# Learning the Quality of Risk Culture in Insurance Firms

**Aparna Gupta**[1]

and

**Abena Owusu**[1]

Lally School of Management, Rensselaer Polytechnic Institute

January 15, 2019

**Abstract:** Insurance firms, similar to banks, help reallocate risk both between agents and over time, and in doing so must excel in their risk management decisions and practices. In this paper, we apply unsupervised machine learning approach to identify the risk culture of insurance firms, and compare it against that of banks (Gupta and Owusu (2019)). Using a K-means clustering technique, we group a sample of 10-K text documents filed by insurance firms into three distinct clusters and label them as good, fair and poor risk culture clusters. Validation of the risk culture cluster labels shows that insurance firms with sound risk culture have higher return on asset, return on equity and Tobin's Q ratios. They also have more independent and female directors on their boards. Comparing the risk culture of insurance firms with that of banks, we find that although both types of firms disclose more about their positive sentiments for leadership, strategy and portfolio, the learning algorithm picks the uncertainty and litigious sentiments for leadership and strategy for distinguishing between firms for their risk culture. Uncertainty and litigious sentiment for risk training, education and recruitment strongly defines risk culture in insurance firms, which indicates the importance of trained risk professionals, like actuaries, in the insurance industry for strong risk decisions, practices and risk management.

**Keywords:** Risk culture, machine learning, text mining, cluster analysis

**JEL Codes**: G23, M14, G32, C55.

## 1. Introduction

After the 2008 financial crisis, policymakers and regulators have put more weight on promoting a sound risk culture and risk management practices within the financial industry (IIF 2009, Stulz

[1] Email address: Aparna Gupta (guptaa@rpi.edu); Abena Owusu (owusua@rpi.edu)

2014, Fergussen 2018). However, the rules, regulations and the overall discussion of issues seems to be focused on banking firms. For instance, after the enactment of the Dodd-Frank Act, many commented that the legislation has many provisions affecting the banking industry, with little attention paid to the insurance industry (Ross and Schwartzman 2010, NAIC 2017). Only one section of the Act, i.e. Title V, mainly applied to the insurance industry, even though some insurance firms were also at the center of the global financial crisis, and many were affected by the financial crisis. Insurance firms, similar to banks, help reallocate massive amounts of risk both between agents and over time, and in doing so must excel in their risk management decisions and practices. Therefore, an evaluation of their risk culture that fundamentally drives their ability to meet their risk management challenges is important.

In this paper, we use unsupervised machine learning techniques to define the risk culture of a sample of 313 insurances firms using 2853 10-K filings from 2000 to 2017. Specifically, we use text mining techniques to extract risk culture features from the firms' 10-K text documents, and apply K-means clustering analysis to group the risk culture features and text documents into three distinct risk culture classes. This machine learning approach enables us to address the challenge of measuring and defining risk culture of insurance firms, while also allows comparing the risk culture of insurance firms with that of banks. In particular, we reference a prior study of risk culture in banks (Gupta and Owusu, 2019) and draw key comparisons between risk culture features and clusters obtained for insurance firms against those for banks.

The Institute of International Finance (IIF 2009) defines risk culture as, "*the norms and traditions of the behavior of individuals and of groups within an organization that determine the way in which they identify, understand, discuss, and act on the risks the organization confronts and the risks it takes.*" This means all employees within the firm have to be conscious of the firm's risk appetite in their day-to-day decisions. Specifically for insurance firms, risk decisions are no longer within the domain of only actuaries (KPMG 2016). Hence in order to effectively supervise risk behaviors within the firm, regulators and firms need to be able to assess and define the overall risk culture of the firm.

One challenge in the studies of risk culture is how to define or measure the risk culture of a firm (Williams 1983, Ghaziani 2009). Traditional approaches using surveys, interviews and questionnaires to measure risk culture, are mostly flawed with low and biased responses, as well as some data collection errors. To overcome these challenges, cultural sociologists are more and more adopting advanced data collection and text analysis techniques, commonly used in the computer science research (Lazer et al. 2009, King 2011, Bail 2014, Evans and Aceves 2016).

In the Internet era, large volumes of data, often referred to as *big data*, are collected on a real time basis, in the form of binary or quantitative data as well as in textual form. These text data are extracted from primary text documents, interview transcripts, comments and discussions between individuals and organizations available on the web, social media and audios. These text data also span a wide time period and geography, which are lacking through traditional sources (Bail 2014). Recent advanced technologies such as automated text mining and machine learning techniques allow for pattern identification in these texts for classification of culture.

We utilize a total of 2,853 regulatory 10-K reports filed annually with the Securities and Exchange Commission (SEC) and available on the SEC EDGAR system, for a sample of 313 insurance firms. Use of 10-K documents for the purpose of the study is favorable as these documents are consistent, regularly and periodically available, are easy to read and a good source of information on financial health of public firms (Li 2008, Bobnaruk et al. 2015). More importantly for our study, they have relevant specific risk sections, namely Item 1A and Item 7, which discuss firm's risk factors, operations and future outlook.

After downloading and preprocessing the text from the corpus of 10-K documents, we use a two-dimensional dictionary built from a comprehensive risk culture framework (Gupta and Owusu 2009) and sentiment dictionary (Loughran and McDonald 2014), to extract features that describe key drivers of risk culture and their associated sentiments. The risk culture framework has seven key drivers that cover Leadership, Strategy, Decision, Control, Recruitment, Reward and Portfolio. We use five sentiment dictionaries that capture Positive, Negative, Uncertainty, Litigious and Constraining sentiments. This leads to a total of 35 risk culture-sentiment features. To allow for easy interpretation of the features and reduce the dimensionality of the feature space, we apply principal component analysis and reduce the 35 features down to 5 principal components. The five principal components correlate strongly with features based on uncertainty, positive and litigious sentiments of several risk culture key driver. These selected features lay the foundation for identifying the risk culture patterns in our dataset.

Unsupervised machine learning uses algorithms to discover patterns and structures within unlabeled data. Unlike supervised machine learning, an unsupervised learning algorithm does not need a known outcome variable or trained data that assists in mapping the unlabeled data to a given set of outcomes. Clustering analysis is commonly used to separate the data into groups with low within cluster variation and a high between cluster variations. Once clustering is achieved, the clusters are labeled to define the category or outcome variable.

We use the k-means cluster analysis (Hartigan and Wong, 1979) to identify three distinct risk culture classes. Using their extracted risk culture-sentiment features, we label the clusters as good, fair and poor risk culture classes. We find that 68% of the insurance firms fall in the good risk culture class, and 18% and 13% fall in the fair and the poor risk culture classes, respectively. To validate the results of our clusters, we compare the financial performance and governance of the three clusters. We find that, insurance firms in the good risk culture cluster have higher ROA, ROE and Tobin's Q financial ratios. They also have a higher number of independent and female directors on their boards and pay higher directors' salaries than insurance firms in the poor risk culture class. Our results thus show that building a strong risk culture among insurance firms improves long-term business goals and meet regulatory requirements.

In the second part of the paper, we seek to compare the risk culture characteristics of insurance firms with those of banks. First, from the content of their 10-K reports, we find that on average firms in both industries have more discussion of their leadership, strategy and portfolio features with a positive sentiment, with fewer paragraphs showing a discussion of constraint sentiment towards recruitment and decisions risk culture key drivers. This indicates that finance industry firms generally express a strong confidence in their risk governance as well as in their access to educated and trained risk personnel with the right skill sets for risk management.

For firms in both industries, uncertainty and litigious sentiments for leadership and strategy, and constraining sentiment for reward and portfolio play a role in distinguishing between firms for defining risk culture. For the insurance industry, we find that in addition to the portfolio risk culture key driver, recruitment key driver is also a strong determinant of the risk culture. Features regarding uncertainty and litigious sentiment for recruitment are strongly picked up by the principal components. This finding shows that the insurance industry is heavily dependent on specialized skilled professionals, such as actuaries, for its risk decisions, practices and management.

Our paper makes three main contributions. First, we contribute to developing literature on using big data to classify risk culture. Our paper is the first to define risk culture of insurance firms from the textual contents within their 10-K filings using unsupervised machine learning techniques. We go beyond the traditional approaches of interviews and surveys, and use advanced data mining and machine learning cluster techniques to define the risk culture for a group of insurance firms, which enables us to discover any hidden pattern or structure within the data.

Second, we propose machine learning as a more advanced approach that can be adopted by regulators to improve their supervisory role within the financial industry. The risk culture of banks and insurance firms change over time and across firms. Thus monitoring and keeping track can be quite challenging for regulators even with the guide of a risk culture framework. With advanced technologies being developed in the computer science and natural language processing fields, activities where cumbersome human effort were required can leverage the use of machine learning algorithms.

Last, we contribute to the ongoing debate in the U.S. to repeal or amend regulations, especially the Dodd-Frank Act, within the financial industry. Given that these regulations impact the risk behavior of firms, it becomes crucial to understand the characteristics of risk culture of firms within the financial industry. Identifying the features that differentiate the banking and insurance industries can be a guide when setting and amending regulations for both industries. For instance, given that leadership plays a role in defining risk culture for firms within the financial industry, rules that supervise the risk behavior of executives and directors must be given priority.

The rest of the paper is organized as follows. We begin Section 2 with a discussion of our data sources including the text data and text features. In Section 3 we discuss the methodology for parsing data and extracting risk culture features from the 10-K reports and present the unsupervised machine learning models for feature selection and clustering. In Section 4, we define the risk culture classes and present our clustering results in Section 5. In Section 6, we compare the differences and similarities in risk culture between insurance firms and banks. Finally, we conclude and discuss future research directions in Section 7.

## 2. Data Description

For our sample of insurance firms, we identify 313 insurance firms between 2000 and 2017, with their available CIKs in the SEC Edgar system. Going by the SIC code, there are eight industrial sectors of the insurance industry. However to ensure each sector has a good number of firms for our machine learning, we merge the 'Surety' and 'Title', and 'Accident and Health' with 'Hospital and Medical' since they are more closely related in roles. We also drop 'Carriers' which we observe to be an outlier with only two firms. Hence we end up with five main industry sectors.

Panel A of Table 1 shows the distribution of the insurance firms across our five insurance sectors. We find that although 'Fire, Marine and Casualty' forms the largest sector by size, 'Life' is the largest sector by total asset with 'Agents' the smallest sector by total assets. From our data, the five largest insurance firms by total asset are American International Group Inc., Metlife Inc., Prudential Financial Inc., Metropolitan Life Insurance Co. and Hartford Financial Services Group Inc. and are all from the Life and Casualty sectors.

[Insert Table 1]

The 10-K reports filed with the SEC provide us with a good source of textual information with a longer time span. In addition to providing a rich content to assess the financial health of the firm, the report has twenty sections, including the Item 1A section, which discusses the risk factors of the firm and Item 7 discussing the operations of the firm and issues faced in operation. 10-K reports have also been used in many textual analysis research studies in finance (Kaplan and Zingales 1997, Li 2008, Bobnaruk et al. 2015). Thus to learn about the risk culture of the firm we download from the SEC Edgar 2,853 10-K files for our sample of insurance firms.

To extract features from the text documents for our machine learning, we use the two-dimensional risk culture dictionary proposed by Gupta and Owusu (2019). The first dimension of the dictionary is from a seven-feature risk culture framework thoroughly built from three different risk culture frameworks (McConnell 2013, Fritz-Morgenthal et al. 2016, Sheedy et al. 20017). Table A1 in the appendix provides reference to proposed risk culture framework. Each feature covers the risk culture attributes as recognized from the IIF's definition of risk culture. The seven features include: Leadership, Strategy, Decision, Control, Reward, Recruitment and Portfolio. Key words are identified for the each of the seven risk culture features. In all there are 600 risk culture words.

The second dimension associates a sentiment to the risk culture features using the Loughran and McDonald (2014) sentiment dictionary. Meaning for each feature, there is a related sentiment to be determined. Five sentiments are used: Positive, Negative, Uncertainty, Litigious and Constraining. Hence we have 35 features in the two-dimensional dictionary.

## 3. Methodology: Textual Analysis and Machine Learning Techniques

In this section, we explain the automated text extraction of risk culture features from the textual content of the 10-K documents and the application of unsupervised machine learning algorithms to identify the risk culture of the firms.

### 3.1. Risk Culture Features Extraction

The 10-K reports are downloaded as raw files and thus require further preprocessing into a uniform and clean text-containing version, called a corpus, before any meaningful text extraction can be done. This is a very important stage of every textual analysis process. The document preprocessing begins with cleaning and removing irrelevant objects such as tables, figures and other quantitative data, which do have any risk culture or sentiment meaning.

After removing irrelevant non-text objects, the next stage is tokenization. In our study, we are interested in analyzing the risk culture at the paragraph level. Hence we adopt a paragraph 'topic' modeling approach because we believe each paragraph discusses or relates to one risk culture feature. As such we first tokenize the corpus into paragraphs before we break each paragraph into words for use in a bag-of-words extraction approach.

Lastly, we remove stop words. In the English language, some common words such as {'a,' 'the,' 'that,' 'about,'} etc. are referred to as stop words in text mining and are commonly removed before text extraction because they do not convey any meaning in the corpus. After taking out the stop words, we stem and reduce each remaining word into a lowercase.

Now that we have a clean corpus, we move on to extract our features. For each paragraph, we extract the frequency of the dictionary words by each dictionary dimension and driver. Note that, we consider a paragraph to be made up of two or more sentences. We then classify a paragraph under a risk culture dimension and sentiment dimension based on the highest frequency of words from the respective dictionary. For instance, a paragraph will be classified as covering 'Positive Leadership' if the highest frequency of words comes from the Leadership feature of the risk culture dimension, and the Positive feature of the sentiment dimension. We do this classification for all paragraphs within the corpus. As an example, Figure 1 shows a plot of the number of paragraphs across features for a corpus representing Aflac Incorporation's 2009 10-K filing.

[Insert Figure 1]

Panel B of Table 1 summarizes the features across the paragraphs. Ranking our data by the most covered paragraph by risk culture feature, we find paragraph discussions on Positive Strategy to be the highest with a mean of 85 paragraphs, followed by Positive Leadership with a mean of 66 paragraphs. Strategy appears three times in the top 5 most discussed paragraph topics, with Litigious and Uncertainty Strategy also showing up in addition to Positive Strategy. In the least risk culture feature discussed, Constraining Decision and Recruitment are low with an average of one paragraph coverage. Constraining Control is more prominent within the Control risk culture attributes.

Thus we conclude that on average, the 10-K reports of insurance firms have more paragraph discussions on their risk strategies, with views that show the strategy to be positive as well as with some uncertainty and litigious sentiments. These reports also show fewer discussions around constrains on their risk decisions and recruitments. This may be due to the fact that insurance companies hire actuaries who provide special skills in risk decisions and modeling, thus, making them more vocal in the strategies they take as well as less concern about constrains on finding the right people for the risk tasks.

Next, we move on to classify our sample of text documents into different risk culture groups.

### 3.2. Unsupervised Machine Learning

For our main machine learning analysis, we use an unsupervised machine learning approach. An unsupervised approach allows for pattern identification and classification of a data set with no labels. We use two unsupervised approaches in our analysis. First we run a feature reduction using a principal component analysis. Second we identify patterns within our data to forms clusters using the reduced features.

#### 3.2.1. Principal Component Analysis

Principal Component Analysis enables us to reduce the dimensionality of feature space. In machine learning, this helps improve cluster analysis and also helps with visualization of the data. Before applying the PCA, we normalize our features by the mean and standard deviation across of features. This step is very important ensuring the unsupervised results are consistent. Figure 2, shows the scree plot of the eigenvalues after running the PCA. Applying Kaiser's Rule, we reduce the 35 features into 5 principal components, since 5 components have eigenvalues greater than 1. Thus we explain approximately 81% of the variation in the data.

[Insert Figure 2]

Next, we use a rotated component matrix to interpret the correlation between the five principal components and the original features. Panel A of Table 2 summarizes the results, with correlation above 0.25 in bold. We find that at lease one risk culture dimension and sentiment dimension is covered by the five components. In Panel B, we document the representation of each component. Component 1 loads more on the Uncertainty Leadership, Strategy, Decision, Recruitment and Negative Recruitment and explains 62% of the variation in the data. Component 2 loads more on the Positive Strategy, Decision, Control and Reward, as well as Negative and Uncertainty Reward with 6% explanation in variation. Component 3 loads on all litigious risk features except Decision and Control. Also Component 4 also loads on constraining risk features with the exception of Decision and Recruitment. Lastly, Component 5 loads on all sentiments relating to the Portfolio risk feature and explains 3% of the variation in the data.

[Insert Table 2]

In summary, the unsupervised PCA reduces the 35 features into 5 components, which loads on 24 of the features shown in Table 3.

[Insert Table 3]

3.2.2.Clustering Analysis

In our second unsupervised approach, we use a two-stage cluster analysis (Ketchen and Shoot 1996; Short and Ketchen 2007) to classify the text documents into different risk culture groups. The two-stage cluster analysis helps to address two main challenges in cluster analysis. First identifying the number of clusters to form, and second determining the starting centroid for generating cluster assignments.

The first stage of the cluster analysis uses a hierarchical clustering technique to determine the number of clusters to form and the starting centroids for the clusters. A hierarchical clustering groups similar objects into clusters and ensures that each cluster is different from the others and members within the clusters are very much similar. We use the Ward's Method (Ward 1963) for the hierarchical clustering. From the dendrogram displayed in Figure 3, we form 3 distinct groups {G1-G3}, {G4-G8}, {G9-G13} of almost similar size. We use the centroids of these groups as the starting centroids in the second clustering stage.

[Insert Figure 3]

In the second stage, a non-hierarchical clustering is analyzed. We use the K-means clustering technique which partitions our sample of banks into ***k***-groups and assigns each observation into one of the ***k*** clusters using the observations distance from each of the ***k*** means centroids. In our studies, we form 3 distinct groups, as determined for the first stage. Results for the k-means clustering are summarized in Table 3.

In Panel A, we find that cluster 3 contains 212 (67%) firms, cluster 1 has 58 (18.53%) and cluster 2 has 43 (13.74%) firms. Panel B summarizes the distribution by industry sector. We find more than 50% of firms within each industry sector to be distributed in cluster 3.

[Insert Table 4]

As a final cluster analysis, we consider clustering within each of the sector separately. Comparing the results in Panel C to that of Panel B, we find both cluster approaches to be comparable. Thus we proceed with interpreting our results at the full data level.

## 4. Results

After classify each firm into a distinct cluster, we move on to define the risk culture characteristics of each cluster and validate our definitions using some firm characteristics that are expected to align with risk culture.

### 4.1. Defining Risk Culture Classes

From Section 3.2.1, the results of the PCA showed that reducing the 35 risk culture features into 5 components correlated with 24 risk culture features. Therefore, we use these 24 risk culture features or characteristics to define the risk culture for each cluster. We find the mean of the risk culture features for each cluster and the overall mean of the risk culture features across clusters. Then we test for the difference in mean between each cluster mean and the overall clusters mean. Table 4 summarizes the results of the t-test.

[Insert Table 5]

Comparing the mean of cluster 3 to the overall mean, we find that cluster 3 has risk features with means lower than the overall mean. These are significant at the 1% significance level. They have lower uncertainty, litigious, constraining and negative risk features. We also find them to have the least mean in positive risk features. Due to its lower mean in litigious and negative risk features, we label cluster 3 as having a good risk culture.

We find the mean features of cluster 1 to be above the overall mean features. However, this is below the mean features of cluster 2, which is also higher t-statistics than that of the overall means. Thus, we label cluster 2 as having a fair risk culture and cluster 2 as having a poor risk culture. Clearly, both clusters have high litigious, uncertainty and negative risk cultures, with cluster 2 being worse off.

Next, we validate our cluster labels using some firm characteristics such as the financial performance and firm governance characteristics.

### 4.2. Validation of Cluster Labels

Previous studies show a positive relationship between a firm's corporate culture and financial performance. Thus we expect firm's with a sound risk culture to have good financial performance with better governance than firms who have a poor risk culture. Thus to validate the results of our clustering and risk culture labels, we use the return on asset (ROA), return on equity (ROE) and Tobin's Q to validate for the financial performance. For the governance characteristics, we consider the number of independent and female directors on the board, the nationality mix, the network size and the salary of the directors.

From Panel A of Table 6, we find that indeed the insurance firms within cluster 3, which was labeled as having a good risk culture, have higher financial performance ratios than banks in clusters 1 and 2 labeled fair and poor respectively. The poor risk culture also shows the lowest financial ratios. Thus we conclude that our risk culture clustering and labeling holds using our first validation tests.

In Panel B, we also compare the board's governance characteristics. Again we expect firms with sound risk culture to also have good governance. Literature on board governance shows that independent directors and female director representation on boards improves governance. Validating this in our clustering analysis, we find that our insurance firms within the good risk culture cluster have more independent and female directors on their board than within the poor risk culture group. The firms in the fair risk culture cluster have the highest number of independent and female directors.

We also consider other characteristics of the board such as the nationality mix of directors and network size. Pan et al. (2016) in their study show that the initial formation of risk culture can be related to the variation in cultural origins. Hence we use the nationality mix of and network size of the board to proxy for a variation in cultural origins. Directors with large network sizes are more likely to bring external influence on their boards. We find that finds within our poor risk culture cluster have no nationality mix on their boards and the lowest network size. On the other hand, firms in the sound risk culture clusters show some variations in nationality of the board members. Which we interpret as a good sign of developing a well-balanced risk culture.

Lastly, we compare the compensation of directors between the three clusters. We find that the sound risk culture clusters have higher director salary than the poor risk culture cluster.

## 5. Comparing Risk Culture of Insurance and Banking Industry

In this last section, we consider the risk culture of the two main players within the financial industry, banks and insurance firms. We reference our earlier work in Gupta and Owusu (2019) for the risk culture characteristics of banks. A summary of results on banks' risk culture is provided in Table A2 in the appendix.

First, we consider the risk culture textual contents of the 10-K reports. We find that, both insurance and banks have more paragraph discussions on the positive risk leadership, strategy and portfolio management. Again, they both have fewer discussions on their constrained risk decision making and recruitment. This conveys some positive confidence from both industries, in setting up right risk appetites and improving risk culture. It is also an indicator that the finance industry has access to and have available educated and trained risk personnel with the right skill sets for risk management.

Next we look at the features relevant for identifying the risk culture in both industries. Comparing Table 3 with Table A1, we find that uncertain and litigious leadership and strategy, as well as constraining reward and portfolio play a role in defining risk culture in both industries. These features are likely to be overlooked if were to simply use a bag-of-words approach to capture how often related key words appear in the document. For instance, from our summary statistics in Table 1, we find that constraining reward and portfolio are not mostly discussed within the documents or not being disclosed often, although they show up as important risk features from our machine learning analysis.

One main difference we find is that, whiles uncertain and litigious recruitment and risk training of staff are identified as important features to define risk culture within insurance companies, these features do not show up for banks. This finding shows that the insurance industry is heavily dependent on specialized skill sets, which are mostly seen in actuaries for its risk decisions and managements. Regulators should therefore pay close attention to employees risk training capabilities when monitoring the firms risk culture. Another difference between the two industries is that, the risk culture of insurance firms can also be identified through their positive risk strategies, decisions and controls as well as both positive and negative risk reward. Again drawing a clear difference between insurance

## Conclusion

This paper uses text analysis and machine learning algorithms to identify the risk culture for a sample of insurance companies. We apply K-means cluster analysis to 2,853 10-K filing documents to form three distinct risk culture clusters. We label the clusters as good, fair and poor risk culture clusters. We find that insurance firms in the good risk culture cluster have high profitability ratios and good governance characteristics compared to insurance firms in the poor risk culture cluster.

Comparing the risk culture of insurance firms to the earlier research study for banks (Gupta and Owusu 2019), we find that uncertainty and litigious sentiments related to risk training, education and recruitment, strongly defines the risk culture of insurance firms. This indicates the importance of having trained risk professionals, like actuaries, within the insurance industry to engage in risk decision and risk management. On the other hand, we find that uncertainty, litigious and constraining risk control and reward features define the risk culture for banks. We also find that both banks and insurance firms disclose more of their positive risk leadership, strategies and portfolio from analysis of the textual contents of their 10-K documents.

Our paper emphasizes the need for regulators to define and understand the characteristics of the risk culture of firms within the financial industry. Identifying the features that define risk culture for financial firms, can help to set effective rules and regulations, and amending existing ones. We also propose that the regulators explore machine learning algorithms and other advanced technologies being developed in computer science and natural language processing areas, to identify hidden patterns or structures within all the regulatory filings documents.

## References


Bail, C. A. (2014). The cultural environment: Measuring culture with big data. Theory and Society, 43(3-4), 465-482.

Bodnaruk, A., Loughran, T. and McDonald, B. (2015). Using 10-k text to gauge financial constraints. Journal of Financial and Quantitative Analysis, 50(4), 623-646.

Evans, J. A., and P. Aceves. (2016). Machine translation: mining text for social theory. Annual Review of Sociology 42, 21-50.

Federal Insurance Office (2014). Annual Report on the Insurance Industry (PDF). Washington, D.C.: U.S. Department of the Treasury. p. 45.

Fergussen, R. W. (2018). Culture and Financial Crisis. Available at: https://www.linkedin.com/pulse/culture-financial-crisis-roger-w-ferguson-jr-/

Ghaziani, A. (2009). An "amorphous mist"? The problem of measurement in the study of culture. Theory and Society, 38 (6), 581-612.

Gupta, A. and Owusu, A. F. (2019). Identifying the risk culture of banks using machine learning. Available on SSRN.

Hartigan, J. A. and Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. Applied Statistics, 28, 100-108.

Institute of International Finance (2009). Reform in the financial services industry: Strengthening Practices for a More Stable System. Institute of International Finance, Washington. Available at: https://www.iif.com/system/files/iifreport_reformfinancialservicesindustry_1209.pdf

Kaplan, S. and Zingales, L. (1997). Do investment-cash flow sensitivities provide useful measures of financing constraints? The Quarterly Journal of Economics, 112(1), 169-215.

Ketchen, D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: an analysis and critique. Strategic management journal, 17(6), 441-458.

King, G. (2011). Ensuring the Data Rich Future of the Social Sciences. Science, 331(11), 719-721.

KPMG (2016). Building a strong insurance risk culture. Available at: https://home.kpmg/content/dam/kpmg/at/images/Themen/corporate-risk-management/building-a-strong-insurance-risk-culture.pdf?logActivity=true

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D. Van Alstyne, M. (2009). Life in the network: the coming of computational social science. Science, 323(5915), 721-723.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics, 45(2), 221-247.

Loughran, T. and McDonald, B. (2014). Measuring Readability in Financial Disclosures. The Journal of Finance, 69(4), 1643-1671.

National Association of Insurance Commissions (2017). Post-Crisis Financial System Reform: Impact on the U.S. Insurance Industry in the Evolving Regulatory Landscape. Available at: https://www.naic.org/capital_markets_archive/170120.htm

Pan, Y., Siegel, S., and Wang, T.Y. (2015). Corporate Risk Culture. Available at SSRN: https://ssrn.com/abstract=2675594

Ross, G. and Schwartzman, J. A. (2010). The Dodd-Frank Act and the insurance industry: Strategic considerations of U.S. financial reform. Milliman, New York. Available at: http://www.milliman.com/insight/insurance/The-Dodd-Frank-Act-and-the-insurance-industry-Strategic-considerations-of-U_S_-financial-reform/

Sheedy, E. and Griffin B. (2018). Risk governance, structures, culture, and behavior: A view from the inside. Corporate Governance: An International Review, 26, 4–22.

Short, J. C., Ketchen Jr, D. J., Palmer, T. B., & Hult, G. T. M. (2007). Firm, strategic group, and industry influences on performance. Strategic management journal, 28(2), 147-167.

Stulz, R. M. (2014). Governance, risk management, and risk-taking in banks. Working Paper 20274. National Bureau of Economic Research, Massachusetts

Ward, J. H., Jr. (1963), "Hierarchical Grouping to Optimize an Objective Function", Journal of the American Statistical Association, 58, 236–244.

Williams, R. (1983). Culture and society, 1780-1950. Columbia University Press.

**Figure 1: Feature Extraction for Aflac Incorporation's 2009 Annual Report**

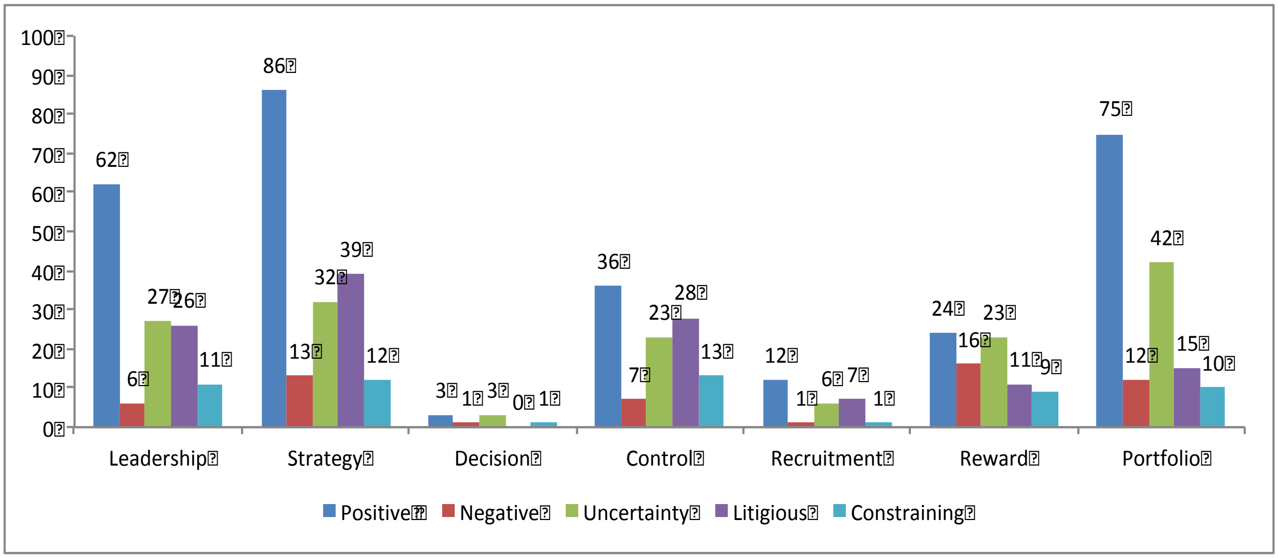

**Figure 2: Principal Component Analysis**

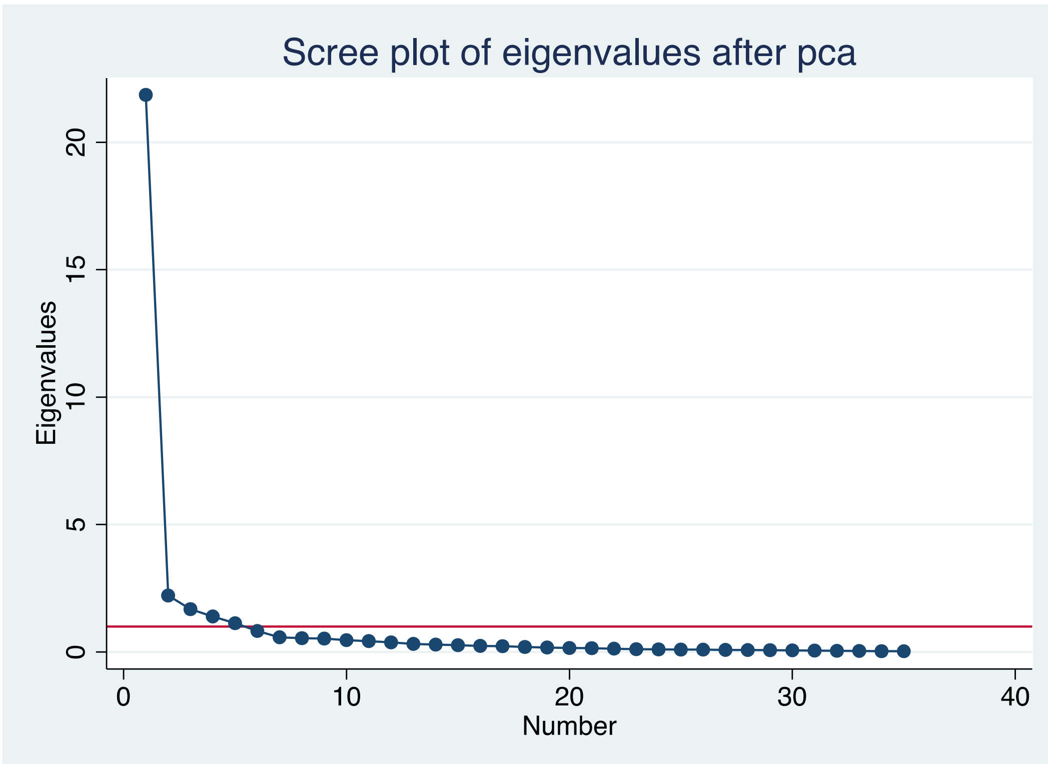

**Figure 3: Ward's Hierarchical Clustering**

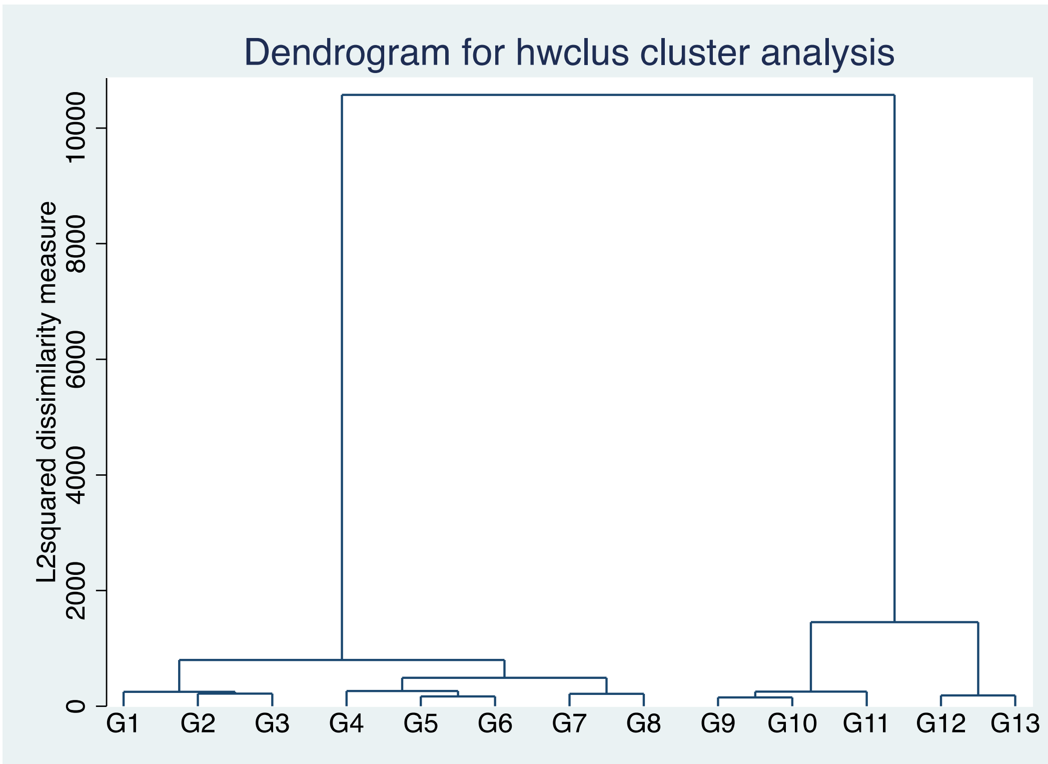

**Table 1: Summary Statistics**

Panel A: Distribution of Insurance Industries

| Industry | SIC Code | Number of Firms | Proportion | Total Assets |
|---|---|---|---|---|
| Fire, Marine, Casualty | 6331 | 134 | 42.81 | 24203.94 |
| Life | 6311 | 54 | 17.25 | 82429.27 |
| Agents | 6411 | 52 | 16.61 | 4007.959 |
| Accident and Health/ Hospital and Medical | 6321/6324 | 40 | 12.78 | 18443.46 |
| Surety and Title | 6351/6361 | 33 | 10.54 | 9913.335 |
| Total | | 313 | 100 | 28699.19 |

Panel B: Summary Statistics of Extracted Risk Culture Features

| Variable | Observation | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|---|
| Number of Paragraph | 2853 | 509.53 | 277.43 | 42 | 1949 |
| Leadership Positive | 2853 | 65.66 | 46.51 | 1 | 482 |
| Leadership Negative | 2853 | 16.01 | 14.35 | 0 | 168 |
| Leadership Uncertainty | 2853 | 38.08 | 29.18 | 0 | 296 |
| Leadership Litigious | 2853 | 28.87 | 21.17 | 1 | 318 |
| Leadership Constraining | 2853 | 11.67 | 9.12 | 0 | 93 |
| Strategy Positive | 2853 | 84.67 | 50.82 | 5 | 476 |
| Strategy Negative | 2853 | 25.77 | 19.24 | 0 | 216 |
| Strategy Uncertainty | 2853 | 56.07 | 37.08 | 1 | 398 |
| Strategy Litigious | 2853 | 54.70 | 40.30 | 2 | 481 |
| Strategy Constraining | 2853 | 13.39 | 9.93 | 0 | 111 |
| Decision Positive | 2853 | 10.70 | 16.55 | 0 | 194 |
| Decision Negative | 2853 | 3.22 | 4.58 | 0 | 50 |
| Decision Uncertainty | 2853 | 11.67 | 19.00 | 0 | 198 |
| Decision Litigious | 2853 | 6.89 | 12.43 | 0 | 167 |
| Decision Constraining | 2853 | 1.08 | 1.91 | 0 | 23 |
| Control Positive | 2853 | 38.79 | 29.19 | 0 | 322 |
| Control Negative | 2853 | 17.64 | 16.95 | 0 | 203 |
| Control Uncertainty | 2853 | 28.44 | 22.95 | 0 | 241 |
| Control Litigious | 2853 | 42.75 | 31.53 | 2 | 435 |
| Control Constraining | 2853 | 21.80 | 25.10 | 0 | 319 |
| Recruitment Positive | 2853 | 10.49 | 14.38 | 0 | 141 |
| Recruitment Negative | 2853 | 3.87 | 6.57 | 0 | 85 |
| Recruitment Uncertainty | 2853 | 10.07 | 14.71 | 0 | 192 |
| Recruitment Litigious | 2853 | 8.67 | 17.73 | 0 | 251 |
| Recruitment Constraining | 2853 | 1.76 | 2.93 | 0 | 49 |

| | | | | | |
|---|---|---|---|---|---|
| Reward Positive | 2853 | 34.75 | 24.87 | 0 | 254 |
| Reward Negative | 2853 | 23.78 | 22.49 | 0 | 255 |
| Reward Uncertainty | 2853 | 20.05 | 13.37 | 0 | 128 |
| Reward Litigious | 2853 | 18.35 | 14.26 | 0 | 167 |
| Reward Constraining | 2853 | 9.56 | 7.41 | 0 | 73 |
| Portfolio Positive | 2853 | 49.22 | 37.93 | 0 | 284 |
| Portfolio Negative | 2853 | 17.13 | 15.64 | 0 | 214 |
| Portfolio Uncertainty | 2853 | 21.83 | 17.37 | 0 | 115 |
| Portfolio Litigious | 2853 | 15.07 | 13.86 | 0 | 147 |
| Portfolio Constraining | 2853 | 8.94 | 8.08 | 0 | 68 |

**Table 2: Principal Component Analysis – Rotated Components on Features**

| Variable | Comp1 | Comp2 | Comp3 | Comp4 | Comp 5 | Unexplained |
|---|---|---|---|---|---|---|
| Leadership Positive | | | | | | 0.1350 |
| Leadership Negative | | | | | | 0.2513 |
| Leadership Uncertainty | 0.2982 | | | | | 0.0854 |
| Leadership Litigious | | | 0.2862 | | | 0.1630 |
| Leadership Constraining | | | | 0.3986 | | 0.2501 |
| Strategy Positive | | 0.288 | | | | 0.1690 |
| Strategy Negative | | | | | | 0.1991 |
| Strategy Uncertainty | 0.3561 | | | | | 0.2361 |
| Strategy Litigious | | | 0.3996 | | | 0.1683 |
| Strategy Constraining | | | | 0.4508 | | 0.1779 |
| Decision Positive | | 0.3362 | | | | 0.0949 |
| Decision Negative | | | | | | 0.2059 |
| Decision Uncertainty | 0.2724 | | | | | 0.1376 |
| Decision Litigious | | | | | | 0.1420 |
| Decision Constraining | | | | | | 0.4413 |
| Control Positive | | 0.3285 | | | | 0.1147 |
| Control Negative | | | | | | 0.1441 |
| Control Uncertainty | | | | | | 0.1403 |
| Control Litigious | | | | | | 0.1850 |
| Control Constraining | | | | 0.3082 | | 0.1238 |
| Recruitment Positive | | | | | | 0.1324 |
| Recruitment Negative | 0.3307 | | | | | 0.1774 |
| Recruitment Uncertainty | 0.266 | | | | | 0.1058 |
| Recruitment Litigious | | | 0.3308 | | | 0.1605 |
| Recruitment Constraining | | | | | | 0.3539 |
| Reward Positive | | 0.4099 | | | | 0.1200 |
| Reward Negative | | 0.3219 | | | | 0.2092 |
| Reward Uncertainty | | 0.2845 | | | | 0.3192 |
| Reward Litigious | | | 0.5101 | | | 0.2545 |
| Reward Constraining | | | | 0.5484 | | 0.2284 |
| Portfolio Positive | | | | | 0.3957 | 0.1597 |
| Portfolio Negative | | | | | 0.2502 | 0.2283 |
| Portfolio Uncertainty | | | | | 0.4056 | 0.2373 |
| Portfolio Litigious | | | 0.4414 | | 0.2796 | 0.2799 |
| Portfolio Constraining | | | | 0.2901 | 0.4424 | 0.2519 |

**Table 3: Selected Risk Culture Features from Principal Component Analysis**

| Component | Selected Features |
|---|---|
| 1 | Uncertainty - Leadership, Strategy, Decision, Recruitment Negative - Recruitment |
| 2 | Positive - Strategy, Decision, Control<br>Reward - Negative, Positive, Uncertainty |
| 3 | Litigious - Leadership, Strategy, Recruitment, Reward, Portfolio |
| 4 | Constraining - Leadership, Strategy, Control, Reward, Portfolio |
| 5 | Portfolio - Positive, Negative, Uncertainty, Litigious, Constraining |

**Table 4: K-Means Clustering Analysis**

Panel A: Cluster Distribution

| Cluster | Freq. | Percent | Cum. |
|---|---|---|---|
| 1 | 58 | 18.53 | 18.53 |
| 2 | 43 | 13.74 | 32.27 |
| 3 | 212 | 67.73 | 100 |
| Total | 313 | 100 | |

Panel B: Distribution by Industry and Cluster

| **Cluster** | **Agents** | **Fire, Marine, Casualty** | **Accident and Health/ Hospital and Medical** | **Life** | **Surety and Title** |
|---|---|---|---|---|---|
| 1 | 11 | 23 | 9 | 10 | 5 |
| 2 | 6 | 15 | 6 | 11 | 5 |
| 3 | 35 | 96 | 25 | 33 | 23 |
| Total | 52 | 134 | 40 | 54 | 33 |

Panel C: K-means Clustering by Industry

| **Cluster** | **Agents** | **Fire, Marine, Casualty** | **Accident and Health/ Hospital and Medical** | **Life** | **Surety and Title** |
|---|---|---|---|---|---|
| 1 | 20 | 26 | 9 | 11 | 15 |
| 2 | 17 | 15 | 6 | 31 | 5 |
| 3 | 15 | 93 | 25 | 12 | 13 |
| Total | 52 | 134 | 40 | 54 | 33 |

**Table 5: Defining Risk Culture for Clusters**

| | | | Cluster 1 [N=58] | | | Cluster 2 [N=43] | | | Cluster 3 [N=212] | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Overall Mean | Mean | T -statistic | P -value | Mean | T -statistic | P -value | Mean | T -statistic | P -value |
| PC1 | Leadership Uncertainty | 0.0888 | 0.1089 | 6.4648 | 0.0000 | 0.1890 | 11.0831 | 0.0000 | 0.0629 | -23.2327 | 0.0000 |
| | Strategy Uncertainty | 0.1216 | 0.1391 | 4.1890 | 0.0001 | 0.2023 | 7.9023 | 0.0000 | 0.1005 | -11.9629 | 0.0000 |
| | Decision Uncertainty | 0.0344 | 0.0474 | 5.3302 | 0.0000 | 0.1079 | 10.0989 | 0.0000 | 0.0159 | -25.0840 | 0.0000 |
| | Recruitment Uncertainty | 0.0275 | 0.0372 | 6.7375 | 0.0000 | 0.0826 | 10.3541 | 0.0000 | 0.0136 | -23.5260 | 0.0000 |
| | Recruitment Negative | 0.0124 | 0.0173 | 5.3013 | 0.0000 | 0.0415 | 7.0522 | 0.0000 | 0.0052 | -27.6331 | 0.0000 |
| PC2 | Strategy Positive | 0.1999 | 0.2352 | 4.0789 | 0.0001 | 0.3700 | 12.5510 | 0.0000 | 0.1557 | -17.7664 | 0.0000 |
| | Decision Positive | 0.0372 | 0.0546 | 5.4317 | 0.0000 | 0.1287 | 12.3503 | 0.0000 | 0.0138 | -39.7526 | 0.0000 |
| | Control Positive | 0.0957 | 0.1214 | 6.1282 | 0.0000 | 0.2138 | 13.2128 | 0.0000 | 0.0647 | -22.9173 | 0.0000 |
| | Reward Positive | 0.0878 | 0.1091 | 3.5679 | 0.0007 | 0.1832 | 14.2813 | 0.0000 | 0.0626 | -18.4153 | 0.0000 |
| | Reward Negative | 0.0666 | 0.0872 | 3.1311 | 0.0027 | 0.1540 | 10.2270 | 0.0000 | 0.0432 | -20.3051 | 0.0000 |
| | Reward Uncertainty | 0.0480 | 0.0573 | 4.2032 | 0.0001 | 0.0864 | 8.5868 | 0.0000 | 0.0377 | -11.1662 | 0.0000 |
| PC3 | Leadership Litigious | 0.0711 | 0.0876 | 4.0969 | 0.0001 | 0.1392 | 7.8714 | 0.0000 | 0.0528 | -13.8180 | 0.0000 |
| | Strategy Litigious | 0.1348 | 0.1701 | 3.5629 | 0.0007 | 0.2619 | 9.1139 | 0.0000 | 0.0994 | -13.9111 | 0.0000 |
| | Recruitment Litigious | 0.0263 | 0.0401 | 4.6893 | 0.0000 | 0.0857 | 6.8672 | 0.0000 | 0.0105 | -25.5513 | 0.0000 |
| | Reward Litigious | 0.0412 | 0.0480 | 2.9105 | 0.0051 | 0.0693 | 6.5302 | 0.0000 | 0.0336 | -7.7164 | 0.0000 |
| | Portfolio Litigious | 0.0305 | 0.0347 | 2.3255 | 0.0236 | 0.0494 | 5.0610 | 0.0000 | 0.0255 | -5.4550 | 0.0000 |
| PC4 | Leadership Constraining | 0.0257 | 0.0305 | 4.1818 | 0.0001 | 0.0476 | 7.2502 | 0.0000 | 0.0199 | -10.9872 | 0.0000 |
| | Strategy Constraining | 0.0326 | 0.0371 | 2.9876 | 0.0041 | 0.0599 | 6.2766 | 0.0000 | 0.0259 | -8.8598 | 0.0000 |
| | Control Constraining | 0.0629 | 0.0826 | 5.5641 | 0.0000 | 0.1713 | 9.4528 | 0.0000 | 0.0355 | -25.8079 | 0.0000 |
| | Reward Constraining | 0.0217 | 0.0233 | 1.2478 | 0.2172 | 0.0351 | 4.1282 | 0.0002 | 0.0185 | -5.7467 | 0.0000 |
| | Portfolio Constraining | 0.0170 | 0.0204 | 2.8011 | 0.0069 | 0.0270 | 4.5221 | 0.0000 | 0.0140 | -6.0616 | 0.0000 |
| PC5 | Portfolio Uncertainty | 0.0422 | 0.0474 | 2.4294 | 0.0183 | 0.0664 | 5.8235 | 0.0000 | 0.0359 | -6.2314 | 0.0000 |
| | Portfolio Positive | 0.0993 | 0.1185 | 3.8929 | 0.0003 | 0.1798 | 8.7230 | 0.0000 | 0.0777 | -10.2567 | 0.0000 |
| | Portfolio Negative | 0.0369 | 0.0488 | 5.0501 | 0.0000 | 0.0756 | 8.5116 | 0.0000 | 0.0258 | -14.8487 | 0.0000 |

**Table 6: Cluster Validation Results**

Panel A: Financial Performance

| Cluster | Return on Asset | Return on Equity | Tobin's Q |
|---|---|---|---|
| 1 - Fair | 0.0335 | 0.0612 | 1.2329 |
| 2 - Poor | 0.0170 | 0.0186 | 1.1679 |
| 3 - Good | 0.0475 | 0.0803 | 1.2628 |
| Total | 0.0434 | 0.0734 | 1.2523 |

Panel B: Board Governance

| Cluster | Number of Directors | Number of Independent Directors | Number of Female | Nationality Mix | Network Size | Director Salary |
|---|---|---|---|---|---|---|
| 1 - Fair | 10.3819 | 8.3308 | 1.3913 | 0.0529 | 822.0530 | 157.7097 |
| 2 - Poor | 9.6714 | 4.5143 | 0.7429 | 0.0000 | 654.5286 | 134.2609 |
| 3 - Good | 10.0698 | 7.1286 | 1.1749 | 0.0382 | 674.5765 | 188.2946 |
| Total | 10.1159 | 7.2807 | 1.2032 | 0.0399 | 699.6853 | 181.8952 |

## Appendix

**Table A1: Risk Culture Framework (Table 1 of GO 2019)**

Panel A: Building Risk Culture Framework

| McConnell (2013) | Fritz-Morgenthal, Hellmuth and Packman (2016) | Sheedy, Griffin and Barbour (2017) | Proposed Risk Culture Framework |
|---|---|---|---|
| Leadership | Avoidance/ Manager | Governance | **Leadership** |
| Strategy | Policy | Business Strategy/ Risk Strategy | **Strategy** |
| Decision Making | Manager/ Valued | Cultural Indicators | **Decision** |
| Control | Policy/Manager | Regulatory Requirements | **Control** |
| Recruitment, Training and Competence | Proactive | Employees | **Recruitment** |
| Reward | Valued | Reputation | **Reward** |
| | | Portfolio | **Portfolio** |

Panel B: Key Drivers of Risk Culture Indicators

| Risk Culture Indicator | Key Drivers |
|---|---|
| Leadership | Core Values, Acting with integrity, Planning and Execution, Communication, People Development, Operational Excellence |
| Strategy | Strategic Perspective, Risk Perspective, Resource, Development of the Organization, Risk Appetite, Risk Framework |
| Decision | Informed, Competent, Structured, Empowered, Open to Challenge, Recorded |
| Control | Define and Implement, Reporting (Management Information), Review, Risk Delegation, Risk Limits, Stress Testing |
| Recruitment | Recruitment, Training, Continuous Development, Feedback, Managing Performance, Risk Education |
| Reward | Salary, Bonus and Profit Share arrangements, Recognition, Risk Aligned, Risk Adjusted, Risk Independence |
| Portfolio | Balance Sheet Risk Factors, Quality of Risk Culture, Management of Portfolio |

**Table A2: Selected Risk Culture Features (Table 3 Panel B in GO 2009)**

| Components | Selected Features |
|---|---|
| 1 | Uncertainty - Leadership, Strategy, Control, Reward |
| 2 | Litigious - Leadership, Strategy, Decision, Control, Reward |
| 3 | Constraining - Recruitment, Reward, Portfolio<br>Portfolio – Positive, Negative |