

Distributional properties of some financial ratios in insurance

Hans van der Heijden

University of Surrey

h.vanderheijden@surrey.ac.uk *

Abstract

This paper presents a distributional analysis of some financial ratios in the insurance industry. Although previous studies have established a reasonably good understanding of the frequency distributions of financial ratios, to date most of our knowledge has been derived from cross-sectoral studies, or from the manufacturing sector in particular. This paper aims to contribute to the literature by further applying and extending this knowledge into the financial services sector.

The focus is on four key profitability ratios: the combined ratio, the expense ratio, the loss ratio and return on assets. A sample of 115 global insurance conglomerates, 64 from the United States and 55 from the EU, is used to test the effects of an outlier-removal procedure and data transformation on the normality of the distributions. The study successfully applies a modest normalisation procedure for all four ratios, highlights the dangers of conducting parametric tests on non-normal data series, and finds that lognormal transformation is beneficial just for the Return on Assets ratio.

1 Introduction

One of the most basic assumptions of cross-sectional financial ratio analysis is that the frequency distribution of any given financial ratio is normally distributed. This assumption is important because it provides the researcher the opportunity to use a range of powerful statistical tests. These include parametric tests such as product-moment correlations and multiple regressions, all of which are pervasive in accounting and finance research.

Yet despite the widespread adoption of this assumption, it is also well recognised in the literature that, unfortunately, data series of financial ratios are rarely normally distributed (Deakin, 1976; Mcleay & Omar, 2000). To resolve the violation of this key assumption, a substantial body of research has been devoted to the normalisation of financial ratio distributions. The results have been a number of guidelines on data transformation as well as outlier treatment.

*Presented at the British Accounting & Finance Association (BAFA) Annual Conference, Birmingham, 12-14 April 2011

This paper aims to contribute to this general body of research by analysing the distributional properties of some key financial ratios in the insurance sector. Although ratios have been analysed in the past for a range of industry sectors (see below for a review), analysis of the insurance industry has been very limited. As with many other industries in the financial services sector, this industry carries with it some unusual characteristics which would not apply to, say, the manufacturing sector. For example, life insurers sell insurance products that customers pay towards over a much longer term (> 20 years) than typical manufacturing products.

This paper will be concentrating on four key profitability ratios: the combined ratio, the expense ratio, the loss ratio and return on assets. All four measures are of critical interest to insurance managers as well as financial analysts. A sample of 115 global insurance conglomerates is used to test the effects of an outlier-removal procedure and data transformation on the normality of the distributions. The study applies successful normalisation procedures for all four ratios and finds that log transformation is beneficial only for the Return on Assets ratio. The study also highlights the dangers of drawing conclusions from parametric tests on non-normalised financial ratios.

This paper is structured as follows. First, I will provide a brief overview of the distributional properties of financial ratios in general, and a summary of the results of normalising data and data transformation. The next section contains an overview of the data used for this study. I will outline a procedure to transform the data such that it approaches the normal distribution. This is followed by a Results section, in which the results of the normalisation procedure are provided. A discussion section concludes the paper.

2 Theory

The research is set against the background of distributional properties of financial ratios in general. This section gives a brief overview of the main results in the literature. [Lev and Sunder \(1979\)](#), [Barnes \(1987\)](#), and [Mcleay and Omar \(2000\)](#) provide more detailed reviews.

As mentioned before, the central concern in this research area is the question whether the distribution of ratios can be assumed to follow the normal (Gaussian) distribution. This assumption is of importance because ratios are often used in statistical tests. The most commonly used tests are employed within a long-standing strand of accounting research, one that has occupied itself with predicting the likelihood of financial distress or bankruptcy. [Beaver \(1966\)](#) and [Altman \(1968\)](#) are pioneering studies in this regard. The statistical methods used in this area, such as multiple discriminant analysis, require the financial ratios to be normally distributed.

It is worth mentioning that several studies have applied these financial distress prediction models to the insurance industry, with reasonable success. A range of financial ratios, some generic, some insurance-specific, have been used to classify solvent and insolvent insurance firms. For example,

BarNiv and Hershberger (1990) used several competing models to classify financially distressed life and health insurance firms. BarNiv (1990) uses logit and other models to forecast the financial distress of property and casualty insurance firms. Other studies in this line of research are Ambrose and Carroll (1994), BarNiv and McDonald (1992), S. H. Lee and Urrutia (1996), and Baranoff, Sager, and Witt (1999).

In these studies, the non-normality of insurance ratios is often highlighted as a problem. For example, BarNiv and Hershberger (1990) include a non-parametric model in response to non-normality. Similarly, Ambrose and Carroll (1994) uses a stepwise logit model in favour of multiple discriminant analysis to deal with the violations of normality. These studies serve to illustrate that insurance data also suffers from non-normality of financial ratios, and that is worth examining this issue in more detail.

A seminal paper that addressed the question of normality head-on is Deakin (1976). Deakin studied 11 commonly used ratios for 1114 companies from 1955 to 1973, and concluded that no less than 10 ratios violated the normality assumption (the exception being the Total Debt to Total Assets ratio). Logarithmic and square-root distributions did improve the situation for some ratios but no generic guidelines could be provided.

In a subsequent review of research on financial ratio distributions, Ezzamel, Mar-Molinero, and Beecher (1987) summarise several key findings. One is the general acknowledgment that financial ratio distributions do not, on the whole, follow the normal distribution. They are instead quite heavily skewed to the right, and they possess a number of extreme outliers. Again, several transformations do seem to improve the normality of the distribution. The authors comment that if data is normally distributed, mean and standard deviation are sufficient statistics to describe the sample, but if they are not, higher moments (such as skewness and kurtosis) must be mentioned also.

A number of reasons have been put forward as to why financial ratio distributions are not normally distributed. These can be split into two categories. The first category rejects the underlying notion (that ratios are normally distributed) altogether and instead points to other reasons why data may not be normal. The second category recognises that ratios may not *initially* be normally distributed, but once the appropriate data transformations and outlier detection mechanisms have been applied, these ratios will eventually be normally distributed.

Three arguments fall into the first category. The first argument as to why ratios may not be normally distributed, even after outlier removal, is because they tend to violate what is known in the literature as the proportionality assumption (Whittington, 1980; Barnes, 1987; Fieldsend, Longford, & Mcleay, 1987). Summarising this assumption briefly, a ratio defined as y/x controls for the effect of the size variable x , but only if the relationship $y = bx$ holds. Most ratios are not strictly proportional and contain an intercept a , as in $y = a + bx$. Skewness in the distribution is a indicator that such an intercept exists.

A second reason why financial ratios may not be normally distributed is because the sample is not homogeneous (C. Lee, 1985; Ezzamel et al., 1987; So, 1987). Ratios have been shown to be dependent on the industry, e.g., with some industries having structurally higher current ratios than others. C. Lee (1985) controlled for this industry effect, and found that confining samples to specific industries enabled more financial ratio series to approach the normal distribution. He did, however, still find that many ratios did not follow this distribution, even after industry was controlled for.

A third reason why financial ratios may not be normally distributed is because they follow a different distribution altogether. Several different distributions have been proposed, including the Gamma distribution (Frecka & Hopwood, 1983), the Cauchy distribution and the Student distribution (McLeay, 1986), and the stable Pareto distribution (So, 1987). All of these are variations of the Gaussian distribution, but allow for the occurrence of more extreme values.

The theoretical and empirical validity of these arguments is acknowledged, but for the remainder of the paper, we will follow the tradition of the second category of arguments. The violation of the proportionality assumption is deemed to be a matter of concern only after normality cannot be reached through a modest amount of data transformation. The heterogeneity of the sample is not deemed to be applicable here because we will be focusing on the insurance industry only, and so the data set is assumed to be sufficiently homogeneous (with the acknowledgment that vastly different business models exist even within the insurance industry). Finally, although the exploration of other statistical distributions is a valid direction of research, it does not resolve the core problem of allowing the use of standard parametric tests, many of which presume the Gaussian distribution.

Consequently, we will assume that a modest amount of data manipulation is sufficiently helpful to reach normality of the distribution. Broadly, we will speak of *normalisation* of a data series if we manipulate the data series such that it follows the normal distribution. The two tools for normalisation, outlier treatment and data transformation will now be discussed briefly.

Outlier treatment. Outlier elimination (or “trimming”) is generally the most frequently used method, although outlier replacement, i.e., with the next available non-outlier, also known as “Wind-sorising”, has also been applied. A well-known study in the literature is Frecka and Hopwood (1983), who removed outliers from Deakin’s data and found that this did improve the normality of the distributions of many of the financial ratios.

The choice of outlier boundaries is subject of some concern. Typically, outliers are identified through the use of acceptable boundaries away from the mean. For example, an extreme value could be characterised as an outlier if it is further away than the mean plus or minus several times the standard deviation. But this rule assumes that data follows the normal distribution, which is what we were examining in the first place. An outlier removal procedure developed by Martikainen, Perttunen, Yli-Olli, and Gunasekaran (1995), described in further detail below, helps to resolve this issue, by not setting boundaries in the first place. In stead of a single pass through the data (screening

outliers according to the boundary values), it gradually removing extreme values, one at a time, until normality is deemed to be reached.

Data transformation. As previously mentioned, it has been reported that ratios follow a lognormal distribution or quadratic distribution, and that consequently, a transformation to natural logs or square roots should return the normality of the data. Although this has been successfully applied on many ratios, several studies also report specific concerns with certain transformations. For example, Ezzamel et al. (1987) note a problem with logarithmic transformation, in that it overcompensates positive skewness, to the extent that it leads to negative skewness (i.e. a distribution leaning to the left). This is a particular problem for percentage ratios, with values between 0 and 1. Because the natural log of x tends to negative infinity as x approaches 0, values of x close to zero become disproportionately negative. They conclude that logarithmic transformation may actually worsen normality for those ratios.

Common sense would suggest that logarithmic transformation be applied only to those ratios that follow the lognormal distribution. Fieldsend et al. (1987) observe that this appears to be the case for basic accounting variables such as Sales and Total Assets which are bounded at zero and represent summaries of transactions. Ratios of lognormal variables, e.g., Sales to Total Assets, may also be lognormal (Lev & Sunder, 1979).

To transform data successfully to a normal distribution, Mcleay and Omar (2000) make a distinction between bounded and unbounded ratios. Bounded ratios are the ones previously discussed, with a lower boundary of zero. Bounded ratios are characterised by high skewness and they consequently benefit from a skewness reducing power transformation. Unbounded ratios are generally characterised by high kurtosis, and they consequently benefit from a kurtosis reducing transformation.

Having summarised this literature, the paper will now proceed with an empirical study of the effectiveness of normalisation, and examine whether a “modest” outlier removal procedure is indeed sufficient to turn insurance profitability ratios into normally distributed variables. We apply this both for the original data series as well as the two common transformations, natural log and square root. The selection and preparation of the data set is discussed next.

3 Data and Method

Data for this study was collected using the Worldscope database, available through ThomsonReuters Datastream (Thomson Financial, 2003). A total of 115 large, publicly listed insurance conglomerates were selected for study. Selection was straightforward and based on a complete set of insurers listed on US and European stock exchanges. 64 EU insurers and 55 US insurers were thus selected. Each insurance conglomerate operates in a range of insurance lines, as evidenced by the multiple SIC codes with which each business was described (6311 – Life Insurance, 6321 – Health Insurance,

6331 – Casualty Insurance, etc).

The four ratios under study are the combined ratio (CR), the expense ratio (ER), the loss ratio (LR), and return on assets (ROA). The loss ratio is defined as the Claim and Loss Expense plus the Long Term Insurance Reserves divided by the Premiums Earned. The expense ratio is defined as the Underwriting Expenses divided by the Total Premiums Written. The combined ratio is the sum of the loss ratio and the expense ratio. These ratios are based on data from the income statement, and define the extent of loss on a premium, i.e., the amount not contributing to profit. The loss ratio is to be interpreted as the amount “lost” on paying out insurance claims. For profitability purposes, the lower this ratio the better. The expense ratio is to be interpreted as the amount “lost” on general expenses. From the perspective of efficiency, the lower this expense ratio the better. The fourth ratio, Return on Assets is defined as net income divided by last years total assets, adjusted for interest and tax.

To normalise a ratio data series, I follow a procedure similar to [Martikainen et al. \(1995\)](#). First, I check whether the data already follows the normal distribution using the Shapiro-Wilk test. If the Shapiro-Wilk W value is not significant at $p = .05$, the conclusion is that the data is normally distributed and stop. If it is not, then I check the skewness of the data, to examine whether the data is left skewed (skewness < 0) or right skewed (skewness > 0). If the data is left skewed, then I remove the minimum extreme value. If the data is right skewed, then I remove the maximum extreme value. After having just removed one outlier, I check for normality again using Shapiro-Wilk. I repeat this procedure of checking skewness and removing extrema until Shapiro-Wilk is finally non-significant.

The benefit of this procedure is that only the minimum number of outliers are removed to make the data normally distributed. A fixed boundary value (such as multiples of standard deviation) would risk the removal of more outliers than strictly necessary.

The normalisation procedure was implemented using the programming language Java and the statistical package R ([R Development Core Team, 2010](#)). Data was exported and stored into a MySQL database. Java was used to retrieve the data from MySQL, apply the algorithm and output the result data, while R provided the Shapiro-Wilk tests. The rJava package was used to integrate R with Java ([Urbanek, 2010](#)).

The four financial ratios were examined for three years, 2007, 2008, and 2009. This period was chosen because they include the year 2008 in which financial services underwent a significant global crisis. Three years were chosen to examine whether normalisation and transformation procedures are just one-off beneficial, or whether there is a pattern over the years.

4 Results

Table 1 presents the results of the normalisation procedure, both of the original data series (first data column) as well as the natural log transformation and the square root transformation (second and third data column).

Table 1: Outliers removed before normality is reached ($N = 115$)

Variable	Original	Natural Log Transformation	Square Root Transformation
2007			
Combined Ratio	5 (7%)	2 (3%)	4 (5%)
Expense Ratio	3 (4%)	4 (5%)	4 (5%)
Loss Ratio	7 (7%)	1 (1%)	4 (4%)
ROA	52 (48%)	0 (0%)	33 (31%)
2008			
Combined Ratio	13 (18%)	9 (13%)	9 (13%)
Expense Ratio	3 (4%)	5 (7%)	3 (4%)
Loss Ratio	13 (13%)	7 (8%)	8 (9%)
ROA	13 (12%)	0 (0%)	5 (6%)
2009			
Combined Ratio	8 (10%)	7 (9%)	7 (9%)
Expense Ratio	3 (4%)	8 (10%)	4 (5%)
Loss Ratio	12 (12%)	8 (8%)	8 (8%)
ROA	32 (29%)	4 (4%)	3 (3%)

Note: Percentages based on available data without missing values.

We can observe that, in the main, both natural log and square root transformations need to remove a few less outliers than the original data series before normality is reached. It is important to recognise, however, that this is in some part due to the technicalities of the conversion. In particular, a negative ratio cannot be transformed to either log or square root, and so such a value does not carry over during transformation. Negative values are rare and they tend to be identified as outliers in the original data series. Thus, the original data series removes some outliers that are also removed during transformation.

We can also see from Table 1 that the first three profitability ratios do not on the whole benefit much from data transformations. The ROA ratio, on the other hand, benefits significantly. Its transformation to log distribution is a striking improvement. For example, for its 2007 data, the original data series needed to be stripped down with 48% observations, indicating that the procedure effectively broke down and that normality cannot be reached in this way without seriously compromising the representativeness of the data set. That same data series, once log transformed, is normally distributed immediately.

Does normalisation and transformation result in any material improvement in our understanding

of the data? To examine this question further, I will focus on the intercorrelations between these four ratios. I will show that normalisation and transformation have implications for the material strength and direction of these intercorrelations.

To demonstrate the effects of normalisation and transformation, Table 2 first presents a correlation matrix of ratios without normalisation and without transformation. These are, in effect, the original data series including outliers. We have focused on the year 2009 for this demonstration; analyses for the years 2007 and 2008 show similar results but are not reported here.

Table 2: Intercorrelations for Non-normalised, Non-transformed Insurance Ratios 2009

Measure	<i>n</i>	<i>M</i>	<i>SD</i>	1	2	3	4
1. Combined Ratio	78	140	148	–			
2. Expense Ratio	79	20	37	.33**	–		
3. Loss Ratio	100	120	146	.97**	.08	–	
4. Return on Assets	112	2.63	3.51	-.15	-.04	-.18	–

Notes: * $p < .05$, ** $p < .01$. Test is Pearson r .

Ratios are $\times 100\%$

Looking at Table 2, three observations are worthy of note. Firstly, the combined and loss ratios are over 100%, which indicates that on average, insurance companies lost more money on premiums than they earned. This is likely as a direct or indirect result of the 2008 financial crisis. Second, we see that expense and loss ratios are related to the combined ratio (as expected), but that expense and loss ratios appear to be unrelated. Third, and most importantly, Table 2 does not appear to suggest a correlation between Return on Assets and the other ratios. There is however a very weak negative relationship, which we would expect because both *low* loss/expense ratios and *high* returns on assets are indications of financial health.

From a purely statistical perspective, Table 2 presents incorrect tests in the sense that parametric test must be applied with caution on data that is demonstrably not normally distributed. It means that the conclusions we draw from these tests may be biased, or simply wrong. Therefore, Table 3 presents the results in the same way, but this time with the normalised data series for each of the four ratios.

Table 3: Intercorrelations for Normalised, Non-transformed Insurance Ratios 2009

Measure	<i>n</i>	<i>M</i>	<i>SD</i>	1	2	3	4
1. N(Combined Ratio)	70	100	27	–			
2. N(Expense Ratio)	76	14	7	-.34**	–		
3. N(Loss Ratio)	88	81	27	.97**	-.50**	–	
4. N(Return on Assets)	80	1.09	1.30	-.43**	.37**	-.54**	–

Notes: * $p < .05$, ** $p < .01$. Test is Pearson r .

Ratios are $\times 100\%$

This table shows a better representation of the correlation between the ratios, because this time,

the parametric test is performed on normalised data. Commenting first on the descriptive results, we can see that the variance has substantially reduced for all ratios. This is to be expected because a number of extreme values have been removed to normalise each series. We can also see that the values have less reasonable averages, although loss ratios of 81% and combined ratios of 100% remain nonetheless indicators of financial difficulty.

Commenting further on the correlations we can see that the expense ratio is still significantly (but weakly) correlated to the combined ratio. This time, however, it is in the opposite direction! This is quite remarkable and provides evidence of the danger of using non-normalised data: without normalising, our conclusion would have been exactly the opposite.

We can also see that this time, the ROA is significantly correlated with the other ratios. There is a negative correlation with the loss ratio, and consequently also with the combined ratio (because loss ratio and combined ratio are almost perfectly correlated). This is perhaps better in line with expectations, because both lower ROAs and higher loss ratios are indicators of financial distress. There is, on the other hand, also a positive correlation with the expense ratio and ROA, which is not entirely consistent with that reasoning. It appears that in the normalisation process, the expense ratio lost three very influential outliers, creating a reversal in sign and consequently a change in direction of expected relationships.

The above results, of course, do not take into account the positive benefit of log-transforming ROA. Table 4 below presents the same results but this time with the lognormal transformation of ROA.

Table 4: Intercorrelations for Normalised and Transformed Insurance Ratios 2009

Measure	<i>n</i>	<i>M</i>	<i>SD</i>	1	2	3	4
1. N(Combined Ratio)	70	100	27	–			
2. N(Expense Ratio)	76	14	7	-.34**	–		
3. N(Loss Ratio)	88	81	27	.97**	-.50**	–	
4. N(ln(Return on Assets))	91	2.05 ¹	2.94 ¹	-.74**	.37**	-.73**	–

Notes: * $p < .05$, ** $p < .01$. Test is Pearson r .

Ratios are $\times 100\%$.

¹ Values are e^x .

The benefit from the log-transformation is that more insurance companies have now returned to the data set, and thus a more complete and more representative sample is used for the statistical tests. In this particular case, 12 insurance companies have returned, and we can see from Table 4 that they are responsible for a higher return on assets (2.05) than the insurance companies from the previous table.

In addition, we can observe that the inclusion of these, perhaps somewhat more extreme insurance companies have strengthened the relationship between ROA and the loss and combined ratio. This relationship is now much stronger and more pronounced than it was previously. Figure 1 il-

illustrates this graphically, visualising the association between combined ratio and return on assets before and after normalisation and transformation.

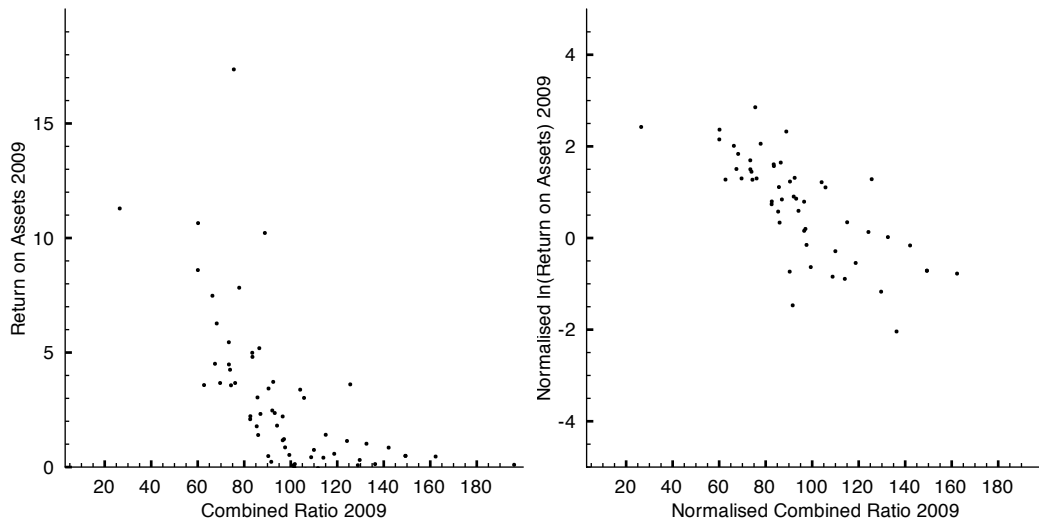


Figure 1: Correlations of combined ratio and return on assets, before and after normalisation and transformation

5 Discussion

Summarising the results of this research, we have found that three profitability ratios, when analysed in cross-sectional fashion, can be normalised with a modest outlier removal procedure. Lognormal and square root transformations did not materially improve the situation. The Return on Assets ratio proved to be substantially more problematic in this respect, but good results were reported following a natural log transformation. Our third finding is that these normalisation and transformation procedures greatly improve the “hidden” strength of correlation between all these ratios.

All four ratios thus benefitted from outlier removal, and one ratio also benefitted from lognormal transformation. These findings are reasonably consistent with earlier results from prior analysis of distributional properties in non-financial services industries. They demonstrate that, consistent with our earlier assumption, cross-sectional financial ratios follow, subject to modest outlier removal and sometimes transformation, the normal distribution. This is encouraging because it means that much of this earlier work can also be further applied to the insurance sector.

An important point to re-emphasise at this stage is the striking differences that normalisation and transformation make in even the most basic of conclusions from the data. Without normalisation, we would have concluded, incorrectly, that expense ratios and combined ratios would be positively related. With normalisation, we can make the correct inference which is that expense ratios and combined ratios are in fact negatively related. Similarly, without normalisation, we would have concluded that loss ratio and combined ratio have no relationship to return on assets. With normalisation

(and further helped by transformation), we can see that these ratios are in fact strongly, negatively correlated.

These entirely contradictory inferences demonstrate that researchers ignore a check of the normality assumption at their own peril. The recommendation must therefore be, in line with previous warnings issued by researchers in this area, that the effects of data transformation and normalisation procedures *always* be studied before conclusions from parametric tests are drawn. This is, of course, by no means a new insight, but the analysis in this paper brings its importance once again sharply into focus.

The study can be extended in three ways. The first area of extension is in the inclusion of more financial ratios. This study has focused on profitability ratios, but other ratios could have been included, particularly those related to solvency. Solvency ratios are of great practical importance to watchdogs and regulators. Following on from the 2008 financial crisis, the EU implemented new capital requirements for insurance companies (the Solvency-II directive). As they continue to be important variables of study, their cross-sectional distribution properties are of high relevance.

A second area of extension is in the check of the proportionality assumption. We have previously seen that the lack of normality in a distribution could be the result of a positive or negative intercept in the relationship between numerator and denominator of the financial ratio (Whittington, 1980; Barnes, 1987). A proportionality check may be included in the normalisation procedure applied in this study.

Finally, further research invariably points in the direction of multi-variate normality assumptions. This paper has confined itself to univariate normal distributions. Yet several tests, including the popular multiple discriminant analysis, rest not just on the individual, univariate normal distribution, but also on the multi-variate normality of ratio data series. Karels and Prakash (1987) and Watson (1990) provides an analysis of multi-variate normality of financial ratios. Future research is to examine these multi-variate normal distributions also in the insurance sector.

References

- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4), 589-609.
- Ambrose, J. M., & Carroll, A. M. (1994). Using Best's ratings in life insurer insolvency prediction. *Journal of Risk and Insurance*, 61(2), 317-327.
- Baranoff, E. G., Sager, T. W., & Witt, R. C. (1999). Industry segmentation and predictor motifs for solvency analysis of the life/health insurance industry. *Journal of Risk and Insurance*, 66(1), 99-123.

- Barnes, P. (1987). The analysis and use of financial ratios: a review article. *Journal of Business Finance and Accounting*, 14(4), 449-461.
- BarNiv, R. (1990). Accounting procedures, market data, cash-flow figures, and insolvency classification: the case of the insurance industry. *The Accounting Review*, 65(3), 578-604.
- BarNiv, R., & Hershbarger, R. A. (1990). Classifying financial distress in the life insurance industry. *Journal of Risk and Insurance*, 57(1), 110-136.
- BarNiv, R., & McDonald, J. B. (1992). Identifying financial distress in the insurance industry: a synthesis of methodological and empirical issues. *Journal of Risk and Insurance*, 59(4), 543-574.
- Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 4, 71-111.
- Deakin, E. B. (1976). Distributions of financial accounting ratios: some empirical evidence. *The Accounting Review*, 51(1), 90-96.
- Ezzamel, M., Mar-Molinero, C., & Beecher, A. (1987). On the distributional properties of financial ratios. *Journal of Business Finance and Accounting*, 14(4), 463-481.
- Fieldsend, S., Longford, N., & Mcleay, S. (1987). Industry effects and the proportionality assumption in ratio analysis: a variance component analysis. *Journal of Business Finance and Accounting*, 14(4), 497-517.
- Frecka, T. J., & Hopwood, W. S. (1983). The effects of outliers on the cross-sectional distributional properties of financial ratios. *The Accounting Review*, 58(1), 115-128.
- Karels, G. V., & Prakash, A. J. (1987). Multivariate normality and forecasting of business bankruptcy. *Journal of Business Finance and Accounting*, 14(4), 573-593.
- Lee, C. (1985). Stochastic properties of cross-sectional financial data. *Journal of Accounting Research*, 23(1), 213-227.
- Lee, S. H., & Urrutia, J. L. (1996). Analysis and prediction of insolvency in the property-liability insurance industry: a comparison of logit and hazard models. *Journal of Risk and Insurance*, 63(1), 121-130.
- Lev, B., & Sunder, S. (1979). Methodological issues in the use of financial ratios. *Journal of Accounting and Economics*, 1(3), 187-210.
- Martikainen, T., Perttunen, J., Yli-Olli, P., & Gunasekaran, A. (1995). Financial ratio distribution irregularities: implications for ratio classification. *European Journal of*

- Operational Research*, 80, 34-44.
- McLeay, S. (1986). Student's t and the distribution of financial ratios. *Journal of Business Finance and Accounting*, 13(2), 209-222.
- McLeay, S., & Omar, A. (2000). The sensitivity of prediction models to the non-normality of bounded and unbounded financial ratios. *British Accounting Review*, 32, 213-230.
- R Development Core Team. (2010). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Available from <http://www.R-project.org> (ISBN 3-900051-07-0)
- So, J. C. (1987). Some empirical evidence on the outliers and the non-normal distribution of financial ratios. *Journal of Business Finance and Accounting*, 14(4), 483-496.
- Thomson Financial. (2003). *Worldscope database: datatype definitions guide*. Thomson Financial.
- Urbanek, S. (2010). rJava: Low-level R to Java interface [Computer software manual]. Available from <http://CRAN.R-project.org/package=rJava> (R package version 0.8-7)
- Watson, C. J. (1990). Multivariate distributional properties, outliers, and transformation of financial ratios. *The Accounting Review*, 65(3), 682-695.
- Whittington, G. (1980). Some basic properties of accounting ratios. *Journal of Business Finance and Accounting*, 7(2), 219-232.