

Consideration of a structural-change point in the chain-ladder method

Hyuk Sung Kwon^{1,a}, Uy Quoc Vu^a

^aDepartment of Statistics and Actuarial Science, Soongsil University, Korea;

Abstract

The chain-ladder method, for which run-off data is employed is popularly used in the rate-adjustment and loss-reserving practices of non-life-insurance and health-insurance companies. The method is applicable when the underlying assumption of a consistent development pattern is in regards to a cumulative loss payment after the occurrence of an insurance event. In this study, a modified chain-ladder algorithm is proposed for when the assumption is considered to be only partially appropriate for the given run-off data. The concept of a structural-change point in the run-off data and its reflection in the estimation of unpaid loss amounts are discussed with numerical illustrations. Experience data from private health insurance coverage in Korea were analyzed based on the suggested method. The performance in estimation of loss reserve was also compared with traditional approaches. We present evidence in this paper that shows that a reflection of a structural-change point in the chain-ladder method can improve the risk management of the relevant insurance products. The suggested method is expected to be utilized easily in actuarial practice as the algorithm is straightforward.

Keywords: structural-change point, chain-ladder method, loss reserve, loss development factor, run-off data, rate adjustment

1. Introduction

A variety of coverages managed by non-life and health insurers provide financial protection for unexpected losses due to accidents or disease. The determination of the benefit payments of such coverages are generally based on the actual loss amount instead of a fixed payment (as in the case of life insurance coverage); therefore, the actuarial pricing and reserving processes are more complicated. Further, it is often necessary to regularly adjust the premium rates for these coverages based on the loss ratio, which is the total contractual benefit payment for an event that is covered by insurance divided by the total net premium income within a certain evaluation period.

It is important to note the total benefit payment is not fully available in the recent experiential data due to the time lag between the time of an insurance event and the time when the insurance benefit payment is completely settled. The two main sources for the time lag are the lag between the occurrence of the insurance event and the acceptance of the claim report as well as the lag between the claim report filing and the complete settlement of the claim. The former is called “incurred but not reported” claims, and the latter is called “reported but not settled” claims in property and casualty insurance. In consideration of this time lag, the appropriate loss reserve must be more than the expected unpaid loss amount and needs to be established by an insurer. Therefore, the estimation of the unpaid loss amount

¹ Corresponding author: Department of Statistics and Actuarial Science, Soongsil University, 369, Sangdo-ro, Dongjak-gu, Seoul 06978, Korea. E-mail: hskwon@ssu.ac.kr

is an important actuarial problem, since it affects both the loss reserving process and the premium rate adjustment in non-life insurance policies. Reserving should also be considered an important concern in life insurance as discussed in Lee *et al.* (2014).

To estimate the unpaid loss amount, actuaries typically use a two-dimensional table wherein the run-off data of the cumulative loss payments of the insurer are organized according to the period of the accident and the development period after each accident period. The unpaid loss amount is then estimated based on the development pattern of the previous cumulative loss payments. This process is the well-known chain-ladder method. An important assumption regarding the chain-ladder method is the consistency of the development pattern over time. This assumption has been applied to a variety of statistical models that have been suggested. Therefore, as is addressed in the Casualty Actuarial Society (2001), the appropriateness of this assumption needs to be verified before the chain-ladder method can be used.

Sometimes the development pattern of recent cumulative loss payments has consistently been considered larger or smaller on the basis of the development pattern of previous periods; however, while similar development patterns are also observed within each group of periods. The accuracy of the estimation would deteriorate in such cases if the standard chain-ladder approach is used since a part of data information regarding the development-pattern is out-of-date. However, a portion of the data can still be utilized even though the standard chain-ladder method cannot be fully applied.

To address the above situation, Lee (2008) introduced the concept of a structural-change point with respect to run-off data and suggested two statistics for the identification of the structural-change point. The author used the p -values from the two-sample t -tests, for which the means of the loss development factors of two subsets from the same development period are compared, to derive the statistics. However, the suggested method depends on an assumption that the loss development factors over different development periods are independent, and this might not be reasonable since a correlation between the loss development factors associated with the same accident period is likely to exist. In addition, the discussion needs to be advanced for the development of a method to estimate of the unpaid loss amount when a structural-change point is identified and for the verification of the impact of a structural-change point on the estimation results.

This paper proposes a modified algorithm for the determination of whether the standard chain-ladder method can be applied by observing the existence of a structural-change point in run-off data and for the reflection of an identified change point in the estimation of unpaid loss amounts. A numerical illustration of the proposed algorithm based on experiential data from Korea is also presented for the verification of the suggested algorithm. It is expected that the proposed approach would contribute practically to the actuarial risk management process in regards to the relevant insurance coverage.

A variety of coverage types for non-life insurance and health insurance are available in Korea. The premium rates for most of these coverage types are typically renewed annually based on the experiential loss ratio of the corresponding coverages that is obtained with the use of the data of the most-recent three years. Recent changes in Korea's insurance industry have been quick along with intensified competition among the corresponding insurers. Innovations in regards to the process to estimate unpaid loss amounts will improve the solvency of relevant insurance products while the competencies of the products are maintained.

The structure of this paper is organized as follows. Past studies of the chain-ladder algorithm are briefly reviewed in Section 2. An algorithm for the identification and reflection of a structural-change point in run-off data is discussed in Section 3. In Section 4, a numerical illustration of the suggested algorithm is presented. Lastly, the study is summarized with concluding remarks in Section 5.

2. Literature review

The estimation of reserves is a very important concern for the management of property and casualty insurance and health insurance. The basic methods that are popularly used in practice are based on the ideas suggested in Bornhuetter and Ferguson (1972) and Tarbell (1934). They employed the chain-ladder algorithm that is well explained in Brown and Lennox (2015) and the Casualty Actuarial Society (2001). Since Kremer (1982), a variety of statistical models underpinned by the chain-ladder method as well as that allow for an analysis of variability associated with reserve estimates, have been discussed.

A generalized linear model was utilized for analyses of the chain-ladder method in Renshaw and Verrall (1998). Mack (1993) introduced a model to calculate the standard error for reserve estimates based on the chain-ladder method when there is an absence of assumptions related to statistical distribution. Additional discussions regarding the model were made in Pesta and Hudecová (2012), Peters *et al.* (2010), Schmidt (1997), and Wüthrich (2008). Wüthrich and Merz (2008) also demonstrated a variety of statistical models associated with the chain-ladder method; in addition, discussions regarding model comparisons that are in terms of various aspects can be found in Mack and Venter (2000), Verrall (2000), Hess and Schmidt (2002), and Taylor (2011). Further, a Bayesian method for an estimation of the parameters of the suggested models, which is used to derive a predictive distribution on the estimated reserve, was studied in England *et al.* (2012), and Taylor (2015).

In addition, the relevant issues regarding loss reserving that is based on the chain-ladder method have been addressed as follows. Wüthrich (2008) and Merz introduced a multivariate model that is used for analyses regarding multiple run-off data that are correlated. Verdonck and Van Wouwe (2011) discussed the effect of an outlier in individual claim data on the chain-ladder reserve estimate. The double chain-ladder method, in which both the claim count data and the loss payment data are considered together in order to obtain robust reserve estimates, was considered in Martínez Miranda *et al.* (2015).

Riegel (2014) recently suggested a model that is applied for loss data associated with coverages that are commonly exposed to large claims. Heberle and Thomas (2014) introduced a method to estimate loss development factors in run-off data for which fuzzy numbers are used. Schiegl (2015) considered three of the aspects (occurrence, report, and payment) of claims for an assessment of the applicability of the chain-ladder method. Verrall and Wüthrich (2015) developed a model for correlations among loss development factors according to the development period.

The focus of previous studies is on cases for which the standard chain-ladder method can be applied. However, as is addressed in the previous section, a variety of solutions and the relevant statistical models need to be studied for cases where the standard chain-ladder algorithm may produce unreasonable results.

3. Method

As has been introduced, actuaries generally use run-off data to estimate unpaid loss amounts. In the run-off data, cumulative loss payments are organized according to accident period and development period after the occurrence of an accident. Table 1 presents a typical run-off data, where row i (accident period) presents development information regarding loss payments associated with claims for which the covered insurance event occurred in the i^{th} period, and column k (development period) represents the k^{th} period after each accident period. The (i, k) -element of the table, denoted by $c(i, k)$, is the total cumulative payments up to the k^{th} development period for the claims in the i^{th} accident period. The numbers in the blank area of the table are unknown and need to be estimated because the

Table 1: Run-off data

Accident period i	Development period k						
	1	2	3	...	$n-2$	$n-1$	n
1	$c(1, 1)$	$c(1, 2)$	$c(1, 3)$...	$c(1, n-2)$	$c(1, n-1)$	$c(1, n)$
2	$c(2, 1)$	$c(2, 2)$	$c(2, 3)$...	$c(2, n-2)$	$c(2, n-1)$	
3	$c(3, 1)$	$c(3, 2)$	$c(3, 3)$...	$c(3, n-2)$		
...			
$n-2$	$c(n-2, 1)$	$c(n-2, 2)$	$c(n-2, 3)$				
$n-1$	$c(n-1, 1)$	$c(n-1, 2)$					
n	$c(n, 1)$						

Table 2: Loss development factors

Accident period i	Successive development period j						
	2	3	4	...	$n-2$	$n-1$	n
1	$f(1, 2)$	$f(1, 3)$	$f(1, 4)$...	$f(1, n-2)$	$f(1, n-1)$	$c(1, n)$
2	$f(2, 2)$	$f(2, 3)$	$f(2, 4)$...	$f(2, n-2)$	$f(2, n-1)$	
3	$f(3, 2)$	$f(3, 3)$	$f(3, 4)$...	$f(3, n-2)$		
...			
$n-3$	$f(n-3, 2)$	$f(n-3, 3)$	$f(n-3, 4)$				
$n-2$	$f(n-2, 2)$	$f(n-2, 3)$					
$n-1$	$f(n-1, 2)$						

Column index j indicates the successive development periods from the $(j-1)^{th}$ period to the j^{th} period after an accident period.

data is usually obtained at the end of accident period n .

In practice, the use of the chain-ladder method by actuaries to estimate the unpaid loss amount of each accident period proceeds as follows. First, the individual loss development factor from a period to the next period, denoted by $f(i, k)$, is first calculated by $c(i, k)/c(i, k-1)$ for each pair of i and k , where $1 \leq i \leq n-1$ and $2 \leq k \leq n-i+1$. $f(i, k)$ indicates the degree of the increment of the cumulative loss payment in successive development periods (from the $(k-1)^{th}$ development period to the k^{th} development period) for claims associated with the accident period i . Then, Table 2 is derived after the run-off data in Table 1 calculates all the individual loss development factors.

The representative value of the loss development factors in each column of Table 2 needs to be determined to estimate the unpaid loss amount of claims associated with each accident period. A number of popular methods help determine the representative value of the loss development factors for the successive development period j , denoted by $f_1(j)$ for $j = 2, \dots, n$, as:

Method 1. Average value of all of the available loss development factors such as

$$f_1(j) = \frac{1}{n-j+1} \sum_{i=1}^{n-j+1} f(i, j).$$

Method 2. Average value of all of the available loss development factors for which the maximum and

minimum values are excluded such as

$$f_2(j) = \frac{1}{n-j-1} \left\{ \sum_{i=1}^{n-j+1} f(i, j) - \max_{1 \leq i \leq n-j+1} f(i, j) - \min_{1 \leq i \leq n-j+1} f(i, j) \right\}.$$

Method 3. Ratio based on all of the available cumulative loss payments that are associated with the $(j-1)^{th}$ period and the j^{th} period such as

$$f_3(j) = \frac{\sum_{i=1}^{n-j} c(i, j)}{\sum_{i=1}^{n-j} c(i, j-1)}.$$

Based on the derived values of $f(j)$ for $j = 2, \dots, n$, the unpaid loss amount for the accident period i , which is denoted by $u(i)$, is estimated as:

$$\hat{u}(i) = \begin{cases} 0, & i = 1, \\ c(i, n-i+1) \cdot \left\{ \prod_{j=n-i+2}^n f(j) - 1 \right\}, & i = 2, \dots, n. \end{cases} \quad (3.1)$$

Last, ultimate total loss amount for the entirety of the accident periods of Table 2 is the sum of $\hat{u}(i)$ for $i = 1, \dots, n$ by equation (3.1). Importantly, it is assumed for the above methods that the pattern of the cumulative loss payments over time that is observed in the past will continue in the future. Therefore, it is inappropriate to use the standard chain-ladder method if this assumption is no longer applicable.

Care needs to be taken if a noticeable change of loss development pattern from a certain accident period is identified. Conceptually, a structural-change point in run-off data represents a certain accident period for which the loss development factors show a different pattern from previous accident periods in some successive development periods; however, the loss development factors of both groups of periods are considered to remain consistent. The structural-change point arises due to factors that include a change of the entire payment process or a change of the relevant regulations.

Methods 1 to 3 can lead to an inaccurate estimate of the unpaid loss amount if there is a structural-change point due to the implicit assumption that the loss development pattern does not change. In addition, the degree of inaccuracy is likely to be large when a structural-change point is identified in the earlier development periods since a loss development factor generally decreases over a development period. For the consideration of the structural-change point in the estimation of the unpaid loss amount, it is necessary to develop a method for the identification of a structural-change point and for the calculation of the representative value of the loss development factors whereby the identified structural-change point is reflected.

The two-sample t -test can be utilized for the characterization of a structural-change point and for the exploration of the existence of a structural-change point, as discussed in Lee (2008). The sample mean of the loss development factors that are ahead of the accident period i in the successive development period j is denoted by $\bar{X}_1(i, j)$, while the sample mean of the loss development factors from the accident period i in the corresponding successive development period is denoted by $\bar{X}_2(i, j)$. Therefore, the following applies:

$$\bar{X}_1(i, j) = \frac{1}{i-1} \sum_{k=1}^{i-1} f(k, j), \quad \bar{X}_2(i, j) = \frac{1}{n-(i+j-1)} \sum_{k=i}^{n-j} f(k, j).$$

Table 3: The p -values of the two-sample t -test

Accident period i	Successive development period j			
	2	3	...	t
3	$p(3, 2)$	$p(3, 3)$...	$p(3, t)$
4	$p(4, 2)$	$p(4, 3)$...	$p(4, t)$
⋮	⋮	⋮	⋮	⋮
s	$p(s, 2)$	$p(s, 3)$...	$p(s, t)$

Based on $\bar{X}_1(i, j)$ and $\bar{X}_2(i, j)$, the null hypothesis for which the loss development factors of the two groups (split by the accident period i) in the successive development period j share the same mean is tested by the test statistic, which is denoted by $t(i, j)$ as:

$$t(i, j) = \frac{\bar{X}_1(i, j) - \bar{X}_2(i, j)}{S_p \sqrt{\frac{1}{i-1} + \frac{1}{n-(i+j-2)}}}, \quad i = 1, 2, \dots, n \text{ \& } j = 2, 3, \dots, n,$$

where

$$S_p = \frac{\sum_{k=1}^{i-1} (f(k, j) - \bar{X}_1(i, j))^2 + \sum_{k=i}^{n-(i+j-1)} (f(k, j) - \bar{X}_2(i, j))^2}{n-j-1}.$$

Assuming that the loss development factors of the two groups are the random samples of normal distribution, $t(i, j)$ follows t -distribution with $n-2$ degrees of freedom. Let $p(i, j)$ be the p -value associated with $t(i, j)$. Since each sample of the test must comprise at least two observations, $((n-4)(n-3))/2$ cases of the possible t -tests in Table 2 can be considered. However, early successive development periods are typically more influential than the later successive development periods since most of the loss payments are made within a short time period after an accident. The loss development factors of a certain accident period should also be considered together since a correlation between them is likely to exist.

Increased efficiency is therefore achieved if the focus is an influential part of the run-off data when a structural change point is to be identified. There are always few data in the later successive development periods (there is only one available loss development factor in the last successive development period in run-off data); therefore, loss development factors associated with the later successive development periods are not likely to contribute to identify a structural change point. In addition, the impact of loss development factors in the later successive development period is almost negligible since those are usually very close to 1.

Actuarial discretion must be involved in the determination of the part of the data selected for close investigation since the determination of the influential part of the data depends on the type of coverage by which the loss development pattern varies. The rectangular area of Table 2 associated with the accident periods from 3 to s and the successive development periods from 2 to t , where both s and t need to be determined, are relevant. A table consisting of the p -values of the t -test in the focused area of the given run-off data is then obtained (Table 3).

Based on Table 3, evidence for the existence of a structural-change point can be investigated. In addition, an appropriate reflection of an identified structural-change point can be considered for the estimation of the unpaid loss amount. If a certain accident period m is considered as a structural-change

point based on the p -values, it is reasonable to reflect the structural-change point in the estimation of the unpaid loss amount. The criterion for the structural-change point needs to be determined according to the discretion of the actuaries and based on factors that include the loss development pattern, the number of significant p -values, and the influential successive development periods.

A possible method for the consideration of an identified structural-change point is the application of the full weight of the loss development factors from accident period m . Consequently, the application of Method 3, $f_4(j)$ is determined with the following equation:

$$f_4(j) = \begin{cases} \frac{\sum_{i=m}^{n-j} c(i, j)}{\sum_{i=m}^{n-j} c(i, j-1)}, & \text{for } 2 \leq j \leq t, \\ \frac{\sum_{i=1}^{n-j} c(i, j)}{\sum_{i=1}^{n-j} c(i, j-1)}, & \text{for } t+1 \leq j \leq n. \end{cases} \quad (3.2)$$

In equation (3.2), all available loss development factors (regardless of the identified structural change point) were used for the calculation of successive development periods from $t+1$ due to the lack of information in run-off data for recent accident periods. However, the impact of the method is expected to be small since t has been selected so that loss development factors after the successive development period t are very close to 1.

Lastly, it is necessary to use the suggested method for the quantification of the variability of the estimated unpaid loss amount. The variance of the estimated unpaid loss amount for each accident period, denoted by $\widehat{\text{MSE}}(\hat{u}(i))$ for $i = 2, 3, \dots, n$, where $f(j)$ is determined by equation (3.2), can be obtained through a simple modification of the formula of Mack (1993) as:

$$\widehat{\text{MSE}}(\hat{u}(i)) = (\hat{u}(i))^2 \left[\sum_{k=n-m} e_k \left\{ \frac{1}{c(i, k)} + \frac{1}{\alpha} \right\} \right], \quad (3.3)$$

where $e_k = \hat{\delta}_k^2 / f(k)^2$, $\alpha = \sum_{j=m}^{n-k} c(j, k)$,

$$\hat{\delta}_k^2 = \frac{1}{n-m-k} \sum_{i=m}^{n-k} \left\{ \frac{c(i, k+1)}{c(i, k)} - f(k) \right\}^2, \quad 1 \leq k \leq n-m-1.$$

Similarly, the variance estimate of the total unpaid loss amount for all accident periods denoted by $\widehat{\text{MSE}}(\hat{u})$ is expressed as:

$$\widehat{\text{MSE}}(\hat{u}) = \sum_{i=2}^n \left[\widehat{\text{MSE}}(\hat{u}(i)) + c(i, n) \sum_{j=i+1}^n c(j, n) \left\{ \sum_{k=n-m} 2 \frac{e_k}{\alpha} + \sum_{k=n-m+1} 2 \frac{e_k}{\alpha} \right\} \right]. \quad (3.4)$$

Using equation (3.4), the standard error of the estimated unpaid loss amount can be calculated. The standard error can be considered to establish the loss reserve for claims that have not been fully paid. A numerical example of the process discussed in this section is illustrated in the following section.

4. Numerical example

4.1. Source

Premium rates of the variety of the coverage types for non-life insurance and health insurance should be adjusted every year through insurance regulation; therefore, the estimation of the unpaid loss

Table 4: Run-off data for analysis

Data	Year of rate adjustment	Data period (accident period)
Data 1	2010	December 2006 – November 2009
Data 2	2011	December 2007 – November 2010
Data 3	2012	December 2008 – November 2011
Data 4	2013	December 2009 – November 2012

amount represents a massive workload for Korean actuaries. The Actuarial Standard of Practice in Korea recommends the use of the most recent three-year experiential data for the rate adjustment process. In this study, four sets of three-year experiential data regarding healthcare coverage payments managed by a non-life insurance company in Korea were used for a numerical illustration of the method suggested in Section 3. The coverage provides an incurred medical cost that is subject to policy deductions, limits and co-payments when the hospitalization of an insured person is due to defined diseases.

The four sets of three-year run-off data were used for the rate adjustments in the years 2010 to 2013, and Table 4 presents the data period for each dataset. Cumulative payments are organized by each quarter according to the accident period and the development period after each accident period. Therefore, for the dataset analyzed in this study, the unit period in Table 1 is a quarter, with $i = 1, 2, \dots, 12$ and $1 \leq k \leq 13 - i$. Further, the actual value of the ultimate payment of the claims of each accident period that is associated with the lower triangle of Table 1 is fully known, allowing for a performance evaluation of the method suggested in this study.

4.2. Investigation of a structural-change point

First, the individual loss development factors in Table 5 are calculated with the use of the four datasets, (Table 2). It is evident that the loss development factors of the earlier development periods are more influential compared to those of later development periods. The impact of a structural-change point (if it exists) on the calculation of the unpaid loss amount is not large for later successive development periods when the loss development factors are relatively small. The increment of the loss development factors becomes less than 1% from the successive development period 6 in Table 5; therefore, it is efficient to focus on the development periods from 2 to 5. Further, there are more available loss development factors in the earlier periods that facilitate the suggested two-sample t -test for the identification of a structural-change point. Therefore, s and t are set as 7 and 5, respectively, in Table 3.

Twenty-four actual values of loss development factor for each successive development period were available because the payments for all accident periods considered in this study had been completed when data were collected. Using the information, we tried to check the existence of structural change point and validate the implicit assumption in the t -test discussed in Section 3 that loss development factors follow normal distribution if there is no structural change point even if the number of observation is limited.

First, all of the individual loss development factors in the entire data period were observed graphically according to successive development periods from 2 to 5. Figure 1 presents the histograms of the loss development factors in each successive development period. Overall, the loss development factors for each successive development period do not seem to show the shape of single normal distribution. The graphs are not symmetric in the successive development periods 2 and 4; in addition, the graphs are not uni-modal in the successive development periods 3 and 5. Therefore, it is reasonable to explore the existence of structural change point.

Table 5: Individual loss development factors based on data

Accident period <i>i</i> (quarter)	Successive development period <i>j</i> (quarter)											
	2	3	4	5	6	7	8	9	10	11	12	
Data 1	1	1.8797	1.1458	1.0495	1.0216	1.0086	1.0086	1.0020	1.0020	1.0011	1.0003	1.0004
	2	1.8810	1.1318	1.0471	1.0247	1.0058	1.0019	1.0019	1.0033	1.0003	1.0003	
	3	1.9029	1.1455	1.0530	1.0226	1.0078	1.0022	1.0019	1.0014	1.0004		
	4	1.9551	1.1356	1.0467	1.0187	1.0079	1.0022	1.0008	1.0016			
	5	1.9217	1.1301	1.0528	1.0273	1.0072	1.0035	1.0022				
	6	1.8884	1.1348	1.0538	1.0247	1.0091	1.0022					
	7	1.8247	1.1285	1.0541	1.0221	1.0078						
	8	1.9175	1.1318	1.0547	1.0215							
	9	1.8585	1.1154	1.0488								
	10	1.8107	1.1218									
	11	1.7812										
Data 2	1	1.9217	1.1301	1.0528	1.0273	1.0072	1.0035	1.0022	1.0014	1.0015	1.0000	1.0000
	2	1.8884	1.1348	1.0538	1.0247	1.0091	1.0022	1.0012	1.0010	1.0001	1.0001	
	3	1.8247	1.1285	1.0541	1.0221	1.0078	1.0026	1.0025	1.0000	1.0002		
	4	1.9175	1.1318	1.0547	1.0215	1.0048	1.0033	1.0000	1.0009			
	5	1.8585	1.1154	1.0488	1.0246	1.0081	1.0031	1.0008				
	6	1.8107	1.1218	1.0452	1.0228	1.0069	1.0015					
	7	1.7812	1.1220	1.0560	1.0214	1.0052						
	8	1.8365	1.1430	1.0483	1.0230							
	9	1.8908	1.1226	1.0530								
	10	1.8319	1.1254									
	11	1.7712										
Data 3	1	1.8585	1.1154	1.0488	1.0246	1.0081	1.0031	1.0008	1.0008	1.0000	1.0000	1.0000
	2	1.8107	1.1218	1.0452	1.0228	1.0069	1.0015	1.0011	1.0012	1.0004	1.0000	
	3	1.7812	1.1220	1.0560	1.0214	1.0052	1.0029	1.0023	1.0011	1.0000		
	4	1.8365	1.1430	1.0483	1.0230	1.0048	1.0029	1.0016	1.0012			
	5	1.8908	1.1226	1.0530	1.0285	1.0104	1.0041	1.0018				
	6	1.8319	1.1254	1.0535	1.0278	1.0097	1.0054					
	7	1.7712	1.1363	1.0587	1.0260	1.0083						
	8	1.8468	1.1511	1.0572	1.0231							
	9	1.8601	1.1349	1.0552								
	10	1.8021	1.1304									
	11	1.7485										
Data 4	1	1.8908	1.1226	1.0530	1.0285	1.0104	1.0041	1.0018	1.0020	1.0003	1.0003	1.0003
	2	1.8319	1.1254	1.0535	1.0278	1.0097	1.0054	1.0016	1.0015	1.0007	1.0009	
	3	1.7712	1.1363	1.0587	1.0260	1.0083	1.0038	1.0017	1.0018	1.0007		
	4	1.8468	1.1511	1.0572	1.0231	1.0064	1.0035	1.0022	1.0016			
	5	1.8601	1.1349	1.0552	1.0349	1.0102	1.0033	1.0019				
	6	1.8021	1.1304	1.0662	1.0323	1.0103	1.0038					
	7	1.7485	1.1423	1.0617	1.0306	1.0085						
	8	1.8422	1.1484	1.0649	1.0306							
	9	1.8733	1.1475	1.0615								
	10	1.7697	1.1452									
	11	1.8609										

Assuming that the two sets of the loss development factors that are separated by a structural-change point follow a normal distribution, the standard unpaired two-sample *t*-test that is discussed in Section 3 was performed. Table 6 presents the *p*-values of the *t*-test for accident periods where two samples with at least two individual loss development factors in all successive development periods from 2 to 5 could be considered. An accident period was considered as a structural-change point if more than one *p*-value from the *t*-test associated with the accident period is significant with a 5% significance level. When multiple change points are present, the point that includes the more-influential

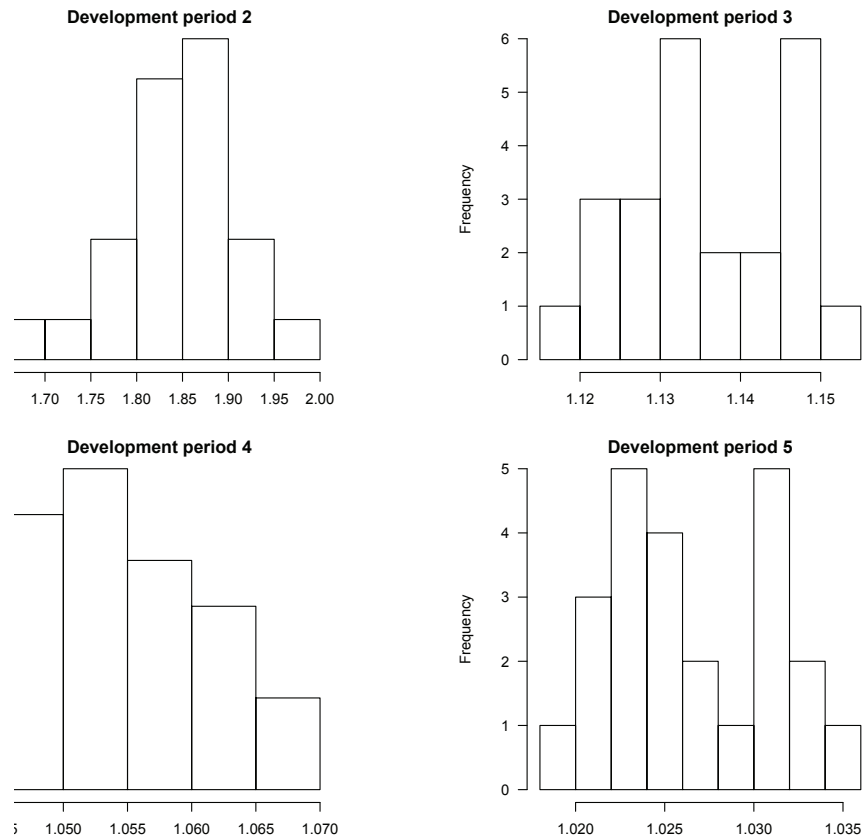


Figure 1: Histogram of the loss development factors in successive development periods 2 to 5.

successive development period was selected as the change point in consideration of development periods and p -values.

In data 1, both accident periods 6 and 7 comprise more than one successive development period in which the p -value of the t -test is less than 0.05. By comparing the p -values, the accident period 7 was selected as the structural-change point. Likewise, structural-change points were explored in the remaining three sets of data. In data 2, a structural-change point that satisfies the criterion was not identified. The accident period 5 was detected as the structural-change point in data 3 where the successive development periods 4 and 5 show significant p -values. Lastly, the accident periods 3 and 5 are the candidate structural-change points of data 4. Since accident period 3 comprise earlier development periods with significant p -values when compared with accident period 5, accident period 3 is considered as the structural-change point.

Lastly, quantile-quantile (q-q) plots for the groups of loss development factors in Table 5 that are divided by the identified structural-change points for each of the successive development periods from 2 to 5 were observed (Figure 2) in order to validate the assumption regarding the distribution of samples in the t -test. Plots in j^{th} row shows q-q plots of loss development factors grouped by the identified structural change point(s) associated with the successive development period $j + 1$.

For example, the two plots in the first row in Figure 2 are q-q plots of the two groups of loss

Table 6: The p -values of the two-sample t -tests

	Accident period i	Successive development period j (quarter)			
		2	3	4	5
Data 1	3	0.7301	0.2814	0.1518	0.8842
	4	0.6328	0.0385	0.4111	0.9540
	5	0.1581	0.0255	0.0679	0.3156
	6	0.0446	0.0477	0.1637	0.9248
	7	0.0258	0.0209	0.3985	0.5443
Data 2	3	0.0843	0.3533	0.5450	0.0205
	4	0.2606	0.3762	0.3353	0.1786
	5	0.0464	0.2394	0.1294	0.5457
	6	0.0392	0.8327	0.3918	0.2917
	7	0.1295	0.8278	0.7578	0.3537
Data 3	3	0.6631	0.0935	0.0235	0.5881
	4	0.8323	0.0375	0.1933	0.1526
	5	0.9976	0.2924	0.0393	0.0495
	6	0.3571	0.1330	0.0411	0.4488
	7	0.2864	0.0590	0.0405	0.9456
Data 4	3	0.2779	0.0113	0.0389	0.6741
	4	0.8657	0.0216	0.0681	0.3251
	5	0.6895	0.2644	0.0392	0.0106
	6	0.4296	0.1843	0.0016	0.2795
	7	0.6285	0.0469	0.1153	0.3345

development factors for the successive development period 2. The accident period 7 in data 1 (June 2009–August 2009) was the only accident period, which was considered as a structural change point associated with the successive development period 2, in the entire data period. Therefore, the first group contains six loss development factors, which are associated with accident periods ahead of the structural change point, and the second group includes the remaining 18 loss development factors (Notice that actual values of loss development factors in all accident periods and all successive development periods are available in this study). Evidence regarding a significant deviation of the underlying distribution of the loss development factors from the normal distribution was not found. However, a further verification of the assumption needs to be made with the use of more experiential data when they are available.

It should be addressed that the criterion for the detection of a structural-change point in run-off data can be modified by the judgment of practitioners. An optimal criterion might depend on the type of coverage since the characteristics of the pattern in the development of the cumulative loss payment differ according to the coverage type. In any case, the appropriateness of a detection method should be monitored with the use of experiential data within the process of an actuarial control cycle.

4.3. Estimation of unpaid loss amount

Using each of the identified structural-change points of the previous section, the representative loss development factor according to the development period was determined using $f_4(j)$ in equation (3.2), with the exception of those that are ahead of the accident period that is associated with the structural-change point; for example with data 1, accident period 7 was detected as the structural-change point. In consideration of the structural-change point, only the cumulative loss payments from the accident period 7 were used to calculate the representative loss development factors up to the successive development period 6. For the successive development periods from 7 to 12, all of the available data were used for the loss development factors because the later successive development periods are not

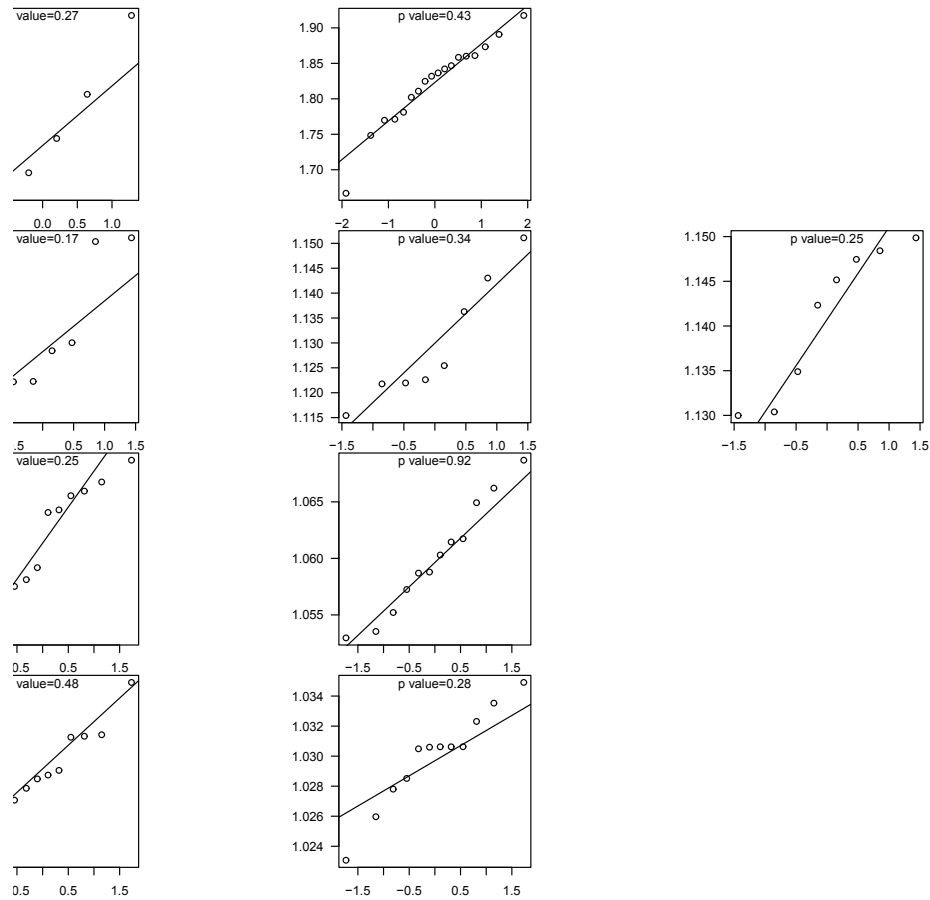


Figure 2: Quantile-quantile plots of homogeneous loss development factors.

influential.

Lastly, the unpaid loss amounts for each accident period were calculated using equation (3.1), while the standard errors of the estimated unpaid loss amounts were calculated using equation (3.4). Table 7 is a summary of the representative loss development factors and estimated unpaid loss amounts. In addition, simulations were performed to understand the distribution of estimated unpaid loss amount due to the randomness in loss development factors. This approach is useful to help determine and appropriate loss reserve that reflects the variability in the estimated unpaid loss amount. For the simulation, it is assumed that the loss development factors follow multivariate normal distribution. Loss development factors in Table 7 were used as the mean vector of multivariate normal distribution. The covariance matrix was also obtained using actual loss development factors from complete payment data. For each data in which a structural change point was identified, 10,000 sets of simulation were performed to estimate the unpaid loss amount. Figure 3 presents the results of simulations for data 1, data 3, and data 4.

Using the complete loss payment data, the performance of the suggested method is now compared with Methods 1 to 3. Table 8 shows the estimated unpaid loss amounts obtained by each method

Table 7: Representative loss development factors that reflect the identified structural-change points

Accident period	1	2	3	4	5	6	7
Data 1	1.8328	1.1237	1.0523	1.0218	1.0078	1.0033	1.0018
Data 2	1.8324	1.1259	1.0510	1.0227	1.0063	1.0026	1.0003
Data 3	1.8187	1.1333	1.0555	1.0264	1.0094	1.0047	1.0018
Data 4	1.8170	1.1419	1.0609	1.0297	1.0088	1.0036	1.0019
Accident period	8	9	10	11	12	Unpaid loss amount (SE)	
Data 1	1.0021	1.0006	1.0003	1.0004	1.0000	32,287 (1,090)	
Data 2	1.0009	1.0006	1.0000	1.0000	1.0000	N/A	
Data 3	1.0011	1.0001	1.0000	1.0000	1.0000	37,173 (1,250)	
Data 4	1.0017	1.0007	1.0006	1.0003	1.0000	43,439 (1,290)	

The unit of the unpaid loss amounts and their standard errors represents millions of Korean won.

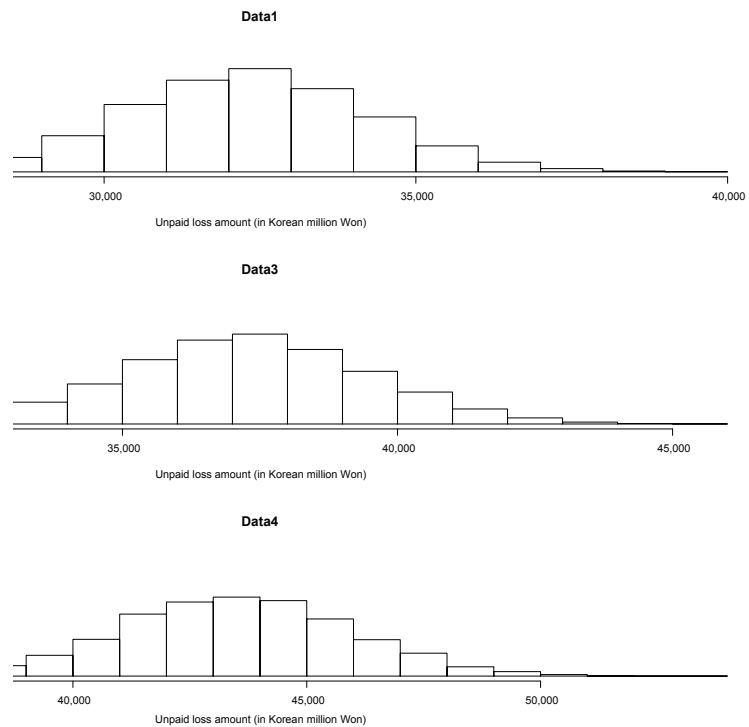


Figure 3: Results of simulation for estimated unpaid loss amount.

with the use of the actual outcomes for the four data. In data 1 and data 3, the estimated unpaid loss amounts according to the suggested method are closer to the actual outcome when compared to those according to Methods 1 to 3. The error rates of the estimated unpaid loss amount are significantly reduced when the standard errors of Table 7 are considered. The performance of the suggested method is not favorable in data 4; however, the differences among the methods are fairly small.

The results imply that the consideration of a structural-change point can improve the accuracy of the estimation of the unpaid loss amount even though the results were derived according to a specific coverage. In addition, the improvement is more noticeable when a structural-change point is identified for a more-recent accident period. The estimated unpaid loss amount is directly associated

Table 8: Comparison of the estimated unpaid loss amounts by estimation method (unit: millions of Korean won)

Method	Data			
	Data 1	Data 2	Data 3	Data 4
Method 1	33,754 (6.17%)	32,817 (7.80%)	35,663 (12.51%)	43,209 (7.94%)
Method 2	33,586 (5.64%)	32,784 (7.89%)	35,597 (12.67%)	43,326 (8.24%)
Method 3	33,337 (4.86%)	32,488 (8.73%)	35,685 (12.46%)	43,186 (7.89%)
The suggest method	32,287 (1.55%)	N/A	37,173 (8.81%)	43,439 (8.52%)
Actual outcome	31,793	35,594	40,763	40,029

Numbers in parenthesis indicate the error rate that is the (estimated value – actual value)/actual value.

with the combined periodic adjustment of the premium rate and loss reserving; therefore, the improved accuracy of the estimation contributes to the competency and solvency of the relevant coverage. The impact of a structural-change point would be greater when the degree of the differential between the loss development factors, separated by the structural-change point is more significant.

This paper proposes that the insurance companies that provide non-life and health coverage should develop a procedure to detect structural-change points considered for the actuarial control cycle. Based on the results of the current study, it is expected that an understanding of the impact of a structural-change point (if it exists) on pricing and reserving will improve the management of actuarial risk management. The suggested method of this study can be implemented simply to be implemented and included in an automated actuarial calculation system or in any of the relevant actuarial software modules.

5. Conclusion

Rate-making and loss-reserving are important actuarial works for the maintenance of the solvencies of non-life insurance and health insurance companies. Moreover, those works involve the estimation of unpaid loss amounts associated with loss events that occur in a certain period since an amount of time typically passes before payments are fully made. Evidently, the competency of the relevant products depend on the accuracy of the estimation. Therefore, an exploration by actuaries of various possible estimation methods to find an optimal method for which relevant experiential data is used is desirable and is especially the case when the standard methods in practice are inappropriate.

Actuaries normally use run-off data that contains up-to-date loss payment information. Run-off data are incomplete when they are utilized; therefore, methods for the estimation of the unpaid loss amount based on the given run-off data have been discussed in numerous papers. For the proposed method, it is implicitly assumed that the loss development pattern is consistent within the entire period will associated with the data. However, it is now known that the unpaid loss amount can be underestimated (or over-estimated) significantly when there is a change of the loss development pattern from a certain time point. In such cases, an innovative method can be explored since a part of the data information can be utilized.

This paper proposes a procedure to identify a structural-change point in the run-off data when the estimation of the unpaid loss amount reflects the identified structural-change point. Further, the estimation result from the use of the proposed method was compared with results derived from existing methods used in practice. For this comparison, we utilized experiential data associated with the complete loss payments of the most popular non-life coverage in Korea.

This research provided evidence on the improvement of the accuracy of the estimation of the unpaid loss amount for which an applicable structural-change point is considered. Furthermore, the impact of the consideration of a structural-change point on the estimation is likely larger when the change point is identified in a more-recent period. The suggested procedure should be utilized through an automated actuarial calculation module because it is simple to implement the suggested method in practice. The suggested procedure can also be modified according to the discretion of a practitioner and with respect to the characteristics of the coverage under consideration.

A specific coverage is explored in this study; therefore, an investigation of other types of coverage would be useful to obtain further insight into the effect of the consideration of a structural-change point in the run-off data on the estimation of unpaid loss amounts. Another possible research area is the development of a statistical model that can accommodate a structural-change point. Such efforts will improve the actuarial risk management of the relevant coverages. However, more refined practical procedure may be useful to accommodate structural change points in run-off data because the size of run-off data usually handled in actuarial practice is inadequate to fit comprehensive statistical models.

References

- Bornhuetter RL and Ferguson RE (1972). The actuary and IBNR, *Proceedings of the Casualty Actuarial Society*, **59**, 181–195.
- Brown RL and Lennox WS (2015). *Introduction to Ratemaking and Loss Reserving for Property and Casualty Insurance* (4th ed), ACTEX Publications, Winsted.
- Casualty Actuarial Society (2001). *Foundations of Casualty Actuarial Science* (4th ed), Casualty Actuarial Society, New York.
- England PD, Verrall RJ, and Wüthrich MV (2012). Bayesian over-dispersed Poisson model and the Bornhuetter & Ferguson claims reserving method, *Annals of Actuarial Science*, **6**, 258–283.
- Heberle J and Thomas A (2014). Combining chain-ladder claims reserving with fuzzy numbers, *Insurance: Mathematics and Economics*, **55**, 96–104.
- Hess KT and Schmidt KD (2002). A comparison of models for the chain-ladder method, *Insurance: Mathematics and Economics*, **31**, 351–364.
- Kremer E (1982). IBNR-claims and the two-way model of ANOVA, *Scandinavian Actuarial Journal*, **1982**, 47–55.
- Lee I, Lee HS, and Kim HT (2014). Analysis of reserves in multiple life insurance using copula, *Communications for Statistical Applications and Methods*, **21**, 23–43.
- Lee SI (2008). Testing structural changes in triangular data, *Communications of the Korean Statistical Society*, **15**, 551–562.
- Mack T (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates, *Astin Bulletin*, **23**, 213–225.
- Mack T and Venter G (2000). A comparison of stochastic models that reproduce chain ladder reserve estimates, *Insurance: Mathematics and Economics*, **26**, 101–107.
- Martínez Miranda MD, Nielsen JP, Verrall R, and Wüthrich MV (2015). Double chain ladder, claims development inflation and zero-claims, *Scandinavian Actuarial Journal*, **2015**, 383–405.
- Pesta M and Hudecová Š (2012). Asymptotic consistency and inconsistency of the chain ladder, *Insurance: Mathematics and Economics*, **51**, 472–479.
- Peters GW, Wüthrich MV, Shevchenko PV (2010). Chain ladder method: Bayesian bootstrap versus classical bootstrap, *Insurance: Mathematics and Economics*, **47**, 36–51.
- Renshaw AE and Verrall RJ (1998). A stochastic model underlying the chain-ladder technique, *British Actuarial Journal*, **4**, 903–923.

- Riegel U (2014). A bifurcation approach for attritional and large losses in chain ladder calculations, *Astin Bulletin*, **44**, 127–172.
- Schiegl M (2015). A model study about the applicability of the chain-ladder method, *Scandinavian Actuarial Journal*, **6**, 482–499.
- Schmidt KD (1997). Non-optimal prediction by the chain ladder method, *Insurance: Mathematics and Economics*, **21**, 17–24.
- Tarbell TF (1934). Incurred but not reported claim reserves. In *Proceedings of the Casualty Actuarial Society*, **20**, 275–280.
- Taylor G (2011). Maximum likelihood and estimation efficiency of the chain ladder, *Astin Bulletin*, **41**, 131–155.
- Taylor G (2015). Bayesian chain ladder models, *Astin Bulletin*, **45**, 75–99.
- Verdonck T and Van Wouwe M (2011). Detection and correction of outliers in the bivariate chain-ladder method, *Insurance: Mathematics and Economics*, **49**, 188–193.
- Verrall RJ (2000). An investigation into stochastic claims reserving models and the chain-ladder technique, *Insurance: Mathematics and Economics*, **26**, 91–99.
- Verrall RJ and Wüthrich MV (2015). Parameter reduction in log-normal chain-ladder models, *European Actuarial Journal*, **5**, 355–380.
- Wüthrich MV (2008). Prediction error in the chain ladder method, *Insurance: Mathematics and Economics*, **42**, 378–388.
- Wüthrich MV and Merz M (2008). *Stochastic Claims Reserving Methods in Insurance*, John Wiley & Sons, West Sussex.

Received November 16, 2016; Revised February 13, 2017; Accepted March 18, 2017