UNIVERSITY OF GOTHENBURG

SCHOOL OF BUSINESS, ECONOMICS AND LAW

How to Differentiate the Difference

A quantitative study on the determinants of discrepancies between expected and actual credit ratings in insurance companies

Master Thesis

M.Sc. Accounting and Financial Management

Spring 2020

Gustav Hellman

Henrik Walldén Persson

Supervisor:

Mari Paananen

Examiner:

Jan Marton

ABSTRACT

We study the determinants of discrepancies between expected and actual credit ratings among insurance companies. We analyze 124 public insurance companies with an assigned credit rating, and model discrepancies as the difference between expected and actual credit ratings. We find relatively strong evidence that embedded value has no association with a positive or negative difference, and that embedded value facilitates more accurate credit ratings. We find weak evidence supporting that earnings management is associated with overestimated ratings relative to financial strength. Firms reporting under IFRS are found to be significantly associated with overestimated ratings. Interpretations suggest that this relationship is explained by the opportunistic and discretionary nature of IFRS 4. Finally, we find that life insurers exhibit overestimated ratings. Life insurers' financial statements are underlined to be difficult to assess, and in that, profitability is hard to derive. Increased complex risk exposure for life insurers might also entail that rating agencies cannot, or do not, acknowledge the actual risk exposure of life insurers. Determinants of discrepancies between expected and actual ratings have not been addressed until now. As such, our findings have apparent benefits for users of financial statements and for rating agencies, as well as users of credit ratings. We contribute not only by filling this gap, but to direct future research towards exploring this framework further.

KEYWORDS: Expected ratings; credit ratings; financial strength; discrepancy; earnings management; embedded value; IFRS; life insurers.

ACKNOWLEDGEMENTS

We would like to acknowledge and thank our supervisor Mari Paananen for her guidance, helpful inputs and for challenging us throughout the entire process. An additional recognition to Söderberg & Partners is also deserved. We are grateful for the opportunity and for the inputs along the way. We would also like to thank our seminar leader, Jan Marton, for the feedback and advice received in the writing process.

Gothenburg, 2020-06-10

Gustav Hellman

Henrik Walldén Persson

TABLE OF CONTENTS

LIST OF TABLES

1. INTRODUCTION

1.1 BACKGROUND

The insurance sector is of high importance to all developed economies, sharing risk and offering individual protection. Through the application of mathematics and statistics, insurance companies transfer and pool risks to protect individuals from risk exposure. This diversification facilitates greater forward-planning certainty, as well as more certainty in consumption and investment behavior (Deutsche Bundesbank, 2019). Le and Xu (2018) emphasize the importance of stakeholders' ability to assess future solvency in insurance companies, underlining that the actual service period of the products often is several years after the purchase, making the assessment of future solvency important, especially for policyholders' purchase decision. Such an instrument to help stakeholders assess future solvency is a credit rating, a widely used indicative measure of financial strength not only in the present but some years ahead. To provide valuable information for investors, the importance of rating agencies and credit ratings to accurately assess the financial position of banks, bonds, insurance companies and other issuers of debt cannot be understated (Partnoy, 2002).

However, an extensive line of research is dedicated to scrutinize the accuracy of credit ratings relative to financial strength, and researchers such as Galil (2003), and Hilscher and Wilson (2016) reach the conclusion that the actual default risk is higher than what is incorporated in ratings. Implying that firms might be assigned higher credit ratings not equivalent to their financial strength. As such, Becker and Milbourn (2011) highlight that firms prefer favorable rating over accurate ratings, derived from the increased access to capital markets that higher ratings entail (Angell et al., 2000). Additionally, rating agencies are profit-driven and have incentives to assign favorable ratings since dissatisfied firms might switch agency (Mutize, 2019, June 23). This conflict of interest was evident when rating agencies, such as Standard and Poor's (S&P) and Moody's, played a major part in the financial crisis and the collapse of the US economy due to inaccurate ratings (Krantz, 2013, September 13).

The Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB) have tried to unify an accepted standard for insurers (El-Gazzar, Jacob & McGregor, 2014), yet the accounting quality in insurance companies is rather dispersed. Correspondingly, Chin (2015) examines the relationship between accounting quality and the accuracy of credit ratings and show that higher accounting quality is associated with higher accuracy of credit ratings. Additionally, for life insurers in particular, Horton (2007) underlines the difficulty of determining what drives profitability and in distinguishing between earnings connected to past and current performance, due to discretionary reporting. As for property and casualty insurers, Cagle (1996) highlights that the expected future claims usually are insurers' largest liability and argues that the uncertainty of these reserves lies in the difficulty in precisely forecasting the value. This implies that getting a fair and representative view of insurance companies is difficult.

Indeed, not only do rating agencies exhibit a form of conflict of incentives, but face both financial reporting subject to uncertainty as well as deceiving issuers contending to receive the most favorable rating possible inequivalent to their financial strength (Liu, Subramanyam, Zhang & Shi, 2018; Alissa, Bonsall, Koharki & Penn Jr, 2013). Evidently, the credit rating process is a complex interaction between rating agencies and the issuers, both with unique agendas and potentially conflicting incentives. The result is that credit ratings risk deviate from the financial strength. Meanwhile, a credit rating is an important indicator of financial strength, and an indicator many stakeholders rely on. It is therefore of interest to examine what factors really drive the discrepancies between actual ratings and financial strength, modelled as expected ratings.

1.2 PROBLEM DISCUSSION

Research on the determinants on credit ratings derived from financial strength are in abundance (e.g. Caporale, Cerrato & Zhang, 2017; Adams, Burton & Hardwick, 2003; Van Gestel et al., 2007; Angell et al., 2000). While most researchers focus only on the determinants affecting credit ratings, broad indications as to the accuracy of ratings relative to financial strength are to be found. Van Gestel et al. (2007), for example, forecast and analyze insurance companies' ratings, in which a set of explanatory variables together form the rating model approach to estimate ratings. Their findings indicate that compared to external ratings, 75 percent of the estimated ratings differ at most one notch. In a similar manner, Alissa et al. (2013) construct an empirically modelled expected credit rating proxy, and a distribution of actual ratings versus expected ratings show that most actual ratings are at-expected. This means that actual and expected ratings are leveled and most predictions within each rating class differ at most one notch from at-expected rating.

Seemingly, several circumstances regarding rating agencies ability to assess insurance companies are indeed subject to uncertainty. Insurers are shown to often deal with long time-horizons and uncertainty regarding potential claims, and as a result, the loss reserve usually is an insurer's largest liability (Cagle, 1996). The particular uncertainty of these reserves stem from the difficulty in precisely forecasting the value, but also the manipulation of accounting statements by managers (1996). Assessing the financial strength of insurance companies, and in that, assigning a correct rating is therefore by no means without uncertainty. This issue is strengthened by the fact that rating agencies mainly rely on the information provided by the firms being rated, an environment that potentially facilitates discretion (Demirtas & Cornaggia, 2013). As such, Beaver, McNichols and Nelson (2003) show that earnings management is common in insurance companies, and in that, particularly understating loss reserves. Indeed, increases in loss reserves may potentially reduce the firm's market value through negative signaling (Cagle, 1996), highlighting managers' aim of being assigned higher ratings (Kisgen, 2006). Likewise, Alissa et al. (2013) state that a firm's credit rating is the focal point in capital structure decisions, as discrepancies from expected ratings may affect a firm's access to capital. Accordingly, findings show a positive association between income-increasing earnings management and future credit ratings. Consequently, as firms use discretion to be portrayed in a favorable way, the transparent overview of the firm is reduced. Another apparent problem concerns IFRS 4 and its guidance regarding the reporting of insurance contracts, where insurers are able to omit relevant information about insurance contracts without reprimands, further reducing transparency and comparability between firms (IASB, 2017).

Evidently, most credit ratings are shown to be rather accurate relative to a firm's financial strength. However, there is still some discrepancy between actual ratings and financial strength. Clearly, apparent incentives show why firms prefer favorable ratings, and there is a lot of uncertainty as to the rating agencies' possibility to accurately assess insurance companies. Whereas research show a good understanding of how determinants of financial strength affect credit ratings, what has not been addressed are the determinants of the *difference* between actual ratings and financial strength, a gap this thesis aims to fill.

1.3 AIM AND PURPOSE

The aim of this study is to shed light on what determines the discrepancies between expected and actual credit ratings among insurance companies, with the purpose to enlighten users with factors to consider when assessing ratings, and to direct future research towards exploring such framework further.

1.4 RESEARCH QUESTION

What are the determinants of discrepancies between expected and actual credit ratings in insurance companies?

1.5 DELIMITATIONS

Our sample is limited to public insurance companies with assigned ratings. Private insurance companies rarely have an assigned credit rating, and data availability is much lower. As such, most existing insurance companies were excluded in our analysis. We also chose to only include companies rated by Standard & Poor's, as they have the broadest coverage of rated companies. As for financial strength determinants, credit rating agencies also incorporate qualitative information in their rating assessment. We focus our estimation of financial strength using only financial variables, and qualitative information is excluded.

2. THEORETICAL FRAMEWORK AND HYPOTHESIS DEVELOPMENT

Highlighted by prior research, the relationship between credit ratings and financial strength is in some settings strong, but in others highlighted to be of some concern. Indeed, credit ratings should reflect financial strength rather well, yet perfect collinearity is both unrealistic and shown to not be the case. The interesting question is therefore what drives the difference between credit ratings and financial strength, or put in our setting, what the determinants of discrepancies are between expected and actual credit rating. These are discussed below.

2.1 EMBEDDED VALUE

In determining what drives the difference between actual and expected rating, the emphasis is not on what is associated with a higher or lower rating, rather on what makes assigning an accurate rating particularly difficult. In that, traditional indicators of firm value have been criticized as they do not account for expected future profits (Wu & Hsu, 2011). Timing differences when recognizing revenue and expenses also cause a mismatch between economic and accounting profits for each reporting period (2011). Thus, life-insurers have increasingly turned to evaluating financial performance using embedded value (Wu & Hsu, 2011), aiming to improve the transparency. Embedded value includes several items, being the present value of future shareholders' cash flow from operations (PVIF) and required capital minus the cost of maintaining the capital plus additional excess allocated to the business (El-Gazzar et al., 2014). In short, it consists of the present value of future net cash flows from in-force life insurance business (McGregor, Jacob & El-Gazzar, 2013). The advancement of embedded value was largely driven by the mismatch of accounting results and performance because of delayed recognition and discrepancy in valuation for assets and liabilities (2013). Thus, it recognizes future changes in accounting earnings, and proponents argue it is a superior base in the valuation of life insurers compared to traditional accounting measures (El-Gazzar et al., 2014). As such, reporting embedded value should increase transparency among insurers.

Furthermore, other forms of insurers have also turned to embedded value, and research have proposed approaches in transferring market-consistent embedded value (MCEV) from life to non-life insurers (Diers, Eling, Kraus & Reuß, 2012). Indeed, research show that embedded value increases the transparency of the firms' performance, thereby highlighting the value relevance of embedded value (Préfontaine, Desrochers & Godbout, 2009; Almezweq & Liu, 2012). Serafeim (2010) further show the value relevance of embedded value, as findings suggest that embedded value reporting is associated with less information asymmetry. Likewise, El-Gazzar et al. (2014) find that embedded value disclosure help investors in assessing life insurers. We therefore expect that firms that report embedded value are more transparent and therefore facilitate a lower difference between actual and expected ratings. The following alternative hypothesis is therefore formulated:

H1: Other things being equal, Embedded Value reporting is not associated with a positive (negative) difference between expected and actual credit ratings.

2.2 EARNINGS MANAGEMENT

Indeed, distorting earnings with discretionary decisions yields a less transparent financial overview of the firm, and an accurate rating may be difficult to assign. Since rating agencies are said to rely on the information provided by the firms being rated, it is problematic if the information potentially facilitates discretion (Demirtas & Cornaggia, 2013). Song (2018) offers a thorough review on earnings management and show several findings. The main ones suggest that insurance companies smooth their earnings using loss reserves, underestimate the reserve when financially weak and engage in income-increasing discretion when managers are exposed to stock price volatility. Similarly, Beaver et al. (2003) find that regardless of insurers' financial strength, earnings management activities are common, and particularly in understating loss reserves. Moreover, increases in loss reserves may have implications on firms as it omits adverse information, potentially reducing the firm's market value through negative signaling (Cagle, 1996) and potentially affecting the rating level as well.

Further, Liu et al. (2018) compare discretionary accruals (DAC) and non-discretionary accruals (NDAC) and study whether a potential downgrade is favorably influenced by earnings management. The authors find that both measures are explanatory, and in comparing high with low levels of DAC, they find that there is a 24 percent lower probability of downgrade in the former. Interpretations of the results reads that DAC and NDAC are not differentiated by rating agencies, which in turn implies that earnings management is favorable in terms of avoiding a rating downgrade. Similarly, Alissa et al. (2013) find a positive association between income-increasing earnings management and future credit ratings. Indeed, the overall quality of a firm should be reflected in the assigned credit rating (Kisgen, 2006). However, if firms use discretion, assigning a rating reflecting the overall quality might be more difficult. We therefore expect that firms that participate in higher levels of earnings management facilitate a higher difference between expected and actual rating. The following alternative hypothesis is formulated:

H2: Other things being equal, accrual Earnings Management is associated with a negative difference between expected and actual ratings.

2.3 REPORTING STANDARD

In addition, reporting standards (i.e. U.S. GAAP vs. IFRS) differ in the level of discretion and freedom of interpretation, and may explain discrepancies between expected and actual ratings. Although attempts have been made between FASB and IASB to develop a unified and accepted accounting standard for life insurers (El-Gazzar et al., 2014), the interpretative discretion still differs between IFRS and U.S. GAAP. The authors examine embedded value disclosure and the valuation effects it has on firms reporting under U.S. GAAP. U.S. capital markets are underlined to be under more scrutiny by regulators and the legal environment and exhibit more transparency. Therefore, they argue that the information provided by embedded value reporting is possibly already reflected under U.S. GAAP measurements (2014).

Literature has highlighted that IFRS facilitate adoption in the individual case and are therefore more subjective than the objective U.S. GAAP (Runesson, Samani & Marton, 2018). Although it is underlined that IFRS facilitate either better reporting in reflecting underlying fundamentals due to more freedom of interpretation, or greater room for earnings management derived from the same reason, managerial discretion is generally higher under IFRS compared to U.S. GAAP (Evans, Houston, Peters & Pratt, 2012). Further, Gerstner, Lohmaier and Richter (2015) state that the fair values of insurance contracts are hard to decipher as they are not traded on the capital markets, instead requiring assumptions and derivation using models. As such, the current standard for insurance contracts, IFRS 4, was developed in order to enhance the accounting quality regarding insurance contracts. However, IASB (2017) underlines the following:

"... IFRS 4 states explicitly that a company is not required to ensure that its accounting policies for insurance contracts are reliable or relevant to the economic decision-making needs of users of its financial statements, such as investors and analysts." (p. 2)

For investors and analysts in particular, this is troublesome as insurers are able to omit important information about insurance contracts, increasing uncertainty and in turn affecting decision-making (IFRS, 2017). Another issue surrounding IFRS 4 is that firms can adopt accounting practices based on their jurisdiction, further complicating comparability and transparency (2017).

Research have further shown that, generally, earnings management is more common in non-U.S. firms than U.S. and U.S. GAAP firms (Alford, Jones, Leftwich & Zmijewski, 1993; Land & Lang, 2002; Lang, Raedy & Yetman, 2003). Although the more effective enforcement for U.S. and U.S. GAAP firms is commonly employed in describing lower earnings management for said firms (Evans et al., 2012), this must not be the case. Evidently, the environment of more effective enforcement encourages real over accruals earnings management, and lower earnings management could be mistaken for a substitution of real instead of accruals earnings management (2012). Accordingly, Barth, Landsman, Lang and Williams's (2012) finding of higher accounting quality in U.S. and U.S. GAAP firms relative to non-U.S. and IFRS firms is questioned as the substitution from accrual to real earnings management, which is harder to detect, might offset this conclusion (Evans et al., 2012). Nonetheless, we expect that IFRS, in facilitating higher discretion, is associated with a higher likelihood of (accruals) earnings management, and therefore making rating agencies less able to assign an accurate rating. The following alternative hypothesis is formulated.

H3: Other things being equal, firms reporting under IFRS are associated with a positive (negative) difference between expected and actual ratings.

2.4 LIFE INSURERS

Although previously argued that embedded value is associated with a lower discrepancy between actual and expected rating, embedded value for life insurers has been criticized of failing to meet the definition of a recognizable asset (Klumpes, 2002). Additionally, voluntary disclosures of embedded value by life insurers tend to be associated with higher future profit expectation than those that do not disclose embedded value (2002). Correspondingly, Klumpes (2002) argue that life insurers also use embedded value as a signaling tool, rather than as an attempt to increase transparency. Research has also highlighted the apparent linkage between life insurance and financial markets (de Bandt & Overton, 2019), as well as sensitivity towards long-term interest rates, personal income and unemployment rate (Browne, Carson, & Hoyt, 1999). Accordingly, in establishing that the health of insurance companies fluctuates with the macroeconomic environment, life insurers exhibit more interconnection with the macroeconomy (EIOPA, 2018). Whereas life insurers previously were seen as a conservative investment with a reputation of unparalleled financial stability, the perception has since vanished (Fenn & Cole, 1993). De Bandt and Overton (2019) acknowledge that decreasing interest rates had led life insurers to seek riskier investment to achieve higher yields, leaving them more vulnerable to macroeconomic shocks. Koijen and Yogo (2017) recognize similar trends in that traditional risks for life insurers have expanded into more complex and opaque ones over the last decade. This is further highlighted by Gerstner et al. (2015), who states that the exact value of life insurers' financial statements is difficult to grasp because of their long-term and complex financial contracts.

Horton (2007) further underlines the difficulties of life insurances, mentioning that under IFRS 4, life insurers tend to delay profit recognition and distribution, understating future premiums and overstating claims. Consequently, insurance policies are initially understated, and as the policies mature, higher profits are realized. A recurring stream of continuously overlapping insurance contracts with different maturities might therefore complicate a fair view of the firm, with a mixture of over- and understated insurance policies. Accordingly, Horton (2007) underlines that users are unable to determine what drives profitability and to distinguish between earnings related to past or current performance. We therefore expect life insurers to be difficult to rate, and therefore hypothesize a higher discrepancy between expected and actual ratings. The following alternative hypothesis is formulated:

H4: Other things being equal, Life insurers are associated with a positive (negative) difference between expected and actual ratings.

3. METHODOLOGY

3.1 SAMPLE SELECTION

Insurance companies were identified using S&P's Capital IQ. Retrieving all insurance companies in the world yielded 47249 companies. We then excluded private companies, a query that yielded 2762 companies. Public insurance companies were then screened for assigned ratings. Total public insurance companies with assigned ratings amounted to 471 companies. We then manually screened for potential misclassified firm types, resulting in 330 firms excluded from the sample. After adjusting the sample for companies exhibiting missing data, our final sample resulted in 124 insurance companies. Table 1 below shows this derivation.

Table 1. Sample Selection

Sample selection	# of companies
Insurance companies identified with Capital IQ	47 249
Private insurance companies	- 44 487
Public insurance companies	2 762
Insurance companies without rating	- 2 291
Public insurance companies with rating	471
Institutions misclassified	- 330
Confirmed public insurance companies with rating	141
Unavailable data	- 17
Final Sample	124

3.2 DATA COLLECTION

Using S&P's Capital IQ and Market Intelligence, we collected fiscal year data from 2009 to 2019 on public insurance companies. To identify the financial strength determinants most commonly used when analyzing credit ratings, prior research was reviewed. The following determinants used to proxy expected ratings were collected using Capital IQ; Return on equity, leverage, solvency ratio, loss ratio and total assets. In collecting historical credit ratings as well as discrepancy determinants EV, EM and IFRS, Market Intelligence was used. Further, discrepancy determinant LIFE was identified using Capital IQ.

3.3 MEASURE OF EXPECTED CREDIT RATING

3.3.1 DETERMINANTS AFFECTING FINANCIAL STRENGTH

As stated by S&P "A company's financial reports are the starting point for the financial analysis of a rated entity" (2008. p, 22). A well-structured proxy for financial strength, therefore, should lay a solid foundation for the predictiveness of financial strength. We motivate our proxies for expected rating using determinants from prior research. Although research covering the relationship between credit ratings and financial strength are in small numbers, studies proxying and forecasting financial strength are many. Accordingly, as we model insurers' financial strength using key fundamental metrics of firm performance, consequently constructing expected ratings, this section reviews commonly employed and explanatory determinants used in previous research. The chosen explanatory variables for analyzing credit ratings are return on equity, leverage, solvency ratio, loss ratio and the natural logarithm of total assets. For each chosen determinant, an expected sign is given.

3.3.1.1 PROFITABILITY

Current research showing the importance of profitability in determining financial strength has done so in different forms and by defining profitability at different levels. Focusing on the firm's core business, operating margin is a common indicator of financial strength. Chen and Wong (2004) show that the operating margin, for both life insurers and property-liability insurers, is an important determinant. The authors are joined by other researchers with a general view on credit ratings, such as Afik, Bouchnik and Galil (2016) and Du (2003). Both Adams et al. (2003) and Caporale et al. (2017) use profitability as an indicator of financial strength. In a similar manner, Van Gestel et al. (2007) use profitability as a measure to explain the company's ratings. Their findings suggest that profitability is a main driver in understanding credit ratings for insurance companies. In assessing insurance firm's financial strength (Yakob, Yusop, Radam & Noriszura, 2012) and in asking why insurers fail (de Bandt & Overton, 2019), return on equity (ROE) is shown to have explanatory power. De Bandt and Overton (2019) show that ROE and ROA is lower for firms that eventually fail. The reverse logic is that higher ROE and ROA imply a lower risk of failure and therefore higher financial strength. Profitability, measured as net income over equity (ROE), is our first determinant of financial strength, and we expect it to be positively associated with ratings.

3.3.1.2 LEVERAGE

Leverage is an important factor and indicator of financial strength, since higher leveraged firms are less able to absorb unexpected losses (Grace et al., 2003). Consequently, the majority of literature on financial strength, credit rating or default risk use leverage as a determinant. Two different major uses of leverage is defining leverage as a firm's debt to equity (see Malik, 2011; Blums, 2003; Boguslauskas, Mileris & Adlytė, 2011), and a firm's total liabilities to total assets (see Afik et al., 2016; Du, 2003; Damodaran, 2015; Shumway, 2001; Hilscher & Wilson, 2016; Ohlson, 1980; Boguslauskas et al., 2011; Giordani, Jacobson, von Schedvin & Villani, 2011). Leverage, measured as total debt to equity, is our second determinant of financial strength, and we expect it to be negatively associated with ratings.

3.3.1.3 LIQUIDITY

Caporale et al. (2017), argue that the financial strength of a firm is affected by liquidity. Since insurance companies deal with high levels of uncertainty, liquidity plays an important factor in determining the financial strength of insurers (Kwon & Wolfrom, 2016). Adams et al. (2003) find that liquidity has a significantly positive effect on ratings. Giordani et al. (2011) also look at liquidity in their attempt to predict firm bankruptcy risk by placing cash and liquid assets in relation to total liabilities. Boguslauskas et al. (2011) further use liquidity in the form of cash in relation to current liabilities in assessing credit risk. Measuring how many times a firm's income, defined at different levels, exceeds its interest expense, the interest coverage ratio, or solvency ratio, provides a view of the firm's riskiness, both in relation to current debt but also for future borrowing. Interest coverage has been used with different definitions. Damodaran (2015), for example, points out that both EBITDA, interest coverage and pre-tax interest coverage are main determinants of a firm's credit rating. Likewise, Du (2003) shows that a firm's pre-tax interest coverage has a positive effect on a firm's credit rating, which intuitively implies a higher financial strength as well. Solvency ratio, measured as EBITDA to interest expense, is our third determinant of financial strength, and we expect it to be positively associated with ratings.

3.3.1.4 LOSS RATIO

The determinants mentioned above are indeed generic metrics of financial strength, but nonetheless important. With the focus on insurance companies, capturing the structure of revenue and cost is important. Insurance companies do not report the generic revenue and cost of goods or services sold, as many other firms do. Instead, what is earned is premiums collected from policyholders, and the main costs are claims from policyholders. Researchers have accounted for the peculiar structure of insurance companies in many different forms, attempting to capture the performance of premiums written or earned, in relation to claims or losses. This is done with great variation. Chen and Wong (2004) look at underwriting results, and measure its performance using the combined ratio. This is defined as the ratio of incurred losses to earned premiums plus incurred expenses to written premiums (2004). The authors are joined by de Bandt and Overton (2009) who look at operating expenses over gross premiums, and Yakob et al. (2012) with a slightly similar measure of net operating expenses to net premiums written. Van Gestel et al. (2007) employ a ratio of net claims over net premiums earned, called loss ratio, measuring the cost of risk. The latter is employed as our fourth determinant, and we expect it to be negatively associated with ratings.

3.3.1.5 FIRM SIZE

The size of the company is another common determinant when measuring financial strength. Caporale et al. (2017) show that firm size is positively related to leverage, whereas it is negative in relation to cash ratio, suggesting that large firms hold less cash and are more leveraged. However, size is argued to be associated with benefits from economies of scale (Adams et al., 2003; van Gestel et al., 2007). Building on Bouzouita and Young's (1998) findings that large insurers are more likely to be rated and have lower insolvency risk than smaller insurers, Adams et al. (2003) hypothesize that the larger the size of the company, the higher the rating of the insurer. Furthermore, Alissa et al. (2013) examine the relationship between earnings management and credit ratings and define size as the natural logarithm of total assets. We follow Alissa et al. (2013) and define firm size as the natural logarithm of total assets as our fifth determinant of financial strength, and we expect it to be positively associated with ratings.

Table 2 below summarizes each variable's definition and expected sign.

Table 2. Summary of Determinants

Determinants	Expected Sign	Definition
RETURN ON EQUITY (ROE)	Positive (+)	Net Earnings/Equity
LEVERAGE (LEV)	Negative (-)	Debt/Equity
SOLVENCY RATIO (SOLV)	Positive (+)	EBITDA/Interest Expense
LOSS RATIO (LOSS)	Negative (-)	Net Claims/Net Premiums Written
FIRM SIZE (SIZE)	Positive (+)	Natural Logarithm of Total Assets

3.3.2 ORDERED PROBIT REGRESSION MODEL

As described above, we follow prior research and employ a selection of financial variables as a proxy for insurance companies' financial strength. As credit ratings are ordered and also discretely specified, an ordered probit model is suitable in the analysis (Amato & Furfine, 2004), accounting for the dependent variable's ability to take on a discrete, ordered format with more than two categories. The ordered probit model calculates a probability for each observation of independent variable to fall into a certain category of the dependent variable. Estimation of the model is made using maximum likelihood, or MLE, where interpretations are based on log-likelihood. The ordered probit model extends the general definition of the probit model by accounting for the ordering of the dependent variable, while constructing N-1 cut-offs, or thresholds, which defines a likelihood category of Y based on values of z. For rating categories n = 1,...., N and with N-1 thresholds, the following general definition applies:

$$f(x) = \begin{cases} 1\ if z_i \leq t_1 \\ 2\ if t_0 < z_i \leq t_2 \\ 3\ if t_1 < z_i \leq t_3 \\ \quad \cdots \\ N\ if z_i > t_{N-1} \end{cases}$$

where z is a vector of control variables generating values of y, in this case ACTUALRATING, and is the following function of the independent variables:

$$ACTUALRATING_{it} = \beta_0 + \beta_1 ROE_{it} + \beta_2 LEV_{it} + \beta_3 SOLV_{it} + \beta_4 LOSS_{it} + \beta_5 SIZE_{it} + \varepsilon_{it} \quad Eq. (1)$$

where ACTUALRATING is an ordinal variable representing S&P's credit ratings and therefore taking on values from 1 to 14, where B- represents 1 and AA represents 14, and where ROE is return on equity, LEV is leverage, SOLV is solvency ratio, LOSS is loss ratio and SIZE is firm size, which collectively represent our determinants of financial strength. Following Alissa et al. (2013), the proxy for expected credit ratings was constructed using the rating category with the highest fitted probability from Eq. (1), and was estimated cross-sectionally by year, avoiding potential look-ahead bias. The resulting estimated variable is *expected rating*. Like *actual rating, expected rating* is classified in an ordinal manner, from the lowest rating 1 (B-) to the highest 14 (AA).

3.4 ORDINARY LEAST SQUARES REGRESSION MODEL

Having estimated each firms' expected rating each year, expected rating was subtracted with the associated actual rating, resulting in the *difference* between them. This was done for four different time intervals, from expected rating$_t$ - actual rating$_t$ up to expected rating$_{t+3}$ - actual rating$_t$, to capture not only the contemporaneous effect but the consistency over time as well. Four discrepancy determinants (EV, EM, IFRS and LIFE) were hypothesized. First, Embedded value (EV) has been found to increase transparency (Wu & Hsu, 2011) and reduce asymmetric information (Serafeim, 2010), which is further strengthened by Beaver et al. (2003) who underline stakeholders' increased ability to assess firms that report EV. Creating a dummy variable equal to 1 if EV is reported, we hypothesized EV to facilitate a more accurate rating assessment, and therefore to not be associated with either a negative of positive *difference*. Earnings management (EM) was hypothesized to be negatively associated with *difference*, as firms are shown to receive favorable ratings when participating in earnings management (Liu et al., 2018; Alissa et al., 2013). Different measures of earnings management have been proposed (see Beaver et al., 2003; Alissa et al., 2013). We define earnings management as the change in total loss reserves over total liabilities prior to year$_{t-1}$, in accordance with anecdotal evidence[1]. IFRS has furthermore been, relative to U.S. GAAP, underlined to be more subjective (Runesson et al., 2018) of higher managerial discretion (Chiu, 2016; Evans et al., 2012), and associated with lower accounting quality than US listed firms reporting under U.S. GAAP (Barth et al., 2012). By creating a dummy variable equal to 1 if firms report under IFRS, the variable was hypothesized to be positively (negatively) associated with *difference*. LIFE was finally hypothesized to be negatively (positively) associated with *difference*, due to increased

[1] Evidence put forward to us in conversation with a representative from the standard setters.

exposure to complex and opaque risk (Koijen & Yogo, 2017), and that users are unable to determine what drives profitability, and are unable to distinguish between earnings related to past or current performance (Horton, 2007). In accordance with Serafeim, 2010), a dummy variable was created equal to 1 if life insurance operations were mentioned in the extended business description. Table 3 below summarizes our hypotheses.

Table 3. Summary of Hypotheses

Variable	Expected Sign	Hypothesis
EMBEDDED VALUE (EV)	No sign	H1
EARNINGS MANAGEMENT (EM)	Negative (-)	H2
IFRS (IFRS)	Positive/Negative (+/-)	H3
LIFE INSURANCE (LIFE)	Positive/Negative (+/-)	H4

To test the relationship between the *difference* between expected and actual ratings with the determinants of discrepancies, ordinary least square (OLS) regressions were performed, following the definition below:

$$DIFF_t = \beta_0 + \beta_1 EV_{it} + \beta_2 EM_{it} + \beta_3 IFRS_{it} + \beta_4 LIFE_{it} + \beta_5 MTB_{it} + \beta_6 CAT_{it}$$
$$+\beta_7 SPV_{it} + \beta_8 INV_{it} + \beta_9 SEG_{it} + \mathcal{E}_{it}$$
$$Eq.(2)$$

where DIFF is the difference between expected and actual rating and control variables include: MTB, the level of market-to-book value of equity; CAT, whether catastrophic events have occurred; SPV, the level of stock price volatility; INV, the level of investment yield; and SEG, the number of business and/or geographic segments. These are specified below. All continuous variables were winsorized at the 1st and 99th percentiles.

3.4.1 CONTROL VARIABLES

Market-to-book (MTB): Measured as price per share over book value, market-to-book is used as a proxy for market valuation, controlling for other factors than company fundamentals. This measure is used by Alissa et al. (2013) as explanatory variable in estimating expected rating, whereas Bonsall et al. (2017) use it as a control variable. We follow the latter since MTB in our estimation of expected credit ratings was dropped due to multicollinearity in the ordered probit model. *Catastrophic events (CAT)*: Njegomir and Marovíc (2012) argue that insurance companies operate with an increased probability of the unexpected to happen, with emphasis on catastrophic events. We therefore construct a dummy variable equal to 1 if a disaster took place, and 0 otherwise, possibly affecting the accuracy of ratings relative to financial strength. *Stock Price Volatility (SPV):* The riskiness of firms indeed affects the level of ratings, but it might also impact the accuracy of ratings relative to financial strength, due to higher uncertainty. Stock return is recurrently shown in research to be associated with financial risk (see Steiger, 2010; Mwaurah, Muturi & Waititu, 2017; Kang & Kang, 2009 and Babi, 2015). In controlling for stock price volatility in proxying credit risk, stock price volatility is calculated as the five-year standard deviation of stock returns, and included as a control variable, similar to Alissa et al. (2013). *Investment yield (INV)*: All insurance companies collect premiums, yet life insurance companies differ from other forms of insurance companies, mainly in their cash flow structure and the duration of assets and liabilities (Diers, Eling, Kraus & Reuß, 2012), and for having particularly long-term maturity on claims. Steady long-term return on investments are therefore crucial to cover long-term claims. Meanwhile, life insurers, compared to non-life insurers, can more easily predict future claims, and a good indicator of performance of life and health insurers is investment yield (van Gestel et al., 2007). These differences are therefore controlled for. Investment yield, defined as investment income over investment assets, is used as a control variable and encompasses an important performance aspect of life insurers. *Segments (SEG)*: Firm complexity could potentially lead to an increased difficulty of assigning an accurate credit rating, and aggregating number of business and/or geographic segments is a common indicator of firm complexity (see Markarian & Parbonetti, 2007; Barinov, Park & Yildizhan, 2016; Gordon, Loeb & Tseng, 2009). Controlling for firm complexity, number of segments for each firm were manually derived from Capital IQ. Country and firm fixed effects are also controlled for.

3.5 RELIABILITY AND VALIDITY

For this study to be reliable, it is of high importance that the same results can be reached if repeated, given that the same condition applies (Bryman & Bell, 2007). Our sample of insurance companies is objectively collected with the definition required in the current study. A key foundation in our study is the estimation of our expected credit ratings, which was constructed in accordance with prior research. Indeed, the methodology and data management have been provided with transparency and objectivity, suggesting that future researchers, if replicating, would reach the same results. For reliability to be fulfilled, the study also needs to be valid. To ensure that, it is necessary that the study captures what it is supposed to (Bryman & Bell, 2007). As such, factors determining discrepancies between expected and actual ratings are motivated based on literature, and we control for other factors possibly affecting the robustness. It is not established what affects the difference, so there is a risk of excluding significant variables. However, research have underlined factors either facilitating or obstructing the credit rating assessment and known factors that affect the accuracy of credit ratings are included in our analysis. Our dependent variable, showing the difference between expected and actual ratings, is furthermore appropriately formulated in measuring what we aim to measure. Time and country effects are also controlled for, ensuring that the results do not depend on factors outside of the variables of interest. Robustness checks were furthermore performed, and both Pearson correlation and VIF tests suggest multicollinearity is not an issue in our study.

3.6 LIMITATIONS

Our sample consists of 124 confirmed public insurance companies with ratings, out of about 47 000 total insurance companies worldwide. This implies that the conclusions drawn from this study indeed are not generalizable to all insurance companies but is limited to those that are public with an assigned rating. However, as they are our focal entity of study, generalizable conclusions drawn are applied to those insurance companies of interest. Additionally, research on financial strength determinants is vast and extensive evidence for several determinants to be explanatory exists. As such, the decision of each determinant of financial strength is indeed based on its frequency in prior research, but also on data availability and ensuring low multicollinearity with other determinants chosen. Correspondingly, the estimated expected ratings might yield other results if other determinants are used, consequently affecting the accuracy of actual credit ratings. Clearly, time and scope of this study also limits the amount of discrepancy determinants motivated in this study.

4. FINDINGS AND ANALYSIS

4.1 DESCRIPTIVES

Table 4 below shows the descriptive statistics over the continuous variables used in the analysis for the ordered probit model and regression model, respectively.

Table 4. Descriptive Statistics, Continuous Variables

Variable	Obs	Mean	Std. Dev.	p25	p50	p75	Min	Max
OPROBIT								
ACTUALRATING	998	9.431	2.002	8.000	10.000	11.000	1.000	14.000
ROE	998	0.097	0.064	0.064	0.097	0.129	-0.312	0.304
LEV	998	0.076	0.075	0.025	0.050	0.107	0.001	0.424
SOLV	998	0.373	1.382	0.055	0.095	0.163	0.011	11.799
LOSS	998	0.915	0.665	0.646	0.801	0.963	0.045	5.372
NLTA	998	10.541	1.897	9.425	10.716	11.914	3.729	13.756
OLS								
DIFF	783	0.309	1.755	-1.000	0.000	1.000	-8.000	9.000
EM	783	0.396	0.090	0.000	0.027	0.061	-0.182	0.669
MTB	783	1.338	0.794	0.824	1.141	1.556	0.249	4.694
SPV	783	0.301	0.157	0.202	0.257	0.356	0.123	1.051
INV	783	0.521	1.064	0.133	0.252	0.411	0.004	8.126
SEG	783	6.785	5.110	4.000	5.000	9.000	1.000	29.000

This table shows summary statistics for all continuous variables used. ACTUALRATING represents the dependent variable in the ordered probit model and indicate each firm's' actual credit rating given by S&P, numerically equivalent to the highest (AA=14) and lowest (B-=1). ROE is the ratio of net income to equity. LEV is the ratio of debt to equity. SOLV is the ratio of EBITDA to interest expense. LOSS is the ratio of net claims to net premiums earned. SIZE is the natural logarithm of total assets. DIFF represents the dependent variable in the OLS model and indicate the difference between expected and actual ratings. EM indicates the level of proxied earnings management. MTB is the ratio of market capitalization to total book value. SPV is the five-year standard deviation of stock returns. INV is the ratio of investment income to investment assets. SEG indicates the number of either geographic or business segments.

Table 5 below shows the descriptive statistics over the dummy variables used in the regression analysis.

Table 5. Descriptive Statistics, Dummy Variables

Variable	Freq.	Percent	Cum.
CAT			
NO	642	81.99	81.99
YES	141	18.01	100.00
COUNTRY			
NORTH AMERICA	377	48.15	48.15
EUROPE	174	22.22	70.37
ASIA-PACIFIC	116	14.81	85.19
AFRICA	6	0.77	85.95
MIDDLE EAST	97	12.39	98.34
LATIN AMERICA	13	1.66	100.00
LIFE			
NO	300	38.31	38.31
YES	483	61.69	100.00
EV			
NO	654	83.52	83.52
YES	129	16.48	100.00
IFRS			
NO	494	63.09	63.09
YES	289	36.91	100.00

This table shows summary statistics for all independent dummy variables used. CAT indicates 1 if catastrophe events and 0 otherwise. COUNTRY indicates each insurance company´s geographical area in which they are listed. LIFE indicates 1 if life insurers and 0 otherwise. EV indicates 1 if the firm reports embedded value, and 0 otherwise. IFRS indicates 1 if firms report said reporting standard, and 0 otherwise.

Table 6 and 7 below show the pairwise correlation of the variables in estimating the expected credit ratings, and in analyzing the determinants of discrepancies between expected and actual ratings, respectively. Even though we observe some correlation between our independent variables, the result indicate that multicollinearity is not an issue in this study.

Table 6. Pairwise Correlation Matrix, Ordered Probit Model

Variable	*ACTUALRATING*	*ROE*	*LEV*	*SOLV*	*LOSS*	*SIZE*
ACTUALRATING	1.000					
ROE	0.196*	1.000				
LEV	0.143*	-0.020	1.000			
SOLV	0.083*	0.064*	-0.120*	1.000		
LOSS	-0.033	-0.085*	0.503*	-0.080*	1.000	
SIZE	0.039*	0.041	0.581*	-0.284*	0.322*	1.000

* p<0.05

This table presents the pairwise correlation matrix for the explanatory variables used in deriving expected credit rating in the ordered probit model. ROE is return on equity. LEV is leverage. SOLV is solvency. LOSS is loss ratio. SIZE is the natural logarithm of total assets.[2]

[2] Highest VIF is LEV at 1.93 and average VIF is 1.42, indicating that multicollinearity is of low concern (O'brien, 2007).

Table 7. Pairwise Correlation Matrix, OLS Model

Variable	*DIFF*	*EV*	*EM*	*IFRS*	*LIFE*	*MTB*	*CAT*	*SPV*	*INV*	*SEG*
DIFF	1.000									
EV	-0.135*	1.000								
EM	-0.087*	-0.004	1.000							
IFRS	-0.238*	0.373*	0.014	1.000						
LIFE	-0.128*	0.299*	-0.037	0.300*	1.000					
MTB	0.007	-0.145*	0.012	-0.158*	-0.191*	1.000				
CAT	0.061	-0.052*	-0.054	0.0580*	0.198*	0.036	1.000			
SPV	0.218*	0.064*	-0.052	-0.107*	0.101*	-0.096*	0.043	1.000		
INV	0.083*	-0.130*	0.015	-0.106*	-0.208*	0.431*	-0.094*	0.059*	1.000	
SEG	0.057	0.201*	-0.073*	-0.012	0.163*	-0.035	-0.029	0.115*	-0.104*	1.000

* $p<0.05$

This table presents the pairwise correlation matrix for the explanatory variables used in the OLS model. EV indicates if the firm reports embedded value, and o otherwise. EM indicates the level of proxied earnings management. IFRS indicates if firms report said reporting standard, and 0 otherwise. LIFE indicates 1 for firms with life insurance operations, and 0 otherwise. MTB is market capitalization to total book value. CAT indicates 1 for catastrophe events, and 0 otherwise. SPV is the five-year standard deviation of stock returns. INV is investment yield. SEG is the number of either geographic or business segments.[3]

4.2 ESTIMATION OF EXPECTED CREDIT RATINGS

The foundation of our analysis concerns deriving expected credit ratings from financial strength determinants. For purposes of deriving said ratings, as mentioned earlier, the determinants motivated in this paper base our proxy for expected ratings. The following ordered probit model show the relationship between the chosen determinants and actual credit ratings.

[3] Highest VIF is MTB at 1.64 and average VIF is 1.37, indicating that multicollinearity is of low concern (O'brien, 2007).

Table 8. Expected Rating Ordered Probit Model

Dependent variable: *ACTUALRATING*	
ROE	2.877***
	(0.514)
LEV	-1.777***
	(0.605)
SOLV	0.196***
	(0.025)
LOSS	-0.238***
	(0.058)
SIZE	0.403***
	(0.024)
Observations	998
LR chi2(5)	363.4
Prob > chi2	0.000
Pseudo R^2	0.0893

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table presents the estimation results for the ordered probit model for all insurance companies from 2009 to 2019. The dependent variable ACTUALRATING represents each firm's' actual credit rating given by S&P, numerically equivalent to the highest (AA=14) and lowest (B-=1). Acquiring expected ratings, the estimated ordered probit model is made by year. Explanatory variables consist of the following. ROE is return on equity. LEV is leverage. SOLV is solvency. LOSS is loss ratio. SIZE is the natural logarithm of total assets. For leverage, total assets over total liabilities was also tested, but was deselected due to collinearity with SIZE. For liquidity, cash over current liabilities were also tested, and no qualitative difference was found.

All variables are significant at the 1 percent level, with the expected sign in accordance with prior research and our expectations. Modelled as an ordered probit model, and with the coefficients formulated in log-odds scale, the strength of the relationship cannot be interpreted, other that there is a positive relationship between ACTUALRATING and ROE, SOLV and SIZE, and a negative relationship between ACTUALRATING and LEV and LOSS. Thus, increases in ROE, SOLV and SIZE is associated with ACTUALRATING being in a higher rating category. FOR LEV and LOSS, the interpretation is that firms with lower LEV and LOSS, ACTUALRATING is likely to be in a higher rating category. As the strength of the relationship cannot be interpreted, marginal effects, showing a probability scale, must be computed, resulting in a distribution of probabilities. Table 9 below show the marginal effects of each determinant on actual rating after the ordered probit model.

Table 9. Marginal Effects after Ordered Probit Model

ACTUAL RATING	*ROE*	*LEV*	*SOLV*	*LOSS*	*SIZE*
AA	0.0271**	-0.0167**	0.00184***	-0.00224**	0.00379***
	(0.0108)	(0.00819)	(0.000702)	(0.000965)	(0.00136)
AA-	0.112***	-0.0692***	0.00763***	-0.00926***	0.0157***
	(0.0277)	(0.0263)	(0.00163)	(0.00277)	(0.00283)
A+	0.215***	-0.133***	0.0146***	-0.0178***	0.0301***
	(0.0457)	(0.0478)	(0.00251)	(0.00481)	(0.00393)
A	0.581***	-0.359***	0.0396***	-0.0481***	0.0815***
	(0.109)	(0.124)	(0.00554)	(0.0120)	(0.00691)
A-	0.202***	-0.125***	0.0138***	-0.0167***	0.0283***
	(0.0485)	(0.0469)	(0.00285)	(0.00487)	(0.00483)
BBB+	-0.249***	0.154***	-0.0169***	0.0206***	-0.0348***
	(0.0526)	(0.0553)	(0.00291)	(0.00546)	(0.00447)
BBB	-0.415***	0.257***	-0.0283***	0.0344***	-0.0582***
	(0.0813)	(0.0893)	(0.00427)	(0.00890)	(0.00573)
BBB-	-0.254***	0.157***	-0.0173***	0.0210***	-0.0356***
	(0.0530)	(0.0560)	(0.00289)	(0.00569)	(0.00441)
BB+	-0.113***	0.0700***	-0.00772***	0.00936***	-0.0159***
	(0.0281)	(0.0269)	(0.00165)	(0.00276)	(0.00290)
BB	-0.0289***	0.0179**	-0.00197***	0.00239**	-0.00405***
	(0.0110)	(0.00856)	(0.000704)	(0.000973)	(0.00138)
BB-	-0.0460***	0.0284**	-0.00314***	0.00381***	-0.00645***
	(0.0143)	(0.0121)	(0.000891)	(0.00132)	(0.00168)
B+	-0.0139**	0.00859*	-0.000948**	0.00115**	-0.00195**
	(0.00648)	(0.00470)	(0.000424)	(0.000565)	(0.000843)
B	-0.0123**	0.00760*	-0.000838**	0.00102*	-0.00172**
	(0.00611)	(0.00438)	(0.000401)	(0.000529)	(0.000803)
B-	-0.00447	0.00276	-0.000305	0.000370	-0.000626
	(0.00347)	(0.00231)	(0.000235)	(0.000295)	(0.000478)
Observations	998	998	998	998	998

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table presents the marginal effect after the estimation results for the ordered probit model for all insurance companies from 2009 to 2019. Computed by firm and year, the table show the mean probability across all variables. Explanatory variables consist of the following. ROE is return on equity. LEV is leverage. SOLV is solvency ratio. LOSS is loss ratio. SIZE is the natural logarithm of total assets.

Each coefficient above shows each variable's marginal effect on the respective rating class, meaning that one-unit increase of each respective variable is associated with a probability of being associated with a rating class. For example, one-unit increase in ROE is associated with being 4.6 percent less likely to be in rating category 'BB-', and 58.1 percent more likely to be in rating category 'A'. The same directional interpretation applies to SOLV and SIZE. For LEV, one-unit increase is associated with being 7 percent more likely to be in rating category 'BB+', and 35.9 percent less likely to be in rating category ‘A’. The same directional interpretation applies to LOSS. The marginal effects (probabilities) are mutually exclusive, where a higher probability at some point entails a lower probability in another. The probabilities for each variable therefore sum up to zero. Estimating the ordered probit model in Table 8 yields an expected credit rating for each firm each year. Table 10 below shows actual ratings as given by S&P vertically, and expected ratings as estimated horizontally, providing an overview of the general equivalence between actual ratings and the empirically modelled expected ratings.

Table 10. Cross-tabulation of Actual Ratings and Expected Ratings

Actual Rating	Expected Rating												
	AA-	*A+*	*A*	*A-*	*BBB+*	*BBB*	*BBB-*	*BB+*	*BB-*	*B+*	*B*	*B-*	Total Actual
AA	0	1	10	0	0	0	0	0	0	0	0	0	11
AA-	0	0	17	14	0	0	0	0	0	0	0	0	31
A+	0	0	29	27	1	0	0	0	0	0	0	0	57
A	0	0	**96**	85	14	6	0	0	0	0	0	0	**201**
A-	1	2	73	**131**	17	18	5	1	0	0	0	0	**248**
BBB+	0	0	37	88	**26**	9	11	0	1	1	0	1	**174**
BBB	0	0	9	61	34	25	6	1	1	2	1	0	140
BBB-	0	0	5	32	16	11	3	0	0	0	0	0	67
BB+	0	0	0	2	8	11	2	1	1	4	0	1	30
BB	0	0	0	1	2	3	0	0	0	1	1	1	9
BB-	0	0	0	1	2	9	0	1	0	1	3	0	17
B+	0	0	0	3	1	2	0	0	0	0	0	0	6
B	0	0	2	3	0	0	0	0	0	0	0	0	5
B-	0	0	0	1	0	0	0	0	0	0	0	1	2
Total Expected	1	3	**278**	**449**	**121**	94	27	4	3	9	5	4	998

This table show the cross-tabulation of actual ratings against expected ratings from the model estimated in Table 8. Actual ratings are shown horizontally, and expected ratings are shown vertically.

Table 10 indicates that although every rating class differs regarding the amount of actual versus expected ratings, the general interpretation is that they, in the majority of the rating classes, reflect each other well. One important note is that a large portion of actual ratings are assigned either an A, A- or BBB+, accounting for roughly 62 percent of ratings. This will have implications on the overall result, as ratings such as B-, B, B+ and AA cumulatively only account for over 2 percent of the sample, meaning that those ratings rarely obtain the highest probability (Van Gestel et al., 2007). Another illustrative note is that for most rating classes, the majority of actual ratings are 'correctly' predicted, or at-expected, and the majority of deviance occurs one notch above or below. For example, for 'A', 96 ratings are at-expected, whereas 73 ratings are at expected rating 'A-', and 29 ratings at expected rating 'A+'. To illustrate this further, Table 11 below shows the distribution of the difference between expected and actual ratings at different time-horizons.

Table 11. Distribution of Difference, Expected vs. Actual Ratings

	DIFF	-9	-8	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9
	Freq.	-	1	2	4	3	41	94	**147**	**283**	**208**	128	53	18	5	2	3	3	3
t	Percent	-	0.10	0.20	0.40	0.30	4.11	9.42	**14.73**	**28.36**	**20.84**	12.83	5.31	1.80	0.50	0.20	0.30	0.30	0.30
	Cum.	-	0.10	0.30	0.70	1.00	5.11	14.53	**29.26**	**57.62**	**78.46**	91.28	96.59	98.40	98.90	99.10	99.40	99.70	100.00
	Freq.	-	1	2	4	2	37	81	**133**	**245**	**194**	120	43	20	3	4	2	2	4
t+1	Percent	-	0.11	0.22	0.45	0.22	4.12	9.03	**14.83**	**27.31**	**21.63**	13.38	4.79	2.23	0.33	0.45	0.22	0.22	0.45
	Cum.	-	0.11	0.33	0.78	1.00	5.13	14.16	**28.99**	**56.30**	**77.93**	91.30	96.10	98.33	98.66	99.11	99.33	99.55	100.00
	Freq.	1	-	1	3	4	29	69	**119**	**213**	**178**	103	42	16	1	8	1	-	4
t+2	Percent	0.13	-	0.13	0.38	0.51	3.66	8.71	**15.03**	**26.89**	**22.47**	13.01	5.30	2.02	0.13	1.01	0.13	-	0.51
	Cum.	0.13	-	0.25	0.63	1.14	4.80	13.51	**28.54**	**55.43**	**77.90**	90.91	96.21	98.23	98.36	99.37	99.49	-	100.00
	Freq.	-	1	2	3	3	24	57	**107**	**173**	**166**	88	40	11	3	4	3	2	1
t+3	Percent	-	0.15	0.29	0.44	0.44	3.49	8.28	**15.55**	**25.15**	**24.13**	12.79	5.81	1.60	0.44	0.58	0.44	0.29	0.15
	Cum.	-	0.15	0.44	0.87	1.31	4.80	13.08	**28.63**	**53.78**	**77.91**	90.70	96.51	98.11	98.55	99.13	99.56	99.85	100.00

This table shows the frequency, percentage and cumulative distribution of the difference between expected and ratings. Difference at 0 indicates that actual ratings are at-expected rating. Difference at 1 indicates that actual and expected ratings differs by on notch, and so forth.

Table 11 above show the difference between expected and actual ratings and indicate actual ratings that are above-expected, at-expected and below-expected ratings for times *t+n.* At time *t*, the amount of ratings at-expected, meaning that actual rating and expected rating corresponds, are 283 out of 998 cases, followed by 245 in *t+1*, 213 in *t+2*, and 173 in *t+3.* Consequently, the amount of actual ratings that are at-expected decreases for each time *t+n*. Controlling for the fact that our sample decreases for every *t+n*, the percent of ratings at-expected for time *t* is about 28.4, for time *t*+1 about 27.3%, for time *t*+2 about 26.9% and for time *t*+3 about 25.1%. The interpretation is such that the longer the time span, the lower the equivalence between actual and expected ratings. In time *t*, actual ratings that are one notch above- or below-expected amounts to 355 out of 998 cases (35.57%). Actual ratings that are either one notch above-, below, or at-expected amounts to 638 out of 998 cases (63.93%).

4.3 HYPOTHESIS TESTING

In order to analyze what drives the difference between expected and actual ratings, OLS regressions were performed. The *difference* was hypothesized to be either positive or negative for firms reporting under IFRS and for LIFE, and negative for EM. EV was hypothesized to have no effect on the *difference*. Table 12 below shows the regression results.

Table 12. Difference, OLS Model

Dependent variable:	$DIFF_t$ (1)	$DIFF_{t+1}$ (2)	$DIFF_{t+2}$ (3)	$DIFF_{t+3}$ (4)
EV	-0.308*	-0.256	-0.233	-0.316
	(0.179)	(0.189)	(0.198)	(0.208)
EM	-0.803	-1.280*	-2.065***	-0.696
	(0.776)	(0.749)	(0.783)	(0.769)
IFRS	-0.498***	-0.484***	-0.546***	-0.516***
	(0.151)	(0.162)	(0.176)	(0.181)
*IFRS*EM*	-0.387	-0.228	1.868	0.776
	(1.351)	(1.445)	(1.498)	(1.606)
LIFE	-0.620***	-0.740***	-0.726***	-0.652***
	(0.137)	(0.146)	(0.156)	(0.162)
MTB	-0.0920	-0.0449	-0.104	-0.238**
	(0.0900)	(0.0993)	(0.109)	(0.116)
CAT	0.306**	0.297*	0.323*	0.514***
	(0.153)	(0.164)	(0.180)	(0.187)
SPV	4.618***	5.237***	5.246***	5.077***
	(0.455)	(0.470)	(0.476)	(0.469)
INV	0.0136	-0.0751	-0.0767	0.00772
	(0.0661)	(0.0753)	(0.0816)	(0.0848)
SEG	0.0448***	0.0318***	0.0236*	0.0186
	(0.0116)	(0.0123)	(0.0130)	(0.0135)
Constant	-1.240***	-1.285***	-1.400***	-0.739**
	(0.349)	(0.362)	(0.365)	(0.360)
COUNTRY FE	YES	YES	YES	YES
YEAR FE	YES	YES	YES	YES
Observations	783	686	596	513
Prob > F	0.000	0.000	0.000	0.000
R^2	0.231	0.265	0.287	0.311

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

All independent variables are lagged *t+n* for each regression except for CAT which is lagged *t-1* at time *t*

This table presents the cross-sectional regressions for all insurance companies from 2009 to 2019. The dependent variable DIFF is the difference between the empirically modelled expected ratings and actual ratings. $DIFF_t$ in Model 1 represents the difference at time t, $DIFF_{t+1}$in model 2 represents the difference between expected ratings time t+1 minus actual ratings time t, followed by time t+2 in model 3 and t+3 in model 4. EV indicates if the firm report Embedded Value, and 0 otherwise. EM indicates the level of proxied earnings management. IFRS indicates if firms report said reporting standard, and 0 otherwise. LIFE indicates 1 for firms with life insurance operations, and 0 otherwise. MTB is market capitalization to total book value. CAT indicates 1 for catastrophe events, and 0 otherwise. SPV is the five-year standard deviation of stock returns. INV is investment yield. SEG is the number of either geographic or business segments.

EV was hypothesized to not be associated with DIFF, built from prior research underlining the fact that firms reporting EV show higher transparency (Wu & Hsu, 2011), and exhibit less asymmetric information (Serafeim, 2010), providing rating agencies with a better view of the firm. As such, the results for time *t+1, t+2* and *t+3* show no statistical significance that EV is associated with either over-or underestimated ratings, suggesting that EV reporting facilitates accurate credit ratings. This is in line with our predictions. The hypothesis that EV is associated with DIFF is therefore rejected for time *t+1, t+2* and *t+3.* This finding is further strengthened by research underlining the value relevance of EV (Préfontaine et al., 2009; Almezweq & Liu, 2012; Horton, 2007; El-Gazzar et al., 2014). The fact that EV is voluntarily disclosed might further strengthen EV moving towards a lower difference, as the disclosure adds another layer of transparency in the signaling value of EV (Klumpes, 2002). On the other hand, the value of EV might be biased in this case, as one could expect that only firms that are in good health should benefit from it. A middle ground might however exist, as life insurers that report EV are shown to be associated with higher future profit expectation than those that do not disclose EV (Klumpes, 2002). Likewise, as highlighted by Klumpes (2002), some firms report EV primarily because it has a positive signaling effect to users of financial statements, and not for its transparency value. Yet, our result indeed suggests that EV reporting facilitates an accurate credit rating. However, what is observed in Table 12 is that EV indeed is associated with overestimated ratings in time *t*, being significant at the 10 percent level, which contradicts our expectations. One possible explanation to why EV is significant is that EV has a positive effect on firm value, which could entail a higher rating. However, as times *t+1*, *t+2* and *t+3* show no significant association, the significance in time *t* is not generalizable. Therefore, our overall findings suggest that EV enables the rating agencies to assign a rating equivalent to insurance companies' financial strength.

For EM, the association was expected to be negative, indicating that firms using EM exhibit overestimated ratings. The motivation behind this is that firms using earnings management is shown to be positively associated with credit ratings (Liu et al., 2018; Alissa et al., 2013). Prior findings suggest that earnings management is common in insurers (Beaver et al., 2003), which is troublesome as rating agencies rely on the information provided by firms (Demirtas & Cornaggia, 2013). EM in time *t+1* and *t+2* show a negative association with DIFF, significant at the 10 and 1 percent level, respectively. This coincides with our predictions, implying that higher EM indeed is associated with overestimated ratings. In time *t+1* and *t+2,* we therefore reject the null hypothesis stating that EM is negatively associated with DIFF. However, the result in time *t* and *t+3* show no statistical significance, and we cannot reject the null hypotheses. Consequently, as EM's association with DIFF is apparent in only two out of four models, this finding should be generalized with caution. Accordingly, what could explain the results is that rating agencies indeed, in some instances, are able to differentiate between discretionary and non-discretionary accruals, thus contradicting Liu et al. (2018). This would imply that rating agencies are able to detect earnings management, and that firms cannot employ earnings management to receive favorable ratings.

Alternatively, the inconclusive result of EM might instead lie in the different forms of earnings management applied by firms under different accounting standards. According to Evans et al. (2012), accrual earnings management is higher in IFRS firms, whereas real earnings management is higher in U.S. GAAP firms, due to its more effective enforcement environment. Since we only examine accrual earnings management, it is possible that U.S. GAAP firms using real earnings management affect the result as 40 percent of our sample contains U.S. GAAP firms. As such, earnings management might still be present, but for U.S. GAAP firms, it consists of real, rather than accruals earnings management, which is harder to detect.

Moreover, IFRS show a negative association with DIFF, significant at the 1 percent level for all regressions. Accordingly, a negative association between IFRS and DIFF suggests that insurers reporting under IFRS exhibit overestimated ratings. Our null hypothesis stating that IFRS is not positively (negatively) associated with DIFF is therefore rejected in all cases. Arguments have been made for IFRS to facilitate the underlying fundamentals of the firm as firms are allowed more freedom of interpretation (Runesson et al., 2018; El-Gazzar et al., 2014). Further, firms reporting under U.S. GAAP are unified under one rules-based accounting standard, whereas IFRS firms are able to interpret the principle-based accounting standard based on the jurisdiction (IASB, 2017). Accordingly, research have shown that, generally, earnings management is more common in non-U.S. firms than U.S. and U.S. GAAP firms (Alford et al., 1993; Land & Lang, 2002; Lang et al., 2003). As such, it is possible that IFRS firms to a higher degree engage in more accrual earnings management, as we should see a bigger accrual earnings management effect in IFRS firms compared to U.S. GAAP firms (Evans et al., 2012). However, the interaction variable between IFRS and EM is not significant. Evidently, the reason IFRS firms exhibit overestimated ratings cannot be explained by higher accrual earnings management. Instead, what is acknowledged by IASB (2017) is the problems surrounding IFRS 4. Accordingly, it is underlined that firms are not required to ensure that the information provided in the financial statements covering insurance contracts are reliable or relevant for decision-makers. This indeed impair the comparability and understandability of insurers' financial statements for users. As such, relevant information can be omitted or obscured without admonitions. Therefore, firms under IFRS might exhibit overestimated ratings due to the discretionary and opportunistic nature of IFRS 4. An implication of this is that the upcoming replacement of IFRS 4 with IFRS 17 (IASB, 2020, March 17), from a user's standpoint, should indeed come sooner rather than later.

Finally, LIFE is significant at all time-intervals at the 1 percent level, where the negative coefficient indicates that life insurers exhibit overestimated ratings. Horton (2007) highlights that life insurers initially understate policies, and as the policies mature, higher profits are realized. The implication is the difficulty of determining the drivers of profitability in life insurances, and to distinguish between earnings related to past or current performance (2007). Additionally, due to decreasing interest rates, life insurers have been forced to seek riskier investments to achieve higher yields, leading to more complex and opaque ones (de Bandt & Overton, 2019; Koijen & Yogo, 2017). As such, a possible explanation for the overestimation of ratings for life insurers might indeed be that the financial reporting is difficult to assess, leading to the difficulty of deriving profitability, and that rating agencies cannot, or do not, acknowledge the actual risk exposure of life insurers. The result is supported by Gerstner et al. (2015), who underlines the difficulty in grasping life insurers' financial statements due to their long-term and complex financial contracts. Life insurers are also underlined to be more vulnerable to macroeconomic shocks than general insurers (Koijen & Yogo, 2017). The implied uncertainty might also explain the overestimation of ratings in life insurers. Thus, Galil's (2003) as well as Hilscher and Wilson's (2016) argument that ratings fail to incorporate higher default risk might indeed hold some merit in this case.

5. CONCLUDING REMARKS

5.1 CONCLUSION

The aim of this study is to examine what determines the discrepancies between expected and actual credit ratings among insurance companies. By estimating empirically modelled expected ratings, corresponding actual ratings were subtracted, yielding the difference between them. In that, four determinants are hypothesized to determine the difference. For three of the four time-intervals analyzed, we find evidence supporting the hypothesis that embedded value is not associated either over- or underestimated ratings, indicating the value relevance of embedded value in facilitating an accurate rating. Earnings management is hypothesized to be positively associated with the difference, supported in two out of four time-intervals, indicating weak evidence supporting that earnings management is associated with overestimated ratings. As for IFRS, we hypothesize a negative association with the difference, motivated by the higher risk of earnings management due to managerial discretion. We find strong evidence suggesting that firms under IFRS indeed exhibit overestimated ratings. However, overestimated ratings under IFRS do not depend on earnings management. Instead, insurers under IFRS are argued to receive overestimated ratings due to the opportunistic and discretionary nature of IFRS 4. Finally, we hypothesize life insurers to be associated with a positive or negative difference. Findings suggest a negative association, indicating that life insurers exhibit overestimated rating. The rationale reads that life insurers' financial statements are hard to interpret, which impairs the ability to derive profitability, and that the increased complex risk exposure entails that the rating agencies cannot, or do not, acknowledge the actual risk exposure of life insurers.

Notably, rating agencies have access to the same, and even more, information than we have. One could therefore question our findings, stating that they are noisy and subject to coincidence. However, despite the difference in access of information, our findings indeed show systematical differences in the accuracy of ratings relative to financial strength. For instance, the type of reporting standard used should not play a role in the accuracy of ratings, yet our findings suggest otherwise. As such, the ability to assign an accurate rating relative to insurers' financial strength is not without discrepancy.

5.2 CONTRIBUTION

The purpose of this study is to enlighten users with factors to consider when assessing ratings, and to direct future research towards exploring such framework further. First, our evidence underlines four factors to consider when assessing ratings. This has apparent benefits for users of financial statements and indeed for rating agencies, as well as users of credit ratings. Second, this study contributes to the credit rating literature. Indeed, determinants of discrepancies between expected and actual ratings have not been addressed until now. Therefore, we contribute not only by filling this gap, but direct future research towards exploring this framework further.

5.3 FUTURE RESEARCH

There is an extensive line of research on credit ratings in connection to financial strength, often determining what factors determine the level of credit rating. Unlike prior areas of focus, this study determines what factors affect the *difference* between ratings and financial strength. Being the first of its kind, several opportunities for further research appear. First, having set the stage as for the methodology in determining factors that affect the discrepancies between expected and actual ratings, it would be interesting to examine other determinants than embedded value, earnings management, firms that report under IFRS, and firms that have life insurance operations. For example, instead of defining LIFE insurers as mentioning life insurance operations in the extended business description, one could determine the effect of insurers having only life operations. Additionally, the analysis could be complemented with macro-economic variables, controlling for the effect that boom and bust times might have on the over- and underestimation of ratings relative to financial strength. Second, applying this methodology to other industries provides a more general understanding of what explains the differences between expected and actual ratings, whether the same results are found in other settings, or whether other industries with a less pronounced discretionary reporting might yield other results. Third, this study only looks at ratings set by S&P. Complementing the analysis comparing different rating agencies is interesting as our findings would either be strengthened, or that there is some discrepancy among rating agencies. Finally, this study could indeed be extended using public as well as private companies.

REFERENCES

Adams, M., Burton, B., & Hardwick, P. (2003). The Determinants of Credit Ratings in the United Kingdom Insurance Industry. *Journal of Business Finance & Accounting, 30*(3-4), 539-572.

Alford, A., J. Jones, R. Leftwich, & M. Zmijewski. 1993. The relative information content of accounting disclosures in different countries. *Journal of Accounting Research 31* (Supplement): 183–223.

Alissa, W., Bonsall IV, S., Koharki, K., & Penn, M. (2013). Firms' use of accounting discretion to influence their credit ratings. *Journal of Accounting and Economics, 55*(2-3), 129-147.

Almezweq, M., & Liu, G. (2012). The value relevance of voluntary European embedded value disclosures: evidence from UK life insurance companies. *International Journal of Accounting and Finance,* 3(4), 343-366.

Amato, J., & Furfine, C. (2004). Are credit ratings procyclical? *Journal of Banking and Finance, 28*(11), 2641-2677.

Angell, K., Archer-Lock, P., Harris, S., Moss, G., Patfield, S., & Sayers, J. (2000). Assessing the financial strength of insurers working party. *General Insurance Convention* 25-28 October.

Afik, Z., Bouhnick, N., & Galil, K. (2016). Have credit ratings become more accurate?. Working paper, Ben-Gurion University of the Negev.

Babi, M. (2015). The effects of financial risks on the relationship between earnings and stock returns. *International Journal of Organizational Leadership, 4*(2), 154-169.

Barinov, A., Shawn, S. P., & Celim, Y. (2019). Form complexity and post-earnings announcement drift.

Barth, M., Landsman, W., Lang, M., & Williams, C. (2012). Are IFRS-based and US GAAP-based accounting amounts comparable? *Journal of Accounting and Economics, 54*(1), 68-93.

Beaver, W. H., McNichols, M. F., & Nelson, K. K. (2003). Management of the loss reserve accrual and the distribution of earnings in the property-casualty insurance industry. *Journal of Accounting and Economics*, 35(3), 347-376.

Becker, B., & Milbourn, T. (2011). How did increased competition affect credit ratings? (Report). *Journal of Financial Economics, 101*(3), 493-514.

Blums, M. (2003). D-Score: Bankruptcy Prediction Model for Middle Market Public Firms.

Boguslauskas, V., Mileris, R., & Adlytė, R. (2011). New internal rating approach for credit risk assessment. *Technological and Economic Development of Economy, 17*(2), 369-381.

Bouzouita, R., & Young, A. J. (1998). A probit analysis of best ratings. *Journal of Insurance Issues*, 23-34.

Browne, M. J., Carson, J. M., & Hoyt, R. E. (1999). Economic and market predictors of insolvencies in the life-health insurance industry. *Journal of Risk and Insurance*, 643-659.

Bryman, A., & Bell, E. (2007). *Business research methods* (2.nd ed.). Oxford: *Oxford University Press*.

Cagle, J. (1996). Insurance company loss reserve adjustments and security prices. *Journal of Insurance Regulation*, 15(1), 124.

Caporale, G. M., Cerrato, M., & Zhang, X. (2017). Analysing the determinants of insolvency risk for general insurance firms in the UK. *Journal of Banking & Finance*, *84*, 107-122.

Chen, R., & Wong, K. (2004). The Determinants of Financial Health of Asian Insurance Companies. *Journal of Risk and Insurance, 71*(3), 469-499.

Chin, M. (2015). Accounting Quality and Credit Ratings' Ability to Predict Default.

Damodaran, A. (2015). Applied corporate finance. *John Wiley & Sons*.

Demirtas, K. O., & Cornaggia, K. R. (2013). Initial credit ratings and earnings management. *Review of Financial Economics*, 22(4), 135-145.

De Bandt, O., & Overton, G. (2019). Why do insurers fail? A comparison of life and non-life insolvencies using a new international database.

Diers, D., Eling, M., Kraus, C., & Reuß, A. (2012). Market-consistent embedded value in non-life insurance: how to measure it and why. *The Journal of Risk Finance*.

Deutsche Bundesbank. (2019). Financial stability review 2019. *Deutsche Bundesbank, Frankfurt am Main*.

Du, Y. (2003). Predicting Credit Rating and Credit Rating Changes: A New Approach.

EIOPA (2018). “Failures and Near Misses in Insurance”. *Publications Office of the European Union.*

El-Gazzar, S. M., Jacob, R. A., & McGregor, S. (2015). The relative and incremental valuation effects of embedded value disclosure by life insurers: evidence from cross-listed firms in the US. *Accounting Horizons*, 29(2), 327-339.

Evans, M. E., Houston, R. W., Peters, M. F., & Pratt, J. H. (2015). Reporting regulatory environments and earnings management: US and non-US firms using US GAAP or IFRS. *The Accounting Review*, 90(5), 1969-1994.

Fenn, G. W., & Cole, R. A. (1994). Announcements of asset-quality problems and contagion effects in the life insurance industry. *Journal of Financial Economics*, 35(2), 181-198.

Galil, K. (2003, October). The quality of corporate credit rating: an empirical investigation. In *EFMA 2003 Helsinki Meetings*.

Gerstner, T., Lohmaier, D., & Richter, A. (2015). Value Relevance of Life Insurers' Embedded Value Disclosure and Implications for IFRS 4 Phase II. *Munich Risk and Insurance Center Working Paper*, (27).

Giordani, P., Jacobson, T., Von Schedvin, E., & Villani, M. (2014). Taking the twists into account: Predicting firm bankruptcy risk with splines of financial ratios. Journal of Financial and Quantitative Analysis, 49(4), 1071-1099.

Gordon, L., Loeb, M., & Tseng, C. (2009). Enterprise risk management and firm performance: A contingency perspective. *Journal of Accounting and Public Policy, 28*(4), 301-327.

Grace, M. F., Klein, R. W., & Phillips, R. D. (2003). Insurance Company Failures: Why Do They Cost So Much?. *Georgia State University Center for Risk Management and Insurance Research Working Paper*, (03-1).

Hilscher, J., & Wilson, M. (2016). Credit ratings and credit risk: Is one measure enough?. *Management science*, *63*(10), 3414-3437.

Horton, J. (2007). The value relevance of 'realistic reporting': Evidence from UK life insurers. *Accounting and Business Research*, 37(3), 175-184,188,191,193,195-197.

International Accounting Standards Board. (2017). WHY CHANGE INSURANCE CONTRACTS ACCOUNTING? *The forthcoming IFRS® insurance contracts Standard.* Retrieved from:
https://www.ifrs.org/-/media/project/insurance-contracts/current-stage/educational-materials/why-change-insurance-contracts-accounting.pdf

International Accounting Standards Board. (2020, March 17). IASB decides on new effective date for IFRS 17 of 1 January 2023. Retrieved from:
https://www.ifrs.org/news-and-events/2020/03/ifrs-17-effective-date/

Kang, C., Kang, H. G. (2009). The Effect of Credit Risk on Stock Returns. *Journal of Economic Research (JER), 14*(1), 49-67.

Kisgen, D. J. (2006). Credit ratings and capital structure. *The Journal of Finance*, 61(3), 1035-1072.

Kisgen, D. J. (2009). Do firms target credit ratings or leverage levels?. *Journal of Financial and Quantitative Analysis*, 44(6), 1323-1344.

Koijen, R. S., & Yogo, M. (2017). Risk of life insurers: Recent trends and transmission mechanisms (No. w23365). *National Bureau of Economic Research*.

Klumpes, P. J. (2002). Incentives facing life insurance firms to report actuarial earnings: evidence from Australia and the UK. *Journal of Accounting, Auditing & Finance*, 17(3), 237-256.

Krantz, M. (2013, September 13). 2008 crisis still hangs over credit-rating firms. *USA TODAY.* Retrieved from:
https://eu.usatoday.com/story/money/business/2013/09/13/credit-rating-agencies-2008-financial-crisis-lehman/2759025/

Kwon, W., & Wolfrom, Leigh. (2016). Analytical tools for the insurance market and macro-prudential surveillance. *Oecd Journal: Financial Market Trends, 2016*(1), 1-47.

Lang, M., J. S. Raedy, & M. Yetman. 2003. How representative are firms that are cross listed in the United States? An analysis of accounting quality. *Journal of Accounting Research* 41: 363–386.

Land, J., & M. Lang. 2002. Empirical evidence on the evolution of international earnings. *The Accounting Review* 77: 115–134.

Le, M., & Xu, J. (2018). Future Solvency Prediction for Property Insurance Companies. *Advances in Social Sciences Research Journal*, *5*(12).

Liu, A. Z., Subramanyam, K. R., Zhang, J., & Shi, C. (2018). Do firms manage earnings to influence credit ratings? Evidence from negative credit watch resolutions. *The Accounting Review*, 93(3), 267-298.

Markarian, G., & Parbonetti, A. (2007). Firm Complexity and Board of Director Composition. *Corporate Governance: An International Review, 15*(6), 1224-1243.

Mischek, M. (2019, June 23). Why credit rating agencies are still getting away with bad behaviour. *The Conversation.* Retrieved from: https://theconversation.com/why-credit-rating-agencies-are-still-getting-away-with-bad-behaviour-117549

Malik, H. (2011). Determinants of insurance companies profitability: an analysis of insurance sector of Pakistan. *Academic Research International*, *1*(3), 315.

McGregor, S., Jacob, R. A., & El-Gazzar, S. M. (2013) 'The valuation effects of embedded value disclosure by life insurers', *Int. J. Economics and Accounting*, Vol. 4, No. 1, pp.26–53.

Mwaurah, I., Muturi, W., & Waititu, A. (2017). The influence of financial risk on stock returns. *International Journal of Scientific and Research Publications*, 7(5), 2250-3153

Njegomir, V., & Marović, B. (2012). Contemporary Trends in the Global Insurance Industry. *Procedia - Social and Behavioral Sciences, 44*(C), 134-142.

O'brien, R. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity, 41*(5), 673-690.

Ohlson, J. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research, 18*(1), 109.

Partnoy, F. (2002). The paradox of credit ratings. In *Ratings, rating agencies and the global financial system* (pp. 65-84). Springer, Boston, MA.

Préfontaine, J., Desrochers, J., & Godbout, L. (2009). The informational content of voluntary embedded value (EV) financial disclosures by Canadian life insurance companies. *International Business & Economics Research Journal* (IBER), 8(12).

Runesson, E., Samani, N., & Marton, J. (2018). *Financial accounting theory: An accounting quality approach* (1st ed.).

Serafeim, G. (2011). Consequences and Institutional Determinants of Unregulated Corporate Financial Statements: Evidence from Embedded Value Reporting. (Report). *Journal of Accounting Research*, 49(2), 529-571.

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *The Journal of Business, 74*(1), 101-124.

Song, I. J. (2018). An Overview of US Property-liability Insurer Earnings Management via Loss Reserves. *Academy of Accounting and Financial Studies Journal*.

Standard and Poor's. (2019, September 18). S&P Global Ratings Definition. *Standard and Poor's.* Retrieved from:
https://www.standardandpoors.com/en_US/web/guest/article/-/view/sourceId/504352
-

Standard & Poor's (2008). 2008 Criteria | Corporates | General: Corporate Criteria: Analytical Methodology. *McGraw Hill Financial.*

Steiger, F. (2010). The Impact of Credit Risk and Implied Volatility on Stock Returns.

Van Gestel, T., Martens, D., Baesens, B., Feremans, D., Huysmans, J., & Vanthienen, J. (2007). Forecasting and analyzing insurance companies' ratings. *International Journal of Forecasting*, *23*(3), 513-529.

Wu, R., & Hsu, A. (2011). Value Relevance of Embedded Value and IFRS 4 Insurance Contracts. *Geneva Papers on Risk and Insurance - Issues and Practice* 36(2): 283-303

Yakob, R. B., Yusop, Z., Radam, A., & Noriszura, I. (2012). Camel Rating Approach to Assess the Insurance Operators Financial Strength. *Jurnal Ekonomi Malaysia*. 46. 3-15