Extension of as-if-Markov modeling to scaled payments

Christiansen, Marcus C.; Furrer, Christian

Published in:
Insurance: Mathematics and Economics

DOI:
10.1016/j.insmatheco.2022.09.001

Publication date:
2022

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Extension of as-if-Markov modeling to scaled payments

Marcus C. Christiansen a, Christian Furrer b, *

a Institut für Mathematik, Carl von Ossietzky Universität Oldenburg, Germany
b Department of Mathematical Sciences, University of Copenhagen, Denmark

ABSTRACT

In multi-state life insurance, as-if-Markov modeling has recently been suggested as an alternative to Markov modeling in case of deterministic sojourn and transition payments. Incidental policyholder behavior, on the other hand, gives rise to duration-dependent payments in the form of so-called scaled payments. The goal of this paper is to establish as-if-Markov modeling also for scaled payments. To this end, we employ change of measure techniques to transfer the added complexity from the payments to an auxiliary probabilistic model. Based hereon, we show how to compute the accumulated cash flow by solving a system of equations comparable to Kolmogorov’s forward equations for Markov chains, but with the transition rates replaced by certain forward transition rates related to the auxiliary probabilistic model. Finally, we provide feasible landmark estimators for these auxiliary forward transition rates subject to entirely random right-censoring.

© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

According to Gatzert (2009), implicit options and the rights to modify contractual conditions are an integral part of product design in life insurance, why proper market consistent valuation and risk management should take implicit options into account. Examples include the free policy (or paid-up) option, which when exercised ceases premiums payments at the expense of down-scaled benefits. In this paper, we extend the as-if-Markov approach of Christiansen (2021) for non-Markov models to allow for so-called scaled payments, which arise naturally in the context of policyholder behavior related to free policy conversion and stochastic retirement.

In the last decade, there has been a significant interest in valuation of life insurance payments in the presence of free policy behavior and stochastic retirement, see Henriksen et al. (2014); Buchardt and Møller (2015); Buchardt et al. (2015); Gad and Nielsen (2016). A common feature of these investigations is that the jump process governing the state of the insured is assumed to be Markovian or semi-Markovian. Actually, Markov chain modeling and semi-Markov modeling constitute the predominant approaches in multi-state life insurance, especially due to the lack of feasible alternatives in regards to both estimation and computation, cf. Chapter 1 in Furrer (2020); recent applications of Markov chain and semi-Markov models, besides some of the aforementioned references, include Guilbert and Planchet (2018); Furrer (2019); de Mol van Otterloo and Alonso-García (2022). But as also explained in Section 1 of Christiansen (2021), forcing Markov assumptions in a situation where the Markov assumption is violated in general gives rise to additional systematic model risk. As an alternative, so-called as-if-Markov modeling has recently been proposed in Christiansen (2021). Under as-if-Markov modeling, life insurance liabilities are evaluated conditional on the present state of the insured – but without actually imposing the Markov assumption. In Christiansen (2021), it is shown that valuation may then essentially proceed using classic formulas for Markov chain models, but with the transition rates replaced by certain so-called forward transition rates. Furthermore, these forward transition rates may be consistently estimated via landmarking, namely by the so-called landmark Nelson-Aalen estimators introduced in Putter and Spitoni (2018).

* Corresponding author.
E-mail address: furrer@math.ku.dk (C. Furrer).

https://doi.org/10.1016/j.insmatheco.2022.09.001
0167-6687 © 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
For multi-state models, the concept of forward transition rates originally emerged from the desire to calculate market values for life insurance liabilities in doubly stochastic Markov models, cf. Norberg (2010); Buchardt et al. (2019), where various proposals are studied and compared. Predating this work are the so-called forward mortalities for survival models introduced by Miltersen and Persson. In Miltersen and Persson, forward mortalities are defined implicitly such that classic formulas that apply for deterministic mortalities produce correct market values even for stochastic mortalities — if the stochastic mortality is replaced by its forward mortality. The shift of perspective from doubly stochastic Markov models to non-Markov models is due to Christiansen (2021).

Landmark estimation is essentially a stratification method, so compared to Markov modeling, as-if-Markov modeling trades systematic model risk for unsystematic estimation risk. Depending on the trade-off in practice, this may make as-if-Markov modeling the preferred choice.

A key requirement in Christiansen (2021) is that the sojourn and transition payments have to be deterministic, which is not the case for scaled payments, since the scaling factor typically depends on for instance the duration since exercise of the underlying policyholder option. In Furrer (2022), it has recently been shown how change of measure techniques may be used to transfer this added complexity in the payments to the probabilistic model. The main contribution of the present paper is a merger of the methods and results of Christiansen (2021) and Furrer (2022) to establish as-if-Markov modeling also for scaled payments. As a secondary contribution, we also unfold and clarify the concise consistency proofs of Christiansen (2021).

The paper is structured as follows. In Section 2, we present the mathematical setup. Section 3 contains a review of the contributions of Christiansen (2021); Furrer (2022) within the setup of the present paper, while Section 4 contains the main results. In Subsection 4.1, we show how the accumulated cash flow and prospective reserve of interest may be found by solving a system of equations comparable to Kolmogorov’s forward equations for Markov chains, but involving certain auxiliary forward transition rates. In Subsection 4.2, we introduce and establish consistency of landmark estimators for the auxiliary forward transition rates and the corresponding accumulated cash flow. Finally, Section 5 discusses implementation and connotations for practice, while Section 6 concludes.

2. Setup

In this section, we describe the mathematical setup by specifying the core components of a multi-state model in life insurance, namely the probabilistic model for the state of the insured, the payments between the insured and insurer, and the (usage of) available information. Furthermore, we introduce the associated expected accumulated cash flows; they constitute our main objects of interest in regards to valuation.

The state of the insured is modeled by a non-explosive jump process $Z = (Z_t)_{t \geq 0}$ on a finite state space $\mathcal{Z}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space. For simplicity, we suppose that $Z_0 = z_0 \in \mathcal{Z}$. The information generated by $Z$ is denoted $\mathbb{F}^Z = (\mathcal{F}^Z_t)_{t \geq 0}$.

We associate with the jump process $Z$ a multivariate counting process $N$ with components $N_{jk} = (N_{jk}(t))_{t \geq 0}$ given by

$$N_{jk}(t) = \#\{s \in (0, t]: Z_{s-} = j, Z_s = k\}, \quad t \geq 0,$$

for $j, k \in \mathcal{Z}$, $j \neq k$. We assume that

$$\mathbb{E} \left[ N_{jk}(t)^2 \right] < \infty, \quad t \geq 0, j, k \in \mathcal{Z}, j \neq k.$$  

This common assumption in particular implies the previously stated non-explosion of $Z$.

We consider basic payments $B = (B(t))_{t \geq 0}$, describing benefits less premiums of some insurance contract. The basic payments are required to satisfy $B(0) \equiv b_0 \in \mathbb{R}$ and

$$B(dt) = \sum_{j \in \mathcal{Z}} 1_{\{Z_{t-} = j}\} B_j (dt) + \sum_{j, k \in \mathcal{Z}} b_{jk} N_{jk}(dt), \quad t > 0,$$

for right-continuous real functions of finite variation $B_j$, $j \in \mathcal{Z}$, and measurable real functions of finite variation $b_{jk}$, $j, k \in \mathcal{Z}$, $j \neq k$. Following Definition 5.1 in Christiansen (2021), the basic payments $B$ thus admit a deterministic cash flow representation.

The inclusion of policyholder behavior such as free policy behavior and stochastic retirement leads to scaled payments which depend on the duration since free policy conversion or retirement and thus do not admit a deterministic cash flow representation, see for instance Furrer (2022). The purpose of this paper is to adapt the main program of Christiansen (2021) for basic payments to scaled payments. To this end, we suppose that $\mathcal{Z} = \mathcal{Z}_0 \cup \mathcal{Z}_1$, that $z_0 \in \mathcal{Z}_0$, and that

$$\mathbb{P}(Z_t \in \mathcal{Z}_0, Z_{t-} \in \mathcal{Z}_1) = 0, \quad 0 \leq t \leq s < \infty,$$

so that the states $\mathcal{Z}_1$ are absorbing for $Z$. Furthermore, we introduce the stopping time $\tau$ as the first hitting time for $Z$ of $\mathcal{Z}_1$, that is

$$\tau = \inf\{s \geq 0: Z_s \in \mathcal{Z}_1\},$$

so that $\tau$ may be interpreted as the exercise time of some policyholder option, and so that $Z_0$ and $\mathcal{Z}_1$ would constitute the states pre- and post-exercise, respectively. The situation without policyholder behavior is recovered by setting $\mathcal{Z}_1 = \emptyset$ in which case $\tau = \infty$. The extension to multiple options is discussed in Section 6. Solely for simplicity, we assume that

$$\mathbb{P}(\tau = t) B_j (dt) = 0$$

for all $j \in \mathcal{Z}_0$; in other words, lump sum payments are assumed not to occur simultaneously with the exercise of the policyholder option. This assumption may be relaxed at the cost of an increase in notational burden. The scaled payments of interest $B^\circ = (B^\circ(t))_{t \geq 0}$ are now given by $B^\circ(0) = B(0)$ and
\[ B^\rho(\, dt) = H(t)B(\, dt), \quad t > 0, \]
\[ H(t) = \rho(t, Z_{t-}, Z_t \mathbb{1}_{t \leq 1}, \quad t \geq 0, \]

for some strictly positive \( \mathbb{F}^2 \)-predictable processes \( \rho(\cdot, j, k), \, j \in \mathbb{Z}_0, \, k \in \mathbb{Z}_1 \), assumed bounded by one. Typically, the scaling factor \( \rho \) is determined as to maintain actuarial equivalence on some technical basis, see e.g. Buchardt and Möller (2015); Buchardt et al. (2015); Christiansen and Djehiche (2020); in practice, the scaling factor \( \rho \) is prevalently deterministic. The extension to scaling factors that are not bounded by one is discussed in Section 6.

As previously mentioned, we aim to adapt the as-if-Markov approach of Christiansen (2021) to the present setting. Consequently, we do not impose any conditions on the intertemporal dependence structure of \( Z \). Instead, we restrict the information used for valuation, which we denote by \( \mathcal{G} \), to be of the form
\[
\mathcal{G}_t = \sigma(Z_t), \quad t \geq 0.
\]

There are multiple reasons for considering reduced, typically non-monotone, information, including the desire to reduce the numerical complexity of computations, an absence of or an insufficient quality of data, and regulatory requirements, cf. the discussions in Section 1 and Section 3 of Christiansen and Furrer (2021) and Section 3 of Christiansen (2021).

If \( Z \) is actually Markovian and the payments admit a deterministic cash flow representation, then the restriction to information \( \mathcal{G} \) in place of \( \mathbb{F}^2 \) is inconsequential. But in many applications, there is a lack of evidence supporting such an assumption on the intertemporal dependence structure of \( Z \). In other words, and following the discussion in Section 9 of Christiansen (2021), the usage of Markov chain modeling is subject to added systematic risk, model risk, which does not vanish for large sample sizes. By restricting the information used for valuation, rather than imposing assumptions on the intertemporal dependence structure of \( Z \), we avoid this model risk at the cost of additional unsystematic risk, estimation risk.

An alternative to the ‘Markovian information’ described by \( \mathcal{G} \) could for instance be \( \mathcal{H} = (\mathcal{H}_t)_{t \geq 0} \) given by
\[
\mathcal{H}_t = \sigma(Z_t) \lor H(t), \quad t \geq 0,
\]
which additionally keeps track of the scaling factor. This alternative is discussed in Section 6. Note that \( \mathcal{G} \) and \( \mathcal{H} \) agree pre-exercise, that is they agree up until the stopping time \( \tau \).

In the context of valuation of future liabilities in life insurance, so-called expected accumulated cash flows are essential. In accordance with Definition 2.2 of Buchardt et al. (2019), given information \( \mathcal{G} \), the expected accumulated cash flows \( A^\rho \) associated with the payments \( B^\rho \) are defined by
\[
A^\rho(t, s) = \mathbb{E}[B^\rho(s) - B^\rho(t) \mid \mathcal{G}_t] = \mathbb{E}[B^\rho(s) - B^\rho(t) \mid Z_t], \quad 0 \leq t \leq s < \infty.
\]

In case of a suitably regular deterministic savings account \( \kappa \) and a finite maximal contract time \( 0 < T < \infty \), the prospective reserve \( V^\rho = (V^\rho(t))_{0 \leq t < T} \) may then be calculated according to
\[
V^\rho(t) = \int_{(t, T]} \frac{\kappa(s)}{\kappa(t)} A^\rho(t, ds), \quad 0 \leq t < T. \tag{4}
\]

Especially in the case of ‘Markovian information’, it is natural to work with state-wise quantities. The state-wise expected accumulated cash flows \( A^\rho_i, \, i \in \mathcal{Z} \), are given by
\[
A^\rho_i(t, s) = \mathbb{E}[B^\rho(s) - B^\rho(t) \mid Z_t = i], \quad 0 \leq t \leq s < \infty,
\]
where the conditional expectation is to be read as the factorized conditional expectation \( \mathbb{E}[B^\rho(s) - B^\rho(t) \mid Z_t = \cdot] \) evaluated in \( i \), cf. Section 2 in Christiansen and Furrer (2021). It holds that
\[
A^\rho_i(t, s) = \sum_{i \in \mathcal{Z}} \mathbb{1}_{[Z_t = i]} A^\rho_i(t, s), \quad 0 \leq t \leq s < \infty. \tag{5}
\]

In the following, we focus on valuation at some fixed point in time \( t_0 \geq 0 \). To be specific, we develop probabilistic and statistical tools to estimate and compute the state-wise expected accumulated cash flows valued at this single point in time, namely \( A^\rho_i(t_0, \cdot) \) for \( i \in \mathcal{Z} \).

3. Background

In this section, we review and provide alternative perspectives on the recent contributions of Christiansen (2021); Furrer (2022) and related literature in the context of the setup of this paper. Merging the methods and results of Christiansen (2021); Furrer (2022) constitutes the starting point of our primary investigation, which is postponed to Section 4.

3.1. As-if-Markov approach

In this subsection we disregard policyholder behavior by letting \( Z_1 = 0 \) leading to \( \tau = \infty \). In this case \( B^\rho = B \), which entails that the payments \( B^\rho \) in particular admit a deterministic cash flow representation and which encourages us to suppress \( \rho \) in the notation. Let
\[
p_{i, j}(t_0, s) = \mathbb{E}[\mathbb{1}_{[Z_t = j]} \mid Z_{t_0} = i], \quad i, j \in \mathcal{Z}, \, t_0 \leq s < \infty,
\]
denote the transition probabilities, and introduce also the quantities
\[ p_{i,j,k}(t_0, s) = \mathbb{E}[N_{jk}(s) \mid Z_{t_0} = i], \quad i, j, k \in \mathcal{Z}, \ j \neq k, \ t_0 \leq s < \infty. \]

Following Christiansen (2021), we define so-called (cumulative) forward transition rates \( \Lambda_{i,j,k}(t_0, \cdot), i, j, k \in \mathcal{Z}, \ j \neq k, \) via \( \Lambda_{i,j,k}(t_0, t_0) = 0 \)
and
\[ \Lambda_{i,j,k}(t_0, ds) = \frac{1_{\{p_{i,j}(t_0, s^-) > 0\}}}{p_{i,j}(t_0, s^-)} p_{i,j,k}(t_0, ds), \quad t_0 < s < \infty. \]

We assume that
\[ \Lambda_{i,j,k}(t_0, s) < \infty, \quad j, k \in \mathcal{Z}, \ j \neq k, \ t_0 < s < \infty. \]

In the spirit of Example 6.7 and Example 6.8 in Christiansen (2021), one may show that for fixed \( i \in \mathcal{Z}, \)
\[ p_{i,j}(t_0, ds) = \sum_{k \in \mathcal{Z}} p_{i,k}(t_0, s^-) \Lambda_{i,k,j}(t_0, ds) - p_{i,j}(t, s^-) \sum_{k \in \mathcal{Z}} \Lambda_{i,j,k}(t_0, ds), \quad j \in \mathcal{Z}, \ t_0 < s < \infty, \tag{6} \]
and also that
\[ A_i(t_0, ds) = \sum_{j \in \mathcal{Z}} p_{i,j}(t_0, s^-) \left( B_j(ds) + \sum_{k \in \mathcal{Z}, k \neq j} b_{jk}(s) \Lambda_{i,j,k}(t_0, ds) \right), \quad t_0 < s < \infty. \tag{7} \]

Thus to compute the state-wise expected accumulated cash flows it essentially suffices to solve (6), which is a system of equations comparable to Kolmogorov’s forward equations for Markov models. If \( Z \) is a Markov process, then the cumulative rates \( \Lambda_{i,j,k}(t_0, \cdot), j, k \in \mathcal{Z}, \ j \neq k, \) do not depend on time \( t_0 \) or state \( i. \) If \( Z \) further admits transition rates \( \mu_{jk}, j, k \in \mathcal{Z}, \ j \neq k, \) then
\[ \Lambda_{i,j,k}(t_0, s) = \int_0^s \mu_{jk}(u) \, du, \quad j, k \in \mathcal{Z}, \ j \neq k, \ 0 \leq s < \infty, \]
and if the transition rates are also continuous, then (6) reduces to the usual Kolmogorov’s forward differential equations by
\[ \frac{d}{ds} p_{i,j}(t_0, s) = \sum_{k \in \mathcal{Z}} p_{i,k}(t_0, s^-) \mu_{kj}(s) - p_{i,j}(t, s^-) \sum_{k \in \mathcal{Z}, k \neq j} \mu_{jk}(s), \quad j \in \mathcal{Z}, \ t_0 < s < \infty. \]

If \( Z \) is a Markov process that does not admit transition rates on \([t_0, \infty), \) say transitions occur only at some given finite time points \( t_0 < t_1 < \cdots < t_n < \infty, \ n \in \mathbb{N}, \) with conditional probabilities \( q_{jk}(t_\ell) := \mathbb{P}(Z_{t_\ell} = k \mid Z_{t_{\ell-1}} = j), \ \ell = 1, \ldots, n, \ j, k \in \mathcal{Z}, \ j \neq k, \) then (6) instead reduces to
\[ p_{i,j}(t_0, s) = p_{i,j}(t_0, t_\ell), \quad j \in \mathcal{Z}, \ t_\ell \leq s < t_{\ell+1}, \ \ell = 0, \ldots, n - 1, \]
as well as
\[ p_{i,j}(t_0, t_{\ell+1}) = \sum_{k \in \mathcal{Z}} p_{i,k}(t_0, t_\ell) q_{kj}(t_{\ell+1}) - p_{i,j}(t_0, t_\ell) \sum_{k \in \mathcal{Z}, k \neq j} q_{jk}(t_{\ell+1}), \quad j \in \mathcal{Z}, \ \ell = 0, \ldots, n - 1. \]

Both the former system of differential equations and this recursion are easy to implement and widely used by practitioners. The system of equations (6) is of a similar form and therefore equally easy to implement.

In Markov models, one may estimate the (cumulative) transition rates and the transition probabilities from empirical data by using the Nelson–Aalen and corresponding Aalen–Johansen estimator; these estimators in particular allow for left-truncation and right-censoring. For non-Markov model, it is proposed in Christiansen (2021) and elsewhere to instead use so-called landmark Nelson–Aalen and corresponding landmark Aalen–Johansen estimators, that retain many of the desirable properties of their regular counterparts, to estimate the (cumulative) forward transition rates and transition probabilities. Following Christiansen (2021), suppose that we observe a portfolio of \( n \in \mathbb{N} \) insureds:
\[ (Z^1_t)_{0 \leq t \leq R_1}, \ldots, (Z^n_t)_{0 \leq t \leq R_n}, \]
where \( R^m, \ m = 1, \ldots, n, \) are non-negative random variables that describe right-censoring, and where \( Z^m, \ m = 1, \ldots, n, \) are replicated of \( Z \). We suppose that the pairs \((Z^1, R^1), \ldots, (Z^n, R^n)\) are mutually independent and identically distributed and that \( R^1 \) is independent of \( Z^1 \), that is to say right-censoring is entirely random. The latter assumption, as well as the inclusion of left-truncation, is discussed in more detail in Section 6. Denote by \( N^m, \ m = 1, \ldots, n, \) the associated multivariate counting processes. For fixed \( i \in \mathcal{Z} \) let
\[ \tilde{N}_{i,j,k}(t_0, s) = \sum_{m=1}^n \mathbb{I}_{[Z^m_0 = i]} N^m_{jk}(s \wedge R^m), \quad t_0 \leq s < \infty, \]
for \( j, k \in \mathcal{Z}, \ j \neq k, \) be the counting processes of the sub-portfolio of insured sojourning in state \( i \) at time \( t_0, \) and let
\[ \tilde{I}_{i,j}(t_0, s) = \sum_{m=1}^{n} \mathbb{I}_{\{Z^m_{t_0} = i\}} \mathbb{I}_{\{s < R^m\}} \mathbb{I}_{\{Z^m = j\}}, \quad t_0 \leq s < \infty, \]

for \( j \in Z \) be the corresponding at-risk processes. The landmark Nelson–Aalen estimator for \( \Lambda_{i,j}(t_0, \cdot) \) is then defined via

\[ \tilde{\Lambda}_{i,j}(t_0, s) = \int_{(t_0,s]} \frac{\mathbb{I}_{\{\tilde{I}_{i,j}(t_0, u-) \geq 0\}}}{\tilde{I}_{i,j}(t_0, u-)} \tilde{N}_{i,j}(t_0, du), \quad t_0 \leq s < \infty, \]

for \( j, k \in Z, j \neq k \), and the corresponding landmark Aalen–Johansen estimator for the transition probabilities is defined as the solution to (6) with the (cumulative) forward transition rates replaced by their landmark Nelson–Aalen estimator, that is

\[ \tilde{\pi}_{i,j}(t_0, ds) = \sum_{k \in Z} \tilde{p}_{i,k}(t_0, s-) \tilde{\Lambda}_{i,k}(t_0, ds) - \tilde{\pi}_{i,j}(t_0, s-) \sum_{k \neq j} \tilde{\Lambda}_{i,k}(t_0, ds), \quad j \in Z, t_0 < s < \infty, \]

and initial values \( \tilde{p}_{i,j}(t_0, 0) = 1_{\{j = i\}} \) for \( j \in Z \), see also Putter and Spitoni (2018); here and in the following we use the convention \( 0/0 := 0 \). In Christiansen (2021), it is shown that these estimators are consistent in the sense of uniform \( L_1 \)-convergence (under certain technical regularity conditions), cf. Theorem 7.3 and Theorem 7.4 of Christiansen (2021). Note that this in particular implies weak consistency of the estimators. Similarly, a consistent estimator of the state-wise expected accumulated cash flow \( A_{i}(t_0, \cdot) \) can be obtained by replacing the (cumulative) forward transition rates and transition probabilities in (7) by their landmark Nelson–Aalen–Johansen estimator. The proofs in Christiansen (2021) are based on Overgaard (2019b), which in particular serves to address some shortcomings in the technical arguments of Datta and Satten (2001); Putter and Spitoni (2018), cf. Section 1 in Nießl et al. (2021). In this paper, we – among other things – further unfold and clarify the proofs of Christiansen (2021).

Landmarking is essentially a stratification method, so one disadvantage of landmarking is data reduction and thus loss of power. If the Markov assumption is actually satisfied, it is obviously preferable to use the regular Nelson–Aalen and Aalen–Johansen estimators, which take into account the whole portfolio and not only the sub-portfolio of insured sojourning in state \( i \) at time \( t_0 \). In case the model is only partially non-Markovian, it has recently been suggested to employ a hybrid concept striking a compromise between regular estimation and landmark estimation, see Maltzhan et al. (2021).

### 3.2. Change of measure techniques

Since the scaled payments \( B^\rho \) include the term \( H \), they do not admit a deterministic cash flow representation. In Furrer (2022), change of measure techniques are in a canonical framework used to transfer the term \( H \) from the payments to the probabilistic model. In the following, we adapt the result of Furrer (2022) to the present setup and cast the relevant change of measure explicitly in terms of a transformation of the jump process.

Let \( U \) be a real-valued random variable, and suppose that \( U \) is independent of \( Z \) and uniformly distributed on the open unit interval \((0, 1]\). Define \( \tilde{Z} = (\tilde{Z}_t)_{t\geq 0} \) according to

\[ \tilde{Z}_t = \mathbb{1}_{\{Z_t \in Z_0\}} Z_t + \mathbb{1}_{\{Z_t \in Z_1\}} \left( \mathbb{1}_{\{U < \rho(t, Z_t, Z_t)\}} Z_t + \mathbb{1}_{\{U \geq \rho(t, Z_t, Z_t)\}} \mathbb{1} \right), \quad t \geq 0, \tag{8} \]

where \( \mathbb{1} \) is an artificial cemetery state. Note that \( \tilde{Z} \) is a non-explosive jump process on the finite state space \( \tilde{Z} := Z \cup \{\mathbb{1}\} \) with \( \tilde{Z}_0 \equiv Z_0 \) and with both \( \{\mathbb{1}\} \) and \( Z_1 \) absorbing.

The information generated by \( \tilde{Z} \) is denoted \( \mathbb{F}^\tilde{Z} = (\mathbb{F}_t^\tilde{Z})_{t \geq 0} \). We associate with the jump process \( \tilde{Z} \) a multivariate counting process \( \tilde{N} \) with components \( \tilde{N}_{jk} = (\tilde{N}_{jk}(t))_{t \geq 0} \) given by

\[ \tilde{N}_{jk}(t) = \#\{s \in (0, t] : \tilde{Z}_{s-} = j, \tilde{Z}_s = k\}, \quad t \geq 0, \]

for \( j, k \in \tilde{Z}, j \neq k \). Since (1) holds, we also obtain

\[ \mathbb{E} \left[ \tilde{N}_{jk}(t)^2 \right] < \infty, \quad j, k \in \tilde{Z}, j \neq k, t \geq 0. \tag{9} \]

Finally, consider auxiliary payments \( \tilde{B} = (\tilde{B}(t))_{t \geq 0} \) given by \( \tilde{B}(0) \equiv b_0 \) and

\[ \tilde{B}(dt) = \sum_{j \in Z} \mathbb{1}_{\{\tilde{Z}_{t-} = j\}} B_j(dt) + \sum_{j,k \in Z} b_{jk}(t) \tilde{N}_{jk}(dt), \quad t > 0, \]

which correspond to the basic payments \( B \) but with \( Z \) and \( N \) replaced by \( \tilde{Z} \) and \( \tilde{N} \), respectively. In particular, the payments \( \tilde{B} \) admit a deterministic cash flow representation with respect to \( \tilde{Z} \) (rather than \( Z \)).

In Furrer (2022), it is presumed that the insurer utilizes full information. The quantity of interest is thus the expected accumulated cash flow associated with the payments \( B^\rho \) given the information generated by \( Z \), corresponding to

\[ \mathbb{E} \left[ B^\rho(s) - B^\rho(t_0) \right| \mathbb{F}_t^Z], \quad t_0 \leq s < \infty. \]

Since \( U \) is assumed to be independent of \( Z \), it does not matter whether one does or does not condition on information generated by \( U \). By closely inspecting the arguments leading to Theorem 3.6 in Furrer (2022), it is possible to show that

\[ \mathbb{E} \left[ \tilde{B}(s) - \tilde{B}(t_0) \right| \mathbb{F}_t^\tilde{Z}] \mathbb{1}_{\{\tilde{Z}_{t_0} \in Z_0\}} = \mathbb{E} \left[ B^\rho(s) - B^\rho(t_0) \right| \mathbb{F}_t^{\tilde{Z}}] \mathbb{1}_{\{\tilde{Z}_{t_0} \in Z_0\}}, \quad t_0 \leq s < \infty. \tag{10} \]
This identity is sensible due to the local property of the conditional expectation, cf. Lemma 8.3 in Kallenberg (2021). Furthermore, since

$$\mathbb{1}_{\{\tilde{Z}_{t_0} \in Z_1\}} \left( \tilde{B}(s) - \tilde{B}(t_0) \right) = \mathbb{1}_{\{\tilde{Z}_{t_0} \in Z_1\}} \frac{B^\rho(s) - B^\rho(t_0)}{H(t_0)}, \quad t_0 \leq s < \infty,$$

and since

$$\mathcal{F}_{t_0} \cap \{\tilde{Z}_{t_0} \in Z_1\} = \left( \mathcal{F}_{t_0} \lor \sigma(\mathbb{1}_{\{U \leq \rho(t, Z_\tau, Z_\tau)\}}) \right) \cap \{\tilde{Z}_{t_0} \in Z_1\},$$

it follows directly from local property of the conditional expectation, cf. Lemma 8.3 in Kallenberg (2021), that

$$\mathbb{E} \left[ \tilde{B}(s) - \tilde{B}(t_0) \middle| \mathcal{F}_{t_0} \right] \mathbb{1}_{\{\tilde{Z}_{t_0} \in Z_1\}} = \mathbb{E} \left[ \frac{B^\rho(s) - B^\rho(t_0)}{H(t_0)} \middle| \mathcal{F}_{t_0} \right] \mathbb{1}_{\{\tilde{Z}_{t_0} \in Z_1\}}, \quad t_0 \leq s < \infty.$$

Referring to the definition of $H$ and the fact that $U$ is independent of $Z$, one may then show that

$$\mathbb{E} \left[ \tilde{B}(s) - \tilde{B}(t_0) \middle| \mathcal{F}_{t_0} \right] \mathbb{1}_{\{\tilde{Z}_{t_0} \in Z_1\}} = \frac{\mathbb{E} \left[ B^\rho(s) - B^\rho(t_0) \middle| \mathcal{F}_{t_0} \right]}{\rho(t, Z_\tau, Z_\tau)} \mathbb{1}_{\{\tilde{Z}_{t_0} \in Z_1\}}, \quad t_0 \leq s < \infty. \quad (11)$$

In combination, the identities (10) and (11) provide a partial link between the expected accumulated cash flow associated with the payments $\tilde{B}$ given the information generated by $\tilde{Z}$ and the expected accumulated cash flow associated with the payments $B^\rho$ given the information generated by $Z$. In the next section, we utilize this partial link and the fact that the payments $\tilde{B}$ admit a deterministic cash flow representation with respect to $\tilde{Z}$ to extend the as-if-Markov approach outlined in Subsection 3.1 to scaled payments, that is to the case $Z_1 \neq \emptyset$.

4. Main results

This section contains the main contributions of the paper. In Subsection 4.1, we derive a system of equations – comparable to Kolmogorov’s forward equations for Markov models – which allows for efficient computation of the state-wise expected accumulated cash flows. The system relies on certain (cumulative) auxiliary forward transition rates. Subsection 4.2 deals with the estimation of the quantities involved, in particular the (cumulative) auxiliary forward transition rates.

4.1. Valuation

We introduce auxiliary state-wise expected accumulated cash flows $\tilde{A}_i(t_0, \cdot)$, $i \in \mathcal{Z}$, according to

$$\tilde{A}_i(t_0, s) = \mathbb{E}[\tilde{B}(s) - \tilde{B}(t_0) \middle| Z_{t_0} = i], \quad t_0 \leq s < \infty.$$

The following result, which builds on the review given in Subsection 3.2, provides a direct link between the state-wise expected accumulated cash flows of interest and the auxiliary state-wise expected accumulated cash flows.

Theorem 1. For each $i \in \mathcal{Z}$ with $\mathbb{P}(Z_{t_0} = i) > 0$ it holds that

$$A^\rho_i(t_0, s) = \tilde{A}_i(t_0, s), \quad t_0 \leq s < \infty. \quad (12)$$

Proof. Fix $s \in [t_0, \infty)$. Let first $i \in \mathcal{Z}_0$, and suppose that $\mathbb{P}(Z_{t_0} = i) > 0$. Then $\{Z_{t_0} = i\} = \{\tilde{Z}_{t_0} = i\}$, so that

$$\tilde{A}_i(t_0, s) = \mathbb{E}[\tilde{B}(s) - \tilde{B}(t_0) \middle| Z_{t_0} = i] = \mathbb{E} \left[ \mathbb{1}_{\{Z_{t_0} = i\}} (\tilde{B}(s) - \tilde{B}(t_0)) \middle| \mathcal{F}_{t_0} \right] = \mathbb{E} \left[ \mathbb{1}_{\{Z_{t_0} = i\}} (\tilde{B}(s) - \tilde{B}(t_0)) \middle| \mathcal{F}_{t_0} \right] \mathbb{P}(Z_{t_0} = i) \mathbb{P}(Z_{t_0} = i).$$

By the law of iterated expectations and (10), we find that

$$\mathbb{E} \left[ \mathbb{1}_{\{Z_{t_0} = i\}} (\tilde{B}(s) - \tilde{B}(t_0)) \middle| \mathcal{F}_{t_0} \right] = \mathbb{E} \left[ \mathbb{1}_{\{Z_{t_0} = i\}} \mathbb{E} \left[ B^\rho(s) - B^\rho(t_0) \middle| \mathcal{F}_{t_0} \right] \mathcal{F}_{t_0} \right]$$

$$= \mathbb{E} \left[ \mathbb{1}_{\{Z_{t_0} = i\}} \mathbb{E} \left[ B^\rho(s) - B^\rho(t_0) \middle| \mathcal{F}_{t_0} \right] \mathcal{F}_{t_0} \right]$$

$$= \mathbb{E} \left[ \mathbb{1}_{\{Z_{t_0} = i\}} \mathbb{E} \left[ B^\rho(s) - B^\rho(t_0) \middle| \mathcal{F}_{t_0} \right] \mathcal{F}_{t_0} \right]$$

$$= \mathbb{E} \left[ \mathbb{1}_{\{Z_{t_0} = i\}} (B^\rho(s) - B^\rho(t_0)) \right].$$

In particular,

$$\tilde{A}_i(t_0, s) = \frac{\mathbb{E} \left[ \mathbb{1}_{\{Z_{t_0} = i\}} (B^\rho(s) - B^\rho(t_0)) \right]}{\mathbb{P}(Z_{t_0} = i)} = \mathbb{E}[B^\rho(s) - B^\rho(t_0) \middle| Z_{t_0} = i] = A^\rho_i(t_0, s)$$

as desired. Let now $i \in \mathcal{Z}_1$, and suppose that $\mathbb{P}(Z_{t_0} = i) > 0$. Note that

$$\left( \tilde{B}(s) - \tilde{B}(t_0) \right) \mathbb{1}_{\{Z_{t_0} = i\}} = (\tilde{B}(s) - \tilde{B}(t_0)) \mathbb{1}_{\{\tilde{Z}_{t_0} = i\}},$$
which implies that

\[ \tilde{A}_i(t_0, s) = \mathbb{E}[\tilde{B}(s) - \tilde{B}(t_0) \mid Z_{t_0} = i] = \mathbb{E}\left[ \mathbb{1}_{[Z_{t_0} = i]}(\tilde{B}(s) - \tilde{B}(t_0)) \right] \frac{\mathbb{P}(Z_{t_0} = i)}{\mathbb{P}(Z_{t_0} = i)}. \]

By similar arguments as in the first part of the proof, but referring to (11) instead of (10), it follows that

\[ \mathbb{E}\left[ \mathbb{1}_{[Z_{t_0} = i]}(\tilde{B}(s) - \tilde{B}(t_0)) \right] = \mathbb{E}\left[ \mathbb{1}_{[Z_{t_0} = i]} \mathbb{E}\left[ B^\rho(s) - B^\rho(t_0) \mid \mathcal{F}^Z_{t_0} \right] \frac{\rho(\tau, Z_{t_0}, Z_t)}{\rho(\tau, Z_{t_0}, Z_t)} \right]. \]

Recall that \( U \) is independent of \( Z \) and uniformly distributed on the open unit interval, so that

\[ \mathbb{E}[\mathbb{1}_{[U \leq \rho(\tau, Z_{t_0}, Z_t)]} \mid \mathcal{F}^Z_{t_0}] = \rho(\tau, Z_{t_0}, Z_t). \]

Repeated usage of the law of iterated expectations then yields

\[ \mathbb{E}\left[ \mathbb{1}_{[Z_{t_0} = i]}(\tilde{B}(s) - \tilde{B}(t_0)) \right] = \mathbb{E}\left[ \mathbb{1}_{[Z_{t_0} = i]} \mathbb{E}\left[ B^\rho(s) - B^\rho(t_0) \mid \mathcal{F}^Z_{t_0} \right] \right] = \mathbb{E}\left[ \mathbb{1}_{[Z_{t_0} = i]}(B^\rho(s) - B^\rho(t_0)) \right]. \]

In conclusion,

\[ \tilde{A}_i(t_0, s) = \mathbb{E}\left[ \mathbb{1}_{[Z_{t_0} = i]}(B^\rho(s) - B^\rho(t_0)) \right] = \mathbb{E}[B^\rho(s) - B^\rho(t_0) \mid Z_{t_0} = i] = A_i^\rho(t_0, s), \]

which serves to complete the proof. \( \square \)

Theorem 1 shows that one may compute the state-wise expected accumulated cash flows of interest by instead computing certain auxiliary state-wise expected accumulated cash flows. The importance of this result lies in the fact that the auxiliary payments admit a deterministic cash flow representation, so that the methods and results of Christiansen (2021) reviewed in Subsection 3.1 may be applied more or less directly. To this end, for \( i \in \mathcal{Z} \) let

\[ \tilde{p}_{i,j}(t_0, s) = \mathbb{E}[\mathbb{1}_{[Z_{t_0} = j]} \mid Z_{t_0} = i], \quad j \in \mathcal{Z}, \ t_0 \leq s < \infty, \]

be a collection of auxiliary probabilities, and let

\[ \tilde{p}_{i,j}(t_0, s) = \mathbb{E}[N_{j,k}(s) \mid Z_{t_0} = i], \quad j, k \in \mathcal{Z}, \ j \neq k, \ t_0 \leq s < \infty. \]

Based hereon, define auxiliary (cumulative) forward transition rates \( \tilde{A}_{i,j}(t_0, \cdot) \), \( i \in \mathcal{Z} \), \( j, k \in \mathcal{Z}, \ j \neq k \), via \( \tilde{A}_{i,j}(t_0, 0) = 0 \) and

\[ \tilde{A}_{i,j}(t_0, ds) = \mathbb{1}_{[\tilde{p}_{i,j}(t_0, s) > 0]} \frac{\tilde{p}_{i,j}(t_0, s)}{\tilde{p}_{i,j}(t_0, s)} \tilde{p}_{i,j}(t_0, ds), \quad t_0 \leq s < \infty \]

Again, we assume that

\[ \tilde{A}_{i,j}(t_0, s) < \infty, \quad j, k \in \mathcal{Z}, \ j \neq k, \ t_0 \leq s < \infty. \quad (13) \]

As before, one may for fixed \( i \in \mathcal{Z} \) show that

\[ \tilde{p}_{i,j}(t_0, ds) = \sum_{k \in \mathcal{Z}} \tilde{p}_{i,k}(t_0, s) \tilde{A}_{i,kj}(t_0, ds) - \tilde{p}_{i,j}(t, s) \sum_{k \in \mathcal{Z}} \tilde{A}_{i,kj}(t_0, ds), \quad j \in \mathcal{Z}, \ t_0 < s < \infty, \quad (14) \]

and also that

\[ \tilde{A}_i(t_0, ds) = \sum_{j \in \mathcal{Z}} \tilde{p}_{i,j}(t_0, s) \left( B_j(ds) + \sum_{k \in \mathcal{Z}} b_{jk}(s) \tilde{A}_{i,kj}(t_0, ds) \right), \quad t_0 < s < \infty. \quad (15) \]

Note that (14) and (15) hold even if the assumption prescribed by (13) is dropped, as long as the involved functions are redefined as measures on the real line.

From (12) and (15) it follows that to compute the state-wise expected accumulated cash flows of interest, it essentially suffices to solve (14), which is a system of equations comparable to Kolmogorov’s forward equations for Markov models. This requires the estimation of the auxiliary (cumulative) forward transition rates and the associated auxiliary probabilities. In the next subsection, we show how this may be done using landmark Nelson–Aalen and corresponding landmark Aalen–Johansen estimators.
4.2. Estimation

In this subsection, we show how to estimate the auxiliary (cumulative) forward transition rates \( \tilde{\lambda}_{i,j}(t_0,-) \), \( i, j \in \mathcal{Z} \), \( j \neq k \), and associated auxiliary probabilities and state-wise expected accumulated cash flows from empirical data. Similar to Christiansen (2021), the results presented here serve only as a proof of concept and far from cover the full range of possible model settings.

Similar to Subsection 3.1, suppose that we observe a portfolio of \( n \in \mathbb{N} \) insured:

\[
(Z^1_t)_{0 \leq t \leq \mathcal{R}^1}, \ldots, (Z^n_t)_{0 \leq t \leq \mathcal{R}^n},
\]

(16)

where \( \mathcal{R}^m \), \( m = 1, \ldots, n \), are non-negative random variables that describe right-censoring. Let \( U^1, \ldots, U^n \) be mutually independent and identically distributed real-valued random variables. Further, suppose that \( (U^m)_{m=1}^n \) is independent of \( (\mathcal{R}^m)_{m=1}^n \), and suppose that \( U^1 \) is uniformly distributed on the open unit interval \((0, 1)\). Define the auxiliary data

\[
(\tilde{Z}^1_t)_{0 \leq t \leq \mathcal{R}^1}, \ldots, (\tilde{Z}^n_t)_{0 \leq t \leq \mathcal{R}^n},
\]

according to

\[
\tilde{Z}^m_t = 1_{\{Z^m_t \leq Z_0 \}} Z_t + 1_{\{Z^m_t > Z_0 \}} \left( \mathbf{1}_{\{U^m \leq \rho(t, Z^m_{\tau_{m-1}}, Z^m_{\tau_{m}})\}} Z^m_t + \mathbf{1}_{\{U^m > \rho(t, Z^m_{\tau_{m-1}}, Z^m_{\tau_{m}})\}} V \right), \quad t \leq \mathcal{R}^m,
\]

where \( m = 1, \ldots, n \). In the following, we denote by \( \tilde{N}^m \) the multivariate counting process associated with \( \tilde{Z}^m \).

For fixed \( i \in \mathcal{Z} \) let

\[
\tilde{N}_{i,j}(t_0, s) = \sum_{m=1}^n I_{\{Z^m_{t_0} = i\}} \tilde{N}^m_{j,k}(s \wedge \mathcal{R}^m), \quad t_0 \leq s < \infty,
\]

for \( j, k \in \mathcal{Z} \), \( j \neq k \), be the auxiliary counting processes of the sub-portfolio of insured sojourning in state \( i \) at time \( t_0 \), and let

\[
\hat{N}_{i,j}(t_0, s) = \sum_{m=1}^n I_{\{Z^m_{t_0} = i\}} I_{\{s < \mathcal{R}^m\}} I_{\{\tilde{Z}^m_{t_0} = j\}}, \quad t_0 \leq s < \infty,
\]

for \( j \in \mathcal{Z} \) be the corresponding at-risk processes. In the following, the landmark \( i \in \mathcal{Z} \) is held fixed.

**Definition 1.** The **landmark Nelson–Aalen estimator** \( \hat{\lambda}_{i,j}(t_0, -) \), \( j, k \in \mathcal{Z} \), \( j \neq k \), for \( \tilde{\lambda}_{i,j}(t_0, -) \), \( j, k \in \mathcal{Z} \), \( j \neq k \), is defined via

\[
\hat{\lambda}_{i,j}(t_0, s) = \int_{(t_0, s]} I_{\{\hat{N}_{i,j}(t_0, u-) > 0\}} \hat{N}_{i,j}(t_0, du), \quad t_0 \leq s < \infty,
\]

for \( j, k \in \mathcal{Z} \), \( j \neq k \).

**Definition 2.** The **landmark Aalen–Johansen estimator** \( \hat{p}_{i,j}(t_0, -) \), \( j \in \mathcal{Z} \), for \( \tilde{p}_{i,j}(t_0, -) \), \( j \in \mathcal{Z} \), is defined as the solution to (14) with the auxiliary (cumulative) forward transition rates replaced by their landmark Nelson–Aalen estimator, that is

\[
\hat{p}_{i,j}(t_0, ds) = \sum_{k \in \mathcal{Z}} I_{k \neq j} \hat{\lambda}_{i,k}(t_0, ds) - \hat{\lambda}_{i,j}(t_0, s-) \sum_{k \in \mathcal{Z}} I_{k \neq j} \hat{\lambda}_{i,k}(t_0, ds), \quad j \in \mathcal{Z}, \ t_0 < s < \infty,
\]

and initial values

\[
\hat{p}_{i,j}(t_0, t_0) = \frac{\hat{N}_{i,j}(t_0, t_0)}{\sum_{j \in \mathcal{Z}} \hat{N}_{i,j}(t_0, t_0)}, \quad j \in \mathcal{Z}.
\]

**Remark 1.** Note that if \( i \in \mathcal{Z} \), then

\[
\hat{p}_{i,j}(t_0, t_0) = I_{j = i}
\]

for \( j \in \mathcal{Z} \), while if \( i \in \mathcal{Z} \), then

\[
\hat{p}_{i,j}(t_0, t_0) = I_{j = i} \frac{\sum_{m=1}^n I_{\{Z^m_{t_0} = i\}} I_{\{t_0 \leq \mathcal{R}^m\}} I_{\{U^m \leq \rho(\tau, Z_{\tau_{m-1}}, Z_{\tau_{m}})\}}}{\sum_{m=1}^n I_{\{Z^m_{t_0} = i\}} I_{\{t_0 \leq \mathcal{R}^m\}}}
\]

for \( j \in \mathcal{Z} \), which is really quite a natural estimator for

\[
\hat{p}_{i,j}(t_0, t_0) = I_{j = i} P(U \leq \rho(\tau, Z_{\tau_{m-1}}, Z_{\tau_{m}}) | Z_{t_0} = i).
\]
The estimator \( \hat{p}_{i,j}(t_0, t_0) \) for \( \tilde{p}_{i,j}(t_0, t_0) \) is consistent in the sense of \( L_1 \)-convergence (as the size of the portfolio increases) and bounded, which are crucial properties for the proof of Theorem 3 below. Alternatively, we may replace \( \hat{p}_{i,j}(t_0, t_0) \) by any other consistent and bounded estimator of \( \tilde{p}_{i,j}(t_0, t_0) \).

In the following, we fix \( i \in Z \) with \( \mathbb{P}(Z_{t_0} = i) > 0 \). For a finite maximal contract time \( T > t_0 \) and \( p \in [1, \infty) \), we define the \( p \)-variation norm \( \| \cdot \|_p := \| \cdot \|_\infty + \| \cdot \|_p \), where \( \| \cdot \|_\infty \) is the supremum norm on \([t_0, T]\) and \( \| \cdot \|_p \) is the total \( p \)-variation on \([t_0, T]\); a compilation of fundamental properties for the \( p \)-variation norm may be found in Appendix A. The next theorem establishes consistency (as the size of the portfolio increases) of the landmark Nelson–Aalen estimators. The theorem is in the same vein as Theorem 7.3 in Christiansen (2021), but it also deals with the transformation from \( Z \) to \( \tilde{Z} \). Recall that \( U^1, \ldots, U^n \) are mutually independent and identically distributed real-valued random variables and that \( (U^m)_{m=1}^\infty \) is independent of \( (Z^m, R^m)_{m=1}^\infty \).

**Theorem 2.** Suppose that

- \( (Z^1, R^1), \ldots, (Z^n, R^n) \) are independent and identically distributed
- \( R^1 \) and \( Z^1 \) are independent
- \( \mathbb{P}(R^1 \geq T) > 0 \).

Let \( p \in (1, 2) \). Then for each \( j, k \in \tilde{Z} \), \( j \neq k \), it holds that

\[
\mathbb{E}\left[ \| \hat{A}_{i,j,k}(t_0, \cdot) - \tilde{A}_{i,j,k}(t_0, \cdot) \|_p \right] \rightarrow 0, \quad n \rightarrow \infty.
\]

**Proof.** Fix \( j, k \in \tilde{Z} \), \( j \neq k \). For an individual insured with state process \( Z \), auxiliary state process \( \tilde{Z} \), and right-censoring time \( R \), we write

\[
N^c_{t,j,k}(t_0, t) = \mathbb{I}(Z_{t_0} = j) \tilde{N}_{j,k}(t \wedge R), \quad t_0 \leq t < \infty,
\]

for the censored landmark auxiliary counting process and

\[
I^c_{t,j,k}(t_0, t) = \mathbb{I}(Z_{t_0} = j) \mathbb{I}(t < R) \mathbb{I}(\tilde{Z}_{t_0} = j), \quad t_0 \leq t < \infty,
\]

for the censored landmark auxiliary state indicator process. Their expectations are denoted by

\[
q_{i,j}(t_0, t) = \mathbb{E}[N^c_{i,j,k}(t_0, t)], \quad t_0 \leq t < \infty,
\]

\[
q_{i,j}(t_0, t) = \mathbb{E}[I^c_{i,j,k}(t_0, t)], \quad t_0 \leq t < \infty.
\]

Based hereon we define a (cumulative) forward transition rate \( \Gamma_{i,j,k}(t_0, \cdot) \) via \( \Gamma_{i,j,k}(t_0, 0) = 0 \) and

\[
\Gamma_{i,j,k}(t_0, dt) = \frac{q_{j,k}(t_0, t_0)}{q_{i,j}(t_0, t_0)} q_{i,j}(t_0, dt), \quad t_0 < t < \infty.
\]

Note that \( \Gamma_{i,j,k}(t_0, t_0) = \hat{A}_{i,j,k}(t_0, t_0) \) by definition. Let now \( t > t_0 \). Under the assumption that \( U \) is independent of \( (Z, R) \) and that \( R \) and \( Z \) are independent, it follows from the law of iterated expectations that

\[
q_{i,j}(t_0, t) - q_{i,j}(t_0, t_0) = \mathbb{E}\left[ \mathbb{E}\left[ \mathbb{I}(Z_{t_0} = i) \int_{(0,t)} \mathbb{I}(u \leq R) \tilde{N}_{j,k}(du) \bigg| Z, U \right] \bigg| Z_{t_0} = i \right],
\]

\[
= \mathbb{E}\left[ \mathbb{I}(Z_{t_0} = i) \int_{(0,t)} \mathbb{E}[\mathbb{I}(u \leq R) \big| Z, U] \tilde{N}_{j,k}(du) \right],
\]

\[
= \mathbb{E}\left[ \mathbb{I}(Z_{t_0} = i) \int_{(0,t)} \mathbb{E}[\mathbb{I}(u \leq R) \big| \tilde{N}_{j,k}(du) \big| Z_{t_0} = i \bigg| Z_{t_0} = i \right],
\]

\[
= \mathbb{P}(Z_{t_0} = i) \int_{(0,t)} \mathbb{E}[\mathbb{I}(u \leq R) \big| \tilde{N}_{j,k}(du)],
\]

where the last line follows from Campbell’s theorem. In similar fashion,

\[
q_{i,j}(t_0, t) = \mathbb{E}\left[ \mathbb{I}(Z_{t_0} = i) \mathbb{I}(t < R) \big| Z, U \right] \mathbb{I}(\tilde{Z}_{t_0} = j) = \mathbb{E}\left[ \mathbb{I}(Z_{t_0} = i) \mathbb{I}(t < R) \mathbb{I}(\tilde{Z}_{t_0} = j) \right] = \mathbb{P}(Z_{t_0} = i) \mathbb{E}[\mathbb{I}(t < R) \big| \tilde{N}_{i,j}(t_0, t)].
\]

In combination, we find that

\[
\Gamma_{i,j,k}(t_0, t) = \hat{A}_{i,j,k}(t_0, t) < \infty, \quad t_0 \leq t < \infty,
\]

where the finiteness follows from (13). So it suffices to show that
\[
\mathbb{E} \left[ \| \hat{A}_{i,j}(t_0, \cdot) - \Gamma_{i,j}(t_0, \cdot) \|_{\mathcal{P}_1} \right] \to 0, \quad n \to \infty.
\]

This is in particular true if
\[
\mathbb{E} \left[ \mathbb{I}_{ \{ N(t_0,T)/n \leq M \} } \| \hat{A}_{i,j}(t_0, \cdot) - \Gamma_{i,j}(t_0, \cdot) \|_{\mathcal{P}_1} \right] \to 0, \quad n \to \infty,
\]
for any \( M \in \mathbb{N} \) and
\[
\lim_{n \to \infty} \mathbb{E} \left[ \mathbb{I}_{ \{ N(t_0,T)/n > M \} } \| \hat{A}_{i,j}(t_0, \cdot) - \Gamma_{i,j}(t_0, \cdot) \|_{\mathcal{P}_1} \right] \to 0, \quad M \to \infty,
\]
where \( \hat{N}_i(t_0, \cdot) \) is given by
\[
\hat{N}_i(t_0, s) = \sum_{j', k' \in \mathcal{Z}} \hat{N}_{i,j'k'}(t_0, s), \quad t_0 \leq s < \infty.
\]

We first prove (19). According to Markov’s inequality, it holds that
\[
\mathbb{E} \left[ \mathbb{I}_{ \{ \hat{N}_i(t_0,T)/n \leq M \} } \right] \leq \frac{\mathbb{E} \left[ \hat{N}_i(t_0,T)/n \right]}{M} \leq \frac{\mathbb{E} \left[ \hat{N}_{j'k'}(T) \right]}{M}.
\]

For \( m = 1, \ldots, n \) we let
\[
\hat{N}_{i,j}^m(t_0, s) := \mathbb{I}_{ \{ Z_{t_0}^m = i \} } \hat{N}_{i,j}^m(s \wedge \mathcal{R}^m), \quad t_0 \leq s < \infty,
\]
as well as
\[
\hat{I}_{i,j}^m(t_0, s) := \mathbb{I}_{ \{ Z_{t_0}^m = i \} } \mathbb{I}_{ \{ s \leq \mathcal{R}^m \} } \mathbb{I}_{ \{ Z_{t_0}^m = j \} }, \quad t_0 \leq s < \infty,
\]
and
\[
\hat{N}_i^m(t_0, s) := \sum_{j' \in \mathcal{Z}} \hat{N}_{i,j'k}(t_0, s), \quad t_0 \leq s < \infty,
\]
so that \( \hat{N}_{i,j}(t_0, \cdot) = \sum_{m=1}^n \hat{N}_{i,j}^m(t_0, \cdot) \), \( \hat{I}_{i,j}(t_0, \cdot) = \sum_{m=1}^n \hat{I}_{i,j}^m(t_0, \cdot) \), and \( \hat{N}_i(t_0, \cdot) = \sum_{m=1}^n \hat{N}_i^m(t_0, \cdot) \).

Since \( \Delta \hat{N}_{i,j}^m(t_0, t) = 1 \) implies \( \hat{p}_{i,j}^m(t_0, t) = 1 \), by applying Jensen’s inequality and by using the mutual independence of the random variables \( \hat{N}_{i,j}^m(t_0, \cdot), \ldots, \hat{N}_{i,j}^m(t_0, \cdot) \), we get
\[
\mathbb{E} \left[ \mathbb{I}_{ \{ \Delta \hat{N}_{i,j}^m(t_0,u) = 0 \} } \frac{\hat{I}_{i,j}^m(t_0,u)}{\hat{I}_{i,j}(t_0,u)} \right] \hat{N}_{i,j}^m(t_0, \cdot) \leq \frac{1}{1 + \mathbb{E} \left[ \hat{I}_{i,j}(t_0,u) - \hat{I}_{i,j}^m(t_0,u) \right]} \leq \frac{1}{1 + (n-1) \hat{q}_{i,j}(t_0,u) - \hat{I}_{i,j}(t_0,u)} \leq \frac{1}{n \hat{q}_{i,j}(t_0,u)}.
\]

Because of this inequality, the law of iterated expectations, Campbell’s theorem, and (17), we obtain
\[
\mathbb{E} \left[ \hat{A}_{i,j}(t_0, T) \right] = \frac{1}{n} \sum_{m=1}^n \mathbb{E} \left[ \int_{(t_0,T]} n \mathbb{E} \left[ \frac{\mathbb{I}_{ \{ \hat{I}_{i,j}(t_0,u) = 0 \} } \hat{I}_{i,j}(t_0,u) }{\hat{I}_{i,j}(t_0,u)} \hat{N}_{i,j}^m(t_0, \cdot) \right] \hat{N}_{i,j}^m(t_0, du) \right]
\]
\[
\leq \frac{1}{n} \sum_{m=1}^n \mathbb{E} \left[ \int_{(t_0,T]} \frac{1}{\hat{q}_{i,j}(t_0,u)} \hat{N}_{i,j}^m(t_0, du) \right]
\]
\[
= \mathbb{E} \left[ \int_{(t_0,T]} \frac{1}{\hat{q}_{i,j}(t_0,u)} \hat{N}_{i,j}^m(t_0, du) \right]
\]
\[
= \int_{(t_0,T]} \frac{1}{\hat{q}_{i,j}(t_0,u)} \hat{N}_{i,j}^m(t_0, du)
\]
\[
\hat{A}_{i,j}(t_0, T) < \infty.
\]

Jensen’s inequality and (21) imply that
\[
\mathbb{E} \left[ \mathbb{I}_{ \{ \Delta \hat{N}_{i,j}^m(t_0,u) = 0 \} } \frac{\mathbb{I}_{ \{ \hat{I}_{i,j}(t_0,u) = 0 \} } \hat{N}_{i,j}^m(t_0, \cdot) }{\hat{I}_{i,j}(t_0,u)^2} \right] \leq \mathbb{E} \left[ \mathbb{I}_{ \{ \Delta \hat{N}_{i,j}^m(t_0,u) = 0 \} } \left( \mathbb{E} \left[ \frac{\mathbb{I}_{ \{ \hat{I}_{i,j}(t_0,u) = 0 \} } \hat{I}_{i,j}(t_0,u) }{\hat{I}_{i,j}(t_0,u)} \hat{N}_{i,j}^m(t_0, \cdot) \right] \right)^2 \right] \leq \frac{1}{n^2 \hat{q}_{i,j}(t_0,u)^2}.
\]
By arguing analogously to (22) but using the latter inequality and the Cauchy-Schwarz inequality for conditional expectations, we can show that

\[
\mathbb{E}\left[ \mathbf{I}_{[\tilde{N}_{i}(t_0,T)/n>M]} \tilde{\Lambda}_{i,j,k}(t_0,T) \right] \\
= \mathbb{E}\left[ \int_{(0,T]} \mathbf{I}_{[\tilde{N}_{i}(t_0,T)/n>M]} \frac{\mathbf{I}_{[\tilde{I}_{i,j}(t_0,u-) > 0]}}{\tilde{I}_{i,j}(t_0,u-)} \tilde{N}_{i,j,k}(t_0,du) \right] \\
= \frac{1}{n} \sum_{m=1}^{n} \mathbb{E}\left[ \int_{(0,T]} n \mathbb{E}\left[ \mathbf{I}_{[\tilde{N}_{i}(t_0,T)/n>M]} \frac{\mathbf{I}_{[\tilde{I}_{i,j}(t_0,u-) > 0]}}{\tilde{I}_{i,j}(t_0,u-)} \tilde{N}_{i,j,k}(t_0,\cdot) \right] \tilde{N}_{i,j,k}^{m}(t_0,du) \right] \\
\leq \frac{1}{n} \sum_{m=1}^{n} \mathbb{E}\left[ \sqrt{\mathbb{E}\left[ \mathbf{I}_{[\tilde{N}_{i}(t_0,T)/n>M]} \tilde{N}_{i,j,k}^{m}(t_0,\cdot) \right]} \int_{(0,T]} \sqrt{\frac{1}{q_{i,j}(t_0,u-)^2}} \tilde{N}_{i,j,k}^{m}(t_0,du) \right] \\
= \mathbb{E}\left[ \sqrt{\mathbb{E}\left[ \mathbf{I}_{[\tilde{N}_{i}(t_0,T)/n>M]} \tilde{N}_{i,j,k}^{m}(t_0,\cdot) \right]} \int_{(0,T]} \frac{1}{q_{i,j}(t_0,u-)} \tilde{N}_{i,j,k}^{m}(t_0,du) \right]. 
\]

(23)

The random variable

\[
\sqrt{\mathbb{E}\left[ \mathbf{I}_{[\tilde{N}_{i}(t_0,T)/n>M]} \tilde{N}_{i,j,k}^{m}(t_0,\cdot) \right]} \int_{(0,T]} \frac{1}{q_{i,j}(t_0,u-)} \tilde{N}_{i,j,k}^{m}(t_0,du),
\]

which appears in the last expectation, is bounded by the random variable

\[
\int_{(0,T]} \frac{1}{q_{i,j}(t_0,u-)} \tilde{N}_{i,j,k}^{m}(t_0,du),
\]

which has expectation \( \tilde{\Lambda}_{i,j,k}(t_0,T) < \infty \), cf. (22). Markov's inequality for conditional expectations, the mutual independence of \( \tilde{N}_{1}^{m}(t_0,T), \ldots, \tilde{N}_{i}^{m}(t_0,T) \), and the conditional dominated convergence theorem yield

\[
\mathbb{E}\left[ \mathbf{I}_{[\tilde{N}_{i}(t_0,T)/n>M]} \tilde{N}_{i,j,k}^{m}(t_0,\cdot) \right] = \mathbb{E}\left[ \mathbf{I}_{[\tilde{N}_{i}(t_0,T)/n>M]} \left[ \tilde{N}_{i,j,k}^{m}(t_0,\cdot) \right]_{j',k' \neq k} \right] \tilde{N}_{i,j,k}^{m}(t_0,\cdot) \\
\leq 1 + \mathbb{E}\left[ \frac{1}{N} \sum_{m=1}^{n} \mathbb{E}[\tilde{N}_{i,j,k}^{m}(t_0,T)] + \frac{1}{n} \tilde{N}_{i,j,k}^{m}(t_0,T) \right] \tilde{N}_{i,j,k}^{m}(t_0,\cdot) \\
\rightarrow 1 + \frac{\mathbb{E}[\tilde{N}_{i,j,k}(t_0,T)]}{M}, \quad n \rightarrow \infty.
\]

(24)

From (23) and (24) as well as the dominated convergence theorem, it thus holds that

\[
\lim_{n \rightarrow \infty} \mathbb{E}\left[ \mathbf{I}_{[\tilde{N}_{i}(t_0,T)/n>M]} \tilde{\Lambda}_{i,j,k}(t_0,T) \right] \leq \sqrt{\frac{\mathbb{E}[\tilde{N}_{i,j,k}(t_0,T)]}{M}} \tilde{\Lambda}_{i,j,k}(t_0,T) \\
\leq \sqrt{\sum_{j',k' \in \mathbb{Z}} \mathbb{E}[\tilde{N}_{j',k'}(T)] \tilde{\Lambda}_{i,j,k}(t_0,T) / M}. 
\]

(25)

Finally, by applying the triangle inequality, (56), (25), and (20), we may conclude that

\[
\lim_{n \rightarrow \infty} \mathbb{E}\left[ \mathbf{I}_{[\tilde{N}_{i}(t_0,T)/n>M]} \| \tilde{\Lambda}_{i,j,k}(t_0,\cdot) - \Gamma_{i,j,k}(t_0,\cdot) \|_{p} \right] \\
\leq \lim_{n \rightarrow \infty} \mathbb{E}\left[ \mathbf{I}_{[\tilde{N}_{i}(t_0,T)/n>M]} \{ 2 \tilde{\Lambda}_{i,j,k}(t_0,T) + 2 \Gamma_{i,j,k}(t_0,\cdot) \} \right] \\
\leq \sqrt{\sum_{j',k' \in \mathbb{Z}} \mathbb{E}[\tilde{N}_{j',k'}(T)] / M} 2 \tilde{\Lambda}_{i,j,k}(t_0,T) + \sqrt{\sum_{j',k' \in \mathbb{Z}} \mathbb{E}[\tilde{N}_{j',k'}(T)] / M} 2 \Gamma_{i,j,k}(t_0,\cdot). 
\]
which converges to zero as $M \to \infty$ due to (9) and (17). This exactly establishes (19).

We now turn our attention to proving (18). Fix $\varepsilon > 0$. By using the triangle inequality, we can show that

$$\begin{align*}
L_{\{N(t_0,T) \leq M\}}[\hat{\Lambda}_{i,j,k}(t_0,\cdot) - \Gamma_{i,j,k}(t_0,\cdot)]_p & \leq L_{\{N(t_0,T) \leq M\}} \left[ \frac{1}{n} \int_{(t_0,t]} \frac{1}{t} \tilde{N}_{i,j,k}(t_0,du) - \int_{(t_0,t]} \frac{1}{q_i,j(t_0,u)} q_{i,j,k}(t_0,du) \right]_p \\
& + \left\| \int_{(t_0,t]} f_{\varepsilon}(\hat{I}_{i,j}(t_0,u)/n) q_{i,j,k}(t_0,du) \right\|_p + \left\| \int_{(t_0,t]} \frac{1}{q_i,j(t_0,u) - \varepsilon} q_{i,j,k}(t_0,du) \right\|_p \\
& \leq 2 \int_{(t_0,t]} \frac{1}{q_i,j(t_0,u)} q_{i,j,k}(t_0,du) \to 0, \quad \varepsilon \to 0.
\end{align*}$$

(26)

for a continuous mapping $f_{\varepsilon} : [0, \infty) \to [0, 1]$ that equals 1 on $[0, \varepsilon]$ and equals 0 on $[2\varepsilon, \infty)$. Actually, using (56) and the monotone convergence theorem, it holds that

$$\int_{(t_0,t]} \frac{1}{q_i,j(t_0,u) - \varepsilon} q_{i,j,k}(t_0,du) \to 0, \quad \varepsilon \to 0.$$

(27)

To prove (18), we are going to show that as $n \to \infty$,

a) the expectation of the fourth line of (26) is bounded by twice the second line of (27) but with $\varepsilon$ replaced by $2\varepsilon$

b) the expectation of the second and third line of (26) goes to zero.

This suffices since a) and b) in conjunction imply that for any $M \in \mathbb{N}$,

$$\lim_{n \to \infty} \mathbb{E}\left[ L_{\{N(t_0,T) \leq M\}}[\hat{\Lambda}_{i,j,k}(t_0,\cdot) - \Gamma_{i,j,k}(t_0,\cdot)]_p \right]$$

is bounded by

$$2 \int_{(t_0,t]} \frac{1}{q_i,j(t_0,u)} q_{i,j,k}(t_0,du) + \int_{(t_0,t]} \frac{1}{q_i,j(t_0,u) - 2\varepsilon} q_{i,j,k}(t_0,du),$$

so that (18) follows from letting $\varepsilon \to 0$ in accordance with (27).

To establish a), we argue analogously to (23) and (22) in order to obtain that

$$\mathbb{E}\left[ \int_{(t_0,T]} \frac{f_{\varepsilon}(\hat{I}_{i,j}(t_0,u)/n)}{I_{i,j}(t_0,u)} \tilde{N}_{i,j,k}(t_0,du) \right] \leq \frac{1}{n} \sum_{m=1}^{n} \mathbb{E}\left[ \int_{(t_0,T]} \sqrt{f_{\varepsilon}(\hat{I}_{i,j}^m(t_0,u)/\varepsilon)} \tilde{N}_{i,j,k}(t_0,du) \right]$$

$$= \int_{(t_0,T]} \mathbb{E}\left[ \frac{1}{q_i,j(t_0,u)} q_{i,j,k}(t_0,du) \right]$$

for $\hat{I}_{i,j}^m(t_0,\cdot) := I_{i,j}(t_0,\cdot) - \hat{I}_{i,j}(t_0,\cdot)$. For any $u \in (t_0,T]$, the strong law of large numbers yields

$$\frac{\hat{I}_{i,j}^m(t_0,u)}{\varepsilon} \to \mathbb{E}[I_{i,j}(t_0,u)] = q_{i,j}(t_0,u), \quad n \to \infty,$$

so by the continuous mapping theorem, we find that

$$\lim_{n \to \infty} \sqrt{\mathbb{E}\left[ \frac{1}{q_i,j(t_0,u) + 1/n^2} q_{i,j,k}(t_0,du) \right]} \leq \mathbb{E}[\hat{I}_{i,j}(t_0,u) - \varepsilon].$$

Invoking (56), collecting results, and applying the dominated convergence theorem, we may conclude that

$$\lim_{n \to \infty} \mathbb{E}\left[ \left\| \int_{(t_0,T]} \frac{f_{\varepsilon}(\hat{I}_{i,j}(t_0,u)/n)}{I_{i,j}(t_0,u)} \tilde{N}_{i,j,k}(t_0,du) \right\|_p \right]$$

$$\leq \lim_{n \to \infty} \mathbb{E}\left[ 2 \int_{(t_0,T]} \frac{f_{\varepsilon}(\hat{I}_{i,j}(t_0,u)/n)}{I_{i,j}(t_0,u)} \tilde{N}_{i,j,k}(t_0,du) \right]$$

$$\leq 2 \int_{(t_0,T]} \frac{1}{q_i,j(t_0,u) - 2\varepsilon} q_{i,j,k}(t_0,du).$$
which establishes a). We now turn our attention to b). By applying the triangle inequality (53), and (56), the second and third line of (26) are bounded by

\[
I_{[\hat{N}(t_0,T)/n \leq M]} M_p \left[ \frac{1}{\hat{I}_{i,j}(t_0,\cdot)/n} \vee \epsilon \right] - \frac{1}{q_{i,j}(t_0,\cdot) \vee \epsilon} \right]_{[p]} + I_{[\hat{N}(t_0,T)/n \leq M]} k_p \left[ \frac{1}{\hat{I}_{i,j}(t_0,\cdot) \vee \epsilon} \right]_{[p]} \left[ n^{-1} \hat{N}_{i,j}(t_0,\cdot) - q_{i,j}(t_0,\cdot) \right]_{[p]}.
\]

(28)

Invoking (9), an application of Theorem 3 in Overgaard (2019a) gives

\[
\mathbb{E}\left[ \left\| n^{-1} \hat{N}_{i,j}(t_0,\cdot) - q_{i,j}(t_0,\cdot) \right\|_{[p]} \right] \to 0, \quad n \to \infty,
\]

(29)

for all \( j', k' \in Z, j' \neq k' \). We establish b) exactly by providing a link between the bound of (28) and \( p \)-variations such as those of (29). To this end, note that the equation \( \frac{1}{\epsilon} - \frac{1}{\epsilon} = \frac{b - \rho}{\epsilon} \) and (52) yield the inequality

\[
\left\| \frac{1}{\hat{I}_{i,j}(t_0,\cdot)/n} \vee \epsilon \right\|_{[p]} \leq \left\| \hat{I}_{i,j}(t_0,\cdot)/n \vee \epsilon \right\|_{[p]} \frac{1}{q_{i,j}(t_0,\cdot) \vee \epsilon} \left\| q_{i,j}(t_0,\cdot) \vee \epsilon \right\|_{[p]}.
\]

(30)

From the identity

\[
I_{[\tilde{Z}=j]} = I_{[\tilde{Z}_0=j]} + \sum_{\ell \in \tilde{Z} : \ell \neq j} (\hat{N}_{i,j}(\cdot) - \hat{N}_{i,j}(\cdot)),
\]

(31)

the triangle inequality, and the monotonicity of \( q_{i,j}(t_0,\cdot), \ell \in \tilde{Z}, \ell \neq j \), we have

\[
\left\| q_{i,j}(t_0,\cdot) \right\|_{v_1} \leq \sum_{\ell \in \tilde{Z}, \ell \neq j} (q_{i,j}(t_0, T) + q_{i,j}(t_0, T))) < \infty,
\]

where the finiteness follows from \( \mathbb{E}[\hat{N}_{j,k}(T)] < \infty \) for all \( j', k' \in \tilde{Z} \), cf. (9). According to (50), (55), and (54), we then find that

\[
\left\| \frac{1}{q_{i,j}(t_0,\cdot) \vee \epsilon} \right\|_{[p]} \leq \frac{1}{\epsilon} + \frac{2}{\epsilon^2} \left\| q_{i,j}(t_0,\cdot) \right\|_{v_1} < \infty.
\]

(32)

The identity (31) moreover implies that

\[
\left\| \hat{I}_{i,j}(t_0,\cdot)/n \right\|_{v_1} \leq \frac{1}{n} \sum_{\ell \in \tilde{Z}, \ell \neq j} (\hat{N}_{i,j}(t_0, T) + \hat{N}_{i,j}(t_0, T)),
\]

so that (50), (55), and (54) yield that

\[
I_{[\hat{N}(t_0,T)/n \leq M]} \left\| \frac{1}{\hat{I}_{i,j}(t_0,\cdot)/n} \vee \epsilon \right\|_{[p]} \leq I_{[\hat{N}(t_0,T)/n \leq M]} \left( \frac{1}{\epsilon} + \frac{2}{\epsilon^2} \left\| \hat{I}_{i,j}(t_0,\cdot)/n \right\|_{v_1} \right)
\]

\[
\leq \frac{1}{\epsilon} + \frac{2}{\epsilon^2} M.
\]

(33)

Combining the identity (30) and the bounds (32) and (33), and using the inequalities (52), (54), and

\[
\left\| \hat{I}_{i,j}(t_0,\cdot)/n \vee \epsilon - q_{i,j}(t_0,\cdot) \vee \epsilon \right\|_{\infty} \leq \left\| \hat{I}_{i,j}(t_0,\cdot)/n - q_{i,j}(t_0,\cdot) \right\|_{\infty},
\]

we find that

\[
I_{[\hat{N}(t_0,T)/n \leq M]} \left\| \frac{1}{\hat{I}_{i,j}(t_0,\cdot)/n} \vee \epsilon \right\|_{[p]} - \frac{1}{q_{i,j}(t_0,\cdot) \vee \epsilon} \right\|_{[p]} \leq k^2 \left\| \hat{I}_{i,j}(t_0,\cdot)/n - q_{i,j}(t_0,\cdot) \right\|_{v_1} \left( \frac{1}{\epsilon} + \frac{2}{\epsilon^2} \left\| q_{i,j}(t_0,\cdot) \right\|_{v_1} \right)
\]

\[
\leq \frac{2}{\epsilon^2} \left\| \hat{I}_{i,j}(t_0,\cdot)/n - q_{i,j}(t_0,\cdot) \right\|_{v_1} \left( \frac{1}{\epsilon} + \frac{2}{\epsilon^2} M \right).
\]

From the identity (31) and the triangle inequality, we finally get

\[
\left\| \hat{I}_{i,j}(t_0,\cdot)/n - q_{i,j}(t_0,\cdot) \right\|_{[p]} \leq \sum_{\ell \in \tilde{Z}, \ell \neq j} \left( \left\| n^{-1} \hat{N}_{i,j}(t_0,\cdot) - q_{i,j}(t_0,\cdot) \right\|_{[p]} + \left\| n^{-1} \hat{N}_{i,j}(t_0,\cdot) - q_{i,j}(t_0,\cdot) \right\|_{[p]} \right)
\]

Collecting results and pointing to (29), we conclude that the expectation of the bound (28) converges to zero as \( n \to \infty \). This establishes b) and thus completes the proof. \( \square \)

The following result establishes also consistency of the corresponding landmark Aalen–Johansen estimators.
Theorem 3. Suppose that the assumptions of Theorem 2 are satisfied. Let \( p \in (1, 2) \). Then for each \( j \in \tilde{Z} \) it holds that

\[
\mathbb{E}\left[ \left\| \hat{\mathbf{p}}_{i,j}(t_0, \cdot) - \tilde{\mathbf{p}}_{i,j}(t_0, \cdot) \right\|_{[p]} \right] \to 0, \quad n \to \infty.
\]

Proof. Fix \( j \in \tilde{Z} \). For \( t \geq t_0 \) let \( \mathbf{p}_i(t_0, t) := (\mathbf{p}_i(t_0, t))_t \) be a row vector and let \( \mathbf{\hat{A}}_i(t_0, t) := \mathbf{\hat{A}}_i(t_0, t)_{t} \) be a matrix with diagonal entries \( \mathbf{\hat{A}}_{i,k}(t_0, t) := \sum_{k \neq l} \mathbf{\hat{A}}_{i,k}(t_0, t) \). The solution to (14) with respect to the landmark Nelson–Aalen estimator and the initial values \( \mathbf{\hat{p}}(t_0, t_0) \) from Definition 2 can be represented as

\[
\mathbf{\hat{p}}(t_0, t) = \mathbf{\hat{p}}(t_0, t_0) \prod_{(t_0, t]} (1 + \mathbf{\hat{A}}_i(t_0, du)), \quad t_0 \leq t < \infty,
\]

where \( \mathbf{1} \) denotes the identity matrix and \( \prod \) the product integral, see also Gill and Johansen (1990). In similar fashion,

\[
\tilde{\mathbf{p}}(t_0, t) = \tilde{\mathbf{p}}(t_0, t_0) \prod_{(t_0, t]} (1 + \tilde{\mathbf{A}}_i(t_0, du)), \quad t_0 \leq t < \infty,
\]

where \( \mathbf{\tilde{p}}(t_0, t) := (\mathbf{\tilde{p}}_i(t_0, t))_t \) is a row vector and \( \mathbf{\tilde{A}}_i(t_0, t) := (\mathbf{\tilde{A}}_{i,k}(t_0, t))_{k \neq l} \) is a matrix with diagonal entries \( \mathbf{\tilde{A}}_{i,k}(t_0, t) := \sum_{k \neq l} \mathbf{\tilde{A}}_{i,k}(t_0, t) \) for \( t \geq t_0 \). Thus by repeated application of the triangle inequality and using (50), we have that

\[
\left\| \mathbf{\hat{p}}_{i,j}(t_0, \cdot) - \tilde{\mathbf{p}}_{i,j}(t_0, \cdot) \right\|_{[p]} \leq c \left\| \mathbf{\hat{p}}(t_0, t_0) - \tilde{\mathbf{p}}(t_0, t_0) \right\|_1 \left\| \prod_{(t_0, t]} (1 + \mathbf{\hat{A}}_i(t_0, du)) \right\|_1
\]

\[
+ c \left\| \mathbf{\hat{p}}(t_0, t_0) \right\|_1 \left\| \prod_{(t_0, t]} (1 + \mathbf{\hat{A}}_i(t_0, du)) - \prod_{(t_0, t]} (1 + \tilde{\mathbf{A}}_i(t_0, du)) \right\|_{[p]},
\]

where \( c \) is a finite constant that depends on the choice of matrix-norm.

The product integral in (34) equals the solution of (14) with respect to the landmark Nelson–Aalen estimator and the initial values \( \mathbf{1}_{\{|x\}}_t, \ell \in \tilde{Z} \). Consequently, the inequality (50), the triangle inequality and (77) of Theorem 12 in Gill and Johansen (1990) yield

\[
\left\| \prod_{(t_0, t]} (1 + \mathbf{\hat{A}}_i(t_0, du)) \right\|_{[p]} \leq \left\| \prod_{(t_0, t]} (1 + \tilde{\mathbf{A}}_i(t_0, du)) \right\|_{[p]} \leq \left\| \mathbf{1} \right\|_1 + 2k_p \sum_{\ell, k \in \tilde{Z}, \ell \neq k} \left\| \mathbf{\hat{A}}_{i,k}(t_0, \cdot) \right\|_{11}.
\]

Since \( \mathbf{\hat{p}}_{i,\ell}(t_0, t_0), \ell \in \tilde{Z} \), are consistent estimators for \( \mathbf{\hat{p}}_{i,\ell}(t_0, t_0), \ell \in \tilde{Z} \) in the sense of \( L_1 \)-convergence, we may then conclude that as \( n \to \infty \),

\[
\mathbb{E}\left[ \mathbf{1}_{\{\sum_{\ell \epsilon \tilde{Z}, k \neq l} \left| \mathbf{\hat{A}}_{i,k}(t_0, \cdot) \right|_{11} \leq M \}} \left\| \mathbf{\hat{p}}(t_0, t_0) - \mathbf{\hat{p}}(t_0, t_0) \right\|_1 \left\| \prod_{(t_0, t]} (1 + \mathbf{\hat{A}}_i(t_0, du)) \right\|_{[p]} \right]\]

converges to zero for any \( M \in \mathbb{N} \).

According to Duhamel’s equation for product integrals followed by two applications of (53), it holds that

\[
\left\| \prod_{(t_0, t]} (1 + \mathbf{\hat{A}}_i(t_0, du)) - \prod_{(t_0, t]} (1 + \tilde{\mathbf{A}}_i(t_0, du)) \right\|_{[p]}
\]

\[
\leq k_p \left\| \prod_{(t_0, t]} (1 + \mathbf{\hat{A}}_i(t_0, du)) \right\|_{[p]} \left\| \mathbf{\hat{A}}_i(t_0, \cdot) - \tilde{\mathbf{A}}_i(t_0, \cdot) \right\|_{[p]} \left\| \prod_{(t_0, t]} (1 + \tilde{\mathbf{A}}_i(t_0, du)) \right\|_{[p]}.
\]

In similar fashion to (37), we also have that

\[
\left\| \prod_{(t_0, t]} (1 + \mathbf{\hat{A}}_i(t_0, du)) \right\|_{[p]} \leq \left\| \mathbf{1} \right\|_1 + 2k_p \sum_{\ell, k \in \tilde{Z}, \ell \neq k} \left\| \mathbf{\hat{A}}_{i,k}(t_0, \cdot) \right\|_{11} < \infty,
\]

where the finiteness follows from (56) and (13). Since \( \mathbf{\hat{p}}_{i,\ell}(t_0, t_0) \leq 1 \) for \( \ell \in \tilde{Z} \) from Theorem 2 it follows that as \( n \to \infty \),

\[
\mathbb{E}\left[ \left\| \mathbf{1}_{\{\sum_{\ell \epsilon \tilde{Z}, k \neq l} \left| \mathbf{\hat{A}}_{i,k}(t_0, \cdot) \right|_{11} \leq M \}} \left\| \mathbf{\hat{p}}(t_0, t_0) \right\|_1 \left\| \prod_{(t_0, t]} (1 + \mathbf{\hat{A}}_i(t_0, du)) - \prod_{(t_0, t]} (1 + \tilde{\mathbf{A}}_i(t_0, du)) \right\|_{[p]} \right]\]

converges to zero. Collecting results, we conclude that

\[
\mathbb{E}\left[ \left\| \mathbf{1}_{\{\sum_{\ell \epsilon \tilde{Z}, k \neq l} \left| \mathbf{\hat{A}}_{i,k}(t_0, \cdot) \right|_{11} \leq M \}} \left\| \mathbf{\hat{p}}_{i,j}(t_0, \cdot) - \tilde{\mathbf{p}}_{i,j}(t_0, \cdot) \right\|_{[p]} \right]\to 0, \quad n \to \infty,
\]

for any \( M \in \mathbb{N} \). Similar to (36), we have

\[
\left\| \mathbf{\hat{p}}_{i,j}(t_0, \cdot) - \tilde{\mathbf{p}}_{i,j}(t_0, \cdot) \right\|_{[p]} \leq c \left\| \mathbf{\hat{p}}(t_0, t_0) \right\|_1 \left\| \prod_{(t_0, t]} (1 + \mathbf{\hat{A}}_i(t_0, du)) \right\|_{[1]} + c \left\| \mathbf{\hat{p}}(t_0, t_0) \right\|_1 \left\| \prod_{(t_0, t]} (1 + \tilde{\mathbf{A}}_i(t_0, du)) \right\|_{[1]}.
\]
Since \(|\hat{\theta}_{j}(t,0) - \bar{\theta}_{j}(t,0)| \leq 1 \) and \(|\hat{\theta}_{j}(t,0) - \bar{\theta}_{j}(t,0)| \leq 1 \) for \( \ell \in \mathbb{Z} \), it follows from (37) and (38) that
\[
\|\hat{\theta}_{j}(t,0) - \bar{\theta}_{j}(t,0)\|_{[p]} \leq c_{1} + c_{2} \sum_{\ell \in \mathbb{Z}, \ell \neq k} \|\hat{\theta}_{\ell,k}(t_{0},\cdot)\|_{[1]}
\]
for suitable real numbers \( c_{1}, c_{2} < \infty \). The triangle inequality and (56) moreover yield
\[
\|\hat{\theta}_{j}(t,0) - \bar{\theta}_{j}(t,0)\|_{[p]} \leq c_{1} + c_{2} \sum_{\ell \in \mathbb{Z}, \ell \neq k} 2\left(\|\hat{\theta}_{\ell,k}(t_{0},\cdot) - \bar{\theta}_{\ell,k}(t_{0},\cdot)\|_{[p]} + \|\hat{\theta}_{\ell,k}(t_{0},\cdot)\|_{[p]}\right).
\]
So, in conjunction with Theorem 2, the dominated convergence theorem, and (56), we get
\[
\lim_{n \to \infty} E\left[\mathbb{I}\left[\sum_{\ell \in \mathbb{Z}, \ell \neq k} |\hat{\theta}_{\ell,k}(t_{0},\cdot) - \bar{\theta}_{\ell,k}(t_{0},\cdot)|_{[1]} \right] \right] \|\hat{\theta}_{j}(t_{0},\cdot) - \bar{\theta}_{j}(t_{0},\cdot)\|_{[p]} 
\leq \lim_{n \to \infty} E\left[\mathbb{I}\left[\sum_{\ell \in \mathbb{Z}, \ell \neq k} |\hat{\theta}_{\ell,k}(t_{0},\cdot) - \bar{\theta}_{\ell,k}(t_{0},\cdot)|_{[1]} \right] \left(c_{1} + c_{2} \sum_{\ell \in \mathbb{Z}, \ell \neq k} 2\|\hat{\theta}_{\ell,k}(t_{0},\cdot)\|_{[p]} \right)\right]
\leq \lim_{n \to \infty} E\left[\mathbb{I}\left[\sum_{\ell \in \mathbb{Z}, \ell \neq k} 2|\hat{\theta}_{\ell,k}(t_{0},T) - \bar{\theta}_{\ell,k}(t_{0},T)|_{[1]} \right] \right] \left(c_{1} + c_{2} \sum_{\ell \in \mathbb{Z}, \ell \neq k} 4\|\hat{\theta}_{\ell,k}(t_{0},T)\|_{[1]} \right)
\]
Markov's inequality and (22) then yield
\[
\lim_{n \to \infty} E\left[\mathbb{I}\left[\sum_{\ell \in \mathbb{Z}, \ell \neq k} |\hat{\theta}_{\ell,k}(t_{0},\cdot) - \bar{\theta}_{\ell,k}(t_{0},\cdot)|_{[1]} \right] \right] \|\hat{\theta}_{j}(t_{0},\cdot) - \bar{\theta}_{j}(t_{0},\cdot)\|_{[p]} \leq \frac{1}{M} \left(\sum_{\ell \in \mathbb{Z}, \ell \neq k} 2\|\hat{\theta}_{\ell,k}(t_{0},T)\|_{[1]} \right) \left(c_{1} + c_{2} \sum_{\ell \in \mathbb{Z}, \ell \neq k} 4\|\hat{\theta}_{\ell,k}(t_{0},T)\|_{[1]} \right),
\]
which converges to zero for \( M \to \infty \), see also (13). Combining this fact with (39) establishes the desired result. \( \square \)

Let \( \hat{\theta}_{j}(t_{0},\cdot) \) be given by (15), but with the auxiliary probabilities \( \bar{\theta}_{j}(t_{0},\cdot) \), \( j \in \mathbb{Z} \), and the auxiliary (cumulative) forward transition rates \( \hat{\theta}_{j,k}(t_{0},\cdot), j, k \in \mathbb{Z}, j \neq k \), replaced by their landmark estimators \( \hat{\theta}_{j,k}(t_{0},\cdot), j \in \mathbb{Z}, \) and \( \hat{\theta}_{j,k}(t_{0},\cdot), j, k \in \mathbb{Z}, \) respectively. The next theorem shows that the consistency results for the landmark Nelson–Aalen estimator and the landmark Aalen–Johnsen estimator carry over to the estimator \( \hat{\theta}_{j}(t_{0},\cdot) \) for the auxiliary-state-wise expected accumulated cash flow \( \hat{\theta}_{j}(t_{0},\cdot) \).

**Theorem 4.** Suppose that the assumptions of Theorem 2 are satisfied. Let \( p \in (1, 2) \). Then it holds that
\[
E\left[\|\hat{\theta}_{j}(t_{0},\cdot) - \bar{\theta}_{j}(t_{0},\cdot)\|_{[p]} \right] \to 0, \quad n \to \infty.
\]

**Proof.** By applying (15), the triangle inequality, (53), and (50), we obtain
\[
\|\hat{\theta}_{j}(t_{0},\cdot) - \bar{\theta}_{j}(t_{0},\cdot)\|_{[p]} \leq \sum_{j \in \mathbb{Z}} k_{p} \|\hat{\theta}_{j}(t_{0},\cdot) - \bar{\theta}_{j}(t_{0},\cdot)\|_{[p]} \|B_{[1]} \|
+ \sum_{j, k \in \mathbb{Z}, j \neq k} k_{p}^{2} \|\hat{\theta}_{j,k}(t_{0},\cdot) - \bar{\theta}_{j,k}(t_{0},\cdot)\|_{[p]} \|B_{jk}[1]\|\|\hat{\theta}_{j,k}(t_{0},\cdot)\|_{[1]}\]
+ \sum_{j, k \in \mathbb{Z}, j \neq k} k_{p}^{2} \|\bar{\theta}_{j,k}(t_{0},\cdot)\|_{[1]} \|B_{jk}[1]\|\|\hat{\theta}_{j,k}(t_{0},\cdot) - \bar{\theta}_{j,k}(t_{0},\cdot)\|_{[p]}\|.
\]
It holds that \( \|B_{j}[1]\| < \infty \) and \( \|B_{jk}[1]\| < \infty \) for all \( j, k \in \mathbb{Z}, j \neq k \), since we assumed that the sojourn and transition payments have finite variation. Furthermore, we have \( \|\bar{\theta}_{j}(t_{0},\cdot)\|_{[1]} < \infty \) for \( j \in \mathbb{Z} \), which follows from (35), the upper bound \( \|\bar{\theta}_{j}(t_{0},t_{0}| \leq 1 \) for all \( \ell \in \mathbb{Z} \), and (38). Therefore, Theorem 2 and Theorem 3 yield that
\[
E\left[\mathbb{I}\left[\sum_{j \in \mathbb{Z}, j \neq k} |\hat{\theta}_{j,k}(t_{0},\cdot) - \bar{\theta}_{j,k}(t_{0},\cdot)|_{[1]} \right] \right] \rightarrow 0, \quad n \to \infty,
\]
for any \( M \in \mathbb{N} \). By applying the triangle inequality, the inequalities (50) and (51), and the fact that \( \int_{[0,1]} g(s) h(ds) \|v_{1} \leq g_{\infty} h \|v_{1} \), we obtain
\[
\|\hat{\theta}_{j}(t_{0},\cdot) - \bar{\theta}_{j}(t_{0},\cdot)\|_{[p]} \leq \|\hat{\theta}_{j}(t_{0},\cdot)\|_{[1]} + \|\bar{\theta}_{j}(t_{0},\cdot)\|_{[1]}
\leq 2 \sum_{j \in \mathbb{Z}} \|\hat{\theta}_{j}(t_{0},\cdot)\|_{\infty} \|B_{j}\|_{v_{1}} + 2 \sum_{j, k \in \mathbb{Z}, j \neq k} \|\hat{\theta}_{j,k}(t_{0},\cdot)\|_{\infty} \|B_{jk}\|_{v_{1}} \|\hat{\theta}_{j,k}(t_{0},\cdot)\|_{v_{1}}
+ 2 \sum_{j \in \mathbb{Z}} \|\hat{\theta}_{j,k}(t_{0},\cdot)\|_{\infty} \|B_{jk}\|_{v_{1}} + 2 \sum_{j, k \in \mathbb{Z}, j \neq k} \|\hat{\theta}_{j,k}(t_{0},\cdot)\|_{\infty} \|B_{jk}\|_{v_{1}} \|\hat{\theta}_{j,k}(t_{0},\cdot)\|_{v_{1}}.
\]
Since $|\hat{p}_{i,\ell}(t_0, t_0)| \leq 1$ and $|\tilde{p}_{i,\ell}(t_0, t_0)| \leq 1$ for all $\ell \in \mathcal{Z}$ and since $\hat{A}_{i,j,k}(t_0, \cdot)$ and $\tilde{A}_{i,j,k}(t_0, \cdot)$ are monotone for all $j, k \in \mathcal{Z}, j \neq k$, we have

$$\|\hat{A}_{i,j,k}(t_0, \cdot) - \tilde{A}_{i,j,k}(t_0, \cdot)\|_{[p]} \leq c_1 + c_2 \sum_{j \neq k} \|\hat{A}_{i,j,k}(t_0, \cdot)\|_{[l]}$$

for suitable real numbers $c_1, c_2 < \infty$. To complete the proof, follow the proof of Theorem 3 starting from (40) but with $\|\hat{p}_{i,j}(t_0, \cdot) - \tilde{p}_{i,j}(t_0, \cdot)\|_{[p]}$ replaced by $\|\hat{A}_{i,j,k}(t_0, \cdot) - \tilde{A}_{i,j,k}(t_0, \cdot)\|_{[p]}$ and (39) replaced by (41). \[ \square \]

The consistency result for the state-wise expected accumulated cash flows $\hat{A}_{i}(t_0, \cdot)$ and $\tilde{A}_{i}(t_0, \cdot)$ carries over to the corresponding prospective reserves. Analogously to formula (4), let $\hat{V}(t_0)$ and $\tilde{V}(t_0)$ be the prospective reserves at time $t_0$ of $\hat{A}_{i}(t_0, \cdot)$ and $\tilde{A}_{i}(t_0, \cdot)$, respectively. The following corollary is then a consequence of Theorem 4 and (51).

**Corollary.** Suppose that the assumptions of Theorem 2 are satisfied. Let $p \in (1, 2)$, and suppose that $\|\frac{1}{n}\|_{[p]} < \infty$. It then holds that

$$\mathbb{E}[\hat{V}(t_0) - \tilde{V}(t_0)] \to 0, \quad n \to \infty.$$  

**5. Implementation**

A central quantity of interest for life insurance actuaries is the state-wise prospective reserve

$$V_i(t_0) := \mathbb{E}\left[ \int_{(t_0, \tau]} \frac{\kappa(s)}{\kappa(t_0)} B(ds) \bigg| Z_{t_0} = i \right] = \int_{(t_0, \tau]} \frac{\kappa(s)}{\kappa(t_0)} A_i(t_0, ds),$$

where $A_i(t_0, s) := \mathbb{E}[B(s) - B(t_0) | Z_{t_0} = i]$. If $B$ has a deterministic cash flow representation, then we can calculate $V_i(t_0)$ from the explicit formula

$$V_i(t_0) = \sum_{j \in \mathcal{Z}} \int_{(t_0, \tau]} \frac{\kappa(t_0)}{\kappa(s)} p_{i,j}(t_0, s) B_j(ds) + \sum_{j \neq k \in \mathcal{Z}} \int_{(t_0, \tau]} \frac{\kappa(t_0)}{\kappa(s)} b_{jk}(s)p_{i,j}(t_0, s) A_{i,j,k}(t_0, ds), \quad (42)$$

cf. (7). If $Z$ is a Markov process, then the rates $A_{i,j,k}(t_0, \cdot)$, $j, k \in \mathcal{Z}, j \neq k$, do not depend on state $i$ and time $t_0$ and can be estimated by the Nelson–Aalen estimator. Suppose that the sequence of time points $t_0 < t_1 < t_2 < \cdots < t_m = \tau$ describes all empirically observed jump times and censoring times in the statistical sample. Then the Nelson–Aalen estimator is a right-continuous step function on the grid $\{t_0, t_1, \ldots, t_m\}$, and it may conveniently be calculated from the recursion schemes

$$\tilde{A}_{i,j,k}(t_0, t_{j+1}) = \tilde{A}_{i,j,k}(t_0, t_j) + \tilde{q}_{i,j,k}(t_0, t_j, t_{j+1})$$

with starting values $\tilde{A}_{i,j,k}(t_0, t_0) = 0$, where

$$\tilde{q}_{i,j,k}(t_0, t_j, t_{j+1}) := \frac{\tilde{N}_{i,j,k}(t_0, t_{j+1}) - \tilde{N}_{i,j,k}(t_0, t_j)}{\tilde{I}_{i,j}(t_0, t_j)}.$$ \[ \text{(43)} \]

Note that $\tilde{q}_{i,j,k}(t_0, t_j, t_{j+1})$ in this case does not depend on state $i$ and time $t_0$. The transition probabilities $p_{i,j}(t_0, \cdot)$, $j \in \mathcal{Z}$, can be estimated by the Aalen–Johansen estimator, which is also a right-continuous step function on the grid $\{t_0, t_2, \ldots, t_m\}$ that conveniently may be calculated from the recursion schemes

$$\tilde{p}_{i,j}(t_0, t_{j+1}) = \sum_{k \in \mathcal{Z}} \tilde{p}_{i,j,k}(t_0, t_j) \tilde{q}_{i,j,k}(t_0, t_j, t_{j+1}) - \tilde{p}_{i,j}(t_0, t_j) \sum_{k \in \mathcal{Z}} \tilde{q}_{i,j,k}(t_0, t_j, t_{j+1})$$ \[ \text{(44)} \]

with starting values $\tilde{p}_{i,j}(t_0, t_0) = 1_{\{j = i\}}$. By plugging the Nelson–Aalen estimator and the Aalen–Johansen estimator into formula (42), we obtain for $V_i(t_0)$ the estimator

$$\tilde{V}_i(t_0) = \sum_{j \in \mathcal{Z}} m_{i,j} \tilde{p}_{i,j}(t_0, t_0) \int_{(t_j, t_{j+1}]} \frac{\kappa(t_0)}{\kappa(s)} B_j(ds) + \sum_{j \neq k \in \mathcal{Z}} m_{i,j} \int_{(t_j, t_{j+1}]} \frac{\kappa(t_0)}{\kappa(s)} b_{jk}(s) \tilde{p}_{i,j}(t_0, t_j) \tilde{q}_{i,j,k}(t_0, t_j, t_{j+1}). \quad (45)$$

If we apply the estimator (43) on a subsample only, namely those insured who are in state $i$ at time $t_0$,

$$\tilde{q}_{i,j,k}(t_0, t_j, t_{j+1}) := \frac{\tilde{N}_{i,j,k}(t_0, t_{j+1}) - \tilde{N}_{i,j,k}(t_0, t_j)}{\tilde{I}_{i,j}(t_0, t_j)},$$ \[ \text{(46)} \]

then the formulas (44) and (45) may still be used for estimating $V_i(t_0)$, even when $Z$ is non-Markovian, which was a key message of Christiansen (2021). In particular, it holds that

$$\mathbb{E}[\tilde{V}_i(t_0) - V_i(t_0)] \to 0, \quad n \to \infty,$$
when \( Z \) is Markovian as well as non-Markovian. The latter result needs the assumptions of Theorem 2 and further regularity assumptions concerning \( \kappa \), which for example are satisfied if \( \kappa \) has finite variation and a bounded reciprocal on \((t_0, T]\).

Let us turn now to the scaled payments \( B^\theta \). Unfortunately, the state-wise prospective reserve

\[
V^\theta_i(t_0) := \mathbb{E} \left[ \int_{(t_0, T]} \frac{\kappa(s)}{\kappa(t_0)} B^\theta(ds) \left| Z_{t_0} = i \right. \right]
\]

does not satisfy (42), regardless of whether \( Z \) is Markovian or not, since the payments \( B^\theta \) do not admit a deterministic cash flow representation. Instead of replacing (42) by a more sophisticated formula, the concept of Furrer (2022) is in a sense to artificially transform the functions \( p_{i,j} \) and \( \Lambda_{i,j,k} \) in such a way that formula (42) finally produces the wanted quantity \( V^\theta_i(t_0) \). This idea may also be used for the purpose of estimation. To put it briefly, in case that \( Z \) is Markovian, apply the estimator (43) on the transformed sample according to (8), that is

\[
\hat{q}_{i,j,k}(t_0, t_\ell, t_{\ell+1}) := \frac{\sum_{l \in \mathcal{Z}} Z_l \cdot \left( \hat{N}_{i,j,k}(t_0, t_{\ell+1}) - \hat{N}_{i,j,k}(t_0, t_\ell) \right)}{\sum_{l \in \mathcal{Z}} \hat{Z}_{i,j}(t_0, t_\ell)}.
\]

Note that \( \hat{q}_{i,j,k}(t_0, t_\ell, t_{\ell+1}) \) in this case does not depend on state \( i \) and time \( t_0 \). By replacing (43) by (47) in the formulas (44) and (45), we obtain an estimator \( \hat{V}_i(t_0) \) that satisfies

\[
\mathbb{E} \left[ \left( \hat{V}_i(t_0) - V^\theta_i(t_0) \right) \right] \to 0, \quad n \to \infty.
\]

This paper combines the concepts of Christiansen (2021) and Furrer (2022). If we use

\[
\hat{q}_{i,j,k}(t_0, t_\ell, t_{\ell+1}) := \frac{\hat{N}_{i,j,k}(t_0, t_{\ell+1}) - \hat{N}_{i,j,k}(t_0, t_\ell)}{\hat{Z}_{i,j}(t_0, t_\ell)}
\]

instead of (43) in the formulas (44) and (45), then we get an estimator \( \hat{V}_i(t_0) \) that satisfies

\[
\mathbb{E} \left[ \left( \hat{V}_i(t_0) - V^\theta_i(t_0) \right) \right] \to 0, \quad n \to \infty,
\]

whether \( Z \) is Markovian or not.

To summarize, the life insurance actuary can retain the classic calculation formulas (44) and (45) and merely needs to adapt the estimator (43) in order to incorporate policyholder options such as free-policy conversion as well as to cope with non-Markovianity. The proper adaptation is given exactly by (48).

6. Closing remarks

In this section, we discuss some extensions, including multiple policyholder options, as well as how to handle missing data arising due to e.g. left-truncation.

Throughout the paper, we have focused on the case of a single policyholder option and thus a single scaling factor. In practice, one may have to consider multiple policyholder options; both the free policy option and stochastic retirement, for instance, are included in Gad and Nielsen (2016). But already in Furrer (2022), change of measure techniques are developed for multiple policyholder options, so it is actually straightforward to appropriately extend the results of the present paper to multiple policyholder options. Similarly, one may deal with scaling factors which are merely bounded, rather than bounded by one, via reparametrization, cf. Subsection 4.2 of Furrer (2022).

This paper focuses on ‘as-if Markov’ modeling, so we have restricted the information used for valuation to \( \mathcal{G}_t = \sigma(Z_t) \), although practitioners could in principle select any suitable information regime (subject to regulatory constraints). In Section 2, we for instance suggested to also keep track of the scaling factor, which entails using instead the information \( \mathcal{H}_t = \sigma(Z_t) \cup H(t) \). The methods and results of this paper are not applicable in the latter information regime. Frankly speaking, we are grappling with the fact that the random variable \( H(t) \) does not take values in a finite set. One solution could be to approximate \( H(t) \) by a random variable with values in a finite set, similar to what has been proposed for semi-Markov models in Remark 7.9 of Christiansen (2021).

Contrary to Christiansen (2021), where consistency of the landmark estimators of interest is established subject to entirely random left-truncation as well as right-censoring, we have instead established consistency subject to only entirely random right-censoring. An important contribution of Christiansen (2021), which we do not advance here, is a time-reversed perspective that is connected to so-called retrospective reserves. In Christiansen (2021), the inclusion of left-truncation is in parts motivated by the time-reversed perspective. In the setup of the present paper, the inclusion of even entirely random left-truncation is rather straightforward, as long as the insurer is able to determine the scaling factor, which usually requires access to further information concerning the health history of the insured.

Declaration of competing interest

The authors declare no conflicts of interest.

Acknowledgements

We thank Martin Bladt for helpful comments that led to a simplification of the proof of Theorem 2. Christian Furrer has carried out this research in association with the project frame InterAct.
Appendix A. \( p \)-variation

This appendix recollects the definition and fundamental properties of the \( p \)-variation norm for \( p \in [1, 2) \). We generally recommend to consult Dudley and Norvaiša (2011). For \( 0 \leq t_0 < T < \infty \) and \( d \in \mathbb{N} \) we consider a matrix-valued function \( h : [t_0, T] \rightarrow \mathbb{R}^{d \times d} \). Let \( \| \cdot \| \) be any norm on \( \mathbb{R}^{d \times d} \). The supremum
\[
\| h \|_\infty := \sup_{t \in [t_0, T]} \| h(t) \|
\]
defines a norm and the \( p \)-variation
\[
\| h \|_{v_p} := \left( \sup_{t_0 < t_1 < \cdots < t_m = T ; m \in \mathbb{N}} \sum_{k=1}^m \| h(t_k) - h(t_{k-1}) \|^p \right)^{1/p}
\]
(49)
defines a semi-norm, so that the sum
\[
\| h \|_{[p]} := \| h \|_\infty + \| h \|_{v_p}
\]
is a norm on the class of bounded multivariate functions with finite \( p \)-variation. Recall that we generally assume that \( p \in [1, 2) \). We have
\[
\| h \|_{v_p} \leq \| h \|_{v_1},
\]
\[
\| h \|_{[p]} \leq \| h \|_{[1]},
\]
(50)
see Lemma 3.45 in Dudley and Norvaiša (2011). Since any function of finite variation can be represented as the difference of two non-decreasing functions, we have
\[
\| h \|_{\infty} \leq \| h(t_0) \| + \| h \|_{v_1},
\]
\[
\| h \|_{[1]} \leq \| h(t_0) \| + 2\| h \|_{v_1}.
\]
(51)
Let \( g : [t_0, T] \rightarrow \mathbb{R}^{d \times d} \) be another matrix-valued function. There exists a finite constant \( k \) such that \( \| AB \| \leq k \| A \| \| B \| \) for all \( A, B \in \mathbb{R}^{d \times d} \). For this constant it holds that
\[
\| hg \|_{[p]} \leq k \| h \|_{[p]} \| g \|_{[p]},
\]
(52)
see Theorem 3.8 in Dudley and Norvaiša (2011). For each \( p \in [1, 2) \) there exists a finite constant \( k_p \) such that
\[
\left| \int_{t_0} g(s) dh(s) \right|_{[p]} \leq k_p \| g \|_{[p]} \| h \|_{[p]},
\]
(53)
see Corollary 3.91 and Theorem 3.92 in Dudley and Norvaiša (2011). If \( h \) is right-continuous and has finite variation, then we may interpret the integral on the left-hand side as a Lebesgue-Stieltjes integral on the interval \( (t_0, \cdot] \).

For a univariate function \( h \) and any constant \( c > 0 \), by using the estimate \( |(h(t_k) \vee c) - (h(t_{k-1}) \vee c)| \leq |h(t_k) - h(t_{k-1})| \) in (49), we get
\[
\| h(\cdot) \vee c \|_{v_p} \leq \| h \|_{v_p},
\]
\[
\| h(\cdot) \vee c \|_{[p]} \leq \| h \|_{[p]} \vee c.
\]
(54)
If \( h \) is a univariate and right-continuous function that has finite variation and satisfies \( h(t) \geq \varepsilon > 0 \) for all \( t \in [t_0, T] \), then the product rule for Lebesgue-Stieltjes integrals yields that
\[
\frac{1}{h(t)} = \frac{1}{h(t_0)} - \int_{(t_0, t]} \frac{1}{h(u)} h(u) - h(du), \quad t \in [t_0, T],
\]
and we can conclude that
\[
\| \frac{1}{h} \|_{v_1} \leq \frac{1}{\varepsilon^2} \| h \|_{v_1},
\]
\[
\| \frac{1}{h} \|_{[1]} \leq \frac{2}{\varepsilon} \| h \|_{v_1}.
\]
(55)
For a univariate, non-negative, non-decreasing function \( h \) it holds that \( h(T) = h(t_0) + \| h \|_{v_1} \), and then (50) and (51) imply that
\[
\| h \|_{[p]} \leq \| h \|_{[1]} \leq 2h(T) \leq 2\| h \|_{\infty} \leq 2\| h \|_{[p]}.
\]
(56)