned}$$

which proves the lemma. ■

In what follows, we assume that the assumptions of the lemma are satisfied and that the common mortality term is the intensity of mortality for autonomous people μ_a at the same current age $x + t$. Let us denote, for $x \geq x_0$

$$\theta(x) = \frac{\lambda_2(x)}{\lambda(x)}.$$

We now have

$$\mu_i(x,t) = \mu_a(x+t) + \Delta_1(x) + \frac{\theta(x)\,[\Delta_2(x) - \Delta_1(x)]}{\theta(x) + [1-\theta(x)]\exp\{[\Delta_2(x) - \Delta_1(x)]\,t\}},$$

and thus

$$\Delta(x,t) = \Delta_1(x) + \frac{\theta(x)\,[\Delta_2(x) - \Delta_1(x)]}{\theta(x) + [1-\theta(x)]\exp\{[\Delta_2(x) - \Delta_1(x)]\,t\}}. \tag{10}$$

The log-likelihood associated with survival in LTC for an individual p with an age of entry in LTC $x_p \geq 0$, an age of exit $y_p > x_p$ and the associated cause

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

of exit $c_p \in \{0, 1\}$ then takes the following expression:

$$
\begin{aligned}
&l_p(\mu_a, \Delta_1, \Delta_2, \theta)\\
&\quad = \log\left[\exp\left(-\int_{x_p}^{y_p} \mu_i(x_p, u - x_p)\mathrm{d}u\right)\mu_i(x_p, y_p - x_p)^{\delta^1_{c_p}}\right]\\
&\quad = \delta^1_{c_p}\log(\mu_i(x_p, y_p - x_p)) - \int_{x_p}^{y_p} \mu_i(x_p, u - x_p)\mathrm{d}u
\end{aligned}
$$

$$
\begin{aligned}
= \delta^1_{c_p}\log\Bigg(&\mu_a(y_p)\\
&+ \Delta_1(x_p) + \frac{\theta(x_p)\left[\Delta_2(x_p) - \Delta_1(x_p)\right]}{\theta(x_p) + \left[1 - \theta(x_p)\right]\exp(\left[\Delta_2(x_p) - \Delta_1(x_p)\right]\left[y_p - x_p\right])}\Bigg)\\
&- \int_{x_p}^{y_p} \mu_a(u)\mathrm{d}u - \Delta_2(x_p)\left[y_p - x_p\right]\\
&+ \log\left\{\theta(x_p) + \left[1 - \theta(x_p)\right]\exp(\left[\Delta_2(x_p) - \Delta_1(x_p)\right]\left[y_p - x_p\right])\right\}.
\end{aligned}
$$

For a given μ_a, the previous log-likelihood may then be computed analytically, which allows for the estimation of $\Delta_1, \Delta_2, \theta$ and then Δ by plugging in Equation (10) using maximum likelihood.

2.4.4. *Intensity of autonomous mortality.* The maximum likelihood method in the previous section requires to know the autonomous mortality beforehand. We therefore need to compute an intermediary estimate of the autonomous mortality whose sole purpose is the estimation of Δ. Indeed, the ultimate autonomous mortality is then computed thanks to Lemma 2.

Once again we rely on the logistic model introduced in Beard (1959, 1971) and Perks (1932)

$$\mu_a(x) = \frac{\exp(a_a x + b_a)}{1 + \exp(a_a x + c_a)} + d_a, \tag{11}$$

with $a_a > 0$, $b_a, c_a \in \mathbb{R} \cup \{-\infty\}$ and $d_a \geq 0$.

For an individual p defined by their age of entry in the portfolio $x_p \geq 0$, their age of exit $y_p > x_p$ and the associated exit cause $c_p \in \{0, 1, 2\}$ the log-likelihood

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

has the following expression similar to Section 2.4.2:

$$
\begin{aligned}
l_p(\mu_a) = &\log\left[\exp\left(-\int_{x_p}^{y_p}\mu_a(u)\mathrm{d}u\right)\mu_a(y_p)^{\delta^1_{c_p}}\right]\\
=&\delta^1_{c_p}\log\left[\frac{\exp(a_a y_p + b_a)}{1+\exp(a_a y_p + c_a)} + d_a\right]\\
&-\frac{\exp(b_a - c_a)}{a_a}\log\left[\frac{1+\exp(a_a y_p + c_a)}{1+\exp(a_a x_p + c_a)}\right] - d_a(y_p - x_p).
\end{aligned}
$$

2.5. Parameters estimation procedure

To estimate the parameters, we use the following procedure

1. We estimate the parameters for the intensity of general mortality $\widehat{\mu_g}$ by using the individuals of both databases and the Brass relational model (as in Brass, 1971) in order to get a robust estimate of the intensity of general mortality with convergence towards a reference mortality table at higher ages (see Section 2.4.1).
2. We estimate the parameters for the intensity of incidence in LTC $\widehat{\lambda}$ (resp. a first-step estimate of the autonomous mortality $\widehat{\mu_a}^{(1)}$), using the contributors database and the Perks logistic model (as in Perks, 1932). More precisely $\widehat{\lambda}$ (resp. $\widehat{\mu_a}^{(1)}$) is the maximum likelihood estimator (MLE) constructed by summing over the individuals the log-likelihoods given in Section 2.4.2 (resp. Section 2.4.4).
3. We estimate the parameters for the additional mortality in LTC $\widehat{\Delta}$ from $\widehat{\mu_a}^{(1)}$ and the annuitant database, using the MLE constructed by summing over the individuals the log-likelihoods given in Section 2.4.3. Several parametric forms for $\widehat{\Delta}$ will be tested in the next section.
4. Thanks to Equation (5), we compute the value of a second-step estimator for the intensity for autonomous mortality $\widehat{\mu_a}^{(2)}$, relying on $\widehat{\lambda}$, $\widehat{\Delta}$, $\widehat{\mu_g}$ and using numerical methods to approximate the outer integrals in (5). This second-step estimator should give more reliable results at higher ages where no data is available.

A summary of the procedure is provided in Figure 4.

Furthermore when we deal with complex models, it can be very interesting to compare them to some of their sub-models to see if the use of many parameters is justified. To this extent, we can rely on the *Bayesian Information Criterion (BIC)*. For a model m_i characterized by a number of parameters k_i and a log-likelihood function l_i maximized at θ_i, the expression of the criterion is as follows:

$$\mathrm{BIC}_i = -2l_i(\theta_i) + k_i\log(n),$$

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

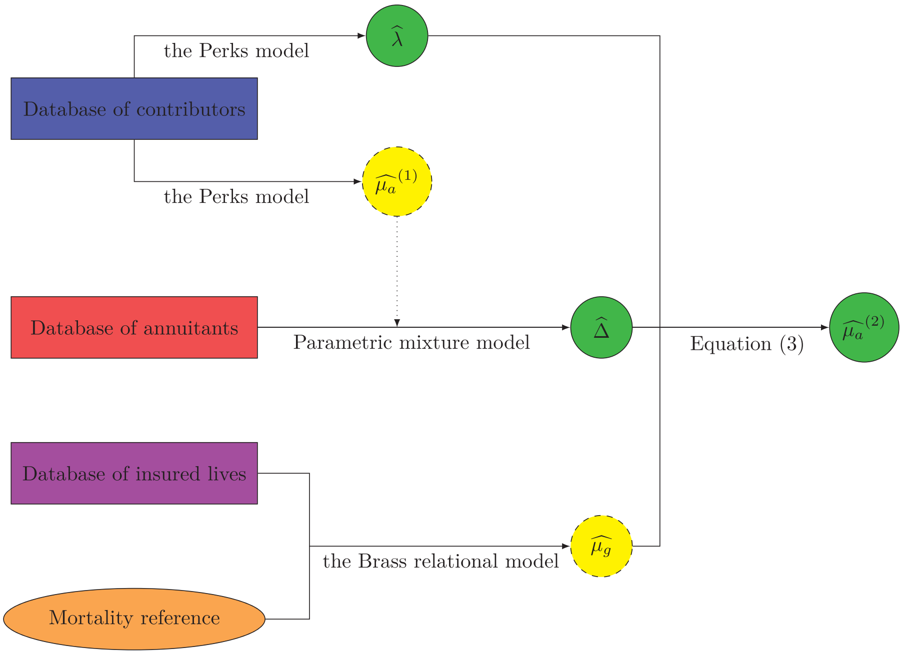

FIGURE 4: Procedure for the estimation of biometric laws. Dashed (resp. plain) circles represent intermediary (resp. final) estimates of biometric laws. Database of insured lives is obtained by merging the contributors and annuitants databases. (Color online)

where n represents the number of observed transitions in the expression of the likelihood. The choice of the coefficient in front of the number of parameters $\log(n)$ differs from the one made in the original Akaike's Information Criterion (AIC) where this coefficient is 1. Also, let us note that in the version of the criterion, n is the number of observed transitions and not the number of individuals as in the original criterion. The interest in introducing this modification in the case of censored data is discussed in Volinsky and Raftery (2000). Using the BIC, we are able to compare models, the model with the lower BIC being the "best" model. In the next section, we use the BIC to challenge the use of more complex parametric models introduced in this section and lower the overall number of parameters.

2.6. Pricing and reserving

We consider a product where the autonomous insured lives pay a fixed amount of premium while they are autonomous. Should they become disabled, the premium is no longer due and they are entitled to an annuity instead. On the French LTC insurance market most products rely on monthly premium and monthly benefit. For the sake of simplicity, we consider continuous-time premiums and annuities instead, the difference with monthly quantities being extremely low.

We denote by τ the continuous-time actuarial interest rate used to compute discounted cash flows.

Let us introduce additional notation

$$A(x, y) = P(Z_y = A | Z_x = A) = \frac{A_y}{A_x} = \exp\left(-\int_x^y [\mu_a(u) + \lambda(u)]\, \mathrm{d}u\right),$$

$$I_x(t, s) = P(Z_{x+s} = I | Z_{x-} = A, Z_x = I, Z_{x+t} = I) = \exp\left(-\int_t^s \mu_i(x, u)\mathrm{d}u\right)$$

and

$$\overline{A}(x, y) = e^{-\tau(y-x)} A(x, y) = \exp\left(-\int_x^y [\mu_a(u) + \lambda(u) + \tau]\, \mathrm{d}u\right),$$

$$\overline{I}_x(t, s) = e^{-\tau(s-t)} I_x(t, s) = \exp\left(-\int_t^s [\mu_i(x, u) + \tau]\, \mathrm{d}u\right)$$

for $x_0 \leq x \leq y$ and $0 \leq t \leq s$ such that A (resp. I) is the survival probability in the state of autonomy (resp. in the disabled state) and $\overline{A}$ (resp. $\overline{I}$) the discounted survival probability in the aforementioned state.

We define the following quantities that are required for the pricing of the product:

- $P(x)$, the expected value of insured liabilities for an autonomous insured life with current age x for a 1€ yearly premium

$$P(x) = \int_x^\infty \overline{A}(x, u)\mathrm{d}u.$$

- $\mathrm{RFC}(x, t)$, the expected value of insurer liabilities for a disabled insured life with an age x at the onset of LTC and a time t spent in the disabled state and a 1€ yearly annuity, also called reserve for claim

$$\mathrm{RFC}(x, t) = \int_t^\infty \overline{I}_x(t, u)\mathrm{d}u.$$

- $\Pi(x)$, the expected value of insurer liabilities for an autonomous insured life with current age x, associated with a 1€ yearly annuity

$$\Pi(x) = \int_x^\infty \lambda(u)\overline{A}(x, u)\mathrm{RFC}(u, 0)\mathrm{d}u.$$

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

- The stability premium $p^*(x)$. It is the value of premium that matches insurer and insured liabilities at the time of subscribing. For an age x at subscribing, we have
$$p^*(x) = \frac{\Pi(x)}{P(x)}.$$
- The reserve for premium (RFP), which is constituted for autonomous people. Its amount is equal to the expected value of future liabilities minus the expected value of premium. For an insured of age at subscribing x_s, current age x, the associated amount of reserve is
$$\mathrm{RFP}(x_s, x) = P(x)\left[p^*(x) - p^*(x_s)\right].$$

3. Results

In this section, we provide an example using aggregated data from several French LTC insurance portfolios. The definition used for LTC is 3ADL4 which means that an insured life is considered disabled if he/she has permanently lost the ability to do on their own at least three out of the four activities of daily living defined by the contract: functional mobility, dressing, bathing and eating. The portfolio we consider contains a very large number of policies and covers a relatively long period. The date of extraction is November 31, 2014 for both contributors and annuitants. We remove the first 3 years spent by contributors in the portfolio and then consider a 12-year observation period between January 1, 2002 and December 31, 2013 for contributors and a 20-year observation period between January 1, 1994 and December 31, 2013 for annuitants. Database of contributors contains over 1.5 million years of exposure with 69.8% of the lines being right censored. Database of annuitants contains close to 45,000 years of exposure and 29.4% of right censored lines. Women account for 65.4% of contributors and 66.7% of annuitants. Separate models are estimated for men and women.

3.1. General mortality

We use the Brass relational model with data for the French population over the years 2010–2013 coming from the Human Mortality Database (University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany), 2015) that we choose as our mortality reference. At the same time, we compute empirical probabilities of death, using the Hoem estimator as described in Planchet and Thérond (2006), as well as 95% confidence intervals under the normal approximation, over the age range where the Cochran criterion is satisfied. Figure 5 displays the logarithm of the CDO for empirical probabilities and the mortality reference and the difference between them. As it is close to a straight line, the underlying assumption of the model is satisfied.

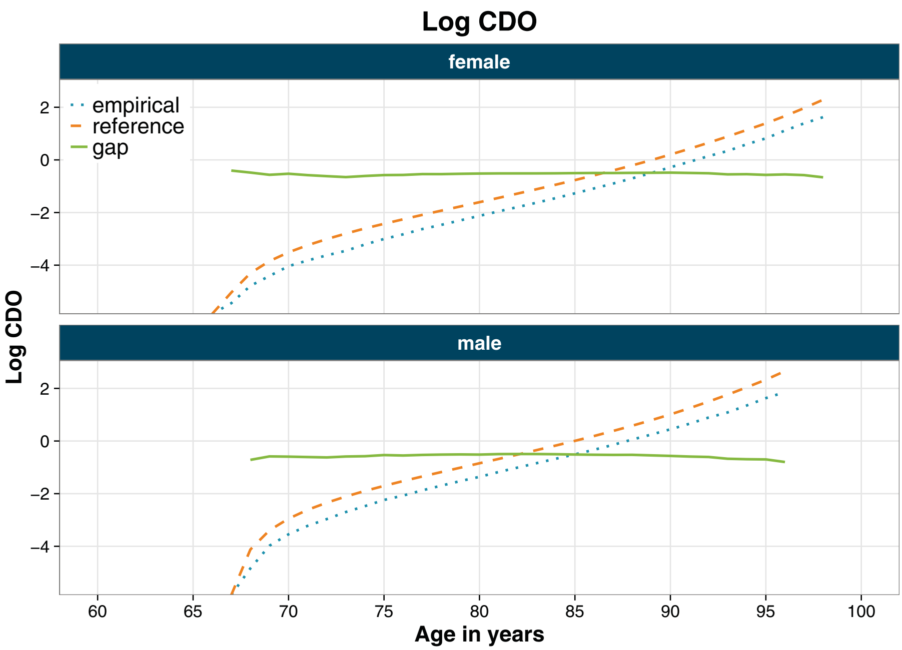

FIGURE 5: Logarithm of cumulative distribution odds (CDO) for observed mortality (dotted), reference mortality (dashed) and their difference (plain). (Color online)

TABLE 2

VALUE OF LOG-LIKELIHOOD l AND BIC OF PREVIOUSLY INTRODUCED MODELS FOR THE INCIDENCE IN LTC.

Model	Intensity	l(males)	BIC(males)	l(females)	BIC(females)
Gompertz	$\lambda(x) = e^{a_\lambda x + b_\lambda}$	−25,099.12	50,223.64	−52,380.25	104,788.21
Makeham	$\lambda(x) = e^{a_\lambda x + b_\lambda} + d_\lambda$	−25,099.12	50,232.11	−52,380.25	104,797.44
Beard	$\lambda(x) = \frac{e^{a_\lambda x + b_\lambda}}{1 + e^{a_\lambda x + c_\lambda}}$	−25,094.61	**50,223.09**	−52,326.58	**104,690.09**
Perks	$\lambda(x) = \frac{e^{a_\lambda x + b_\lambda}}{1 + e^{a_\lambda x + c_\lambda}} + d_\lambda$	−25,093.74	50,229.81	−52,325.83	104,697.83

Figure 6 represents the observed and reference mortality, as well as the mortality fitted using the Brass relational model. The latter mortality is close to the observed mortality for ages where enough data is available and then smoothly converges towards the reference at higher ages, where no data is available.

3.2. Incidence in LTC

The results for the estimation of incidence in LTC can be found in Table 2. The Gompertz (resp. the Beard) model performs better than the Makeham (resp. the Perks) model according to the BIC, which means that the use of an extra parameter which represents an initial level of incidence present at all ages is not required. In addition, the Beard logistic model is a better fit to the data than

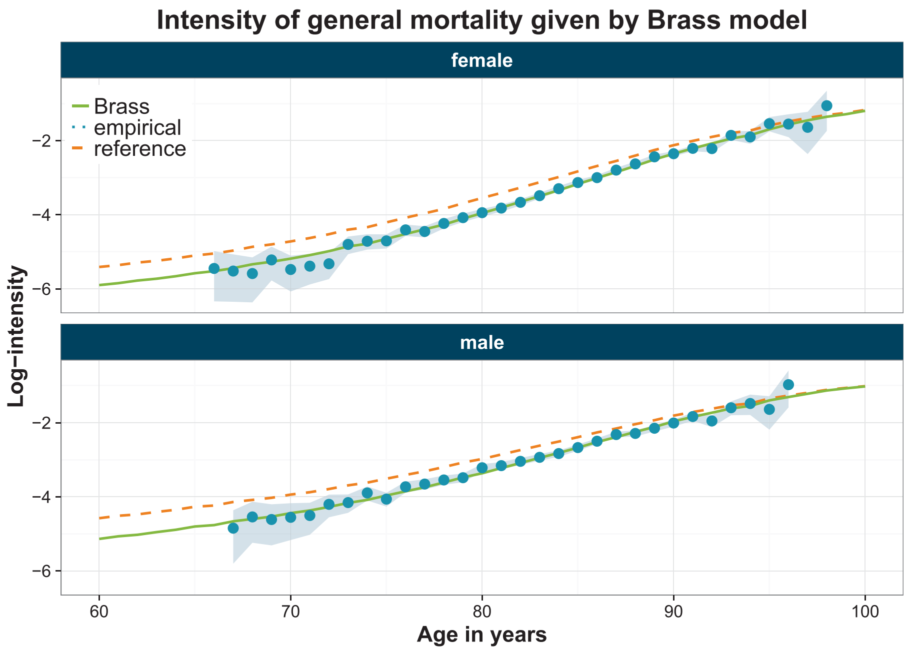

FIGURE 6: Intensity of mortality estimated from the data (dots) with 95% confidence intervals, from the mortality reference (dashed) and given by the Brass relational model (plain). (Color online)

the Gompertz exponential model. One can come to this conclusion by looking at Figure 7 that represents the empirical incidence as well as the inferred incidence for the Gompertz and the Beard models. The empirical incidence in LTC increases exponentially at first but at higher ages there is a slowing down in this increase, more pronounced for females than for males, which makes the Beard logistic model a better fit.

3.3. Mortality in LTC

We rely on the results from Section 2.4.3 and define the mortality in LTC by providing a parametric model for Δ_1, Δ_2 and θ. In this section, we only focus on a handful of well-known models that in our opinion are the most obvious candidates. Furthermore, for the sake of simplicity, we only consider models where Δ_1 and Δ_2 take the same parametric form. For $\Delta_1(x)$ and $\Delta_2(x)$, we consider constant, the Gompertz and the Makeham exponential models as well as the Beard and the Perks logistic models (so five different models in total). For $\theta(x)$, we have the additional constraint that we should have $0 \leq \theta(x) \leq 1$ for all ages. We consider a constant model and then four logistic models with increasing degrees of freedom. Indeed, the full logistic law has four parameters and therefore four degrees of freedom. By setting the ultimate values for $x = -\infty$ and $x = +\infty$ respectively to 0 and 1, we obtain a logistic model with only two

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

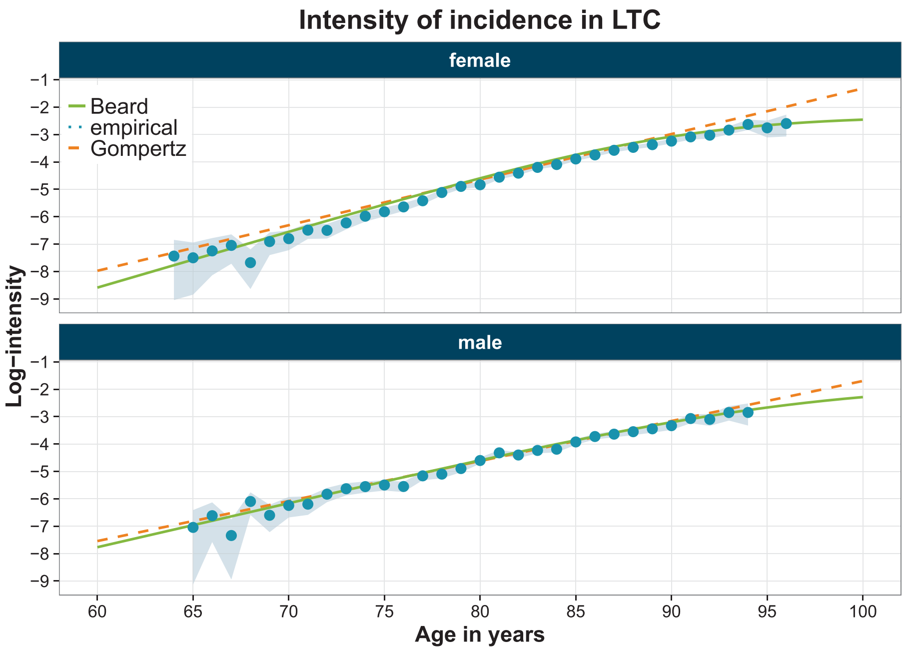

FIGURE 7: Estimates of the incidence in LTC. Dots and ribbon represent empirical estimates with 95% confidence intervals. Plain (resp. dashed) line represents the Gompertz (resp. the Beard) model fitted to the data. (Color online)

parameters. We may relax either of those constraints by introducing additional parameters $0 \leq \alpha \leq \beta \leq 1$ so that α (resp. β) is the ultimate value of $\theta(x)$ when $x = -\infty$ (resp. $x = +\infty$). Hence, we estimate $5 \times 5 = 25$ combinations of models. Results are available in Table A1 in the appendix.

Figure 8 represents each of the models on the angle of the number of parameters and maximum log-likelihood. As the BIC is a linear combination of the two aforementioned components, contour curves of increasing BIC correspond to parallel lines of increasing intercept in this representation. The model with the best BIC is such that there is no other model in the upper half two-dimensional space delimited by the associated contour curve. Distance from any model to this contour curve is proportional to the difference in BIC between that model and the best model. The criterion selects model 9 for males (models 7 and 10 being close contenders) and model 10 for females. All those models rely on the Gompertz law for the specific mortality terms Δ_1 and Δ_2. As regards θ, model 10 uses the full four parameters logistic model, whereas model 9 only uses three parameters, the asymptotic value for the prevalence of high mortality (group 2) pathologies at lower ages being set to 100%. From this point all results are based on the parameters inferred for models 9 for males and model 10 for females.

Figure 9 displays the value of $\theta(x)$ that represents the prevalence of high mortality pathologies among newly disabled people inferred by the model. This

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

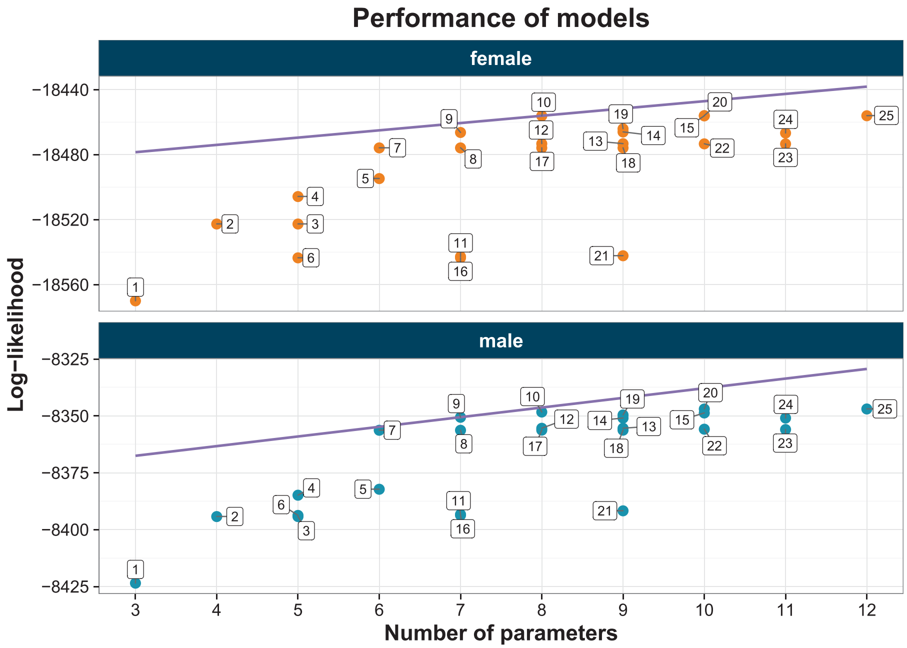

FIGURE 8: Representation of the models (dots) in the plan of log-likelihood and number of parameters. The plain line represents the contour curve for the model with the highest BIC. (Color online)

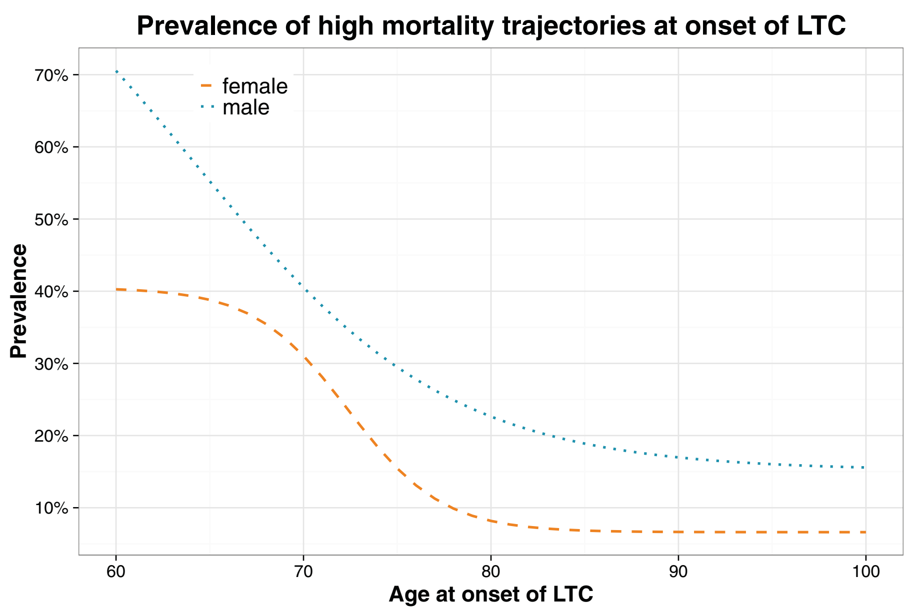

FIGURE 9: Prevalence of high mortality trajectories in the population of newly disabled inferred by the model. (Color online)

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

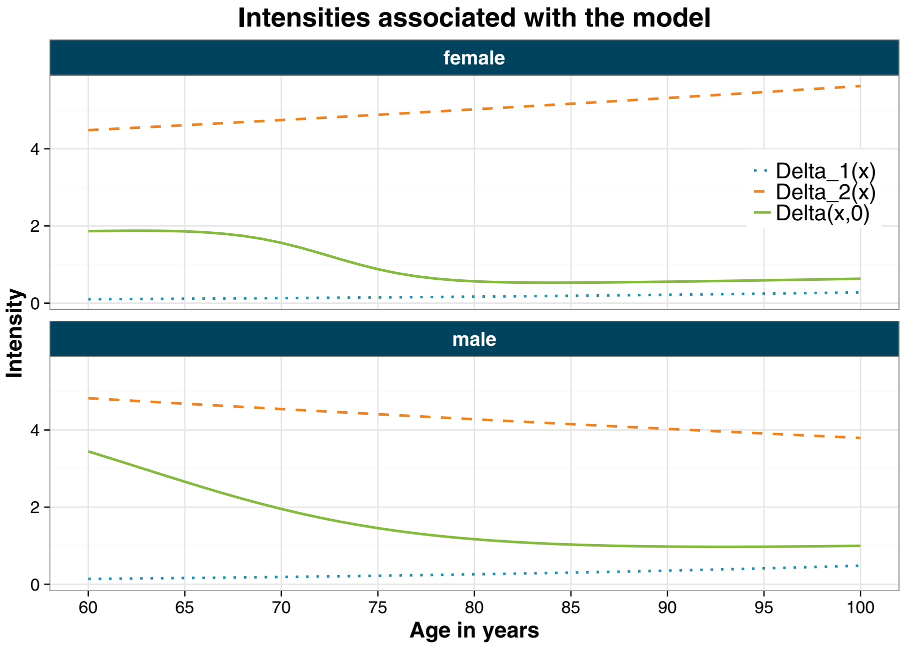

FIGURE 10: Specific mortality terms for both populations in the mixture (dotted and dashed lines), and resulting mortality at the onset of LTC (plain line). (Color online)

prevalence decreases with age and is much higher for males (70% at age 60 and 17% at age 90) than for females (40% at age 60 and 6% at age 90). Figure 10 represents the specific mortality terms $\Delta_1(x)$ and $\Delta_2(x)$ and the resulting mortality term for the newly disabled $\Delta(x, 0)$, which is the weighted mean of $\Delta_1(x)$ and $\Delta_2(x)$ with weights $1-\theta(x)$ and $\theta(x)$ respectively. We observe that $\Delta_2(x)$ is way higher than $\Delta_1(x)$. Besides, the initial mortality $\Delta(x, 0)$ decreases with age until 85 and then remains stable. Let us remind that for higher durations, $\Delta(x, t)$ converges towards the lower value between $\Delta_1(x)$ and $\Delta_2(x)$ as the weight of the population with higher mortality in the mixture decreases to 0. Those results seem compatible with our interpretation in terms of cancer for the group of high mortality pathologies and dementia as well as cardiovascular and neurological diseases for the other group. However, one should keep in mind that pathologies are not actually observed in the data, and Figures 9 and 10 only represent the underlying distribution inferred by the model.

Figure 11 represents annual death probabilities associated with the empirical data on the one hand and given by the model on the other hand. We compute empirical annual probabilities by grouping disabled people, according to their age of entry in LTC with 5-year age bands between 65 and 90. For each age band, we then compute annual death probabilities by duration under the assumption that the intensity of mortality is constant over intervals of 1 year for the duration. We represent 95% confidence intervals for those probabilities

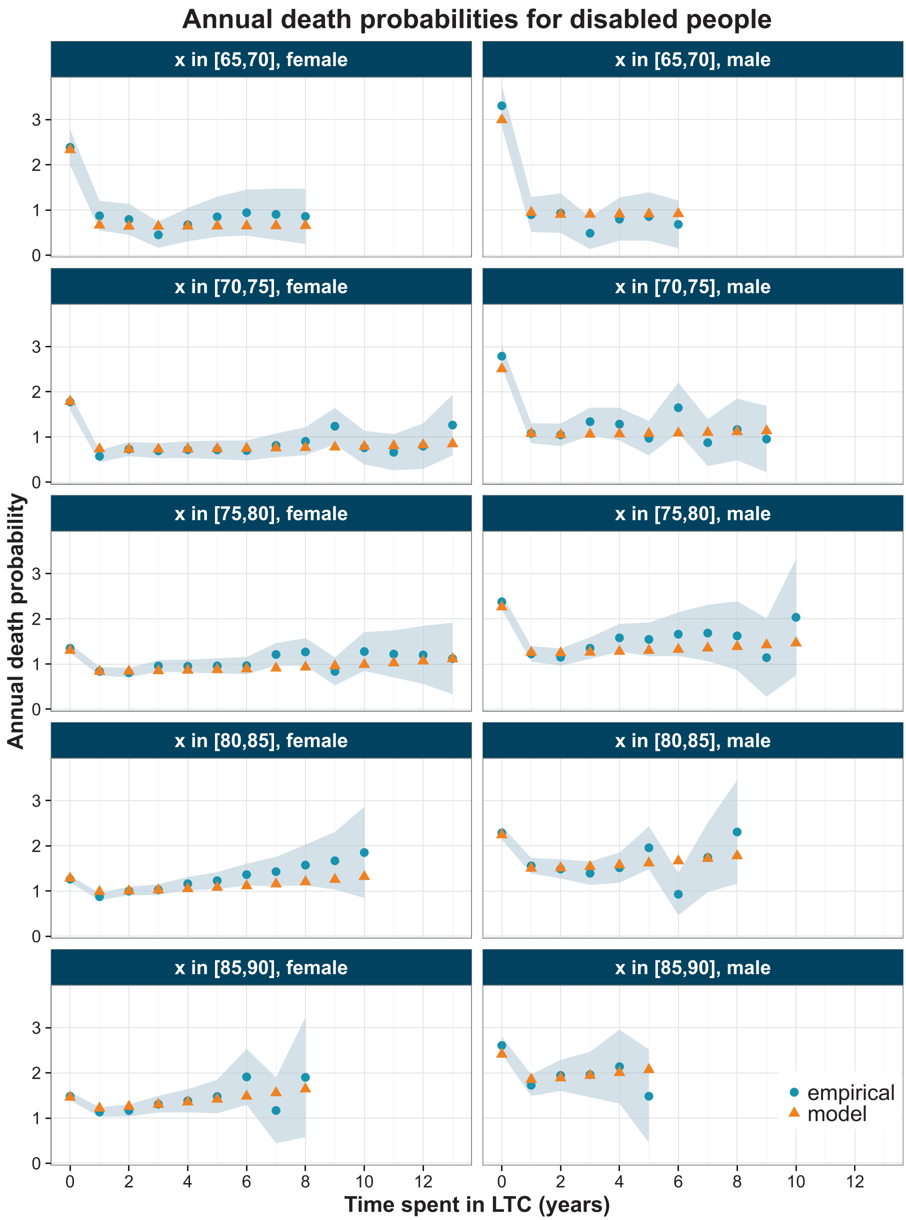

FIGURE 11: Consecutive death probabilities for disabled people according to the model (triangles) with empirical probabilities (circles) and associated 95% confidence intervals. The y-scale has been re-normalized to preserve confidentiality of results. (Color online)

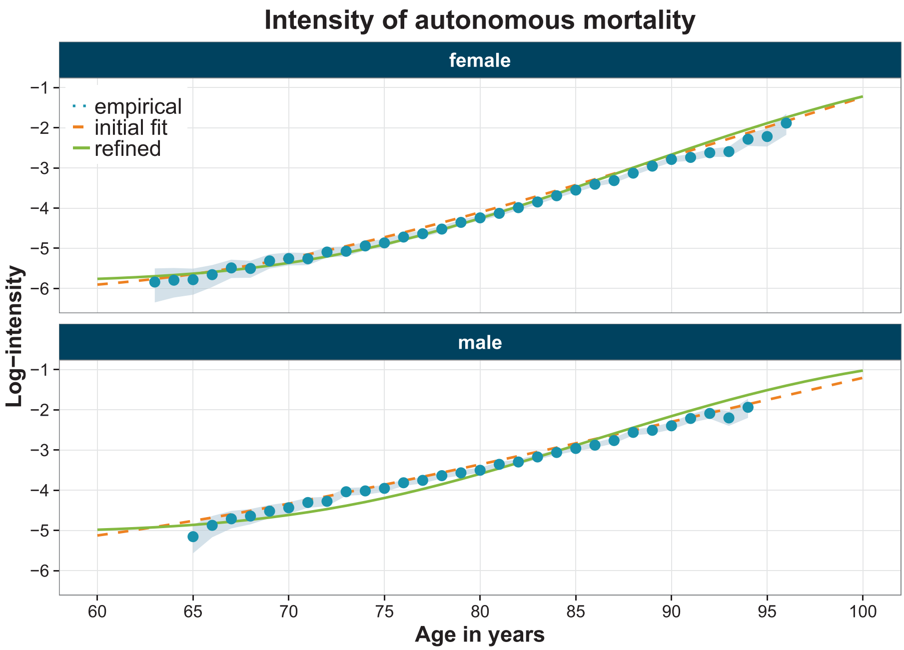

FIGURE 12: Intensity of autonomous mortality. Dots: empirical rates; dashed line: direct fit of the Perks model; plain line: refined intensity from Equation (5). (Color online)

under the normal approximation (as in Planchet and Thérond, 2006). We also compute annual death probabilities given by the model for individuals of ages 67.5, 72.5,..., 87.5 at onset of LTC. Confidence intervals are still very wide, especially for men as well as at lower/higher age at onset of LTC and/or high duration in LTC. Nonetheless, the annual probabilities computed using the model appear to match the empirical probabilities very well for both males and females when data is available. It appears that by taking into account a mixture component in the model, we were able to reproduce the evolution of death probabilities with respect to time spent in LTC. Nevertheless, in each component of the mixture, time spent in LTC only appears in the autonomous mortality term, through the current age $x + t$.

3.4. Autonomous mortality

Figure 12 represents the initial intensity of autonomous mortality we get from the first-step estimator $\widehat{\mu_a}^{(1)}$ and the refined intensity from the second-step estimator $\widehat{\mu_a}^{(2)}$. The refined intensity remains close to the empirical intensity for females but for males there are some divergences which can be explained by the lower volume of data.

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

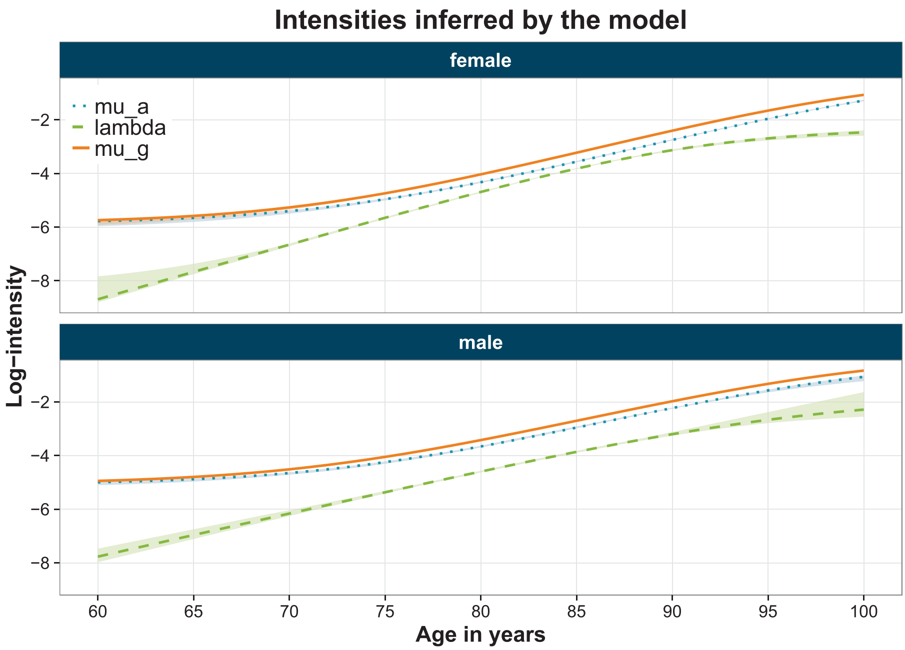

FIGURE 13: Intensities of general mortality (plain), autonomous mortality (dotted) and incidence in LTC (dashed), with 95% confidence intervals obtained by bootstrap. (Color online)

3.5. Summary of intensities and prevalence of LTC

In order to assess the robustness of the estimation performed, we use a non-parametric quantile bootstrap method. From the initial database of insured lives, we build 200 new samples by drawing, with replacement, as many individuals as in the initial observation database. For each sample, we then run all the steps of the estimation procedure, including the choice of the best model according to the BIC. We then use the inferred parameters to compute the final intensities of transition as well as the prevalence of LTC. Finally, for each age, we select the 2.5% and 97.5% quantiles of the empirical distribution of those quantities in order to get bootstrap confidence interval.

Figure 13 represents the intensity of mortality for the general population, the intensity of mortality for autonomous people as well as the incidence in LTC. Confidence intervals are very tight for autonomous and general mortality. For the incidence in LTC they are larger, especially at lower or higher ages, and for males as the data is scarcer. Figure 14 represents the prevalence of LTC among the general population inferred by the model. The prevalence increases almost exponentially with respect to age at first, with a slowing down at higher ages. Prevalence is initially close for males and females, but from the age of 80 it becomes much higher for females. Overall, the confidence intervals are very large especially for males at higher ages where the number of survivors is limited.

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

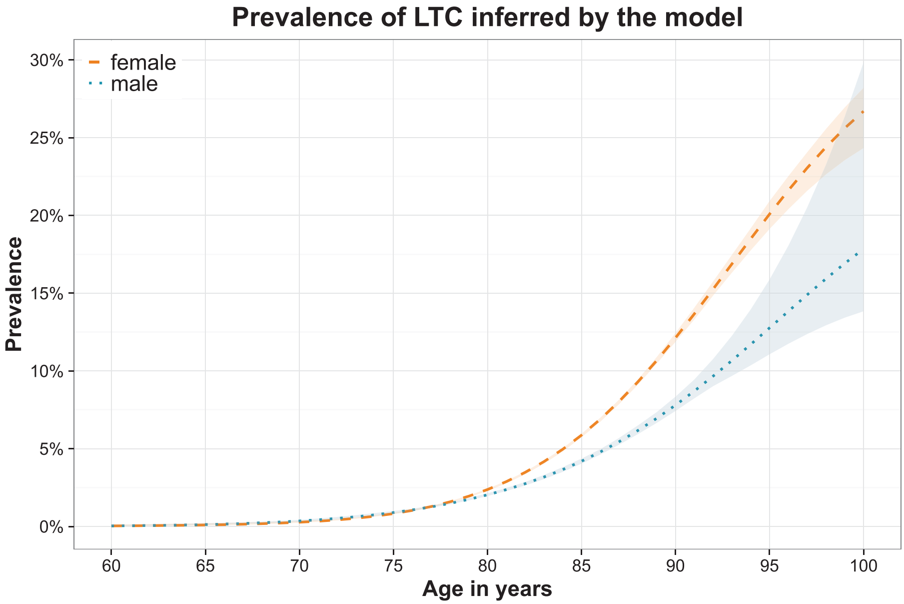

FIGURE 14: Prevalence of LTC by age in the general population (plain line), with 95% confidence intervals obtained by bootstrap. (Color online)

3.6. Results of pricing and reserving

We consider a LTC insurance product where autonomous policyholders pay a monthly level premium, whose amount is set based on their age of subscription. Should they become disabled, they would stop paying the premium and instead receive a monthly annuity of 1,000€ until they die. We use an actuarial interest rate of 1% for the pricing of the product. Figure 15 shows the required level of premium according to the model for ages at subscribing from 50 and 80, as well as confidence intervals obtained by bootstrap, using the methodology described in the previous section. The premium increases exponentially with age and is twice as expensive for women than for men. Confidence intervals are relatively tight given the uncertainty on the underlying biometric laws. The method therefore proves quite robust.

We also compute the average reserve of premium in Figure 16. We define it as the product between the probability $A(x_s, x)$ for the individual to remain autonomous between the age of subscription x_s and the current age x and the associated amount of reserve for premium $\mathrm{RFP}(x_s, x)$ at that age. The reserve for premium reaches a maximum between ages 78 and 88, depending on the age of subscription and then decreases when the cost associated with the claims starts to outweigh the amount of premium. We also compute the average reserve for claim in Figure 17. We define it as the product between the survival probability in LTC $I_x(0, t)$ at the age of claim x for the given duration t and the associated amount of reserve for claim $\mathrm{RFC}(x, t)$. This reserve decreases by

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

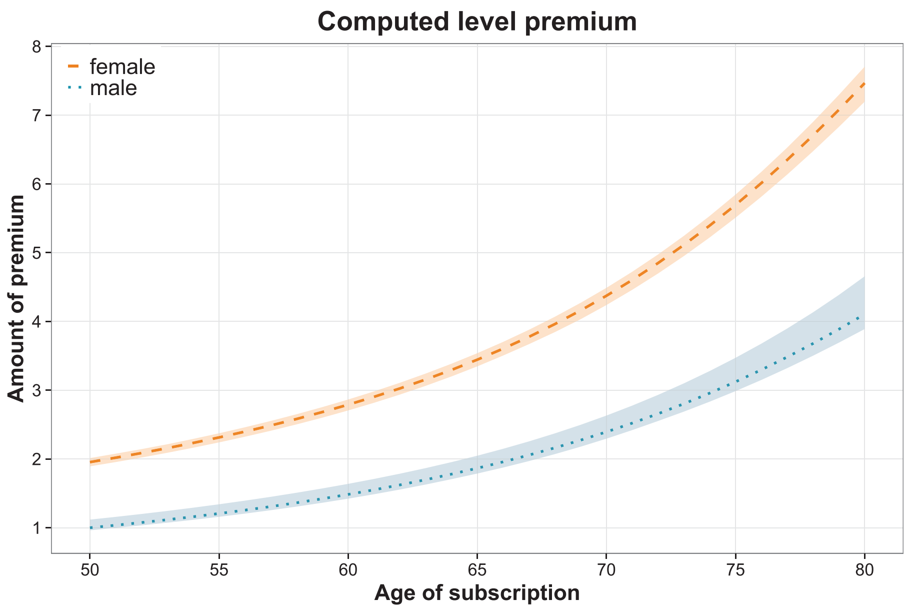

FIGURE 15: Amount of monthly premium required according to the model, with 95% confidence intervals obtained by bootstrap. The y-scale has been re-normalized to preserve confidentiality of results. (Color online)

duration as the number of survivors does. The initial amount of reserve for claim reaches its maximum for claim inception under 60 for women and between 70 and 80 for men. Indeed, for males, the incidence of cancer is very high for ages under 70. Therefore, men under 70 have very high death probabilities for the first year following the onset of LTC, whereas men over 80 have very high death probabilities for the subsequent years, because mortality from other causes of death get higher with ages. For women, this second phenomenon carries more weight, and the most expensive claims are made before age 60.

4. DISCUSSION

In this paper, we introduce a method to estimate biometric laws associated with a LTC insurance portfolio. This method relies on a continuous-time semi-Markov model, as opposed to discrete-time methods used by practitioners in most countries (with the notable exception of Denmark, see Ramlau-Hansen (1991)). This model relies on three transition intensities: incidence in LTC, autonomous mortality and mortality in LTC. We suggest parametric models for the transition intensities. The Brass relational model is used for the intensity of general mortality and the Perks logistic model is used for incidence in LTC, as well as for the first-step estimate of autonomous mortality. As regards mortality in LTC we introduce a mixture model. The aim is to model the underlying heterogeneity in the population caused by the very different pathologies that may lead to

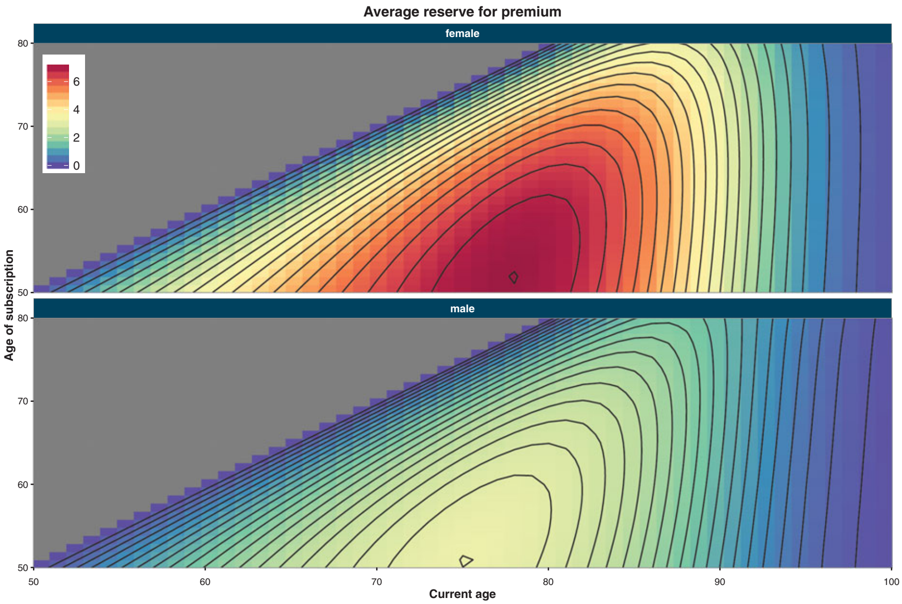

FIGURE 16: Expected value of reserve for premium by age at subscribing and current age, assessed at subscribing. The z-scale has been re-normalized to preserve confidentiality of results. (Color online)

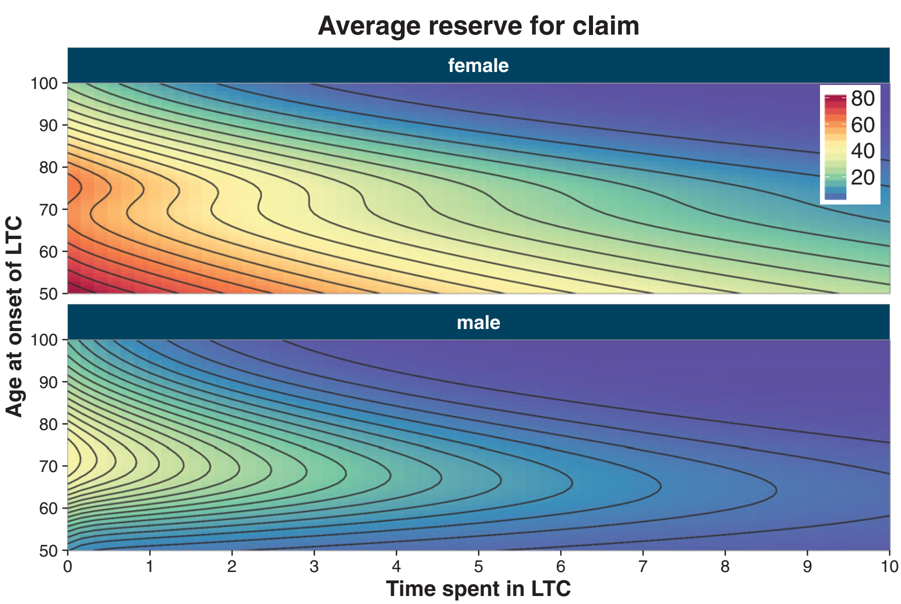

FIGURE 17: Expected value of reserve for claim by age at entry in LTC and time spent in LTC, assessed at claim inception. The z-scale has been re-normalized to preserve confidentiality of results. (Color online)

https://doi.org/10.1017/asb.2016.41 Published online by Cambridge University Press

LTC. Inference of parameters relies on the maximum likelihood method. We then introduce a formula to include general mortality of the portfolio in the model (on which we expect to have more reliable knowledge) and use it to get a second-step estimator of the autonomous mortality, which should prove more reliable at higher ages. We also provide adequate formulas for continuous-time pricing and reserving based directly on the transition intensities. Let us remind that there is only few data available at higher ages on for high duration in LTC. Therefore, parametric methods are compulsory to extrapolate biometric laws at higher ages. Using parametric models from the start is very convenient for the practitioner as it allows to derive biometric laws in a single step, while non-parametric methods requires to find adequate age bands to perform empirical estimations, smooth the empirical probabilities and finally extrapolate the results for higher ages.

We then apply our methodology to data from a real LTC insurance portfolio. Empirical probabilities demonstrates that mortality during the first year following the onset of LTC is way higher than for the subsequent years. A semi-Markov model which takes into account both the age at the onset of LTC and the duration in the LTC state is therefore required in order to explain this phenomenon. By taking into account heterogeneity in the trajectories through a mixture model, we obtain such a model for the mortality in LTC which proves very close to empirical estimations. This may indicate that most of the effect of duration on the mortality actually comes from the heterogeneity of causes.

In this article, we take into account several potential sources of error. As we use a parametric approach, there is a significant risk of modelling error that we try to mitigate by comparing the results of the model with the empirical annual probabilities obtained using a classic non-parametric approach. We also consider several sub-models and remain parsimonious in the number of parameters we introduce by using the Bayesian Information Criterion to compare models. Furthermore, the robustness of estimation is also assessed using a non-parametric quantile bootstrap method.

The parametric form we introduce for mortality in LTC is based on the assumption that pathologies can be sorted in two main groups of homogeneous mortality. This assumption may be tested by focusing on the study of the pathologies causing LTC. Data containing information about pathologies is however extremely scarce and kept private by most insurers. Another limit to our estimation approach is that it is stationary and does not consider that biometric laws are changing over time. The estimation of drifts would indeed prove very difficult because of the limited observation period, and lack of consistency in definition of LTC as well as changes in underwriting and claim management policies over time. Also, most products in France allow the insurer to increase the level of premium in order to account for drifts in the underlying risk. Although this may justify not to consider any trend in the model, a sensitivity approach would in any case prove very useful. We could consider several scenarios for the improvement of incidence and mortality rates and look at the impact on the insurer technical result. Nevertheless, to the best of our knowledge, neither the data nor the theoretical framework associated with this issue exist. Finally, the model only considers one level of LTC, when most individual LTC products currently sold provide several levels of benefits according to the severity of the disability state. Extending the model to consider several levels of LTC as in Lepez *et al.* (2013) or Biessy (2015) would therefore prove very useful. Once again, finding adequate data to perform estimation of parameters is very challenging.

Acknowledgments

I would like to acknowledge the help of Catherine Matias, research director at the CNRS[1] who oversaw my research as a PhD candidate. She provided invaluable help through the writing of this article. I am also grateful to two anonymous referees for their insight on an earlier version of the paper that led to many improvements.

Note

1. CNRS: Centre National de la Recherche Scientifique, France's largest public organism for scientific research.

References

Beard, R. (1959) Note on some mathematical mortality models. In *The Lifespan of Animals. Ciba Foundation Colloquium on Ageing* (eds. G.E.W. Wolstenholme and M. O'Connor), pp. 302–311. Boston: Little, Brown and Company.

Beard, R. (1971) Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. In *Biological Aspects of Demography* (ed. W. Brass). London: Taylor and Francis.

Biessy, G. (2015) Long-term care insurance: A multi-state semi-Markov model to describe the dependency process in elderly people. *Bulletin Français d'Actuariat*, **15**(29), 41–73.

Brass, W. (1971) Mortality models and their uses in demography. *Transactions of the Faculty of Actuaries* **33**, 123–142.

Brass, W. (1974) Perspectives in population prediction: Illustrated by the statistics of england and wales. *Journal of the Royal Statistical Society* **137**(4), 532–583.

Christiansen, M.C. (2012) Multistate models in health insurance. *Advances in Statistical Analysis*, **96**(2), 155–186.

Cinlar, E. (1969) Markov renewal theory. *Advances in Applied Probability*, **1**, 123–187.

Fong, J.H., Shao, A.W. and Sherris, M. (2015) Multistate actuarial models of functional disability. *North American Actuarial Journal*, **19**(1), 41–59.

Gompertz, B. (1825) On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, **115**, 513–583.

Guibert, Q. and Planchet, F. (2014) Construction de lois d'expérience en présence d'évènements concurrents: Application à l'estimation des lois d'incidence d'un contrat dépendance. *Bulletin Français d'Actuariat*, **14**(27), 5–28.

Guibert, Q. and Planchet, F. (2015) Non-parametric inference of transition probabilities based on Aalen–Johansen integral estimators for semi-competing risks data: Application to

LTC insurance. Conference Paper, LIFE Section Colloquium of the International Actuarial Association, Oslo.
HABERMAN, S. and PITACCO, E. (1998) *Actuarial Models for Disability Insurance*, 1st edition. London, UK: Chapman and Hall/CRC.
HANNERZ, H. (2001) An extension of relational methods in mortality estimation. *Demographic Research*, **4**(10), 337–368.
KLEIN, J.P. (1991) Small sample moments of some estimators of the variance of the Kaplan-Meier and Nelson-Aalen estimators. *Scandinavian Journal of Statistics*, 333–340.
LEPEZ, V., ROGANOVA, S. and FLAHAULT, A. (2013) A semi-Markov model to investigate the different transitions between states of dependency in elderly people. In *Colloquium of the International Actuarial Association*, Lyon.
MAKEHAM, W.M. (1867) On the law of mortality. *Journal of the Institute of Actuaries*, **13**, 325–358.
PERKS, W. (1932) On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries*, **63**(1), 12–57.
PITACCO, E. (2014) *Health Insurance – Basic Actuarial Models*. Cham, Switzerland: Springer International Publishing.
PITACCO, E. (2015) *Actuarial values for long-term care insurance products. A sensitivity analysis*. ARC Centre of Excellence in Population Ageing Research, Working Paper.
PLANCHET, F. and THÉROND, P. (2006) *Modèles de durée. Applications actuarielles*. Collection (Assurance, Audit, Actuariat), Economica, Paris, France.
PRITCHARD, D.J. (2006) Modeling disability in long-term care insurance. *North American Actuarial Journal*, **10**(4), 48–75.
RAMLAU-HANSEN, H. (1991) Distribution of surplus in life insurance. *ASTIN Bulletin*, **21**, 57–71.
RICKAYZEN, B.D. and WALSH, D.E.P. (2002) A multi-state model of disability for the United Kingdom: Implications for future need for long-term care for the elderly. *British Actuarial Journal*, **8**, 341–393.
ROBINSON, J. (1996) A long-term care status transition model. In *The Old-Age Crisis – Actuarial Opportunities: The 1996 Bowles Symposium*. New York, USA: Georgia State University.
SCOR (1995) L'assurance dépendance. Dossiers SCOR, SCOR Tech.
University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany) (2015) Human mortality database. Available at www.mortality.org; www.humanmortality.de.
VOLINSKY, C.T. and RAFTERY, A.E. (2000) Bayesian information criterion for censored survival models. *Biometrics*, **56**(1), 256–262.

GUILLAUME BIESSY
SCOR Global Life SE
5 avenue Kléber
75795 Paris Cedex 16, France
Laboratoire de Mathématiques et Modélisation d'Évry, Université d'Évry Val d'Essonne
UMR CNRS 8071
USC INRA
Évry, France
E-Mail: gbiessy@scor.com

APPENDIX

TABLE A1

VALUE OF LOG-LIKELIHOOD l AND BIC OF MODELS FOR MORTALITY IN LTC.

Model	Δ_1, Δ_2	θ	k	l(males)	BIC(males)	l(females)	BIC(females)
1	Constant	Constant	3	−8,423.49	16,872.43	−18,569.90	37,166.71
2	Constant	Logistic(0, 1)	4	−8,394.21	16,822.36	−18,522.60	37,081.07
3	Constant	Logistic(0, β)	5	−8,394.21	16,830.84	−18,522.60	37,090.04
4	Constant	Logistic(α, 1)	5	−8,384.81	16,812.05	−18,505.70	37,056.24
5	Constant	Logistic(α, β)	6	−8,382.20	16,815.30	−18,494.63	37,043.07
6	Gompertz	Constant	5	−8,393.70	16,829.81	−18,543.44	37,131.71
7	Gompertz	Logistic(0, 1)	6	−8,356.27	16,763.44	−18,475.74	37,005.29
8	Gompertz	Logistic(0, β)	7	−8,356.27	16,771.93	−18,475.74	37,014.26
9	Gompertz	Logistic(α, 1)	7	−8,350.55	**16,760.50**	−18,466.26	36,995.30
10	Gompertz	Logistic(α, β)	8	−8,348.25	16,764.38	−18,456.02	**36,983.79**
11	Makeham	Constant	7	−8,393.26	16,845.91	−18,542.66	37,148.10
12	Makeham	Logistic(0, 1)	8	−8,355.39	16,778.65	−18,473.17	37,018.08
13	Makeham	Logistic(0, β)	9	−8,355.39	16,787.14	−18,473.17	37,027.04
14	Makeham	Logistic(α, 1)	9	−8,350.91	16,778.18	−18,465.94	37,012.60
15	Makeham	Logistic(α, β)	10	−8,348.66	16,782.18	−18,456.04	37,001.75
16	Beard	Constant	7	−8,393.62	16,846.64	−18,543.43	37,149.63
17	Beard	Logistic(0, 1)	8	−8,356.21	16,780.30	−18,475.74	37,023.23
18	Beard	Logistic(0, β)	9	−8,356.22	16,788.81	−18,475.74	37,032.20
19	Beard	Logistic(α, 1)	9	−8,349.54	16,775.44	−18,463.98	37,008.67
20	Beard	Logistic(α, β)	10	−8,346.92	16,778.69	−18,455.87	37,001.41
21	Perks	Constant	9	−8,391.68	16,859.73	−18,542.14	37,165.00
22	Perks	Logistic(0, 1)	10	−8,355.79	16,796.42	−18,473.26	37,036.20
23	Perks	Logistic(0, β)	11	−8,355.95	16,805.23	−18,473.30	37,045.24
24	Perks	Logistic(α, 1)	11	−8,350.91	16,795.15	−18,466.62	37,031.88
25	Perks	Logistic(α, β)	12	−8,346.90	16,795.62	−18,455.96	37,019.53