CLAIMS RESERVING WITH A STOCHASTIC VECTOR PROJECTION*

LUÍS PORTUGAL[1, 2], ATHANASIOS A. PANTELOUS[1,†] AND HIRBOD ASSA[1]

[1]DEPARTMENT OF MATHEMATICAL SCIENCES AND INSTITUTE FOR RISK AND UNCERTAINTY, UNIVERSITY OF LIVERPOOL, LIVERPOOL, UNITED KINGDOM

[2]PARTNER AND MANAGING DIRECTOR OF THE ACTUARIAL GROUP, LISBON, PORTUGAL

ABSTRACT

In the last three decades, a variety of stochastic reserving models has been proposed in the general insurance literature mainly using (or reproducing) the eminent Chain-Ladder claims reserving estimates. In practice, when the data doesn't satisfy the Chain-Ladder assumptions, high prediction errors might occur. Thus, in this paper, a combined methodology is proposed which is based on the stochastic vector projection method and uses the regression through the origin approach of Murphy (1994), but with heteroscedastic errors instead, and different to those that used by Mack (1993, 1994). Furthermore, the Mack (1993) distribution-free model appears to have higher prediction errors when it is compared with the proposed one, particularly, for data sets with increasing (regular) trends. Finally, three empirical examples with irregular and regular data sets illustrate the theoretical findings, and the concepts of best estimate and risk margin are reported.

JEL CLASSIFICATION : G22, C13, C18, C35

KEYWORDS: Stochastic Reserving; Chain-Ladder Distribution-Free; Vector Projection; Best Estimate; Risk Margin; Link Ratios; Loss Development Factors; Homoscedastic and Heteroscedastic Errors; Prediction Errors.

---

* **Acknowledgement**: We are extremely grateful to the anonymous reviewers who have afforded us considerable assistance in enhancing both the quality of the findings and the clarity of their presentation. The authors would like to acknowledge the gracious support of this work by the Actuarial Group (Lisbon, Portugal), and the EPSRC and ESRC Centre for Doctoral Training on Quantification and Management of Risk & Uncertainty in Complex Systems & Environments (EP/L015927/1). This paper benefited from the comments of participants (Bilal Ayyub, Michael A. H. Dempster, Ioannis Kyriakou, Colin O'Hare, Aihua Zhang) at the 2nd Quantitative Finance and Risk Analysis Symposium, Rhodes, Greece. Warm thanks are due to Greg Taylor and Richard Verrall who read very carefully preliminary versions of our paper and who have afforded us considerable assistance in enhancing both the quality of the findings and the clarity of their presentation. Particularly, Remark 2.2 and example 2 are due to Greg Taylor. All remaining errors are ours.

† **Corresponding Author**: Athanasios A. Pantelous is within the Department of Mathematical Sciences and the Institute for Risk and Uncertainty, University of Liverpool, Liverpool L697ZL, Liverpool, United Kingdom, e-mail: A.Pantelous@liverpool.ac.uk

Luis Portugal (luis.portugal@actuarial.pt, sglportu@liverpool.ac.uk);

Hirbod Assa (H.Assa@liverpool.ac.uk).

CLAIMS RESERVING WITH A STOCHASTIC VECTOR PROJECTION

ABSTRACT

In the last three decades, a variety of stochastic reserving models has been proposed in the general insurance literature mainly using (or reproducing) the eminent Chain-Ladder claims reserving estimates. In practice, when the data doesn't satisfy the Chain-Ladder assumptions, high prediction errors might occur. Thus, in this paper, a combined methodology is proposed which is based on the stochastic vector projection method and uses the regression through the origin approach of Murphy (1994), but with heteroscedastic errors instead, and different to those that used by Mack (1993, 1994). Furthermore, the Mack (1993) distribution-free model appears to have higher prediction errors when it is compared with the proposed one, particularly, for data sets with increasing (regular) trends. Finally, three empirical examples with irregular and regular data sets illustrate the theoretical findings, and the concepts of best estimate and risk margin are reported.

JEL CLASSIFICATION : G22, C13, C18, C35

KEYWORDS: Stochastic Reserving; Chain-Ladder Distribution-Free; Vector Projection; Best Estimate; Risk Margin; Link Ratios; Loss Development Factors; Homoscedastic and Heteroscedastic Errors; Prediction Errors.

1. INTRODUCTION

1.1 THE IMPORTANCE OF RESERVING

In this paper, we want to show that the Mack (1993) Stochastic Chain-Ladder distribution-free model for claims reserving can be further improved by using the *same* framework as before, however, with different loss development factors instead. In our approach, the factors are given by the vector projection (or regression through the origin) method in a *similar* way with Murphy (1994), but with *heteroscedastic* errors instead. It should be emphasised here that the *difference* between the incurred claims and the ultimate cost makes the choice of reserving strategy as one of the most critical aspect in the insurance industry. This fact is related mainly to Non-Life (General) insurance policies, but it can also affect some products of the Life insurance industry as well. In practice, the reserving policy can have various effects on the insurer; for instance, it can affect: a) the *profitability* (as well as the dividends for the shareholders or the bonuses for the managers), b) the *financial strength* of the insurer (i.e., insurer solvency), c) the *insurer's value* (its appraisal value), d) the paid *taxes* (or/and tax deductions) and e) the *capital requirements* on risk based capital or even traditional solvency systems.

Insurers work in a context of *risk* and *uncertainty*, and they need to know in advance the ultimate cost of claims. This demand arises due to three main factors:

- The existence of claims occurred but not yet notified.
- Lack of information on each claim when the notification is done.
- And the need to re-open some claims after being settled.

For many years (and even today, but in a less extent), there was not a commonly agreeable way of reserving among the insurers. That has also been recognized by a European Community document (European Community, 1999) which establishes as a priority the harmonization of the insurer's technical reserves' calculation on the framework of a new solvency regime, the *Solvency II*. For instance, the reserves should be calculated with an appropriate methodology for each liability and with the same term structure of interest rates on discounting. The aforementioned regime is suggested by a European Union directive (2009/138/CE)[1] that defines the fair value reserves, i.e., the amount that should allow a liability to be trans-

[1] Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II) (Text with EEA relevance) can be found on http://data.europa.eu/eli/dir/2009/138/oj

ferred to another insurer or/and reinsurer. The directive officially came in force on the January 1st, 2016, but it has already influenced the claims reserving strategy in several countries for some time now. In the same direction, in U.S., a similar legislation came into power on the January 1st, 2015 requiring insurers to regularly perform an *Own Risk and Solvency Assessment* (ORSA)[2]. This obliges insurance companies to issue their own assessment of their current and future risk through an internal risk self-assessment process and it allows regulators to form an enhanced view of an insurer's ability to withstand financial stress.

All these legislations in Europe, U.S. and other countries (such as U.K. and Australia among the others) not only oblige insurers to assess all the liabilities from all the lines of business, but also require a better estimation for the total reserves on the entire portfolio. Practically, this means that it is desirable to have estimates of reserves with a lower prediction error as possible.

The next parts of the paper are organized as follows. In the next subsections of the introduction, the methodology to get the best reserves estimate, and the vector projection approach are summarized. In Section 2, the stochastic distribution-free vector projection technique for estimating claims reserves is developed in details by using the Mack (1993, 1994) model's framework, i.e., having heteroscedastic errors, but now with a regression through the origin as this was proposed by Murray (1994), to estimate the loss development factors and the claims variance. Moreover, the calculation of mean squared errors is given, and thus, the prediction error is derived. In Section 3, the new approach is evaluated further with the use of three numerical examples with irregular and regular developments of data. Moreover, a summary of a use test with 114 triangles is provided. Finally, some conclusions are presented in Section 4.

1.2 METHODOLOGY TO GET THE BEST ESTIMATE

The most popular technique, which has the best reserve estimates (and risk margins), aggregates data on homogeneous groups of claims, and uses a triangle of past information that is used to estimate another triangle related with the future evolution of claims. The content may be the cumulative payments or the incurred claims (i.e., the cumulative payments plus the reserves). It is also possible to use these triangles to estimate the number of open claims, the

---

[2] Information regarding Own Risk and Solvency Assessment (ORSA) can be found on National Association of Insurance Commissioners webpage: http://www.naic.org/cipr_topics/topic_own_risk_solvency_assessment.htm. See, also Pooser, and Walker (2015).

number of settled claims, the number of outstanding claims or even the evolution of reimbursements.

A common example is the one presented below (see **Table 1**). Let us define *T* be the number of years of information, *i* the origin year (for instance, the accident year or the underwriting year), and *j* the development year of the claims after the claim origin year. With these definitions, we may also define the accumulated payments of each year of origin, and each year of development as $C_{i,j}$. Having these and assuming that *T* = 10 years, we may have the following triangle of past information. The origin year in this example is an accident year, and the upper triangle D_u is the set of the accumulated payments, where it is denoted by $D_u = \{C_{i,j}: i + j - 1 \leq T,\ T = 1, \dots, 10\}$.

Table 1: Triangle of Cumulative Payments (Mack 1993, 1994; England and Verrall 2002)

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	5012	8269	10907	11805	13539	16181	18009	18608	18662	18834
2	106	4285	5396	10666	13782	15599	15496	16169	16704	
3	3410	8992	13873	16141	18735	22214	22863	23466		
4	5655	11555	15766	21266	23425	26083	27067			
5	1092	9565	15836	22169	25955	26180				
6	1513	6445	11702	12935	15852					
7	557	4020	10946	12314						
8	1351	6947	13112							
9	3133	5395								
10	2063									

This represents the past history of claims. The data in **Table 1** has been used by several authors in the literature of claims reserving and it is a triangle of cumulative payments (Mack 1993, 1994; England and Verrall 2002). The upper triangle gives information from the past and the technique assumes that we may use it to forecast the future and estimate the lower triangle D_l given by $D_l = \{C_{i,j}: i + j - 1 > T, T = 1, \dots, 10\}$. In this case, we get a matrix that joins together the lower and upper triangles. The last column of the matrix gives the ultimate costs of each origin year. To estimate the lower triangle, we need to have the loss development factors. This statistic allows us to estimate each cell on the lower triangle, and relays on the method we consider for that. Before we go further, it is convenient to see the link ratios and the ratio of cumulative payments between adjacent cells on the upper triangle and for the same origin year. The link ratios and the loss development factor are presented on the following paragraphs.

In practice, to calculate the loss development factors several actuaries use the Chain-Ladder (CL) method. See for instance the following 2013 discussion about these techniques at the Institute of Actuaries (IoA) with T.A.G. Marcuson to say: "… *there is a reason why established techniques such as the Chain-Ladder and Bornhuetter-Ferguson (B-F) are so well-entrenched in actuarial reserving … it is because they are robust (certainly the B-F, but with suitable care the CL as well), common-sense approach to a problem. They apply to aggregate data, which means we can overcome some data efficiencies, and, most importantly, they are relatively easy to communicate to non-actuaries*" (Parodi, 2014).

The model summarizes all the link ratios $F_{i,j+1}$ (also called age-to-age factors) in one figure, the loss development factor that must be estimated. The CL loss development factor estimator, $\hat{b}_j^{CL}$ is given by the second component of Eq. (1.1)

$$F_{i,j+1} = \frac{C_{i,j+1}}{C_{i,j}} \text{ and } \hat{b}_j^{CL} = \frac{\sum_{i=1}^{T-j} C_{i,j+1}}{\sum_{i=1}^{T-j} C_{i,j}}. \tag{1.1}$$

However, the CL assumes that there is some stability of the payments over time, and very often this assumption is violated in reality. This is translated such as if this method is used, and the correspondent *prediction error* is estimated, then the latter will be high, which means also that the model is not accurate enough on estimating the reserves.

Let us see this with another example before coming back to **Table 1** later on. Assume that we have a steady growth of claims every year with a constant development factor, but with some diagonal effects that are accumulated year over year (e.g., claims inflation), see **Table 2**.

Table 2*:* Matrix with Trends

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	1000	1980	4039	8402	17643	37051	77066	157214	311283	591438
2	1100	2376	5251	11762	26465	59281	131012	282985	591438	
3	1320	3089	7351	17643	42344	100778	235821	537671		
4	1716	4324	11027	28229	71984	181401	448059			
5	2402	6486	17643	47990	129572	344661				
6	3603	10378	29993	86381	246187					
7	5766	17643	53988	164124						
8	9802	31758	102578							
9	17643	60340								
10	20063									

Then, we get the link ratios Table 3.

Table 3: Intermediate Link Ratios

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	1,98	2,04	2,08	2,10	2,10	2,08	2,04	1,98	1,90
2	2,16	2,21	2,24	2,25	2,24	2,21	2,16	2,09	
3	2,34	2,38	2,40	2,40	2,38	2,34	2,28		
4	2,52	2,55	2,56	2,55	2,52	2,47			
5	2,70	2,72	2,72	2,70	2,66				
6	2,88	2,89	2,88	2,85					
7	3,06	3,06	3,04						
8	3,24	3,23							
9	3,42								

It is clear from **Table 3** that there is no stability, and trends appear on almost all the link ratios. Thus, the CL method forecasts 3.12 for the first development factor, something that belongs to the past and don't reflect the current evolution (and the same is happening to the other loss development factors). Now, if there is no reason to believe that the factor tends to the past average, it is difficult to sustain the use of the CL. This also means that if the insurer increases (decreases) their claims payments velocity, then more (less) reserve is required by the CL method. Obviously, it is exactly the opposite of what we should expect and it is difficult to trust on the estimated reserves from the CL. This situation happens in practice very often, and it is due to several reasons, such as the speed of paying and settling the claims, the changes in underwriting and claims policies oblige the consideration of the most recent experience, and not so much the one from the past. In the relevant literature, papers do still rely on the CL method. Examples may be seen in several old, but also recent papers (Verrall 1989, 1990, 1991a, 1991b, 2000; Renshaw 1989; Wright 1990; Mack 1993, 1994; Zhang 2010; Miranda et al. 2012; Wüthrich and Merz 2014).

Other methods, like the B-F method, were also influenced by the CL features but in this case there are some disadvantages. For instance, the B-F requires a good benchmark to each accident year to be applicable (England et al. 2012). However, very often such benchmarks are not applicable or don't even exist. Despite this, there are also in the general insurance literature alternatives to overcome this problem; for example, someone can consult Taylor (1977, 1987, 2000), Barnett and Zehnwirth (1999), and Brydon and Verrall (2009).

In the present paper, a different approach is followed by combining together the Mack distribution-free model framework with a change on the way the loss development factors are estimated by implementing the regression through the origin approach of Murphy (1994). The target is to reduce the reserves' prediction error as it is defined by Mack (1993) for particular classes of data sets. As long as our approach is non-parametric, we won't refer to the numerous parametric[3] methods already known in the corresponding literature; see for instance the several alternative parametric methods to the Chain-Ladder presented in England and Verrall (2002).

1.3 THE VECTOR PROJECTION

In this subsection, we investigate that it might be possible to obtain much better results, if the development factors considered between the two columns are provided by the *Vector Projection* (VP) method. As it is known from Straub (1988), the CL is just an approximation to the least square solution[4]. The author shows that the loss development factor that gives us the minimum square of errors is given by a regression without intercept. The latter may be seen as a vector projection between two adjacent columns of our upper triangle. For example, in **Table 1**, this could be the column 2 projected using the column 1, for both columns with origin years 1 to 9.

The projection of the vector y onto the vector x is a new vector that corresponds to the y estimate, $\hat{y}$, see for example, Gentle (2007) for more details.

$$\hat{y} = \frac{<x,y>}{\|x\|^2} x, \tag{1.2}$$

where $< x, y >$ is the inner product of x and y and $\| x \|^2$ is the 2-norm of x. This means that we are getting a new vector, which is based on the previous one, y, but now projected in the x direction. This VP is a regression between two variables without intercept term and it is an

[3] It should be pointed out here that parametric models may help for a better understanding of the model and the development of its statistical properties. Another good example may be seen in Kuang, Nielsen and Nielsen (2009), where the importance of a good parameterisation is explained and maximum likelihood estimators of the canonical parameters are derived.

[4] Straub (1988) shows that the CL is just an approximation to the minimization of squares of residuals. He did it without having any assumptions in respect of the residuals.

alternative approach to estimate any loss development factor between two development years (where we have at least two observations).

Following Murphy (1994), and Barnett and Zehnwirth (1999), it is known that the *link ratios* approach for reserving may be also seen as a regression without intercept term for each of the loss development factors. For each observation used in the calculation, they show that we are going to have

$$C_{i,j+1} = b_j C_{i,j} + \varepsilon_{i,j}. \tag{1.3}$$

They assume that the variance of each residual $\varepsilon_{i,j}$ is given by a constant σ_j^2 weighted by the power δ on each $C_{i,j}$ observation,

$$Var(\varepsilon_{i,j}) = \sigma_j^2 C_{i,j}^{\delta}. \tag{1.4}$$

The b_j is estimated from the data and corresponds to our loss development factor[5]. Its value depends on the parameter δ and on the history of payments (if $\delta \neq 0$). Indeed if $\delta = 1$, we get the CL, i.e., a weighted average of the link ratios and if $\delta = 2$, the average link ratios method is derived, i.e., a simple average of the link ratios. Other methods are obtained if we have other $\delta \neq 0$, including also the subjective use of the link ratios.

These results have something in common, they always have *heteroscedastic*[6] errors when $\delta \neq 0$, that is the errors variance is not constant on each observation of the regression. Now, if $\delta = 0$ is assumed, the ordinary least squares regression without intercept is derived, i.e., a VP with constant variance σ_j^2 across all the observations of each regression. This result is different from the cases where $\delta \neq 0$, because with constant variance, the errors are homo-

[5] Indeed, it makes a good sense to estimate the loss development factors as a regression through the origin, i.e., through the VP method. The loss development factors may be seen as weighted averages of link ratios and the latter are the ratios of two cumulative payments from different development years. These ratios are estimated by the slope of the line that summarizes the relation between the two accumulated payments. Straub (1988) shows that the slope that minimizes the sum of squares of errors is not the CL, but the regression through the origin, i.e., the VP in our case.

[6] In Murphy (1994) and Barnett and Zehnwirth (1999), the general regression trough the origin by using assumptions to the residuals is developed and they demonstrated that several models might be contacted as a consequence of it. Obviously, some of the models are heteroscedastic, such as the simple average link ratios and the CL. In Murphy (1994), the VP method is introduced, but with homoscedastic errors. In the present paper, as it is clear in the text, the stochastic VP method of Murphy (1994) with heteroscedastic errors instead is introduced to compete with Mack (1993, 1994)'s stochastic heteroscedastic CL approach.

scedastic inside each equation for all the observations of the regression. However, this distinction between the cases of zero and non-zero deltas does not appear to be so strong, as it will be clearer in the next section (see Remark 2.2).

Now, in VP, the loss development factors b_j, with $j = 1, \dots, T-1$, are obtained by

$$\hat{b}_j^{VP} = \frac{\sum_{i=1}^{T-j} C_{i,j} C_{i,j+1}}{\sum_{i=1}^{T-j} C_{i,j}^2}, \tag{1.5}$$

and the cumulative payments may be obtained, like in the CL (Mack 1993, 1994), using the following relation

$$\hat{C}_{i,j}^{VP} = C_{i,j-1} \hat{b}_j^{VP}. \tag{1.6}$$

This means that in our example (see **Tables 2** and **3**), the loss development factor for the first column is 3.32, which is more consistent with the recent increase of the link ratios. Indeed, it is known that the use of the regression models has several advantages in claims reserving (Taylor 1978; Barnett and Zehnwirth 1999; Frees 2010).

Furthermore, let us consider the situation where we have a CL over a perfect upper triangle. The latter means that on each column the link ratios are always equal to the loss development factor. As a consequence of this, after estimating the lower triangle, we get a zero prediction error as it is defined in Mack (1993). A triangle like this can be the one presented in **Table 4**. In this theoretical case, the loss development factors from CL and VP are exactly the same and the prediction error is zero.

As we can easily verify, the VP gives the same results as the CL under some conditions. Indeed, this happens because the link ratios are totally stable on each development year. However, if this is not the case, the VP gives different results from the CL, and in some cases, it adapts better to the evolution of the link ratios to situations like the one presented in **Table 2**. However, as it will become clearer in the application part, we may have triangles where the VP approach may perform worse than the traditional CL.

Consequently, in this paper, we introduce an alternative to the distribution-free stochastic CL method of Mack (1993), the stochastic VP, using also the well-known regression through the origin approach proposed by Murray (1994), but with *heteroscedastic* errors instead following similar arguments with Mack (1993, 1994) approach that is the variance does not be constant over all the observations from each regression. Moreover, the eminent stochastic

Mack distribution-free framework may be further improved using this technique to some particular sets of data.

Table 4: Perfect Chain-Ladder Matrix of Cumulative Payments

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	1000	1800	3060	4896	7344	10282	13366	16039	17643	17643
2	1100	1980	3366	5386	8078	11310	14703	17643	19408	
3	1331	2396	4073	6517	9775	13685	17790	21348		
4	1772	3189	5421	8674	13010	18214	23679			
5	2594	4669	7937	12699	19048	26668				
6	4177	7519	12782	20452	30678					
7	7400	13320	22645	36232						
8	14421	25958	44128							
9	30913	55643								
10	72890									

2. A STOCHASTIC DISTRIBUTION-FREE VECTOR PROJECTION

In this section, the combined technique for estimating outstanding claims based on the VP methodology is proposed and developed in details. As our model is formulated on the Mack (1993, 1994) distribution-free model framework, we also assume that our VP has heteroscedastic errors inside each accident year, but the errors now are proportional to the square of payments. The latter is a consequence of the way we estimate the loss development factors, i.e., a weighted average of the link ratios with weights given by the square of payments. In the traditional CL, the weights are given by the payments.

Thus, the *cumulative payments* $C_{i,j}$ on each origin year are independent and there exist some *loss development* factors b_j such that $j = 1, \dots, T-1$, where we have for $1 \leq i \leq T$ and $1 \leq j \leq T$:

$$\mathbb{E}\left(C_{i,j+1} | C_{i,1}, \dots, C_{i,j}\right) = \mathbb{E}\left(C_{i,j+1} | C_{i,j}\right), \tag{2.1}$$

$$\mathbb{E}\left(C_{i,j+1} | C_{i,j}\right) = b_j C_{i,j}, \tag{2.2}$$

$$Var\left(C_{i,j+1} | C_{i,j}\right) = \sigma_j^2 C_{i,j}^2. \tag{2.3}$$

In this paper, the way that the loss development factors are estimated is changed using the VP approach instead of the typical CL estimator, i.e.,

$$\hat{b}_j^{VP} = \frac{\sum_{i=1}^{T-j} C_{i,j} C_{i,j+1}}{\sum_{i=1}^{T-j} C_{i,j}^2}. \quad (2.4)$$

Additionally, the parameter of the variance is estimated and it differs with the derived results of the Mack (1993, 1994) distribution-free model. This is due, to the way that the loss development factors are estimated, to the different weights considered, and to the need to have an unbiased estimator

$$\hat{\sigma}_j^{2,VP} = \frac{1}{T-j-1} \sum_{i=1}^{T-j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{b}_j^{VP} \right)^2. \quad (2.5)$$

The following result has initially been derived by Mack (1993), and it is also valid in our VP approach. In other words, Lemma 2.1 makes clear that Eq. (2.2), and the independency between the *cumulative payments* $C_{i,j}$ on each origin year are indeed implicit assumptions of the VP as well as the CL method.

Lemma 2.1 [Mack 1993] *Under the model assumption Eq.* (2.2) *and for* $D_u = \{C_{i,j}: i + j - 1 \leq T\}$, *a recursive algorithm is derived for the calculation of the ultimate cost on an origin year* i *based on the upper triangle information* D_u*:*

$$\mathbb{E}(C_{i,T} | D_u) = \mathbb{E}(C_{i,T} | C_{i,T-i+1}) = b_{T-1} \cdots b_{T-i+1} C_{i,T-i+1}. \quad (2.6)$$

■

The VP estimator of the unknown loss development factors is given by Eq. (2.3), and the ultimate cost estimator is

$$\hat{C}_{i,T}^{VP} = C_{i,T-i+1} \hat{b}_{T-i+1}^{VP} \cdots \hat{b}_{T-1}^{VP}. \quad (2.7)$$

Several properties from our VP estimator Eq. (2.4) are presented in the next Lemma:

Lemma 2.2 *For the VP estimators Eq. (2.4),* $\hat{b}_j^{VP}$ *for* $j = 1, \ldots, N$ *with* $N = T - 1$, *the following properties are derived.*

a) *They are unbiased.*
b) *Under the model assumption Eq.* (2.2), *they are uncorrelated.*
c) *Given* D_u, *the estimator of the ultimate costs is an unbiased estimator of the true value,* $\mathbb{E}(\hat{C}_{i,T}^{VP} | D_u) = \mathbb{E}(C_{i,T} | D_u)$.
d) *They are weighted average of the intermediate link ratios* $F_{i,j+1} = \frac{C_{i,j+1}}{C_{i,j}}$, *with the weights to be given by the square of the payments.*

$$\hat{b}_j^{VP} = \frac{\sum_{i=1}^{T-j} C_{i,j}^2 F_{i,j+1}}{\sum_{i=1}^{T-j} C_{i,j}^2}.$$

e) *The conditional variance of* $\hat{b}_j^{VP}$ *which has minimal condition variance among all unbiased linear combinations of the unbiased estimators* $\left(F_{i,j+1}\right)_{1\le i\le T-j}$ *for* b_j *conditional on* D_u *is given by*

$$Var\left(\hat{b}_j^{VP}\middle|D_u\right) = \left(\sum_{i=1}^{T-j} \frac{1}{\sigma_j^2}\right)^{-1}.$$

Similarly, the covariance is given by

$$Cov\left(\left(\frac{C_{i,j+1}}{C_{i,j}}, \hat{b}_j^{VP}\right)\middle|D_u\right) = \sigma_j^2 \frac{C_{i,j}^2}{\sum_{i=1}^{T-j} C_{i,j}^2}.$$

Proof. (a) It is almost straightforward to show that $\mathbb{E}\left(\hat{b}_j^{VP}|D_u\right) = \hat{b}_j^{VP}$, *for* $j = 1, \dots, N$ *with* $N = T - 1$, i.e.,

$$\mathbb{E}\left(\hat{b}_j^{VP}|D_u\right) = \frac{\sum_{i=1}^{T-j} C_{i,j}\mathbb{E}(C_{i,j+1}|D_u)}{\sum_{i=1}^{T-j} C_{i,j}^2} = \hat{b}_j^{VP}.$$

(b) The proof is similar to the one obtained by Mack (1993), so it is omitted. A proof can also be found in Wüthrich and Merz (2008).

(c) Let's have

$$\begin{aligned}\mathbb{E}\left(\hat{C}_{i,T}^{VP}|C_{i,T-i+1}\right) &= \mathbb{E}\left(C_{i,T-i+1}\hat{b}_{T-i+1}^{VP}\cdots\hat{b}_{T-1}^{VP}|C_{i,T-i+1}\right)\\ &= b_{T-1}\mathbb{E}\left(C_{i,T-i+1}\hat{b}_{T-i+1}^{VP}\cdots\hat{b}_{T-2}^{VP}|C_{i,T-i+1}\right) = b_{T-1}\mathbb{E}\left(\hat{C}_{i,T-1}^{VP}|C_{i,T-i+1}\right).\end{aligned}$$

This means that $\mathbb{E}\left(\hat{C}_{i,T-1}^{VP}|C_{i,T-i+1}\right) = b_{T-2}\mathbb{E}\left(\hat{C}_{i,T-2}^{VP}|C_{i,T-i+1}\right)$.

Continuing the iteration, we get

$$\mathbb{E}\left(\hat{C}_{i,T}^{VP}|C_{i,T-i+1}\right) = b_{T-1}b_{T-2}\cdots b_1\mathbb{E}\left(\hat{C}_{i,1}^{VP}|C_{i,T-i+1}\right) = \mathbb{E}\left(C_{i,T}|D_u\right),$$

As the latter doesn't depend on the loss development factor calculation, the result is also similar to the one obtained by the Mack (1993, 1994) distribution-free model. A proof may also be found in Wüthrich and Merz (2008).

(d) $\hat{b}_j^{VP} = \frac{\sum_{i=1}^{T-j} C_{i,j}C_{i,j+1}}{\sum_{i=1}^{T-j} C_{i,j}^2} = \frac{\sum_{i=1}^{T-j} C_{i,j}^2 F_{i,j+1}}{\sum_{i=1}^{T-j} C_{i,j}^2}$.

(e) Considering the two Lemmas 3.3 and 3.4 in Wüthrich and Merz (2008), and the fact that

$$Var\left(F_{i,j+1}\middle|D_u\right) = Var\left(F_{i,j+1}\middle|C_{i,j}\right) = \sigma_j^2,$$

then

$$Var\left(\hat{b}_j^{VP}\middle|D_u\right) = \left(\sum_{i=1}^{T-j}\frac{1}{\sigma_j^2}\right)^{-1} = \frac{\sigma_j^2}{T-j}$$

and the covariance is given by

$$Cov\left(\left(\frac{C_{i,j+1}}{C_{i,j}}, \hat{b}_j^{VP}\right)\middle|D_u\right) = \frac{C_{i,j}^2}{\sum_{i=1}^{T-j} C_{i,j}^2} Var\left(\frac{C_{i,j+1}}{C_{i,j}}\middle|D_u\right) = \sigma_j^2 \frac{C_{i,j}^2}{\sum_{i=1}^{T-j} C_{i,j}^2}.$$

■

Remark 2.1 Based on the Gauss-Markov theorem, the VP *loss development factors*, $\hat{b}_j^{VP}$, are the best *linear unbiased estimator*; see Fomby et al. (1984). In more details, the estimators are *unbiased*, *linear* on the observations (i.e., a weighted average on the link ratios), and the loss development factor's variances and covariance are a function of σ^2, and the past observations, thus, $\hat{b}_j^{VP}$ are the ones with a lower variance.

Remark 2.2 Now, the loss development factors in Eq. (2.4) can also be given alternatively. Let us consider the general Murphy (1994) case given by Eq. (1.4), i.e., $Var\left(C_{i,j+1}|C_{i,j}\right) = \sigma_j^2 C_{i,j}^{\delta}$. The best linear unbiased estimate of $b_j = \mathbb{E}\left(\left(C_{i,j+1}/C_{i,j}\right)|C_{i,j}\right)$, which is linear in terms of the $C_{i,j+1}/C_{i,j}$ is therefore $\hat{b}_j^{Mu} = \sum_i w_{ij}\left(C_{i,j+1}/C_{i,j}\right)$ with weights inversely proportional to variance: $w_{ij} = C_{i,j}^{2-\delta}/\sum_i C_{i,j}^{2-\delta}$, and where we shall take summation limits to be obvious through. Then, the estimator is given by $\hat{b}_j^{Mu} = \frac{\sum_i C_{i,j}^{1-\delta} C_{i,j+1}}{\sum_i C_{i,j}^{2-\delta}} = \frac{\sum_i C_{i,j}^{1-\frac{1}{2}\delta}\left(C_{i,j+1}/C_{i,j}^{\frac{1}{2}\delta}\right)}{\sum_i\left(C_{i,j}^{1-\frac{1}{2}\delta}\right)^2}$. So, this estimator may be expressed as a projection coefficient: $\hat{b}_j^{Mu} = \frac{<V_j, U_{j+1}>}{\|V_j\|^2}$, the projection of U_{j+1} on V_j , where V_j is the vector whose components are $C_{i,j+1}/C_{i,j}^{\frac{1}{2}\delta}$. If one sets $\delta = 0$, the VP estimator is derived, and equivalently, $\hat{b}_j^{VP} \equiv \hat{b}_j^{Mu}$.

Lemma 2.3 *Under the model assumptions the estimator which is given by Eq.* (2.5) *is an unbiased estimator of the true parameter for* σ_j^2.

Proof. Let's calculate $\mathbb{E}\left(\hat{\sigma}_j^{2,VP}|D_u\right)$,

$$\mathbb{E}\left(\hat{\sigma}_j^{2,VP}|D_u\right) = \frac{1}{T-j-1}\sum_{i=1}^{T-j}\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{b}_j^{VP}\right)^2|D_u\right].$$

Then, the expected value is provided. It may be decomposed as follows

$$\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{b}_j^{VP}\right)^2|D_u\right]$$
$$= \mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - b_j\right)^2|D_u\right] + 2\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - b_j\right)\left(b_j - \hat{b}_j^{VP}\right)|D_u\right]$$
$$+ \mathbb{E}\left[\left(\hat{b}_j^{VP} - b_j\right)^2|D_u\right].$$

Developing the first component on the right side,

$$\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - b_j\right)^2|D_u\right] = Var\left(\frac{C_{i,j+1}}{C_{i,j}}\middle|D_u\right) = \sigma_j^2.$$

Using Lemma 2.2, the second component is derived

$$\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - b_j\right)\left(\hat{b}_j^{VP} - b_j\right)|D_u\right] = Cov\left(\left(\frac{C_{i,j+1}}{C_{i,j}}, \hat{b}_j^{VP}\right)\middle|D_u\right) = \sigma_j^2\frac{C_{i,j}^2}{\sum_{i=1}^{T-j}C_{i,j}^2}.$$

This means that

$$-2\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - b_j\right)\left(\hat{b}_j^{VP} - b_j\right)|D_u\right] = -2\sigma_j^2\frac{C_{i,j}^2}{\sum_{i=1}^{T-j}C_{i,j}^2}.$$

Because for the last component, we also get from Lemma 2.2 that

$$\mathbb{E}\left[\left(\hat{b}_j^{VP} - b_j\right)^2|D_u\right] = Var\left(\hat{b}_j^{VP}|D_u\right) = \frac{\sigma_j^2}{T-j}.$$

Adding all the three components together, we get

$$\mathbb{E}\left(\hat{\sigma}_j^{2,VP}\right) = \frac{1}{T-j-1}\sum_{i=1}^{T-j}\left(\sigma_j^2 - 2\sigma_j^2\frac{C_{i,j}^2}{\sum_{i=1}^{T-j}C_{i,j}^2} + \frac{\sigma_j^2}{T-j}\right) =$$

$$\frac{1}{T-j-1}\sigma_j^2(T-j-2+1) = \sigma_j^2.$$ ∎

With the following results, the calculation of the *mean squared error* (mse) is provided. First, the necessary Lemma for the connection between the *mse* of the estimated reserves and claims is given.

Lemma 2.4 [Mack, 1993] *The mse of the estimated reserves and claims is equal.* ■

Theorem 2.1 *Under the assumptions* (2.1), (2.2) *and* (2.3), *where all the origin years are independent and there are unbiased estimators for the loss development factor and the variance parameter, the mean squared error for each origin year reserve, $mse(\hat{R}_i)$, can be estimated by*

$$\widehat{mse(\hat{R}_i)} = \hat{C}_{i,T}^{2,VP} \sum_{j=T-i+1}^{T-1} \frac{\hat{\sigma}_j^{2,VP}}{\hat{b}_j^{2,VP}} \left(\frac{1}{\hat{C}_{i,j}^{VP^2}} + \frac{1}{T-j} \right), \tag{2.8}$$

where $\hat{\sigma}_j^{2,VP}$, $\hat{b}_j^{2,VP}$ and $\hat{C}_{i,T}^{VP}$ are given by Eqs. (2.4), (2.5) and (2.7).

Proof. Let's start with

$$mse(\hat{R}_i) = Var(C_{i,T}|D_u) + [\mathbb{E}(C_{i,T}|D_u) - \hat{C}_{i,T}^{VP}]^2.$$

As the $Var(C_{i,T-i+1}) = 0$ and $\sigma_{T-i+1}^2 = 0$, the development of the first component, $Var(C_{i,T}|D_u)$ is given by (using the estimators of the loss development factors)

$$\begin{aligned} Var(C_{i,T}|D) &= \mathbb{E}_i\left(Var(C_{i,T}|C_{i,1}, \dots, C_{i,T-1})\right) + Var_i\left(\mathbb{E}(C_{i,T}|C_{i,1}, \dots, C_{i,T-1})\right) = \mathbb{E}_i(C_{i,T-1}^2) \\ &\quad + \hat{b}_{T-1}^{2,VP} Var_i(C_{i,T-1}) \\ &= [\sigma_{T-2}^2 \sigma_{T-1}^2 + \hat{b}_{T-2}^{2,VP} \sigma_{T-1}^2 + \hat{b}_{T-1}^{2,VP} \sigma_{T-2}^2] \mathbb{E}_i(C_{i,T-2}^2) + \hat{b}_{T-1}^{2,VP} \hat{b}_{T-2}^{2,VP} Var_i(C_{i,T-2}) \\ &= C_{i,T-r+1}^2 \sum_{k=T-i+1}^{T-1} \hat{b}_{T-i+1}^{2,VP} \cdots \hat{b}_{k-1}^{2,VP} \sigma_k^2 \hat{b}_{k+1}^{2,VP} \cdots \hat{b}_{T-1}^{2,VP} \\ &= \hat{C}_{i,T}^{2,VP} \sum_{j=T-i+1}^{T-1} \frac{\hat{\sigma}_j^{2,VP}}{\hat{b}_j^{2,VP}} \frac{1}{\hat{C}_{i,j}^{VP^2}}. \end{aligned}$$

Moreover, considering the vector projection approach and the Mack (1993)'s proof, the second component is given by

$$[\mathbb{E}(C_{i,T}|D_u) - \hat{C}_{i,j}^{VP}]^2 = C_{T-i+1}^2 (b_{T-i+1} \cdots b_{T-1} - \hat{b}_{T-i+1}^{VP} \cdots \hat{b}_{T-1}^{VP})^2 = C_{T-i+1}^2 F^2,$$

with

$$F = b_{T-i+1} \dots b_{T-1} - \hat{b}_{T-i+1}^{VP} \cdots \hat{b}_{T-1}^{VP} = S_{T-i+1} + \cdots + S_{T-1},$$

and

$$S_j = \hat{b}_{T-i+1}^{VP} \cdots \hat{b}_{j-1}^{VP} (b_j - \hat{b}_j^{VP}) b_{j+1} \cdots b_{T-1}.$$

Hence

$$F^2 = \sum_{j=T-i+1}^{T-j} S_j^2 + 2\sum_{i<j} S_i S_j = \mathbb{E}\left(S_j^2 \middle| D_u\right) + 2\mathbb{E}\left(S_i S_j \middle| D_u\right).$$

As the estimator for the loss development factor is unbiased, we have that $\mathbb{E}\left(S_i S_j \middle| D_u\right) = 0$. Consequently, we get $\mathbb{E}\left(S_j^2 \middle| D_u\right)$, thus we just need to remember the variance of the estimator, i.e.,

$$\mathbb{E}\left(\left(b_j - \hat{b}_j^{VP}\right)^2 \middle| D_u\right) = Var\left(\hat{b}_j^{VP} \middle| D_u\right) = \frac{\sigma_j^2}{T-j}.$$

We get now

$$\mathbb{E}\left(S_j^2 \middle| D_u\right) = \frac{\hat{b}_{T-i+1}^{2,VP} \cdots \hat{b}_{j-1}^{2,VP} \sigma_j^2 b_{j+1}^2 \cdots b_{T-1}^2}{T-j}.$$

Using $F^2 = \sum_{j=T-i+1}^{T-j} S_j^2$, the estimators of the loss development factors and of the variance parameter, we get

$$\left[\mathbb{E}\left(C_{i,T} \middle| D_u\right) - \hat{C}_{i,j}^{VP}\right]^2 = \hat{C}_{i,T}^{2,VP} \sum_{j=T-i+1}^{T-1} \frac{\hat{\sigma}_j^{2,VP}}{\hat{b}_j^{2,VP}} \left(\frac{1}{T-j}\right).$$

And finally, we have

$$mse\left(\hat{R}_i\right) = \hat{C}_{i,T}^{2,VP} \sum_{j=T-i+1}^{T-1} \frac{\hat{\sigma}_j^{2,VP}}{\hat{b}_j^{2,VP}} \left(\frac{1}{\hat{C}_{i,j}^{VP\,2}} + \frac{1}{T-j}\right).$$

This leads to the estimator stated in the Theorem. ■

From Theorem 2.1, we observe that the mean squared error given by Eq. (2.8) for each accident year is similar to the Mack (1993)'s CL prediction error, i.e., it depends on the level of square of estimated ultimate claims, and on the sum of estimator of variance for each development factor weighted by the estimator of development factor and a multiplicative factor. However, there are also some differences. Obviously, the estimators are obtained through the VP instead of the CL method, and the multiplicative factor has now two different components. The first component is the inverse of square of estimated payments. In Mack (1993, 1994)'s model, this component is the inverse of estimated payments. The reasoning behind this difference lies on the new assumption (2.3) of the VP model, where the conditional variance of payments depends on the square[7] of payments. The second component is totally dif-

[7]In the Mack (1993, 1994)'s model, it depends on the payments.

ferent between the VP and the CL models. In the former, it doesn't depend on the inverse of sum of payments, like in the CL does, but just on the inverse of number of years to develop each accident year, i.e., $1/(T-j)$. This is due to the conditional variance of VP loss development factor which is given by the product of the σ_j^2 by this factor. In the CL method, this factor is the inverse of sum of future payments. The next Corollary completes the theoretical findings.

Corollary 2.1 *With the assumptions and notations of Theorem 2.1, the mse of the overall reserve estimate,* $\hat{R} = \hat{R}_2 + \hat{R}_3 + \cdots + \hat{R}_T$ *can be given by*

$$msep(\hat{R}) = \sum_{i=2}^{T} \left\{ mse(\hat{R}_i) + \hat{C}_{i,T}^{VP} \left(\sum_{j=i+1}^{T} C_{j,T} \right) \sum_{j=T-i+1}^{T-1} \frac{2\hat{\sigma}_j^{2,VP} / \hat{b}_j^{2,VP}}{T-j} \right\}. \quad (2.9)$$

Proof. The result comes immediately following Mack (1993) using

$$\mathbb{E}\left(S_j^2 \middle| D_u\right) = \frac{\hat{b}_{T-i+1}^{2,VP} \cdots \hat{b}_{j-1}^{2,VP} \sigma_j^2 b_{j+1}^2 \cdots b_{T-1}^2}{T-j}. \quad \blacksquare$$

As it was before, the mean square error of prediction structure for the overall reserve is similar to the Mack CL prediction error. However, the estimators of payments, the mean square error of prediction, and the ratio of variance to the square of loss development factors are obtained by using the VP results. Additionally, as it was the case above, the last component is totally different between the VP and CL models. In the former, it doesn't depend on the inverse of sum of payments, like in the CL does, but just on the inverse of number of years to develop each accident year, i.e., $1/(T-j)$.

3. Numerical Examples

3.1. Irregular Development of Data

In this section, data from the **Table 1** with cumulative payments $C_{i,j}$ are used to illustrate the comparison between the two claims reserving methodologies. The particular data set is used by several authors, including Mack (1993, 1994), and England and Verrall (2002).

Indeed, using **Table 1**, which has very irregular (extreme) development of data, we can observe that both the VP and CL have high prediction errors, and the former is not able to decrease the latter prediction error (see **Tables 5** and **6**). The prediction errors are, 52% for the CL and 63% for the VP.

Table 5: Stochastic Vector Projection with Irregular Data

Development Year	Loss Development Factors	Variance
2	2,217	192,637
3	1,569	0,243
4	1,261	0,104
5	1,162	0,005
6	1,100	0,007
7	1,041	0,003
8	1,032	0,000
9	1,016	0,000
10	1,009	0,000

Origin Year	Estimated Reserves	Prediction Error
2	154	97%
3	593	71%
4	1 577	33%
5	2 648	33%
6	3 344	26%
7	5 013	18%
8	10 151	25%
9	9 623	24%
10	10 670	250%
Total	**43 772**	**63%**

Table 6: Mack (1993, 1994) Stochastic Distribution-Free Model with Irregular Data

Development Year	Loss Development Factors	Variance
2	2,999	27 883,479
3	1,624	1 108,526
4	1,271	691,443
5	1,172	61,230
6	1,113	119,439
7	1,042	40,820
8	1,033	1,343
9	1,017	7,883
10	1,009	1,343

Origin Year	Estimated Reserves	Prediction Error
2	154	134%
3	617	101%
4	1 636	46%
5	2 747	53%
6	3 649	55%
7	5 435	41%
8	10 907	49%
9	10 650	59%
10	16 339	150%
Total	**52 135**	**52%**

Due to the strong irregular development of data, we don't have a good fit for both methods and the VP doesn't improve the CL results. However, the CL obliges 19% more of reserves than the VP does. As the VP weights with the square of payments and the latter are higher for older years the fit is not so good for the most recent years, as the behaviour seems to be different. Mack (1993, 1994)'s data on **Table 1** has abnormally high link ratios on accident years 2 and 7 for the first development year. The VP gives a very high weight to these link ratios (due to the use of the square of the payments as weights), something that doesn't happen so much with the CL (as the payments are just used as weights). Practically, this means that the CL smooths more the effect of these two link ratios outliers and the variance of loss development factors and the variance parameter are lower, producing a lower prediction error. In the following subsections regular development of data are used.

3.2 Regular Development of Data

Example 1: Data from Taylor and Ashe (1983)

In this subsection, we consider now a different set of data also used by Mack (1993, 1994), and originally from Taylor and Ashe (1983). It's has an increasing (regular) trend comparing with the previous one and more adapted to the CL environment.

With this more *regular* triangle both the VP and the CL have lower prediction errors. However, the VP presents a smaller prediction error than the CL, i.e., 9% for the former and 13% for the latter, see **Tables 7** and **8**.

Table 7: Stochastic Vector Projection with Regular Data from Example 1

Parameters	
Loss Development Factors	**Variance**
3,418	0,472
1,749	0,029
1,462	0,020
1,167	0,005
1,097	0,005
1,087	0,002
1,055	0,000
1,078	0,000
1,018	0,000

Origin Year	Estimated Reserves	Prediction Error
2	94 634	63%
3	478 103	18%
4	723 104	12%
5	1 002 041	13%
6	1 408 034	14%
7	2 131 332	12%
8	3 885 296	10%
9	4 255 237	9%
10	4 501 720	10%
Total	**18 479 500**	**9%**

Table 8: Mack (1993, 1994) Stochastic Distribution-Free Model with Regular Data from Example 1

Parameters	
Loss Development Factors	**Variance**
3,491	160280,327
1,747	37736,855
1,457	41965,213
1,174	15182,903
1,104	13731,324
1,086	8185,772
1,054	446,617
1,077	1147,366
1,018	446,617

Origin Year	Estimated Reserves	Prediction Error
2	94 634	80%
3	469 511	26%
4	709 638	19%
5	984 889	27%
6	1 419 459	29%
7	2 177 641	26%
8	3 920 301	22%
9	4 278 972	23%
10	4 625 811	29%
Total	**18 680 856**	**13%**

We may also report that the difference of estimated reserves, between the VP and the CL, is just of -1%. The stochastic estimators are also important because they allow us to have the best estimate added of a risk margin[8]. For a certain degree of confidence, for example 99.5%, the risk margin is the value that is added to the reserve best estimate and it gives us maximum reserve level on 99.5% of the samples we might produce. Following Mack' (1993, 1994) suggestion of having normal distribution confidence intervals when the coefficient of variation of the reserves is lower than 15%, when we calculate the percentile 99.5% of reserves to the VP than CL. Due to the low prediction error from the VP, its stochastic reserve is much lower than the CL, around -10% for 99.5% confidence level. Thus, analytically, we get 22 624 853 to the VP and 24 984 154 money-units to the CL.

Example 2: Data from Taylor and McGuire (2016)

We consider now a different set of very regular data used recently by Taylor and McGuire (2016). It is even more regular than the previous one. With the new set of more regular development of data, the VP and the CL have lower prediction errors. However, the VP presents a smaller prediction error than the CL, 1.3% for the former and 2.9% for the latter, see **Tables 9** and **10**.

Table 9: Stochastic Vector Projection with Regular Data from Example 2

Development Year	Loss Development Factors	Variance
2	1,812	0,008
3	1,260	0,000
4	1,158	0,000
5	1,088	0,000
6	1,056	0,000
7	1,039	0,000
8	1,030	0,000
9	1,025	0,000
10	1,021	0,000

[8] The risk margin is important not only to give us a measure of uncertainty of our estimate, but also because it's one of the components of claims reserving on the new Solvency II regime when internal models are considered. The Fair Value of reserves is the sum of best estimate with a risk margin.

Origin Year	Estimated Reserves	Prediction Error
2	3 398	0,0%
3	8 155	0,1%
4	14 608	1,6%
5	22 719	1,8%
6	32 025	2,0%
7	45 870	1,9%
8	60 175	1,5%
9	80 926	1,4%
10	105 594	2,5%
Total	**373 469**	**1,3%**

Table 10: Mack (1993, 1994) Stochastic Distribution-Free Model with Regular Data from Example 2

Development Year	Loss Development Factors	Variance
2	1,815	449,408
3	1,261	22,347
4	1,158	8,575
5	1,088	7,547
6	1,055	4,294
7	1,039	1,887
8	1,030	0,576
9	1,025	0,001
10	1,021	0,000

Origin Year	Estimated Reserves	Prediction Error
2	3 398	0,0%
3	8 155	0,2%
4	14 579	2,8%
5	22 645	3,7%
6	31 865	4,3%
7	45 753	4,3%
8	60 093	3,8%
9	80 983	3,9%
10	105 874	8,7%
Total	**373 346**	**2,9%**

We may also see that the difference of estimated reserves, between the VP and the CL, is also very small, but due to the low prediction error from the VP, its stochastic reserve will also be lower than the one from CL about -4%.

Remark 3.1 It should be mentioned here that the prediction error is not the only quantitative criterion to follow on analysing the triangle's results. Other items should also be addressed, such as the retrospective residuals and the back-testing, when these models are considered. A recent and good example of this may be seen in Taylor and McGuire (2016). However, further discussion on alternative quantitative criteria is out of the scope of this paper. A following research paper it might be more appropriate.

3.3 Use Test with 114 triangles

In this subsection, in order to interpret more the concluded remarks derived by the previously described three numerical examples, which are standard in the reserving literature, and to provide also a "business orientated" analysis, we select 114[9] triangles randomly[10] with paid claims and 10 years of information to be comparable directly to the previous cases. **Table 11** reports the derived results. Additionally, **Figure 1** provides a comparison between the prediction errors calculated based on CL and VP methods, respectively. Among the 114 triangles studied, the VP has a lower prediction error and lower reserve estimation on 65 cases. In the 32 other cases, despite the lower prediction error, the reserve estimation of the VP is higher. The CL has just a lower prediction error in 17 cases and in 16 of them produced a higher level of reserves. Thus, based on those data, we can conclude that the VP has a lower prediction error in 85% (97 out of 114) of those cases. The VP reserves are lower on 71% (81 out of 114) of the cases, with an average reduction on reserves of 2%.

Table 11: Summary of Results of the Use Test

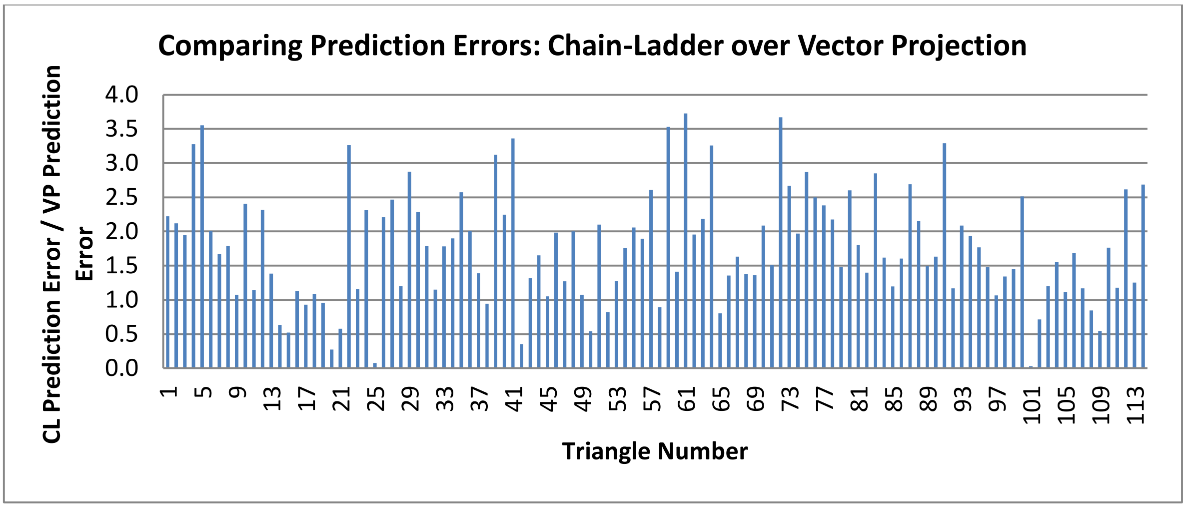
	VP Prediction Error Lower	CL Prediction Error Lower	Total
VP Reserves Lower	65	16	81
CL Reserves Lower	32	1	33
Total	97	17	114

[9] The data set used is from the consulting company, Actuarial Group, Lisbon Portugal. Obviously, it's not possible to disclose any further information about the triangles used and their orientation. **Table 11** and **Figure 1** are for illustration properties and useful for our "business orientated" analysis. Definitely, the interesting readers and particularly the practitioners can use their own data to evaluate and to reconfirm our findings.

[10] Practically, we mean that the dataset is chosen by using different companies and for different periods of businesses.

Moreover, we should also report that the features of the triangles with the lower CL prediction errors are as follows. In most of these 17 cases, there are some special situations, such as cells with zeros (five triangles) or cells with negative accumulated payments (nine triangles). The other three triangles appear to share something similar, i.e., the payments increase with the accident year until a certain point and then start decreasing.

Figure 1: Chain Ladder Prediction Error / Vector Projection Prediction Error



4. Conclusions

The assessment of financial strength in the insurance industry includes a thorough analysis of outstanding claims reserves, including an assessment of possible variability in the reserves. Failure to do so might result in the insolvency or lack of competitiveness of some insurers (England and Verrall 2002). Methods of analysis, which help with the reserve estimation as well as provide insight into the variability of those reserves, are always very welcomed, particularly if they are able to reduce the prediction errors. Let alone if they are also flexible to control the prediction errors level.

In this paper, the stochastic Vector Projection methodology has been proposed using the regression through the origin approach of Murphy (1994), but with heteroscedastic errors instead, in order to develop it comparably with the Mack (1993, 1994) stochastic distribution-free framework and test it empirically with the Chain-Ladder method. Interestingly, the equation for the loss development factors which is formulated on the Mack (1993, 1994) heteroscedastic errors framework, although, it is assumed that the errors are proportional to the square of payments instead, see Eq. (2.3); it can also be derived straightforwardly from the Murphy (1994) homoscedastic errors framework; see Eq. (1.4), and Remark 2.2.

Obviously, the prediction error is not the only measure to have when the claims reserves are estimated, additionally, other items should be also addressed, such as the retrospective residuals, the back-testing and so on and so forth. However, in real world applications in the insurance industry, it is almost impossible to tolerate a model which might have a higher or even very high prediction error. Due to the stability assumption of the Chain-Ladder model, this is often the case with it.

Finally, the three empirical examples have been applied. We observe that when the data has irregular developments both the Chain-Ladder and Vector Projection approaches generate high prediction errors, and thus, they cannot be considered as the best approaches to handle with this class of data. Additionally, we show that the Vector Projection on that environment is not able to outperform the Chain-Ladder. On the other hand, however, when more regular data is considered, like in Examples 1 and 2, the prediction error for both methods is improved, and the Vector Projection outperforms the Chain-Ladder. In these regular cases, the risk margins of the Vector Projection are also lower comparing with those derive from the Chain Ladder. Practically, this also implies a lower fair value of reserves for the Vector Projection method. The results are tested and confirmed by using 114 triangles with paid claims and 10 years of information to be comparable directly to the previous cases, where 85% of them appear to give a lower prediction error when the Vector Projection method is used. As a natural continuation of our paper, different to prediction error measurements should be used, and the multivariate approach as well as the mixture of a payments triangle with extra information coming from an incurred claims triangle will be also considered in a forthcoming paper.

References

Barnett, G. and Zehnwirth, B. 1999. *Best Estimates for Reserves*, CAS, 1 – 54.

Brydon, D. and Verrall, R. 2009. *Calendar Year Effects, Claims Inflation and the Chain-Ladder Technique*, A.A.S. 4, II, 287-301

England, P. and Verrall, R. 2002. Stochastic Claims Reserving in General Insurance, *British Actuarial Journal*, **8** (III), 443 – 544.

England, P., Verrall, R. and Wüthrich, M. 2012. Bayesian Overdispersed Poisson Model and the Bornhuetter-Ferguson Claim Reserving Method, *Annals of Actuarial Science,* **6**(2), 258 – 283.

European Community 1999. *Solvency Margin Review*, DIV 9049 (06/99).

Frees, E. 2010. *Regression Modelling with Actuarial and Financial Applications*, Cambridge University Press.

Fomby, T., Hill, R. and Johnson, S. 1984. *Advanced Econometric Methods,* 2nd Edition, Springer, New York.

Gentle, J. 2007. *Matrix Algebra - Theory, Computations, and Applications in Statistics*, Springer, New York.

Kuang, D., Nielsen, B. and Nielsen, J.P. 2009. Chain-Ladder as Maximum Likelihood Revisited, *Annals of Actuarial Science*, **4**(1): 105-121.

Mack, T. 1993. Distribution-Free Calculation of the Standard Error of the Chain Ladder Method Reserves Estimates, *ASTIN Bulletin*, **23**(2), 213 – 225.

Mack, T. 1994. Which Stochastic Model is Underlying the Chain Ladder Method, *Insurance Mathematics and Economics* **15**(2/3), 133 – 138.

Mack, T. 1999. The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor, *ASTIN Bulletin*, **29**(2), 361 – 366

Miranda, M, Nielsen, J. and Verrall, R. 2012. Double Chain Ladder, *ASTIN Bulletin,* **42**(1), 59 – 76.

Murphy, D. 1994. *Unbiased Loss Development Factors,* Casualty Actuarial Society Proceedings, 183 – 246.

Parodi, P. 2014. Triangle-free reserving A non-traditional framework for estimating reserves and reserve uncertainty Abstract of the London discussion, *British Actuarial Journal*, **19**(1), 219 – 233.

Pooser, D.M. and Walker P.L. 2015. Own Risk and Solvency Assessment: Origins and Implications for Enterprise Risk Management, *Journal of Insurance Regulation*, **34**(9), 1-19.

Renshaw, A., 1989. *Chain-Ladder and Interactive Modelling*, Journal of the Institute of Actuaries, **116**, 559 – 587.

Straub, E. 1988. *Non-Life Insurance Mathematics*, Springer-Verlag, Berlin.

Taylor, G. and Ashe, F. 1983. Second Moments of Estimates of Outstanding Claims, *Journal of Econometrics*, **23**, 37-61.

Taylor, G. 1977. Separation of Inflation and other Effects from the Distribution of Non-Life Insurance Claims Delays, *ASTIN Bulletin*, **9**, 217-230.

Taylor, G. 1987. *Regression Models in Claims Analysis I: Theory,* Casualty Actuarial Society, 353 – 384.

Taylor, G. 2000. *Loss Reserving: an Actuarial Perspective*, Huebner International Series on Risk, Insurance and Economic Security, Vol. 21, Springer New York, USA.

Taylor, G. and McGuire, G. 2016. *Stochastic Loss Reserving using Generalized Linear Models,* Ed. Casualty Actuarial Society

Verrall, R. 1989. A State Space Representation of the Chain Ladder Model, *Journal of the Institute of Actuaries*, **116**, 589 – 609.

Verrall, R. 1990. Bayes and Empirical Bayes Estimation of the Chain Ladder Model, *ASTIN Bulletin*, **20**, 217 – 243.

Verrall, R. 1991a. On the Estimation of Reserves from Loglinear Models, *Insurance Mathematics and Economics*, **10**, 75 – 80.

Verrall, R. 1991b. Chain Ladder and Maximum Likelihood, *Journal of the Institute of Actuaries*, **118**, 489 – 499.

Verrall, R. 2000. An Investigation into Stochastic Claims Reserving Models and the Chain-Ladder Technique, *Insurance Mathematics and Economics*, **26**, 91 – 99.

Wright, T.S. 1990. A Stochastic Method for Claims Reserving in General Insurance, *Journal of the Institute of Actuaries*, **117**, 677 – 731.

Wüthrich, M. and Merz, M. 2008. *Stochastic Claims Reserving Methods in Insurance*, Wiley.

Wüthrich, M. and Merz, M. 2014. *Modified Munich Chain Ladder Method*, Swiss Finance Institute Research Paper nº 14-65.

Zhang, Y. 2010. A General Multivariate Chain-Ladder, *Insurance Mathematics and Economics*, **46**, 588 – 599.