ORIGINAL SCIENTIFIC PAPER
UDK: 658.149.3:368.1(497.11)"2010/2019"
005.52:005.33
DOI: 10.5937/EKOPRE2106306V
Date of Receipt: March 17, 2021

Vladimir Vasić
University of Belgrade
Faculty of Economics
Department of Statistics and Mathematics

Jelena Kočović
University of Belgrade
Faculty of Economics
Department of Statistics and Mathematics

Marija Koprivica
University of Belgrade
Faculty of Economics
Department of Statistics and Mathematics

PRINCIPAL COMPONENT ANALYSIS IN DETERMINING REPRESENTATIVE FINANCIAL RATIOS WITHIN NON-LIFE INSURANCE SECTOR IN SERBIA

Analiza glavnih komponenata u određivanju reprezentativnih finansijskih pokazatelja u sektoru neživotnih osiguranja u Srbiji

Abstract

The paper deals with the application of principal component analysis in determining financial ratios that are representative within non-life insurance sector. Starting from many financial indicators found in the literature in the field of insurance, the purpose of the study is to identify a smaller set of ratios that are most relevant for assessing the financial position and performance of non-life insurance companies in Serbia with a minimum loss of information. On the basis of financial reports of non-life and composite insurers in the period 2010-2019, we calculated 38 financial ratios, grouped into seven categories (capital adequacy, asset quality, reinsurance risk and performance, adequacy of technical reserves, profitability, liquidity and management soundness). Using parallel analysis and Velicer's minimum average partial test, we found that it is possible to explain 85% of variability of the initial set of ratios with six financial ratios. The obtained results can be used for the purposes of efficient financial analysis of individual insurance companies and the entire non-life insurance sector in Serbia.

Keywords: *non-life insurance, financial ratios, principal component analysis, Velicer's minimum average partial test, parallel analysis, principal component loadings.*

Sažetak

Predmet rada je primena analize glavnih komponenata u određivanju reprezentativnih finansijskih pokazatelja u sektoru neživotnih osiguranja. Cilj istraživanja je da se, polazeći od mnoštva finansijskih pokazatelja koji se susreću u literaturi u oblasti osiguranja, identifikuje manji skup pokazatelja koji su najrelevantniji za ocenu finansijskog položaja i performansi kompanija koje se bave poslovima neživotnih osiguranja u Srbiji, uz minimalan gubitak informacija. Na osnovu finansijskih izveštaja neživotnih i kompozitnih osiguravača tokom perioda 2010-2019. godine, izračunato je 38 finansijskih pokazatelja, koji su razvrstani u sedam kategorija (adekvatnost kapitala, kvalitet imovine, rizik i performanse reosiguranja, adekvatnost tehničkih rezervi, profitabilnost, likvidnost i kvalitet menadžmenta). Primenom paralelne analize i Velicerovog minimalnog prosečnog delimičnog testa, utvrđeno je da je sa svega šest finansijskih pokazatelja moguće objasniti 85% varijabiliteta inicijalnog skupa pokazatelja. Dobijeni rezultati mogu biti korišćeni u svrhe efikasne finansijske analize pojedinačnih osiguravajućih kompanija i celokupnog sektora neživotnih osiguranja u Srbiji.

Ključne reči: *neživotno osiguranje, finansijski pokazatelji, analiza glavnih komponenata, Velicerov minimalni prosečni delimični test, paralelna analiza, opterećenja glavnih komponenata.*

Introduction

Financial ratios as quantitative indicators calculated on the basis of data from corporate financial reports are widely used in the field of insurance. In the analysis of financial statements, they are applied for the purposes of assessing the liquidity, profitability and solvency of insurance companies, as well as projecting their future financial position and performance. Ratio analysis enables the identification of "strengths" and "weaknesses" of the company, as a basis for business and strategic decision-making. As a rough measure of risks to which insurance companies are exposed, including insurance risks, financial and operational risks, these indicators contribute to adequate risk management. Financial ratios are indispensable analytical tools in assessing the rating of insurers, as well as in the process of regulation and supervision of their business. They provide information on the soundness and performance not only of individual companies, but also of the entire insurance market, as well as its segments. Hence, a number of insurance stakeholders are interested in these indicators, including current and potential policyholders, investors, creditors, management, employees, business partners and government authorities.

The insurance business is characterized by a pronounced complexity, stochastic nature and strict regulation. Therefore, in addition to financial ratios that are common in other sectors, special ratios which take into account the specific characteristics of activities of insurance companies are applied in insurance. So far, a number of sets of financial ratios have been proposed for insurance companies by researches, rating agencies, insurance regulators and supervisors, as well as international institutions (such as CARAMELS set of indicators that are developed by the International Monetary Fund). Thereby, specific indicators are defined for life and non-life insurers, taking into account their substantially different risk exposures and business models. Nevertheless, there is a significant degree of overlap between many indicators in terms of their meaning and interpretation. On the other hand, it is logical to assume that all defined indicators are not equally relevant in all insurance markets, given the level of development and the structure of these markets. Also, the relative importance of individual indicators changes over time due to changes in the macroeconomic environment and regulatory regulations. This raises the question of how to choose from a multitude of financial ratios those that are relevant to a particular insurance market in a given period.

The paper deals with the application of principal component analysis in determining financial ratios that are representative within non-life insurance sector. This multivariate statistical technique permits explanation of relationships existing between a large number of ratios with respect to their common underlying factors. Starting from many financial ratios that are found in the literature in the field of insurance, the aim of this paper is to identify a smaller set of ratios that are most relevant for assessing the financial health of non-life insurance companies in Serbia while retaining the maximum amount of information.

The remainder of the paper is structured as follows. The first section provides an overview of the literature related to the use of factor and principal component analysis with financial ratios in different sectors. Research methodology is described in the second section, followed by explanation of sample selection, data sources and descriptive statistics of research variables in the third section. The research results are presented and discussed in the fourth section.

Literature review

Attempts to identify representative financial ratios are found in several empirical researches conducted in different sectors. Pinches et al. [22] were the first to employ factor analysis in order to develop empirically-based classifications of financial ratios used in industrial organizations. Their pioneering study served as a starting point for later research aimed at grouping of financial ratios and reducing their number for the purposes of more efficient and focused analysis of corporate financial statements in particular sectors.

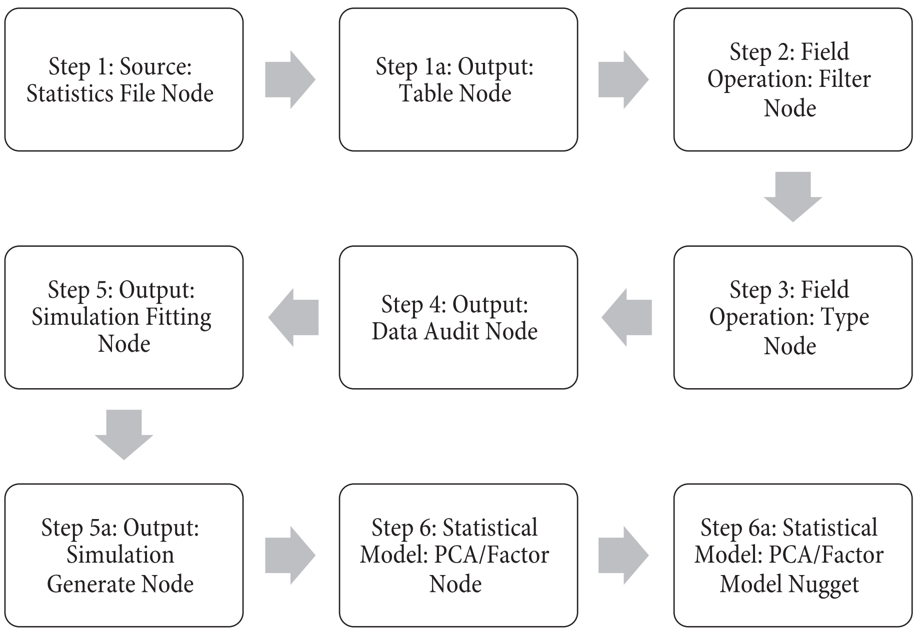
Öcal et al. [19] used factor analysis to determine interrelationships between various financial ratios on a sample of Turkish construction companies during the period 1997-2001. Starting from 25 ratios, they found five underlying factors. Similarly, Vergara & Serna [27] conducted factor analysis on the financial ratios of companies in the Colombian construction sector for the period 2000-2014. They identified four underlying factors explaining 88% of the total variability of 13 conventional financial ratios. De et al. [4] applied factor analysis on 25 financial ratios of selected companies from Indian cement industry to derive eight underlying factors. Yap et al. [29] investigated the application of principal component analysis in the selection of financial ratios that are specific for the consumer product and trading and services sectors in Malaysia. From an initial set of 28 ratios, seven and nine ratios, explaining around 80% of the ratio variances, were selected for the two sectors. Lukason & Laitinen [16] used the factor analysis on 11 financial ratios of bankrupted manufacturing firms from 15 European countries and found five underlying factors. Yoshino & Taghizadeh-Hesary [30] used principal component analysis to obtain a set of ratios useful for predicting a probability of default for a sample of small and medium-sized enterprises in Asia. In order to construct a synthetic financial performance index, Sanz et al. [24] applied principal component analysis on financial ratios of EU-28 companies operating in the publishing sector.

Comparable research in the financial services sector is relatively rare. On the example of the Greek banking sector, Dimitropoulos et al. [5] obtained four risk factors from 11 financial ratios by principal component analysis. A study carried out by Armeanu & Lache [1] derived three principal components from eight financial variables measuring financial strength of insurance companies in Romania. Erdemir & Tatlidil [6] used principal component analysis to reduce the number of input and output variables in data envelopment analysis of efficiency of Turkish insurance companies. To our knowledge, principal component analysis has not been applied to financial ratios of insurance companies in Serbia, and therefore this study is the first attempt.

Research methodology

The process of applying the principal components analysis is a complex data mining procedure [28, p.520], which can be shown in Figure 1. Inside it a stream consisting of several steps is presented. The first step defines the data source as well as the format in which the data is stored. In the next step, the loaded data is presented using a table, so that it is possible to perform a visual screening of the data, in terms of the first data check. In the next step of the data mining stream, the operation of filtering variables is applied, in the sense that only the variables from the data set that refer to the financial ratios go into

Figure 1: Data Mining Principal Component Modelling Stream



Source: The result of the analysis conducted by the authors the further course of the analysis. In step 3, the appropriate measurement scale is determined, for the financial ratios to be analysed. The basic descriptive statistical analysis of all analysed financial ratios is performed in the next step: data audit node.

Due to the fact that the analysis of the principal components is based on the correlation matrix of the analysed financial ratios, we construct it using step number 5, which consists of two nodes (simulation fitting, as well as simulation generate node) [15, p.99]. In the last sixth step, we set up the procedure for data reduction in data mining. Within this step, care should be taken, because both the principal component analysis and the factor analysis are within the same procedure. So one should pay attention, in the sense, that the principal component analysis is actually applied. The result of statistical modelling is given in the last step, and based on it, it will be possible to determine which are the representative financial ratios.

Representative financial ratios can be identified using principal component analysis for dimensional reduction purposes. In the process of dimension reduction, it is determined which financial ratio has a high loading on the extracted principal components. For this reason, in the principal component analysis process, a principal component loadings matrix needs to be formed. The first stage in the analysis of principal components is, of course, detecting whether a sufficient sample size is available, as well as whether the collected data are suitable for the use of principal component analysis.

For sample size, it is recommended that the number of observation should be at least five times larger than the number of variables, while the absolute minimum is 100 observations. When examining whether it is appropriate to apply a dimensional reduction procedure to the data collected, a Kaiser-Mayer-Olkin sample adequacy measure could be used. The given measure is defined in the range from 0 to 1, and if the value of the Kaiser-Mayer-Olkin measure is greater than 0.5, then it is appropriate to apply a dimensional reduction procedure. The given measure is calculated by the formula

$$KMO = \frac{\sum\sum_{i \neq j} \mathrm{r}_{ij}^{2}}{\sum\sum_{i \neq j} \mathrm{r}_{ij}^{2} + \sum\sum_{i \neq j} \mathrm{partial\_r}_{ij}^{2}} \tag{1}$$

where r_{ij} denotes the correlation coefficient between the i^{th} and j^{th} analysed financial ratios, and $partial\_r_{ij}$ denotes the partial correlation coefficient between the the i^{th} and j^{th} analysed financial ratios. Otherwise, the partial correlation coefficient between two financial ratios represents the (ordinary) correlation coefficient when the influences of all other financial ratios are excluded.

Also, to check whether the given data are suitable for the use of the data reduction procedure, Bartlett's test of sphericity can be used, which tests the null hypothesis that the correlation matrix of financial ratios is an identity matrix. The formula for test statistics is

$$x^2 = \frac{2p-6n+11}{6} \cdot \ln|\mathrm{R}| \tag{2}$$

where p represents the number of financial ratios, n is the number of observations, and $|R|$ is the determinant of the correlation matrix of financial ratios.

The next step, which is crucial, is to determine the number of extracted principal components. This is important because the correct statistical methodology that determines the number of extracted principal components is usually not implemented in well-known and popular statistical softwares. They usually have two standard approaches: Kaiser's criterion "eigenvalue greater than 1" or Cattel's scree plot. However, these two approaches usually yield a larger number of extracted principal components. For that reason, Horn's parallel analysis will be used, as well as Velicer's minimum average partial test. In parallel analysis, the initial eigenvalues of the principal components over the observed sample data are compared with the eigenvalues over the random data sets. Namely, a large number (for example 1000) of data sets of random numbers is formed which are of the same dimensions as the actual data set. Then, for each data set with random numbers, eigenvalues are calculated for each principal component, and then these calculated eigenvalues (which has, for example, 1000 for each principal component) are averaged. Finally, the values of real eigenvalues are compared with mean random data eigenvalues. The rule is: the principal components are extracted, where the initial eigenvalues are greater than the mean random data eigenvalues. Otherwise, when calculating eigenvalues, it comes down to a typical problem of eigenstructure.

The task is to create an eigenstructure over the sample data matrix. If the matrix of realized data is denoted by X and its standardized form by Z, then finding the eigenstructure of the matrix X is reduced to finding the eigenstructure of its correlation matrix R, because $R = \frac{1}{n-1} Z^T Z$ where the symbol ^{T} denotes operation transposed matrix. The eigenstructure of the correlation matrix is $R = UDU^T$, U represents the matrix of eigenvectors of R, and D the matrix of eigenvalues of R.

Velicer's minimum average partial procedure is the calculation of partial correlation coefficients, as follows. First, the correlation matrix R is calculated based on the sample data. Then, based on the given correlation matrix, we perform an analysis of the principal components with only one extracted principal component. Subsequently, a matrix of partial correlation coefficients is created, which is obtained when the influence of the first principal component is excluded from the observed variables. Then, the average of the squared partial correlation coefficients is calculated. The given procedure is repeated, but now two principal components are extracted, and so on. Finally, the number of extracted principal components is determined by finding the minimum average of the squared partial correlation coefficients. So, it is determined for which number of extracted principal components the minimum value of the average of the squared partial correlation coefficients is obtained.

Finally, let us point out how principal components loadings are calculated, which is essential for determination of representative financial ratios. Thereby, it should be noted that in the analysis of principal components, principal components are not correlated with each other, and financial ratios are presented in a standardized form. Under these conditions, principal components loadings represent the correlation coefficients between financial ratios and principal components. Thus, the principal components loadings matrix can be calculated by formula

$$A = \frac{1}{n-1} \cdot Z^T C \tag{3}$$

where C represents the matrix of standardized scores of principal components, which are calculated according to formula $C = ZUD^{-1/2}$ [26, p.275]. If we replace the matrix C with the formula for the matrix A, we get $A = \frac{1}{n-1} ZTZUD^{-1/2}$ and that is (based on the previous equations) equal to $A = RUD^{-1/2}$. Since R is represented as $UDUT$, it follows that the matrix A is now equal to $UDUTUD^{-1/2}$. Because the matrix U is orthogonal, ie UTU, it follows that

$$A = UDID^{-1/2} = UDD^{-1/2} = UD^{1/2}. \tag{4}$$

So with the help of already calculated eigenvectors and eigenvalues of the correlation matrix, we are able to calculate the principal components loadings matrix.

Sample selection and descriptive statistics

The non-life insurance sector accounted for 77% percent of the total insurance premium generated on the Serbian insurance market in 2019. Units of observation in our analysis were all 12 insurance companies that constitute this sector, including 6 companies providing exclusively non-life insurance, and 6 composite insurers, providing life and non-life insurance. Thus, the study covered the entire population, not just the sample.

Initially we identified 38 financial ratios which are generally considered relevant to insurance companies, including core set of CARAMELS indicators [3], ratios recommended in the relevant literature [2, 21], as well as ratios used worldwide for supervisory purposes [14] or for assigning credit ratings [17]. The ratios are classified in seven main categories, as shown in Table 1.

Capital adequacy ratios show whether the insurer's capital is sufficient to cover the risks that threaten its business. They link equity to an appropriate position in the balance sheet or income statement that reflects risk exposure. In non-life insurance, written premium on net or gross basis is used as a measure of insurance risk exposure (s1, s4), while the value of total assets approximates exposure to financial risks (s2). In addition, capital adequacy may be impaired with excessive growth of gross premium written (s3), as well as with the oscillations of capital itself (s5).

Asset quality ratios provide a deeper insight into insurer's exposure to investment risks. They are calculated as the share of high-risk assets in insurer's capital (aq1, aq5, aq6) or total assets (aq2, aq3, aq4).

Reinsurance is certainly the most important risk management instrument for insurance companies. However, it implies the possibility that the reinsurer will fail to meet its obligations to the insurer. Low share of net written premium in gross premium written (rpr1), as well as high share of reinsurance reserves in capital (rpr2) or total assets (rpr4) indicate the presence of the reinsurance credit risk. The share of claims paid by reinsurers in total claims (rpr3) is used as a measure of reinsurance performance.

Adequacy of technical reserves, and primarily of loss reserves, is a prerequisite for timely settlement of insurers' obligations to policyholders. From the aspect of preserving the solvency of insurers, it is desirable to have as large technical reserves as possible compared to capital (atr1), premiums (atr2, atr3, atr4) and claims paid (atr5).

The key sources of insurer profit are underwriting business and investment business. The basic underwriting performance indicator is combined ratio (p1) as the sum of loss ratio (net incurred claims / net earned premium) and expense ratio (operating expenses / net earned premium). The share of loss adjustment expenses in net losses paid is also important for non-life insurers' profitability (p9). Investment performance indicators take into account the relative size of net investment income (p2, p3), or investment expenses (p8). Finally, general profitability indicators - return on equity (p4), return on revenue (p5) and derived from it return on premiums (p6), as well as return on assets (p7) are also relevant for insurance companies.

Liquidity ratios reflect the amount of the liquid assets in relation to liabilities (l2, l3, l5, l6, l7), total assets (l4) or claims and expenses paid (l1). Thereby, liquid asset is defined in different ways, starting from cash and cash equivalents, through invested assets to current assets less inventories.

Management soundness indicators represent a rough measure of insurers' exposure to operational risks, among which weaknesses and failures of management stand out. Quantification of this risk is difficult due to lack of data, so it is approximated by the share of expenses in gross premiums written (ms1, ms2).

Financial ratios from the initial set are calculated on the basis of balance sheets and income statements of analysed insurance companies for years 2010 to 2019, which are gathered from the website of the National Bank of Serbia [18]. Table 2 shows the descriptive statistics of all 38 financial ratios included in the analysis of principal components. From the statistical indicators, the mean values as well as the standard deviations are shown. Also, the

Table 1: Financial ratios used in Data Mining Principal Component Analysis

Category/ Name	Formula
CAPITAL ADEQUACY	
s1	Net written premium / Capital
s2	Capital / Total assets
s3	Growth in net written premium
s4	Gross premium written / Capital
s5	Change in Shareholders' equity
ASSET QUALITY	
aq1	Affiliated investments / Capital
aq2	(Real estate + unquoted equities + debtors) / Total assets
aq3	Equities / Total assets
aq4	Real estate / Total assets
aq5	(Equities + real estate) / Capital
aq6	Intangibles / Capital
REINSURANCE RISK AND PERFORMANCE	
rpr1	Net written premium / Gross premium written
rpr2	Reinsurance reserves / Capital
rpr3	Claims paid by reinsurers / Total claims
rpr4	Reinsurance reserves / Total assets
ADEQUACY OF TECHNICAL RESERVES	
atr1	Technical reserves / Capital
atr2	Loss reserves / Net premiums earned
atr3	(Capital + technical reserves) / Net written premium
atr4	Technical reserves / Gross premium written
atr5	Net technical reserves / Average of net claims paid in last 3 years
PROFITABILITY	
p1	Combined ratio (loss ratio + expense ratio)
p2	Net investment income / Net earned premium
p3	Net investment income / Average invested assets
p4	Return on equity - ROE (Net income / Capital)
p5	Return on revenue (Net income / (Premium income + investment income + other income))
p6	Return on premiums (Net income / Gross premiums written)
p7	Return on assets - ROA (Net income / Total assets)
p8	Investment expenses / Gross premiums written
p9	Loss adjustment expenses / Net losses paid
LIQUIDITY	
l1	(Cash + invested assets) / Claims and expenses paid
l2	Liabilities / (Current assets less inventories)
l3	(Current assets less inventories) / Current liabilities
l4	(Current assets less inventories) / Total assets
l5	(Cash + invested assets) / (Unearned premium reserve + loss reserve)
l6	Cash and cash equivalents / Current liabilities
l7	(Current assets less inventories) / Technical reserves
MANAGEMENT SOUNDNESS	
ms1	Operating expenses / Gross premiums written
ms2	Personnel expenses / Gross premiums written

Source: The result of the analysis conducted by the authors.

sample size for each financial ratio is shown. It is natural to expect the sample size to be 120, as we have collected data for all 12 insurance companies for the last 10 years. However for one insurance company, there are data for the last 8 years, since it started operating in 2012, so the sample size for each financial ratio is 118.

Analysing the last column in Table 2, which shows the missing data, we see that for 3 financial ratios, there is 1 missing data, which is less than 1%. Regarding the modern methodology for solving missing data in the analysis of principal components, it is given in [7, p.48]. Also in the material [25, p.97] it is suggested that if the percentage of missing data is less than 5%, the given problem can be solved by traditional methods, that is, either by deleting the incomplete observation, or replacing the missing value with the sample average of financial ratio.

Table 2: Descriptive Statistics of Financial ratios

	Mean	Std. Deviation	Analysis N	Missing N
s1	204.1396%	177.60902%	118	0
s2	29.4879%	18.37983%	118	0
s3	11.1799%	32.50709%	118	1
s4	237.1124%	188.84746%	118	0
s5	18.7251%	65.59689%	118	1
aq1	3.7448%	8.54564%	118	0
aq2	22.6023%	15.96164%	118	0
aq3	2.7376%	6.90531%	118	0
aq4	14.0186%	13.16855%	118	0
aq5	58.5875%	57.00308%	118	0
aq6	3.6087%	8.24254%	118	0
rpr1	84.8310%	18.24678%	118	0
rpr2	20.5800%	37.22054%	118	0
rpr3	10.8755%	12.86955%	118	0
rpr4	4.7781%	7.68097%	118	0
atr1	319.3209%	252.36706%	118	0
atr2	47.6803%	27.53626%	118	0
atr3	279.5808%	184.78260%	118	0
atr4	90.9365%	25.40672%	118	0
atr5	250.0197%	89.26436%	118	1
p1	98.9872%	31.79686%	118	0
p2	7.5288%	27.11162%	118	0
p3	3.9473%	16.90049%	118	0
p4	-3.1904%	64.13051%	118	0
p5	1.1236%	23.98857%	118	0
p6	0.2397%	30.52644%	118	0
p7	0.3366%	8.95121%	118	0
p8	5.0227%	27.03663%	118	0
p9	13.7598%	16.89776%	118	0
l1	300.9626%	325.97969%	118	0
l2	144.4977%	147.03405%	118	0
l3	378.0983%	919.22841%	118	0
l4	63.5955%	23.19596%	118	0
l5	737.5748%	1757.91613%	118	0
l6	17.7281%	30.83298%	118	0
l7	119.9347%	91.27676%	118	0
ms1	35.7607%	11.39627%	118	0
ms2	5.6707%	4.65506%	118	0

Note: For each variable, missing values are replaced with the variable mean.
Source: The result of the analysis conducted by the authors using IBM SPSS Statistics 26 software.

Empirical results

Before applying the analysis for data reduction, it is necessary to first examine whether the assumptions for the given analysis are met. For this purpose, the Kaiser-Mayer-Olkin measure of sample adequacy is applied [8, p.137], as well as Bartlett's test of sphericity [23, p.341]. As Table 3 shows, the value of Kaiser-Meyer-Olkin statistics is 0.6 (ie greater than 0.5), so it is appropriate to apply the data reduction analysis to the given data.

Table 3: Kaiser-Meyer-Olkin Measure and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.601
Bartlett's Test of Sphericity	Approx. Chi-Square	7756.355
	df	703
	Sig.	0.000

Source: The result of the analysis conducted by the authors using IBM SPSS Statistics 26 software.

In Bartlett's test, the null hypothesis represents that the analysed variables are uncorrelated with each other (because it is assumed that the correlation matrix is equal to the unit matrix), so then it makes no sense to perform an analysis to reduce the data. In Table 3, we note that the *p-value* in a given test is less than 5%, so we reject the null hypothesis. After determining that it is appropriate to apply data reduction analysis, ie principal component analysis, over the analysed data, the next task would be to determine the exact number of principal components to be extracted using Velicer's Minimum Average Partial Test, as well as Parallel Analysis [10, p.242].

To apply the parallel analysis, it is necessary to compare the value of initial eigenvalues which is in the second column called Total in Table 4; with the value mean of random data eigenvalues located in the second column of Table 5.

Table 4: Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.586	19.964	19.964	7.586	19.964	19.964
2	6.457	16.991	36.955	6.457	16.991	36.955
3	4.490	11.815	48.770	4.490	11.815	48.770
4	3.884	10.221	58.992	3.884	10.221	58.992
5	2.814	7.404	66.396	2.814	7.404	66.396
6	2.222	5.848	72.244	2.222	5.848	72.244
7	1.599	4.209	76.453			
8	1.356	3.569	80.022			
9	1.088	2.862	82.884			
10	0.971	2.554	85.438			
11	0.888	2.337	87.775			
12	0.869	2.286	90.062			
13	0.831	2.186	92.247			
14	0.578	1.522	93.769			
15	0.483	1.272	95.041			
16	0.320	0.842	95.883			
17	0.282	0.741	96.624			
18	0.249	0.654	97.279			
19	0.201	0.529	97.808			
20	0.196	0.515	98.323			
21	0.148	0.391	98.714			
22	0.111	0.291	99.005			
23	0.089	0.233	99.238			
24	0.077	0.204	99.442			
25	0.045	0.119	99.561			
26	0.034	0.090	99.651			
27	0.033	0.087	99.737			
28	0.029	0.076	99.814			
29	0.017	0.045	99.859			
30	0.016	0.042	99.901			
31	0.011	0.029	99.930			
32	0.008	0.022	99.952			
33	0.005	0.014	99.966			
34	0.004	0.012	99.978			
35	0.004	0.010	99.988			
36	0.003	0.008	99.996			
37	0.001	0.002	99.998			
38	0.001	0.002	100.000			

Note: Extraction Method: Principal Component Analysis.
Source: The result of the analysis conducted by the authors using IBM SPSS Modeler 18.0 software.

Table 5: Random Data Eigenvalues and Average Partial Correlations

Root/	Random Data Eigenvalues		Average Partial Correlations	
Eigenvalues	Means	95 Percentile	squared	power4
0			0.0851	0.0355
1	2.2820	2.4607	0.0943	0.0354
2	2.1108	2.2279	0.1033	0.0380
3	1.9885	2.0977	0.0875	0.0246
4	1.8786	1.9684	0.0875	0.0253
5	1.7831	1.8665	0.0751	0.0181
6	1.6979	1.7679	0.0682	0.0151
7	1.6184	1.6875	0.0594	0.0116
8	1.5436	1.6090	0.0601	0.0121
9	1.4722	1.5327	0.0634	0.0128
10	1.4076	1.4645	0.0608	0.0124
11	1.3435	1.3983	0.0599	0.0136
12	1.2836	1.3397	0.0696	0.0169
13	1.2254	1.2800	0.0756	0.0189
14	1.1681	1.2212	0.0758	0.0179
15	1.1154	1.1630	0.0821	0.0201
16	1.0636	1.1130	0.0852	0.0221
17	1.0141	1.0595	0.0896	0.0230
18	0.9659	1.0142	0.0970	0.0255
19	0.9190	0.9599	0.0973	0.0286
20	0.8753	0.9177	0.1153	0.0363
21	0.8317	0.8740	0.1230	0.0430
22	0.7889	0.8309	0.1387	0.0513
23	0.7484	0.7891	0.1291	0.0439
24	0.7079	0.7477	0.1228	0.0412
25	0.6689	0.7090	0.1324	0.0443
26	0.6314	0.6698	0.1475	0.0528
27	0.5940	0.6329	0.1465	0.0594
28	0.5585	0.5941	0.1501	0.0552
29	0.5232	0.5584	0.1599	0.0645
30	0.4884	0.5232	0.1873	0.0857
31	0.4544	0.4900	0.2151	0.1060
32	0.4213	0.4575	0.2325	0.1205
33	0.3892	0.4226	0.2822	0.1671
34	0.3553	0.3883	0.4103	0.2750
35	0.3229	0.3570	0.6408	0.5179
36	0.2900	0.3217	0.5161	0.4219
37	0.2544	0.2860	1.0000	1.0000
38	0.2144	0.2506		

Source: The result of the analysis conducted by the authors using IBM SPSS Modeler 18.0 software.

The rule in parallel analysis is: extract all those principal components for which the value of initial eigenvalues from column two of Table 4 is greater than the value mean of random data eigenvalues (which is given in column two of Table 5).

The Velicer's minimum average partial test creates the average of the squares of the partial correlation coefficients. Partial correlation coefficients are created in such a way that the influence of the first principal component is excluded from the analysed financial ratios. Then, the averages of the squares of partial correlation coefficients are formed. Now, the partial correlation coefficients are created in such a way that the influence of the first and second principal components is excluded from the analysed financial ratios. And so on. The obtained averages of the squares of the partial correlation coefficients are given in column 4 of Table 5. This test determines the optimal number of extracted principal components, by determining the minimum in the fourth column. The minimum in the mentioned column is in the row where eigenvalue equals 7, which represents the number of extracted principal components.

So, by applying the previous two rules, the number of extracted principal components would be 6 (or 7). If the values of initial eigenvalue for each principal component are displayed on the graph, we get a scree plot, which is represented by Figure 2. On the scree plot we see that the values of the initial eigenvalues of the first 6 (or 7) principal components are on the steep slope of the broken line. If we apply the Kaiser's criterion, then we would extract the first 9 principal components; while there is no clear solution for the implementation of the scree test here, because the line connecting the values of the initial eigenvalues is broken in several places (and it needs to be broken only in one place).

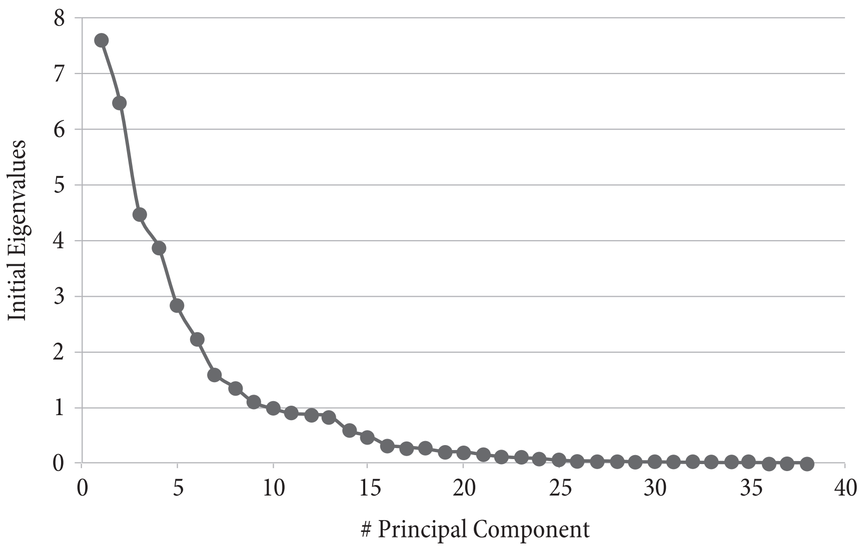
After determining how many principal components to extract, the next step would be to create principal component loadings matrix (Table 6). Values within the given matrix represent the intensity of loadings of each extracted principal component according to each financial ratio included in the analysis. Representative financial ratios are determined by detecting those financial ratios, which are the maximum loaded by the extracted principal components.

The extracted first principal component loads the financial ratio ms1 the most, so that it represents the first representative financial ratio. It should be noted that the maximum load is observed in absolute terms. The extracted second principal component loads the financial ratio s2 the most (in the absolute sense), so that it represents the second representative financial ratio. We arrive at the third representative financial ratio, in a similar way that is, by determining the financial ratio that is most (in absolute terms) loaded by the third principal component. The given financial ratio is aq4. The extracted fourth principal component (in absolute terms) loads the financial ratio rpr4 the most, which also becomes the fourth representative financial ratio. The penultimate extracted principal component loads the p1 financial ratio the most (in absolute amount). The last extracted principal component loads the financial ratio l4 the most.

Table 7 shows the representative financial ratios, which were obtained by applying the analysis of principal components. By analysing the first column of Table 7, it can be seen that all representative financial ratios are from different categories. We also see that the six representative financial ratios, compared to a total of 38 financial ratios analysed, represent a large reduction in data, amounting to almost 85%.

The representativeness of ratio of operating expenses to gross premiums written (ms1) can be explained by the fact that motor third-party liability insurance has a

Figure 2: Scree Plot



Source: The result of the analysis conducted by the authors.

Table 6: Principal Component Loadings Matrix

	Component					
Ratios	1	2	3	4	5	6
s1	-0.331	0.711	-0.369	0.178	0.358	0.133
s2	-0.194	-0.777	0.329	-0.279	0.195	0.073
s3	0.138	0.097	-0.013	-0.034	-0.102	0.166
s4	-0.278	0.715	-0.435	0.082	0.393	0.072
s5	0.127	0.020	-0.012	0.007	-0.278	0.013
aq1	-0.092	0.154	0.250	0.097	-0.062	-0.340
aq2	-0.710	-0.050	0.578	0.024	0.032	-0.034
aq3	-0.281	-0.249	0.491	0.104	0.222	-0.197
aq4	-0.690	-0.124	0.586	0.136	-0.007	0.052
aq5	-0.648	0.375	0.215	0.253	0.288	0.129
aq6	-0.200	0.340	-0.348	-0.045	0.435	0.027
rpr1	-0.361	0.280	0.278	0.667	-0.307	0.139
rpr2	0.125	0.251	-0.327	-0.633	-0.122	-0.160
rpr3	0.302	-0.007	-0.372	-0.517	0.241	-0.279
rpr4	0.261	-0.018	-0.262	-0.765	-0.046	-0.150
atr1	0.033	0.611	-0.565	0.285	0.311	-0.096
atr2	-0.198	0.605	-0.075	-0.355	-0.387	0.055
atr3	0.566	-0.637	-0.294	0.025	0.177	0.016
atr4	-0.383	0.612	0.040	-0.233	-0.384	0.144
atr5	0.069	0.263	0.000	-0.144	-0.168	0.463
p1	-0.241	-0.250	-0.115	0.010	0.512	-0.116
p2	0.729	0.424	0.358	0.096	0.303	0.060
p3	0.560	0.527	0.475	-0.039	0.358	0.074
p4	0.462	-0.278	0.475	-0.165	-0.495	-0.187
p5	0.753	0.350	0.497	-0.106	0.142	0.035
p6	0.723	0.392	0.513	-0.037	0.194	0.064
p7	0.720	0.293	0.576	-0.120	0.031	0.014
p8	-0.522	-0.495	-0.524	0.143	-0.377	-0.104
p9	0.402	-0.120	-0.117	0.545	-0.188	0.205
l1	0.657	-0.417	-0.291	0.415	0.078	-0.064
l2	0.182	0.138	-0.109	0.422	0.028	-0.708
l3	0.468	-0.315	-0.203	0.472	-0.072	0.393
l4	0.191	-0.011	-0.235	-0.245	-0.113	0.765
l5	0.540	-0.312	-0.231	0.638	-0.044	0.110
l6	-0.061	-0.543	0.128	0.108	0.464	0.136
l7	0.018	-0.608	-0.081	-0.443	0.319	0.401
ms1	-0.814	0.072	0.276	0.026	0.113	0.211
ms2	-0.378	-0.675	0.193	-0.221	0.368	0.008

Note: Extraction Method: Principal Component Analysis. 6 components extracted.
Source: The result of the analysis conducted by the authors using IBM SPSS Modeler 18.0 software.

predominant share in the non-life insurance premium in the analysed market. In this type of insurance, there is a problem of excessive acquisition costs that exceed the overhead allowance, as part of the gross premium intended to cover them. Hence, the relative amount of operating expenses is especially important from the aspect of measuring management soundness.

The choice of a combined ratio (p1) as representative of the non-life insurance sector is expected, as it is a summary indicator of non-life underwriting profitability, reflecting the sufficiency of net earned premiums to cover net incurred claims and operating expenses. Any analysis of the performance of non-life insurers without the combined ratio would be incomplete. It is also logical to find a liquidity ratio (l4) among the representative indicators for the non-life insurance sector, having in mind the short-term nature of the sources of funds and liabilities of non-life insurers.

Table 7: Representative financial ratios based on Data Mining Principal Component Analysis

Category	Name	Formula
Management soundness	ms1	Operating expenses / Gross premiums written
Capital adequacy	s2	Capital / Total assets
Asset quality	aq4	Real estate / Total assets
Reinsurance risk and performance	rpr4	Reinsurance reserves / Total assets
Profitability	p1	Combined ratio (loss ratio + expense ratio)
Liquidity	l4	(Current assets - inventories) / Total assets

Source: The result of the analysis conducted by the authors.

The representativeness of the ratio of real estate to total assets (aq4) stems from the negligibly low presence of other forms of high-risk assets (such as equities) in insurers' assets, primarily due to the underdevelopment of the domestic capital market. Therefore, it is not reasonable to expect that the calculation of other asset quality ratios will significantly contribute to the assessment of the financial position of non-life insurers operating in Serbia.

The share of net written premium in gross premium written is commonly used to measure reinsurance risk. However, the interpretation of this ratio is complicated and multi-dimensional: a low value indicates a pronounced reinsurance risk, while a high value implies a pronounced insurance risk. The conducted analysis shows that the share of reinsurance reserves in total assets (rpr4) as a more direct indicator of reinsurance credit risk, is more relevant for the Serbian non-life insurance sector.

Finally, it is interesting that from the aspect of measuring capital adequacy to cover risks, the emphasis is on the insurers' assets (s2). Insurance risks, which are approximated by the premium, are generally considered to be the most important for non-life insurers. However, the obtained results show that in the case of the non-life insurance sector in Serbia it is necessary to pay special attention to financial risks, arising from the assets side of the balance sheet of insurers.

Conclusion

Over the years, a large number of financial ratios have been developed and applied in the field of insurance. The calculation of all these ratios in the implementation of the financial analysis of insurance companies would be not only impractical, but also of little use, due to the interrelationships that exist between different ratios. In addition, not every ratio is equally suitable for every insurance market and in every period.

In this paper we applied principal component analysis in determining financial ratios that are representative within non-life insurance sector in Serbia. Starting from numerous financial ratios that are generally considered relevant for insurance companies, we identified a smaller number of ratios which can capture almost the same quantity of information available in the original larger set. Analysis has been applied over audited financial data of all 12 companies constituting non-life insurance sector in Serbia for the period 2010-2019. Initially 38 variables (financial ratios) were selected for the study and classified in seven categories.

Six ratios, operating expenses to gross premiums written, capital to total assets, real estate to total assets, reinsurance reserves to total assets, liquid assets to total assets and combined ratio were found to be representative for the sector. These ratios explain 85% of the total variability of all analyzed financial ratios.

The obtained results can serve as an input for further research based on financial indicators of non-life insurers. At the same time, the results can be useful for an efficient and purposeful analysis of the financial position and performance of individual insurance companies and the entire non-life insurance sector in Serbia. Thus, they can be valuable at the micro-level for business and investment decision making, as well as at the macro-level, in market surveillance and policymaking for the insurance sector.

The limitation of this research is certainly the small number of companies that participated in the analysis. However, since we limited the research to the example of the Republic of Serbia, the analysis included all companies in the non-life insurance sector. The time horizon for observing the financial ratios of the analyzed non-life insurance companies is the last available 10 years. The direction of future research would certainly, among other things, be to check the consistency of the obtained results on the basis of the latest collected data. Additionally, principal component analysis could be applied in determining financial ratios relevant for life insurance sector in Serbia.

References

1. Armeanu, D., & Lache, L. (2008). Application of the Model of Principal Components Analysis on Romanian Insurance Market. *Theoretical and Applied Economics, 6*(523), 11-20.
2. Arterton, A. J., Dean, J. W., Michaelides, N., Silverman, A. H., Vince, A. R. (1993). What ratios really matter? *General Insurance Convention 1993,* Institute and Faculty of Actuaries.
3. Das, U., Davies, N., & Podpiera, R. (2003). Insurance and Issues in Financial Soundness. *IMF Working Paper*, No. 03/138, 1-43.
4. De, A., Bandyopadhyay, G., & Chakraborty, B. N. (2011). Application of the Factor Analysis on the Financial Ratios and Validation of the Results by the Cluster Analysis: An Empirical Study on the Indian Cement Industry. *Journal of Business Studies Quarterly, 2*(3), 13-31.
5. Dimitropoulos, P. E., Asteriou, D., & Koumanakos, E. (2010). The relevance of earnings and cashflows in a heavily regulated industry: Evidence from the Greek banking sector. *Advances in Accounting, 26*(2), 290-303.
6. Erdemir, Ö. K., & Tatlidil, H. (2017). Data Reduction in Data Envelopment Analysis: A Research on Efficiency of Insurance Companies. *Journal of Current Researches on Business and Economics, 7*(1), 65-77.
7. Geraci, M., & Farcomeni, A. (2018) Principal Component Analysis in the Presence of Missing Data. In Naik, G.R. (Eds.), *Advances in Principal Component Analysis* (pp. 47-70). Singapore: Springer Nature Singapore Pte Ltd.
8. Hair Jr, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (Eighth Edition). Andover: Cengage Learning, EMEA.
9. Hardle, W. K., & Simar, L. (2015). *Applied Multivariate Statistical Analysis* (Fourth Edition). Berlin: Springer-Verlag.
10. Ho, R. (2014). *Handbook of Univariate and Multivariate Data Analysis with IBM SPSS* (Second Edition). Boca Raton, FL: CRC Press.
11. IBM (2017). *IBM SPSS Statistics Algorithms*. Armonk, NY: IBM Corporation.
12. Johnson, R., & Wichern, D. (2014). *Applied Multivariate Statistical Analysis* (Sixth Edition). Harlow, England: Pearson Education Limited.
13. Jolliffe, I.T. (2002). *Principal Component Analysis* (Second Edition). New York: Springer-Verlang New York, Inc.
14. Kwon, W. J., & Wolfrom, L. (2016). Analytical tools for the insurance market and macro-prudential surveillance. *OECD Journal: Financial Market Trends,* 2016(1).

15. Larose, D. T., & Larose, C. D. (2015). *Data Mining and Predictive Analytics* (Second Edition). Hoboken, New Jersey: John Wiley & Sons, Inc.
16. Lukason, O., & Laitinen, E. K. (2016). Failure processes of old manufacturing firms in different European countries. *Investment Management and Financial Innovations, 13*(2), 310-321.
17. Moody's Investors Service (2019). *Property and Casualty Insurers Methodology.* New York: Moody's Investors Service.
18. National Bank of Serbia, *Insurance undertakings operations,* available at https://nbs.rs/sr/finansijske-institucije/osiguranje/poslovanje/, accessed 15.03.2021.
19. Öcal, M. E., Oral, E. L., Erdis, E., & Vural, G. (2007). Industry financial ratios - application of factor analysis in Turkish construction industry. *Building and Environment, 42,* 385-392.
20. O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, *32*(3), 396-402.
21. Outreville, J. F. (1998). *Theory and Practice of Insurance.* Boston, MA: Springer.
22. Pinches, G. E., Mingo, K. A., & Caruthers, J. K. (1973). The Stability of Financial Patterns in Industrial Organizations. *The Journal of Finance, 28*(2), pp. 389-396.
23. Pituch, K. A., & Stevens, J. P. (2016). *Applied Multivariate Statistics for the Social Sciences – Analyses with SAS and IBM's SPSS* (Sixth edition). New York: Routledge.
24. Sanz, J. Á., Bedate, A. M., & Durántez, M. (2018). Determining some factors of the financial situation in the European Union publishing sector. *Review of Economic Perspectives, 18*(1), 25-43.
25. Tabachnick, B. G., & Fidell, L. S. (2014). *Using Multivariate Statistics* (Sixth Edition). Harlow, England: Pearson Education Limited.
26. Tacq, J. (1997). *Multivariate Analysis Techniques in Social Science Research.* London: Sage Publications.
27. Vergara, J., & Serna, M. (2018). Factor Analysis to Evaluate the Financial Performance of the Construction Industry in an Emerging Market: The Case of Colombia. *Revista de Métodos Cuantitativos para la Economía y la Empresa, 178*(1), 52-70.
28. Wendler, T., & Grottrup, S. (2016). *Data Mining with SPSS Modeler.* Switzerland: Springer International Publishing Switzerland.
29. Yap, B. C.-F., Mohamad, Z., & Chong, K.-R. (2013). The Application of Principal Component Analysis in the Selection of Industry Specific Financial Ratios. *British Journal of Economics, Management & Trade, 3*(3), 242-252.
30. Yoshino, N., & Taghizadeh-Hesary, F. (2015). Analysis of Credit Ratings for Small and Medium-Sized Enterprises: Evidence from Asia. *Asian Development Review, 32*(2), 18-37.

Vladimir Vasić

was born in Belgrade in 1969. He graduated from the Faculty of Economics in 1994, completed a Master course in Statistical Analysis in 1999 and acquiring a PhD title in Statistics at the Faculty of Economics in 2003. He is employed as Assistant Professor at the University of Belgrade – Faculty of Economics (Department of Statistics and Matematics) at courses: Multivariate Analysis and Nonparametric Statistical Methods. He has dealing with the implementation of data mining and data science in the business economy. He is involved in many commercial projects in the region. He is the author of over 20 papers in journals on the SCI and SSCI lists.

Jelena Kočović

is Full Professor at the the University of Belgrade – Faculty of Economics, where she teaches Financial and Actuarial Mathematics, Insurance and Insurance Tariffs. She is Head of the Department of Statistics and Mathematics. She has published over 300 scientific papers in the field of insurance, actuarial science and investment. She is a member of the Philosophical Society of Lomonosov MSU and the Scientific Association of Economists of Serbia. She served as director of the Centre for Scientific Research of the Faculty of Economics. She is a certified actuary, a court expert in the field of finance and actuarial science and a UN expert for catastrophic risk management. She was president of the Serbian Actuarial Association. She is a member of several committees of the International Actuarial Association. Professor Kočović has organised a number of international symposia and managed many scientific research and commercial projects, as well as innovative courses on financial mathematics, insurance and actuarial science

Marija Koprivica

(née Jovović) is Associate Professor at the University of Belgrade – Faculty of Economics, where she teaches Financial and Actuarial Mathematics and Insurance Tariffs. She participated in numerous domestic and international scientific conferences, as well as innovative courses in the field of insurance, actuarial and financial mathematics. Her scientific interests are mostly related to the application of modern actuarial and models of financial mathematics in insurance and financial markets. She has published a significant number of scientific papers in these fields in monographs, journals and conference proceedings. She is a member of the Society of Economists of Belgrade, of the Serbian Actuarial Association and of the International Actuarial Association.