Summer 2018

On modeling quantities for insurer solvency against catastrophe under some Markovian assumptions

Daniel Jefferson Geiger

Department: Mathematics and Statistics

ON MODELING QUANTITIES FOR INSURER SOLVENCY AGAINST CATASTROPHE UNDER SOME MARKOVIAN ASSUMPTIONS

by

DANIEL JEFFERSON GEIGER

A DISSERTATION

Presented to the Graduate Faculty of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

MATHEMATICS WITH STATISTICS EMPHASIS

2018

Approved by

Akim Adekpedjou, Advisor
Gregory M. Gelles
Ilene H. Morgan
Gayla R. Olbricht
V. A. Samaranayake

ABSTRACT

Insurance companies sometimes face catastrophic losses, yet they must remain solvent enough to meet the legal obligation of covering all claims. Catastrophes can result in large damages to the policyholders, causing the arrival of numerous claims to insurance companies at once. Furthermore, the severity of an event could impact the time until the next occurrence. An insurer needs certain levels of startup capital to meet all claims, and then must have adequate reserves on a continual basis, even more so when catastrophes occur. This work examines two facets of these matters: for an infinite time horizon, we extend and develop models for insurer bankruptcy-related quantities accounting for the reality of large claims occurring. Meanwhile, for finite time horizons, we model the present value of claims that have been incurred but not yet reported, so-called "IBNR" claims. In the former, we show how our method for "Gerber-Shiu" functions works in a recently proposed dependency structure allowing insurers to charge clients different premiums depending on their riskiness. In the latter, we build upon a recent method which allowed claims to arrive in batches; besides permitting discounting to be time-dependent, we allow the insurer to adjust the assumed distribution of the time until the next event by comparing the number of claims from the current event to any number of random intervals. We provide numerical studies for both scenarios.

Keywords: Solvency, Time value of ruin, Gerber-Shiu functions, Reserves, Incurred but not reported claims (IBNR), Heavy-tailed risks

ACKNOWLEDGMENTS

I dedicate this dissertation in actuarial science to Jesus of Nazareth, who failed once so that I might not fail on an infinite time horizon. My advisor, Dr. Akim Adekpedjou, has been an academic father-figure to me over the many years that I have worked with him; he is truly a wise man. I would like to thank my advisory committee for vetting my sometimes eccentric ideas, and Dr. Ilene Morgan in particular for all the times she provided advice on random (but perhaps not stochastic) matters. I thank the Missouri University of Science and Technology for the Chancellor's Fellowship, and the Department of Mathematics and Statistics for supporting me with a Graduate Assistantship. I thank Dr. David Grow for having been a helpful resource in the more technical matters of mathematical analysis regarding the first part of this dissertation.

Through my time as a graduate student at Missouri S&T, I have grown far beyond what I could have envisioned. Thank you again to those who have mentored me in diverse ways, helping me pursue my potential as a professional researcher. I have been cheered on by a familial network both near and far. In my immediate family, my parents have been such enthusiastic supporters. These six years have been a challenging time, and by no means always pleasant, but have been a good six years, through which I have become a more mature person, and for which I am forever grateful.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Table | Page

NOMENCLATURE

Symbol	Description
GSF	Gerber-Shiu function
CPTA	Corrected phase-type approximations
פ	The Hebrew letter “pe”
ק	The Hebrew letter “qof”
ב	The Hebrew letter “beth”
IBNR	Incurred but not reported
IR	Incurred and reported
rv	Random variable
df	Distribution function
iid	Independent and identically distributed
LT	Laplace transform
LST	Laplace-Stieltjes transform

1. INTRODUCTION

1.1. BACKGROUND, LITERATURE REVIEW, AND SCOPE OF DISSERTATION

Insurance companies are in the business of covering risks of their policy-holding clients, who expect to be financially compensated for their losses in return for having paid premiums. The insurers sometimes face catastrophic losses, yet they must remain solvent enough to meet the legal obligation of covering all claims. Some examples of events which intuitively cause possibly large losses include tornadoes, forest fires, hurricanes, typhoons, floods, earthquakes, etc. Epidemics, wars or civil upheaval could also produce larger losses; this dissertation concentrates on insurers' losses in general. If an insurer receives more claims than they have the ability to pay, they will become bankrupt, or experience "ruin" in technical terms. Besides fees from legal proceedings and interest on unpaid claims, "ruin" stands to harm the insurer's reputation, both with clients and otherwise potential clients.

The money which insurers must have ready at hand to pay every claim received is known as the companies' "reserves." Intuitively, someone who desires to open an insurance company must have some startup capital before reserves can even come into consideration. Without much imagination needed, one can see that the amount of startup capital affects a new insurance company's solvency even long-term. The probability (likelihood) of ultimate ruin has received much attention from actuarial researchers over the decades, and likewise for the time of ruin in more recent decades. Other ways exist to consider the "what-if" scenario of an insurer becoming bankrupt: for example, the deficit sustained by a ruined insurer has commonly been called the "severity" of ruin in the literature.

We want to develop models for estimating the necessary startup capital and subsequent reserves an insurer needs in order to remain solvent. Large, rare claims present extra challenges in modeling; we need a unified manner of modeling quantities relevant to insurer solvency which recognizes that catastrophes happen. But there is more: not all catastrophes qualitatively fall within the same level of cataclysmic fallout. Even just in the context of the USA, comparing tornadoes with several 2017 US hurricanes may make the point, let alone the potential for the New Madrid seismic region to have an earthquake. If we have not made the point clear enough, let us mention the Yellowstone supervolcano. Tornadoes, hurricanes, earthquakes and supervolcanoes as a simple "thought experiment" demonstrate the need for a multiplicity of qualitative levels of risk. Considering all the uncertainty associated with the severity of a natural disaster or health epidemic even within a single "risk class," the concept of comparing an event's severity to a random threshold or random intervals naturally arises; we make this idea precise later on.

In addition to startup capital, day-to-day reserves must be modeled adequately in a world with catastrophes. On an individual level, claims might not have been reported to an insurer, but the insurer must pay each and every valid claim all the same. After totaling their car, a policyholder might well be hospitalized, and hence unable to inform the insurer of the loss until sufficiently recovered. Or a home owner may be displaced after a hurricane; they might not be sure of the damage their home sustained until returning there. The greater the number of claims instigated by a catastrophic event, the more there are which may not come to the insurer's attention immediately. These claims are known as "incurred but not reported," or IBNR claims. Upon receiving notice of a claim, the insurer might not pay (settle on) the claim immediately. Perhaps a database of the insurer contains a discrepancy on the policyholder's address, for one reason or another. Or the insurer may have reasons to suspect fraud in a claim, in which case due diligence in verifying such a matter lies in the insurer's financial interests. Such situations, with a lag between the claim being reported and the time when the policyholder receives the compensation they expect, are known as instances of "reported but not settled" claims, or RBNS.

The insurer's risk process refers to the expression $U(t) = u + C(t) - \sum_{i=1}^{N(t)} X_i$, where $U(t)$ represent's the insurer's funds at time t. First, $U(0) = u$ is the insurer's aforementioned initial capital. The premium collection process $C(t)$ is the total premiums the insurer has received by time t; classically this is set to be $C(t) = ct$. $N(t)$ is the number of claims that have occurred by time t, and the X_i terms are the severity (amount) of each individual claim. This risk process $U(t)$ is the classical context for considering questions like the probability of ultimate ruin, the severity of (deficit at) ruin, or the surplus before ruin.

The problem of a unified framework for ruin-theoretic quantities was first addressed by Gerber and Shiu (1998), in which they introduced the "expected discounted penalty function," where a random penalty is due at the time of ruin. The penalty in that paper was a function of the insurer's surplus (funds) immediately before ruin and the deficit at ruin (severity of the insurer's bankruptcy), as well as possibly being discounted with respect to the time of ruin. These "EDPFs" quickly became known as "Gerber-Shiu" functions (GSFs). Whereas Gerber and Shiu (1998) introduced GSFs in the classical risk model (compound Poisson, also called "Cramér-Lundberg"), myriad extensions quickly emerged.

Some papers in the years following Gerber and Shiu (1998) included Ahn and Badescu (2007), Pitts and Politis (2007), Tang and Wei (2010), Li and Sendova (2015), and Chau *et al.* (2015). An early approach to the unit GSF in the classical risk model with heavy-tailed claim sizes appeared in Šiaulys and Asanavičiūtė (2006), who assumed claims had a subexponential distribution (to be defined below); that paper gave an asymptotic approximation. Ahn and Badescu (2007) observed that their phase-type setup could fit GSFs arbitrarily closely, but would in general be poor under heavy-tailed claims. An approach via functional analysis to the problem of approximating GSFs was given by Pitts and Politis (2007) also in the classical risk model, but did not handle heavy-tailed claims. In the renewal generalization of the classical risk model allowing interclaim times to have an arbitrary distribution, Tang and Wei (2010) comprehensively established the asymptotic behavior of GSFs with claims being either subexponential or convolution-equivalent; they commented that their formulas could perform rather poorly for smaller initial capital under subexponential claims. A more recent approach to approximating Gerber-Shiu functions, without modeling individual claims (or hence dependencies at that level), came from Chau *et al.* (2015). Their numerical examples did not include heavy-tailed distributions.

In Albrecher and Boxma (2004), the proposal was to compare each claim to a random threshold, upon which the Poisson arrival rate parameter (or, the distribution of the time until the next claim) could be adjusted. However, they only allowed one premium rate for both risk classes, besides only treating the survival probability. The authors of Li and Sendova (2015) generalized Albrecher and Boxma (2004) to Gerber-Shiu functions with distinct premium rates for each risk class. However, they left out how to handle heavy-tailed claims. The premium rate assumed only a single value in Albrecher and Boxma (2005), although the model for GSFs there subsumed that of Albrecher and Boxma (2004). A promising approach to feasibly modeling heavy-tailed claims appeared in Vatamidou *et al.* (2013), but they only addressed the ultimate ruin probability and the aggregate total losses in the classical risk model. Likewise, Vatamidou *et al.* (2014a) only addressed the equivalent of the ruin probability in a (more involved) queuing model.

Herein lies the starting point for our work in Section 2. We generalize the corrected phase-type approximations (CPTA) of Vatamidou *et al.* (2013) to arbitrary penalty functions $w(\cdot,\cdot)$ (nonnegative functions on $\mathbb{R}^+ \times \mathbb{R}^+$) and discount rate $\delta \geq 0$ in the risk model of Li and Sendova (2015), utilizing a couple parts of Vatamidou *et al.* (2014a) in the process. Synthesizing results found in Tang and Wei (2010), we show that CPTA of GSFs behave “as desired” (to be made precise later) for large capital in the classical risk model, much like Vatamidou *et al.* (2013, 2014a) did so for the ruin probability only in the classical risk model.

The starting point for our work in Section 3 is the paper by Landriault *et al.* (2017). As far as the distribution of aggregate claims is concerned, Léveillé and Garrido (2001a,b) assumed a constant discount rate for claims arriving according to a renewal process, both in finite time and in the asymptotic limit as time goes to infinity. Therefore, they assumed claims to arrive one at a time, disregarding catastrophic events causing multiple claims. Additionally, they assumed claims to be reported upon occurrence, meaning without any reporting lag, thus omitting IBNR setups. Likewise, the work by Léveillé and Adékambi (2011) allowing the discount rate to be stochastic also ignored reporting lags and multiple claims from single events, and they only considered up to second moments. The aforementioned work by Landriault *et al.* (2017) found the finite-time moments of IBNR claims and the joint moments of incurred and reported (IR) claims and IBNR claims possibly at a later time. They allowed multiple claims to be incurred at once, specifically mentioning this could handle catastrophes. While they thus addressed this catastrophe-related shortcoming of the earlier literature, they assumed the sizes of these batches of claims were independent of all other model quantities and of each other, thus continuing to neglect the possibility of the magnitude of one catastrophe affecting the time until the next potentially catastrophic event. Lastly, they continued to assume a constant discount rate.

Where we improve matters is in extending Landriault *et al.* (2017) to allow multiple choices of interevent time distributions and general deterministic time-discounting of claim severities. We use the semi-Markov dependency structure of Albrecher and Boxma (2004); Li and Sendova (2015) applied to the number of claims caused by an event ("batch sizes"). However, instead of just two classes (levels) of riskiness as in those papers, we allow any positive number of such classes. As a consequence, the random thresholds now become random intervals on the positive integers, to which we compare the claim batch sizes to see in which interval one such batch size falls. Then the interval into which the batch size of claims from the current event falls determines the distribution of the time until the next event. Thus we have Markov renewal processes (e.g. Janssen and Manca (2006)) where previously Landriault *et al.* (2017) had renewal processes. Furthermore, we show that some particular cases of time-varying discount rates can produce different results than the constant discount rate assumption.

1.2. SOME TECHNICAL CONCEPTS

Now we shall lay out some technical concepts which arise in both parts of our work. For limiting relationships in which a parameter tends towards $+\infty$, we follow some notation from Tang and Wei (2010). Let $\alpha(x)$ and $\beta(x)$ be two (eventually) positive functions, i.e. there exists x_0 such that $x > x_0$ implies $\alpha(x) > 0$, and likewise for $\beta(x)$. Denote c_* and c^* to be values for which $c_* \le \liminf_{x\to\infty} \frac{\alpha(x)}{\beta(x)} \le \limsup_{x\to\infty} \frac{\alpha(x)}{\beta(x)} \le c^*$. Then, $\alpha(x) = O(\beta(x))$ and $\alpha(x) = o(\beta(x))$ respectively mean $c^* < \infty$ and $c^* = 0$. The relation $\alpha(x) \asymp \beta(x)$ means $0 < c_* \le c^* < \infty$; Klüppelberg (1989) called this "weak asymptotic equivalence." Further, $\alpha(x) \lesssim \beta(x)$ and $\alpha(x) \gtrsim \beta(x)$ respectively mean $c^* = 1$ and $c_* = 1$. When $c^* = c_* = 1$, we write the usual $\alpha(x) \sim \beta(x)$.

In this work, we mean Lebesgue-Stieltjes integration by

$$\int_A f(x)\, \mathrm{d}\alpha(x) \equiv \int_A f(x)\alpha(\mathrm{d}x),$$

for A a subset of the domain of $\alpha(x)$. If $\alpha(x) = x$ is the identity function, then we mean Lebesgue integration by $\int_A f(x)\, \mathrm{d}x$. We understand $\int_a^b$ to mean integrating over the set $(a, b]$, and $\int_{a-}^b$ to mean over the set $[a, b]$: that is, in the latter we include any atom at a, while in the former we omit any such atom. If $b = \infty$, we mean the usual limiting sense: $\lim_{R\to\infty} \int_a^R$. As in Tang and Wei (2010), we call a nonnegative function $f(x)$ on $\mathbb{R}^+$ locally integrable if $\int_0^{x_0} f(x)\, \mathrm{d}x$ is finite for all $x_0 > 0$, and also globally integrable if $\int_0^\infty f(x)\, \mathrm{d}x$ is finite. The following integral transform arises throughout our work in both sections.

Definition 1 (Laplace-Stieltjes Transform, Widder (1941))**.** *Let $\alpha(t)$ be a real-valued function of bounded variation in $t \in [0, R]$ for all $R > 0$. Let also s be a complex variable. The Laplace-Stieltjes transform of $\alpha(t)$ evaluated at s is given by $\tilde{\alpha}(s) = \int_{0-}^\infty e^{-st}\, \mathrm{d}\alpha(t) = \lim_{R\to\infty} \int_{0-}^R e^{-st}\, \mathrm{d}\alpha(t)$, as long as the limit exists.*

Widder (1941) defined the LST with $\alpha(t)$ being a complex-valued function of the real variable t; letting $\alpha(t)$ be real-valued suffices for our purposes. When $\alpha(t)$ is absolutely continuous (possesses a derivative for $t \geq 0$), if $\beta(t) = \alpha^{(1)}(t)$, we call $\hat{\beta}(s) = \int_0^\infty e^{-st}\beta(t)\,\mathrm{d}t$ the Laplace transform (LT) of $\beta(t)$ evaluated at $s \in \mathbb{C}$, (given convergence at s of course). The immediately following comments about convolutions and Laplace transforms may be found in detail in Widder (1941): the convolution of two integrable functions $\alpha(x)$ and $\beta(x)$ on $\mathbb{R}^+$ is $\gamma(x) = \int_0^x \alpha(x-y)\beta(y)\,\mathrm{d}y = \int_0^x \beta(x-y)\alpha(y)\,\mathrm{d}y$. If $\alpha(x)$ and $\beta(x)$ are locally integrable, then $\hat{\gamma}(s) = \hat{\alpha}(s)\hat{\beta}(s)$ given existence of the two Laplace transforms on the right-hand side; this is (Widder, 1941, Theorem 2.12.1a). This "product theorem" generalizes to Stieltjes convolutions ($\int_{0-}^x \alpha(x-y)\,\mathrm{d}\beta(y)$) when suitable regularity conditions are placed on $\alpha(x)$ and $\beta(x)$; see (Widder, 1941, Section 2.11) for details. Besides the Laplace and Laplace-Stieltjes transforms, the Dickson-Hipp transform (or operator) arises frequently in Section 2.

Definition 2 (Dickson-Hipp Transform, Dickson and Hipp (2001); Li and Garrido (2004))**.** *Let $f(t)$ be a real-valued function and integrable. Let $s \in \mathbb{C}$ have $\Re(s) \geq 0$. Then, for $t \geq 0$, $\mathrm{T}_s f(t) = \int_t^\infty e^{-s(x-t)} f(x)\,\mathrm{d}x$ is the Dickson-Hipp transform of $f(t)$.*

Notice that $\mathrm{T}_r f(0) = \hat{f}(r)$, such that the translation operator generalizes the Laplace transform. Along with this, some other properties given in Li and Garrido (2004) will be used in our work:

$$\mathrm{T}_{s_1}\mathrm{T}_{s_2} f(t) = \frac{\mathrm{T}_{s_1} f(t) - \mathrm{T}_{s_2} f(t)}{s_2 - s_1}, s_1 \neq s_2 \in \mathbb{C}, t \geq 0;$$
$$\mathrm{T}_{s_1}^n f(t) = \frac{(-1)^{n-1}}{(n-1)!}\frac{\partial^{n-1}}{\partial s_1^{n-1}}\mathrm{T}_{s_1} f(t), t \geq 0.$$

The Dickson-Hipp transform also appears frequently in our work on Gerber-Shiu functions, where its usage is as standard as that of the Laplace transform.

Because of its fundamental importance in the derivation of corrected phase-type approximations of Gerber-Shiu functions with time discounting or multiple Lundberg roots, we quote for completeness (Vatamidou *et al.*, 2014a, Theorem A.3) (from ArXiv:1405.4853, licensed under CC-BY-NC-SA 3.0):

Lemma 1. *Let r be a simple root of an analytic function $f(s)$. For some function $h(s,\epsilon)$ and for all small real values ϵ, we define the perturbed function*

$$F(s,\epsilon) = f(s) + h(s,\epsilon).$$

If $h(s,\epsilon)$ is analytic in s and ϵ near $(r,0)$, then $F(s,\epsilon)$ has a unique simple root $(x(\epsilon),\epsilon)$ near $(r,0)$ for all small values of ϵ. Moreover, $x(\epsilon)$ is an analytic function in ϵ, and if $\frac{\partial^n}{\partial s^n} h(s,0) \equiv 0$, $n = 0,1,\ldots$, then it holds

$$x(\epsilon) = r - \epsilon \frac{\frac{\partial}{\partial \epsilon} h(r,0)}{f^{(1)}(r)} + O\left(\epsilon^2\right).$$

In our demonstration of how corrected phase-type approximations correctly capture the tail behavior of the exact value for general $w(\cdot,\cdot)$ and $\delta \geq 0$ in the compound Poisson risk model, we use some properties which may be found, for example, in Tang and Wei (2010):

Definition 3 (The density classes $\mathscr{L}_d(\alpha)$ and $\mathscr{S}_d(\alpha)$)**.** *A function $f : [0,\infty) \to [0,\infty)$, measurable and eventually positive, belongs to $\mathscr{L}_d(\alpha)$, $\alpha \geq 0$, if $\lim_{x\to\infty} \frac{f(x-y)}{f(x)} = e^{\alpha y}$, $\forall y \in \mathbb{R}$. If also $\lim_{x\to\infty} \frac{f^{*2}(x)}{f(x)} = 2\hat{f}(-\alpha)$, then $f \in \mathscr{S}_d(\alpha)$.*

We make such frequent usage of (Tang and Wei, 2010, Lemma 4.3 (1)) that we regard it worth recalling; see Tang and Wei (2010) for the original statement.

Lemma 2. *Let f_1 and f_2 be locally integrable functions from $[0,\infty)$ to $[0,\infty)$. Suppose there exist some $\gamma \geq 0$ and some $\tilde{\gamma} > \gamma$ such that $f_1 \in \mathscr{L}_d(\gamma)$ and $f_2(x) = O\left(e^{-\tilde{\gamma} x}\right)$. Then as $x \to \infty$, it is the case that $f_1 * f_2(x) \sim f_1(x)\hat{f}_2(-\gamma)$.*

Concerning phase-type distributions, we give the following definition and basic properties; see for example Asmussen and Albrecher (2010) or Bladt and Nielsen (2017). Throughout this dissertation, an underscore signifies a vector or matrix; e.g., $\underline{\aleph}$.

Definition 4 (Phase-type distributions)**.** *Let $\underline{T}$ be a subintensity matrix of finite dimensions $p \times p$, and $\underline{\alpha}$ a $1 \times p$ vector of probabilities such that $\underline{\alpha}\,\underline{e} \leq 1$, where $\underline{e}$ is a $p \times 1$ vector with each entry* 1*. Then we call $F(x) = 1 - \underline{\alpha} e^{\underline{T}x} \underline{e}$ a phase-type distribution, which has density $f(x) = \underline{\alpha} e^{\underline{T}x} \underline{t}$ on $(0, \infty)$ and possibly an atom at* 0*. Equivalently, $X \sim PH(\underline{\alpha}, \underline{T})$ is the time until absorption of a Markov jump process with $p < \infty$ transient states (with initial probability vector $\underline{\alpha}$) and one absorbing state, where the intensity matrix of this process has the form $\begin{pmatrix} \underline{T} & \underline{t} \\ \underline{0} & 0 \end{pmatrix}$, with $\underline{t} = -\underline{T}\,\underline{e}$.*

Some important properties which may be found in the aforementioned references are that phase-type distributions are closed under convolution, are light-tailed, and have rational LT $\underline{\alpha}(s\underline{I} - \underline{T})^{-1}\underline{t}$, where $\underline{I}$ is the identity matrix.

The "Bell polynomial" will arise in the material pertaining to IBNR claims. In Johnson (2002), this polynomial is given as

$$B_{m,k}(x_1, x_2, \ldots, x_{m-k+1}) = \frac{1}{k!} \sum_{\substack{j_1+\cdots+j_k=m \\ j_i \geq 1}} \binom{m}{j_1, \ldots, j_k} x_{j_1} \cdots x_{j_k},$$

or, equivalently, in Landriault *et al.* (2017) as

$$B_{m,k}(x_1, x_2, \ldots, x_{m-k+1}) = \sum_{\substack{\sum_i j_i = k, \sum_i i j_i = m \\ i=1,\ldots,m-k+1}} \frac{m!}{j_1! j_2! \cdots j_{m-k+1}!} \left(\frac{x_1}{1!}\right)^{j_1} \left(\frac{x_2}{2!}\right)^{j_2} \cdots \left(\frac{x_{m-k+1}}{(m-k+1)!}\right)^{j_{m-k+1}}.$$

Lastly, we recall some notions from renewal theory which arise in our work. One reference is that of Janssen and Manca (2006) (plenty of other references exist on renewal theory). The basic concepts are that of a "renewal process" and of a "renewal equation."

Definition 5 (Renewal process)**.** *Let* $\tau_k \overset{\text{iid}}{\sim} F(\cdot)$ *for* $k \in \mathbb{N}^+ \equiv \{1, 2, 3, \ldots\}$. *Assume* $\tau_0 \overset{\text{as}}{=} 0$. *Then* $T_n = \sum_{k=1}^{n} \tau_k$ *is a "renewal process."*

A "renewal equation" is one of the form

$$m(u) = \phi \int_0^u m(u-x)k(x)\,\mathrm{d}x + v(u),$$

for suitably constrained nonnegative functions $k(\cdot)$ and $v(\cdot)$. If $\phi \in (0, 1)$, the renewal equation is called "defective," and if $\phi = 1$, called "proper"; the "excessive" case of $\phi > 1$ is outside of our area of concern.

2. CORRECTED PHASE-TYPE APPROXIMATIONS OF GERBER-SHIU FUNCTIONS IN A HEAVY-TAILED RISK MODEL WITH BOTH INTERCLAIM TIMES AND PREMIUMS DEPENDING ON CLAIM SIZES

2.1. OVERVIEW OF SECTION

We model heavy-tailed Gerber-Shiu functions by making claims be a mixture of phase-type and heavy-tailed components, weighted more heavily towards the former. We do so in a recently introduced risk model where both interclaim times and the premiums collected depend on the claim sizes. First, we find the Lundberg roots of the full mixture model as perturbation of those in the phase-type base models. From there, we proceed to find the approximations for general penalty functions in the dependent risk model, then simplifying these to the compound Poisson risk model. The first term of our approximations is the Gerber-Shiu function with the phase-type claims, and the "correction" term (multiplied by ϵ) contains the heavy-tailed component at most once per summand. Calling our expressions "corrected phase-type approximations" like the extant literature, we generalize these from the ultimate ruin probability in the classical risk model to Gerber-Shiu functions in the aforementioned dependent risk model. Without being asymptotic expressions themselves, our corrected phase-type approximations continue to capture the heavy-tailed behavior of the true value, which we make specific in the classical risk model. We numerically study the approximations' relative errors for some specific penalty functions and claims distributions, and finally give an application. This section is an expanded form of Geiger and Adekpedjou (2018).

2.2. GENERAL DISCUSSION

In the famous paper by Gerber and Shiu (1998), Hans Gerber and Elias Shiu introduced the functions which now bear their names as a framework for modeling insurer ruin-related quantities. However, throughout their paper, they implicitly assumed the existence of a negative root of what they called the "Lundberg equation" ($l(s) = 0$, to be introduced formally in Section 2.3); that root fails to exist for heavy-tailed claims distributions. Even for the special case of ultimate ruin probabilities in the classical risk model, the computationally tractable phase-type distributions are known to handle heavy-tailed behavior badly; see Vatamidou *et al.* (2014b) for an extensive discussion of that. For heavy-tailed claims, one may find the asymptotic tail behavior of Gerber-Shiu functions in Tang and Wei (2010), where renewal risk models were considered. The paper by Vatamidou *et al.* (2013) proposed a non-asymptotic method of approximation for ruin probabilities which properly captures heavy-tailed behavior, retains computational tractability, and has quantifiable error. They hinted at their method being broadly applicable to risk theory; in Vatamidou *et al.* (2014a) they applied their approach to a more complicated queuing model.

In this section we take that comment in a rather different direction, namely approximating Gerber-Shiu functions with heavy-tailed claims. Furthermore, we do so in a more general setting than the classical risk model Gerber and Shiu (1998) considered: the dependent risk model introduced in Li and Sendova (2015) allowing multiple classes of insureds with differing premium rates. Letting $J(t)$ be the class of the insured at time t, they proposed a dependent risk model, where

$$U(t) = u + \sum_{i=1}^{2} c_i \int_0^t \mathrm{I}\,(J(s) = i)\,\mathrm{d}s - \sum_{j=1}^{N(t)} X_j \tag{2.1}$$

is the insurer's capital on hand at time t, X_j are the iid claim sizes, and $N(t)$ is the number of claims in the interval $[0, t]$. With $Q_j \overset{\text{iid}}{\sim} H(y)$, the interclaim times are specified by

$$W_{j+1}\Big|\Big(X_j > Q_j\Big) \sim \text{Exp}(\lambda_1),$$
$$W_{j+1}\Big|\Big(X_j < Q_j\Big) \sim \text{Exp}(\lambda_2).$$

If $i \in \{1, 2\}$ is the initial class of the insured, $T_i = \inf\{t \geq 0 \mid U(t) < 0\}$ denotes the time of ruin for that class. We consider an unspecified penalty function $w(x, y) \geq 0, x \geq 0, y \geq 0$ with the discount rate $\delta \geq 0$. That is, we denote the Gerber-Shiu function analyzed in Li and Sendova (2015) by

$$m_i(u) = \text{E}\Big(e^{-\delta T_i} w(U(T_i-), |U(T_i)|)\, \text{I}(T_i < \infty) \mid U(0) = u\Big), u \geq 0, i \in \{1, 2\}.$$

Later on, we also consider the compound Poisson risk model with constant premium rate of 1, mostly following the notation used by Gerber and Shiu (1998).

The key idea presented in Vatamidou *et al.* (2013) was as follows: to model the claims distribution as a mixture of phase-type and heavy-tailed components (respectively $B(\cdot)$ and $C(\cdot)$), namely $P_\epsilon(x) = (1-\epsilon)B(x) + \epsilon C(x)$, and to view $\epsilon \in [0, 1)$ as a perturbation parameter. This "mixture model" then had two associated phase-type base models: the "discard" and the "replace." The former had as claims law $P_\epsilon^\bullet(x) = (1-\epsilon)B(x) + \epsilon$, while the latter had $P_0(x) = B(x)$. What Vatamidou *et al.* (2013) called collectively "corrected phase-type approximations" of the ruin probability consisted of the "phase-type approximations" thereof (with claims laws $P_\epsilon^\bullet(x)$ and $P_0(x)$), and a "correction term" containing the heavy-tailed component only once.

Whereas we propose approximating heavy-tailed Gerber-Shiu functions by a substantial generalization of Vatamidou *et al.* (2013), some other approaches were recently proposed in the literature. First, while it only tackles approximating ruin probabilities like Vatamidou *et al.* (2013), a method of a somewhat similar spirit may be found in Peralta *et al.* (2018); Rojas-Nandayapa and Xie (2017), and references therein. As for Gerber-Shiu functions, one early, functional analytic approach appeared in Pitts and Politis (2007); we note that their illustrative example used (non-Erlang) gamma claim sizes. The article by Chau *et al.* (2015) applied Fourier-cosine series expansion techniques to approximate Gerber-Shiu functions for a risk process with claims modeled in the aggregate by a Lévy subordinator. Their approach provided approximations of linear complexity, seemingly comparable to the computational complexity in Vatamidou *et al.* (2013). We were not convinced, however, that their Fourier-cosine method would capture heavy-tailed behavior of individual claims in compound laws; indeed, none of their numerical examples came from the subexponential density class. For an aggregate approach to heavy-tailedness in particular, the recent paper by Kolkovska and Martín-González (2016) derived Gerber-Shiu functions for the compound Poisson risk model with an α-stable motion as a perturbation term. After building up several propositions about the scale functions of their considered risk process, those authors presented a form for the corresponding Gerber-Shiu function as an infinite series of convolutions. Finally, for several types of heavy-tailed individual claims distributions (i.e. in the compound Poisson part of the perturbed process), they provided asymptotic formulas for the joint tail distribution of the deficit at and surplus prior to ruin; we note their Theorem 2 was in the non-discounted $\delta = 0$ case.

Like the authors in Vatamidou *et al.* (2013, 2014a), we seek results which provide benefits not only in the asymptote, but for all initial capital u; they showed this is possible for infinite-time ruin probabilities, finite-time aggregate losses, and waiting times in queues, tractably capturing heavy-tailed behavior directly in their approximations of these. The asymptotes in general were as initial capital $u \to \infty$; this is as in Tang and Wei (2010), where the authors asserted that good asymptotic formulas somewhat alleviate needing very large initial capital amounts. However, Vatamidou *et al.* (2013) went beyond this, demonstrating the existence of a bound on the relative error for arbitrary initial surplus, when using the corrected discard approximation to the probability of ultimate ruin. They conjectured this because they demonstrated that the corrected discard approximation always underestimates the true ultimate ruin probability in the compound Poisson risk model. At least for specific penalty function choices (such as $w(\cdot,\cdot) = 1$ with $\delta \geq 0$), it might be possible in the classical risk model to find such a bound for all $u \in [a(\delta), \infty)$, where seemingly $a(0) = 0$ and $a(\delta) > 0$ for $\delta > 0$. However, desiring to promote our generalization of Vatamidou *et al.* (2013) for more than just the classical risk model, we have not pursued establishing our above-suggested generalization of (Vatamidou *et al.*, 2013, Theorem 8). Specifically, for some parameter choices in the Li-Sendova risk model, the corrected discard approximation of the unit GSF with $\delta > 0$ can overestimate the exact value evidently for all large u.

We generalize the other key results of Vatamidou *et al.* (2013) to Gerber-Shiu functions in the risk model (2.1) of Li and Sendova (2015). We follow the basic idea used in Vatamidou *et al.* (2013) of modeling individual claims by $P_\epsilon(x) = (1-\epsilon)B(x) + \epsilon C(x)$. Now, in time-discounted quantities (which they left untouched), the Lundberg root depends on the claims law; this must be accounted for in perturbation expansions of Gerber-Shiu functions. It turns out that the correct way to handle the Lundberg roots depending on the claims law is by perturbation expansions in terms of the "base model" Lundberg root, the perturbation parameter, and some "pieces" of Gerber-Shiu functions. Using a theorem from Vatamidou *et al.* (2014a), we derive in Lemma 3 below the Lundberg roots for the Li-Sendova risk model in perturbed form, for all non-negative δ; we then proceed to use these in finding perturbation expansions for $m_i(u)$ with general penalties $w(\cdot,\cdot)$. These provide the basis for our corrected phase-type approximations, given after Theorem 1.

To get the analogues in the compound Poisson risk model, we let $c_1 = c_2 = 1$ and $\lambda_1 = \lambda_2 = \lambda$, and provide in Corollary 1 the perturbation expansions for Gerber-Shiu functions in this special case of (2.1). We note that, for either the Li-Sendova risk model, or the classical compound Poisson special case, our approximations may be made precise in the limit $\epsilon \to 0$ following the proof of (Vatamidou *et al.*, 2014a, Proposition 3.8), albeit with an extra step for the corrected discard approximation. Due to the method of handling the mixture-model Lundberg roots (where extending (Vatamidou *et al.*, 2014a, Theorem A.3) to find a complete series expansion in the perturbation parameter would require evaluating subsequent derivatives of Laplace transforms of heavy-tailed distributions at a base-model Lundberg root), we expect this limiting form of convergence to propagate through any Gerber-Shiu functions when Lundberg root perturbation occurs.

We establish in the compound Poisson risk model that, under mild regularity conditions, $m_{d,\epsilon}(u)$ and $m_{r,\epsilon}(u)$ capture the tail behavior of $m_{\epsilon}(u)$ up to constant scalars, basing this upon results in Tang and Wei (2010). Namely, their Corollary 3.2 and Lemma 4.3 are fundamental to establishing our Propositions 1, 2, and 3, analogously to the role basic subexponential properties in Asmussen and Albrecher (2010) played in the proofs of (Vatamidou *et al.*, 2013, Theorems 5–7). An initial venture into finding the asymptotic tail behavior of Gerber-Shiu functions in the Li-Sendova risk model indicated the basic process is largely analogous to the ideas in Tang and Wei (2010), just noticing that (for example) with threshold df $H(y) = 1 - e^{-\nu y}$, $\xi_2(y) = \bar{H}(y)p(y)$ is in the density class $\mathscr{S}_d(\nu)$ provided $p(y) \in \mathscr{S}_d(0)$. Vatamidou *et al.* (2014a) found corrected phase-type approximations for a queuing model in a Markovian Arrival Process (MArP), a more complicated setup than the compound Poisson risk model. In that paper, they briefly noted that they showed in Vatamidou *et al.* (2013) that a single (linear in convolution) appearance of the heavy-tailed component captures the correct tail behavior, up to a constant, of the true value; however, they did not formally show this for the MArP. Likewise, we explicitly show our generalizations for Gerber-Shiu functions to capture the heavy-tailed behavior in the compound Poisson risk model only. See also Asmussen and Albrecher (2010) for comments on expecting heavy-tailed dependent risk models generally to behave similarly in the asymptote in a manner to reduce the impact of dependency. We observe now that we do not explicitly consider finite time horizons in this section, unlike Vatamidou *et al.* (2013), but as Gerber and Shiu (1998) noticed, proper choices of the penalty function can retrieve various finite-time quantities.

We organize the section as follows: first we give some preliminary notation in Section 2.3. Next, in Section 2.4 we provide the perturbation expansions of some claims law-dependent quantities, followed by those of $\hat{m}_{\epsilon,i}(s)$ as the basis for $m^i_{d,\epsilon}(u)$ and $m^i_{r,\epsilon}(u)$, and likewise in the classical risk model. Simplifying to said model in Section 2.5, we show that the corrected phase-type approximations of Gerber-Shiu functions capture the exact value's heavy-tailed behavior, up to multiplication by logical scalar constants; we comment on how the correction terms impact the error of the phase-type approximations. We return to the dependent risk model in Section 2.6 for a numerical illustration, and we close the section with an application in the classical risk model in Section 2.7.

2.3. NOTATION

With $\aleph$ denoting a generic entity, the meanings of $\aleph_\epsilon$, $\aleph^\bullet_\epsilon$, and $\aleph_0$ are the same as in Vatamidou *et al.* (2013). That is, for an entity in the full mixture model (which depends on ϵ), we write $\aleph_\epsilon$. For the "discard" base case, we write $\aleph^\bullet_\epsilon$; for the "replace" base case, we write $\aleph_0 \equiv \aleph_{\epsilon|\epsilon=0}$, as the "replace" case corresponds to setting $\epsilon = 0$ in the full model (Vatamidou *et al.* (2013)). If $F(x) = \Pr(\aleph \leq x)$, then we write $\bar{F}(x) = \Pr(\aleph > x)$. We denote the Laplace-Stieltjes transform of $\aleph$ by $\tilde{F}(s)$, and the Laplace transform of the density of $\aleph$, $f(x) = F^{(1)}(x)$, by $\hat{f}(s)$. Like Vatamidou *et al.* (2013), we assume claims have the "mixture-model" distribution $\Pr(X_\epsilon \leq x) = (1-\epsilon)\Pr(B \leq x) + \epsilon \Pr(C \leq x)$. We assume that the phase-type (see Bladt and Nielsen (2017) for a recent overview) generic random variable (rv) B has an absolutely continuous density $b(x)$, and likewise for the heavy-tailed generic rv C, for which $\eta_c = \mathrm{E}(C) < \infty$. Furthermore, we assume that $B(x) = C(x) = 0$ for $x \leq 0$ and that $\lim_{x\to\infty} B(x) = \lim_{x\to\infty} C(x) = 1$. We use $\beth$ to denote any of the claims laws related to the full mixture model.

When we consider the classical risk model, we generally follow the notational style used by Gerber and Shiu (1998), and we set the constant premium income rate to 1. We denote the Gerber-Shiu function introduced in Gerber and Shiu (1998) by $m_\epsilon(u)$.

When the penalty $w(\cdot,\cdot) \equiv 1$, we write $\phi_\epsilon(u)$ instead, and if the discount rate $\delta = 0$ as well, we write $\psi_\epsilon(u)$. These special cases of $m_\epsilon(u)$ will appear frequently in Section 2.5. We assume the usual condition $1 > \lambda\eta_\epsilon = (1-\epsilon)\lambda\eta_b + \epsilon\lambda\eta_c$ for the mixture model, from which $1 > \lambda\eta_\epsilon^\bullet \equiv (1-\epsilon)\lambda\eta_b$ follows immediately, and $1 > \lambda\eta_0$ follows if we assume $\eta_c \geq \eta_b$. As in Tang and Wei (2010), we let $\bar{\Omega}_{\text{פ}}(u) = \int_u^\infty \omega_{\text{פ}}(x)\,\mathrm{d}x$, where $\omega_{\text{פ}}(u) = \int_u^\infty w(u, x-u)\mathrm{d}P_{\text{פ}}(x)$ as in Gerber and Shiu (1998).

For the dependent risk model of Li and Sendova (2015), the subscript $i \in \{1,2\}$ denotes the initial class of insured. We assume that $H(y)$, the proper distribution function of the random thresholds Q, has $H(y) = 0$ for $y \leq 0$. We define, as in Albrecher and Boxma (2004):

$$\chi_{\text{פ},1}(s) = \mathrm{E}\left(e^{-sX_{\text{פ}}}\,\mathrm{I}\left(X_{\text{פ}} > Q\right)\right) = \int_{y\in[0,\infty)} e^{-sy}H(y)\,\mathrm{d}P_{\text{פ}}(y),$$
$$\chi_{\text{פ},2}(s) = \mathrm{E}\left(e^{-sX_{\text{פ}}}\,\mathrm{I}\left(X_{\text{פ}} \leq Q\right)\right) = \int_{y\in[0,\infty)} e^{-sy}\bar{H}(y)\,\mathrm{d}P_{\text{פ}}(y).$$

We set $\xi_{\text{פ},1}(y) = H(y)p_{\text{פ}}(y)$ and $\xi_{\text{פ},2}(y) = \bar{H}(y)p_{\text{פ}}(y)$; note that $\chi_{\text{פ},2}(s) = \hat{\xi}_{\text{פ},2}(s) + P_{\text{פ}}(0)$. We use the shorthand notation ${}_\circ\chi_{\text{פ},i}(s) = \chi_{\text{פ},i}(s) - \mathrm{I}(i=2)$. The "Lundberg functions" $l_{\text{פ}}(s) = \hat{v}(s) - \hat{\mu}_{\text{פ}}(s)$ have roots $r_{\text{פ}}, \rho_{\text{פ}} \geq 0$ (see (Li and Sendova, 2015, Lemmas 3.1, 3.2)); here $\hat{v}(s) = \left(s - \frac{\delta+\lambda_1}{c_1}\right)\left(s - \frac{\delta+\lambda_2}{c_2}\right)$, and $\hat{\mu}_{\text{פ}}(s) = \frac{\lambda_1}{c_1}\chi_{\text{פ},1}(s)\left(\frac{\delta+\lambda_2}{c_2} - s\right) + \frac{\lambda_2}{c_2}\chi_{\text{פ},2}(s)\left(\frac{\delta+\lambda_1}{c_1} - s\right)$.

With the notation $R_{\text{פ}} = \sum_{i=1}^2 \frac{c_i}{\lambda_i}\chi_{\text{פ},i}(0)$ and ${}_\circ R_{\text{פ}} = \sum_{i=1}^2 \frac{c_i}{\lambda_i}{}_\circ\chi_{\text{פ},i}(0)$, we assume the positive security loading condition $R_\epsilon > \eta_\epsilon$ of Li and Sendova (2015). With the assumption $R_b - \eta_b \geq R_c - \eta_c$, we also get $R_0 > \eta_0$; ${}_\circ R_c < \eta_c$ implies $R_\epsilon^\bullet > \eta_\epsilon^\bullet$. The intuition of $R_b - \eta_b \geq R_c - \eta_c$ is that, to the contrary, $R_\epsilon > \eta_\epsilon$ would imply $\epsilon < 0$; ${}_\circ R_c < \eta_c$ is just $\eta_c > 0$ in the classical risk model.

From Li and Garrido (2004), we use some properties of the translation operator T (introduced by Dickson and Hipp (2001)):

$$\mathrm{T}_r f(x) = \int_x^\infty e^{-r(y-x)} f(y)\,\mathrm{d}y,$$

where r is a possibly complex number and has non-negative real part. We denote the compound geometric rv associated with the solution of defective renewal equations (see Lin and Willmot (1999)) by $G_{פ}$ for $\delta > 0$ and by $M_{פ}$ for $\delta = 0$.

2.4. CORRECTED PHASE-TYPE APPROXIMATIONS OF GERBER-SHIU FUNCTIONS

We now derive corrected phase-type approximations of Gerber-Shiu functions. The first step is to find perturbation expansions for the Lundberg roots; using these expansions, we obtain perturbation expansions for another term which depends on the claims law. Finally, we derive CPTA in the Li-Sendova dependent risk model, afterwards reducing these CPTA to the simpler compound Poisson risk model.

2.4.1. Lundberg Root Perturbation Expansions. For the existence, uniqueness and distinctness of the Lundberg roots, see (Li and Sendova, 2015, Lemmas 3.1, 3.2). As observed before, the Lundberg roots depend on the claims distribution; therefore, we need to express the mixture-model Lundberg roots as perturbation expansions in ϵ of the base-model Lundberg roots. The following simple lemma is crucial to deriving $m^i_{d,\epsilon}(u)$ and $m^i_{r,\epsilon}(u)$; it turns out that the correction terms of these contain the "coefficients" of ϵ.

Lemma 3. *Let $\epsilon > 0$ and $\delta \geq 0$. If r_ϵ is either Lundberg root of the Li-Sendova dependent risk model with claims law $P_\epsilon(x)$, it follows that* (i) $r_\epsilon = r^\bullet_\epsilon + \epsilon \frac{{}_\circ\hat{\mu}_c(r^\bullet_\epsilon)}{l^{\bullet(1)}_\epsilon(r^\bullet_\epsilon)} + O\left(\epsilon^2\right)$*; and* (ii) $r_\epsilon = r_0 + \epsilon \frac{{}_\circ\hat{\mu}_c(r_0) - {}_\circ\hat{\mu}_b(r_0)}{l^{(1)}_0(r_0)} + O\left(\epsilon^2\right)$.

Proof of Lemma 3. We let $r_{פ} > \rho_{פ} \geq 0$ without loss of generality (wlog). We first note that (i) and (ii) trivially hold by construction for ρ_ϵ when $\delta = 0$. For, by (Li and Sendova, 2015, Lemma 3.1), $\rho_\epsilon = \rho^\bullet_\epsilon = \rho_0 = 0$ also; because we have assumed $\tilde{C}(0) = \tilde{B}(0) = 1$, it follows that ${}_\circ\hat{\mu}_c(0) = {}_\circ\hat{\mu}_b(0) = 0$. Now we handle the case of r_ϵ with $\delta \geq 0$ (or ρ_ϵ with $\delta > 0$). The expansion (ii) follows by applying (Vatamidou *et al.*, 2014a, Theorem A.3), the conditions of which are easily verified, to the mixture-model Lundberg function $l_\epsilon(s) = l_0(s) - \epsilon({}_\circ\hat{\mu}_c(s) - {}_\circ\hat{\mu}_b(s))$. As for (i), by (Vatamidou *et al.*, 2014a, Theorem A.3),

r_ϵ and ρ_ϵ are analytic in ϵ (the same being true for $r_\epsilon^\bullet$ and $\rho_\epsilon^\bullet$); this allows us to say, in what follows, that $\epsilon(r_\epsilon - r_\epsilon^\bullet)_\circ\hat{\mu}_c^{(1)}(r_\epsilon^\bullet) = O\left(\epsilon^2\right)$, for by applying that theorem to $l_\epsilon^\bullet(s) = l_0(s) + \epsilon_\circ\hat{\mu}_b(s)$, we see that $r_\epsilon - r_\epsilon^\bullet = O(\epsilon)$. Next, recalling that $l_\epsilon(s) = l_\epsilon^\bullet(s) - \epsilon_\circ\hat{\mu}_c(s)$, we use Taylor's expansion of $l_\epsilon(r_\epsilon)$ about $r_\epsilon^\bullet$: $0 = l_\epsilon(r_\epsilon^\bullet) + (r_\epsilon - r_\epsilon^\bullet)l_\epsilon^{(1)}(r_\epsilon^\bullet) + O\left(\epsilon^2\right) = -\epsilon_\circ\hat{\mu}_c(r_\epsilon^\bullet) + (r_\epsilon - r_\epsilon^\bullet)l_\epsilon^{\bullet(1)}(r_\epsilon^\bullet) + O\left(\epsilon^2\right)$. As noted in the proof of (Vatamidou *et al.*, 2014a, Theorem A.3), $l_\epsilon^{\bullet(1)}(r_\epsilon^\bullet) \neq 0$ because $r_\epsilon^\bullet$ is a simple root of $l_\epsilon^\bullet(s)$; hence, (i) follows.

Now we give the details of applying (Vatamidou *et al.*, 2014a, Theorem A.3) to establish (ii) with $\delta \geq 0$: $-\epsilon({}_\circ\hat{\mu}_c(s) - {}_\circ\hat{\mu}_b(s))$ is analytic by $(s, \epsilon) = (r_0, 0)$, since wlog $r_0 > 0$, and the Laplace transforms in $-\epsilon({}_\circ\hat{\mu}_c(s) - {}_\circ\hat{\mu}_b(s))$ are well-known to be analytic in s for $\Re(s) > 0$. Thus, by (Vatamidou *et al.*, 2014a, Theorem A.3), $(r_\epsilon(\epsilon), \epsilon) := r_\epsilon$ is a unique simple root of $l_\epsilon(s)$ near $(r_\epsilon, 0)$ for all small ϵ, and is analytic in ϵ. Obviously $-\epsilon({}_\circ\hat{\mu}_c(s) - {}_\circ\hat{\mu}_b(s))|_{\epsilon=0} \equiv 0$ satisfies the condition of (Vatamidou *et al.*, 2014a, Theorem A.3) for expressing r_ϵ as an expansion of ϵ, so furthermore, (ii) holds.

Lastly, we elaborate how to establish (i) (showing (ii) works the same) for ρ_ϵ when $\delta = 0$; recall that now $\rho_\epsilon = \rho_\epsilon^\bullet = \rho_0 = 0$. Thus,

$$\rho_\epsilon = 0 = 0 + \epsilon\frac{{}_\circ\hat{\mu}_c(0)}{l_\epsilon^{\bullet(1)}(0)} = \rho_\epsilon^\bullet + \epsilon\frac{{}_\circ\hat{\mu}_c(\rho_\epsilon^\bullet)}{l_\epsilon^{\bullet(1)}(\rho_\epsilon^\bullet)} = \rho_\epsilon^\bullet + \epsilon\frac{{}_\circ\hat{\mu}_c(\rho_\epsilon^\bullet)}{l_\epsilon^{\bullet(1)}(\rho_\epsilon^\bullet)} + O\left(\epsilon^2\right).$$

□

Using (i) allows direct Lundberg root-based analysis for generalizing (Vatamidou *et al.*, 2013, Definition 1) from $\psi_{d,\epsilon}(u)$ to $m^i_{d,\epsilon}(u)$, where $w(\cdot,\cdot)$ is arbitrary and $\delta \geq 0$. In Vatamidou *et al.* (2014a), the derivations of corrected discard approximations linked through the "replace" base model, rather than directly using the "discard" base model. Vatamidou *et al.* (2013) commented that the "discard" case was simpler than the "replace" case for ruin probabilities, and in a similar sense we find this for Gerber-Shiu functions, after a few additional steps in deriving r_ϵ as perturbation of $r_\epsilon^\bullet$.

2.4.2. Zero Initial Surplus. The expressions Li and Sendova (2015) gave for $m_{\epsilon,i}(u)$ contained the term $\frac{\lambda_2}{c_2}m_{\epsilon,1}(0) - \frac{\lambda_1}{c_1}m_{\epsilon,2}(0)$, which depends on the claims law; therefore, we need perturbation expansions for that term before proceeding to derive the expansions from which we define $m^i_{d,\epsilon}(u)$ and $m^i_{r,\epsilon}(u)$. Now, we shall denote $\Delta_{\text{פ},i}(u) = \frac{\lambda_{3-i}}{c_{3-i}}m_{\text{פ},i}(u) - \frac{\lambda_i}{c_i}m_{\text{פ},3-i}(u)$. Then, corresponding to (Li and Sendova, 2015, Eqn 4.3), we have that $\Delta_{\text{פ},i}(0) = \frac{\lambda_i}{c_i}\frac{\gamma_{\text{פ},3-i}(r_{\text{פ}},\rho_{\text{פ}})}{\zeta_{\text{פ},3-i}(r_{\text{פ}},\rho_{\text{פ}})}$; here

$$
\begin{aligned}
\gamma_{\text{פ},i}(r,\rho) &= \left(\rho - \frac{\delta+\lambda_i}{c_i}\right)\left(r - \frac{\delta+\lambda_i}{c_i}\right)(\hat{\omega}_{\text{פ}}(r) - \hat{\omega}_{\text{פ}}(\rho)),\\
\zeta_{\text{פ},i}(r,\rho) &= \chi_{\text{פ},i}(r)\left(\rho - \frac{\delta+\lambda_i}{c_i}\right) - \chi_{\text{פ},i}(\rho)\left(r - \frac{\delta+\lambda_i}{c_i}\right).
\end{aligned}
$$

Next, we denote:

$$
\nabla^c_{\text{פ}}\aleph(r,\rho) = \frac{{}_\circ\hat{\mu}_c(r)}{l^{(1)}_{\text{פ}}(r)}\frac{\partial}{\partial r}\aleph(r,\rho) + \frac{{}_\circ\hat{\mu}_c(\rho)}{l^{(1)}_{\text{פ}}(\rho)}\frac{\partial}{\partial \rho}\aleph(r,\rho),
$$

and

$$
\nabla^b_{\text{פ}}\aleph(r,\rho) = \frac{{}_\circ\hat{\mu}_b(r)}{l^{(1)}_{\text{פ}}(r)}\frac{\partial}{\partial r}\aleph(r,\rho) + \frac{{}_\circ\hat{\mu}_b(\rho)}{l^{(1)}_{\text{פ}}(\rho)}\frac{\partial}{\partial \rho}\aleph(r,\rho).
$$

Using the Lundberg root perturbation expansions given in Lemma 3 above, one may show that

$$
\begin{aligned}
\Delta_{\epsilon,i}(0) &= \Delta^{\bullet}_{\epsilon,i}(0) + \epsilon\kappa^{\bullet c}_{\epsilon,i}(0) + O\left(\epsilon^2\right),\\
\Delta_{\epsilon,i}(0) &= \Delta_{0,i}(0) + \epsilon\left(\kappa^c_{0,i}(0) - \kappa^b_{0,i}(0)\right) + O\left(\epsilon^2\right),
\end{aligned}
\tag{2.2}
$$

in which, we have:

$$
\begin{aligned}
\kappa^c_{\text{פ},i}(0) = {}& \frac{\lambda_i}{c_i}\frac{1}{\zeta_{\text{פ},3-i}(r_{\text{פ}},\rho_{\text{פ}})}(\gamma_{c,3-i}(r_{\text{פ}},\rho_{\text{פ}}) + \nabla^c_{\text{פ}}\gamma_{\text{פ},3-i}(r_{\text{פ}},\rho_{\text{פ}}))\\
&- \frac{\lambda_i}{c_i}\frac{\gamma_{\text{פ},3-i}(r_{\text{פ}},\rho_{\text{פ}})}{(\zeta_{\text{פ},3-i}(r_{\text{פ}},\rho_{\text{פ}}))^2}({}_\circ\zeta_{c,3-i}(r_{\text{פ}},\rho_{\text{פ}}) + \nabla^c_{\text{פ}}\zeta_{\text{פ},3-i}(r_{\text{פ}},\rho_{\text{פ}})),
\end{aligned}
\tag{2.3}
$$

$$
\kappa^b_{\text{פ},i}(0) = \frac{\lambda_i}{c_i}\frac{1}{\zeta_{\text{פ},3-i}(r_{\text{פ}},\rho_{\text{פ}})}\left(\gamma_{b,3-i}(r_{\text{פ}},\rho_{\text{פ}}) + \nabla^b_{\text{פ}}\gamma_{\text{פ},3-i}(r_{\text{פ}},\rho_{\text{פ}})\right)
\tag{2.4}
$$

$$
- \frac{\lambda_i}{c_i} \frac{\gamma_{\text{פ},3-i}(r_{\text{פ}}, \rho_{\text{פ}})}{\left(\zeta_{\text{פ},3-i}(r_{\text{פ}}, \rho_{\text{פ}})\right)^2} \Big({}_{\circ}\zeta_{b,3-i}(r_{\text{פ}}, \rho_{\text{פ}}) + \nabla^{b}_{\text{פ}} \zeta_{\text{פ},3-i}(r_{\text{פ}}, \rho_{\text{פ}}) \Big).
$$

To see the validity of the first expansion in (2.2), note that

$$
\begin{aligned}
\gamma_{\epsilon,i}(r_\epsilon, \rho_\epsilon) &= \left(\rho_\epsilon - \frac{\delta + \lambda_i}{c_i}\right)\left(r_\epsilon - \frac{\delta + \lambda_i}{c_i}\right)(\hat{\omega}_\epsilon(r_\epsilon) - \hat{\omega}_\epsilon(\rho_\epsilon)) \\
&= \gamma^{\bullet}_{\epsilon,i}(r_\epsilon, \rho_\epsilon) + \epsilon \gamma_{c,i}(r_\epsilon, \rho_\epsilon) \\
&= \gamma^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + (r_\epsilon - r^{\bullet}_\epsilon)\frac{\partial}{\partial r}\gamma^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + (\rho_\epsilon - \rho^{\bullet}_\epsilon)\frac{\partial}{\partial \rho}\gamma^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) \\
&\quad + \epsilon\left(\gamma_{c,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + (r_\epsilon - r^{\bullet}_\epsilon)\frac{\partial}{\partial r}\gamma_{c,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + (\rho_\epsilon - \rho^{\bullet}_\epsilon)\frac{\partial}{\partial \rho}\gamma_{c,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon)\right) \\
&\quad + O\left(\epsilon^2\right) \\
&= \gamma^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + \epsilon\Big(\nabla^{\bullet c}_{\epsilon} \gamma^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + \gamma_{c,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon)\Big) + O\left(\epsilon^2\right). \qquad (2.5)
\end{aligned}
$$

The first step follows from $\hat{\omega}_\epsilon(s) = \hat{\omega}^{\bullet}_\epsilon(s) + \epsilon\hat{\omega}_c(s)$; in the second step, we use the bivariate Taylor expansion; in the third step, we use Lemma 3 (i) and the definition of $\nabla^{\bullet c}_{\epsilon}$. Also, as $\chi_{\epsilon,i}(s) = \chi^{\bullet}_{\epsilon,i}(s) + {}_{\circ}\chi_{c,i}(s)$,

$$
\begin{aligned}
\zeta_{\epsilon,i}(r_\epsilon, \rho_\epsilon) &= \chi_{\epsilon,i}(r_\epsilon)\left(\rho_\epsilon - \frac{\delta + \lambda_i}{c_i}\right) - \chi_{\epsilon,i}(\rho_\epsilon)\left(r_\epsilon - \frac{\delta + \lambda_i}{c_i}\right) \\
&= \zeta^{\bullet}_{\epsilon,i}(r_\epsilon, \rho_\epsilon) + \epsilon {}_{\circ}\zeta_{c,i}(r_\epsilon, \rho_\epsilon) \\
&= \zeta^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + (r_\epsilon - r^{\bullet}_\epsilon)\frac{\partial}{\partial r}\zeta^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + (\rho_\epsilon - \rho^{\bullet}_\epsilon)\frac{\partial}{\partial \rho}\zeta^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) \\
&\quad + \epsilon\left({}_{\circ}\zeta_{c,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + (r_\epsilon - r^{\bullet}_\epsilon)\frac{\partial}{\partial r}{}_{\circ}\zeta_{c,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + (\rho_\epsilon - \rho^{\bullet}_\epsilon)\frac{\partial}{\partial \rho}{}_{\circ}\zeta_{c,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon)\right) \\
&\quad + O\left(\epsilon^2\right) \\
&= \zeta^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + \epsilon\Big(\nabla^{\bullet c}_{\epsilon} \zeta^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + {}_{\circ}\zeta_{c,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon)\Big) + O\left(\epsilon^2\right).
\end{aligned}
$$

Then

$$
\frac{1}{\zeta_{\epsilon,i}(r_\epsilon, \rho_\epsilon)} = \frac{1}{\zeta^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + \epsilon\Big(\nabla^{\bullet c}_{\epsilon} \zeta^{\bullet}_{\epsilon,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon) + {}_{\circ}\zeta_{c,i}(r^{\bullet}_\epsilon, \rho^{\bullet}_\epsilon)\Big) + O(\epsilon^2)}
$$

$$= \frac{1}{\zeta^{\bullet}_{\epsilon,i}(r^{\bullet}_{\epsilon}, \rho^{\bullet}_{\epsilon})} \left(1 - \frac{1}{\zeta^{\bullet}_{\epsilon,i}(r^{\bullet}_{\epsilon}, \rho^{\bullet}_{\epsilon})} \epsilon \left(\nabla^{\bullet c}_{\epsilon} \zeta^{\bullet}_{\epsilon,i}(r^{\bullet}_{\epsilon}, \rho^{\bullet}_{\epsilon}) + {}_{\circ}\zeta_{c,i}(r^{\bullet}_{\epsilon}, \rho^{\bullet}_{\epsilon}) \right) + O\left(\epsilon^2\right) \right);$$

combining this with (2.5) gives the first line of (2.2).

2.4.3. GSF Laplace Transform Perturbation Expansions. Using Lemma 3 and the perturbation expansions for $\Delta_{\epsilon,i}(0)$ given in (2.2), in Theorem 1 we give two perturbation expansions of the Laplace transforms $\hat{m}_{\epsilon,i}(s)$ for which the base term is respectively $\hat{m}^{\bullet}_{\epsilon,i}(s)$ and $\hat{m}_{0,i}(s)$, and the next term contains $C(x)$ linearly at most once per convolution in any component. Our Theorem 1 giving the Laplace transforms is more similar to (Vatamidou *et al.*, 2014a, Propositions 3.6, 4.5) than (Vatamidou *et al.*, 2013, Theorems 1, 2). Unlike either pair of existing results, we leave $w(\cdot,\cdot)$ and $\delta \geq 0$ unspecified. Consequently, we do not make explicit every part of the correction terms. Like in Vatamidou *et al.* (2014a), we only specify the expansions up to $O\left(\epsilon^2\right)$; this is one of the alluded consequences of Lemma 3.

Theorem 1. *We have the following "discard" and "replace" perturbation expansions for the Laplace transform of the general Gerber-Shiu function in the Li-Sendova dependent risk model:*

$$\text{(i) } \hat{m}_{\epsilon,i}(s) = \hat{m}^{\bullet}_{\epsilon,i}(s) + \epsilon \tfrac{\tilde{G}^{\bullet}_{\epsilon}(s)}{1 - {}_{\delta}\upsilon^{\bullet}_{\epsilon}} \mathrm{T}_s \, \mathrm{T}_{r^{\bullet}_{\epsilon}} \, \mathrm{T}_{\rho^{\bullet}_{\epsilon}} \, \text{ק}^{\bullet}_{\epsilon,i}(0) + O\left(\epsilon^2\right), \textit{ and}$$
$$\text{(ii) } \hat{m}_{\epsilon,i}(s) = \hat{m}_{0,i}(s) + \epsilon \tfrac{\tilde{G}_0(s)}{1 - {}_{\delta}\upsilon_0} \mathrm{T}_s \, \mathrm{T}_{r_0} \, \mathrm{T}_{\rho_0} \left(\text{ק}_{0,i} - \text{ב}_{0,i} \right)(0) + O\left(\epsilon^2\right).$$

In (i) *and* (ii)*, the term* ${}_{\delta}\upsilon_{\text{פ}} = \mathrm{T}_0 \, \mathrm{T}_{r_{\text{פ}}} \, \mathrm{T}_{\rho_{\text{פ}}} \, \mu_{\text{פ}}(0)$*, and* $\tilde{G}_{\text{פ}}(s) = \frac{1 - {}_{\delta}\upsilon_{\text{פ}}}{1 - \mathrm{T}_s \, \mathrm{T}_{r_{\text{פ}}} \, \mathrm{T}_{\rho_{\text{פ}}} \, \mu_{\text{פ}}(0)}$ *is the Laplace-Stieltjes transform of the common compound geometric distribution arising from (Li and Sendova, 2015, Theorem 4.1) with claims df* $P_{\text{פ}}(y)$*. With* $\kappa^c_{\text{פ},i}(0)$ *and* $\kappa^b_{\text{פ},i}(0)$ *given by* (2.3) *and* (2.4)*, the functions* $\text{ק}_{\text{פ},i}(\cdot)$ *and* $\text{ב}_{\text{פ},i}(\cdot)$ *are specified by the Laplace transforms*

$$\hat{\text{ק}}_{\text{פ},i}(s) = \frac{\lambda_i}{c_i} \left(\frac{\delta + \lambda_{3-i}}{c_{3-i}} - s \right) \hat{\omega}_c(s) + {}_{\circ}\hat{\mu}_c(s) \hat{m}_{\text{פ},i}(s) + \Delta_{\text{פ},i}(0) \hat{\xi}_{c,3-i}(s) + \kappa^c_{\text{פ},i}(0) \hat{\xi}_{\text{פ},3-i}(s),$$
$$\hat{\text{ב}}_{\text{פ},i}(s) = \frac{\lambda_i}{c_i} \left(\frac{\delta + \lambda_{3-i}}{c_{3-i}} - s \right) \hat{\omega}_b(s) + {}_{\circ}\hat{\mu}_b(s) \hat{m}_{\text{פ},i}(s) + \Delta_{\text{פ},i}(0) \hat{\xi}_{b,3-i}(s) + \kappa^b_{\text{פ},i}(0) \hat{\xi}_{\text{פ},3-i}(s).$$

Proof of Theorem 1. The techniques used in deriving (i) apply to deriving (ii); therefore, we simply demonstrate (i). With our usage of $\Delta_{\mathfrak{D},i}(0)$ in which we use the equation with $i \in \{1, 2\}$ to derive the equivalent of (Li and Sendova, 2015, Eqn (4.3)) (they only used that of $i = 1$ or Eqn (3.9)), we may show (i) simultaneously for $i = 1$ and $i = 2$. First, as in Li and Sendova (2015) define $\hat{\beta}_{\mathfrak{D},i}(s) = \frac{\lambda_i}{c_i}\left(\frac{\delta+\lambda_{3-i}}{c_{3-i}} - s\right)\hat{\omega}_{\mathfrak{D}}(s) + \Delta_{\mathfrak{D},i}(0)\chi_{\mathfrak{D},3-i}(s)$. Next, we use $l_\epsilon(s) = l^\bullet_\epsilon(s) - \epsilon_\circ\hat{\mu}_c(s)$ and (Li and Sendova, 2015, Eqns (4.15, 4.16)), to obtain

$$\hat{m}_{\epsilon,i}(s) = \frac{\frac{(s-r_\epsilon)(s-\rho_\epsilon)}{(s-r^\bullet_\epsilon)(s-\rho^\bullet_\epsilon)}\,\mathrm{T}_s\,\mathrm{T}_{r_\epsilon}\,\mathrm{T}_{\rho_\epsilon}\,\beta_{\epsilon,i}(0)}{1 - \mathrm{T}_s\,\mathrm{T}_{r^\bullet_\epsilon}\,\mathrm{T}_{\rho^\bullet_\epsilon}\,\mu^\bullet_\epsilon(0) - \epsilon\frac{{}_\circ\hat{\mu}_c(s)}{(s-r^\bullet_\epsilon)(s-\rho^\bullet_\epsilon)}}. \tag{2.6}$$

First we handle the numerator of (2.6). Note that

$$\begin{aligned}&\frac{(s-r_\epsilon)(s-\rho_\epsilon)}{(s-r^\bullet_\epsilon)(s-\rho^\bullet_\epsilon)}\,\mathrm{T}_s\,\mathrm{T}_{r_\epsilon}\,\mathrm{T}_{\rho_\epsilon}\,\beta_{\epsilon,i}(0)\\ &= \frac{(r_\epsilon - s)\left(\hat{\beta}_{\epsilon,i}(s) - \hat{\beta}_{\epsilon,i}(\rho_\epsilon)\right) - (\rho_\epsilon - s)\left(\hat{\beta}_{\epsilon,i}(s) - \hat{\beta}_{\epsilon,i}(r_\epsilon)\right)}{(r^\bullet_\epsilon - s)(\rho^\bullet_\epsilon - s)(r_\epsilon - \rho_\epsilon)}.\end{aligned}$$

From the definitions of the respective quantities and (2.2), we have

$$\begin{aligned}\hat{\beta}_{\epsilon,i}(s) &= \frac{\lambda_i}{c_i}\left(\frac{\delta+\lambda_{3-i}}{c_{3-i}} - s\right)\left(\hat{\omega}^\bullet_\epsilon(s) + \epsilon\hat{\omega}_c(s)\right) + \left(\Delta^\bullet_{\epsilon,i}(0) + \epsilon\kappa^{\bullet c}_{\epsilon,i}(0) + O\left(\epsilon^2\right)\right)\\ &\quad\times\left(\chi^\bullet_{\epsilon,3-i}(s) + \epsilon_\circ\chi_{c,3-i}(s)\right)\\ &= \frac{\lambda_i}{c_i}\left(\frac{\delta+\lambda_{3-i}}{c_{3-i}} - s\right)\hat{\omega}^\bullet_\epsilon(s) + \Delta^\bullet_{\epsilon,i}(0)\chi^\bullet_{\epsilon,3-i}(s)\\ &\quad+\epsilon\left(\frac{\lambda_i}{c_i}\left(\frac{\delta+\lambda_{3-i}}{c_{3-i}} - s\right)\hat{\omega}_c(s) + \Delta^\bullet_{\epsilon,i}(0)_\circ\chi_{c,3-i}(s) + \kappa^{\bullet c}_{\epsilon,i}(0)\chi^\bullet_{\epsilon,3-i}(s)\right) + O\left(\epsilon^2\right)\\ &= \hat{\beta}^\bullet_{\epsilon,i}(s) + \epsilon\left(\hat{\mathfrak{p}}^\bullet_{\epsilon,i}(s) - {}_\circ\hat{\mu}_c(s)\hat{m}^\bullet_{\epsilon,i}(s) - \mathrm{I}(i=1)\Delta^\bullet_{\epsilon,i}(0)\right) + O\left(\epsilon^2\right).\end{aligned} \tag{2.7}$$

For, $\epsilon\chi^\bullet_{\epsilon,3-i}(s) + O\left(\epsilon^2\right) = \epsilon\hat{\xi}^\bullet_{\epsilon,3-i}(s) + O\left(\epsilon^2\right)$ and $\chi_{c,3-i}(s) = \hat{\xi}_{c,3-i}(s)$ for $i \in \{1, 2\}$. Observe that

$$\hat{\beta}_{\epsilon,i}(\rho_\epsilon) = \hat{\beta}^\bullet_{\epsilon,i}(\rho_\epsilon) + \epsilon\left(\hat{\mathfrak{p}}^\bullet_{\epsilon,i}(\rho_\epsilon) - {}_\circ\hat{\mu}_c(\rho_\epsilon)\hat{m}^\bullet_{\epsilon,i}(\rho_\epsilon) - \mathrm{I}(i=1)\Delta^\bullet_{\epsilon,i}(0)\right) + O\left(\epsilon^2\right)$$

$$
\begin{aligned}
&= \hat{\beta}^{\bullet}_{\epsilon,i}(\rho^{\bullet}_{\epsilon}) + (\rho_{\epsilon} - \rho^{\bullet}_{\epsilon})\hat{\beta}^{\bullet(1)}_{\epsilon,i}(\rho^{\bullet}_{\epsilon}) \\
&\quad + \epsilon\Big(\hat{\text{ק}}^{\bullet}_{\epsilon,i}(\rho^{\bullet}_{\epsilon}) - {}_{\circ}\hat{\mu}_c(\rho^{\bullet}_{\epsilon})\hat{m}^{\bullet}_{\epsilon,i}(\rho^{\bullet}_{\epsilon}) - \mathrm{I}(i=1)\Delta^{\bullet}_{\epsilon,i}(0)\Big) + O\Big(\epsilon^2\Big) \\
&= \hat{\beta}^{\bullet}_{\epsilon,i}(\rho^{\bullet}_{\epsilon}) + \epsilon\Bigg(\frac{{}_{\circ}\hat{\mu}_c(\rho^{\bullet}_{\epsilon})}{l^{\bullet(1)}_{\epsilon}(\rho^{\bullet}_{\epsilon})}\hat{\beta}^{\bullet(1)}_{\epsilon,i}(\rho^{\bullet}_{\epsilon}) + \hat{\text{ק}}^{\bullet}_{\epsilon,i}(\rho^{\bullet}_{\epsilon}) \qquad (2.8) \\
&\quad - {}_{\circ}\hat{\mu}_c(\rho^{\bullet}_{\epsilon})\hat{m}^{\bullet}_{\epsilon,i}(\rho^{\bullet}_{\epsilon}) - \mathrm{I}(i=1)\Delta^{\bullet}_{\epsilon,i}(0)\Bigg) + O\Big(\epsilon^2\Big).
\end{aligned}
$$

For, e.g., $\epsilon(\rho_{\epsilon} - \rho^{\bullet}_{\epsilon})\hat{\text{ק}}^{\bullet(1)}_{\epsilon,i}(\rho^{\bullet}_{\epsilon}) = O\Big(\epsilon^2\Big)$ by Lemma 3 (i). Clearly, (2.8) also holds with r_{ϵ} and $r^{\bullet}_{\epsilon}$ instead of ρ_{ϵ} and $\rho^{\bullet}_{\epsilon}$. So, we have that

$$
\begin{aligned}
&(r_{\epsilon} - s)\Big(\hat{\beta}_{\epsilon,i}(s) - \hat{\beta}_{\epsilon,i}(\rho_{\epsilon})\Big) \qquad (2.9) \\
&= (r^{\bullet}_{\epsilon} - s)(\rho^{\bullet}_{\epsilon} - s)\Big\{\mathrm{T}_s\,\mathrm{T}_{\rho^{\bullet}_{\epsilon}}\,\beta^{\bullet}_{\epsilon,i}(0) + \epsilon\Big(\mathrm{T}_s\,\mathrm{T}_{\rho^{\bullet}_{\epsilon}}\Big(\text{ק}^{\bullet}_{\epsilon,i} - {}_{\circ}\mu_c * m^{\bullet}_{\epsilon,i}\Big)(0) \\
&\quad - \frac{{}_{\circ}\hat{\mu}_c(\rho^{\bullet}_{\epsilon})}{l^{\bullet(1)}_{\epsilon}(\rho^{\bullet}_{\epsilon})}\frac{\hat{\beta}^{\bullet(1)}_{\epsilon,i}(\rho^{\bullet}_{\epsilon})}{\rho^{\bullet}_{\epsilon} - s} + \frac{{}_{\circ}\hat{\mu}_c(r^{\bullet}_{\epsilon})}{l^{\bullet(1)}_{\epsilon}(r^{\bullet}_{\epsilon})}\frac{\mathrm{T}_s\,\mathrm{T}_{\rho^{\bullet}_{\epsilon}}\,\beta^{\bullet}_{\epsilon,i}(0)}{r^{\bullet}_{\epsilon} - s}\Big) + O\Big(\epsilon^2\Big)\Big\},
\end{aligned}
$$

and we also have

$$
\begin{aligned}
&(\rho_{\epsilon} - s)\Big(\hat{\beta}_{\epsilon,i}(s) - \hat{\beta}_{\epsilon,i}(r_{\epsilon})\Big) \qquad (2.10) \\
&= (\rho^{\bullet}_{\epsilon} - s)(r^{\bullet}_{\epsilon} - s)\Big\{\mathrm{T}_s\,\mathrm{T}_{r^{\bullet}_{\epsilon}}\,\beta^{\bullet}_{\epsilon,i}(0) + \epsilon\Big(\mathrm{T}_s\,\mathrm{T}_{r^{\bullet}_{\epsilon}}\Big(\text{ק}^{\bullet}_{\epsilon,i} - {}_{\circ}\mu_c * m^{\bullet}_{\epsilon,i}\Big)(0) \\
&\quad - \frac{{}_{\circ}\hat{\mu}_c(r^{\bullet}_{\epsilon})}{l^{\bullet(1)}_{\epsilon}(r^{\bullet}_{\epsilon})}\frac{\hat{\beta}^{\bullet(1)}_{\epsilon,i}(r^{\bullet}_{\epsilon})}{r^{\bullet}_{\epsilon} - s} + \frac{{}_{\circ}\hat{\mu}_c(\rho^{\bullet}_{\epsilon})}{l^{\bullet(1)}_{\epsilon}(\rho^{\bullet}_{\epsilon})}\frac{\mathrm{T}_s\,\mathrm{T}_{r^{\bullet}_{\epsilon}}\,\beta^{\bullet}_{\epsilon,i}(0)}{\rho^{\bullet}_{\epsilon} - s}\Big) + O\Big(\epsilon^2\Big)\Big\}.
\end{aligned}
$$

Therefore, substituting (2.9) and (2.10) into the numerator of (2.6),

$$
\begin{aligned}
&\frac{(s - r_{\epsilon})(s - \rho_{\epsilon})}{(s - r^{\bullet}_{\epsilon})(s - \rho^{\bullet}_{\epsilon})}\mathrm{T}_s\,\mathrm{T}_{r_{\epsilon}}\,\mathrm{T}_{\rho_{\epsilon}}\,\beta_{\epsilon,i}(0) \\
&= \mathrm{T}_s\,\mathrm{T}_{r^{\bullet}_{\epsilon}}\,\mathrm{T}_{\rho^{\bullet}_{\epsilon}}\,\beta^{\bullet}_{\epsilon,i}(0) + \epsilon\Bigg\{\frac{1}{r^{\bullet}_{\epsilon} - \rho^{\bullet}_{\epsilon}}\Bigg(\frac{{}_{\circ}\hat{\mu}_c(\rho^{\bullet}_{\epsilon})}{l^{\bullet(1)}_{\epsilon}(\rho^{\bullet}_{\epsilon})} - \frac{{}_{\circ}\hat{\mu}_c(r^{\bullet}_{\epsilon})}{l^{\bullet(1)}_{\epsilon}(r^{\bullet}_{\epsilon})}\Bigg)\mathrm{T}_s\,\mathrm{T}_{r^{\bullet}_{\epsilon}}\,\mathrm{T}_{\rho^{\bullet}_{\epsilon}}\,\beta^{\bullet}_{\epsilon,i}(0) \\
&\quad + \mathrm{T}_s\,\mathrm{T}_{r^{\bullet}_{\epsilon}}\,\mathrm{T}_{\rho^{\bullet}_{\epsilon}}\Big(\text{ק}^{\bullet}_{\epsilon,i} - {}_{\circ}\mu_c * m^{\bullet}_{\epsilon,i}\Big)(0) + \frac{{}_{\circ}\hat{\mu}_c(r^{\bullet}_{\epsilon})}{l^{\bullet(1)}_{\epsilon}(r^{\bullet}_{\epsilon})}\frac{\hat{\beta}^{\bullet(1)}_{\epsilon,i}(r^{\bullet}_{\epsilon}) + \mathrm{T}_s\,\mathrm{T}_{\rho^{\bullet}_{\epsilon}}\,\beta^{\bullet}_{\epsilon,i}(0)}{(r^{\bullet}_{\epsilon} - s)(r^{\bullet}_{\epsilon} - \rho^{\bullet}_{\epsilon})} \\
&\quad - \frac{{}_{\circ}\hat{\mu}_c(\rho^{\bullet}_{\epsilon})}{l^{\bullet(1)}_{\epsilon}(\rho^{\bullet}_{\epsilon})}\frac{\hat{\beta}^{\bullet(1)}_{\epsilon,i}(\rho^{\bullet}_{\epsilon}) + \mathrm{T}_s\mathrm{T}_{r^{\bullet}_{\epsilon}}\beta^{\bullet}_{\epsilon,i}(0)}{(\rho^{\bullet}_{\epsilon} - s)(r^{\bullet}_{\epsilon} - \rho^{\bullet}_{\epsilon})}\Bigg\} + O\Big(\epsilon^2\Big), \qquad (2.11)
\end{aligned}
$$

because Lemma 3 implies

$$(r_\epsilon - \rho_\epsilon)^{-1} = (r^\bullet_\epsilon - \rho^\bullet_\epsilon)^{-1}\left(1 - \epsilon\frac{1}{r^\bullet_\epsilon - \rho^\bullet_\epsilon}\left(\frac{{}_\circ\hat{\mu}_c(r^\bullet_\epsilon)}{l^{\bullet(1)}_\epsilon(r^\bullet_\epsilon)} - \frac{{}_\circ\hat{\mu}_c(\rho^\bullet_\epsilon)}{l^{\bullet(1)}_\epsilon(\rho^\bullet_\epsilon)}\right) + O\left(\epsilon^2\right)\right).$$

Now, considering the property $\mathrm{T}_r\,\mathrm{T}_r\,\beta(0) = -\hat{\beta}^{(1)}(r)$ found in Li and Garrido (2004),

$$\begin{aligned}\frac{\hat{\beta}^{(1)}(r) + \mathrm{T}_s\,\mathrm{T}_\rho\,\beta(0)}{(r-s)(r-\rho)} &= \frac{1}{r-\rho}\,\mathrm{T}_s\,\mathrm{T}_r\,\mathrm{T}_\rho\,\beta(0) + \frac{1}{r-s}\,\mathrm{T}_r\,\mathrm{T}_r\,\mathrm{T}_\rho\,\beta(0), \text{ and}\\ \frac{\hat{\beta}^{(1)}(\rho) + \mathrm{T}_s\,\mathrm{T}_r\,\beta(0)}{(\rho-s)(r-\rho)} &= \frac{1}{r-\rho}\,\mathrm{T}_s\,\mathrm{T}_r\,\mathrm{T}_\rho\,\beta(0) - \frac{1}{\rho-s}\,\mathrm{T}_\rho\,\mathrm{T}_r\,\mathrm{T}_\rho\,\beta(0)\end{aligned}$$

follow by adding $\mathrm{T}_r\,\mathrm{T}_\rho\,\beta(0) - \mathrm{T}_r\,\mathrm{T}_\rho\,\beta(0)$ to the numerators on the left-hand side. Therefore, (2.11) becomes

$$\begin{aligned}&\frac{(s-r_\epsilon)(s-\rho_\epsilon)}{(s-r^\bullet_\epsilon)(s-\rho^\bullet_\epsilon)}\,\mathrm{T}_s\,\mathrm{T}_{r_\epsilon}\,\mathrm{T}_{\rho_\epsilon}\,\beta_{\epsilon,i}(0)\\ &= \mathrm{T}_s\,\mathrm{T}_{r^\bullet_\epsilon}\,\mathrm{T}_{\rho^\bullet_\epsilon}\,\beta^\bullet_{\epsilon,i}(0) + \epsilon\Bigg\{\mathrm{T}_s\,\mathrm{T}_{r^\bullet_\epsilon}\,\mathrm{T}_{\rho^\bullet_\epsilon}\left(\text{ק}^\bullet_{\epsilon,i} - {}_\circ\mu_c * m^\bullet_{\epsilon,i}\right)(0)\\ &\quad + \frac{{}_\circ\hat{\mu}_c(r^\bullet_\epsilon)}{l^{\bullet(1)}_\epsilon(r^\bullet_\epsilon)}\,\frac{\mathrm{T}_{r^\bullet_\epsilon}\,\mathrm{T}_{r^\bullet_\epsilon}\,\mathrm{T}_{\rho^\bullet_\epsilon}\,\beta^\bullet_{\epsilon,i}(0)}{r^\bullet_\epsilon - s} + \frac{{}_\circ\hat{\mu}_c(\rho^\bullet_\epsilon)}{l^{\bullet(1)}_\epsilon(\rho^\bullet_\epsilon)}\,\frac{\mathrm{T}_{\rho^\bullet_\epsilon}\,\mathrm{T}_{r^\bullet_\epsilon}\,\mathrm{T}_{\rho^\bullet_\epsilon}\,\beta^\bullet_{\epsilon,i}(0)}{\rho^\bullet_\epsilon - s}\Bigg\} + O\left(\epsilon^2\right).\end{aligned} \tag{2.12}$$

We easily see for the Lundberg roots that

$$\begin{aligned}l^{(1)}_{\text{פּ}}(r_{\text{פּ}}) &= (r_{\text{פּ}} - \rho_{\text{פּ}})\left(1 - \mathrm{T}_{r_{\text{פּ}}}\,\mathrm{T}_{r_{\text{פּ}}}\,\mathrm{T}_{\rho_{\text{פּ}}}\,\mu_{\text{פּ}}(0)\right), \text{ and}\\ l^{(1)}_{\text{פּ}}(\rho_{\text{פּ}}) &= (\rho_{\text{פּ}} - r_{\text{פּ}})\left(1 - \mathrm{T}_{\rho_{\text{פּ}}}\,\mathrm{T}_{r_{\text{פּ}}}\,\mathrm{T}_{\rho_{\text{פּ}}}\,\mu_{\text{פּ}}(0)\right).\end{aligned}$$

For,

$$\begin{aligned}l^{(1)}_{\text{פּ}}(s) &= \frac{\partial}{\partial s}(s - r_{\text{פּ}})(s - \rho_{\text{פּ}})\left(1 - \mathrm{T}_s\,\mathrm{T}_{r_{\text{פּ}}}\,\mathrm{T}_{\rho_{\text{פּ}}}\,\mu_{\text{פּ}}(0)\right)\\ &= (2s - (r_{\text{פּ}} + \rho_{\text{פּ}}))\left(1 - \mathrm{T}_s\,\mathrm{T}_{r_{\text{פּ}}}\,\mathrm{T}_{\rho_{\text{פּ}}}\,\mu_{\text{פּ}}(0)\right)\\ &\quad - (s - r_{\text{פּ}})(s - \rho_{\text{פּ}})\frac{\partial}{\partial s}\,\mathrm{T}_s\,\mathrm{T}_{r_{\text{פּ}}}\,\mathrm{T}_{\rho_{\text{פּ}}}\,\mu_{\text{פּ}}(0).\end{aligned}$$

Hence,

$$
\begin{aligned}
&\frac{(s-r_\epsilon)(s-\rho_\epsilon)}{(s-r_\epsilon^\bullet)(s-\rho_\epsilon^\bullet)}\,\mathrm{T}_s\,\mathrm{T}_{r_\epsilon}\,\mathrm{T}_{\rho_\epsilon}\,\beta_{\epsilon,i}(0) \\
&= \mathrm{T}_s\,\mathrm{T}_{r_\epsilon^\bullet}\,\mathrm{T}_{\rho_\epsilon^\bullet}\,\beta^\bullet_{\epsilon,i}(0) + \epsilon\Bigg\{\mathrm{T}_s\,\mathrm{T}_{r_\epsilon^\bullet}\,\mathrm{T}_{\rho_\epsilon^\bullet}\Big(\text{ק}^\bullet_{\epsilon,i} - {}_\circ\mu_c * m^\bullet_{\epsilon,i}\Big)(0) + \frac{{}_\circ\hat{\mu}_c(r_\epsilon^\bullet)\hat{m}^\bullet_{\epsilon,i}(r_\epsilon^\bullet)}{(r_\epsilon^\bullet-\rho_\epsilon^\bullet)(r_\epsilon^\bullet-s)} \\
&\quad - \frac{{}_\circ\hat{\mu}_c(\rho_\epsilon^\bullet)\hat{m}^\bullet_{\epsilon,i}(\rho_\epsilon^\bullet)}{(r_\epsilon^\bullet-\rho_\epsilon^\bullet)(\rho_\epsilon^\bullet-s)}\Bigg\} + O\Big(\epsilon^2\Big) \\
&= \mathrm{T}_s\,\mathrm{T}_{r_\epsilon^\bullet}\,\mathrm{T}_{\rho_\epsilon^\bullet}\,\beta^\bullet_{\epsilon,i}(0) + \epsilon\Bigg\{\mathrm{T}_s\,\mathrm{T}_{r_\epsilon^\bullet}\,\mathrm{T}_{\rho_\epsilon^\bullet}\,\text{ק}^\bullet_{\epsilon,i}(0) - \frac{{}_\circ\hat{\mu}_c(s)\hat{m}^\bullet_{\epsilon,i}(s)}{(r_\epsilon^\bullet-s)(\rho_\epsilon^\bullet-s)}\Bigg\} + O\Big(\epsilon^2\Big),
\end{aligned}
\tag{2.13}
$$

by the identity

$$
\mathrm{T}_s\,\mathrm{T}_r\,\mathrm{T}_\rho\,a(0) = \frac{\hat{a}(r)}{(r-s)(r-\rho)} - \frac{\hat{a}(\rho)}{(\rho-s)(r-\rho)} + \frac{\hat{a}(s)}{(r-s)(\rho-s)}. \tag{2.14}
$$

Reassembling $\hat{m}_{\epsilon,i}(s)$, we get

$$
\begin{aligned}
\hat{m}_{\epsilon,i}(s) &= \Bigg\{\mathrm{T}_s\,\mathrm{T}_{r_\epsilon^\bullet}\,\mathrm{T}_{\rho_\epsilon^\bullet}\,\beta^\bullet_{\epsilon,i}(0) + \epsilon\Bigg(\mathrm{T}_s\,\mathrm{T}_{r_\epsilon^\bullet}\,\mathrm{T}_{\rho_\epsilon^\bullet}\,\text{ק}^\bullet_{\epsilon,i}(0) - \frac{{}_\circ\hat{\mu}_c(s)\hat{m}^\bullet_{\epsilon,i}(s)}{(r_\epsilon^\bullet-s)(\rho_\epsilon^\bullet-s)}\Bigg) + O\Big(\epsilon^2\Big)\Bigg\} \\
&\qquad \times\Bigg(1 - \mathrm{T}_s\,\mathrm{T}_{r_\epsilon^\bullet}\,\mathrm{T}_{\rho_\epsilon^\bullet}\,\mu^\bullet_\epsilon(0) - \epsilon\frac{{}_\circ\hat{\mu}_c(s)}{(s-r_\epsilon^\bullet)(s-\rho_\epsilon^\bullet)}\Bigg)^{-1} \\
&= \Bigg\{\hat{m}^\bullet_{\epsilon,i}(s) + \epsilon\frac{\tilde{G}^\bullet_\epsilon(s)}{1-{}_\delta\upsilon^\bullet_\epsilon}\Bigg(\mathrm{T}_s\,\mathrm{T}_{r_\epsilon^\bullet}\,\mathrm{T}_{\rho_\epsilon^\bullet}\,\text{ק}^\bullet_{\epsilon,i}(0) - \frac{{}_\circ\hat{\mu}_c(s)\hat{m}^\bullet_{\epsilon,i}(s)}{(r_\epsilon^\bullet-s)(\rho_\epsilon^\bullet-s)}\Bigg) + O\Big(\epsilon^2\Big)\Bigg\} \\
&\qquad \times\Bigg(1 + \epsilon\frac{{}_\circ\hat{\mu}_c(s)}{(s-r_\epsilon^\bullet)(s-\rho_\epsilon^\bullet)}\frac{\tilde{G}^\bullet_\epsilon(s)}{1-{}_\delta\upsilon^\bullet_\epsilon} + O\Big(\epsilon^2\Big)\Bigg) \\
&= \hat{m}^\bullet_{\epsilon,i}(s) + \epsilon\frac{\tilde{G}^\bullet_\epsilon(s)}{1-{}_\delta\upsilon^\bullet_\epsilon}\,\mathrm{T}_s\,\mathrm{T}_{r_\epsilon^\bullet}\,\mathrm{T}_{\rho_\epsilon^\bullet}\,\text{ק}^\bullet_{\epsilon,i}(0) + O\Big(\epsilon^2\Big).
\end{aligned}
$$

The second equality follows by multiplying the numerator and denominator of the first equality by $\frac{\tilde{G}^\bullet_\epsilon(s)}{1-{}_\delta\upsilon^\bullet_\epsilon} = \frac{1}{1-\mathrm{T}_s\,\mathrm{T}_{r_\epsilon^\bullet}\,\mathrm{T}_{\rho_\epsilon^\bullet}\,\mu^\bullet_\epsilon(0)}$, then expanding $\frac{1}{1-z}$ with $z = \epsilon\frac{1}{1-{}_\delta\upsilon^\bullet_\epsilon}\frac{{}_\circ\hat{\mu}_c(s)\tilde{G}^\bullet_\epsilon(s)}{(s-r_\epsilon^\bullet)(s-\rho_\epsilon^\bullet)}$. $\square$

We use Theorem 1 as the basis for the corrected phase-type approximations $m^i_{d,\epsilon}(u)$ and $m^i_{r,\epsilon}(u)$. That is, truncating the $O\Big(\epsilon^2\Big)$-terms in Theorem 1 (i) and Theorem 1 (ii), then inverting with respect to s, we have the following definition:

Definition 6 (Corrected Phase-Type Approximations of Gerber-Shiu Functions)**.** *In the Li-Sendova dependent risk model, the corrected "discard" and "replace" approximations for initial insured classes $i = 1, 2$ with general penalty function $w(\cdot, \cdot)$ are*

$$m^i_{d,\epsilon}(u) = m^\bullet_{\epsilon,i}(u) + \frac{\epsilon}{1 - {}_\delta\upsilon^\bullet_\epsilon} \int_{0-}^{u} \mathrm{T}_{r^\bullet_\epsilon}\, \mathrm{T}_{\rho^\bullet_\epsilon}\, \text{ק}^\bullet_{\epsilon,i}(u-x)\, \mathrm{d}G^\bullet_\epsilon(x),$$
$$m^i_{r,\epsilon}(u) = m_{0,i}(u) + \frac{\epsilon}{1 - {}_\delta\upsilon_0} \int_{0-}^{u} \mathrm{T}_{r_0}\, \mathrm{T}_{\rho_0} \left(\text{ק}_{0,i} - \text{ב}_{0,i}\right)(u-x)\, \mathrm{d}G_0(x).$$

By essentially the procedure used to derive (Li and Sendova, 2015, Eqn 4.19), one may get an explicit expression for the correction terms. Now we simplify Theorem 1 to the compound Poisson risk model.

Corollary 1. *The Laplace transform $\hat{m}_\epsilon(s)$ may be expressed as:*

$$\hat{m}_\epsilon(s) = \hat{m}^\bullet_\epsilon(s) + \frac{\epsilon\lambda}{1 - \phi^\bullet_\epsilon(0)} \tilde{G}^\bullet_\epsilon(s)\, \mathrm{T}_s\, \mathrm{T}_{\rho^\bullet_\epsilon}(\omega_c + c * m^\bullet_\epsilon - m^\bullet_\epsilon)(0) + O\left(\epsilon^2\right),$$
$$\hat{m}_\epsilon(s) = \hat{m}_0(s) + \frac{\epsilon\lambda}{1 - \phi_0(0)} \tilde{G}_0(s)\, \mathrm{T}_s\, \mathrm{T}_{\rho_0}(\omega_c + c * m_0 - \omega_b - b * m_0)(0) + O\left(\epsilon^2\right).$$

Proof of Corollary 1. While Li and Sendova (2015) did not explicitly state this, it is clear that when $\lambda_1 = \lambda_2 = \lambda$ and $c_1 = c_2 = 1$, we have $r_{\text{פ}} = \delta + \lambda$, so that $\Delta_{\text{פ},i}(0) \equiv 0$. Also, $\gamma_{c,i}(r_{\text{פ}}, \rho_{\text{פ}}) = \gamma_{b,i}(r_{\text{פ}}, \rho_{\text{פ}}) \equiv 0$ in this case, such that $\kappa^c_{\text{פ},i}(0) = \kappa^b_{\text{פ},i}(0) \equiv 0$ as well. Then we have

$$\hat{\text{ק}}_{\text{פ}}(s) = \lambda(\delta + \lambda - s)(\hat{\omega}_c(s) + (\hat{c}(s) - 1)\hat{m}_{\text{פ}}(s)),$$
$$\hat{\text{ב}}_{\text{פ}}(s) = \lambda(\delta + \lambda - s)\left(\hat{\omega}_b(s) + \left(\hat{b}(s) - 1\right)\hat{m}_{\text{פ}}(s)\right).$$

By (2.14), we readily see that $\mathrm{T}_s\, \mathrm{T}_{r_{\text{פ}}}\, \mathrm{T}_{\rho_{\text{פ}}}\, \text{ק}_{\text{פ}}(0) = \lambda\, \mathrm{T}_s\, \mathrm{T}_{\rho_{\text{פ}}}(\omega_c + c * m_{\text{פ}} - m_{\text{פ}})(0)$ and $\mathrm{T}_s\, \mathrm{T}_{r_{\text{פ}}}\, \mathrm{T}_{\rho_{\text{פ}}}\, \text{ב}_{\text{פ}}(0) = \lambda\, \mathrm{T}_s\, \mathrm{T}_{\rho_{\text{פ}}}(\omega_b + b * m_{\text{פ}} - m_{\text{פ}})(0)$. We make this use of (2.14) explicit in showing ${}_\delta\upsilon_{\text{פ}} = \phi_{\text{פ}}(0)$. Denoting $\hat{g}_{\text{פ}}(s) = \mathrm{T}_s\, \mathrm{T}_{r_{\text{פ}}}\, \mathrm{T}_{\rho_{\text{פ}}}\, \mu_{\text{פ}}(0)$, we observe that ${}_\delta\upsilon_{\text{פ}} = \mathrm{T}_0\, \mathrm{T}_{r_{\text{פ}}}\, \mathrm{T}_{\rho_{\text{פ}}}\, \mu_{\text{פ}}(0) = \hat{g}_{\text{פ}}(0)$. Now, in the compound Poisson risk model with unit premium rate, $\hat{\mu}_{\mathfrak{D}}(s) = \lambda(\delta + \lambda - s)\tilde{P}_{\mathfrak{D}}(s)$; thus by identity (2.14),

$$\begin{aligned}
\hat{g}_{\mathfrak{D}}(s) &= \frac{\hat{\mu}_{\mathfrak{D}}(r_{\mathfrak{D}})}{(r_{\mathfrak{D}} - s)(r_{\mathfrak{D}} - \rho_{\mathfrak{D}})} - \frac{\hat{\mu}_{\mathfrak{D}}(\rho_{\mathfrak{D}})}{(\rho_{\mathfrak{D}} - s)(r_{\mathfrak{D}} - \rho_{\mathfrak{D}})} + \frac{\hat{\mu}_{\mathfrak{D}}(s)}{(r_{\mathfrak{D}} - s)(\rho_{\mathfrak{D}} - s)} \\
&= 0 - \frac{\lambda(\delta + \lambda - \rho_{\mathfrak{D}})\tilde{P}_{\mathfrak{D}}(\rho_{\mathfrak{D}})}{(\rho_{\mathfrak{D}} - s)(\delta + \lambda - \rho_{\mathfrak{D}})} + \frac{\lambda(\delta + \lambda - s)\tilde{P}_{\mathfrak{D}}(s)}{(\delta + \lambda - s)(\rho_{\mathfrak{D}} - s)} \\
&= \lambda \frac{\tilde{P}_{\mathfrak{D}}(s) - \tilde{P}_{\mathfrak{D}}(\rho_{\mathfrak{D}})}{\rho_{\mathfrak{D}} - s} \\
&= \lambda \, \mathrm{T}_s \, \mathrm{T}_{\rho_{\mathfrak{D}}} \, p_{\mathfrak{D}}(0).
\end{aligned}$$

Therefore also $\tilde{G}_{\mathfrak{D}}(s) = \frac{1-\delta \upsilon_{\mathfrak{D}}}{1-\mathrm{T}_s \, \mathrm{T}_{r_{\mathfrak{D}}} \, \mathrm{T}_{\rho_{\mathfrak{D}}} \, \mu_{\mathfrak{D}}(0)} = \frac{1-\phi_{\mathfrak{D}}(0)}{1-\lambda \, \mathrm{T}_s \, \mathrm{T}_{\rho_{\mathfrak{D}}} \, p_{\mathfrak{D}}(0)}$. □

2.5. QUANTIFYING THE UTILITY OF CORRECTED PHASE-TYPE APPROXIMATIONS OF GERBER-SHIU FUNCTIONS

Whereas in Section 2.4 we gave the basis for CPTA of Gerber-Shiu functions, now we turn to demonstrating that our approximations are useful in more than the special case of GSFs which Vatamidou *et al.* (2013) considered. We examine this from two angles: first, the asymptotic tail behavior of our CPTA for large initial capital u, and then the error of the phase-type approximations (meaning without the respective correction terms) as the perturbation parameter ϵ goes to 0.

2.5.1. Asymptotic Tail Behavior. Because phase-type approximations of heavy-tailed ruin probabilities inherently fail in capturing the correct behavior in the tail (see for example Vatamidou *et al.* (2014b)), Vatamidou *et al.* (2013) proposed adding the correction term. In their Theorems 5, 6, and 7, they showed that the correction term would properly capture the heavy-tailedness of the exact value (up to multiplication by a constant). We consider such a property important for the correction term to hold any usefulness when claims are heavy-tailed. Therefore, we establish this in our Propositions 1, 2, and 3 below.

In Vatamidou *et al.* (2013, 2014a), they formally showed corrected phase-type approximations to capture the heavy-tailed behavior only in the classical risk model. Similarly, in this subsection, we set $\lambda_1 = \lambda_2 = \lambda$ and $c_1 = c_2 = 1$. Then for all $\delta \geq 0$, we may use Corollary 1 to reduce Definition 6 to:

$$m_{d,\epsilon}(u) = m_\epsilon^\bullet(u) + \frac{\epsilon\lambda}{1-\phi_\epsilon^\bullet(0)} \int_{0-}^{u} \mathrm{T}_{\rho_\epsilon^\bullet}(\omega_c + c * m_\epsilon^\bullet - m_\epsilon^\bullet)(u-x)\,\mathrm{d}G_\epsilon^\bullet(x),$$
$$m_{r,\epsilon}(u) = m_0(u) + \frac{\epsilon\lambda}{1-\phi_0(0)} \int_{0-}^{u} \mathrm{T}_{\rho_0}(\omega_c + c * m_0 - \omega_b - b * m_0)(u-x)\,\mathrm{d}G_0(x).$$

To derive the asymptotic tail behavior of $m_{d,\epsilon}(u)$ and $m_{r,\epsilon}(u)$, we place assumptions both on the claims laws $P_{\mathfrak{D}}(x)$, and on the choice of penalty $w(\cdot,\cdot)$ through the function $\omega_{\mathfrak{D}}(x)$. We will speak of the quantities based upon the rv C in terms of the density classes $\mathscr{L}_d(\alpha)$ and $\mathscr{S}_d(\alpha)$ as used in Tang and Wei (2010). Namely, a function $f : [0,\infty) \to [0,\infty)$, measurable and eventually positive, belongs to $\mathscr{L}_d(\alpha)$, $\alpha \geq 0$, if $\lim_{x\to\infty} \frac{f(x-y)}{f(x)} = e^{\alpha y}, \forall y \in \mathbb{R}$. If also $\lim_{x\to\infty} \frac{f^{*2}(x)}{f(x)} = 2\hat{f}(-\alpha)$, then $f \in \mathscr{S}_d(\alpha)$. Assumption 1 gives the constraints that we place on the distributions of B and C, as well as on their associated functions $\omega_b(u)$ and $\omega_c(u)$.

Assumption 1. *The distributions $B(u)$ and $C(u)$ are both absolutely continuous, with bounded densities $b(u)$ and $c(u)$, $b(u)$ phase-type and $c(u)$ heavy-tailed, respectively. The functions $\omega_b(u)$ and $\omega_c(u)$ are locally integrable, and also globally integrable if $\delta = \alpha = 0$.*

1. *When $\delta > 0$, $c(u) \in \mathscr{S}_d(0)$ and $\omega_c(u) \in \mathscr{L}_d(\alpha)$. When furthermore $\alpha = 0$: if $\omega_c(u) \in \mathscr{S}_d(0)$, then $c(u) = O(\omega_c(u))$; if $\omega_c(u) \in \mathscr{L}_d(0)$ only, then $\omega_c(u) = O(c(u))$.*

2. *When $\delta = 0$, $c(u)$ is eventually non-increasing, and $\bar{C}(u) \in \mathscr{S}_d(0)$. When $\alpha > 0$, $\omega_c(u) \in \mathscr{L}_d(\alpha)$. When instead $\alpha = 0$: $\omega_c(u)$ is eventually non-increasing, and $\bar{\Omega}_c(u) \in \mathscr{L}_d(0)$. If $\bar{\Omega}_c(u) \in \mathscr{S}_d(0)$ also, then $\bar{C}(u) = O\Big(\bar{\Omega}_c(u)\Big)$; otherwise, $\bar{\Omega}_c(u) = O\Big(\bar{C}(u)\Big)$.*

By some discussion in (Asmussen and Albrecher, 2010, Section IX.1), $b(u) \sim Au^k e^{-\gamma u}$ with $k \in \{0, 1, 2, \dots\}$ and $A, \gamma > 0$. Therefore, $b(u) \in \mathscr{L}_d(\gamma)$. By (Tang and Wei, 2010, Lemma 4.1 (2)), $b(u) = O\left(e^{-\tilde{\gamma} u}\right)$ for all $\tilde{\gamma} \in (0, \gamma)$: just set $\tilde{\gamma} := \gamma - \eta, \eta \in (0, \gamma)$, in which $\eta > 0$ has the role which ϵ does in (Tang and Wei, 2010, Lemma 4.1 (2)). From (Foss *et al.*, 2013, Lemma 2.17), $g(u) = o(h(u))$ for $g(u) = O(e^{-\gamma u})$ and $h(u) \in \mathscr{L}_d(0)$, so $b(u) = o(c(u))$ when $\delta > 0$. Since $b(u)$ is phase-type, basic phase-type properties (Asmussen and Albrecher (2010) or Bladt and Nielsen (2017)) mean that $\bar{B}(u)$ is phase-type also, with the same matrix $\underline{T}$; in other words, $\bar{B}(u) \sim A' u^k e^{-\gamma u}$, with $k \in \{0, 1, 2, \dots\}$, and $A', \gamma > 0$ and so $\bar{B}(u) \in \mathscr{L}_d(\gamma)$. So under Assumption 1, when $\delta = 0$, $\bar{B}(u) = o\left(\bar{C}(u)\right)$ follows by the same reasoning as above for $b(u) = o(c(u))$ (but see Vatamidou *et al.* (2013) for an alternate demonstration).

Observe that $b(u) = o(c(u))$ implies $\omega_b(u) = o(\omega_c(u))$. Letting $k_0 > 0$ be arbitrary, choose u_0 such that $u > u_0$ implies $b(u) \leq k_0 c(u)$. Thus, for $u > u_0$, since $x > 0$ means also $u + x > u_0$,

$$\omega_b(u) = \int_0^\infty w(u, x) b(u + x)\,\mathrm{d}x \leq \int_0^\infty w(u, x) k_0 c(u + x)\,\mathrm{d}x = k_0 \omega_c(u).$$

So the same k_0 and u_0 also correspond to the statement $\omega_b(u) = o(\omega_c(u))$. It also holds that $\bar{\Omega}_b(u) = o\left(\bar{\Omega}_c(u)\right)$ because $b(u) = o(c(u))$. Letting $k_0 > 0$, choose u_0 such that $u > u_0$ implies $b(u) \leq k_0 c(u)$. Then, for $u > u_0$,

$$\bar{\Omega}_b(u) = \int_0^\infty \omega_b(u + x)\,\mathrm{d}x \leq \int_0^\infty k_0 \omega_c(u + x)\,\mathrm{d}x = k_0 \bar{\Omega}_c(u).$$

We will frequently use some basic properties about the operator T on functions in $\mathscr{L}_d(\alpha)$ in proving Propositions 1, 2, and 3; the following lemma establishes said properties. Note that $\alpha \vee \rho = \max(\alpha, \rho)$.

Lemma 4. *Let $\rho \geq 0$ and $\alpha \geq 0$; assume that $h(u) \in \mathscr{L}_d(\alpha)$. We have: if $\rho > 0$, then $\mathrm{T}_\rho\, h(u) \in \mathscr{L}_d(\alpha)$; and if $\alpha \vee \rho > 0$, then $\mathrm{T}_\rho\, h(u) \sim \frac{1}{\rho + \alpha} h(u)$.*

Proof of Lemma 4. First we show $\mathrm{T}_\rho\, h(u) \in \mathscr{L}_d(\alpha)$. For $\rho > 0$, let $y \in \mathbb{R}$. Canceling $e^{\rho u}$, then using L'Hopital's rule, followed by applying $h(u) \in \mathscr{L}_d(\alpha)$, we easily see that $\lim\limits_{u\to\infty} \frac{\mathrm{T}_\rho\, h(u-y)}{\mathrm{T}_\rho\, h(u)} = e^{\alpha y}$. Namely,

$$\begin{aligned}
\lim_{u\to\infty} \frac{\mathrm{T}_\rho\, h(u-y)}{\mathrm{T}_\rho\, h(u)} &= \lim_{u\to\infty} \frac{\int_{u-y}^{\infty} e^{-\rho(v-(u-y))} h(v)\,\mathrm{d}v}{\int_u^\infty e^{-\rho(v-u)} h(v)\,\mathrm{d}v} \\
&= \lim_{u\to\infty} \frac{\int_u^\infty e^{-\rho(w-u)} h(w-y)\,\mathrm{d}w}{\int_u^\infty e^{-\rho(v-u)} h(v)\,\mathrm{d}v} \\
&= \lim_{u\to\infty} \frac{\int_u^\infty e^{-\rho w} h(w-y)\,\mathrm{d}w}{\int_u^\infty e^{-\rho v} h(v)\,\mathrm{d}v} \\
&= \lim_{u\to\infty} \frac{-e^{-\rho u} h(u-y)}{-e^{-\rho u} h(u)} \\
&= \lim_{u\to\infty} \frac{h(u-y)}{h(u)} = e^{\alpha y}.
\end{aligned}$$

For the second assertion, notice that $\lim\limits_{u\to\infty} \frac{\mathrm{T}_\rho\, h(u)}{h(u)} = \lim\limits_{u\to\infty} \int_0^\infty e^{-\rho v} \frac{h(u+v)}{h(u)}\,\mathrm{d}v$. By (Tang and Wei, 2010, Lemma 4.1 (1)), for all $\eta > 0$, there exist constants $c_0 > 0$ and $u_0 > 0$ such that for all $u \geq u_0$ and $y \geq 0$, $\frac{h(u+y)}{h(u)} \leq c_0 e^{-(\alpha-\eta)y}$. So, choosing $\eta \in (0, \rho+\alpha)$, we have $\int_0^\infty e^{-\rho v} \frac{h(u+v)}{h(u)}\,\mathrm{d}v \leq \int_0^\infty e^{-\rho v} c_0 e^{-(\alpha-\eta)v}\,\mathrm{d}v < \infty$. Then by the dominated convergence theorem, $\lim\limits_{u\to\infty} \frac{\mathrm{T}_\rho\, h(u)}{h(u)} = \int_0^\infty e^{-\rho v} \lim\limits_{u\to\infty} \frac{h(u+v)}{h(u)}\,\mathrm{d}v = \int_0^\infty e^{-\rho v} e^{-\alpha v}\,\mathrm{d}v$. $\square$

Either $\alpha = 0$ or $\alpha > 0$ may occur in Propositions 1, 2, and 3. For, consider the penalty function $w(x, y) = e^{-s_1 x - s_2 y}$, where $s_1 \in [0, \infty)$ and $s_2 \in (0, \infty)$. The authors in Tang and Wei (2010) discussed that $p_\epsilon(u) \in \mathscr{L}_d(0)$ implies $\omega_\epsilon(u) \sim \frac{1}{s_2} e^{-s_1 u} p_\epsilon(u)$; from this it follows that $\omega_\epsilon(u) \in \mathscr{L}_d(s_1)$. Each asymptotic relation in Propositions 1, 2, and 3 is meant in the limit $u \to \infty$.

Proposition 1. *Let Assumption 1 hold; then*

$$m_\epsilon(u) \sim \begin{cases} \dfrac{\epsilon\lambda^2}{\delta^2}\rho_\epsilon \hat{\hat{\Omega}}_\epsilon(\rho_\epsilon)c(u) & \text{if } \delta > 0, \alpha > 0 \\ \dfrac{\epsilon\lambda^2}{\delta^2}\rho_\epsilon \hat{\hat{\Omega}}_\epsilon(\rho_\epsilon)c(u) + \dfrac{\epsilon\lambda}{\delta}\omega_c(u) & \text{if } \delta > 0, \alpha = 0 \\ \dfrac{\epsilon\lambda^2}{(1-\psi_\epsilon(0))^2}\hat{\hat{\Omega}}_\epsilon(0)\bar{C}(u) & \text{if } \delta = 0, \alpha > 0 \\ \dfrac{\epsilon\lambda^2}{(1-\psi_\epsilon(0))^2}\hat{\hat{\Omega}}_\epsilon(0)\bar{C}(u) + \dfrac{\epsilon\lambda}{1-\psi_\epsilon(0)}\bar{\Omega}_c(u) & \text{if } \delta = 0, \alpha = 0 \end{cases}$$

holds.

Proof of Proposition 1. To establish Proposition 1, we need to show that our assumptions on $c(u)$ and $\omega_c(u)$ imply the conditions (Tang and Wei, 2010, Corollary 3.2 (2, 3, 5, 6)) placed on $p_\epsilon(u)$ and $\omega_\epsilon(u)$ are met. Then our result rather easily follows by way of $p_\epsilon(u) \sim \epsilon c(u)$ and $\omega_\epsilon(u) \sim \epsilon\omega_c(u)$. That the quantity $\lambda \int_0^\infty \int_0^\infty e^{-\rho_{\text{פ}}} p_{\text{פ}}(x+y)\mathrm{d}x\mathrm{d}y$ in equation (3.19) of Tang and Wei (2010) is less than 1 obviously holds in the mixture model. For $\delta \geq 0$, by $\omega_b(u) = o(\omega_c(u))$, $\omega_c(u) \in \mathscr{L}_d(\alpha)$ implies $\omega_\epsilon(u) \in \mathscr{L}_d(\alpha)$ for all $\alpha \geq 0$. Letting $y \in \mathbb{R}$,

$$\begin{aligned} \lim_{u\to\infty} \frac{\omega_\epsilon(u-y)}{\omega_\epsilon(u)} &= \lim_{u\to\infty} \frac{(1-\epsilon)\omega_b(u-y) + \epsilon\omega_c(u-y)}{(1-\epsilon)\omega_b(u) + \epsilon\omega_c(u)} \\ &= \lim_{u\to\infty} \frac{(1-\epsilon)o(\omega_c(u-y)) + \epsilon\omega_c(u-y)}{(1-\epsilon)o(\omega_c(u)) + \epsilon\omega_c(u)} \\ &= \lim_{u\to\infty} \frac{((1-\epsilon)o(1) + \epsilon)\omega_c(u-y)}{((1-\epsilon)o(1) + \epsilon)\omega_c(u)} = e^{\alpha y}. \end{aligned} \tag{2.15}$$

First consider $\delta > 0$; by (Klüppelberg, 1989, Lemma 1.2), $p_\epsilon(u) \in \mathscr{S}_d(0)$ follows from $c(u) \in \mathscr{S}_d(0)$ and $b(u) = o(c(u))$. Because $b(u) = o(c(u))$ and $c(u) \in \mathscr{L}_d(0)$, $p_\epsilon(u) \in \mathscr{L}_d(0)$ in the sense of (2.15). As probability density functions, $c(u)$ and $p_\epsilon(u)$ are globally, and thus locally, integrable. Then for $\alpha > 0$, Proposition 1 follows from (Tang and Wei, 2010, Corollary 3.2 (2)). That is,

$$m_\epsilon(u) \sim \frac{\lambda^2}{\delta^2}\rho_\epsilon \hat{\hat{\Omega}}_\epsilon(\rho_\epsilon)p_\epsilon(u)$$

$$= \frac{\lambda^2}{\delta^2}\rho_\epsilon \hat{\hat{\Omega}}_\epsilon(\rho_\epsilon)((1-\epsilon)o(c(u)) + \epsilon c(u))$$
$$\sim \frac{\epsilon\lambda^2}{\delta^2}\rho_\epsilon \hat{\hat{\Omega}}_\epsilon(\rho_\epsilon)c(u).$$

On the other hand, for $\delta > 0$ and $\alpha = 0$, because $\omega_b(u) = o(\omega_c(u))$ and we assumed $\omega_b(u)$ and $\omega_c(u)$ locally integrable, $\omega_c(u) \in \mathscr{S}_d(0)$ implies $\omega_\epsilon(u) \in \mathscr{S}_d(0)$ by (Klüppelberg, 1989, Lemma 1.2). Specifically, $\lim_{u\to\infty} \omega_\epsilon(u)/\omega_c(u) = \epsilon$ because $\omega_b(u) = o(\omega_c(u))$. The relation $c(u) = O(\omega_c(u))$ implies $p_\epsilon(u) = O(\omega_\epsilon(u))$ because $b(u) = o(c(u))$. Choose $k_0 > 0$ and u_0 such that $u > u_0$ implies $c(u) \le k_0\omega_c(u)$; also choose u_1 such that $u > u_1$ implies $b(u) \le \frac{\epsilon}{1-\epsilon}c(u)$. Letting $u > u_0 \vee u_1$,

$$p_\epsilon(u) = (1-\epsilon)b(u) + \epsilon c(u) \le 2\epsilon c(u) \le 2k_0\epsilon\omega_c(u)$$
$$\le 2k_0((1-\epsilon)\omega_b(u) + \epsilon\omega_c(u)) = 2k_0\omega_\epsilon(u).$$

Likewise, $\omega_c(u) = O(c(u))$ implies $\omega_\epsilon(u) = O(p_\epsilon(u))$ because $\omega_b(u) = o(\omega_c(u))$. Choose $k_0 > 0$ and u_0 such that $u > u_0$ implies $\omega_c(u) \le k_0 c(u)$; also choose u_1 such that $u > u_1$ implies $\omega_b(u) \le \frac{\epsilon}{1-\epsilon}\omega_c(u)$. Letting $u > u_0 \vee u_1$,

$$\omega_\epsilon(u) = (1-\epsilon)\omega_b(u) + \epsilon\omega_c(u) \le 2\epsilon\omega_c(u) \le 2k_0\epsilon c(u)$$
$$\le 2k_0((1-\epsilon)b(u) + \epsilon c(u)) = 2k_0 p_\epsilon(u).$$

Therefore, Proposition 1 follows from (Tang and Wei, 2010, Corollary 3.2 (3)). That is,

$$m_\epsilon(u) \sim \frac{\lambda}{\delta}\omega_\epsilon(u) + \frac{\lambda^2}{\delta^2}\rho_\epsilon \hat{\hat{\Omega}}_\epsilon(\rho_\epsilon)p_\epsilon(u)$$
$$= (1-\epsilon)o(1)\left(\frac{\lambda}{\delta}\omega_c(u) + \frac{\lambda^2}{\delta^2}\rho_\epsilon \hat{\hat{\Omega}}_\epsilon(\rho_\epsilon)c(u)\right) + \epsilon\left(\frac{\lambda}{\delta}\omega_c(u) + \frac{\lambda^2}{\delta^2}\rho_\epsilon \hat{\hat{\Omega}}_\epsilon(\rho_\epsilon)c(u)\right)$$
$$\sim \frac{\epsilon\lambda}{\delta}\omega_c(u) + \frac{\epsilon\lambda^2}{\delta^2}\rho_\epsilon \hat{\hat{\Omega}}_\epsilon(\rho_\epsilon)c(u).$$

Now we let $\delta = 0$: the density $p_\epsilon(u)$ is eventually non-increasing because we assume so for $c(u)$, and $b(u)$ is phase-type. As commented previously, $\bar{B}(u) = o\big(\bar{C}(u)\big)$; thus $\bar{C}(u) \in \mathscr{S}_d(0)$ implies $\bar{P}_\epsilon(u) \in \mathscr{S}_d(0)$ by (Klüppelberg, 1989, Lemma 1.2). In the sense of (2.15), $\bar{P}_\epsilon(u) \in \mathscr{L}_d(0)$, and $\lim\limits_{u\to\infty} \bar{P}_\epsilon(u)\bar{C}(u)^{-1} = \epsilon$. Because we assumed $\eta_\epsilon < \infty$ and $\eta_c < \infty$, the densities $\bar{P}_\epsilon(u)$ and $\bar{C}(u)$ are globally integrable, and locally as well. Then for $\alpha > 0$, Proposition 1 follows from (Tang and Wei, 2010, Corollary 3.2 (5)). Specifically,

$$
\begin{aligned}
m_\epsilon(u) &\sim \frac{\lambda^2}{(1-\psi_\epsilon(0))^2}\hat{\bar{\Omega}}_\epsilon(0)\bar{P}_\epsilon(u) \\
&= \frac{\lambda^2}{(1-\psi_\epsilon(0))^2}\hat{\bar{\Omega}}_\epsilon(0)\Big((1-\epsilon)o\Big(\bar{C}(u)\Big) + \epsilon\bar{C}(u)\Big) \\
&\sim \frac{\epsilon\lambda^2}{(1-\psi_\epsilon(0))^2}\hat{\bar{\Omega}}_\epsilon(0)\bar{C}(u).
\end{aligned}
$$

Meanwhile, when $\delta = \alpha = 0$, having assumed $\bar{\Omega}_c(u) \in \mathscr{L}_d(0)$, we have $\bar{\Omega}_\epsilon(u) \in \mathscr{L}_d(0)$ because $\bar{\Omega}_b(u) = o\big(\bar{\Omega}_c(u)\big)$. Again, this holds in the sense of (2.15) with $\alpha = 0$. Clearly, $\omega_\epsilon(u)$ is eventually non-increasing given the same property for $b(u)$, $c(u)$, and $\omega_c(u)$. The relation $\bar{C}(u) = O\big(\bar{\Omega}_c(u)\big)$ implies $\bar{P}_\epsilon(u) = O\big(\bar{\Omega}_\epsilon(u)\big)$ because $\bar{B}(u) = o\big(\bar{C}(u)\big)$. Choose $k_0 > 0$ and u_0 such that $u > u_0$ implies $\bar{C}(u) \le k_0\bar{\Omega}_c(u)$; also choose u_1 such that $u > u_1$ implies $\bar{B}(u) \le \frac{\epsilon}{1-\epsilon}\bar{C}(u)$. Letting $u > u_0 \vee u_1$,

$$
\begin{aligned}
\bar{P}_\epsilon(u) &= (1-\epsilon)\bar{B}(u) + \epsilon\bar{C}(u) \le 2\epsilon\bar{C}(u) \le 2k_0\epsilon\bar{\Omega}_c(u) \\
&\le 2k_0((1-\epsilon)\bar{\Omega}_b(u) + \epsilon\bar{\Omega}_c(u)) = 2k_0\bar{\Omega}_\epsilon(u).
\end{aligned}
$$

Likewise, $\bar{\Omega}_c(u) = O\big(\bar{C}(u)\big)$ implies $\bar{\Omega}_\epsilon(u) = O\big(\bar{P}_\epsilon(u)\big)$ because $\bar{\Omega}_b(u) = o\big(\bar{\Omega}_c(u)\big)$. Choose $k_0 > 0$ and u_0 such that $u > u_0$ implies $\bar{\Omega}_c(u) \le k_0\bar{C}(u)$; also choose u_1 such that $u > u_1$ implies $\bar{\Omega}_b(u) \le \frac{\epsilon}{1-\epsilon}\bar{\Omega}_c(u)$. Letting $u > u_0 \vee u_1$,

$$
\begin{aligned}
\bar{\Omega}_\epsilon(u) &= (1-\epsilon)\bar{\Omega}_b(u) + \epsilon\bar{\Omega}_c(u) \le 2\epsilon\bar{\Omega}_c(u) \le 2k_0\epsilon\bar{C}(u) \\
&\le 2k_0((1-\epsilon)\bar{B}(u) + \epsilon\bar{C}(u)) = 2k_0\bar{P}_\epsilon(u).
\end{aligned}
$$

Finally, $\bar{\Omega}_b(u) = o\Big(\bar{\Omega}_c(u)\Big)$ with $\omega_b(u)$ and $\omega_c(u)$ globally integrable implies by (Klüppelberg, 1989, Lemma 1.2) that $\bar{\Omega}_\epsilon(u) \in \mathscr{S}_d(0)$ when $\bar{\Omega}_c(u) \in \mathscr{S}_d(0)$. The fact that $\omega_{\text{פ}}(u)$ is globally integrable implies that for all $u_0 > 0$, $\int_0^{u_0} \bar{\Omega}_{\text{פ}}(u)\,\mathrm{d}u = \int_0^{u_0}\int_0^\infty \omega_{\text{פ}}(u+x)\,\mathrm{d}x\,\mathrm{d}u = \int_0^\infty\int_0^{u_0} \omega_{\text{פ}}(u+x)\,\mathrm{d}u\,\mathrm{d}x < \infty$, as discussed in the proof of (Tang and Wei, 2010, Lemma 5.2). In other words, $\bar{\Omega}_b(u)$ and $\bar{\Omega}_c(u)$ are locally integrable. Therefore, Proposition 1 follows from (Tang and Wei, 2010, Corollary 3.2 (6)). That is,

$$
\begin{aligned}
m_\epsilon(u) &\sim \frac{\lambda}{1-\psi_\epsilon(0)}\bar{\Omega}_\epsilon(u) + \frac{\lambda^2}{(1-\psi_\epsilon(0))^2}\hat{\bar{\Omega}}_\epsilon(0)\bar{P}_\epsilon(u) \\
&= (1-\epsilon)o(1)\left(\frac{\lambda}{1-\psi_\epsilon(0)}\bar{\Omega}_c(u) + \frac{\lambda^2}{(1-\psi_\epsilon(0))^2}\hat{\bar{\Omega}}_\epsilon(0)\bar{C}(u)\right) \\
&\quad + \epsilon\left(\frac{\lambda}{1-\psi_\epsilon(0)}\bar{\Omega}_c(u) + \frac{\lambda^2}{(1-\psi_\epsilon(0))^2}\hat{\bar{\Omega}}_\epsilon(0)\bar{C}(u)\right) \\
&\sim \frac{\epsilon\lambda}{1-\psi_\epsilon(0)}\bar{\Omega}_c(u) + \frac{\epsilon\lambda^2}{(1-\psi_\epsilon(0))^2}\hat{\bar{\Omega}}_\epsilon(0)\bar{C}(u).
\end{aligned}
$$

$\square$

When we consider the corrected discard and corrected replace approximations of the exact-valued Gerber-Shiu function $m_\epsilon(u)$, the asymptotic tail behavior retains the same form. The one difference is that the constant coefficients change their associated claims distribution: instead of $P_\epsilon(u)$, we find $P_\epsilon^\bullet(u)$ in Proposition 2 and $P_0(u)$ in Proposition 3. That is, for $m_\epsilon(u)$, in Proposition 1 the relevant coefficients follow the mixture-model distribution $P_\epsilon(u)$, whereas for the corrected phase-type approximations, in Propositions 2 and 3 those coefficients follow the corresponding base-model phase-type distribution.

Proposition 2. *Let Assumption 1 hold; then*

$$m_{d,\epsilon}(u) \sim \begin{cases} \dfrac{\epsilon\lambda^2}{\delta^2}\rho_\epsilon^\bullet \hat{\bar{\Omega}}_\epsilon^\bullet(\rho_\epsilon^\bullet)c(u) & \text{if } \delta > 0, \alpha > 0 \\ \dfrac{\epsilon\lambda^2}{\delta^2}\rho_\epsilon^\bullet \hat{\bar{\Omega}}_\epsilon^\bullet(\rho_\epsilon^\bullet)c(u) + \dfrac{\epsilon\lambda}{\delta}\omega_c(u) & \text{if } \delta > 0, \alpha = 0 \\ \dfrac{\epsilon\lambda^2}{(1-\psi_\epsilon^\bullet(0))^2}\hat{\bar{\Omega}}_\epsilon^\bullet(0)\bar{C}(u) & \text{if } \delta = 0, \alpha > 0 \\ \dfrac{\epsilon\lambda^2}{(1-\psi_\epsilon^\bullet(0))^2}\hat{\bar{\Omega}}_\epsilon^\bullet(0)\bar{C}(u) + \dfrac{\epsilon\lambda}{1-\psi_\epsilon^\bullet(0)}\bar{\Omega}_c(u) & \text{if } \delta = 0, \alpha = 0 \end{cases}$$

holds.

Proof of Proposition 2. By (Tang and Wei, 2010, Lemmas 5.1, 5.2), $m_\epsilon^\bullet(u)$ is locally integrable. For phase-type density $b(u)$, it is easy to show that $\mathrm{T}_\rho\, b(u)$ is also phase-type. Let $\Re(\rho) \geq 0$. Suppose $b(u) = \underline{\alpha} e^{\underline{T}u}\underline{t}$; then

$$\begin{aligned} \mathrm{T}_\rho\, b(u) &= \int_u^\infty e^{-\rho(y-u)}\underline{\alpha} e^{\underline{T}y}\underline{t}\,\mathrm{d}y \\ &= e^{\rho u}\int_u^\infty \underline{\alpha} e^{-(\rho\underline{I}-\underline{T})y}\underline{t}\,\mathrm{d}y \\ &= e^{\rho u}\underline{\alpha}(\rho\underline{I}-\underline{T})^{-1}e^{-(\rho\underline{I}-\underline{T})u}\underline{t} \\ &= \underline{\alpha}(\rho\underline{I}-\underline{T})^{-1}e^{\underline{T}u}\underline{t}. \end{aligned}$$

That means, for a phase-type distribution $b(u)$ with representation $PH(\underline{\alpha}, \underline{T})$, $\mathrm{T}_\rho\, b(u)$ has representation $PH\left(\underline{\alpha}(\rho\underline{I}-\underline{T})^{-1}, \underline{T}\right)$. Since all eigenvalues of $\underline{T}$ are located strictly in the negative half plane (e.g., (Bladt and Nielsen, 2017, Corollary 3.1.15)) and $\Re(\rho) \geq 0$, the matrix $(\rho\underline{I}-\underline{T})^{-1}$ is well-defined. It is known (e.g., (Lin and Willmot, 1999, Section 2)) that the base distribution of the compound geometric df $G_{\mathfrak{D}}(u) = 1 - m_{\mathfrak{D}}(u)$ is proportional to the form $\mathrm{T}_\rho\, b(u)$, so by (Bladt and Nielsen, 2017, Theorem 3.1.28), $G_{\mathfrak{D}}(u)$ is phase-type for $\delta \geq 0$ whenever claims are as well. By (Tang and Wei, 2010, Lemma 4.3 (1)),

$\int_{0-}^{u} h(u-x)\,\mathrm{d}G_{\mathfrak{D}}(x) \sim h(u)$ for $h(u) \in \mathscr{S}_d(0)$, or locally integrable $h(u) \in \mathscr{L}_d(0)$. For,

$$\int_{0-}^{u} h(u-x)\,\mathrm{d}G_{\mathfrak{D}}(x) = \int_{0}^{u} h(u-x)\left\{-\bar{G}_{\mathfrak{D}}^{(1)}(x)\right\}\mathrm{d}x + h(u)G_{\mathfrak{D}}(0)$$
$$\sim h(u)\Big(-\Big(0 \cdot \hat{\bar{G}}_{\mathfrak{D}}(0) - \bar{G}_{\mathfrak{D}}(0)\Big) + G_{\mathfrak{D}}(0)\Big).$$

It follows from (Gerber and Shiu, 1998, Equation 4.7) that $m_{\epsilon}^{\bullet}(u) = O\Big(e^{-R_{\epsilon}^{\bullet}u}\Big)$, in which $-R_{\epsilon}^{\bullet} < 0$ is the negative root of $l_{\epsilon}^{\bullet}(s) = 0$. The significance is that $m_{\epsilon}^{\bullet}(u)$ is $o(\cdot)$ of any function in $\mathscr{L}_d(0)$.

First we let $\delta > 0$: by (Tang and Wei, 2010, Lemma 4.3 (1)), $c * m_{\epsilon}^{\bullet}(u) \in \mathscr{L}_d(0)$, and $c * m_{\epsilon}^{\bullet}(u) \sim c(u)\hat{m}_{\epsilon}^{\bullet}(0)$; by (Klüppelberg, 1989, Lemma 1.2), $c * m_{\epsilon}^{\bullet}(u) \in \mathscr{S}_d(0)$ also. Specifically, if $y \in \mathbb{R}$, $c * m_{\epsilon}^{\bullet}(u-y) \sim c(u-y)\hat{m}_{\epsilon}^{\bullet}(0) \sim c(u)\hat{m}_{\epsilon}^{\bullet}(0) \sim c * m_{\epsilon}^{\bullet}(u)$. The latter half (that $c(u)\hat{m}_{\epsilon}^{\bullet}(0) \sim c * m_{\epsilon}^{\bullet}(u)$) demonstrates the subexponentiality of $c * m_{\epsilon}^{\bullet}(u)$ due to that of $c(u)$. By Lemma 4 and (Klüppelberg, 1989, Lemma 1.2), $\mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, c * m_{\epsilon}^{\bullet}(u) \in \mathscr{S}_d(0)$ itself. For, $\mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, c * m_{\epsilon}^{\bullet}(u) \in \mathscr{L}_d(0)$ and $\mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, c * m_{\epsilon}^{\bullet}(u) \sim \frac{1}{\rho_{\epsilon}^{\bullet}} c * m_{\epsilon}^{\bullet}(u)$. One may show that $\int_{0-}^{u} \mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, m_{\epsilon}^{\bullet}(u-x)\,\mathrm{d}G_{\epsilon}^{\bullet}(x) = O\Big(e^{-R_{\epsilon}^{\bullet}u}\Big)$. It follows from $m_{\epsilon}^{\bullet}(u) = O\Big(e^{-R_{\epsilon}^{\bullet}u}\Big)$ that $\mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, m_{\epsilon}^{\bullet}(u) = O\Big(e^{-R_{\epsilon}^{\bullet}u}\Big)$ also. Choose $k_0 > 0$ and u_0 such that $u > u_0$ implies $m_{\epsilon}^{\bullet}(u) \le k_0 e^{-R_{\epsilon}^{\bullet}u}$; then

$$\mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, m_{\epsilon}^{\bullet}(u) = \int_{0}^{\infty} e^{-\rho_{\epsilon}^{\bullet}v} m_{\epsilon}^{\bullet}(u+v)\,\mathrm{d}v \le k_0 \int_{0}^{\infty} e^{-\rho_{\epsilon}^{\bullet}v} e^{-R_{\epsilon}^{\bullet}(u+v)}\,\mathrm{d}v = \frac{k_0}{\rho_{\epsilon}^{\bullet} + R_{\epsilon}^{\bullet}} e^{-R_{\epsilon}^{\bullet}u}.$$

Since $-\bar{G}_{\epsilon}^{\bullet(1)}(u) \in \mathscr{L}_d(\gamma)$ for some $\gamma > 0$ by having a phase-type distribution, by Lemma 4 and (Tang and Wei, 2010, Lemma 4.1 (2)) we have that $-\bar{G}_{\epsilon}^{\bullet(1)}(u) := \mathrm{g}_{\epsilon}^{\bullet}(u) = O\Big(e^{-\tilde{\gamma}u}\Big)$ for some $\tilde{\gamma} \in (0, \gamma)$. Now, we will choose $\tilde{\gamma} \in (R_{\epsilon}^{\bullet}, \gamma)$. Choose $k_0 > 0$ and u_0 such that $u > u_0$ implies $\mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, m_{\epsilon}^{\bullet}(u) \le k_0 e^{-R_{\epsilon}^{\bullet}u}$, and choose $k_1 > 0$ and u_1 such that $u > u_1$ implies $\mathrm{g}_{\epsilon}^{\bullet}(u) \le k_1 e^{-\tilde{\gamma}u}$. In what immediately follows, we let $u > 2(u_0 \vee u_1)$:

$$\int_{0-}^{u} \mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, m_{\epsilon}^{\bullet}(u-x)\,\mathrm{d}G_{\epsilon}^{\bullet}(x) = G_{\epsilon}^{\bullet}(0)\,\mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, m_{\epsilon}^{\bullet}(u) + \int_{0}^{u} \mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, m_{\epsilon}^{\bullet}(u-x)\mathrm{g}_{\epsilon}^{\bullet}(x)\,\mathrm{d}x$$
$$= G_{\epsilon}^{\bullet}(0)\,\mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, m_{\epsilon}^{\bullet}(u) + \left(\int_{0}^{u-(u_0\vee u_1)} + \int_{u-(u_0\vee u_1)}^{u}\right) \mathrm{T}_{\rho_{\epsilon}^{\bullet}}\, m_{\epsilon}^{\bullet}(u-x)\mathrm{g}_{\epsilon}^{\bullet}(x)\,\mathrm{d}x$$

$$
\begin{aligned}
&\leq k_2 e^{-R^\bullet_\epsilon u} + \int_0^{u-(u_0\vee u_1)} k_0 e^{-R^\bullet_\epsilon (u-x)} \mathfrak{g}^\bullet_\epsilon(x)\,\mathrm{d}x + \int_{u-(u_0\vee u_1)}^{u} k_1 e^{-\tilde{\gamma}x}\,\mathrm{T}_{\rho^\bullet_\epsilon}\, m^\bullet_\epsilon(u-x)\,\mathrm{d}x \\
&= k_2 e^{-R^\bullet_\epsilon u} + k_0 e^{-R^\bullet_\epsilon u} \int_0^{u-(u_0\vee u_1)} e^{R^\bullet_\epsilon x} \mathfrak{g}^\bullet_\epsilon(x)\,\mathrm{d}x + k_1 \int_0^{(u_0\vee u_1)} e^{-\tilde{\gamma}(u-v)}\,\mathrm{T}_{\rho^\bullet_\epsilon}\, m^\bullet_\epsilon(v)\,\mathrm{d}v \\
&\leq \left(k_2 + k_0 \hat{\mathfrak{g}}^\bullet_\epsilon(-R^\bullet_\epsilon)\right) e^{-R^\bullet_\epsilon u} + k_1 e^{-\tilde{\gamma}u} \int_0^{u_0\vee u_1} e^{\tilde{\gamma}v}\,\mathrm{T}_{\rho^\bullet_\epsilon}\, m^\bullet_\epsilon(v)\,\mathrm{d}v \leq k_3 e^{-R^\bullet_\epsilon u}. \qquad (2.16)
\end{aligned}
$$

Therefore, $\int_{0-}^{u} \mathrm{T}_{\rho^\bullet_\epsilon}\, m^\bullet_\epsilon(u-x)\,\mathrm{d}G^\bullet_\epsilon(x) = O\left(e^{-R^\bullet_\epsilon u}\right)$. Now, suppose $\alpha > 0$: by Lemma 4, $\mathrm{T}_{\rho^\bullet_\epsilon}\, \omega_c(u) \in \mathscr{L}_d(\alpha)$; so by (Tang and Wei, 2010, Lemma 4.1 (2)), $\mathrm{T}_{\rho^\bullet_\epsilon}\, \omega_c(u) = O\left(e^{-\tilde{\alpha}u}\right)$, $\forall \tilde{\alpha} \in (0,\alpha)$. Then choose $\tilde{\alpha} \in (0, \alpha \wedge \gamma)$ where $-\bar{G}^{\bullet(1)}_\epsilon(u) = O(e^{-\gamma u})$; in the sense of (2.16), it follows that $\int_{0-}^{u} \mathrm{T}_{\rho^\bullet_\epsilon}\, \omega_c(u-x)\,\mathrm{d}G^\bullet_\epsilon(x) = O\left(e^{-\tilde{\alpha}u}\right)$. Therefore, $m_{d,\epsilon}(u) \sim \frac{\epsilon\lambda}{1-\phi^\bullet_\epsilon(0)}\,\mathrm{T}_{\rho^\bullet_\epsilon}\, c * m^\bullet_\epsilon(u)$; Proposition 2 follows for $\delta > 0$, $\alpha > 0$ by Lemma 4, (Tang and Wei, 2010, Lemma 4.3 (1)), and the relations $\delta = \rho_{\mathfrak{D}}(1-\phi_{\mathfrak{D}}(0))$ and $\hat{m}_{\mathfrak{D}}(0) = \frac{\lambda}{\delta}\rho_{\mathfrak{D}}\hat{\hat{\Omega}}_{\mathfrak{D}}(\rho_{\mathfrak{D}})$. In particular,

$$
\begin{aligned}
m_{d,\epsilon}(u) &\sim \frac{\epsilon\lambda}{1-\phi^\bullet_\epsilon(0)}\,\mathrm{T}_{\rho^\bullet_\epsilon}\, c * m^\bullet_\epsilon(u) \sim \frac{\epsilon\lambda}{1-\phi^\bullet_\epsilon(0)}\frac{1}{\rho^\bullet_\epsilon} c * m^\bullet_\epsilon(u) \\
&\sim \frac{\epsilon\lambda}{\delta}\hat{m}^\bullet_\epsilon(0)c(u) = \frac{\epsilon\lambda^2}{\delta^2}\rho^\bullet_\epsilon\hat{\hat{\Omega}}^\bullet_\epsilon(\rho^\bullet_\epsilon)c(u).
\end{aligned}
$$

Now suppose $\alpha = 0$: $\mathrm{T}_{\rho^\bullet_\epsilon}\, \omega_c(u) \in \mathscr{L}_d(0)$ by Lemma 4. To get Proposition 2 when $\delta > 0$ and $\alpha = 0$, apply Lemma 4 again; whether $\omega_c(u) = O(c(u))$ or $c(u) = O(\omega_c(u))$, one may show that $\omega_c(u) + c * m^\bullet_\epsilon(u) \sim \omega_c(u) + c(u)\hat{m}^\bullet_\epsilon(0)$ by (Tang and Wei, 2010, Lemma 4.3 (1)). That is,

$$
\begin{aligned}
m_{d,\epsilon}(u) &\sim \frac{\epsilon\lambda}{1-\phi^\bullet_\epsilon(0)}\,\mathrm{T}_{\rho^\bullet_\epsilon}(\omega_c + c * m^\bullet_\epsilon)(u) \sim \frac{\epsilon\lambda}{\rho^\bullet_\epsilon(1-\phi^\bullet_\epsilon(0))}(\omega_c + c * m^\bullet_\epsilon)(u) \\
&= \frac{\epsilon\lambda}{\delta}(\omega_c + c * m^\bullet_\epsilon)(u) \sim \frac{\epsilon\lambda}{\delta}(\omega_c(u) + \hat{m}^\bullet_\epsilon(0)c(u)) \\
&= \frac{\epsilon\lambda}{\delta}\omega_c(u) + \frac{\epsilon\lambda^2}{\delta^2}\rho^\bullet_\epsilon\hat{\hat{\Omega}}^\bullet_\epsilon(\rho^\bullet_\epsilon)c(u).
\end{aligned}
$$

For the assertion that $\omega_c(u) + c * m^\bullet_\epsilon(u) \sim \omega_c(u) + \hat{m}^\bullet_\epsilon(0)c(u)$,

$$\left|\frac{\omega_c(u) + c * m^\bullet_\epsilon(u)}{\omega_c(u) + \hat{m}^\bullet_\epsilon(0)c(u)} - 1\right| = \left|\frac{\frac{c*m^\bullet_\epsilon(u)}{c(u)\hat{m}^\bullet_\epsilon(0)} - 1}{\frac{\omega_c(u)}{c(u)\hat{m}^\bullet_\epsilon(0)} + 1}\right| \leq \left|\frac{c * m^\bullet_\epsilon(u)}{c(u)\hat{m}^\bullet_\epsilon(0)} - 1\right| \to 0 \text{ as } u \to \infty.$$

The inequality follows because the denominator of the second step is greater than 1 for all large u by virtue of $c(u), \omega_c(u) \in \mathscr{L}_d(0)$ being eventually positive. The zero limit of course follows from $c * m^\bullet_\epsilon(u) \sim c(u)\hat{m}^\bullet_\epsilon(0)$.

Now we let $\delta = 0$: we have

$$m_{d,\epsilon}(u) = m^\bullet_\epsilon(u) + \frac{\epsilon\lambda}{1 - \psi^\bullet_\epsilon(0)} \int_{0-}^{u} \left(\bar{\Omega}_c + \bar{C} * m^\bullet_\epsilon\right)(u - x)\, \mathrm{d}M^\bullet_\epsilon(x). \tag{2.17}$$

For, setting $\delta = 0$ also produces $\rho^\bullet_\epsilon = 0$. Then, taking the Laplace transform of $m_{d,\epsilon}(u)$, $\hat{m}_{d,\epsilon}(s) = \hat{m}^\bullet_\epsilon(s) + \frac{\epsilon\lambda}{1-\psi^\bullet_\epsilon(0)} \mathrm{T}_s \mathrm{T}_0(\omega_c + c * m^\bullet_\epsilon - m^\bullet_\epsilon)(0)\tilde{M}^\bullet_\epsilon(s)$; noting that $\hat{c}(0)\hat{m}^\bullet_\epsilon(0) - \hat{m}^\bullet_\epsilon(0) = 0$ and $\frac{1-\hat{c}(s)}{s} = \hat{\bar{C}}(s)$, (2.17) follows. Analogously to the $\delta > 0$ cases, $\bar{C} * m^\bullet_\epsilon(u) \in \mathscr{S}_d(0)$. That is, if $y \in \mathbb{R}$, then $\bar{C} * m^\bullet_\epsilon(u - y) \sim \bar{C}(u - y)\hat{m}^\bullet_\epsilon(0) \sim \bar{C}(u)\hat{m}^\bullet_\epsilon(0) \sim \bar{C} * m^\bullet_\epsilon(u)$. Suppose $\alpha > 0$: by (Tang and Wei, 2010, Lemma 4.1 (2)), $\bar{\Omega}_c(u) = O\left(e^{-\tilde{\alpha}u}\right)$, $\forall \tilde{\alpha} \in (0, \alpha)$. Let $\eta \in (0, \alpha)$; choose u_0 and k_1 such that $u > u_0$ implies $\omega_c(u) \leq k_1 e^{-(\alpha-\eta)u}$. Then, for $u > u_0$,

$$\bar{\Omega}_c(u) = \int_u^\infty \omega_c(y)\, \mathrm{d}y \leq \int_u^\infty k_1 e^{-(\alpha-\eta)y}\, \mathrm{d}y = \frac{k_1}{\alpha - \eta} e^{-(\alpha-\eta)u}.$$

Now set $\tilde{\alpha} = \alpha - \eta$. Let $\gamma > 0$ be the largest number for which $-\bar{M}^{\bullet(1)}_\epsilon(u) = O(e^{-\gamma u})$; choosing $\tilde{\alpha} \in (0, \alpha \wedge \gamma)$ we similarly have $\int_{0-}^u \bar{\Omega}_c(u - x)\, \mathrm{d}M^\bullet_\epsilon(x) = O\left(e^{-\tilde{\alpha}u}\right)$. Choose $k_0, u_0 > 0$ such that $u > u_0$ implies $\bar{\Omega}_c(u) \leq k_0 e^{-\tilde{\alpha}u}$, and $k_1, u_1 > 0$ such that $u > u_1$ implies $\mathfrak{m}^\bullet_\epsilon(u) := -\bar{M}^{\bullet(1)}_\epsilon(u) \leq k_1 e^{-\gamma u}$. Then for $u > 2(u_0 \vee u_1)$,

$$\int_{0-}^u \bar{\Omega}_c(u - x)\, \mathrm{d}M^\bullet_\epsilon(x) = M^\bullet_\epsilon(0)\bar{\Omega}_c(u) + \int_0^u \bar{\Omega}_c(u - x)\mathfrak{m}^\bullet_\epsilon(x)\, \mathrm{d}x$$

$$
\begin{aligned}
&= M^\bullet_\epsilon(0)\bar{\Omega}_c(u) + \left(\int_0^{u-(u_0\vee u_1)} + \int_{u-(u_0\vee u_1)}^{u}\right)\bar{\Omega}_c(u-x)\mathfrak{m}^\bullet_\epsilon(x)\,\mathrm{d}x \\
&\le k_2e^{-\tilde{\alpha}u} + \int_0^{u-(u_0\vee u_1)} k_0e^{-\tilde{\alpha}(u-x)}\mathfrak{m}^\bullet_\epsilon(x)\,\mathrm{d}x + \int_{u-(u_0\vee u_1)}^{u} k_1e^{-\gamma x}\bar{\Omega}_c(u-x)\,\mathrm{d}x \\
&= k_2e^{-\tilde{\alpha}u} + k_0e^{-\tilde{\alpha}u}\int_0^{u-(u_0\vee u_1)} e^{\tilde{\alpha}x}\mathfrak{m}^\bullet_\epsilon(x)\,\mathrm{d}x + k_1\int_0^{(u_0\vee u_1)} e^{-\gamma(u-v)}\bar{\Omega}_c(v)\,\mathrm{d}v \\
&\le \big(k_2 + k_0\hat{\mathfrak{m}}^\bullet_\epsilon(-\tilde{\alpha})\big)e^{-\tilde{\alpha}u} + k_1e^{-\gamma u}\int_0^{u_0\vee u_1} e^{\gamma v}\bar{\Omega}_c(v)\,\mathrm{d}v \le k_3e^{-\tilde{\alpha}u}.
\end{aligned} \tag{2.18}
$$

In the second term of the third line, $x \in (0, u-(u_0\vee u_1))$ means $u - x \in (u_0 \vee u_1, u)$, and in the third term, $x > u - (u_0 \vee u_1) > u_0 \vee u_1$. So $m_{d,\epsilon}(u) \sim \frac{\epsilon\lambda}{1-\psi^\bullet_\epsilon(0)}\bar{C} * m^\bullet_\epsilon(u)$; Proposition 2 follows for $\delta = 0$, $\alpha > 0$ by (Tang and Wei, 2010, Lemma 4.3 (1)) and the relation $\hat{m}_{\mathfrak{D}}(0) = \frac{\lambda}{1-\psi_{\mathfrak{D}}(0)}\hat{\bar{\Omega}}_{\mathfrak{D}}(0)$. That is,

$$
\begin{aligned}
m_{d,\epsilon}(u) &\sim \frac{\epsilon\lambda}{1-\psi^\bullet_\epsilon(0)}\int_{0-}^{u}\bar{C} * m^\bullet_\epsilon(u-x)\,\mathrm{d}M^\bullet_\epsilon(x) \sim \frac{\epsilon\lambda}{1-\psi^\bullet_\epsilon(0)}\bar{C} * m^\bullet_\epsilon(u) \\
&\sim \frac{\epsilon\lambda}{1-\psi^\bullet_\epsilon(0)}\hat{m}^\bullet_\epsilon(0)\bar{C}(u) = \frac{\epsilon\lambda^2}{\left(1-\psi^\bullet_\epsilon(0)\right)^2}\hat{\bar{\Omega}}^\bullet_\epsilon(0)\bar{C}(u).
\end{aligned}
$$

Now suppose $\alpha = 0$: by assuming $\omega_c(u)$ globally integrable, $\bar{\Omega}_c(u)$ is locally integrable (see the proof of (Tang and Wei, 2010, Lemma 5.2)), and we assumed $\bar{\Omega}_c(u) \in \mathscr{L}_d(0)$. Then $m_{d,\epsilon}(u) \sim \frac{\epsilon\lambda}{1-\psi^\bullet_\epsilon(0)}\left(\bar{\Omega}_c + \bar{C} * m^\bullet_\epsilon\right)(u)$; whether $\bar{\Omega}_c(u) = O\left(\bar{C}(u)\right)$ or $\bar{C}(u) = O\left(\bar{\Omega}_c(u)\right)$, one may show by (Tang and Wei, 2010, Lemma 4.3 (1)) that $\bar{\Omega}_c(u)+\bar{C} * m^\bullet_\epsilon(u) \sim \bar{\Omega}_c(u) + \bar{C}(u)\hat{m}^\bullet_\epsilon(0)$, and Proposition 2 follows for $\delta = 0$, $\alpha = 0$. For the assertion that $\bar{\Omega}_c(u) + \bar{C} * m^\bullet_\epsilon(u) \sim \bar{\Omega}_c(u) + \hat{m}^\bullet_\epsilon(0)\bar{C}(u)$,

$$
\left|\frac{\bar{\Omega}_c(u) + \bar{C} * m^\bullet_\epsilon(u)}{\bar{\Omega}_c(u) + \hat{m}^\bullet_\epsilon(0)\bar{C}(u)} - 1\right| = \left|\frac{\frac{\bar{C}*m^\bullet_\epsilon(u)}{\bar{C}(u)\hat{m}^\bullet_\epsilon(0)} - 1}{\frac{\bar{\Omega}_c(u)}{\bar{C}(u)\hat{m}^\bullet_\epsilon(0)} + 1}\right| \le \left|\frac{\bar{C} * m^\bullet_\epsilon(u)}{\bar{C}(u)\hat{m}^\bullet_\epsilon(0)} - 1\right| \to 0 \text{ as } u \to \infty.
$$

The inequality follows because the denominator of the second step is greater than 1 for all large u by virtue of $\bar{C}(u), \bar{\Omega}_c(u) \in \mathscr{L}_d(0)$ being eventually positive. The zero limit of course follows from $\bar{C} * m^\bullet_\epsilon(u) \sim \bar{C}(u)\hat{m}^\bullet_\epsilon(0)$. And finally,

$$m_{d,\epsilon}(u) \sim \frac{\epsilon\lambda}{1-\psi^\bullet_\epsilon(0)} \int_0^u \left(\bar{\Omega}_c + \bar{C} * m^\bullet_\epsilon\right)(u-x)\,\mathrm{d}M^\bullet_\epsilon(x) \sim \frac{\epsilon\lambda}{1-\psi^\bullet_\epsilon(0)}\left(\bar{\Omega}_c + \bar{C} * m^\bullet_\epsilon\right)(u)$$
$$\sim \frac{\epsilon\lambda}{1-\psi^\bullet_\epsilon(0)}\left(\bar{\Omega}_c(u) + \bar{C}(u)\hat{m}^\bullet_\epsilon(0)\right) = \frac{\epsilon\lambda}{1-\psi^\bullet_\epsilon(0)}\bar{\Omega}_c(u) + \frac{\epsilon\lambda^2}{\left(1-\psi^\bullet_\epsilon(0)\right)^2}\hat{\bar{\Omega}}^\bullet_\epsilon(0)\bar{C}(u).$$

□

Proposition 3. *Let Assumption 1 hold; then*

$$m_{r,\epsilon}(u) \sim \begin{cases} \dfrac{\epsilon\lambda^2}{\delta^2}\rho_0\hat{\bar{\Omega}}_0(\rho_0)c(u) & \text{if } \delta > 0, \alpha > 0 \\ \dfrac{\epsilon\lambda^2}{\delta^2}\rho_0\hat{\bar{\Omega}}_0(\rho_0)c(u) + \dfrac{\epsilon\lambda}{\delta}\omega_c(u) & \text{if } \delta > 0, \alpha = 0 \\ \dfrac{\epsilon\lambda^2}{\left(1-\psi_0(0)\right)^2}\hat{\bar{\Omega}}_0(0)\bar{C}(u) & \text{if } \delta = 0, \alpha > 0 \\ \dfrac{\epsilon\lambda^2}{\left(1-\psi_0(0)\right)^2}\hat{\bar{\Omega}}_0(0)\bar{C}(u) + \dfrac{\epsilon\lambda}{1-\psi_0(0)}\bar{\Omega}_c(u) & \text{if } \delta = 0, \alpha = 0 \end{cases}$$

holds.

Establishing Proposition 3 is mostly analogous to establishing Proposition 2, so we omit the details. We simply point out that $\frac{\lambda}{1-\phi_0(0)}\int_{0-}^u \mathrm{T}_{\rho_0}\,\omega_b(u-x)\,\mathrm{d}G_0(x)$ is precisely $m_0(u)$, and showing that $(b * m_0)(u) = O\left(e^{-R_0 u}\right)$ works similarly to (2.16) and (2.18). Also, when $\delta = 0$, we find that $m_{r,\epsilon}(u)$ becomes

$$m_{r,\epsilon}(u) = m_0(u) + \frac{\epsilon\lambda}{1-\psi_0(0)}\int_{0-}^u \left(\bar{\Omega}_c + \bar{C} * m_0 - \bar{\Omega}_b - \bar{B} * m_0\right)(u-x)\,\mathrm{d}M_0(x).$$

2.5.2. The Contribution of the Correction Terms. Another important aspect of determining the helpfulness of the generalized correction terms is the approximation errors with and without those terms. Due to Lemma 3, finding bounds on the errors of $m^i_{d,\epsilon}(u)$

and $m^i_{r,\epsilon}(u)$ appears not straightforward. However, the vantage point of (Vatamidou *et al.*, 2014a, Proposition 3.8) does apply, with the same interpretation. Those authors did not explicitly formulate how to use that approach for the corrected discard approximation, in which the correction term depends on ϵ; we now illustrate this.

Recall that $m_{\epsilon,i}(u)$ is the exact value of the Gerber-Shiu function, while $m^{\bullet}_{\epsilon,i}(u)$ and $m_{0,i}(u)$ are respectively the discard and replace phase-type approximations thereof. We have:

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon}\Big(m_{\epsilon,i}(u) - m^{\bullet}_{\epsilon,i}(u)\Big) = \frac{1}{1 - {}_{\delta}\upsilon_0} \int_{0-}^{u} \mathrm{T}_{r_0}\, \mathrm{T}_{\rho_0}\, \text{ק}_{0,i}(u-x)\, \mathrm{d}G_0(x). \tag{2.19}$$

Similar to the derivation of Theorem 1 (i) and Theorem 1 (ii), we may show that $\hat{m}^{\bullet}_{\epsilon,i}(s) = \hat{m}_{0,i}(s) - \frac{\epsilon}{1-{}_{\delta}\upsilon_0}\, \mathrm{T}_s\, \mathrm{T}_{r_0}\, \mathrm{T}_{\rho_0}\, \text{ב}_{0,i}(0)\tilde{G}_0(s) + O\big(\epsilon^2\big)$. Subtracting this expansion of $\hat{m}^{\bullet}_{\epsilon,i}(s)$ from the "replace" expansion of $\hat{m}_{\epsilon,i}(s)$, we get $\hat{m}_{\epsilon,i}(s) = \hat{m}^{\bullet}_{\epsilon,i}(s) + \frac{\epsilon}{1-{}_{\delta}\upsilon_0}\, \mathrm{T}_s\, \mathrm{T}_{r_0}\, \mathrm{T}_{\rho_0}\, \text{ק}_{0,i}(0)\tilde{G}_0(s) + O\big(\epsilon^2\big)$. From here, establishing (2.19) works just like the method used in proving the existing Proposition 3.8 of Vatamidou *et al.* (2014a).

Notice that the right-hand side of (2.19) is simply the correction term of $m^i_{d,\epsilon}(u)$ with $\epsilon := 0$. That is, as the perturbation parameter ϵ tends to 0, for a given value of initial capital u, the error of the discard approximation $m^{\bullet}_{\epsilon,i}(u)$ converges to the correction term in $m^i_{d,\epsilon}(u)$ with $\epsilon := 0$. The obvious equivalent of (2.19) holds in the case of the corrected replace approximation $m^i_{r,\epsilon}(u)$; then, however, one does not need intermediate perturbation expansions such as $\hat{m}^{\bullet}_{\epsilon,i}(s)$ in terms of $\hat{m}_{0,i}(s)$.

2.6. NUMERICAL ILLUSTRATIONS

We give a numerical illustration of the effectiveness of corrected phase-type approximations of Gerber-Shiu functions in the Li-Sendova risk model. Like Vatamidou *et al.* (2013) we use a specific claims distribution for which the exact values of Gerber-Shiu functions may be found (we employ multiprecision numerical Laplace transform inversion algorithms given in Abate and Valkó (2004); Trefethen *et al.* (2006)), and we retain $\eta_b = \frac{1}{3}$ and $\eta_c = \frac{1}{2}$ such that $\eta_b < \eta_c$. Unlike Vatamidou *et al.* (2013), here we choose the heavy-tailed component to be the more commonly used (and less heavy-tailed) Weibull with shape parameter $\alpha = \frac{1}{2}$; see equation 29.3.118 of Abramowitz and Stegun (1965) for the Laplace transform of this distribution. Since the insured class $i = 1$ is selected when a claim is larger than the random threshold, we can consider class 1 "high-risk" and class 2 "low-risk"; therefore we set $\frac{c_1}{\lambda_1} > \frac{c_2}{\lambda_2}$; specifically we set $c_1 = 8$, $\lambda_1 = 6$, $c_2 = 4.5$, $\lambda_2 = 4$. We set the threshold $H(u)$ to be exponential with rate $\nu = 0.5$, which gives $\int_0^\infty u \, \mathrm{d}H(u) > \eta_c$. In the exponential/Weibull mixture, we set $\epsilon = 0.001$ to explore a "worst-case" scenario similar to that in Vatamidou *et al.* (2013). We seek our demonstration to go beyond the (non time-discounted) ultimate ruin probability, so we choose a penalty mentioned in Gerber and Shiu (1998), namely $w(\cdot, y) = \frac{1-e^{-\rho_{פ} y}}{\delta}$ with $\delta = 50$; this penalty choice is interesting by including a claims law-dependent parameter.

Table 2.1. "Annuity" penalty, initial class $i = 1$

u	$m_{\epsilon,1}(u)$	$m^{\bullet}_{\epsilon,1}(u)$	$m_{0,1}(u)$	$m^{1}_{d,\epsilon}(u)$	$m^{1}_{r,\epsilon}(u)$
0	1.068774e−3	1.067958e−3	1.068981e−3	1.068775e−3	1.068774e−3
2	4.049021e−6	3.929837e−6	3.935074e−6	4.049019e−6	4.049022e−6
4	5.308547e−8	1.485092e−8	1.487655e−8	5.308383e−8	5.308518e−8
6	1.588628e−8	5.646873e−11	5.658844e−11	1.588554e−8	1.588599e−8
8	7.500555e−9	2.150154e−13	2.155560e−13	7.500197e−9	7.500384e−9
10	3.876637e−9	8.189702e−16	8.213521e−16	3.876455e−9	3.876543e−9

Table 2.2. "Annuity" penalty, initial class $i = 2$

u	$m_{\epsilon,2}(u)$	$m^{\bullet}_{\epsilon,2}(u)$	$m_{0,2}(u)$	$m^{2}_{d,\epsilon}(u)$	$m^{2}_{r,\epsilon}(u)$
0	8.392050e−4	8.385886e−4	8.393871e−4	8.392053e−4	8.392049e−4
2	3.168762e−6	3.083271e−6	3.087359e−6	3.168761e−6	3.168763e−6
4	3.881568e−8	1.164948e−8	1.166950e−8	3.881452e−8	3.881548e−8
6	1.123994e−8	4.429368e−11	4.438728e−11	1.123941e−8	1.123973e−8
8	5.289782e−9	1.686549e−13	1.690778e−13	5.289529e−9	5.289662e−9
10	2.728609e−9	6.423867e−16	6.442507e−16	2.728481e−9	2.728542e−9

In discussing their numerical studies, Vatamidou *et al.* (2013) observed that the correction terms greatly improved the accuracy of $\psi^{\bullet}_{\epsilon}(u)$ and $\psi_0(u)$ even when $\epsilon \ll 1$. As C in their test distribution only had the first-order moment finite, one might wonder how the correction terms fare for lighter-tailed subexponential distributions such as Weibull. In Tables 2.1 and 2.2, the phase-type approximations of $m_{\epsilon,i}(u)$ initially estimate these well, like through $u \leq 2$; quickly, however, they decay to give relative errors close to 1, by about $u = 6$. On the other hand, the decay of $m^i_{d,\epsilon}(u)$ and $m^i_{r,\epsilon}(u)$ clearly follows that of $m_{\epsilon,i}(u)$ more closely; specifically, for $i = 1$, the maximal absolute relative errors for the given values are $4.773e - 5$ and $2.438e - 5$, respectively. We see, therefore, that for such a choice of C, $m^i_{d,\epsilon}(u)$ and $m^i_{r,\epsilon}(u)$ continue to approximate $m_{\epsilon,i}(u)$ well for small ϵ. Furthermore, whereas Vatamidou *et al.* (2013) set ${}_{\delta}\upsilon_{\epsilon} = 0.5$, in our example with $\delta = 50$, ${}_{\delta}\upsilon_{\epsilon} \approx 0.0637 < 0.5$. In the setup of Propositions 1, 2, and 3, it follows that $\omega_c(u) \in \mathscr{L}_d(0)$ and $c(u) = o(\omega_c(u))$. (In this $\omega_c(u)$, the Lundberg root is $\rho^{\bullet}_{\epsilon}$, that of the "discard" base model). With our numerical illustration, we thus show that the comments of Vatamidou *et al.* (2013) about the advantage of adding the correction terms in a "worst-case scenario" can still apply after introducing considerable time discounting, selecting a much lighter heavy-tailed component, allowing for two classes of insureds, and modeling a quantity dependent upon the claims law.

Beyond the previous example, our code implements two phase-type distributions and two heavy-tailed distributions, allowing thus four choices of the mixture-model distribution. The phase-type distributions we implement are the exponential and Erlang-2, respectively with LSTs $\tilde{B}(s) = \frac{\frac{1}{\eta_b}}{\frac{1}{\eta_b}+s}$ and $\tilde{B}(s) = \left(\frac{\frac{2}{\eta_b}}{\frac{2}{\eta_b}+s}\right)^2$. On the other hand, for the heavy-tailed distributions, we implement the Abate-Whitt distribution with LST $\tilde{C}(s) = 1 - \frac{s}{\left(\frac{1}{\eta_c}+\sqrt{s}\right)\left(1+\sqrt{s}\right)}$ (see Abate and Whitt (1999)), and the Weibull with shape parameter $\frac{1}{2}$, which has LST $\tilde{C}(s) = \sqrt{\frac{\pi}{2\eta_c s}}\mathrm{erfcx}\left(\frac{1}{\sqrt{2\eta_c s}}\right)$ (see Eqn (29.3.118) of Abramowitz and Stegun (1965)). In all of these LSTs, we parametrized them to make the mean be the parameter. As in our specific example above, we set $\eta_b = \frac{1}{3}$ and $\eta_c = \frac{1}{2}$ to correspond to the values chosen by Vatamidou *et al.* (2013).

The four penalties we implement are respectively $w(\cdot,\cdot) = 1$, $w(x,\cdot) = e^{-\sigma x}$, $w(\cdot, y) = e^{-\tau y}$, and the aforementioned $w(\cdot, y) = \frac{1-e^{-\rho_{\text{פ}} y}}{\delta}$. Our reason for using multiprecision Laplace transform inversion is to allow our code to handle multiple penalty functions (one may find the unit Gerber-Shiu function explicitly for the exponential/Abate-Whitt mixture as a generalization of (Vatamidou *et al.*, 2013, Theorem 9); we have omitted such details). As we have given an example of results from the fourth penalty (the "annuity": see Gerber and Shiu (1998)), and the unit penalty is a special case of the Laplace transforms of the surplus before and the deficit at ruin, in the following we shall give examples from these latter two LT penalties. In addition, we have found the qualitative conclusions about the tail relative error of the phase-type approximations in Tables 2.1 and 2.2 to hold quite broadly across choices of penalty and model parameters; we recall also the findings in Vatamidou *et al.* (2014b) about the special case of GSFs considered in Vatamidou *et al.* (2013). In the asymptote (of initial capital u becoming large), we found $m^i_{d,\epsilon}(u)$ had the same relative error for both initial classes of insured $i \in \{1, 2\}$, and likewise for $m^i_{r,\epsilon}(u)$. Again, our Propositions 1, 2, and 3 establish that adding the correction term to $m^\bullet_\epsilon(u)$ and $m_0(u)$ captures the heavy-tailed behavior of $m_\epsilon(u)$ (with the nuances of allowing general $w(\cdot,\cdot)$) in the compound Poisson risk model. So we will focus more on comparing $m_{d,\epsilon}(u)$ and $m_{r,\epsilon}(u)$ to the asymptotic result given by our Proposition 1, specifically in terms of the relative errors for small initial capital u (in the compound Poisson risk model, of course).

However, we will first finish summarizing our general numerical observations. For the compound Poisson risk model specifically, we observed that $m_{d,\epsilon}(u)$ would initially overestimate $m_\epsilon(u)$ when $\delta > 0$, but underestimate $m_\epsilon(u)$ for all u when $\delta = 0$. This is in contradistinction to Vatamidou *et al.* (2013), who merely stated that the corrected discard approximation underestimated the true value for all $u \geq 0$. More like Vatamidou *et al.* (2013), we found that $m_{r,\epsilon}(u)$ generally gave better numerical estimates of $m_{\epsilon}(u)$ than $m_{d,\epsilon}(u)$, by which we mean smaller maximum error (in absolute value) and lesser relative error in the tail. When $\delta = 0$, we have found that $m_{d,\epsilon}(u)$ and $m_{r,\epsilon}(u)$ do not always reach a maximum value (seemingly for larger values of ϵ and the safety loading θ, when $C(x)$ is Abate-Whitt); this type of potential limit to the utility of CPTA was present even in the extant cases of $\psi_{d,\epsilon}(u)$ and $\psi_{r,\epsilon}(u)$. On the other hand, when we relax the compound Poisson risk model to the Li-Sendova risk model, evidently much more can happen. Namely, we found that $m^{i}_{d,\epsilon}(u)$ could overestimate $m_{\epsilon,i}(u)$ for small u even with $\delta = 0$, or for all large u under some parameter combinations.

For illustrative examples of comparing the performance of CPTA against asymptotic approximations, we choose the penalties $w(x,\cdot) = e^{-50x}$, and $w(\cdot, y) = e^{-50y}$. We set the discount rate $\delta = 1.5$ and the safety loading $\theta = 0.15$. We do so with perturbation parameter choices $\epsilon = 0.1$ and $\epsilon = 0.001$. We use a mixture of Erlang-2 and Abate-Whitt; Tables 2.3-2.6 show the relative errors of $m_{d,\epsilon}(u)$, $m_{r,\epsilon}(u)$, and the Proposition 1 for some small u. We point out here that for these values the asymptotic results perform considerably worse than the CPTA, but better for $\epsilon = 0.1$ than for $\epsilon = 0.001$.

Table 2.3. $w(x,\cdot) = e^{-50x}$, $\epsilon = 0.001$

u	$m_{d,\epsilon}(u)$	$m_{r,\epsilon}(u)$	$m_{\text{asy}}(u)$
0.2	9.9077e−8	1.1796e−7	9.9805e−1
0.4	1.3389e−7	2.9383e−7	9.9892e−1
0.6	2.0918e−7	5.1306e−7	9.9911e−1
0.8	3.1932e−7	7.6210e−7	9.9914e−1
1	4.7262e−7	1.0180e−6	9.9909e−1
1.2	6.8244e−7	1.2556e−6	9.9898e−1
1.4	9.6717e−7	1.4464e−6	9.9881e−1
1.6	1.3517e−6	1.5561e−6	9.9857e−1
1.8	1.8695e−6	1.5423e−6	9.9824e−1
2	2.5660e−6	1.3516e−6	9.9780e−1
2.2	3.5025e−6	9.1627e−7	9.9719e−1
2.4	4.7625e−6	1.4926e−7	9.9638e−1
2.6	6.4592e−6	−1.0623e−6	9.9528e−1
2.8	8.7472e−6	−2.8656e−6	9.9378e−1
3	1.1837e−5	−5.4533e−6	9.9174e−1

Table 2.4. $w(x,\cdot) = e^{-50x}$, $\epsilon = 0.1$

u	$m_{d,\epsilon}(u)$	$m_{r,\epsilon}(u)$	$m_{\text{asy}}(u)$
0.2	1.3289e−3	1.5472e−3	8.0616e−1
0.4	1.7787e−3	3.6039e−3	8.8900e−1
0.6	2.7145e−3	6.1578e−3	9.0579e−1
0.8	4.0625e−3	9.0255e−3	9.0620e−1
1	5.8986e−3	1.1875e−2	8.9848e−1
1.2	8.3369e−3	1.4323e−2	8.8450e−1
1.4	1.1511e−2	1.5933e−2	8.6443e−1
1.6	1.5561e−2	1.6226e−2	8.3780e−1
1.8	2.0610e−2	1.4717e−2	8.0397e−1
2	2.6737e−2	1.0980e−2	7.6248e−1
2.2	3.3936e−2	4.7384e−3	7.1330e−1
2.4	4.2080e−2	−4.0236e−3	6.5716e−1
2.6	5.0905e−2	−1.4990e−2	5.9568e−1
2.8	6.0028e−2	−2.7505e−2	5.3131e−1
3	6.9002e−2	−4.0675e−2	4.6698e−1

Table 2.5. $w(\cdot, y) = e^{-50y}$, $\epsilon = 0.001$

u	$m_{d,\epsilon}(u)$	$m_{r,\epsilon}(u)$	$m_{\text{asy}}(u)$
0.2	−3.5039e−8	−9.0050e−8	9.9807e−1
0.4	−4.1447e−8	1.7812e−8	9.9888e−1
0.6	−4.4369e−8	1.8125e−7	9.9906e−1
0.8	−3.7896e−8	3.8235e−7	9.9908e−1
1	−1.5953e−8	6.0340e−7	9.9903e−1
1.2	2.9202e−8	8.2484e−7	9.9891e−1
1.4	1.0781e−7	1.0240e−6	9.9872e−1
1.6	2.3357e−7	1.1740e−6	9.9847e−1
1.8	4.2493e−7	1.2423e−6	9.9811e−1
2	7.0680e−7	1.1889e−6	9.9763e−1
2.2	1.1129e−6	9.6409e−7	9.9698e−1
2.4	1.6890e−6	5.0533e−7	9.9610e−1
2.6	2.4975e−6	−2.6655e−7	9.9491e−1
2.8	3.6229e−6	−1.4526e−6	9.9329e−1
3	5.1805e−6	−3.1823e−6	9.9109e−1

Table 2.6. $w(\cdot, y) = e^{-50y}$, $\epsilon = 0.1$

u	$m_{d,\epsilon}(u)$	$m_{r,\epsilon}(u)$	$m_{\text{asy}}(u)$
0.2	−3.3945e−4	−8.3636e−4	8.0730e−1
0.4	−3.5400e−4	4.0504e−4	8.8400e−1
0.6	−3.0661e−4	2.2732e−3	8.9997e−1
0.8	−1.2633e−4	4.5457e−3	8.9978e−1
1	2.5634e−4	6.9749e−3	8.9124e−1
1.2	9.2211e−4	9.2669e−3	8.7616e−1
1.4	1.9632e−3	1.1080e−2	8.5466e−1
1.6	3.4784e−3	1.2036e−2	8.2629e−1
1.8	5.5622e−3	1.1751e−2	7.9047e−1
2	8.2861e−3	9.8891e−3	7.4680e−1
2.2	1.1675e−2	6.2365e−3	6.9543e−1
2.4	1.5684e−2	7.7836e−4	6.3730e−1
2.6	2.0182e−2	−6.2470e−3	5.7425e−1
2.8	2.4964e−2	−1.4353e−2	5.0890e−1
3	2.9779e−2	−2.2883e−2	4.4428e−1

2.7. AN APPLICATION

To illustrate how one might use corrected phase-type approximations of Gerber-Shiu functions, we approximate $\mathrm{E}(T_\epsilon\, \mathrm{I}(T_\epsilon < \infty) \mid U_\epsilon(0) = u)$ in the compound Poisson risk model; the basic inspiration comes from Section 8.6.5 of Dickson (2017). To begin, denote $H^{[n]}_{\mathfrak{D}}(u) = \frac{\partial^n}{\partial \delta^n}\phi_{\mathfrak{D}}(u)|_{\delta=0}$, and $\varphi_{\mathfrak{D}}(u) = -H^{[1]}_{\mathfrak{D}}(u) = \mathrm{E}(T_{\mathfrak{D}}\, \mathrm{I}(T_{\mathfrak{D}} < \infty) \mid U_{\mathfrak{D}}(0) = u)$. Now, one may show that $\hat{\phi}_{d,\epsilon}(s) = \hat{\phi}^{\bullet}_{\epsilon}(s) + \frac{\epsilon\lambda\eta_c(s-\rho^{\bullet}_{\epsilon})}{l^{\bullet}_{\epsilon}(s)}\, \mathrm{T}_s\, \mathrm{T}_{\rho^{\bullet}_{\epsilon}}(c^e * G^{\bullet}_{\epsilon})(0)$, where $c^e(u) = \frac{1}{\eta_c}\bar{C}(u)$. Rearranging this form of $\hat{\phi}_{d,\epsilon}(s)$ gives $l^{\bullet}_{\epsilon}(s)\hat{\phi}_{d,\epsilon}(s) = l^{\bullet}_{\epsilon}(s)\hat{\phi}^{\bullet}_{\epsilon}(s) + \epsilon\lambda\eta_c(\hat{c}^e(\rho^{\bullet}_{\epsilon})\tilde{G}^{\bullet}_{\epsilon}(\rho^{\bullet}_{\epsilon}) - \hat{c}^e(s)\tilde{G}^{\bullet}_{\epsilon}(s))$. Note that $l^{\bullet}_{\epsilon}(s) = \delta - (s - \lambda + \lambda\tilde{P}^{\bullet}_{\epsilon}(s)) = \delta - s(1 - \lambda\eta^{\bullet}_{\epsilon}\tilde{P}^{\bullet e}_{\epsilon}(s)) = \delta - s\frac{1-\psi^{\bullet}_{\epsilon}(0)}{\tilde{M}^{\bullet}_{\epsilon}(s)}$. Then we have

$$\begin{aligned} \delta\Big(\hat{\phi}_{d,\epsilon}(s) - \hat{\phi}^{\bullet}_{\epsilon}(s)\Big) = s\frac{1-\psi^{\bullet}_{\epsilon}(0)}{\tilde{M}^{\bullet}_{\epsilon}(s)}\Big(\hat{\phi}_{d,\epsilon}(s) - \hat{\phi}^{\bullet}_{\epsilon}(s)\Big) \\ +\epsilon\lambda\eta_c\Big(\hat{c}^e(\rho^{\bullet}_{\epsilon})\tilde{G}^{\bullet}_{\epsilon}(\rho^{\bullet}_{\epsilon}) - \hat{c}^e(s)\tilde{G}^{\bullet}_{\epsilon}(s)\Big). \end{aligned} \tag{2.20}$$

Now, it holds that $\tilde{G}_{\mathfrak{D}}(s) = 1 - s\hat{\phi}_{\mathfrak{D}}(s)$; we differentiate (2.20) with respect to δ, and set $\delta = 0$, which results in

$$\begin{aligned} \hat{H}^{[0]}_{d,\epsilon}(s) - \hat{H}^{\bullet[0]}_{\epsilon}(s) = s\frac{1-\psi^{\bullet}_{\epsilon}(0)}{\tilde{M}^{\bullet}_{\epsilon}(s)}\Big(\hat{H}^{[1]}_{d,\epsilon}(s) - \hat{H}^{\bullet[1]}_{\epsilon}(s)\Big) \\ +\epsilon\lambda\eta_c\left(\frac{\partial}{\partial\delta}\hat{c}^e(\rho^{\bullet}_{\epsilon})\tilde{G}^{\bullet}_{\epsilon}(\rho^{\bullet}_{\epsilon})|_{\delta=0} + s\hat{c}^e(s)\hat{H}^{\bullet[1]}_{\epsilon}(s)\right). \end{aligned} \tag{2.21}$$

By setting $s = 0$, we see that $\epsilon\lambda\eta_c\frac{\partial}{\partial\delta}\hat{c}^e(\rho^{\bullet}_{\epsilon})\tilde{G}^{\bullet}_{\epsilon}(\rho^{\bullet}_{\epsilon})|_{\delta=0} = \hat{H}^{[0]}_{d,\epsilon}(0) - \hat{H}^{\bullet[0]}_{\epsilon}(0)$. Therefore, (2.21) becomes

$$\begin{aligned} \frac{1}{s}\Big(\hat{H}^{[0]}_{d,\epsilon}(s) - \hat{H}^{[0]}_{d,\epsilon}(0) - \Big(\hat{H}^{\bullet[0]}_{\epsilon}(s) - \hat{H}^{\bullet[0]}_{\epsilon}(0)\Big)\Big) \\ = \frac{1-\psi^{\bullet}_{\epsilon}(0)}{\tilde{M}^{\bullet}_{\epsilon}(s)}\Big(\hat{H}^{[1]}_{d,\epsilon}(s) - \hat{H}^{\bullet[1]}_{\epsilon}(s)\Big) + \epsilon\lambda\eta_c\hat{c}^e(s)\hat{H}^{\bullet[1]}_{\epsilon}(s). \end{aligned} \tag{2.22}$$

The left-hand side of (2.22) may be expressed as $\mathrm{T}_s\,\mathrm{T}_0\Big(H_{\epsilon}^{\bullet[0]} - H_{d,\epsilon}^{[0]}\Big)(0)$, where $H_{\epsilon}^{\bullet[0]}(u) - H_{d,\epsilon}^{[0]}(u) = -\frac{\epsilon\lambda\eta_c}{1-\psi_{\epsilon}^{\bullet}(0)}(\Pr(M_{\epsilon,0}^{\bullet} + M_{\epsilon,1}^{\bullet} + C^e > u) - \Pr(M_{\epsilon,0}^{\bullet} > u))$ (see also Definition 1 of Vatamidou *et al.* (2013)). Let $\Xi_{\mathfrak{D}}^{c}(u) = \frac{\lambda\eta_c}{(1-\psi_{\mathfrak{D}}(0))^2}\int_{0-}^{u}\mathrm{T}_0(\Pr(M_{\mathfrak{D},0} + M_{\mathfrak{D},1} + C^e > u - x) - \Pr(M_{\mathfrak{D},0} > u - x))\,\mathrm{d}\Pr(M_{\mathfrak{D},2} \le x) - \frac{\lambda\eta_c}{1-\psi_{\mathfrak{D}}(0)}\int_{0-}^{u}\varphi_{\mathfrak{D}}(u-x)\,\mathrm{d}\Pr(M_{\mathfrak{D},0} + C^e \le x)$, and let $\Xi_{\mathfrak{D}}^{b}(u)$ be the equivalent with B instead of C; in both, $M_{\mathfrak{D},i}(x)$ are iid with tail $\phi_{\mathfrak{D}}(x)$. Rearranging (2.22) and inverting the Laplace transforms gives the corrected discard approximation $\varphi_{d,\epsilon}(u) = \varphi_{\epsilon}^{\bullet}(u) + \epsilon\Xi_{\epsilon}^{\bullet c}(u)$. Likewise, we also have the corrected replace approximation $\varphi_{r,\epsilon}(u) = \varphi_0(u) + \epsilon(\Xi_0^{c}(u) - \Xi_0^{b}(u))$.

2.8. CONCLUDING DISCUSSION

We have demonstrated how to generalize the method of Vatamidou *et al.* (2013) to Gerber-Shiu functions; this works not only in the classical risk model, but also in the dependent generalization thereof considered by Li and Sendova (2015). We expect our generalization of the extant technique to apply in other dependency structures as well; possible future work could involve working out our results in the general semi-Markov framework considered by Albrecher and Boxma (2005). Corrected phase-type approximations of Gerber-Shiu functions continue to have the proper heavy-tailed behavior like in the particular case Vatamidou *et al.* (2013) considered. The correction terms possess an intuitive, precisely quantifiable interpretation of their improvement upon the error of phase-type approximations. In the future, we might extend Propositions 1, 2, and 3 to the Li-Sendova risk model. We could also look more closely at choices of the threshold distribution $H(y)$ and the resulting impact on $m_{d,\epsilon}^{i}(u)$ and $m_{r,\epsilon}^{i}(u)$. Lastly, it could be interesting to add additional classes of insureds.

3. RESERVING FOR INCURRED CLAIMS: RECURSIONS UNDER A MARKOV RENEWAL MODEL AND NON-CONSTANT DISCOUNT RATE

3.1. OVERVIEW OF SECTION

We model the finite-time moments of IBNR and IR claims, extending a recently proposed model. Our proposed extensions are to allow the time between claim-causing events to depend on the severity of the previous event, and to allow the discount rate to vary continuously with time. In the former, we adapt a dependency structure from the risk-theory literature which involved comparing claim sizes to a random threshold, now instead comparing the number of claims caused by an event to any number of random intervals. In the latter, we demonstrate the use of an unspecified, deterministic non-constant discount rate within our proposed dependency structure. We give the special cases where either one of our extensions are omitted, and we discuss some particular examples of the latter extension regarding discount rates.

3.2. GENERAL DISCUSSION

Whereas in Section 2 we considered problems involving startup capital of an insurer on an infinite time horizon, now we turn to finite time horizons and examine certain problems pertaining to the day-to-day reserves an insurer needs to stay solvent. In real life, events which cause claims may trigger more than one claim in a single event, like in a pandemic or an earthquake. Whether or not these claims have been reported to the insurer at the time of the incident's occurrence, the insurer must pay every single claim meeting the policy's criteria. A health epidemic could easily leave numerous policyholders hospitalized or quarantined for some time, and a major-enough earthquake could displace thousands, if not millions, of people, especially in densely populated regions. A policyholder with a severe illness might be unable to report their expenses quickly, whether related to worker's compensation, the bill for a hospital stay, or otherwise. After an earthquake, homeowners or automobile owners might not return home immediately, and hence be unaware of how much damage their properties sustained, whether from the seismic activity or after-effects like tsunamis or simply people looting. The point of these examples is that catastrophes can cause many IBNR claims, conceivably in a manner proportional to the event's severity.

Like our previous work on Gerber-Shiu functions, we model individual claims, rather than in the aggregate. There is good reason for this; as we mentioned before, claims may neither come to the insurer's attention immediately upon being incurred, nor be settled (paid) by the insurer immediately upon being reported. Nevertheless, traditionally insurers modeled reserves in the aggregate; for a survey of classical methods (such as the chain ladder) based on "run-off triangles" used by practicing actuaries to determine the reserves needed for making all loss payments, one may consult Schmidt and Zocher (2008). In fact, per Bornhuetter and Ferguson (1972), for much of the 20th century, the actuarial literature mostly overlooked the situation of IBNR claims; they responded sharply and with an "alarm-sounding." In the early 1990s, Gile (1994) proposed for practicing actuaries a particular stochastic model for IBNR, allowing some dependency between loss severity and the reporting lag. However, that paper aggregated claims per reporting period and the like, thus ignoring the severity of individual claims, the so-called "micro-level." Additionally, only the mean and variance of IBNR liabilities were considered, rather than the overall distribution.

On the front of compound random sums, Léveillé and Garrido (2001a,b) found the moments of what they called a "compound renewal present value process." They did so for an ordinary renewal process, a delayed renewal process, and a stationary renewal process, in all cases with a constant discount rate. Meanwhile, Léveillé and Adékambi (2011) modeled much the same, but with a stochastic discount rate. Besides continuing to neglect reporting lags and multiple claims from single events, they only considered up to second moments.

The paper by Landriault *et al.* (2017) considered a compound renewal claims process, with time value of money at a constant discount rate, reporting lags, batches of claims per event; and dependencies among the claim time, reporting lag, and claim severity for a given claim from a given event. They primarily investigated the discounted IBNR claim amounts under such a setup, also incorporating discounted incurred and reported (IR) claim amounts; later they specialized these results to examine the number of IBNR claims. Their basic methodology was deriving expressions for the Laplace transforms of these quantities, differentiating to get a renewal equation of the recursive moments, then solving that equation.

In this section, we build upon their main results in two directions; since the IBNR claim count is just a special case of the quantities Landriault *et al.* (2017) considered, we focus on extending their Theorems 1 and 3. We relax their assumption that the batch sizes of claims are all independent of each other and the time to the next event; consider the 2017 US hurricane season, in particular Harvey and Irma, as a real-world suggestion that such an assumption may prove faulty. So, we propose allowing the time to the next event to depend on the number of claims produced by the current event. In a ruin-theoretic context, Albrecher and Boxma (2004) and Li and Sendova (2015) compared each individual claim (without batches of claims) to a random threshold and then adjusted the claim arrival rate accordingly; they only modeled two such classes, albeit Li and Sendova (2015) hinting at the possibility of more classes. Previously, we showed how to approximate such quantities under one type of catastrophic assumption, by a generalization of Vatamidou *et al.* (2013); here we apply a similar sort of dependency to a different type of catastrophic assumption, namely claim numbers, and we demonstrate one way to have more than two categories, such that the random threshold becomes random intervals.

The other direction we take things is to let the discount rate vary with time in a deterministic manner. In settings other than what we consider here, a few papers have examined stochastic discount rates. The first in the context of Gerber-Shiu functions was apparently Wang and Ling (2017), which only considered the compound Poisson risk model with light-tailed claims. In a discrete risk model, Deng *et al.* (2017) studied Gerber-Shiu functions with random discounts as well. An earlier study of stochastic discount rates in compound renewal sums for first and second moments was conducted in Léveillé and Adékambi (2011), and more recently in Rabehasaina and Woo (2016). For our Markovian generalization of Landriault *et al.* (2017), we show how the formulas work with deterministic general real $D(x)$, which includes of course $D(x) = \delta x$. Kennedy (1992) contains a discussion of the benefits of modeling the discount rate as a deterministic function of time, even if a stochastic discount rate could be still more realistic. In the context of time-varying discount rates, we choose some particular cases of $\delta(u)$, and in an example considered by Landriault *et al.* (2017) we numerically contrast the effects of these cases with the existing choice of $\delta(\cdot) \equiv \delta$.

We organize the rest of the section as follows: in Section 3.3, we articulate the model under consideration and the associated notation. In Section 3.4, we give a recursive relation for the finite-time IBNR moments under our proposed generalization of the model of Landriault *et al.* (2017). We likewise give a recursive relation for the finite-time IBNR and IR joint moments in Section 3.5. We provide some numerical illustrations of time-varying discount rates in Section 3.6, and we close the section in Section 3.7.

3.3. NOTATION AND MODEL

We use the model of Landriault *et al.* (2017), mostly following their notation. We also extend their model, one main way being that we let $\tau_1 \sim F_l(t)$, for $l \in \Im$, where $\Im = \{1, 2, \ldots, \mathfrak{m}\}$. These τ_k represent the interevent times, for $k \geq 2$; τ_1 is the time until the first claim-causing event. Also, we assume $\tau_0 = 0$ almost surely (a.s.). For each $l \in \Im$, we assume $F_l(0) = 0$, meaning no atom at 0; this makes intuitive sense because we allow claims to occur in batches, much like Landriault *et al.* (2017). The number of claims random variables $C_t \overset{\text{iid}}{\sim} \Pr(C = c)$ with "generic rv" C. That is, if an event at time t causes claims, C_t is the number of such claims associated with that event. Unlike Landriault *et al.* (2017), at event time $T_k = \sum_{j=1}^{k} \tau_j$, $k \in \{1, 2, 3, \dots\}$, we compare C_{T_k} to $\mathfrak{m}$ random intervals (over $\mathbb{N}$). To this end, let $Q_{l,k} \overset{\text{iid}}{\sim} R_l(q)$, with $\{Q_{l,k}\}_{l=1}^{\mathfrak{m}-1}$ independent of each other and of all other random variables in the model. Then we say $\tau_{k+1} \mid C_{T_k} \in \mathscr{I}_l \sim F_l(t)$, in which $\mathscr{I}_1 = (Q_{1,k}, \infty)$, $\mathscr{I}_l = (Q_{l,k}, Q_{l-1,k}]$, $l = 2, \dots, \mathfrak{m} - 1$, and $\mathscr{I}_{\mathfrak{m}} = (0, Q_{\mathfrak{m}-1,k}]$. The intuition for excluding $Q = 0$ is that T_k are claim-causing event times, i.e. which produce one or more claims. Note that the functions $R_l(q)$ need to be stochastically increasing for well-definedness of $\Pr(C \in \mathscr{I}_l)$. Now, as a discrete analog to the functions introduced by Albrecher and Boxma (2004), we set $\chi_l(z) = \mathrm{E}\Big(z^C\, \mathrm{I}(C \in \mathscr{I}_l)\Big)$, so that

$$\chi_l(z) = \sum_{c=1}^{\infty} z^c \left(R_l(c-1) - R_{l-1}(c-1)\right) \Pr(C = c),$$

where $R_0(\cdot) \equiv 0$, and $R_{\mathfrak{m}}(\cdot) \equiv 1$. Our idea is that, viewing the number of claims caused by a particular event as a "metric" of the severity of that event, after such an event an insurer might wish to reevaluate the assumed distribution of the time until the next claim-causing event. Then we can interpret $l \in \mathfrak{I}$ as a level of event severity in terms of how many claims resulted.

Now we recall quantities from Landriault *et al.* (2017) which we use without generalizing much further. A policyholder might not immediately file a claim over a loss incurred from an incident, which leads to the notion of a "reporting lag." If k is the number of the event, and i is the number of the claim resulting from said event, then we denote the corresponding reporting lag by $W_{i,k} \overset{\text{iid}}{\sim} K(\cdot)$; furthermore, $l(\cdot)$ is some nonnegative function of this lag. With the same meaning for the indices i and k, $X_{i,k}$ is the (non-negative) deflated claim severity (valued at the time of the event k). We assume $X_{i,k} \overset{\text{iid}}{\sim} P(\cdot)$. If $\aleph(\cdot)$ is a df, then we write the Laplace-Stieltjes transform as $\tilde{\aleph}(s) = \int_{0-}^{\infty} e^{-sx} \aleph(\mathrm{d}x)$. For example, the LST corresponding to $X_{i,k}$ is $\tilde{P}(s)$. In a slight extension of the notation of Landriault *et al.* (2017), we say the random vectors $(\tau_k, W_{i,k}, X_{i,k})$ jointly follow $J_l(\cdot,\cdot,\cdot)$ in $k \in \mathbb{N}^+$; given τ_k, we assume independence of $X_{i,m}$ and $W_{j,n}$ whenever either $i \neq j$ or $m \neq n$. In other words, denoting $W_{i,k}|\tau_k \sim K_{W|\tau}(w|t)$ and $X_{i,k}|(\tau_k, W_{i,k}) \sim P_{X|\tau,W}(x|t,w)$ like Landriault *et al.* (2017), our extended form of their Eqn (2.1) is

$$J_l(t,w,x) = F_l(t)K_{W|\tau}(w|t)P_{X|\tau,W}(x|t,w).$$

Some functions will appear later in much the same way they do in Landriault *et al.* (2017): with $\mu_n(t,w) = \mathrm{E}\Big(X_{i,k}^n|\tau_k = t, W_{i,k} = w\Big)$, for $n \in \mathbb{N}$,

$$\xi_i(x,t) = \int_{t-x}^{\infty} l(w)^i \mu_i(x,w)K_{W|\tau}(\mathrm{d}w|x), 0 \le x \le t, \tag{3.1}$$

$$\eta_i(x,t) = \int_0^{t-x} l(w)^i \mu_i(x,w)K_{W|\tau}(\mathrm{d}w|x), 0 \le x \le t. \tag{3.2}$$

Now we elaborate on the Markovian nature of our generalization of the model of Landriault *et al.* (2017), one reference on Markov renewal theory being Janssen and Manca (2006). We can formulate the τ_k setup as a semi-Markov chain; define the bivariate process $\{(J_n, \tau_n), n \ge 0\}$, where $\tau_n \mid J_{n-1} = l \sim F_l(t)$, for $l \in \mathfrak{I}$ and $n > 0$. For $n \ge 0$, J_n tracks the distribution of τ_{n+1}, the time until event $n+1$; in other words, the event $\{J_n = l\}$ is the same as the event $\{C_{\tau_n} \in \mathscr{I}_l \mid \tau_n\}$. As in (Janssen and Manca, 2006, Section (4.2)), we have assumed $\Pr(\tau_0 = 0) = 1$ a.s.; writing $\Pr(J_0 = l) = p_l$, then we suppose $\sum_{l\in\mathfrak{I}} p_l = 1$. For our model, we can see that for all $n > 0$ and $i, j \in \mathfrak{I}$, we have

$$Q_{ij}(t) = \Pr\big(J_n = j, \tau_n \le t \mid (J_k, \tau_k), k = 0, \dots, n-1; J_{n-1} = i\big) = \chi_j(1)F_i(t), \tag{3.3}$$

a so-called "semi-Markov matrix." Then $(J, \tau) = \{(J_n, \tau_n), n \ge 0\}$ is a semi-Markov chain with state space $\mathfrak{I} \times \mathbb{R}^+$ as per (Janssen and Manca, 2006, Definition 4.2.2), defined by $\Big(\underset{\sim}{p}, \underset{\sim}{Q}(t)\Big)$, and is homogeneous. By (Janssen and Manca, 2006, Proposition 3.1), the embedded Markov chain J_n has transition matrix $\underset{\sim}{Q}(\infty)$. As $T_k = \sum_{i=1}^k \tau_i$, $\{(J_n, T_n), n \ge 0\}$ is the Markov renewal process of kernel $\underset{\sim}{Q}(t)$.

We define $N_l(t) = \sup\left\{k \geq 0 : \sum_{i=1}^{k} \tau_i \leq t \mid J_0 = l\right\}$, the number of claim-causing events by time t, given that the initial class is l. We denote the "Markov renewal function" associated with $N_l(t)$ by $H_l(t) = \mathrm{E}(N_l(t))$; using iterated expectation conditioning on τ_1, we get

$$\begin{aligned}
H_l(t) &= \int_0^\infty \mathrm{E}(N_l(t) \mid \tau_1 = v) F_l(\mathrm{d}v) \\
&= \int_0^\infty \sum_{n=0}^{\infty} n \Pr(N_l(t) = n \mid \tau_1 = v) F_l(\mathrm{d}v) \\
&= \int_0^\infty \sum_{n=0}^{\infty} n \Pr\left(\sum_{i=1}^{n+1} \tau_i > t \mid \tau_1 = v\right) F_l(\mathrm{d}v) \\
&= \int_0^\infty \sum_{n=0}^{\infty} n \Pr\left(\sum_{i=2}^{n+1} \tau_i > t - v\right) F_l(\mathrm{d}v) \\
&= \int_0^t \sum_{m \in \Im} \sum_{n=0}^{\infty} n \Pr\left(\sum_{i=2}^{n+1} \tau_i > t - v \mid J_1 = m\right) \chi_m(1) F_l(\mathrm{d}v).
\end{aligned}$$

By regenerativity, then

$$\begin{aligned}
H_l(t) &= \int_0^t \sum_{m \in \Im} \chi_m(1) \sum_{n=0}^{\infty} n \Pr(N_m(t - v) = n - 1) F_l(\mathrm{d}v) \\
&= \int_0^t \sum_{m \in \Im} \chi_m(1)\, \mathrm{E}(N_m(t - v) + 1) F_l(\mathrm{d}v) \\
&= F_l(t) + \sum_{m \in \Im} \chi_m(1) \int_0^t H_m(t - v) F_l(\mathrm{d}v).
\end{aligned}$$

With $\tilde{H}_l(s) = \int_0^\infty e^{-st} H_l(t)\mathrm{d}t$ and $\tilde{F}_l(s) = \int_0^\infty e^{-st} F_l(\mathrm{d}t)$, we have $\tilde{H}_l(s) = \tilde{F}_l(s) + \sum_{m \in \Im} \chi_m(1) \tilde{H}_m(s) \tilde{F}_l(s)$. In matrix form, this is $\underline{\tilde{H}}(s) = \underline{\tilde{F}}(s) + \underline{\chi}(1)^\top \underline{\tilde{H}}(s) \underline{\tilde{F}}(s) = \underline{\tilde{F}}(s) + \underline{\tilde{F}}(s) \underline{\chi}(1)^\top \underline{\tilde{H}}(s)$; following some matrix arithmetic, we get

$$\underline{\tilde{H}}(s) = \left(\underline{I} - \underline{\tilde{F}}(s) \underline{\chi}(1)^\top\right)^{-1} \underline{\tilde{F}}(s).$$

Since $\left(\underline{I} - \tilde{\underline{F}}(s)\underline{\chi}(1)^{\top}\right)^{-1} = \sum_{n=0}^{\infty}\left(\tilde{\underline{F}}(s)\underline{\chi}(1)^{\top}\right)^{n}$,

$$\tilde{\underline{H}}(s) = \sum_{n=0}^{\infty}\left(\tilde{\underline{F}}(s)\underline{\chi}(1)^{\top}\right)^{n}\tilde{\underline{F}}(s) = \tilde{\underline{F}}(s)\sum_{n=0}^{\infty}\left(\underline{\chi}(1)^{\top}\tilde{\underline{F}}(s)\right)^{n} = \tilde{\underline{F}}(s)\left(1 - \underline{\chi}(1)^{\top}\tilde{\underline{F}}(s)\right)^{-1}.$$

Thus $\tilde{H}_l(s) = \left(1 - \underline{\chi}(1)^{\top}\tilde{\underline{F}}(s)\right)^{-1}\tilde{F}_l(s)$.

We will let the discount rate vary with time, such that $\delta(x) \equiv \delta$ might not hold. The cumulative force of discount at time $x \geq 0$ will be $D(x) = \int_0^x \delta(u)\,\mathrm{d}u$; see Section 3.6 for some further discussion. We allow $D(x)$ to assume values in $(-\infty, \infty]$ except on $A \subset (0, \infty)$, with A at most countably infinite; this guarantees that $e^{-D(x)} \in [0, \infty)$ except on the same A. In subsequent sections, we will use the notation $F_{n\delta|l}(\mathrm{d}x) = e^{-nD(x)}F_l(\mathrm{d}x)$ and $F_{n\delta}(\mathrm{d}x) = \underline{\chi}(1)^{\top}\underline{F}_{n\delta}(\mathrm{d}x) = e^{-nD(x)}\underline{\chi}(1)^{\top}\underline{F}(\mathrm{d}x)$.

3.4. MOMENTS OF IBNR CLAIMS

Denote

$$Z_l(t) = \sum_{k=1}^{N_l(t)} Y_k(t), \tag{3.4}$$

$$Y_k(t) = \sum_{i=1}^{C_{T_k}} e^{-(D(T_k+\tau_0)-D(\tau_0))} l(W_{i,k})\,\mathrm{I}(W_{i,k} + T_k > t) X_{i,k}.$$

This is the same as (Landriault *et al.*, 2017, Eqn (2.2)), except the subscript l tracks the distribution of τ_1, and we have $D(x)$ instead of simply δx. Then we write the Laplace transform of $Z_l(t)$ as $\tilde{L}_{\gamma|l}(t)$. Paralleling (Landriault *et al.*, 2017, Eqn (3.1)), we get that

$$\tilde{\underline{L}}_{\gamma}(t) = \bar{\underline{F}}(t) + \int_0^t \underline{\chi}(\zeta(\gamma; t|x))^{\top}\tilde{\underline{L}}_{\gamma e^{-D(x)}}(t - x)\underline{F}(\mathrm{d}x).$$

Now, however, we have also conditioned on the number of claims caused by the event at τ_1. The function $\zeta(\gamma; t|x)$ is given by $1+\int_{t-x}^{\infty}\Big(\tilde{P}_{X|\tau,W}\Big(\gamma e^{-D(x)}l(w)|x,w\Big)-1\Big)K_{W|\tau}(\mathrm{d}w|x)$. Here we give the details of finding each component of $\underline{\tilde{L}}_{\gamma}(t)$ as given several lines above; first,

$$
\begin{aligned}
\tilde{L}_{\gamma|l}(t) &= \mathrm{E}(\exp(-\gamma Z_l(t))) \\
&= \mathrm{E}(\mathrm{E}(\exp(-\gamma Z_l(t))|\sigma(N_l(t)+1))) \\
&= \mathrm{E}\left(\prod_{k=1}^{N_l(t)} \exp(-\gamma Y_k(t))\right) \\
&= \int_0^{\infty} \mathrm{E}\left(\prod_{k=1}^{N_l(t)} \exp(-\gamma Y_k(t))\middle|\tau_1 = x\right) F_l(\mathrm{d}x) \\
&= \bar{F}_l(t) + \int_0^{t} \mathrm{E}\left(\prod_{k=1}^{N_l(t)} \exp(-\gamma Y_k(t))\middle|\tau_1 = x\right) F_l(\mathrm{d}x).
\end{aligned}
$$

In the last line, if $\tau_1 = x > t$, the product inside $\mathrm{E}(\cdot)$ becomes 1 because $Z_l(t) = N_l(t) = 0$. Taking the integrand in the second term,

$$
\begin{aligned}
&\mathrm{E}\left(\prod_{k=1}^{N_l(t)} \exp(-\gamma Y_k(t))\middle|\tau_1 = x\right) \\
&= \mathrm{E}\left(\mathrm{E}\left(\exp(-\gamma Y_1(t)) \prod_{k=2}^{N_l(t)} \exp(-\gamma Y_k(t))\middle|\tau_1 = x, C_x = C\right)\right) \\
&= \sum_{m\in\mathfrak{I}} \mathrm{E}\left(\mathrm{E}\left(\exp(-\gamma Y_1(t)) \prod_{k=2}^{N_l(t)} \exp(-\gamma Y_k(t))\, \mathrm{I}(C \in \mathscr{I}_m)\middle|\tau_1 = x, C_x = C\right)\right) \\
&= \sum_{m\in\mathfrak{I}} \mathrm{E}(\mathrm{E}(\exp(-\gamma Y_1(t))|\tau_1 = x, C_x = C) \\
&\quad \times \mathrm{E}\left(\prod_{k=2}^{N_l(t)} \exp(-\gamma Y_k(t))\, \mathrm{I}(C \in \mathscr{I}_m)\middle|\tau_1 = x, C_x = C\right)\right),
\end{aligned}
$$

since, given $\tau_1 = x$ and $C_x = C$, the random variates

$$\exp(-\gamma Y_1(t)) \text{ and } \prod_{k=2}^{N_l(t)} \exp(-\gamma Y_k(t))\, \mathrm{I}(C \in \mathscr{I}_m)$$

are independent. Now,

$$\begin{aligned}
&\mathrm{E}(\exp(-\gamma Y_1(t))|\tau_1 = x, C_x = C) \\
&= \mathrm{E}\left(\exp\left(-\gamma \sum_{i=1}^{C} e^{-D(x)} l(W_{i,1})\, \mathrm{I}(W_{i,1} > t-x) X_{i,1}\right)\middle|\tau_1 = x, C_x = C\right) \\
&= \mathrm{E}\left(\prod_{i=1}^{C} \exp\left(-\gamma e^{-D(x)} l(W_{i,1})\, \mathrm{I}(W_{i,1} > t-x) X_{i,1}\right)\middle|\tau_1 = x, C_x = C\right) \\
&= \left(\mathrm{E}\left(\exp\left(-\gamma e^{-D(x)} l(W_{1,1})\, \mathrm{I}(W_{1,1} > t-x) X_{1,1}\right)\middle|\tau_1 = x\right)\right)^C \\
&= (\zeta(\gamma; t|x))^C.
\end{aligned}$$

Next, by regenerativity of $N_l(t)$ at $\tau_1 = x$,

$$\begin{aligned}
&\mathrm{E}\left(\prod_{k=2}^{N_l(t)} \exp(-\gamma Y_k(t))\, \mathrm{I}(C \in \mathscr{I}_m)\middle|\tau_1 = x, C_x = C\right) \\
&= \mathrm{E}\left(\prod_{k=2}^{N_l(t)} \exp\left(-\gamma \sum_{i=1}^{C_{x+\sum_{j=2}^k \tau_j}} e^{-\left(D\left(x+\sum_{j=2}^k \tau_j\right)\right)} l(W_{i,k})\, \mathrm{I}\left(W_{i,k} + x + \sum_{j=2}^{k} \tau_j > t\right) X_{i,k}\right)\right. \\
&\qquad \times \mathrm{I}(C \in \mathscr{I}_m)|\tau_1 = x, C_x = C) \\
&= \mathrm{E}\left(\prod_{k=2}^{N_m(t-x)+1} \exp\left(-\gamma e^{-D(x)} \sum_{i=1}^{C_{\sum_{j=2}^k \tau_j}} e^{-\left(D\left(x+\sum_{j=2}^k \tau_j\right)-D(x)\right)}\right.\right. \\
&\qquad \left.\left.\times\, l(W_{i,k})\, \mathrm{I}\left(W_{i,k} + \sum_{j=2}^{k} \tau_j > t-x\right) X_{i,k}\right)\middle|\tau_1 = x, C_x = C\right) \mathrm{I}(C \in \mathscr{I}_m) \\
&= \mathrm{E}\left(\prod_{k=2}^{N_m(t-x)+1} \exp\left(-\gamma e^{-D(x)} \sum_{i=1}^{C_{\sum_{j=2}^k \tau_j}} e^{-\left(D\left(\sum_{j=2}^k \tau_j + \tau_1\right)-D(\tau_1)\right)}\right.\right.
\end{aligned}$$

$$
\begin{aligned}
&\times l(W_{i,k})\,\mathrm{I}\bigg(W_{i,k}+\sum_{j=2}^{k}\tau_j>t-x\bigg)X_{i,k}\bigg)\bigg|\tau_1=x,C_x=C\bigg)\,\mathrm{I}(C\in\mathscr{I}_m)\\
&=\mathrm{E}\Bigg(\prod_{k=2}^{N_m(t-x)+1}\exp\bigg(-\gamma e^{-D(x)}\sum_{i=1}^{C_{T_k}}e^{-(D(T_k+\tau_1)-D(\tau_1))}\\
&\qquad\times l(W_{i,k})\,\mathrm{I}(W_{i,k}+T_k>t-x)X_{i,k}\bigg)\bigg|\tau_1=x,C_x=C\Bigg)\,\mathrm{I}(C\in\mathscr{I}_m)\\
&=\mathrm{E}\Bigg(\prod_{k=2}^{N_m(t-x)+1}\exp\Big(-\gamma e^{-D(x)}Y_k(t-x)\Big)\bigg|\tau_1=x\Bigg)\mathrm{I}(C\in\mathscr{I}_m)\\
&=\tilde{L}_{\gamma e^{-D(x)}|m}(t-x)\,\mathrm{I}(C\in\mathscr{I}_m).
\end{aligned}
$$

In the second to third lines, $C_x \overset{\mathrm{d}}{=} C_y$, and $N_l(t)|(J_1=m) \overset{\mathrm{d}}{=} N_m(t-x)+1$. The key observation for allowing general $D(x)=\int_0^x \delta(u)\,\mathrm{d}u$ instead of $\delta(u)\equiv\delta$ is that after conditioning on $\tau_1=x$, $\tau_1 \overset{\mathrm{as}}{=} x$ takes the same role as $\tau_0 \overset{\mathrm{as}}{=} 0$ with respect to regenerativity of the process $N_l(t)$, and hence in the definition of $Y_k(t)$. Reassembling the previous two runs of expressions,

$$
\begin{aligned}
\mathrm{E}\Bigg(\prod_{k=1}^{N_l(t)}\exp(-\gamma Y_k(t))\bigg|\tau_1=x\Bigg)&=\sum_{m\in\mathfrak{I}}\mathrm{E}\Big(\big(\zeta(\gamma;t|x)\big)^{C}\,\mathrm{I}(C\in\mathscr{I}_m)\tilde{L}_{\gamma e^{-D(x)}|m}(t-x)\Big)\\
&=\sum_{m\in\mathfrak{I}}\chi_m(\zeta(\gamma;t|x))\tilde{L}_{\gamma e^{-D(x)}|m}(t-x)=\underline{\chi}(\zeta(\gamma;t|x))^{\top}\underline{\tilde{L}}_{\gamma e^{-D(x)}}(t-x).
\end{aligned}
$$

Thus, $\underline{\tilde{L}}_{\gamma}(t)=\underline{\bar{F}}(t)+\int_0^t\underline{\chi}(\zeta(\gamma;t|x))^{\top}\underline{\tilde{L}}_{\gamma e^{-D(x)}}(t-x)\underline{F}(\mathrm{d}x)$ as claimed.

Taking $\frac{\partial^n}{\partial\gamma^n}\tilde{L}_{\gamma|l}(t)|_{\gamma=0}$, we get the analog of (Landriault *et al.*, 2017, Eqn (3.2)):

$$
\mathrm{E}\Big(Z_l^n(t)\Big)=\int_0^t e^{-nD(x)}\sum_{m\in\mathfrak{I}}\chi_m(1)\,\mathrm{E}(Z_m^n(t-y))F_l(\mathrm{d}x)+v_{n|l}(t);\tag{3.5}
$$

$$
v_{n|l}(t)=\sum_{q=1}^{n}\binom{n}{q}\int_0^t e^{-nD(x)}\sum_{m\in\mathfrak{I}}\mathrm{E}\Big(Z_m^{n-q}(t-x)\Big)\sum_{k=1}^{q}\chi_m^{(k)}(1)B_{q,k}\Big(\underline{\xi}(x,t)^{\top}\Big)F_l(\mathrm{d}x).
$$

The dimension of $\underline{\xi}(x,t)$ is $q-k+1$, $B_{q,k}(\cdots)$ is the Bell polynomial, and $\xi_i(x,t)$ is given in Section 3.3. To verify (3.5), note that $\frac{\partial^n}{\partial\gamma^n}a(\gamma)b(\gamma) = \sum_{i=0}^n \binom{n}{i} a^{(i)}(\gamma)b^{(n-i)}(\gamma)$ somewhat like the binomial expansion. This applies to each summand in $\underline{\chi}(\zeta(\gamma;t|x))^\top \underline{\tilde{L}}_{\gamma e^{-D(x)}}(t-x)$, where $a(\gamma) = \chi_m(\zeta(\gamma;t|x))$ and $b(\gamma) = \tilde{L}_{\gamma e^{-D(x)}|m}(t-x)$. Using "Faà di Bruno's formula" (e.g. (Johnson, 2002, Eqn (2.2))), we get

$$\frac{\partial^i}{\partial\gamma^i}\chi_m(\zeta(\gamma;t|x)) = \sum_{q=0}^{i} \chi_m^{(q)}(\zeta(\gamma;t|x))B_{i,q}\Big(\underline{\nabla}\zeta(\gamma;t|x)^\top\Big).$$

Here, $\underline{\nabla}\zeta(\gamma;t|x)$ has dimension $i-q+1$, and the pth entry is $\frac{\partial^p}{\partial\gamma^p}\zeta(\gamma;t|x)$. Therefore, $\frac{\partial^n}{\partial\gamma^n}\tilde{L}_{\gamma|l}(t)$ becomes

$$\begin{aligned}(-1)^n \operatorname{E}\Big(Z_l^n(t)e^{-\gamma Z_l(t)}\Big) = \sum_{m\in\Im}\int_0^t \sum_{i=0}^{n}\sum_{q=0}^{i} \chi_m^{(q)}(\zeta(\gamma;t|x))B_{i,q}\Big(\underline{\nabla}\zeta(\gamma;t|x)^\top\Big)\\ \times\Big(-e^{-D(x)}\Big)^{n-i}\operatorname{E}\Big(Z_m^{n-i}(t)e^{-\gamma e^{-D(x)}Z_m(t-x)}\Big)F_l(\mathrm{d}x).\end{aligned}$$

Setting $\gamma=0$, followed by some basic algebra noting $B_{0,0}(\cdot)\equiv 1$ and that $0<i=\sum_{j=1}^{i-q+1} jb_j$ requires at least one $b_j>0$, gives (3.5).

The following theorem generalizes (Landriault *et al.*, 2017, Theorem 1).

Theorem 2. *Let $n\in\mathbb{N}^+$. Assume $D(x)$ is such that $\int_0^\infty e^{-nD(x)}F_l(\mathrm{d}x)\in(0,1]$ for all $l\in\Im$ and $n\in\mathbb{N}^+$. Then, for each $l\in\Im$, we have:*

$$\operatorname{E}\Big(Z_l^n(t)\Big) = v_{n|l}(t) + \int_0^t \underline{\chi}(1)^\top\underline{v}_n(t-x)\sum_{q=0}^{\infty}\Big(F_{n\delta|l} * F_{n\delta}^{*q}\Big)(\mathrm{d}x). \tag{3.6}$$

Proof of Theorem 2. The assumption on $D(x)$ ensures that (3.5) is a defective or proper renewal equation for all $l\in\Im$. Furthermore, $\int_0^\infty F_{n\delta}(\mathrm{d}x) = \underline{\chi}(1)^\top\int_0^\infty e^{-nD(x)}\underline{F}(\mathrm{d}x) \le \underline{\chi}(1)^\top\underline{1} = 1$, because of the assumption on $D(x)$. Setting $M_{n|l}(t) = \operatorname{E}\Big(Z_l^n(t)\Big)$, we write (3.5) in matrix form as $\underline{M}_n(t) = \int_0^t \underline{\chi}(1)^\top\underline{M}_n(t-x)\underline{F}_{n\delta}(\mathrm{d}x) + \underline{v}_n(t)$. Taking Laplace transforms with respect to t, we get

$$\tilde{\underline{M}}_n(s) = \underline{\chi}(1)^\top \tilde{\underline{M}}_n(s) \tilde{\underline{F}}_{n\delta}(s) + \tilde{\underline{v}}_n(s) = \tilde{\underline{F}}_{n\delta}(s) \underline{\chi}(1)^\top \tilde{\underline{M}}_n(s) + \tilde{\underline{v}}_n(s),$$

as $\underline{\chi}(1)^\top \tilde{\underline{M}}_n(s)$ is a scalar. Here, $\tilde{M}_{n|l}(s) = \int_0^\infty e^{-st} M_{n|l}(t)\,\mathrm{d}t$ and likewise $\tilde{v}_{n|l}(s) = \int_0^\infty e^{-st} v_{n|l}(t)\,\mathrm{d}t$. Now, since $\tilde{\underline{F}}_{n\delta}(s)\underline{\chi}(1)^\top$ is a matrix, rearranging gives $\tilde{\underline{M}}_n(s) = \left(\underline{I} - \tilde{\underline{F}}_{n\delta}(s)\underline{\chi}(1)^\top\right)^{-1} \tilde{\underline{v}}_n(s)$. Then

$$\begin{aligned}
\tilde{\underline{M}}_n(s) &= \sum_{q=0}^{\infty} \left(\tilde{\underline{F}}_{n\delta}(s)\underline{\chi}(1)^\top\right)^q \tilde{\underline{v}}_n(s) \\
&= \tilde{\underline{v}}_n(s) + \sum_{q=1}^{\infty} \left(\tilde{\underline{F}}_{n\delta}(s)\underline{\chi}(1)^\top\right)^q \tilde{\underline{v}}_n(s) \\
&= \tilde{\underline{v}}_n(s) + \tilde{\underline{F}}_{n\delta}(s) \sum_{q=1}^{\infty} \left(\underline{\chi}(1)^\top \tilde{\underline{F}}_{n\delta}(s)\right)^{q-1} \underline{\chi}(1)^\top \tilde{\underline{v}}_n(s) \\
&= \tilde{\underline{v}}_n(s) + \frac{\underline{\chi}(1)^\top \tilde{\underline{v}}_n(s) \tilde{\underline{F}}_{n\delta}(s)}{1 - \underline{\chi}(1)^\top \tilde{\underline{F}}_{n\delta}(s)}.
\end{aligned}$$

In componentwise form, $\tilde{M}_{n|l}(s) = \tilde{v}_{n|l}(s) + \underline{\chi}(1)^\top \tilde{\underline{v}}_n(s) \frac{\tilde{F}_{n\delta|l}(s)}{1-\underline{\chi}(1)^\top \tilde{\underline{F}}_{n\delta}(s)}$; (3.6) follows upon Laplace transform inversion. □

We can simplify Theorem 2 if we let either the discount rate be constant, or if we assume there is only one risk level (or possible df for each $\tau_{k+1} \mid C_{T_k}$). Both cases of Corollary 2 still generalize (Landriault *et al.*, 2017, Theorem 1).

Corollary 2. *Let $n \in \mathbb{N}^+$.*

1. *Let the discount rate be constant, namely $\delta(\cdot) = \delta$. Then, for each $l \in \Im$:*

$$\mathrm{E}\left(Z_l^n(t)\right) = v_{n|l}(t) + \int_0^t e^{-n\delta x} \sum_{m \in \Im} \chi_m(1) v_{n|m}(t-x) H_l(\mathrm{d}x). \tag{3.7}$$

Here, $H_l(t) = \mathrm{E}(N_l(t))$ is the Markov renewal function associated with initial risk class l.

2. *Let* $\tau_k \overset{\text{iid}}{\sim} F(\cdot)$ *for* $k \in \mathbb{N}^+$*, that is,* $|\Im| = 1$*. Also assume* $\int_0^\infty e^{-nD(x)} F(\mathrm{d}x) \in (0, 1]$ *for all* $n \in \mathbb{N}^+$*. Then,*

$$E(Z^n(t)) = v_n(t) + \int_0^t v_n(t-x) \sum_{q=1}^{\infty} F_{n\delta}^{*q}(\mathrm{d}x). \tag{3.8}$$

Here, $F_{n\delta}^{*q}(\mathrm{d}x)$ *is the q-fold convolution of* $F_{n\delta}(\mathrm{d}x) = e^{-nD(x)} F(\mathrm{d}x)$ *with itself.*

In either case, we set $\mathrm{E}\Big(Z_l^0(t)\Big) \equiv 1, t > 0$.

In Corollary 2 (1), note that $\mathrm{E}\Big(Z_l^n(t)\Big)$ depends on $\mathrm{E}\Big(Z_m^q(t)\Big)$, $q \in \{1, \ldots, n-1\}$, $m \in \Im$, which includes all possible initial distributions of τ_1. Writing (3.7) of Corollary 2 in matrix form results in $\underline{M}_n(t) = \underline{v}_n(t) + \int_0^t e^{-n\delta x} \underline{\chi}(1)^\top \underline{v}_n(t-x) \underline{H}(\mathrm{d}x)$. When $|\Im| = \mathfrak{m} = 1$, $\underline{\chi}(1)^\top$ becomes $B(1) = 1$, and $\underline{M}_n(\cdot)$, $\underline{v}_n(\cdot)$, and $\underline{H}(\cdot)$ revert to each one's scalar form, retrieving (Landriault *et al.*, 2017, Theorem 1). In Corollary 2 (2), we also retrieve (Landriault *et al.*, 2017, Theorem 1). For, setting $\delta(\cdot) = \delta$, $D(x) = \delta x$, and since $\int_0^\infty e^{-sx} F_{n\delta}(\mathrm{d}x) = \tilde{F}(s + n\delta)$, it follows that $\int_0^\infty e^{-sx} \sum_{q=1}^\infty F_{n\delta}^{*q}(\mathrm{d}x) = \frac{\tilde{F}(s+n\delta)}{1-\tilde{F}(s+n\delta)}$, which is precisely $\tilde{H}(s + n\delta) = \int_0^\infty e^{-sx} e^{-n\delta x} H(\mathrm{d}x)$.

Proof of Corollary 2. For both cases, the methodology involves Laplace-transforming (3.8), rearrangement, and Laplace inversion.

1. Taking Laplace transforms in (3.5) and writing things in matrix form, we have:

$$\tilde{\underline{M}}_n(s) = \tilde{\underline{F}}(s + n\delta) \underline{\chi}(1)^\top \tilde{\underline{M}}_n(s) + \tilde{\underline{v}}_n(s).$$

Solving this in the manner of deriving $\tilde{H}_l(s)$ above, we get

$$\tilde{\underline{M}}_n(s) = \tilde{\underline{v}}_n(s) + \underline{\chi}(1)^\top \tilde{\underline{v}}_n(s) \tilde{\underline{H}}(s + n\delta).$$

Inverting each component Laplace transform in this latter gives the expression (3.7).

2. When $\Im = \{1\}$, (3.5) becomes

$$\mathrm{E}(Z^n(t)) = \int_0^t e^{-nD(x)}\,\mathrm{E}(Z^n(t-x))F(\mathrm{d}x) + v_n(t);$$
$$v_n(t) = \sum_{q=1}^{n} \binom{n}{q} \int_0^t e^{-nD(x)}\,\mathrm{E}(Z^{n-q}(t-x)) \sum_{k=1}^{q} B^{(k)}(1) B_{q,k}\left(\underline{\xi}(x,t)^\top\right) F(\mathrm{d}x).$$

Taking Laplace transforms again, now $\tilde{M}_n(s) = \tilde{M}_n(s)\tilde{F}_{n\delta}(s) + \tilde{v}_n(s)$; rearranged, this is $\tilde{M}_n(s) = \tilde{v}_n(s) + \tilde{v}_n(s)\sum_{q=1}^{\infty}\left(\tilde{F}_{n\delta}(s)\right)^q$. Inverting this Laplace-transformed equation gives (3.8).

□

3.5. JOINT MOMENTS OF IBNR AND IR CLAIMS

Now we shall generalize (Landriault *et al.*, 2017, Theorem 3) in the same manner. For incurred and reported claims, we define

$$Z_{ir|l}(t) = \sum_{k=1}^{N_l(t)} Y_{ir|k}(t), \tag{3.9}$$
$$Y_{ir|k}(t) \sum_{i=1}^{C_{T_k}} e^{-(D(T_k+\tau_0)-D(\tau_0))} l(W_{i,k})\,\mathrm{I}(W_{i,k}+T_k \le t) X_{i,k}.$$

Letting $\Delta \ge 0$, for $u, v \ge 0$ we define a joint Laplace transform:

$$\tilde{L}_{u,v|l}(t,\Delta) = \mathrm{E}(\exp(-uZ_{ir|l}(t) - vZ_l(t+\Delta))). \tag{3.10}$$

Next we find an expression for (3.10) which will be suitable for the generalized form of (Landriault *et al.*, 2017, Eqn (3.17)). Starting,

$$\begin{aligned}\tilde{L}_{u,v|l}(t,\Delta) &= \mathrm{E}(\exp(-uZ_{ir|l}(t) - vZ_l(t+\Delta)))\\ &= \mathrm{E}(\mathrm{E}(\exp(-uZ_{ir|l}(t) - vZ_l(t+\Delta))|\sigma(N_l(t+\Delta)+1)))\end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \mathrm{E}\left(\prod_{k=1}^{N_l(t)} \exp(-uY_{ir|k}(t)) \prod_{k=1}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \middle| \tau_1 = x \right) F_l(\mathrm{d}x) \\
&= \bar{F}_l(t+\Delta) + \int_t^{t+\Delta} \mathrm{E}\left(\prod_{k=1}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \middle| \tau_1 = x \right) F_l(\mathrm{d}x) \\
&\quad + \int_0^t \mathrm{E}\left(\prod_{k=1}^{N_l(t)} \exp(-uY_{ir|k}(t)) \prod_{k=1}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \middle| \tau_1 = x \right) F_l(\mathrm{d}x).
\end{aligned}
$$

Taking the integrand in the second term,

$$
\begin{aligned}
&\mathrm{E}\left(\prod_{k=1}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \middle| \tau_1 = x \right) \\
&= \mathrm{E}\left(\mathrm{E}\left(\exp(-vY_1(t+\Delta)) \prod_{k=2}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \middle| \tau_1 = x, C_x = C \right) \right) \\
&= \sum_{m \in \Im} \mathrm{E}\left(\mathrm{E}\left(\exp(-vY_1(t+\Delta)) \prod_{k=2}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \, \mathrm{I}(C \in \mathscr{I}_m) \middle| \tau_1 = x, C_x = C \right) \right) \\
&= \sum_{m \in \Im} \mathrm{E}\big(\mathrm{E}(\exp(-vY_1(t+\Delta)) | \tau_1 = x, C_x = C) \\
&\quad \times \mathrm{E}\left(\prod_{k=2}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \, \mathrm{I}(C \in \mathscr{I}_m) \middle| \tau_1 = x, C_x = C \right) \big).
\end{aligned}
$$

Now,

$$
\begin{aligned}
&\mathrm{E}(\exp(-vY_1(t+\Delta)) | \tau_1 = x, C_x = C) \\
&= \mathrm{E}\left(\exp\left(-v \sum_{i=1}^{C} e^{-D(x)} l(W_{i,1}) \, \mathrm{I}(W_{i,1} > t + \Delta - x) X_{i,1} \right) \middle| \tau_1 = x, C_x = C \right) \\
&= \mathrm{E}\left(\prod_{i=1}^{C} \exp\left(-v e^{-D(x)} l(W_{i,1}) \, \mathrm{I}(W_{i,1} > t + \Delta - x) X_{i,1} \right) \middle| \tau_1 = x, C_x = C \right) \\
&= \left(\mathrm{E}\left(\exp\left(-v e^{-D(x)} l(W_{1,1}) \, \mathrm{I}(W_{1,1} > t + \Delta - x) X_{1,1} \right) \middle| \tau_1 = x \right) \right)^C \\
&= (\zeta(v; t + \Delta | x))^C.
\end{aligned}
$$

Then, continuing,

$$
\begin{aligned}
&\mathrm{E}\left(\prod_{k=1}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \middle| \tau_1 = x\right) \\
&= \sum_{m\in\mathfrak{I}} \mathrm{E}\left((\zeta(v; t+\Delta|x))^C \, \mathrm{E}\left(\prod_{k=2}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \, \mathrm{I}(C \in \mathscr{I}_m) \middle| \tau_1 = x, C_x = C\right)\right).
\end{aligned}
$$

Now, by regenerativity of $N_l(t+\Delta)$ at $\tau_1 = x$,

$$
\begin{aligned}
&\mathrm{E}\left(\prod_{k=2}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \, \mathrm{I}(C \in \mathscr{I}_m) \middle| \tau_1 = x, C_x = C\right) \\
&= \mathrm{E}\left(\prod_{k=2}^{N_m(t+\Delta-x)+1} \exp\left(-ve^{-D(x)} Y_k(t+\Delta-x)\right) \middle| \tau_1 = x\right) \mathrm{I}(C \in \mathscr{I}_m) \\
&= \tilde{L}_{ve^{-D(x)}|m}(t+\Delta-x) \, \mathrm{I}(C \in \mathscr{I}_m).
\end{aligned}
$$

That is,

$$
\mathrm{E}\left(\prod_{k=1}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \middle| \tau_1 = x\right) = \sum_{m\in\mathfrak{I}} \chi_m(\zeta(v; t+\Delta|x)) \tilde{L}_{ve^{-D(x)}|m}(t+\Delta-x).
$$

Defining

$$
\begin{aligned}
&\zeta(u, v; t, \Delta|x) \\
&= \mathrm{E}\left(\exp\left(-e^{-D(x)} l(W_{1,1}) X_{1,1} (u \, \mathrm{I}(W_{1,1} \le t-x) + v \, \mathrm{I}(W_{1,1} > t+\Delta-x))\right) \middle| \tau_1 = x\right),
\end{aligned}
$$

we may similarly show

$$
\begin{aligned}
&\int_0^t \mathrm{E}\left(\prod_{k=1}^{N_l(t)} \exp(-uY_{ir|k}(t)) \prod_{k=1}^{N_l(t+\Delta)} \exp(-vY_k(t+\Delta)) \middle| \tau_1 = x\right) F_l(\mathrm{d}x) \\
&= \int_0^t \underline{\chi}(\zeta(u, v; t, \Delta|x))^\top \underline{\tilde{L}}_{ue^{-D(x)}, ve^{-D(x)}}(t-x, \Delta) F_l(\mathrm{d}x),
\end{aligned}
$$

finally getting

$$\tilde{L}_{u,v|l}(t,\Delta) = \bar{F}_l(t+\Delta) + \int_t^{t+\Delta} \underline{\chi}\big(\zeta(v;t+\Delta|x)\big)^\top \underline{\tilde{L}}_{ve^{-D(x)}}(t+\Delta-x)F_l(\mathrm{d}x) \qquad (3.11)$$
$$+ \int_0^t \underline{\chi}\big(\zeta(u,v;t,\Delta|x)\big)^\top \underline{\tilde{L}}_{ue^{-D(x)},ve^{-D(x)}}(t-x,\Delta)F_l(\mathrm{d}x).$$

The first term corresponds to $\tau_1 = x \in (t+\Delta,\infty)$, the second to $\tau_1 = x \in (t,t+\Delta]$, and the third to $\tau_1 = x \in (0,t]$. In the first term, $N_l(t+\Delta) = N_l(t) = 0$; in the second term, $N_l(t) = 0$. In these cases, then, having zero claim-causing events means the total claims, $Z_l(t+\Delta)$ and $Z_{ir|l}(t)$, are zero also.

Denoting $M_{m,n|l}(t;\Delta) = E\Big(Z^m_{ir|l}(t)Z^n_l(t+\Delta)\Big)$, analogously to (Landriault *et al.*, 2017, Eqn (3.17)) we get that

$$M_{m,n|l}(t;\Delta) = \int_0^t e^{-(m+n)D(x)}\underline{\chi}(1)^\top \underline{M}_{m,n}(t-x;\Delta)F_l(\mathrm{d}x) + v_{m,n|l}(t;\Delta); \qquad (3.12)$$
$$v_{m,n|l}(t;\Delta) = \sum_{\substack{j=0\\ i+j>0}}^n \sum_{i=0}^m \binom{m}{i}\binom{n}{j}\int_0^t e^{-(m+n)D(x)}\underline{M}_{m-i,n-j}(t-x;\Delta)^\top \underline{B}^*_{i,j}(x;t,t+\Delta)F_l(\mathrm{d}x),$$
$$\underline{B}^*_{i,j}(x;t_1,t_2) = \sum_{k=1\wedge i}^{i}\sum_{l=1\wedge j}^{j}\underline{\chi}^{(k+l)}(1)B_{i,k}\Big(\underline{\eta}(x,t_1)^\top\Big)B_{j,l}\Big(\underline{\xi}(x,t_2)^\top\Big), 0 \le t_1 \le t_2.$$

The dimension of the vector $\underline{\eta}(x,t_1)$ is $i-k+1$, and likewise $\underline{\xi}(x,t_2)$ has dimension $j-l+1$. The p-th entry of the latter is given by (3.1), and the p-th entry of the former is given by (3.2). To get (3.12), in (3.11) we only need to concern ourselves with the third term on the right-hand side, since we will take the $\frac{\partial^{m+n}}{\partial u^m \partial v^n}$ derivative, with $m \ge 1$. That is,

$$\frac{\partial^{m+n}}{\partial u^m \partial v^n}\tilde{L}_{u,v|l}(t,\Delta) = \int_0^t \frac{\partial^{m+n}}{\partial u^m \partial v^n}\underline{\chi}\big(\zeta(u,v;t,\Delta|x)\big)^\top \underline{\tilde{L}}_{ue^{-D(x)},ve^{-D(x)}}(t-x,\Delta)F_l(\mathrm{d}x).$$

Now, note that $\frac{\partial^{m+n}}{\partial u^m \partial v^n}\zeta(u,v;t,\Delta|x) = 0$ if $m \wedge n \ge 1$, because for all $\Delta \ge 0$,

$$\{W \le t-x\} \cap \{W > t+\Delta-x\} = \emptyset.$$

Therefore, for $\iota \in \Im$,

$$\frac{\partial^{m+n}}{\partial u^m \partial v^n}\chi_\iota(\zeta(u,v;t,\Delta|x))$$
$$= \frac{\partial^m}{\partial u^m}\left(\frac{\partial^n}{\partial v^n}\chi_\iota(\zeta(u,v;t,\Delta|x))\right) = \frac{\partial^n}{\partial v^n}\left(\frac{\partial^m}{\partial u^m}\chi_\iota(\zeta(u,v;t,\Delta|x))\right).$$

First, though, the multivariate product rule (mentioned for example in Constantine and Savits (1996)) gives that

$$\frac{\partial^{m+n}}{\partial u^m \partial v^n}\chi_\iota(\zeta(u,v;t,\Delta|x))\tilde{L}_{ue^{-D(x)},ve^{-D(x)}|\iota}(t-x,\Delta)$$
$$= \sum_{i=0}^{m}\sum_{j=0}^{n}\binom{m}{i}\binom{n}{j}\frac{\partial^{i+j}}{\partial u^i \partial v^j}\chi_\iota(\zeta(u,v;t,\Delta|x))\frac{\partial^{m+n-(i+j)}}{\partial u^{m-i}\partial v^{n-j}}\tilde{L}_{ue^{-D(x)},ve^{-D(x)}|\iota}(t-x,\Delta).$$

Then, using the so-called “Faà di Bruno” formula twice,

$$\frac{\partial^{i+j}}{\partial u^i \partial v^j}\chi_\iota(\zeta(u,v;t,\Delta|x)) = \frac{\partial^i}{\partial u^i}\left(\frac{\partial^j}{\partial v^j}\chi_\iota(\zeta(u,v;t,\Delta|x))\right)$$
$$= \frac{\partial^i}{\partial u^i}\left(\sum_{p=0}^{j}\chi_\iota^{(p)}(\zeta(u,v;t,\Delta|x))B_{j,p}(\underline{\nabla}^v\zeta(u,v;t,\Delta|x))\right)$$
$$= \sum_{p=0}^{j}\sum_{r=0}^{i}\binom{i}{r}\frac{\partial^r}{\partial u^r}\chi_\iota^{(p)}(\zeta(u,v;t,\Delta|x))\frac{\partial^{i-r}}{\partial u^{i-r}}B_{j,p}(\underline{\nabla}^v\zeta(u,v;t,\Delta|x))$$
$$= \sum_{p=0}^{j}\sum_{q=0}^{i}\chi_\iota^{(p+q)}(\zeta(u,v;t,\Delta|x))B_{i,q}(\underline{\nabla}^u\zeta(u,v;t,\Delta|x))B_{j,p}(\underline{\nabla}^v\zeta(u,v;t,\Delta|x)).$$

The notation $\underline{\nabla}^u$ means $\left(\frac{\partial^s}{\partial u^s}\right)$, $s = 1,\ldots,i-q+1$, and $\underline{\nabla}^v$ means $\left(\frac{\partial^s}{\partial v^s}\right)$, $s = 1,\ldots,j-p+1$. In the third to fourth lines above, we use that for $r = 0,\ldots,i-1$, $\frac{\partial^{i-r}}{\partial u^{i-r}}B_{j,p}(\underline{\nabla}^v\zeta(u,v;t,\Delta|x)) = 0$. Thus it holds that

$$\frac{\partial^{m+n}}{\partial u^m \partial v^n}\tilde{L}_{u,v|l}(t,\Delta) = \int_0^t \sum_{\iota\in\Im}\sum_{i=0}^{m}\sum_{j=0}^{n}\binom{m}{i}\binom{n}{j}\sum_{p=0}^{j}\sum_{q=0}^{i}\chi_\iota^{(p+q)}(\zeta(u,v;t,\Delta|x))$$
$$B_{i,q}(\underline{\nabla}^u\zeta(u,v;t,\Delta|x))B_{j,p}(\underline{\nabla}^v\zeta(u,v;t,\Delta|x))\frac{\partial^{m+n-(i+j)}}{\partial u^{m-i}\partial v^{n-j}}\tilde{L}_{ue^{-D(x)},ve^{-D(x)}|\iota}(t-x,\Delta)F_l(\mathrm{d}x).$$

Setting $(u,v)=(0,0)$ and recalling $j=\sum_{s=1}^{j-p+1} sj_s$ and $i=\sum_{s=1}^{i-q+1} sj_s$, (3.12) follows. Now we are ready to generalize (Landriault *et al.*, 2017, Theorem 3).

Theorem 3. *Let $n \in \mathbb{N}$, $m \in \mathbb{N}^+$. Assume $D(x)$ is such that $\int_0^\infty e^{-qD(x)} F_l(\mathrm{d}x) \in (0,1]$ for all $l \in \Im$ and all $q \in \mathbb{N}^+$. Then for each $l \in \Im$, we have:*

$$\begin{aligned}&\mathrm{E}\Big(Z^m_{ir|l}(t) Z^n_l(t+\Delta)\Big) \qquad (3.13)\\ &= v_{m,n|l}(t;\Delta) + \int_0^t \underline{\chi}(1)^\top \underline{v}_{m,n}(t-x;\Delta) \sum_{q=0}^{\infty} \Big(F_{(m+n)\delta|l} * F^{*q}_{(m+n)\delta}\Big)(\mathrm{d}x).\end{aligned}$$

Proof of Theorem 3. We may express each component of (3.12) as

$$M_{m,n|l}(t;\Delta) = \int_0^t \underline{\chi}(1)^\top \underline{M}_{m,n}(t-x;\Delta) F_{(m+n)\delta|l}(\mathrm{d}x) + v_{m,n|l}(t;\Delta),$$

which is a renewal equation by our assumptions on $D(x)$. So

$$\begin{aligned}\underline{M}_{m,n}(t;\Delta) &= \int_0^t \underline{\chi}(1)^\top \underline{M}_{m,n}(t-x;\Delta) \underline{F}_{(m+n)\delta}(\mathrm{d}x) + \underline{v}_{m,n}(t;\Delta)\\ &= \int_0^t \underline{Q}_{(m+n)\delta}(\mathrm{d}x) \underline{M}_{m,n}(t-x;\Delta) + \underline{v}_{m,n}(t;\Delta)\end{aligned}$$

is a Markov renewal equation. That is,

$$\big[Q_{(m+n)\delta}(\mathrm{d}x)\big]_{ij} = \chi_j(1) F_{(m+n)\delta|i}(\mathrm{d}x) = e^{-(m+n)D(x)} \chi_j(1) F_i(\mathrm{d}x);$$

recall (3.3) given above. Taking Laplace transforms of (3.12) wrt t gives $\tilde{\underline{M}}_{m,n}(s;\Delta) = \underline{\chi}(1)^\top \tilde{\underline{M}}_{m,n}(s;\Delta) \tilde{\underline{F}}_{(m+n)\delta}(s) + \tilde{\underline{v}}_{m,n}(s;\Delta)$. Now, $\underline{\chi}(1)^\top \tilde{\underline{M}}_{m,n}(s;\Delta)$ being a scalar, we may write

$$\tilde{\underline{M}}_{m,n}(s;\Delta) = \tilde{\underline{F}}_{(m+n)\delta}(s) \underline{\chi}(1)^\top \tilde{\underline{M}}_{m,n}(s;\Delta) + \tilde{\underline{v}}_{m,n}(s;\Delta),$$

which after matrix arithmetic becomes $\underline{\tilde{M}}_{m,n}(s;\Delta) = \left(\underline{I} - \underline{\tilde{F}}_{(m+n)\delta}(s)\underline{\chi}(1)^\top\right)^{-1}\underline{\tilde{v}}_{m,n}(s;\Delta)$. Then,

$$\begin{aligned}
\underline{\tilde{M}}_{m,n}(s;\Delta) &= \left(\underline{I} - \underline{\tilde{F}}_{(m+n)\delta}(s)\underline{\chi}(1)^\top\right)^{-1}\underline{\tilde{v}}_{m,n}(s;\Delta) \\
&= \sum_{i=0}^{\infty}\left(\underline{\tilde{F}}_{(m+n)\delta}(s)\underline{\chi}(1)^\top\right)^{i}\underline{\tilde{v}}_{m,n}(s;\Delta) \\
&= \underline{\tilde{v}}_{m,n}(s;\Delta) + \sum_{i=1}^{\infty}\left(\underline{\tilde{F}}_{(m+n)\delta}(s)\underline{\chi}(1)^\top\right)^{i}\underline{\tilde{v}}_{m,n}(s;\Delta) \\
&= \underline{\tilde{v}}_{m,n}(s;\Delta) + \underline{\tilde{F}}_{(m+n)\delta}(s)\sum_{i=1}^{\infty}\left(\underline{\chi}(1)^\top\underline{\tilde{F}}_{(m+n)\delta}(s)\right)^{i-1}\underline{\chi}(1)^\top\underline{\tilde{v}}_{m,n}(s;\Delta) \\
&= \underline{\tilde{v}}_{m,n}(s;\Delta) + \underline{\tilde{F}}_{(m+n)\delta}(s)\frac{1}{1-\underline{\chi}(1)^\top\underline{\tilde{F}}_{(m+n)\delta}(s)}\underline{\chi}(1)^\top\underline{\tilde{v}}_{m,n}(s;\Delta) \\
&= \underline{\tilde{v}}_{m,n}(s;\Delta) + \underline{\chi}(1)^\top\underline{\tilde{v}}_{m,n}(s;\Delta)\frac{\underline{\tilde{F}}_{(m+n)\delta}(s)}{1-\underline{\chi}(1)^\top\underline{\tilde{F}}_{(m+n)\delta}(s)}.
\end{aligned}$$

Componentwise Laplace transform inversion of the final line gives (3.13). □

We may also specialize our Theorem 3 in the same two ways we did so with our Theorem 2; both cases in the following Corollary 3 still generalize (Landriault *et al.*, 2017, Theorem 3).

Corollary 3. *Let $n \in \mathbb{N}$, $m \in \mathbb{N}^+$.*

1. *Let $\delta(\cdot) = \delta$, that is, a constant discount rate. Then, for each $l \in \Im$,*

$$\mathrm{E}\left(Z^m_{ir|l}(t)Z^n_l(t+\Delta)\right) = v_{m,n|l}(t;\Delta) + \int_0^t e^{-(m+n)\delta x}\underline{\chi}(1)^\top\underline{v}_{m,n}(t-x;\Delta)H_l(\mathrm{d}x).$$

2. *Let $\tau_k \overset{\text{iid}}{\sim} F(\cdot)$ for $k \in \mathbb{N}^+$, such that $|\Im| = 1$. Also assume $\int_0^\infty e^{-qD(x)}F(\mathrm{d}x) \in (0,1]$ for all $q \in \mathbb{N}^+$. Then,*

$$\mathrm{E}\left(Z^m_{ir}(t)Z^n(t+\Delta)\right) = v_{m,n}(t;\Delta) + \int_0^t v_{m,n}(t-x;\Delta)\sum_{q=1}^{\infty}F^{*q}_{(m+n)\delta}(\mathrm{d}x).$$

In either case, we set $\mathrm{E}\Big(Z^0_{ir|l}(t)\Big) \equiv 1, t > 0.$

The meanings of $H_l(\cdot)$ and $F^{*q}_{(m+n)\delta}(\cdot)$ are the same here as in Corollary 2. We omit the details of proving Corollary 3, since they are essentially the same as the proof of Corollary 2. Likewise, the two cases in Corollary 3 retrieve (Landriault *et al.*, 2017, Theorem 3) in the same ways that the cases of Corollary 2 retrieve (Landriault *et al.*, 2017, Theorem 1). In our Theorem 3 and Corollary 3 (1), the expressions given for the joint moments of $Z_{ir|l}(t)$ and $Z_l(t+\Delta)$ depend on the lower joint moments of $Z_{ir|m}(t)$ and $Z_m(t+\Delta)$ for all risk classes $m \in \mathfrak{I}$.

3.6. NUMERICAL EXAMPLES OF TIME-VARYING DISCOUNT RATES

Now we numerically illustrate the failure of the existing discount rate model to capture other possible scenarios. Kennedy (1992) remarked that the “valuation function” $v(x) \equiv e^{-D(x)}$ satisfies the differential equation $\frac{\mathrm{d}}{\mathrm{d}x}v(x) + \delta(x)v(x) = 0$ with initial condition $v(0) = 1$. Saying $v(0) = 1$ of course is equivalent to saying $D(x) = 0$. Trivially, $\delta(x) = -\frac{b}{a+bx}$ and $v(x) = a + bx$ satisfy said equation, and the initial condition gives $a = 1$. If we require $\delta(u) \geq 0$ for some $u \in [0, \infty)$, then we must have $b \leq 0$ and $u \in \Big[0, -\frac{1}{b}\Big)$, where we say $\lim_{b \uparrow 0} -\frac{1}{b} = \infty$. With this choice of $\delta(u)$ so constrained, it holds that $D(x) = \int_0^x \delta(u)\,\mathrm{d}u = -\log(1+bx)$ assumes values in $\mathbb{R}$. On the other hand, if we regard the operations in $v(x) = e^{-\int_0^x \delta(u)\,\mathrm{d}u}$ as taking place in $\mathbb{C}$, then the only constraint we have is that of $\int_0^\infty (v(x))^n F_l(\mathrm{d}x) \in (0, 1]$ for $l \in \mathfrak{I}$ (so that the renewal equations in Theorems 2 and 3 are proper or defective). In the case of $v(x) = 1 + bx$, this implies $\sum_{q=0}^n \binom{n}{q} b^q \int_0^\infty x^q F_l(\mathrm{d}x) \leq 1$, from which it follows that $b \leq 0$ still, albeit without only holding on the interval $t \in \Big[0, -\frac{1}{b}\Big)$.

To show the effects of having $v(x) = 1 + bx$ instead of $v(x) = e^{-\delta x}$, we examine a special case of Theorem 2 considered by Landriault *et al.* (2017). Namely, we let there be just one class of claims produced ($|\mathfrak{I}| = 1$), and we let the batch sizes be 1 ($B(z) = z$). We assume the dependency $J(t, w, x) = F(t)K(w)P_{X|W}(x|w)$, where $F(\mathrm{d}x) = \lambda^2 x e^{-\lambda x}\mathrm{d}x$ is Erlang-2, $K(\mathrm{d}w) = \theta e^{-\theta w}$ is exponential, and the function $l(w) = e^{-\epsilon w}$. Landriault *et al.* (2017) assumed the following parameter values: $\lambda = 3$, $\theta = 0.5$, $\mu_1 = 1$, and $\epsilon = 0.06$. They let $D(x) = \delta x$ with $\delta = 0.05$, whereas we let $D(x) = -\log(1 + bx)$ with multiple values of b: -2δ, $-\delta$, and $-\frac{\delta}{2}$. The quantity we use to illustrate our point is $\mathrm{E}(Z(t)) = M_1(t)$ (which Landriault *et al.* (2017) also gave explicitly). As we see, multiple departures may happen from the case of $\delta(\cdot) = \delta$. Again, we have used the algorithms of Abate and Valkó (2004); Trefethen *et al.* (2006). We point out that by choosing $b = -\delta$ here, a constant discount-rate assumption slightly over-projects the values of $\mathrm{E}(Z(t))$ given by $v(x) = 1-\delta x$, while such an assumption ($v(x) = e^{-\delta x}$) greatly over-projects $\mathrm{E}(Z(t))$ for $b = -2\delta$, and greatly under-projects $\mathrm{E}(Z(t))$ for $b = -\frac{\delta}{2}$.

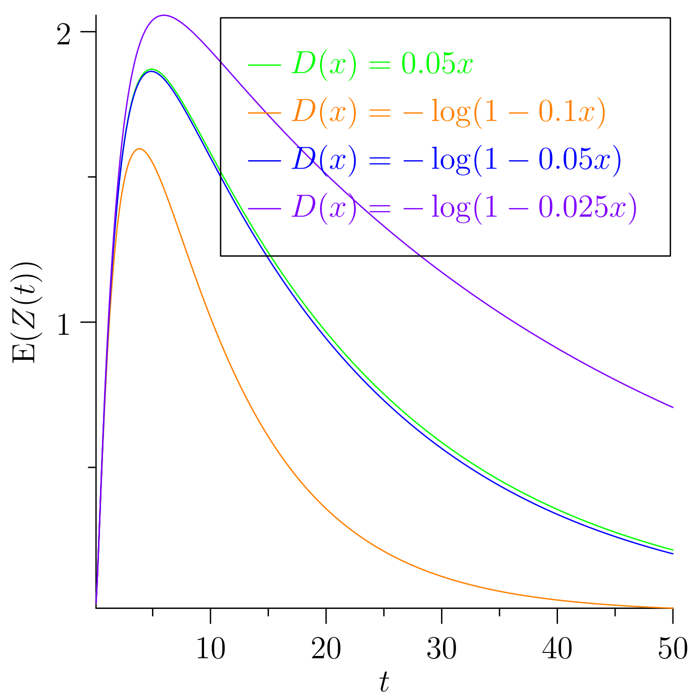

Figure 3.1. Time-varying discounted IBNR first moments, $v(x) = 1 + bx$

More to the point, we now examine the "Stoodley formula" (see Kennedy (1992); Stoodley (1934)), $\delta(u) = p + \frac{c}{1+qe^{cu}}$, which clearly subsumes the constant discount-rate assumption. To such an end, we examine the same model as that of Figure 3.1, and we set $p = \delta = 0.05$. Elementary calculations show that now $v(x) = \frac{1}{1+q}e^{-(p+c)x} + \frac{q}{1+q}e^{-px}$. We choose two values for c: -0.04 and 0.08, and two values for q: 0.25 and 0.6.

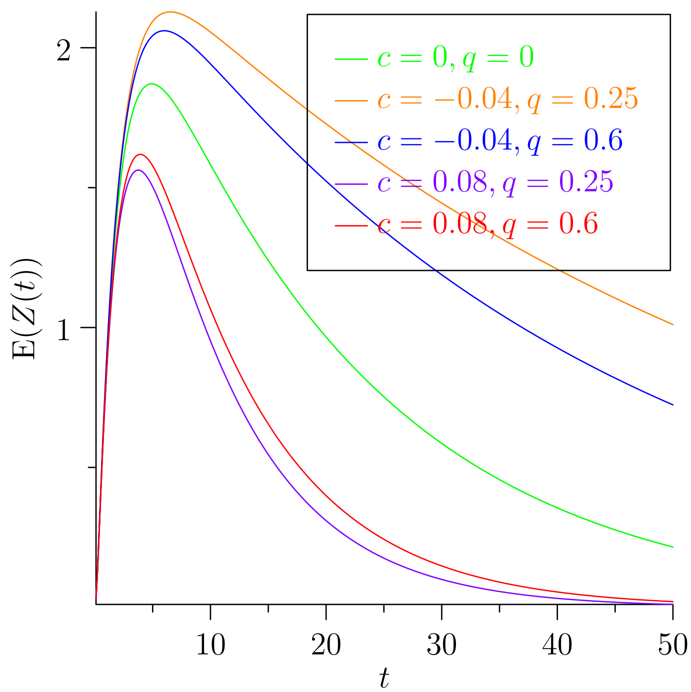

Figure 3.2. Time-varying discounted IBNR first moments, $v(x) = \frac{1}{1+q}e^{-(p+c)x} + \frac{q}{1+q}e^{-px}$

In Figure 3.2, evidently failing to include the term with c and q in $\delta(u)$ leads to an underestimation of $\mathrm{E}(Z(t))$ when $c = -0.04$, and an overestimation when $c = 0.08$. Further still, the value of q impacts the amount of over- or underestimation.

3.7. CONCLUDING DISCUSSION

We have extended the methodology of Landriault *et al.* (2017) both to allow insurers to reassess the distribution of the time until the next event given how many claims arise from the current event, and to allow the IBNR and IR claims to be discounted at a non-constant rate. Both the work of Landriault *et al.* (2017) and our generalization thereof implicitly assume finiteness of the integer moments of claims; for heavy-tailed claims distributions such as the Pareto or the Abate-Whitt (Abate and Whitt (1999)) with higher moments all infinite, modifications could be needed. Given the topic of our Section 2, a natural future direction for this section would be bringing the CPTA approach of Vatamidou *et al.* (2013) which we extended in Section 2 to bear upon our extension of Landriault *et al.* (2017) in this section. Another further step would be closer examination of the effects of particular choices of the function $D(x)$, or of particular distributional assumptions for the random intervals of batch sizes. The dependence structure we assumed between interevent times and the batch sizes could be relaxed further; again, the 2017 US hurricane season could provide motivation for such.

4. CONCLUSION

4.1. OUR CONTRIBUTIONS

This dissertation has examined two different senses of catastrophic risk modeling for insurance companies, under two variations of a Markovian dependence scheme. In the first, we examined Gerber-Shiu functions (on an infinite time horizon) with heavy-tailed individual claims. In the second, we studied reserving for IBNR claims when the time until the next event depends on the number of claims from the current event and money is discounted at a time-dependent rate. In both senses, we have added to the tools at an insurance company's disposal for estimating the funds needed in order to remain a healthy enterprise.

We have extended the non-asymptotic approximation approach of Vatamidou *et al.* (2013, 2014a) from ruin probabilities to the Gerber-Shiu function framework in the model of Li and Sendova (2015). Even after introducing the generality of time-discounting and nonnegative penalty functions $w(\cdot,\cdot)$, under mild regularity conditions we show the method of corrected phase-type approximations still captures the proper tail behavior of the functions being approximated. For the same case of the compound Poisson risk model, we numerically demonstrate CPTA performing better for small initial capital than asymptotic approximations. From our theoretical results, we illustrate the derivation of CPTA for the mean ruin time. Our work in Section 2 allows one to capture heavy-tailed behavior of Gerber-Shiu functions for both small and large initial capital without needing asymptotic expressions. Further, insurance companies may incorporate potentially catastrophic claims into a risk model charging different premiums depending on the insured.

We have also extended the results of Landriault *et al.* (2017) to allow the number of IBNR claims from an event to impact the time until the next event. Whereas in Albrecher and Boxma (2004); Li and Sendova (2015) insurers had only two interevent time distributions available, we allow any number of such distributions. As well, when calculating aspects of reserves, insurers are not constrained to assume an economy with a constant discount rate. We show numerically that indeed such an assumption may fail to predict the true valuation of IBNR claims, in a number of ways.

4.2. FUTURE WORK

We plan to use CPTA to relax the assumption implicit in Section 3 of all moments of the claim severity distribution existing. This will benefit reserving for IBNR claims by introducing catastrophic assumptions to the individual, micro-level claims. In the context of ruin theory, a natural next step would be extending our Propositions 1, 2, and 3 to the Li-Sendova risk model. We wish further to allow an arbitrary number of insured classes (with corresponding premium rates) in our Theorem 1. Doing so would involve more matrix theory, and for the asymptotic tail behavior, Markov renewal theory as expounded for example in Janssen and Manca (2006) seems like a tool we expect to use.

Again in the reserving context, we may allow τ_1 to have an arbitrary distribution $F_0(\cdot)$ rather than one of the $F_l(\cdot)$ (where $l \in \Im$), which would give rise to a "delayed Markov renewal process" (see again Janssen and Manca (2006)) in place of the Markov renewal process considered in Section 3 above. That way, the insurer does not have to categorize a "phantom" (current) event as falling within a certain catastrophic level, rather only doing so after an actual event occurs. As Léveillé and Garrido (2001b) showed the ordinary renewal case to be embedded in the delayed renewal case, we expect similar outcomes in our more general Markovian setup.

In the longer-term, we aspire to "combine" our Sections 2 and 3 to study Gerber-Shiu types of functionals of the insurer's surplus evolving over time, allowing for both senses of catastrophe in the modeling. Ahn *et al.* (2018) tackle a different, and in some ways simpler, case of this, with the usual limitations of claims arriving one at a time, according to a Poisson process. They further constrain claims to be phase-type distributed, with the associated drawbacks of light-tailed distributions. We would like to bring our catastrophic-relevant Markovian assumptions to bear in such a bridging of Gerber-Shiu theory and stochastic claims reserving, making our generalization of the corrected phase-type approximation methodology of Vatamidou *et al.* (2013, 2014a) relevant to more than risk theory alone, and that much closer to use by practicing actuaries.

REFERENCES

Abate, J. and Valkó, P. P., 'Multi-precision Laplace transform inversion,' International Journal for Numerical Methods in Engineering, 2004, **60**(5), pp. 979–993.

Abate, J. and Whitt, W., 'Explicit M/G/1 waiting-time distributions for a class of long-tail service-time distributions,' Operations Research Letters, 1999, **25**(1), pp. 25–31.

Abramowitz, M. and Stegun, I. A., *Handbook of Mathematical Functions*, volume 55 of *Applied Mathematics Series*, National Bureau of Standards, 10th printing, December 1972, with corrections edition, 1965.

Ahn, S. and Badescu, A. L., 'On the analysis of the Gerber-Shiu discounted penalty function for risk processes with Markovian arrivals,' Insurance: Mathematics and Economics, 2007, **41**(2), pp. 234–249.

Ahn, S., Badescu, A. L., Cheung, E. C. K., and Kim, J.-R., 'An IBNR-RBNS insurance risk model with marked Poisson arrivals,' Insurance: Mathematics and Economics, 2018, **79**(1), pp. 26–42.

Albrecher, H. and Boxma, O. J., 'A ruin model with dependence between claim sizes and claim intervals,' Insurance: Mathematics and Economics, 2004, **35**(2), pp. 245–254.

Albrecher, H. and Boxma, O. J., 'On the discounted penalty function in a Markov-dependent risk model,' Insurance: Mathematics and Economics, 2005, **37**(3), pp. 650–672.

Asmussen, S. and Albrecher, H., *Ruin Probabilities*, volume 14, World Scientific, 2010.

Bladt, M. and Nielsen, B. F., *Matrix-Exponential Distributions in Applied Probability*, volume 81 of *Probability Theory and Stochastic Modelling*, Springer Nature, 1st edition, 2017.

Bornhuetter, R. L. and Ferguson, R. E., 'The actuary and IBNR,' in 'Proceedings of the Casualty Actuarial Society,' volume 59, 1972 pp. 181–195.

Chau, K. W., Yam, S. C. P., and Yang, H., 'Fourier-cosine method for Gerber-Shiu functions,' Insurance: Mathematics and Economics, 2015, **61**, pp. 170–180.

Constantine, G. and Savits, T., 'A multivariate Faà di Bruno formula with applications,' Transactions of the American Mathematical Society, 1996, **348**(2), pp. 503–520.

Deng, Y., Liu, J., Huang, Y., Li, M., and Zhou, J., 'On a discrete interaction risk model with delayed claims and stochastic incomes under random discount rates,' Communications in Statistics–Theory and Methods, 2017, pp. 1–17.

Dickson, D. C. M., *Insurance Risk and Ruin*, Cambridge University Press, 2nd edition, 2017.

Dickson, D. C. M. and Hipp, C., 'On the time to ruin for Erlang (2) risk processes,' Insurance: Mathematics and Economics, 2001, **29**(3), pp. 333–344.

Foss, S., Korshunov, D., and Zachary, S., *An Introduction to Heavy-Tailed and Subexponential Distributions*, volume 6, Springer, 2nd edition, 2013.

Geiger, D. J. and Adekpedjou, A., 'Corrected phase-type approximations of Gerber-Shiu functions in a heavy-tailed risk model with both interclaim times and premiums depending on claim sizes,' Submitted, 2018.

Gerber, H. U. and Shiu, E. S. W., 'On the time value of ruin,' North American Actuarial Journal, 1998, **2**(1), pp. 48–72.

Gile, B. S., 'A statistical approach to IBNR reserves,' Journal of Actuarial Practice, 1994, **2**(2), pp. 257–272.

Janssen, J. and Manca, R., *Applied Semi-Markov Processes*, Springer Science & Business Media, 2006.

Johnson, W. P., 'The curious history of Faà di Bruno's formula,' The American Mathematical Monthly, 2002, **109**(3), pp. 217–234.

Kennedy, G., 'Modelling time-dependent interest rates by mathematical formula,' IMA Journal of Management Mathematics, 1992, **4**(4), pp. 317–330.

Klüppelberg, C., 'Subexponential distributions and characterizations of related classes,' Probability Theory and Related Fields, 1989, **82**(2), pp. 259–269.

Kolkovska, E. T. and Martín-González, E. M., 'Gerber-Shiu functionals for classical risk processes perturbed by an α-stable motion,' Insurance: Mathematics and Economics, 2016, **66**, pp. 22–28.

Landriault, D., Willmot, G. E., and Xu, D., 'Analysis of IBNR claims in renewal insurance models,' Scandinavian Actuarial Journal, 2017, **2017**(7), pp. 628–650.

Léveillé, G. and Adékambi, F., 'Covariance of discounted compound renewal sums with a stochastic interest rate,' Scandinavian Actuarial Journal, 2011, **2011**(2), pp. 138–153.

Léveillé, G. and Garrido, J., 'Moments of compound renewal sums with discounted claims,' Insurance: Mathematics and Economics, 2001a, **28**(2), pp. 217–231.

Léveillé, G. and Garrido, J., 'Recursive moments of compound renewal sums with discounted claims,' Scandinavian Actuarial Journal, 2001b, **2001**(2), pp. 98–110.

Li, S. and Garrido, J., 'On ruin for the Erlang (n) risk process,' Insurance: Mathematics and Economics, 2004, **34**(3), pp. 391–408.

Li, Z. and Sendova, K. P., 'On a ruin model with both interclaim times and premiums depending on claim sizes,' Scandinavian Actuarial Journal, 2015, **2015**(3), pp. 245–265.

Lin, X. S. and Willmot, G. E., 'Analysis of a defective renewal equation arising in ruin theory,' Insurance: Mathematics and Economics, 1999, **25**(1), pp. 63–84.

Peralta, O., Rojas-Nandayapa, L., Xie, W., and Yao, H., 'Approximation of ruin probabilities via erlangized scale mixtures,' Insurance: Mathematics and Economics, 2018, **78**, pp. 136–156.

Pitts, S. M. and Politis, K., 'Approximations for the Gerber-Shiu expected discounted penalty function in the compound Poisson risk model,' Advances in Applied Probability, 2007, **39**(2), pp. 385–406.

Rabehasaina, L. and Woo, J.-K., 'On a multivariate renewal-reward process involving time delays: Applications to IBNR process and infinite server queues,' arXiv preprint arXiv:1611.10202, 2016.

Rojas-Nandayapa, L. and Xie, W., 'Asymptotic tail behaviour of phase-type scale mixture distributions,' Annals of Actuarial Science, 2017, pp. 1–21.

Schmidt, K. D. and Zocher, M., 'The Bornhuetter-Ferguson principle,' Variance, 2008, **2**(1), pp. 85–110.

Šiaulys, J. and Asanavičiūtė, R., 'On the Gerber-Shiu discounted penalty function for subexponential claims,' Lithuanian Mathematical Journal, 2006, **46**(4), pp. 487–493.

Stoodley, C. L., 'The effect of a falling rate of interest on the values of certain actuarial functions,' Transactions of the Faculty of Actuaries, 1934, **14**(1), pp. 137–175.

Tang, Q. and Wei, L., 'Asymptotic aspects of the Gerber–Shiu function in the renewal risk model using Wiener–Hopf factorization and convolution equivalence,' Insurance: Mathematics and Economics, 2010, **46**(1), pp. 19–31.

Trefethen, L. N., Weideman, J. A. C., and Schmelzer, T., 'Talbot quadratures and rational approximations,' BIT Numerical Mathematics, 2006, **46**(3), pp. 653–670.

Vatamidou, E., Adan, I. J. B. F., Vlasiou, M., and Zwart, B., 'Corrected phase-type approximations of heavy-tailed risk models using perturbation analysis,' Insurance: Mathematics and Economics, 2013, **53**(2), pp. 366–378.

Vatamidou, E., Adan, I. J. B. F., Vlasiou, M., and Zwart, B., 'Corrected phase-type approximations of heavy-tailed queueing models in a Markovian environment,' arXiv e-prints arXiv:1405.4853, 2014a.

Vatamidou, E., Adan, I. J. B. F., Vlasiou, M., and Zwart, B., 'On the accuracy of phase-type approximations of heavy-tailed risk models,' Scandinavian Actuarial Journal, 2014b, **2014**(6), pp. 510–534.

Wang, H. and Ling, N., 'On the Gerber–Shiu function with random discount rate,' Communications in Statistics–Theory and Methods, 2017, **46**(1), pp. 210–220.

Widder, D. V., *The Laplace Transform*, Princeton University Press, Princeton, NJ, Dover (2010) edition, 1941.

VITA

Daniel Jefferson Geiger loved math and breaking out of the box from an early age, whether that involved developing fictional languages, inventing new time signatures in music, or finding the sum of integers squared while falling asleep one night in junior high school. After graduating from the Illinois Mathematics and Science Academy, he came to the Missouri University of Science and Technology, staying for the next decade. With an inquisitive mind prone to asking “What if?”, actuarial science was a perfect fit, while taking classes like actuarial modeling and life contingencies during his early undergraduate years at S&T. In May 2012, he received a B.S. in Applied Mathematics with Actuarial Science emphasis, a B.A. in Economics, and a minor in Spanish, graduating summa cum laude.

Perhaps Daniel asked himself sometime during those undergraduate years, “What if I created something new in actuarial science?” So, naturally he found himself in graduate school at S&T. During this time, he received the Chancellor’s Fellowship from 2012 to 2017. In May 2014, he received a non-thesis M.S. in Applied Mathematics with Statistic emphasis, thus graduating a second time from S&T. The following four years challenged him in ways like never before, yet he grew through it all, producing this dissertation in actuarial science under the guidance of Dr. Akim Adekpedjou.

While as a youth Daniel often carried clipboard, pencil and paper with in-process math problems, as a PhD student he still often had a pencil and paper at hand, but instead with notebooks and research articles or monographs. Always having been a man of many interests, after receiving his M.S., he served as a Graduate Research Assistant for a time in the S&T Department of Economics, in a faculty’s study sponsored by the United States Environmental Protection Agency. On May 8, 2018, Daniel defended his dissertation, paving the way to graduate a third and final time from S&T. In July 2018, he received his PhD in Mathematics with Statistics emphasis from Missouri S&T.