Claim Reserving via Inverse Probability Weighting
A Micro-Level Chain-Ladder Method

Sebastián Calcetero Vanegas*, Andrei L. Badescu, X. Sheldon Lin
Department of Statistical Sciences
University of Toronto
Toronto, Ontario
*sebastian.calcetero@mail.utoronto.ca

July 5, 2023

Abstract

Claim reserving is primarily accomplished using macro-level or aggregate models, with the Chain-Ladder method being the most popular one. However, these methods are heuristically constructed, rely on oversimplified data assumptions, neglect the heterogeneity of policyholders, and so lead to a lack of accuracy. In contrast, micro-level reserving leverages on stochastic modeling with granular information for improved predictions, but usually comes at the cost of more complex models that are unattractive to practitioners. In this paper, we introduce a simplistic macro-level type approach that can incorporate granular information at the individual level. We do so by considering a novel framework in which we view the claim reserving problem as a population sampling problem and propose an estimator using inverse probability weighting techniques, with weights driven by policyholder attributes. The framework provides a statistically sound method for aggregate claim reserving in a frequency and severity distribution-free fashion, while also incorporating the capability to utilize granular information via a regression-type framework. The resulting reserve estimator has the attractiveness of resembling the Chain-Ladder claim development principle, but applied at the individual claim level, and so it is easy to interpret by actuaries, and more appealing to practitioners as an extension of a macro-models.

Keywords— Claim reserving, Survey Sampling, Inverse Probability Weighting, Chain Ladder, Survival modeling

1 Introduction

Claim reserving is a crucial aspect of insurance and risk management, and is vital for ensuring solvency, assessing risk, and setting appropriate premiums. The insurance industry employs several types of reserves but is primarily interested in the reserve for outstanding claims, which cover the estimated costs of unsettled and non-reported claims, representing the insurer's liability for future payments related to already occurred accidents. This reserve can be split into subcomponents depending on the source of the claim i.e. whether it is from a reported claim or not, and it is of interest to create this distinction for accounting purposes. These reserves play a vital role in maintaining financial stability and ensuring the availability of funds for future claim payments.

Reserving in general insurance is one of the most studied problems in actuarial research. See for e.g. Schmidt (2011) for an extended list of most of the research covering this topic. Taha et al. (2021) and references therein provide a friendly overview of methods used in insurance reserving.

Briefly, two primary approaches to reserving, namely micro and macro approaches, have been widely studied in the actuarial literature. The theoretical foundations of these methods can be found in the literature of stochastic claim reserving, such as Wüthrich and Merz (2008).

On one hand, the macro approach to reserving focuses on estimating claim payments at an aggregate level. Among the macro approaches, Chain-Ladder-based techniques are widely employed in the insurance industry due to their ease of implementation, interpretation, and reliance on intuitive assumptions. See for e.g Mack (1994), Quarg and Mack (2004), Martínez-Miranda et al. (2012). These methods avoid the need for complex mathematical concepts, such as predictive models or stochastic processes, instead relying on simple operations that can be implemented using spreadsheets. Additionally, they only require estimating development factors, which can be easily obtained from aggregate data without the need for specialized software. Consequently, the Chain-Ladder method is favored by insurance companies and regulators, with more than 90% of insurers relying on it as their primary reserving method (Francis (2016)).

However, these aggregate methods overlook the actual composition and heterogeneity of the insurance portfolio. The Chain-Ladder assumes homogeneity among claims within a given group, disregarding valuable insights that can be gained by considering factors such as relevant attributes associated with the risk of each policyholder (Wüthrich (2018b)). In fact, the most recent literature on claim reserving (e.g Crevecoeur et al. (2022) and literature therein) highlights the importance of using all the information available (i.e the granular data) for the estimation of accurate reserves, and how ineffective is ignoring it. Consequently, the Chain-Ladder and similar macro-level models exhibit clear limitations and modest accuracy of estimation of the reserves when compared to models that do account for granular information i.e micro-level models.

On the other hand, the micro approach to reserving involves estimating individual claim payments by considering detailed characteristics such as policyholder information, claim type, severity, and other relevant factors (Boumezoued and Devineau (2017)). Micro-level reserving methods utilize probabilistic models that directly capture the behavior of policyholders and their impact on reserves, resulting in accurate forecasts. See for e.g Antonio and Plat (2014), Fung et al. (2021), Taylor et al. (2008), Wüthrich (2018a).

However, micro-level models pose challenges in terms of complexity, portfolio heterogeneity, and size, making them difficult to implement in practice. These models incorporate both stochastic and predictive modeling, adding layers of complexity that may hinder understanding by practitioners and regulators. Moreover, micro-level models often require assumptions about model components, such as distributions and simplifications of reality, which raise concerns about their validity. Consequently, these models are not widely adopted by actuarial practitioners due to the additional implementation effort required and the lack of consensus on modeling practices from a regulatory standpoint, even though these provide more reliable estimations. Indeed, according to Francis (2016) and related studies, micro-level reserving methods are virtually absent among insurance companies worldwide, with close to 0% utilizing them as either their primary method or for informational purposes.

A main obstacle to the consideration of micro-level reserving by practitioners and regulators is the significant disparity in methodologies with respect to macro-level models, in addition to the associated effort required for its construction. Macro-level models, such as the Chain-Ladder, differ significantly from micro-level models in terms of how the reserve estimation is derived. Consequently, the transition from a macro-level to a micro-level model represents a substantial and challenging undertaking for any insurance company. Furthermore, regulators face difficulties in validating and accepting a micro-level model when its underlying principles deviate significantly from the familiar idea of Chain-Ladder and the construction of the reserves via development factors. Therefore, the substantial gap between these two principal reserving methodologies hinders

the adoption of micro-level modeling by practitioners.

In this paper, we focus on reducing the gap between macro and micro models by providing a methodology that enables the use of individual information in a macro model, and so improves its performance while retaining most of its simplicity and interpretability. To do so, in this paper, we consider a novel approach to claim reserving by viewing the problem as a survey sampling problem. By treating the reported claims as a sampling from a larger population of claims, we develop a statistically justified macro approach based on an inverse probability weighting (IPW) method that accommodates the introduction of individual claim information via a regression-like model on the sampling probabilities, similar to how it is achieved to propensity scores. Just as is the case of traditional aggregate models, the IPW approach to claim reserving only requires the modeling of the development of the claims (i.e reporting and payment delays), without the need for explicit models for claim frequency or severity. As a result, the modeling efforts are focused on estimating policyholder-specific inclusion probabilities based on the observed distribution of the delays.

The resulting IPW estimator exhibits a functional form reminiscent of the Chain-Ladder method and its development factors. However, it distinguishes itself by having claim-specific factors that depend on the attributes of the claims. As a result, our methodology can be viewed as a "micro-level version" of the Chain-Ladder, where the development of each claim up to its ultimate value is performed at the individual level. Hence, our proposed approach represents a natural extension of traditional aggregate methods, tailored to incorporate individual claims information in a statistically justified manner, but in a much friendlier fashion. It is important to note that our approach is motivated independently of the Chain-Ladder method and differs from other attempts to account for heterogeneity in macro-level models, such as Wüthrich (2018b) or Carrato and Visintin (2019). In these approaches, a classification of claims into homogeneous classes is conducted, followed by the application of the Chain-Ladder method within each class. In contrast, our methodology seeks to integrate individual claims information without relying on such classification procedures or applying the run-off triangle development principle.

The IPW method represents an improvement over aggregate claim reserving models based on the Chain-Ladder, while also providing a cost-effective alternative to traditional micro-level reserving models. It maintains the desirable practicality and interpretability of macro-level models, making it a more appealing choice for both practitioners and regulators. This approach may serve as an initial step to encourage practitioners, who typically rely on macro models, to explore the potential benefits and insights obtained from incorporating individual information in the reserving process. Ultimately, it paves the way for practitioners and regulators to consider tailored-made models based on micro-level techniques.

This paper is structured as follows: Section 2 introduces the reserving problem as a sampling problem, and shows the derivation of IPW estimator for the outstanding claims, and its properties. Section 3 extends the methodology to consider other types of reserves such as the incurred but not reported (IBNR) and reported but not settled (RBNS) reserves. Section 4 discusses how to estimate the required inputs of the model. Section 5 provides a numerical study with real data. Lastly, Section 6 provides the conclusion and future research directions.

2 Claim reserving via inverse probability weighting

In this section, we present the claim reserving problem and demonstrate how it can be effectively tackled using inverse probability weighting methods. Since there are various types of reserves in general insurance, in this section we provide the overall idea of the methodology for the total reserve of outstanding claims. Section 3 will delve into the specific details of the methodology for the most

3

prevalent and significant reserves in general insurance, namely RBNS and IBNR reserves.

2.1 The claim reserving problem

Suppose an insurance company is analyzing its total liabilities associated with claims whose accident times occur between $t = 0$ and $t = \tau$, where τ is the valuation time of analysis as defined by the actuary. In general insurance, accidents are often not immediately reported to the insurance company for various reasons, resulting in a significant delay between the occurrence of a claimable accident and the time the insurance company is notified. As a result, at a given valuation time τ, the insurance company only has information on the claims reported by τ and is unaware of the unreported claims. Furthermore, the complexity of the problem increases due to another delay in the payment process. When a claim is reported, it is common for it to be paid in several sub-payments over time rather than as a lump sum. This is because the impact of an accident can evolve, requiring additional payments until it is fully settled. Therefore, at a given valuation time τ, the insurance company is only aware of the claims that were reported on time, and for each one, it may have paid only a partial amount of the associated claim size, rather than the entire amount owed.

As a result, the insurance company is interested in estimating the total claim amount of these unreported claims, as well as the remaining payments of the reported claims, to construct the overall reserve of outstanding claims. This reserve is also known in the insurance jargon as the Incurred But Not Settled (IBNS), and it's usually decomposed into further subcomponents depending on whether the payment is associated with a reported or not reported claim. For simplicity, here we consider the estimation of the overall reserve of outstanding claims without referring to the components.

That said, let's describe the payment process as follows:

- Let $N(\tau)$ represent the total number of different payments associated with all the claims whose accident time is before the valuation time τ.

- Let Y_i, $i = 1, \ldots, N(\tau)$ denote the sequence of amounts associated to each payment. Note that several of these payments may belong to the same claim/accident, but we will not make any distinction regarding this fact.

- Let T_i, $i = 1, \ldots, N(\tau)$ denote the sequence of accident times associated with the claim underlying each payment; let R_i, $i = 1, \ldots, N(\tau)$ denote the sequence of the associated reporting times; let S_i, $i = 1, \ldots, N(\tau)$ denote the sequence of the associated times in which the payments takes place. Clearly, $T_i < R_i < S_i$ and note that the values T_i, R_i would be the same for payments associated with the same claim, but the S_i would differ.

- Let $U_i = R_i - T_i$, $i = 1, \ldots, N(\tau)$ be the sequence of the reporting delay times associated with the claim underlying each payment, and $V_i = S_i - R_i$, $i = 1, \ldots, N(\tau)$ be the sequence of the associated payment delay time of each payment. Note that U_i is the same for all the payments associated with the same claim.

- Let X_i, $i = 1, \ldots, N(\tau)$ be the sequence of information/attributes of relevance, that is associated with the accident, the type of claim, the policyholder attributes, or the characteristics of the payment itself.

- Let $N^P(\tau)$ the number of payments made by valuation time τ out of the total $N(\tau)$ i.e the number of payments made to the claims reported by τ.

4

Along those lines, the total liability of the insurance company associated with accidents occurring before the valuation time τ, which we will denote as $L(\tau)$, is given by:

$$L(\tau) = \sum_{i=1}^{N(\tau)} Y_i$$

Similarly, the portion of liability that is known to the insurance company (i.e the so-called paid amount) by valuation time τ, which we will denote as $L^P(\tau)$, is:

$$L^P(\tau) = \sum_{j=1}^{N^P(\tau)} Y_j$$

We note that the indices of the payments made might not have the same order as all the payments, but we write it this way using the index j for the sake of simplicity of the notation.

Along these lines, an actuary is interested in estimating the remaining liability i.e the outstanding claims. We will denote this quantity as $L^O(\tau)$, and it is given by just the difference:

$$L^O(\tau) = L(\tau) - L^P(\tau)$$

This value is what the insurance company requires to set up the reserve for outstanding claims, either non-reported, non-settled, or both, and is our goal for estimation. For further details of the claim reserving problem, we refer the reader to Wüthrich and Merz (2008).

2.2   A survey sampling framework for claim reserving

Our proposal in this paper is based on a simple yet novel idea that allows us to frame the reserving problem in the context of *survey sampling*, enabling us to leverage techniques from this field to our advantage. Survey sampling is a statistical technique used to estimate population totals based on a smaller sample, especially in contexts where data collection from the entire population is impractical. The *sampling design* is the systematic process of selecting individuals or units from the population to be included in the sample. Different sampling methods are used depending on the research objectives and resources available. By using statistical techniques based on the sampling design, researchers can make reliable inferences about the population based on observations from the sample.

Applying this concept to our reserving problem, we can consider all $N(\tau)$ payments as the population under study, while the current $N^P(\tau)$ payments made by the valuation date serve as the selected sample for understanding this population. It is important to note that the sampling design and the actual sampling process are not determined or performed by the investigator but are purely driven by the randomness associated with whether a payment is made or not by the valuation date. Thus, the sample is given rather than being selected by the actuary. This is one of the distinctions between our setup and typical survey sampling situations.

The sampling mechanism based on the payment data can be conceptualized as a two-stage sampling process (Thompson (2012)). In the first stage, a *Poisson sampling* design without replacement is employed to sample the reported claims. This means, for each of the claims in the population, a Bernoulli experiment is conducted, where success is defined as the claim being reported by the valuation time, and failure occurs if it is not reported. Refer to Särndal et al. (2003) for more details on the Poisson sampling.

Moving to the second stage, we focus on the payments associated with each of the sampled claims from the previous stage (i.e., the reported claims). In this case, another sampling procedure

is carried out to determine which payments of a claim are made before the valuation time and which are not. This is also achieved by a Bernoulli like experiments, however do note that these are not independent because of the ordering of the payments e.g a second payment of a claim can be sampled as long as the first payment is sampled.

As a result of the sampling, we can assign a dicotomic random variable $\mathbf{1}_i(\tau)$, $i = 1, \ldots, N(\tau)$ with success probability $\pi_i(\tau)$, to each payment in the population. Such a variable takes the value of 1 or 0, indicating whether the payment Y_i belongs to the sample of payments made or not, respectively, by a given valuation time τ. These variables are referred as the *membership indicators* of the payments and are determined based on the delay in reporting (for the first stage of sampling) and the delay in payment (for the second stage) by the valuation time. Mathematically,

$$\mathbf{1}_i(\tau) = \mathbf{1}_{\{S_i \leq \tau\}} = \mathbf{1}_{\{T_i + U_i + V_i \leq \tau\}} = \mathbf{1}_{\{U_i \leq \tau - T_i\}} \mathbf{1}_{\{V_i \leq \tau - R_i\}}$$

where the indicators in the product on the right-hand side are the indicators of the first and second stages of sampling, respectively.

The probabilities $\pi_i(\tau)$ are known as *inclusion probabilities* and can be interpreted as the likelihood of payment Y_i belonging to the sample or, equivalently, being paid by the valuation time τ. Mathematically, these are given by

$$\pi_i(\tau) = P(U_i \leq \tau - T_i) \times P(V_i \leq \tau - R_i) = \pi_i^U(\tau) \times \pi_i^V(\tau) \tag{1}$$

where $\pi_i^U(\tau) = P(U_i \leq \tau - T_i)$ and $\pi_i^V(\tau) = P(V_i \leq \tau - R_i)$ are the inclusion probabilities of the first and second stage of sampling, respectively. These probabilities are dependent on the valuation time and are likely to vary across payments due to the different attributes X_i associated with each payment. While a more formal notation would be $\pi(\tau; Y_i, X_i)$ to highlight this dependency, we simplify it as $\pi_i(\tau)$ to streamline the notation and emphasize that the indexation on i corresponds to the probabilities being specific for each payment, and determined based on their attributes. In the literature on survey sampling, this is known as the sampling being *informative* as the actual values of the payments may be associated with the sampling design. It is important to note that these probabilities are not predefined and are therefore unknown to the investigator. We will delve into this matter further in Section 4.

Finally, note that the sample size in the design is not a fixed quantity. The sample size, which in our case is equivalent to the number of payments currently made $N^P(\tau)$, is a random variable defined as $N^P(\tau) = \sum_{i=1}^{N(\tau)} \mathbf{1}_i(\tau)$, which we can identify as the thinning of the counting process of the total number of payments.

2.3 Point estimation using the Horvitz-Thompson estimator

As motivated by the population sampling literature (Thompson (2012) or Särndal et al. (2003)), a well-established unbiased estimator of the population total of payments (i.e $L(\tau)$) under the aforementioned sampling design is provided by the *Horvitz-Thompson* (HT) estimator described as follows:

$$\hat{L}(\tau) = \sum_{j=1}^{N^P(\tau)} \frac{Y_j}{\pi_j(\tau)}$$

and therefore an unbiased estimator of the outstanding claims is the difference between the estimated total and the currently paid liability:

6

$$\hat{L}^O(\tau) = \hat{L}(\tau) - L^P(\tau) = \sum_{j=1}^{N^P(\tau)} \frac{Y_j}{\pi_j(\tau)} - \sum_{j=1}^{N^P(\tau)} Y_j = \sum_{j=1}^{N^P(\tau)} \frac{1 - \pi_j(\tau)}{\pi_j(\tau)} Y_j \tag{2}$$

The intuition behind the HT estimator lies in the fact that only a portion of all payments Y_j is reported, proportionally to $\pi_j(\tau)$, and so each payment is "augmented" by a factor of $1/\pi_j(\tau)$ to approximate the actual total amount. It is important to note that the HT estimator is non-parametric, meaning that it does not require any distributional assumptions about the underlying distribution of the number of claims (frequency) or the distribution of claim sizes (severity). Additionally, we emphasise on the fact that even though the estimator is based on the population level (i.e a macro level scale), the inclusion probabilities are dependent on the individual attributes of policyholders, claims, and payments. Therefore the estimator incorporates granular information as part of the estimation.

The HT estimator is widely recognized as one of the most influential estimators in the statistics literature, having been extensively studied for over 70 years in the field of population sampling (e.g Arnab (2017)). Consequently, the HT estimator has a solid theoretical foundation and possesses numerous desirable properties that directly inherit to the claim-reserving problem, including consistency, unbiasedness, and sufficiency, among others. More recently, it has also been applied in inverse probability weighting (IPW) methods for estimation under missing data and in causal inference (e.g Seaman and White (2013) and Yao et al. (2021)), including applications in fairness in insurance, and so the terminology "IPW estimator" is more extended in and outside the statistics literature. We will mostly refer to the estimator of the reserve as the IPW estimator, and reserve the naming of HT estimator when referring to the general concept.

A specific case of interest arises when we set $Y_i = 1$. In this scenario, all the sums above simplifies to a count of the number of payments, allowing us to obtain an unbiased estimator for the number of payments yet to make as:

$$\hat{N}^O(\tau) := \hat{N}(\tau) - N^P(\tau) = \sum_{j=1}^{N^P(\tau)} \frac{1}{\pi_j(\tau)} - \sum_{j=1}^{N^P(\tau)} 1 = \sum_{j=1}^{N^P(\tau)} \frac{1 - \pi_j(\tau)}{\pi_j(\tau)} \tag{3}$$

It is worth noting that this particular expression coincides with the one utilized by Fung et al. (2022) for the specific case of the number of incurred but not reported (IBNR) claims. In their work, they derived this expression under the assumption that the number of unreported claims follows a geometric distribution and demonstrated its unbiasedness when the number of claims is driven by a Poisson process. However, it is important to emphasize that within the framework of the HT estimator, this result is immediate and does not require of the assumption of the geometric distributions.

A "Micro-level" Chain-Ladder method

From an actuarial standpoint, the IPW estimator can be perceived as an individual-level adaptation of the Chain-Ladder method. In fact, by expressing the estimator as:

$$\hat{L}(\tau) = \sum_{j=1}^{N^P(\tau)} f_j(\tau) Y_j$$

we can interpret $f_i(\tau) := 1/\pi_i(\tau)$ as an individual development factor assigned to each payment Y_i. These factors serve to project the payment to its ultimate value which aligns with the

7

fundamental principle of the Chain-Ladder method. As the factors $f_i(\tau)$ are influenced by the policyholder's attributes, we can think of this methodology as a "micro-level" version of the Chain-Ladder method, as it applies the development on an individual level while retaining the essential characteristics of the Chain-Ladder. Indeed, we note that if no information of attributes is incorporated in the inclusion probabilities, then the development factors would be uniform across all claims. Consequently, the ultimate liability would be determined solely by multiplying the current paid amount by the development factor, which is nothing but the Chain-Ladder method. The IPW estimator and the Chain-Ladder method have a very entangled connection, but we will deepen this discussion in Calcetero-Vanegas et al. (2023).

2.4 Confidence interval of the estimation

Non-parametric confidence intervals for the reserve can be constructed based on the sampling distribution of the HT estimator, as discussed by Arnab (2017). In summary, under minimal regularity conditions, the HT estimator follows approximately a normal distribution under the two-stage sampling design for large populations (Chauvet and Vallée (2018)). Thus, an approximate $1 - \alpha$ confidence interval can be constructed using normal quantiles. However, as explained by Särndal et al. (2003), the accuracy of the normal approximation relies on the sample size and the distribution of Y_i, which tends to exhibit skewness and heavy-tails. Consequently, the normal distribution might provide a suboptimal approximation for our reserving application.

Alternatively, one can construct a confidence interval by applying a log transformation to the liability. This approach utilizes the delta method to construct an interval for the logarithm of the liability, which tends to exhibit behavior closer to normality. Subsequently, the interval is transformed back to the original scale using the reverse transformation. This log-transformed confidence interval can be a more appropriate choice, considering the distribution of the data and its potential skewness and heavy-tailed characteristics.

Therefore, an approximate $1 - \alpha$ confidence interval for $L^O(\tau)$ can be constructed as:

$$\left(\exp\left(\log\left(\hat{L}^O(\tau) \right) - Z_{\alpha/2} \frac{\sqrt{Var(\hat{L}^O(\tau))}}{\hat{L}^O(\tau)} \right), \exp\left(\log\left(\hat{L}^O(\tau) \right) + Z_{\alpha/2} \frac{\sqrt{Var(\hat{L}^O(\tau))}}{\hat{L}^O(\tau)} \right) \right) \quad (4)$$

where $Z_{\alpha/2}$ is the $\alpha/2$ quantile of the standard normal distribution, and the variance of the estimator is traditionally estimated using the expression:

$$\hat{Var}_1(\hat{L}^O(\tau)) = \sum_{j=1}^{N^P(\tau)} \frac{(1 - \pi_j(\tau))^3}{\pi_j^2(\tau)} Y_j^2 + \sum_{j=1}^{N^P(\tau)} \sum_{k=1, k \neq j}^{N^P(\tau)} \chi_{jk} \frac{\pi_{max(j,k)} - \pi_j \pi_k}{\pi_{max(j,k)}} \frac{1 - \pi_j(\tau)}{\pi_j(\tau)} \frac{1 - \pi_k(\tau)}{\pi_k(\tau)} Y_j Y_k$$

where $\chi_{jk} = 1$ if the j-th and k-th payment in the data are generated from the same claim, and $\chi_{jk} = 0$ otherwise.

As noted by Thompson (2012), the computation of the expression for the variance above can be quite laborious. This is because, in the second term, there is a combinatorial component associated with the covariance between payments belonging to the same claim. To address this challenge, alternative estimators of the variance can be employed, as suggested by Berger (1998). They propose the use of a simpler estimator, given by:

$$\hat{Var}_2(\hat{L}^O(\tau)) = \frac{\sum_{j=1}^{N^P(\tau)} \left(N^P(\tau) \frac{1 - \pi_j(\tau)}{\pi_j(\tau)} Y_j - \hat{L}^O(\tau) \right)^2}{N^P(\tau)(N^P(\tau) - 1)} \quad (5)$$

8

The expression above can be viewed as the jackknife estimator of the variance of the HT estimator (see for e.g Efron and Stein (1981)). This formulation is computationally simpler to obtain and, as commented by Thompson (2012), it tends to provide a more conservative estimate with a larger value compared to the actual variance. Therefore, we consider this estimator of the variance to be highly preferable over the previously derived one.

Lastly, we note that another approach for the construction of confidence intervals can be obtained using a bootstrapping of the HT estimator as described in Särndal et al. (2003). This approach, although computationally more expensive, is data-driven and could be a desirable alternative.

3 Calculation of RBNS, IBNR, incremental claims and other reserves

The reserve for outstanding claims, as discussed earlier, accounts for unreported and partially paid claims. While Equation (2) provides an estimator for the total reserve, it doesn't specify the allocation of the reserve to different types of payments. However, for accounting purposes, cash management, and risk assessment, actuaries need to specify the components of the overall reserve, commonly known as the IBNR (Incurred But Not Reported) reserve, the RBNS (Reported But Not Settled) reserve, and incremental payments over specific time periods.

In this section, we present how the survey sampling framework can be adapted to decompose the estimation of the total reserve (Equation (2)) into these sub-components as per the actuary's requirements. We introduce the "change of population principle" as a general approach to accomplish this decomposition within the IPW framework. We then demonstrate its application in deriving the RBNS, IBNR, incremental claims, and potentially other relevant calculations.

3.1 The change of population principle

In Section 2, we used the fact the currently paid amount can be considered as a sample of the total amount of payments. As a result, we defined a sampling design within the total amount of payments along with its corresponding inclusion probabilities. However, it is important to recognize a simple yet crucial fact: the currently paid amount can also be regarded as a sample from various sub-populations within the total amount of payments.

Figure 1 demonstrates a method of partitioning the total liability at a given valuation time τ into sub-populations associated with specific reserves of interest. This figure provides a visual representation, akin to a run-off triangle, distinguishing reported and non-reported payments at τ. The x-axis represents the development time, which goes from $t = T$ (i.e. the accident time) up to time $t = T + \omega$, being ω the maximum settlement time of a claim. This figure can be thought of as a screenshot of the classification of all the payments at a given valuation date τ.

To illustrate the concept, let's examine Figure 1. Combining all regions (A to G) yields the population discussed in Section 2, representing total liabilities. Focusing on the lower half of the figure, regions A to D, encompasses payments associated with reported claims at τ, where the current paid amount serves as a sample. Additionally, by narrowing our focus to the lower half of the figure and considering payments up to a specific time, such as $t = \tau_2$ (the union of regions A, B, and C), we obtain a truncated version of payments. Specifically, this includes total payments made for currently reported claims, excluding those made after $t = \tau_2$. Note that the current paid amount is a sample from all of this subpopulation.
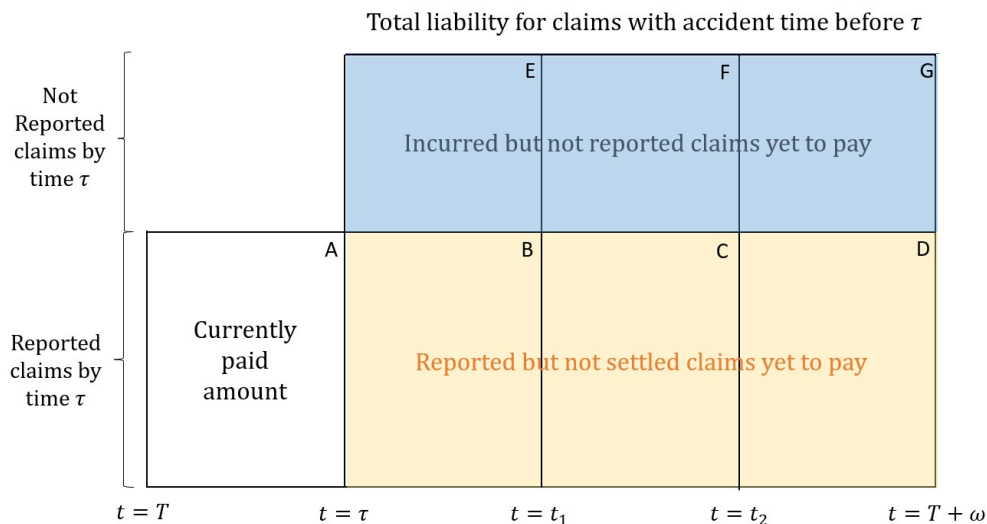
Figure 1: Decomposition of the total liability as of valuation time τ into sub-populations

Adopting this approach, estimating the total liability for a specific subpopulation involves treating it as the main population from which current payments are sampled. Consequently, the IPW estimator, discussed previously, can be employed to estimate the liability. We refer to this approach as the change of population principle.

This change in the subpopulation leads to a different sample design and, consequently, different inclusion probabilities. However, these probabilities can be easily determined using elementary conditional probability arguments. Specifically, when we limit the analysis to a subpopulation $\mathcal{S}$, we denote the inclusion probability under this restriction as $\pi_j^{\mathcal{S}}(\tau)$. This probability represents the likelihood of payment Y_j being reported at τ, given its membership in subpopulation $\mathcal{S}$. Bayes' rule allows expressing this probability as

$$\pi_j^{\mathcal{S}}(\tau) = \frac{\pi_j(\tau)}{P_j(\mathcal{S})}$$

Here, $P_j(\mathcal{S})$ denotes the probability of payment j being sampled in subpopulation $\mathcal{S}$ according to the original sampling design, influenced by factors like reporting delay and claim evolution. It is important to note that in this context, we assume the subpopulation $\mathcal{S}$ is a subpopulation encompassing the current payments (region A in Figure 1).

We will observe that for the reserves of interest, these probabilities can be straightforwardly expressed in terms of the probabilities associated with the previously defined delay time random variables U_j and V_j. Therefore, no additional estimations are necessary.

3.2 Calculation of the RBNS reserve

The Reported But Not Settled (RBNS) reserve represents payments that are yet to be made for claims already reported at valuation time τ. This reserve corresponds to the combined regions B, C, and D in Figure 1.

To estimate the reserve, we apply the change of population principle and define the population $\mathcal{S}$ as the total payments associated with reported claims at τ. This population corresponds to the lower half of Figure 1, specifically $\mathcal{S} = A \cup B \cup C \cup D$. In reserving terminology, this corresponds to the ultimate of incurred losses for claims reported prior to τ.

10

Next, we determine the inclusion probabilities. The selection of the subpopulation depends on claim reporting, which occurs with probability $P_j(\mathcal{S}) = P(U_j \leq \tau - T_j)$. Utilizing the previously mentioned result derived from Bayes' rule, the new inclusion probabilities for this population are:

$$\frac{\pi_j(\tau)}{P_j(\mathcal{S})} = \frac{P(U_j \leq \tau - T_j) \times P(V_j \leq \tau - R_j)}{P(U_j \leq \tau - T_j)} = P(V_j \leq \tau - R_j) = \pi_j^V(\tau)$$

This result is intuitive since the region only considers claims already reported at τ, and the remaining randomness pertains to the evolution of payment occurrences only. Consequently, we can utilize the IPW estimator to obtain an unbiased estimator for the total payments of reported claims as

$$\sum_{j=1}^{N^P(\tau)} \frac{Y_j}{\pi_j^V(\tau)}$$

Hence, the RBNS reserve of interest can be obtained by subtracting this quantity from the current paid amount:

$$\hat{L}^{RBNS}(\tau) = \sum_{j=1}^{N^P(\tau)} \frac{Y_j}{\pi_j^V(\tau)} - L^P(\tau) = \sum_{j=1}^{N^P(\tau)} \frac{1 - \pi_j^V(\tau)}{\pi_j^V(\tau)} Y_j \qquad (6)$$

3.3 Calculation of the pure IBNR reserve

To estimate the Incurred But Not Reported (IBNR) reserve, we cannot directly apply the change of population principle since the current paid amount is not a subpopulation of the not reported claims population (Figure 1). However, we can easily overcome this by considering the current paid amount as the difference between two populations: the total payments (all regions in Figure 1) and the total payments of currently reported claims (lower half of Figure 1). Estimations for the liabilities associated with these populations have been discussed in Sections 2 and 3.2, respectively. Therefore, the IBNR liability can be estimated as the difference between these two estimations.

$$\hat{L}^{IBNR}(\tau) = \hat{L}^O(\tau) - \hat{L}^{RBNS}(\tau) = \sum_{j=1}^{N^P(\tau)} \left(\frac{1}{\pi_j(\tau)} - \frac{1}{\pi_j^V(\tau)} \right) Y_j = \sum_{j=1}^{N^P(\tau)} \left(\frac{1 - \pi_j^U(\tau)}{\pi_j^U(\tau)} \right) \frac{Y_j}{\pi_j^V(\tau)} \quad (7)$$

We would like to highlight that this approach to estimating the IBNR is analogous to the conventional actuarial method using run-off triangles, where the total reserve is estimated using the incurred claims triangle and subtracting the reserve obtained from the paid claims triangle. Unlike aggregate approaches that may yield negative reserve estimates, our method ensures non-negative estimations.

3.4 Calculation of Cumulative and incremental payments

The estimator we have presented so far provides the ultimate amount of liabilities, but insurance companies require projections of the reserve payments over specific periods. These payments, known as incremental claims, can be estimated within our framework as follows: we utilize the change of population principle to estimate cumulative claims for different periods and then calculate the incremental claims as the difference between these cumulative claims. Although we will demonstrate this analysis for the total reserve, it can be similarly applied to the RBNS. Notably, our model is

11

continuous rather than discrete, allowing for the accommodation of any desired periodicity for incremental claims.

Let's consider an insurance company assessing claims incurred before the valuation time τ and interested in estimating the incremental claims associated with a future period between t_1 and t_2 ($\tau < t_1 < t_2$), denoted as $L(\tau, t_1, t_2)$. Visually, $L(\tau, t_1, t_2)$ corresponds to regions C and F in Figure 1.

We start by considering the population of cumulative claims up to time $t_1 \geq \tau$, where only payments made up to t_1 are included i.e $L(\tau, 0, t_1)$ (regions A, B, and E in Figure 1). Using the change of population principle, a payment belongs to this population if its payment time is before t_1, which occurs with probability $P_j(\mathcal{S}) = \pi_j(t_1)$. Thus, the inclusion probability is:

$$\frac{\pi_j(\tau)}{P_j(\mathcal{S})} = \frac{\pi_j(\tau)}{\pi_j(t_1)}$$

and so the IPW estimator is given by:

$$\hat{L}(\tau, T, t_1) = \sum_{j=1}^{N^P(\tau)} \frac{\pi_j(t_1)}{\pi_j(\tau)} Y_j \tag{8}$$

The incremental claims between t_1 and t_2 are then given by $L(\tau, t_1, t_2) = L(\tau, 0, t_2) - L(\tau, 0, t_1)$, and so an unbiased estimator for incremental claims is:

$$\hat{L}(\tau, t_1, t_2) = \hat{L}(\tau, 0, t_2) - \hat{L}(\tau, 0, t_1) = \sum_{j=1}^{N^P(\tau)} \frac{\pi_j(t_2) - \pi_j(t_1)}{\pi_j(\tau)} Y_j \tag{9}$$

This is a very intuitive expression: The denominator, $\pi_j(\tau)$, scales the observed claims Y_j to the total amount, while the difference in probabilities in the numerator, $\pi_j(t_2) - \pi_j(t_1)$, captures the proportion of the total observed between t_1 and t_2.

3.5 Others applications

The IPW framework and the change of population principle extends beyond the reserves discussed thus far, allowing for the estimation of other types of reserves based on the specific needs of the actuary. For instance, the incurred but not paid (IBNP) reserve can be estimated as a portion of the current RBNS estimation. In this case, the payments themselves serve as the population for analysis using the change of population principle, with an inclusion probability linked to the occurrence of the first payment in the sequence.

Another example, though less explored, is the unearned premium reserve (UPR). This reserve pertains to payments for claims where the accident occurs after the valuation time τ, but only for policies in force at τ. To estimate the UPR, the change of population principle can be applied by defining a larger superpopulation consisting of all payments associated with claims from policies in force at τ, regardless of when the accidents occurred.

Finally, it is important to note that the IPW framework provides estimations without assuming specific meaning for Y_i. The quantity of interest represented by Y_i can be diverse. For example, setting $Y_i = 1$ provides an estimation of the number of payments. Alternatively, the actuary can define Y_i as fees, commissions, policy management costs, etc., enabling a cost decomposition analysis of the reserve estimates.

12

4 Estimation of the model

In order to implement the IPW estimator, the key input required is the unknown inclusion probabilities $\pi_i(\tau)$. These probabilities are associated with the evolution of a claim, including reporting and settlement delays and depend on various attributes of the payment, the claim, and the policyholder, denoted as X_i, as well as the claim amount Y_i itself. In this section, we outline a data-driven approach to estimating these values. As explained in Section 2, the inclusion probabilities consist of two separate components: the probability of reporting and the probability of settlement. Each one is estimated separately, so we discuss the different strategies to achieve this in Sections and .

4.1 Estimation of the reporting delay times probabilities $P(U_j \leq \tau - T_j)$

To estimate probabilities, it is common to assume that the reporting delay times, conditioned on claim attributes X_i, follow a common distribution function (see, for example, Verrall and Wüthrich (2016)), that we will denote with $F_{U|X_i}(u)$ and that will be the target of estimation in some type of regression framework to include the dependence on covariates.

The variable of interest, U_i, is a time-to-event random variable commonly studied in survival modeling. Therefore, existing approaches in survival modeling can be utilized to estimate the overall distribution function and the desired probabilities. The most popular methods are based on the family of proportional hazard models, also known as Cox regression models (e.g George et al. (2014)). These methods aim to directly model the log-hazard function of the random time variable while accounting for the attributes in the modeling using a linear regression-like model as follows:

$$\log\big(\lambda_{U|X}(u)\big) = \log(\lambda_0(u)) + \langle X, \beta \rangle + \varepsilon \tag{10}$$

Here, $\lambda_0(u)$ represents the baseline hazard function, which can be chosen from a parametric family or modeled nonparametrically e.g. using B-spline representation. $\langle X, \beta \rangle$ represents the regression formula involving the covariates X with corresponding regression parameters β, and ε captures unobserved effects as an error term, also knows as a random effect or frailty. Depending on the analysis, various structures for the random effect can be considered, such autoregressive structure to capture trends and dependencies over time, or correlated effects to account for dependencies between claims that evolve together. Further guidance on specifying models for the hazard function can be found in Argyropoulos and Unruh (2015).

Consequently, the desired probability can be derived using the relationship:

$$\pi_i^U(\tau) = Pr(U_i \leq \tau - T_i) = F_{U|X_i}(\tau - T_i) = 1 - \exp\bigg(-\int_0^{\tau-T_i} \lambda_{U|X_i}(u)du\bigg) \tag{11}$$

A crucial aspect in the estimation of the model above is accounting for the right-truncation of the data. Indeed, due to the delay on the reporting times, our observations are limited to the conditional random variables: $U_i|U_i \leq \tau - T_i$, and ignoring this fact would result in a downward bias in the overall distribution. Fortunately, the literature on survival analysis has widely explored this issue and provided solutions that the user can adopt for the estimation of the model. See for e.g Shakur et al. (2021), Gui and Li (2005)Dempster et al. (1977), Verbelen et al. (2015), Fung et al. (2022)).

It is worth noting that not all survival models use linear regression structures or aim to describe the hazard function, and alternative approaches can offer different and flexible structures inspired by the machine learning literature (e.g Wang et al. (2019)). For instance, Wiegrebe et al. (2023) consider non-linear regression on covariates via deep learning approaches, Fung et al. (2022) propose

a flexible model based on a mixture of experts, while other approaches utilize survival trees such as Bou-Hamad et al. (2011). These alternatives provide increased flexibility compared to proportional hazard models but may require additional expertise for model fitting and interpretation. The choice of the model must be achieved in a data driven fashion aiming for the best fit to the data. Regardless of the chosen methodology, careful consideration should be given to estimation under the right truncation of the data.

Lastly, due to the popularity of survival analysis in statistics applications, most of the methods described above have already been implemented in statistical software packages and are readily available for its use in our applications. For example, in R, there are various implementations, including Cox models as in Bender et al. (2018), mixture of experts as in Tseung et al. (2021), and deep survival models, survival trees, forests, and more as in Sonabend et al. (2021).

4.2 Estimation of the payment times probabilities $P(V_j \leq \tau - R_j)$

Similarly to the previous case, we assume that the payment times, conditioned on claim attributes X_i (including the reporting time), follow a common distribution function denoted as $F_{V|X_i}(v)$ that we aim to estimate via a regression framework. While estimating this probability might seem similar to the previous case, a significant difference arises due to the recurrent nature of payment events for a given claim, as opposed to the one-time event of claim reporting. This *recurrent event process* (e.g Cook et al. (2007)) necessitates an appropriately adapted modeling approach. In this section, we will discuss two closely related yet distinct approaches to address this estimation.

Counting processes

Recurrent events are closely related to counting processes, where the former focuses on event times and the latter on the number of events. In our case, we can consider the number of payments over time to be governed by a stochastically defined point process. Numerous works in insurance have explored modeling such processes in the context of reserving (e.g Antonio and Plat (2014), Maciak et al. (2021), Yanez and Pigeon (2021)).

Let's define $M(t)$, where $t \in (0, \omega)$, as the counting process associated with the number of payments for a single claim. This is the counting process associated with the payments times V_j for a given claim. For the sake of readability, we will omit the dependence on covariates in the notation, although it is important to acknowledge that all these quantities are dependent on them.

Since our objective is to determine probabilities of the form $P(V \leq t)$, our goal is to express the desired probability in terms of the process $M(t)$. The following proposition establishes this connection:

Proposition 1. *Under the above definitions, we have that:*

$$P(V \leq t) = E\left(\frac{M(t)}{M(\omega)}\right) \tag{12}$$

Proof. Reverse the time of the process $\tilde{V} = \omega - V$. The reversed time process can be seen as a mortality process (see Section 4.2 for more details), where the initial number of lives is $M(\omega)$ and the lifetime random variable of a newborn follows the same distribution as $\tilde{V}$. Then

$$P(V \leq t) = P(\tilde{V} \geq \omega - t) = E\left(P(\tilde{V} \geq \omega - t | M(\omega))\right) = E\left(\frac{E(M(t)|M(\omega))}{M(\omega)}\right) = E\left(\frac{M(t)}{M(\omega)}\right)$$

where the second last equality holds by the traditional life table relationship $_tp_0 = l_t/l_0$, and the last equality is the tower property of conditional expectations.

<center>14</center>

Equation 12 reveals that the desired probability possesses an intuitive expression associated to the evolution of a claim up to settlement. In essence, the right-hand side of Equation (12) represents the probability is the expected proportion of payments made by time t out of total of payments. Another equivalent interpretations of this quantity is as the inverse of a development factor for the number of payments from time t to the ultimate value at time ω. It is worth noting that this expression can be analytically computed only for certain processes. One such example is the widely used Poisson process (and some extensions), as illustrated in Example 1.

Example 1. *Suppose the $M(t)$ is a non-homogeneous Poisson process (NHPP) with intensity rate $\mu(t)$, then*

$$P(V \leq t) = \frac{\int_0^t \mu(s)ds}{\int_0^\omega \mu(s)ds}$$

Furthermore, the result holds even if we consider an NHPP with frailty e^ε i.e when conditional on the random variable e^ε, the process is NHPP with intensity rate $e^\varepsilon \mu(t)$.

Proof. This is a widely known result but can be proven right away using the theorem above. To do so, use conditional expectation on $M(\omega)$. It is wide known for NHPP that $M(t)|M(\omega) \sim$ Binom $\left(n = M(\omega), p = \frac{\int_0^t \mu(s)ds}{\int_0^\omega \mu(s)ds}\right)$, and so we simply have:

$$P(V \leq t) = E\left(\frac{M(t)}{M(\omega)}\right) = E\left(\frac{E(M(t)|M(\omega))}{M(\omega)}\right) = E\left(\frac{M(\omega)\frac{\int_0^t \mu(s)ds}{\int_0^\omega \mu(s)ds}}{M(\omega)}\right) = \frac{\int_0^t \mu(s)ds}{\int_0^\omega \mu(s)ds}$$

For the frailty, note that conditional on e^ε, the probability above does not depend on e^ε as it cancels out in both the numerator and denominator, so the same expression holds right away.

□

Along those lines, the actuary must select in a data-driven fashion an appropriate counting process (that incorporates the use of attributes) to model the number of payments per claim, and then proceed to compute the desired probability using Equation (12). Fitting counting processes could be complex task and can vary depending on the approach used. For a comprehensive discussion on this matter, we refer to Andersen et al. (1985).

Reversed time counting process

By reversing the time of the counting process for the number of payments using the transformation $\tilde{V} = \omega - V$, we can interpret the resulting process as a mortality process, as discussed by Hiabu (2017). This analogy allows us to describe $\tilde{V}$ using a mortality model, which simplifies the fitting of the counting process. Reversing the time is a well-studied approach in the survival modeling literature Klein and Moeschberger (2003).

Most mortality models belong to the class of survival models, and so can be embedded into the framework described in Section 4.1. The advantage of working with the reversed time process and mortality models instead of the counting process directly is the wider range of options available in terms of statistical modeling, including several readily implementable semi-parametric models (Mulayath Variyath and Sankaran (2014)).

In this case, we assume that the reversed hazard function (e.g Block et al. (1998)) of the time random variable $\tilde{V} = \omega - V$, denoted as $\tilde{\lambda}_{V|X}(t)$, is described using a Cox regression-like model that incorporates attribute information:

$$\log\left(\tilde{\lambda}_{V|X}(t)\right) = \log\left(\tilde{\lambda}_0(t)\right) + \langle X, \alpha \rangle + \varepsilon$$

As before, $\tilde{\lambda}_0(t)$ represents a baseline reversed hazard function, $\langle X, \alpha \rangle$ represents a regression formula involving the covariates X with parameters α, and ε represents a random effect. This modeling approach is analogous to the one described in Section 4.1, so we refer the reader to it.

As a result, the desired probability can be derived as:

$$\pi_i^V(\tau) = P(V_i \leq \tau - R_i) = P(\tilde{V}_i \geq \omega - (\tau - R_i)) = \exp\left(-\int_0^{\omega-(\tau-R_i)} \tilde{\lambda}_{V|X_i}(t)dt\right) \qquad (13)$$

Similar to the case of the reporting delay time, we face a right truncation problem when considering payments occurring after the valuation date. However, when reversing the time, this issue transforms into a left truncation problem. Observations of the variables are only available if $\tilde{V}_i|\tilde{V}_i \geq \omega - (\tau - R_i)$. Therefore, it is important to estimate the survival model for the reversed time random variable using a right truncation algorithm. Fortunately, modern implementations of survival models often include this capability as discussed in Section 4.1.

4.3 Goodness of fit and other considerations

Since the inclusion probabilities are the sole inputs for the IPW estimator, it is essential to have a well-fitted model for optimal performance. In this section, we discuss the assessment of the model's goodness of fit using pseudo residuals. Additionally, we comment on the possible instability of the resulting estimator and discuss ways of addressing such an issue.

Pseudo-residuals

To assess the accuracy of the resulting predictive distribution from the models in Section 4.1, one approach is to use uniform pseudo-residuals based on the probability integral transform (Rüschendorf (2009)). These pseudo-residuals are constructed by evaluating the fitted distribution function at the observed values. They are widely used for goodness of fit assessment in various model families (e.g Buckby et al. (2020)). Considering that the observations come from a truncated distribution, the truncated version of the distribution should be taken into account. These pseudo residuals can be expressed as:

$$r_i^U = \frac{\hat{F}_{U|X}(U_i)}{\hat{F}_{U|X}(\tau - T_i)} \qquad r_i^V = \frac{\hat{F}_{V|X}(V_i)}{\hat{F}_{V|X}(\tau - R_i)}$$

The uniform pseudo-residuals should exhibit approximate uniformity if the fitted model adequately represents the data. The uniform pseudo-residuals can be transformed to the normal scale using the quantile function of the standard normal distribution, denoted as Φ^{-1}:

$$\tilde{r}_i^U = \Phi^{-1}(r_i^U) \qquad \tilde{r}_i^V = \Phi^{-1}(r_i^V)$$

These transformed normal pseudo-residuals allow for easier visualization and detection of deviations from the expected distribution when compared with uniform scale, nevertheless they are equivalent. Note that the *Cox-Snell* residuals, commonly used in survival analysis, are obtained by

employing the quantile function of an exponential distribution instead of the normal distribution. See for e.g Klein and Moeschberger (2003).

The normal pseudo-residuals can be utilized to assess the goodness of fit of the model through graphical analysis, such as scatter plots, etc, in the same fashion as with ordinary residuals in linear regression. The focus of the assessment is to determine whether the distribution of these residuals resembles a normal distribution, which can be achieved through QQ and PP plots, or hypothesis testing techniques.

Adjustments to the IPW estimate

The inclusion probabilities can vary significantly impacting the stability of the estimator. A extreme case is when the estimated inclusion probability of claim is close to 0, which mostly occurs when a claim is recently reported. Such circumstances can lead to instability in the estimator, potentially resulting in abnormally high values of the reserve when compared with experience on previous reserving exercises. This behavior has been widely documented in survey sampling literature of the HT estimator. See for e.g., Hulliger (1995), Chen et al. (2017) and references therein.

Trimming the inclusion probabilities is a method proposed to address the behavior of extreme values. In this approach, if the inclusion probability is too small, it can be replaced with a larger value to get rid of the instability (Chen et al. (2017)). For reserving applications, the probabilities can be trimmed by artificially assuming a slightly later valuation date, which would increase the inclusion probabilities. Alternatively, data-driven methods, such as the algorithm proposed by Zong et al. (2018), offer a more systematic approach. They propose algorithm 1 as illustrated below, and demonstrate that the mean square error of the IPW estimator behaves more favorably than when not performing any correction.

Algorithm 1 Algorithm by Zong et al. (2018) to trimm inclusion probabilities for the IPW estimator

---

- Obtain the ordered inclusion probabilities $\left\{\pi_{(1)}, \pi_{(2)}, \ldots, \pi_{(N^P(\tau))}\right\}$ from smallest to largest.

for $j = 1, \ldots, N^P(\tau)$ **do**

    if $\pi_{(j)} \leq \frac{1}{j+1}$ and $\pi_{(j+1)} > \frac{1}{j+2}$ **then**

        Modify inclusion probabilities as:

$$\boldsymbol{\pi}^* = \{\underbrace{\pi_{(j)}, \ldots, \pi_{(j)}}_{j-1}, \pi_{(j)}, \pi_{(j+1)}, \ldots, \pi_{(N^P(\tau))}\}$$

    end

end

It is important to note that replacing the inclusion probabilities with larger values may introduce a downward bias in the reserve estimation. Therefore, it is recommended to perform the adjustment only if the estimation displays sensitivity to the changes in the inclusion probabilities. If the estimation remains nearly unchanged after the adjustment, retaining the original estimation would be preferred over the trimmed one.

5 Numerical study with real data

In this section, we showcase the application of the IPW estimator using a real data set obtained from a European automobile insurance company. The data-set comprises information on Body Injury (BI) claims from January 2009 to December 2012. The insurance policy had considerable structural

17

changes in terms of coverage, premiums, and administrative handling in the years thereafter, so the information is not considered. The date in consideration represents the point at which we possess information regarding all settled claims within this time frame.

In line with our methodology's primary objective of serving as an alternative to traditional macro-level models while being simpler than fitting a micro-level model, we maintain a simplified approach to emphasize the practicality of the method in real-world applications.

5.1 Data description

The dataset contains detailed information about claim settlements, policyholder attributes, and automobile characteristics within the aforementioned time period. This information encompasses various factors such as car weight, engine displacement, engine power, fuel type (gasoline or diesel), car age, policyholder age, and region (based on a general classification). Furthermore, the dataset includes details related to the accidents themselves, such as the time of occurrence and the type of accident (albeit based on a general classification). Additionally, information regarding the progression of claim payments is available, including reporting time, the corresponding settlement amounts, and the corresponding occurrence times.

Our statistical modeling study focuses on the evolution of claims, specifically related to reporting delay time and payment times. We illustrate some characteristics of these quantities, such as the distributions in Figure 2 and summarize key statistics in Table 1.
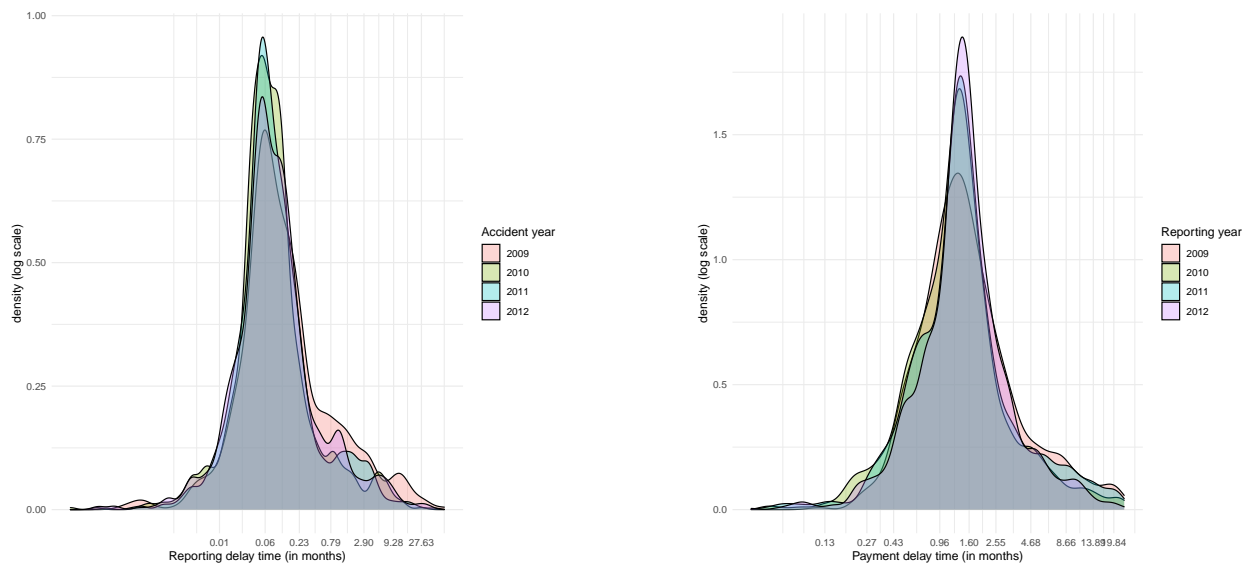
Figure 2: Density functions by year of the reporting delay time (left panel) and the payment time (right panel) in months. Plots are in the log scale.

Upon reviewing the information presented in Figure 2 and Table 1, it is evident that the reporting delay tends to be relatively short, with an average duration of less than a month. However, there is notable variability in the tail behavior of this variable. In contrast, the progression of claim payments typically spans a few months on average, but there are instances where settlement times can extend over several years. It is important to note that the distributions of these variables exhibit complexities that are challenging to capture using parametric models. Specifically, they display significant temporal fluctuations, indicating that historical data may not adequately represent

Variable	Year	Mean	Std. Dev.	Min.	1st Qu.	Median	3rd Qu.	Max.
	2009	1.03	4.09	0.00	0.05	0.10	0.28	43.28
Reporting	2010	0.45	1.85	0.00	0.04	0.07	0.13	25.66
delay time	2011	0.50	2.16	0.00	0.04	0.07	0.17	24.09
	2012	0.57	2.34	0.00	0.04	0.08	0.18	32.69
	2009	3.32	7.04	0.10	0.93	1.43	2.52	48.05
Payment	2010	2.31	4.37	0.04	0.81	1.31	1.92	35.37
delay time	2011	2.82	5.41	0.03	0.91	1.41	2.14	35.38
	2012	2.49	4.95	0.04	0.98	1.44	2.14	40.23

Table 1: Summary statistics per year of the reporting delay time and the payment delay time (both in months)

future events. To account for this, we define the maximum development time for future analysis as $\omega = 24$ months, with approximately 99.95% of claims being settled within this timeframe.

5.2   Model fitting

To model the reporting delay time, we employ a Cox regression-type model, as outlined in Section 4.1. Likewise, we adopt the reversed-time approach discussed in Section 4.2 to develop a model for claim evolution. Due to the limited length of our dataset and the observed variations across different years, we fit the model using only the last two years (i.e., ω) preceding a given valuation date. Specifically, we utilize the time window of 2009 and 2010 for training purposes, while evaluating the models using the years 2011 and 2012.

For the sake of illustration, we calculate reserves on a monthly basis. To ensure that the estimation data captures the recent evolution of claims as much as possible, we employ a rolling window approach, where only the last two years are used to fit the model for each month. This approach aligns with the practices employed by industry professionals in their daily work. We do not consider a time-series model via correlated frailties due to the short time window of the training set i.e only 2 years. Although there may be some variations in model parameters across different periods, the overall fit behaves similarly across time. We proceed to present the fitting processes for the first two years of data in the training set.

Reporting delay time

For the reporting delay time, U, we fit a Cox regression model as in Equation (10) using the attributes of the policyholder as covariates:

$$\log\big(\lambda_{U|X}(u)\big) = \log(\lambda_0(u)) + \beta_1\text{Car-Weight} + \beta_2\text{Engine-Power} + \beta_3\text{Fuel-Type} + \beta_4\text{Age} + \beta_5\text{Car-Age}$$
$$+ \beta_6\text{Accident-type} + \beta_7(\text{Claim-Amount}) + S_8(\text{Accident-day}) + \beta_9\text{Region}$$

The baseline hazard function, denoted as $\lambda_0(u)$, is estimated using a B-Spline representation. Additionally, we incorporate a non-linear effect of the covariate "Accident-day" with the term $S_8(\text{Accident-day})$, which is also estimated using a B-Spline representation. This inclusion of a non-linear effect associated with calendar time provides the model with a dynamic alike structure, allowing it to account for some temporal changes.

To estimate the parameters while considering the right truncation of the data, we utilize a generalized additive model implementation via the piece-wise exponential modeling approach, as described by Bender et al. (2018). This approach is readily available in various R packages such

as `flexsurvreg`, `pammtools` or `GJRM`. The results of the estimation are presented in Table 2 and Figure 3.

Coefficient	β_1	β_2	β_3	β_4	β_5	β_6	β_7
Value	$-3.6*10^{-4}$	$7.37*10^{3}$	0.231	$-4.72*10^{-3}$	$2.05*10^{-2}$	0.398	$-1.93*10^{-5}$
Std. Err.	$3.88*10^{-6}$	$4.27*10^{-5}$	$1.39*10^{-3}$	$4.53*10^{-5}$	$2.56*10^{-4}$	$1.905*10^{-3}$	$5.98*10^{-8}$
p-val	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001

Table 2: Estimated coefficients of the Cox regression model for the reporting delay time

Table 2 shows that all the policyholder attributes included in the model are statistically relevant to describe the behaviour of the reporting delay time. The same conclusion applies to the categorical variable "Region", although the detailed results are not presented in the table due to its numerous categories. The right panel of Figure 3 illustrates the non-linear effect associated with the accident day, revealing a quarterly seasonal pattern. Specifically, the hazard rate, indicating the reporting speed, is higher in the second and fourth quarters compared to the other two quarters. Furthermore, the left panel of Figure 3 displays the baseline hazard function, indicating a large hazard rate during the initial months, suggesting a concentration of reporting delays within this period. The hazard rate then decreases rapidly but remains nonzero for large delay times, indicating a heavy tail.
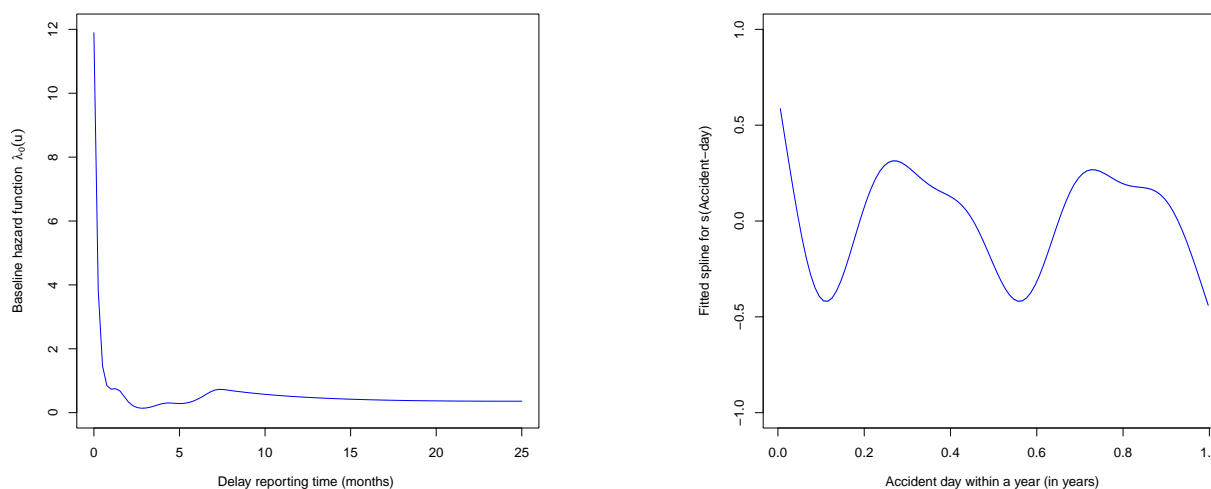
Figure 3: Fitted baseline hazard function for the reporting delay time (left panel) and fitted effect of the accident day on the hazard of the reporting delay time (right panel)

To assess the adequacy of the model fit, we employ normal pseudo residuals, as introduced in Section 4.3. Figure 4 presents QQ and PP plots, comparing these normal pseudo residuals against the theoretical normal distribution. From both plots, it is evident that the normal pseudo residuals exhibit no significant deviations from their expected theoretical counterparts. Hence, there is no evidence suggesting a lack of fit in the fitted distribution function.

Payment delay time

Next, we present the model for claim evolution utilizing the reversed time-counting process estimation approach, in conjunction with a Cox regression model similar to the one previously described.

Figure 4: Q-Q plot (left panel) and P-P plot (right panel) of the normal pseudo-residuals for the Cox model for the reporting delay time.

In this case, we consider a maximum settlement time of $\omega = 24$ months, as mentioned earlier, and proceed to model the reversed time random variable $\tilde{V} = \omega - V$. The Cox regression model for the reversed hazard function is defined as follows:

$$
\begin{aligned}
\log\left(\tilde{\lambda}_{V|X}(v)\right) = \log\left(\tilde{\lambda}_0(v)\right) &+ \alpha_1\text{Car-Weight} + \alpha_2\text{Engine-Power} + \alpha_3\text{Fuel-Type} + \alpha_4\text{Age} + \alpha_5\text{Car-Age} \\
&+ \alpha_6\text{Accident-type} + \alpha_7(\text{Payment-Amount}) + S_8(\text{Reporting-day}) + \alpha_9\text{Region} \\
&+ \alpha_{10}\text{Reporting-delay-time}
\end{aligned}
$$

the baseline reversed hazard function, denoted as $\tilde{\lambda}_0(v)$, is estimated using a B-Spline representation. The non-linear effects of the covariate "Reporting-day" are captured through the term $S_8(\text{Reporting-day})$, also estimated using a B-Spline representation.

This modeling approach mirrors the methodology presented in the previous section, and thus, we refrain from delving into further details. However, we highlight two key distinctions in this regression model. Firstly, we incorporate the reporting day as a non-linear effect. Secondly, we include the observed reporting delay time as a covariate. Notably, these probabilities are based on all the information available at the payment time, which encompasses any additional information obtained during the reporting process.

The fitted model is presented in Table 3 and Figure 5. The results exhibit similarities to the previous case, as shown in Table 3, where all policyholder attributes included in the model are statistically significant in describing the behavior of the reversed payment time. The right panel of Figure 5 displays the non-linear effect associated with the reporting day, indicating a quarterly seasonal pattern. However, the pattern is not as distinct as observed for the reporting delay time. The right panel of Figure 5 illustrates the baseline reversed hazard function, which should be interpreted in reverse. In this plot, time 24 months corresponds to time 0 in the original scale, and time 0 in the plot represents 24 months in the original scale. It is evident that the hazard function is initially high during the first couple of months (in the original scale), implying a significant number of payments occurring within this period. Subsequently, the hazard rate decreases rapidly, approaching zero, indicating the occurrence of some payments.

21

Coefficient	α_1	α_2	α_3	α_4	α_5	α_6	α_7	$\alpha_1 0$
Value	$-6.609*10^{-5}$	$6.18*10^{-3}$	0.188	$-1.036*10^{-3}$	$-4.56*10^{-2}$	-0.293	$-7.86*10^{-6}$	$-2.57*10^{-2}$
Std. Err.	$3.69*10^{-6}$	$4.47*10^{-5}$	$1.43*10^{-3}$	$4.71*10^{-5}$	$2.81*10^{-4}$	$2.01*10^{-3}$	$5.74*10^{-8}$	$2.74*10^{-4}$
p-val	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Table 3: Estimated coefficients of the Cox regression model for the reversed payment time
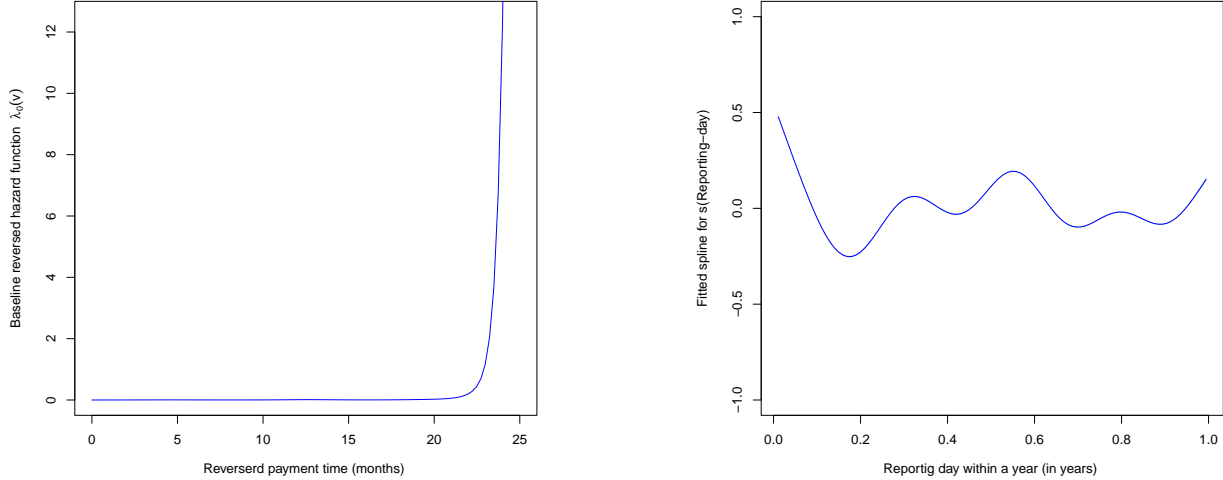
Figure 5: Fitted baseline hazard function for the reporting delay time (left panel) and fitted effect of the accident day on the hazard of the reporting delay time (right panel)

To assess the goodness of fit, we employ again the normal pseudo residuals and validate them accordingly. Figure 6 displays QQ and PP plots, comparing these normal pseudo residuals against the theoretical normal distribution. Similar to the previous analysis, there is no significant deviation observed between the normal pseudo residuals and their expected theoretical counterparts in both plots. Hence, there is no evidence indicating a lack of fit in the fitted distribution function.

5.3 Estimation of the reserve for a single date

Here we show the estimation of the outstanding claims (IBNS), RBNS and IBNR reserves for the first date of the testing period. To ease the visualization and comparison, we present the estimation of the reserves in the classical run-off triangle format in Tables 5 and 6, yet however, recall that our method doesn't rely on a given periodicity for its calculation or the construction of a triangle.

The inclusion probabilities $\pi_i(\tau)$ and $\pi_i^V(\tau)$ for the outstanding not settled claims are estimated directly from the models from the previous section using Equations (11), (13) and (1). Figure 7 displays the histogram of such probabilities and Table 4 displays some summary statistics. Briefly, we observe that the inclusion probabilities vary drastically from one claim to another due to the heterogeneity of the claims. Note that the probabilities tend to be closer to 1 than to 0 due to the low average reporting delay and payment time, and therefore only the most recently reported claims have a small probability.

Tables 5 and 6 present cumulative run-off triangles for total outstanding claims and reported but not settled claims, respectively, as of the valuation date. The incurred but not reported claims reserve estimation is derived from the difference between these triangles, which is not shown to
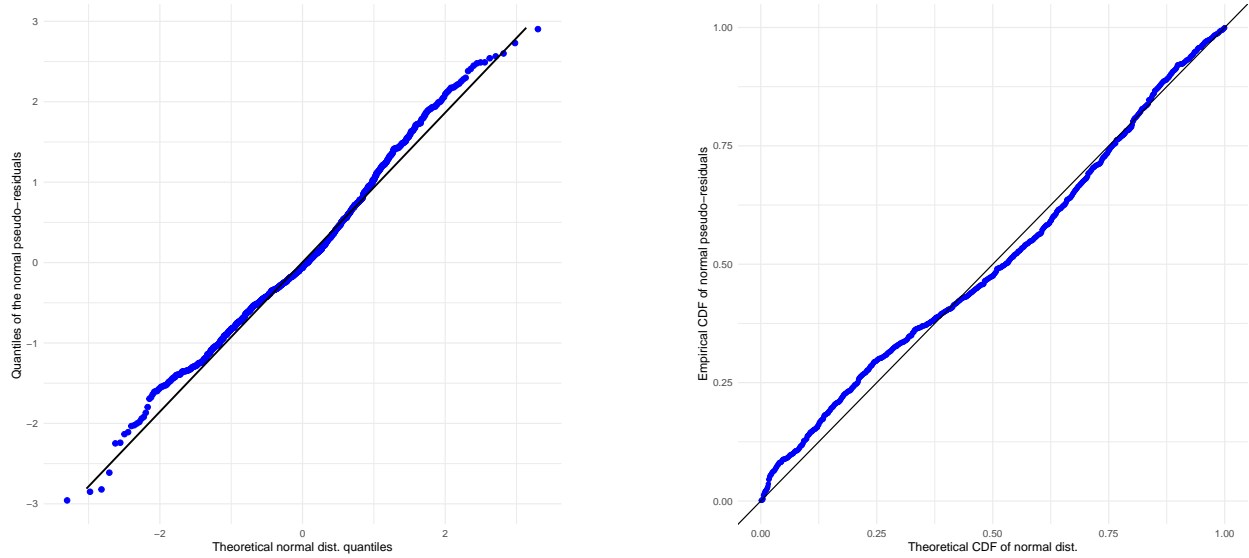
22
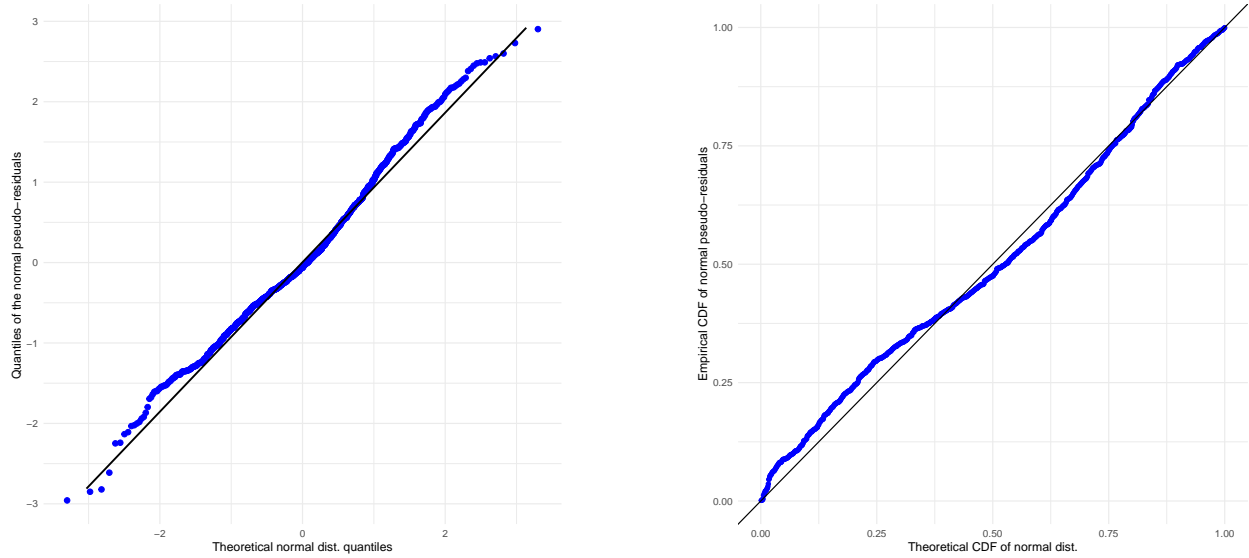
Figure 6: Q-Q plot (left panel) and P-P plot (right panel) of the normal pseudo-residuals for the Cox model for the reversed payment time.
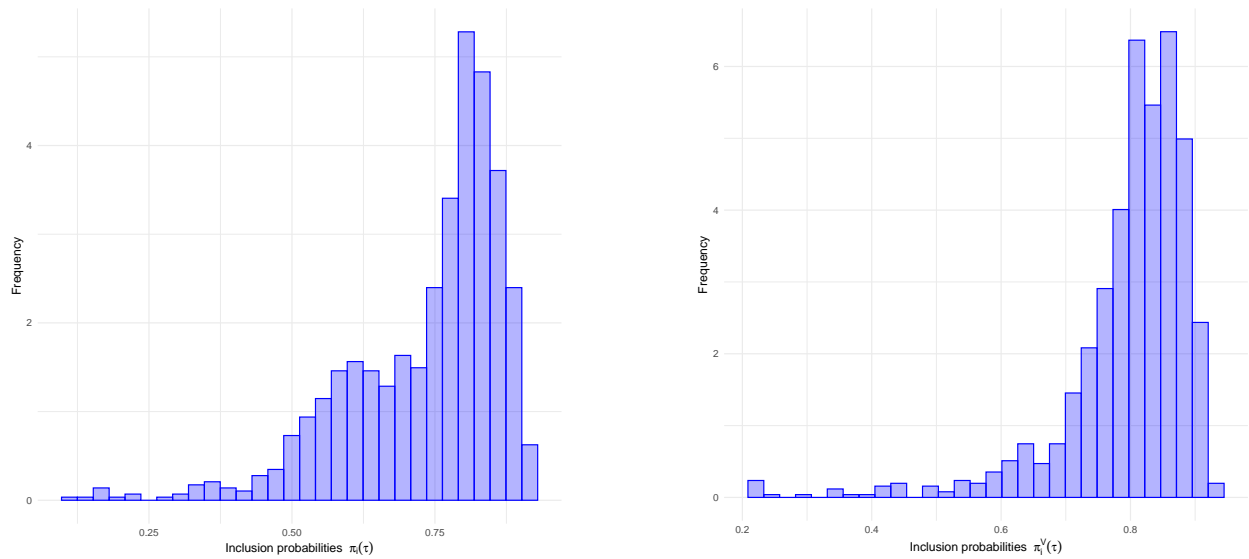
Figure 7: Histogram of the inclusion probabilities used for the IPW estimator of the total reserve (left panel) and the RBNS reserve (right panel)

Probability	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
$\pi(\tau)$	0.116	0.653	0.781	0.736	0.833	0.921
$\pi^V(\tau)$	0.216	0.766	0.820	0.797	0.860	0.927

Table 4: Summary statistics of the inclusion probabilities

avoid redundancy. We completed the lower half of the triangles using observed true reserve values, the IPW estimator (using Equation (8)), and the Chain-Ladder method for comparison purposes, employing a monthly periodicity. To maintain readability and practicality, the table is limited to

23

13 months, representing approximately 98% of settled claims within this period. To differentiate between the RBNS and IBNR claims components in the Chain-Ladder method, we utilize the double Chain-Ladder method Martínez-Miranda et al. (2012).

Analyzing the lower half of the triangles in Tables 5 and 6, we observe that the IPW estimator provides cumulative payment estimations that exhibit similar trends and magnitudes as the actual cumulative claims. No evident patterns of under or over estimation are observed. Additionally, the IPW estimator does not consistently exhibit similar behavior to the Chain-Ladder method, in terms of both over and under estimations or the occurrence of abnormal values.

M	1	2	3	4	5	6	7	8	9	10	11	12	13	ULT
1	14,134	14,134	20,211	20,211	26,314	27,371	27,371	27,371	28,059	28,059	28,059	28,059	28,059	28,059 32,170 29,864
2	30,727	177,519	177,519	220,701	239,987	260,631	262,010	262,010	262,010	264,832	266,273	266,273	266,273 268,495 266,369	268,401 323,763 285,454
3	87,856	158,587	170,758	170,758	231,602	231,602	235,675	235,675	247,039	247,039	247,039	247,039 249,269 247,265	247,908 272,571 252,047	297,541 309,415 270,071
4	164,017	205,623	248,124	264,613	264,613	282,013	286,108	286,108	286,108	289,410	289,410 291,061 289,716	290,675 308,302 294,460	290,675 320,097 299,458	297,467 360,788 320,600
5	119,065	225,436	279,176	296,804	300,808	306,261	314,102	314,102	320,340	320,340 322,443 320,489	322,591 355,560 325,380	322,591 373,943 330,774	322,591 388,044 335,167	332,056 435,457 358,063
6	12,518	99,082	146,828	156,780	167,589	169,668	170,493	181,229	181,229 182,423 181,283	212,492 194,660 182,229	212,492 202,168 184,534	212,639 207,354 187,098	212,639 211,442 190,504	257,430 222,585 204,108
7	37,174	132,479	209,631	218,619	232,507	241,423	245,323	245,323 247,042 245,494	273,696 264,234 254,273	273,696 273,743 255,601	278,910 284,226 258,672	294,019 291,805 262,410	294,909 296,200 267,209	346,304 314,629 286,328
8	40,234	191,089	227,720	241,797	285,608	285,608	285,608 290,212 286,667	290,118 327,506 293,846	295,826 346,422 302,802	296,713 369,539 304,561	312,941 389,189 308,474	312,941 398,289 312,988	315,764 408,089 318,639	321,604 442,873 340,750
9	14,251	123,802	176,626	217,705	229,304	229,304 232,096 230,700	233,399 257,759 234,301	234,422 272,238 239,968	234,592 285,793 247,470	235,953 324,709 248,875	235,953 344,994 252,014	258,863 374,998 255,766	259,236 386,755 260,134	782,474 434,518 278,538
10	22,408	188,162	274,916	363,112	372,005 368,019 367,423	408,322 412,041 385,109	411,377 437,451 390,692	413,752 447,627 399,462	413,806 496,137 413,096	415,288 549,596 415,207	417,411 629,822 421,599	417,411 658,452 428,037	417,411 672,707 434,747	483,132 788,018 465,082
11	22,017	256,833	312,025	312,148 319,151 315,773	324,473 371,166 337,649	345,710 397,213 354,696	409,621 408,778 359,395	424,520 422,950 367,446	426,909 461,395 380,638	426,909 492,821 382,364	427,083 496,838 387,143	435,422 503,133 392,530	436,328 513,386 400,024	454,346 546,347 428,665
12	11,371	105,591	157,762 114,398 109,022	258,713 172,447 125,655	281,683 195,043 134,820	286,280 204,176 140,940	291,383 217,368 143,118	291,383 245,751 146,415	310,758 282,399 151,346	329,514 287,515 152,163	329,940 290,146 154,078	346,858 295,399 156,346	348,634 300,999 159,074	359,294 321,814 170,367
13	24,662	35,614 30,208 24,662	169,336 78,756 64,017	194,656 96,315 69,236	207,267 102,162 72,943	230,055 119,705 74,363	245,208 138,875 75,853	249,372 168,485 78,469	249,372 172,777 79,224	249,372 174,132 80,102	249,372 176,301 81,156	250,671 178,786 82,654	250,671 181,360 83,441	250,922 190,887 88,919

Table 5: Monthly cumulative run-off triangle for all outstanding claims in the valuation date. In black is the actual value, in blue is the estimation using the IPW estimator, and in green is the traditional Chain-Ladder

The overall findings indicate that, for the given valuation date, the IPW estimator generally offers a superior approximation of the reserve compared to the traditional Chain-Ladder method. However, it is important to note that the IPW estimator does not consistently outperform the Chain-Ladder method in all cells of the triangle; the latter remains a more accurate method for certain dates. Although the IPW estimator can provide detailed estimations at the cell level, it may not possess the same level of precision as the estimation of the reserve as a whole. This limitation arises from the reliance on population sampling, which necessitates a large and representative sample for accurate estimation. Consequently, the more granular the estimation (i.e the smaller the subpopulation of interest), the lower the level of accuracy.

24

M	1	2	3	4	5	6	7	8	9	10	11	12	13	ULT
1	14,134	14,134	20,211	20,211	26,314	27,371	27,371	27,371	28,059	28,059	28,059	28,059	28,059	28,059 32,048 29,744
2	30,727	177,519	177,519	220,701	239,987	260,631	262,010	262,010	262,010	264,832	266,273	266,273	266,273 268,495 266,352	268,401 316,876 283,936
3	87,856	158,587	170,758	170,758	231,602	231,602	235,675	235,675	247,039	247,039	247,039	247,039 249,269 247,233	247,908 272,571 251,331	295,920 303,666 267,896
4	164,017	205,623	248,124	264,613	264,613	282,013	286,108	286,108	286,108	289,410	289,410 291,061 289,716	290,675 308,302 294,290	290,675 320,097 298,562	296,975 335,713 318,018
5	119,065	225,436	279,176	296,804	300,808	306,261	314,102	314,102	320,340	320,340 322,443 320,456	322,591 355,527 324,742	322,591 373,910 329,669	322,591 383,230 333,381	332,056 402,049 354,482
6	12,518	99,082	146,828	156,780	167,589	169,668	170,493	181,229	181,229 182,423 181,265	212,492 194,408 183,790	212,492 201,907 186,288	212,639 205,999 189,194	212,639 209,165	257,430 219,408 201,648
7	37,174	132,479	209,631	218,619	232,507	241,423	245,323	245,323 247,042 245,468	273,696 263,487 252,902	273,696 272,968 253,831	278,910 277,245 256,254	282,049 280,291 259,887	282,049 282,707 263,961	333,444 297,815 281,367
8	40,234	191,089	227,720	241,797	285,608	285,608	285,608 290,212 286,252	290,118 327,506 291,115	295,826 346,422 298,601	296,713 354,353 299,853	310,655 358,311 303,047	310,655 360,353 307,302	313,478 362,343 312,034	319,318 391,250 332,059
9	14,251	123,802	176,626	217,705	229,304	229,304 232,096 230,261	233,399 257,759 232,668	234,422 272,238 236,447	234,592 277,845 242,690	234,592 279,895 243,679	234,592 280,390 246,201	257,501 281,496 249,730	257,874 284,773 253,379	259,624 309,678 269,920
10	22,408	188,162	274,916	363,112	372,005 368,019 366,874	407,354 412,041 378,973	410,409 437,451 382,687	412,784 447,627 388,215	412,838 451,060 399,427	414,320 451,101 400,865	416,443 451,161 406,052	416,443 454,506 411,896	416,443 464,323 417,404	438,292 499,305 444,322
11	22,017	256,833	312,025	312,148 319,151 315,615	324,473 371,166 336,465	340,741 397,213 348,095	404,394 408,778 351,263	419,293 413,575 356,156	421,682 413,949 366,992	421,682 413,949 368,139	421,856 413,969 371,767	424,104 419,168 376,883	424,104 427,627 383,052	439,958 452,425 408,325
12	11,371	105,591	157,762 114,398 108,806	258,713 172,447 124,708	280,666 195,043 133,214	284,373 204,176 137,359	288,514 208,635 144,880	288,514 209,811 145,436	303,401 209,811 146,924	314,932 209,812 149,019	315,358 211,616 151,245	322,242 215,341 161,149	324,018 219,281	333,471 231,872
13	24,662	35,614 30,208 24,662	169,336 77,486 61,364	184,855 94,635 66,154	193,175 100,349 68,776	193,175 103,022 69,700	205,684 103,997 70,561	209,848 104,078 72,564	209,848 104,140 73,113	209,848 104,880 73,740	209,848 106,127 74,677	209,848 107,566 75,884	209,848 109,055 76,486	210,100 113,674 81,161

Table 6: Monthly cumulative run-off triangle for only reported claims in the valuation date. In black is the actual value, in blue is the estimation using the IPW estimator, and in green is the double Chain-Ladder estimation

Reserve Type	Method	Ultimate	Reserve	Error	% Error
IBNS	True	4,479,030	1,605,716	-	-
	IPW	4,723,264	1,849,950	-244,234	-15.2%
	CL	3,526,809	653,495	952,221	59.3%
RBNS	True value	3,813,048	939,733	-	-
	IPW	3,905,779	1,032,465	-92,732	-9.9%
	CL	3,434,026	560,712	379,022	40.3%
IBNR	True value	665,983	665,983	-	-
	IPW	817,485	817,485	-151,502	-22.7%
	CL	92,783	92,783	573,199	86.1%

Table 7: Error metrics for the estimates of the reserves in the first date of the testing period.

In line with this approach, instead of focusing on cell-level comparisons, our emphasis lies on the aggregation of cells to determine the actual reserve value, which is the ultimate objective of estimation. It is noteworthy that the IPW estimator directly provides an estimation of the total reserves using Equations (2), (6) and (7), eliminating the need for constructing the run-off triangle. Table 7 presents the aggregated reserve values obtained by summing the lower half of

25

the triangles, along with the corresponding estimation errors. Our findings reveal that the IPW estimator yields reserve values that closely align with their true counterparts for all reserve types, exhibiting significantly lower estimation errors compared to the Chain-Ladder method.

Furthermore, to evaluate the predictive quality of these estimates from a probabilistic standpoint, Figure 8 illustrates the predictive distribution of the reserves based on the sampling distribution of the IPW estimators, juxtaposed with the actual observed values. Notably, we observe that the true values consistently fall within the central region of the distribution, closely aligning with the corresponding modes, which represent the predicted reserve values. Consequently, the IPW-based predictions exhibit consistency with the observed reality, further validating their reliability.
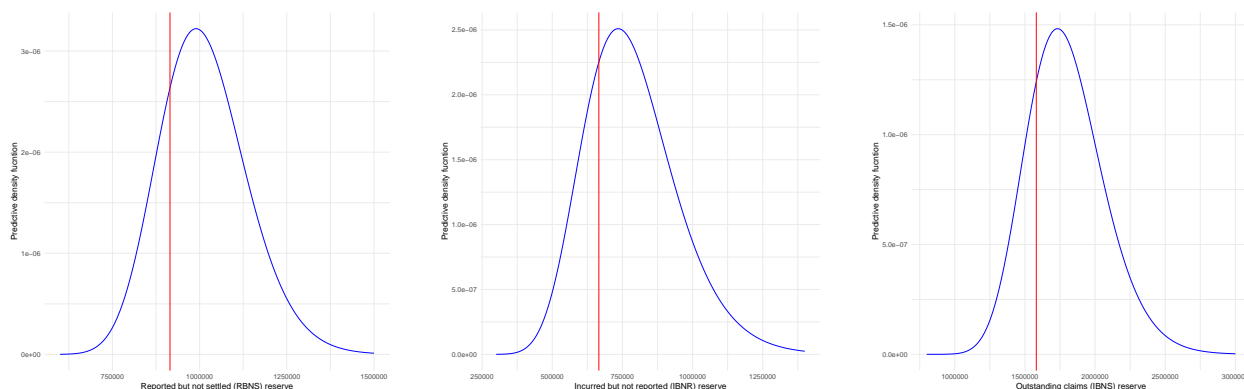
Figure 8: Predictive distributions of the IPW method for the RBNS, IBNR and IBNS reserves. In red is the true observed value.

5.4 Estimation of the reserves for several dates

Here we present the estimation of reserves for all 24 months in the testing period. Figure 9 illustrates the estimations for the outstanding claims compared to the true value of the reserve at the corresponding month. Additionally, Figures 10 and 11 depict the estimation for RBNS and IBNR, respectively. To provide a comprehensive analysis, we include 95% confidence intervals for the estimations and include Chain-Ladder (CL) method estimates for comparison. Furthermore, Table 8 presents error metrics to assess the disparities between the estimations across all dates.

Reserve Type	Method	ME	RMSE	MAE	MAPE
IBNS	IPW	68,779	215,368	228,874	14%
	CL	540,239	662,345	574,588	40%
RBNS	IPW	183,197	189,087	228,939	17%
	CL	140,470	279,493	286,723	33%
IBNR	IPW	-114,418	198,679	209,339	37%
	CL	399,769	328,503	399,769	76%

Table 8: Error metrics in the testing period. ME: Mean error, RMSE: Root mean square error, MAE: Mean absolute error, MAPE: Mean absolute percentage error

With respect to the total reserve, Figure 9 demonstrates that the IPW estimator produces predictions that closely align with the true value of the reserve for the majority of the observed periods, exhibiting no discernible pattern of under or overestimations. Additionally, the actual
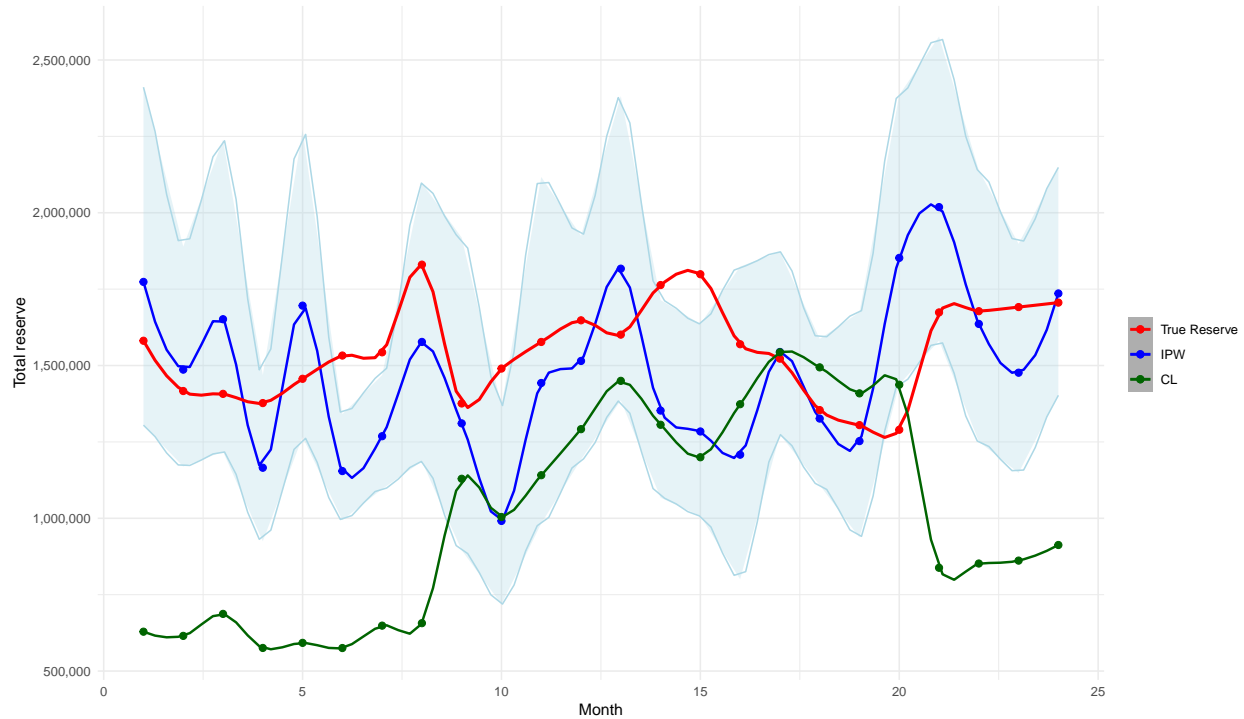
26

Figure 9: Estimation for the reserve of outstanding claims per month

reserve value consistently falls within the associated confidence intervals, indicating a good fit with the predicted value. Notably, the IPW prediction proves to be more accurate than the traditional Chain-Ladder method during the considered period. This observation is further supported by the results in Table 8, where the error metrics for the IPW over the 24-month period outperform those of the Chain-Ladder method.

With respect to the RBNS, we observe in Figure 10 that the IPW estimator provides highly accurate predictions for the majority of the first year within the time window and for a portion of the second half of the second year. However, during the intermediate period (8th month to 17th month), the IPW underestimates the reserve, although some data points in this range still fall within the confidence band. It is worth noting that this period exhibits relatively higher reserve levels compared to the rest of the considered time window, which may be attributed to management-related actions of the insurance company that lead to larger reserves. In such cases, the IPW estimation takes longer to capture these changes, as the distribution estimation relies on the preceding two years of data. Consequently, it takes several months for the most recent data to have a significant impact on the estimation. On the other hand, the Chain-Ladder method appears to be more adept at capturing this particular change. However, outside of this specific period, the Chain-Ladder method demonstrates considerable underperformance. Finally, as indicated in Table 8, the IPW consistently outperforms the Chain-Ladder method on average throughout the entire period, as evidenced by the lower error metrics.

Finally, regarding the IBNR reserve, we observe in Figure 11 that the IPW estimator provides a reasonable estimation for almost the entire considered period, fluctuating around the true reserve. It is worth noting that the IPW estimator exhibits a more variable behavior compared to previous scenarios. This variability is expected because the IBNR, in our case, represents a smaller proportion of the subpopulation due to the relatively low reporting delay time. Generally, as the size of
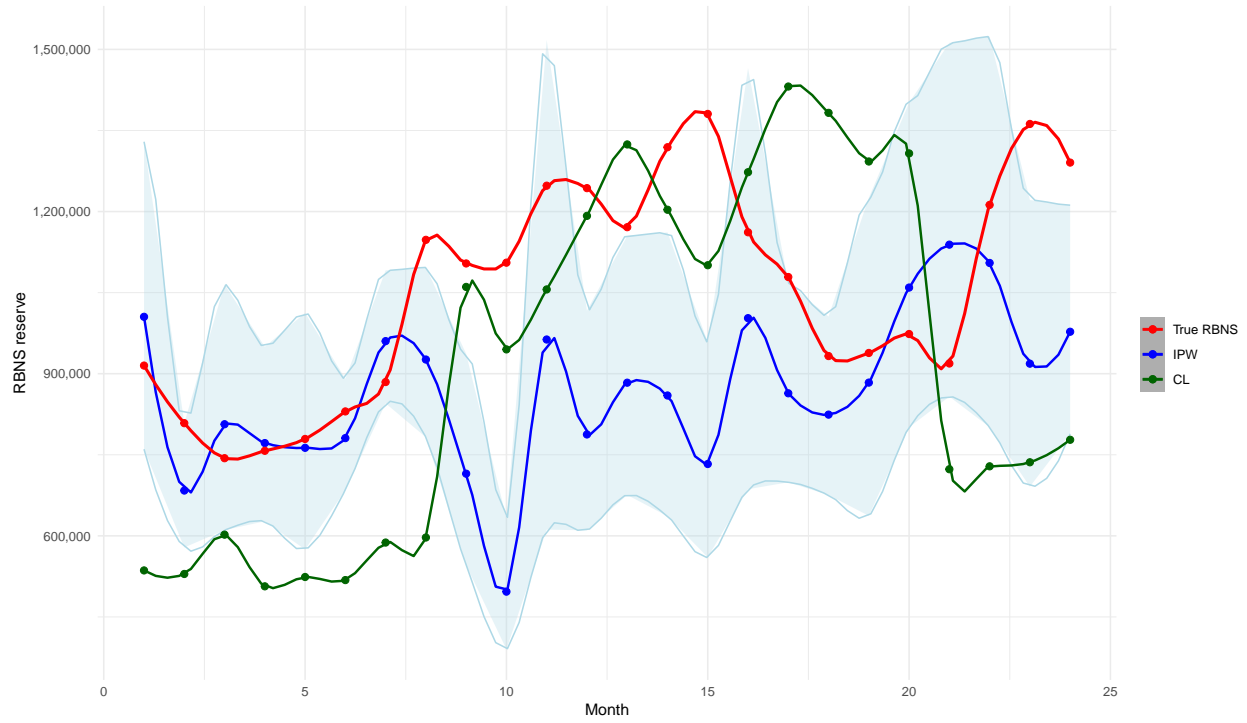
27

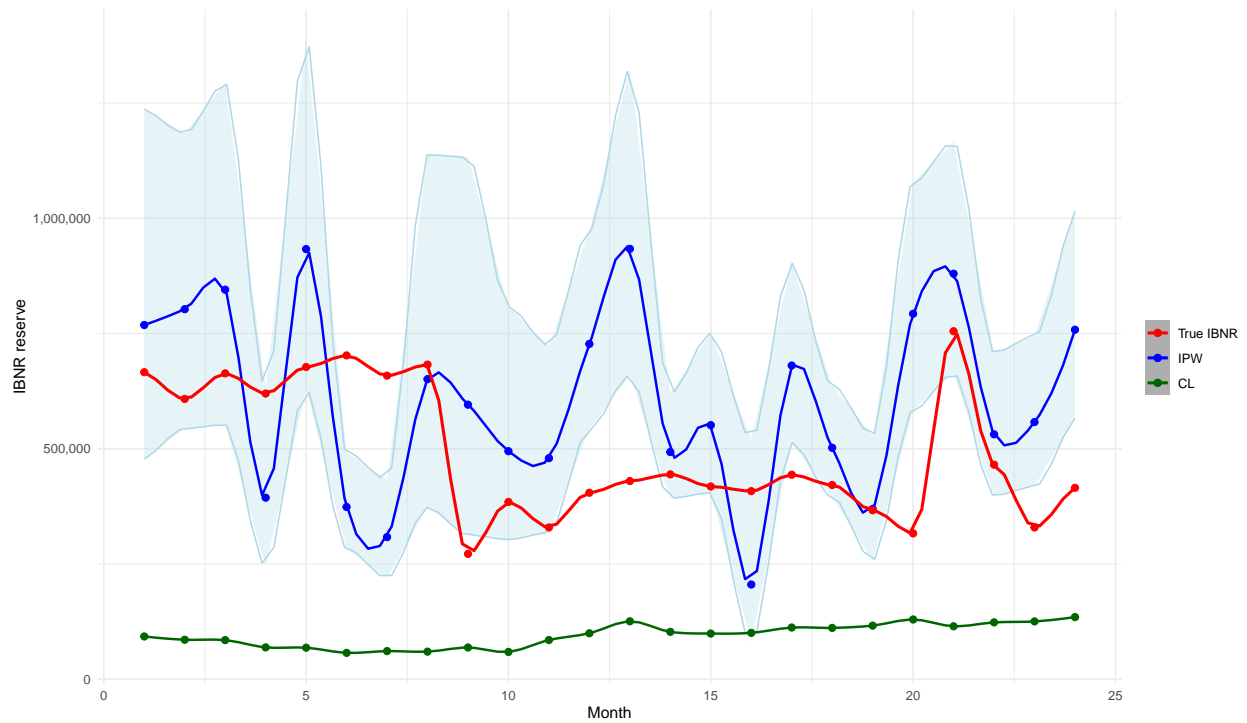Figure 10: Estimation for the RBNS reserve per month

Figure 11: Estimation for the IBNR reserve per month

the subpopulation decreases, the estimation variance of the IPW increases. Despite this variability,

28

the IPW estimator consistently outperforms the traditional Chain-Ladder method over the entire 24-month period, as indicated in Table 8.

Observation 1. *We encountered instability in the behavior of the IPW estimator (i.e absurdly abnormal large reserves) when performing the estimation on some dates, along the same lines as the behavior described in Section 4.3. To address this issue, we implemented the adjusted version of the IPW estimator, as described in algorithm 1, and compared it to the raw IPW estimator. If the percentual difference between the two estimators exceeded a certain threshold (e.g., more than 3%), we retained the adjusted estimation. For cases where the difference was not significant, we kept the original estimation. As a result, the estimations become more stable across dates.*

Note that the adjustment can be implemented regardless of the threshold rule, but it may introduce a systematic downward bias. Therefore, in practical applications, it is advisable to apply the adjustment only when necessary.

6 Conclusions

Macro-level reserving models, particularly the Chain-Ladder method, overlook the underlying heterogeneity within the portfolio of policyholders, treating all claims equally, and providing modest estimations. Therefore, the estimation of the reserve does not benefit from the use of the individual attributes of the policyholders, which have been shown to provide a significant improvement in the accuracy of the methods.

In this paper, we address the limitation of macro-level reserving models by proposing a statistically justified macro-level reserve estimator based on Inverse Probability Weighting (IPW). Unlike traditional macro-level models, our method incorporates individual-level information in the weights to improve the accuracy of reserve estimation. Moreover, such incorporation is achieved within a less complex framework compared to micro-level models, in the sense that no explicit assumptions on claim frequency or severity are made.

The IPW estimator serves as a hybrid approach that bridges the gap between macro and micro-level methods. It assigns attribute-driven weights to each claim, allowing for a development factor specific to each claim's settlement, similar to the familiar principles of the Chain-Ladder method when applied at the granular level. This estimator represents an initial step towards obtaining more precise reserves from macro-level models and serves as an intermediate stage in the development of a customized micro-level reserving model.

We believe that the IPW estimator offers a possibly seamless transition from macro to micro-level reserving for insurance companies. We hope practitioners find this method appealing as it is a natural extension of the traditional Chain-Ladder method, accounting for portfolio heterogeneity in a statistically justified fashion.

Future research should explore alternative approaches for estimation, potentially through the development of tailored models specifically designed for inclusion probabilities as in the development factor in Equation (12), which is simpler to interpret. Additionally, investigating the connection between claim reserving and population sampling techniques holds promise for further advancements in estimating reserves. We are currently engaged in related research, Calcetero-Vanegas et al. (2023), delving deeper into the implications of survey sampling theory on the claim reserving problem.

Acknowledgments

References

Andersen, P. K., Borgan, Ø., Hjort, N. L., Arjas, E., Stene, J., and Aalen, O. (1985). Counting process models for life history data: A review [with discussion and reply]. *Scandinavian Journal of Statistics*, pages 97–158.

Antonio, K. and Plat, R. (2014). Micro-level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal*, 2014(7):649–669.

Argyropoulos, C. and Unruh, M. L. (2015). Analysis of time to event outcomes in randomized controlled trials by generalized additive models. *PLoS One*, 10(4):e0123784.

Arnab, R. (2017). *Survey sampling theory and applications*. Academic Press.

Bender, A., Groll, A., and Scheipl, F. (2018). A generalized additive model approach to time-to-event analysis. *Statistical Modelling*, 18(3-4):299–321.

Berger, Y. G. (1998). Rate of convergence for asymptotic variance of the horvitz–thompson estimator. *Journal of Statistical Planning and Inference*, 74(1):149–168.

Block, H. W., Savits, T. H., and Singh, H. (1998). The reversed hazard rate function. *Probability in the Engineering and Informational Sciences*, 12(1):69–90.

Bou-Hamad, I., Larocque, D., and Ben-Ameur, H. (2011). A review of survival trees.

Boumezoued, A. and Devineau, L. (2017). Individual claims reserving: a survey.

Buckby, J., Wang, T., Zhuang, J., and Obara, K. (2020). Model checking for hidden markov models. *Journal of Computational and Graphical Statistics*, 29(4):859–874.

Calcetero-Vanegas, S., Badescu, A., and Lin, S. (2023). A survey sampling framework for claim reserving methods in general insurance. *In preparation*.

Carrato, A. and Visintin, M. (2019). From the chain ladder to individual claims reserving using machine learning techniques. In *ASTIN Colloquium, Cape Town, South Africa, April*, pages 2–5.

Chauvet, G. and Vallée, A.-A. (2018). Consistency of estimators and variance estimators for two-stage sampling.

Chen, Q., Elliott, M. R., Haziza, D., Yang, Y., Ghosh, M., Little, R. J. A., Sedransk, J., and Thompson, M. (2017). Approaches to improving survey-weighted estimates. *Statistical Science*, 32(2):227–248.

Cook, R. J., Lawless, J. F., et al. (2007). *The statistical analysis of recurrent events*. Springer.

Crevecoeur, J., Robben, J., and Antonio, K. (2022). A hierarchical reserving model for reported non-life insurance claims. *Insurance: Mathematics and Economics*, 104:158–184.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)*, 39(1):1–22.

Efron, B. and Stein, C. (1981). The jackknife estimate of variance. *The Annals of Statistics*, pages 586–596.

Francis, L. (2016). Astin working party releases report on reserving practices for general insurance worldwide. https://ar.casact.org/astin-working-party-releases-report-on-reserving-practices-for-general-insurance-worldwide/. Accessed: 2023-06-04.

Fung, T. C., Badescu, A. L., and Lin, X. S. (2021). A new class of severity regression models with an application to ibnr prediction. *North American Actuarial Journal*, 25(2):206–231.

Fung, T. C., Badescu, A. L., and Lin, X. S. (2022). Fitting censored and truncated regression data using the mixture of experts models. *North American Actuarial Journal*, 26(4):496–520.

George, B., Seals, S., and Aban, I. (2014). Survival analysis and regression models. *Journal of nuclear cardiology*, 21(4):686–694.

Gui, J. and Li, H. (2005). Threshold gradient descent method for censored data regression with applications in pharmacogenomics. In *Biocomputing 2005*, pages 272–283. World Scientific.

Hiabu, M. (2017). On the relationship between classical chain ladder and granular reserving. *Scandinavian Actuarial Journal*, 2017(8):708–729.

Hulliger, B. (1995). Outlier robust horvitz-thompson estimators. *Survey Methodology*, 21(1):79–87.

Klein, J. P. and Moeschberger, M. L. (2003). *Survival analysis: techniques for censored and truncated data*, volume 1230. Springer.

Maciak, M., Okhrin, O., and Pešta, M. (2021). Infinitely stochastic micro reserving. *Insurance: Mathematics and Economics*, 100:30–58.

Mack, T. (1994). Which stochastic model is underlying the chain ladder method? *Insurance: mathematics and economics*, 15(2-3):133–138.

Martínez-Miranda, M. D., Nielsen, J. P., and Verrall, R. (2012). Double chain ladder. *ASTIN Bulletin: The Journal of the IAA*, 42(1):59–76.

Mulayath Variyath, A. and Sankaran, P. (2014). Parametric regression models using reversed hazard rates. *Journal of Probability and Statistics*, 2014.

Quarg, G. and Mack, T. (2004). Munich chain ladder. *Blätter der DGVFM*, 26(4):597–630.

Rüschendorf, L. (2009). On the distributional transform, sklar's theorem, and the empirical copula process. *Journal of statistical planning and inference*, 139(11):3921–3927.

Särndal, C.-E., Swensson, B., and Wretman, J. (2003). *Model assisted survey sampling*. Springer Science & Business Media.

Schmidt, K. D. (2011). A bibliography on loss reserving.

Seaman, S. R. and White, I. R. (2013). Review of inverse probability weighting for dealing with missing data. *Statistical methods in medical research*, 22(3):278–295.

Shakur, A. H., Huang, S., Qian, X., and Chang, X. (2021). Survfit: doubly sparse rule learning for survival data. *Journal of Biomedical Informatics*, 117:103691.

Sonabend, R., Király, F. J., Bender, A., Bischl, B., and Lang, M. (2021). mlr3proba: an r package for machine learning in survival analysis. *Bioinformatics*, 37(17):2789–2791.

Taha, A., Cosgrave, B., Rashwan, W., and McKeever, S. (2021). Insurance reserve prediction: Opportunities and challenges. In *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 290–295. IEEE.

Taylor, G., McGuire, G., and Sullivan, J. (2008). Individual claim loss reserving conditioned by case estimates. *Annals of Actuarial Science*, 3(1-2):215–256.

Thompson, S. K. (2012). *Sampling*, volume 755. John Wiley & Sons.

Tseung, S. C., Badescu, A. L., Fung, T. C., and Lin, X. S. (2021). Lrmoe. jl: a software package for insurance loss modelling using mixture of experts regression model. *Annals of Actuarial Science*, 15(2):419–440.

Verbelen, R., Gong, L., Antonio, K., Badescu, A., and Lin, S. (2015). Fitting mixtures of erlangs to censored and truncated data using the em algorithm. *ASTIN Bulletin: The Journal of the IAA*, 45(3):729–758.

Verrall, R. J. and Wüthrich, M. V. (2016). Understanding reporting delay in general insurance. *Risks*, 4(3):25.

Wang, P., Li, Y., and Reddy, C. K. (2019). Machine learning for survival analysis: A survey. *ACM Computing Surveys (CSUR)*, 51(6):1–36.

Wiegrebe, S., Kopper, P., Sonabend, R., and Bender, A. (2023). Deep learning for survival analysis: A review. *arXiv preprint arXiv:2305.14961*.

Wüthrich, M. V. (2018a). Machine learning in individual claims reserving. *Scandinavian Actuarial Journal*, 2018(6):465–480.

Wüthrich, M. V. (2018b). Neural networks applied to chain–ladder reserving. *European Actuarial Journal*, 8:407–436.

Wüthrich, M. V. and Merz, M. (2008). *Stochastic claims reserving methods in insurance*. John Wiley & Sons.

Yanez, J. S. and Pigeon, M. (2021). Micro-level parametric duration-frequency-severity modeling for outstanding claim payments. *Insurance: Mathematics and Economics*, 98:106–119.

Yao, L., Chu, Z., Li, S., Li, Y., Gao, J., and Zhang, A. (2021). A survey on causal inference. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 15(5):1–46.

Zong, X., Zhu, R., and Zou, G. (2018). Improved horvitz-thompson estimator in survey sampling. *arXiv preprint arXiv:1804.04255*.