# Sound Risk Culture in Banks

Aparna Gupta and Abena Owusu*

*Lally School of Management, Rensselaer Polytechnic Institute*

January 13, 2018

**Abstract**

In this paper, we examine the effect of different risk cultures on a banks performance, risk-taking incentives and corporate governance. We extend beyond banks' risk culture evaluation and classification using interviews and surveys, and apply textual analysis and cluster analysis to banks' annual 10-K reports. Benefiting from the sentiment dictionary built by Loughran and McDonald (2011) and risk culture framework proposed by McConnell (2013), we develop a two-dimensional dictionary and extract paragraph level features reflecting sentiment and risk culture topics from 10-K reports. We use legal expense data in a regression analysis based supervised learning for feature reduction and implement a k-means clustering to categorize the 10-K reports into three risk culture classes: high, moderate and low risk cultures. For our sample of U.S. bank holding companies, we find that banks in the high and moderate risk culture classes have better performance, lower total compensation for executives and fewer institutional ownership, than banks in the low risk culture cluster. We also find that, the boards of banks in the high and moderate risk are more diverse, have more meetings in a year and decrease the number of options received by non-employee directors. Our results suggest that promoting a sound risk culture can improve financial performance and risk governance within the banking industry.

**Keywords:** Sound risk culture, text mining, cluster analysis, unsupervised learning, supervised learning

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

## 1. Introduction

After the recent global financial crisis, governments and financial regulators are faced with an important role of improving risk management within the banking system and the financial industry as a whole. Efforts by the Financial Stability Board (FSB), the U.K. Financial Conduct Authority and the U.S.'s Financial Industry Regulatory Authority, have focused on the risk culture of firms and how to ensure a 'good' or 'sound' risk culture that it is in line with the risk appetite set by the board of directors (IIF, 2009; FSB, 2011, 2014). Now, changing the risk culture of a bank can be complex and difficult to achieve, and past research studies have proposed risk culture frameworks to guide and assist firms and regulators in their effort to change risk culture (FSA, 2007; McConnell, 2013). However, there is little empirical studies to test the effect of different risk cultures on the performance, risk-taking and risk management of a bank.

In this paper, we take advantage of cluster analysis techniques in machine learning and define three classes of risk culture for a sample of U.S. bank holding companies (BHC) from 2000 to 2016: high, moderate and low risk cultures. We examine how the different risk culture classes impact financial performance, risk-taking incentive structures and corporate governance of banks. The term culture itself is complex, and as Williams (1973, 1983) clearly states "Culture is one of the two or three most complicated words in the English dictionary." Thus to define risk culture in our study, we adopt the Institute of International Finance (IIF) (2009) definition of risk culture as "the norms and traditions of the behavior of individuals and of groups within an organization that determines the way in which they identify, understand, discuss, and act on the risks the organization confronts and the risks it takes."

Again, risk culture, like many other concepts of culture, is intangible and therefore can be difficult to measure and quantify (McConnell, 2013; Ashby, Power and Palermo, 2015; Sheedy, Griffin and Barbour, 2018). To construct our risk culture measure, we go beyond the traditional ways of constructing risk culture data through interviews and surveys, and explore the massive corporate textual data made available to the public. Specifically, we apply textual analysis to the annual 10-K filings available on the Securities and Exchange Commissions' (SEC) EDGAR system. We focus on extracting from the 10-K text documents, key words that are indicators of the risk culture of BHCs in our sample.

To capture the risk culture of a banks from their SEC filings, we extract two dimensions of information; the aspects of risk culture that are mentioned in the reports and what the bank's attitude is on that risk culture aspect. Thus we build two dictionaries, a sentiment dictionary and a risk culture dictionary to create the basis for extracting the risk culture information from the SEC filings. We build the risk culture dictionary using the Risk Culture Framework (RCF) proposed by McConnell (2013). The RCF defines risk culture under six key drivers: 'Leadership', 'Strategy', 'Control', 'Decision Making', 'Recruitment, Training and Competence' and 'Reward'. This framework reflects the managers' values, behavior, management systems, and employees' individual activities that are related to risk culture. In forming our risk culture dictionary, we add a seventh key driver, 'Portfolio' to capture quantitative and financial words on the balance sheet that are related to risk culture.

For the sentiment dictionary, Loughran and McDonald (2014) make a valuable contribution to the literature on text analysis by creating a timely new dictionary re-identifying properties of words in the finance domain. In their latest version, 85,131 terms that have appeared in financial reports are included and a portion of them is categorized into nine sentiments. From the nine sentiments, we choose five sentiment categories relevant to our risk culture measurement objective: 'Negative', 'Positive', 'Uncertainty', 'Litigious' and 'Constraining'. If a sentiment is accompanied by a negation term, then a problem is created such that the sentiment's intent is reversed (Prollochs et al., 2015; Loughran and McDonald, 2014). We address this issue of negation in our study. For each negative and positive word in a sentence, we apply the sentiment property of the TextBlob library in Python to the sentence in which the word appears. We assign positive or negative sentiment to the word based on the dominating sentiment of the sentence.

Now using raw word frequency at the document level causes a large amount of noise. On the other hand, a paragraph usually discusses only one topic and the attitude or sentiment associated with that topic. Thus we introduce a paragraph topic and sentiment model to extract the metrics at the paragraph level. We categorize each paragraph in a corpus into a unique two-dimensional category of sentiment and risk culture. We then count the paragraph frequencies by each two-dimension category. Since there are seven categories in the risk culture dictionary and five categories in the sentiment dictionary, a corpus has in total 35 features, each showing the frequency of paragraphs that belong to each joint category.

Since risk culture is a newly defined concept, no single metric is able to perfectly “supervise" whether the features picked can illustrate the risk culture of a bank. However, McNulty and Akhigbe (2015) econometrically show that high legal expenses incurred by a bank predict weak risk culture to some extent. Hence, we utilize legal expense as a guidance measure in a supervised learning approach to inform us what features significantly uncover aspects of risk culture of a bank. In a linear regression based supervised learning approach, we apply a step-wise regression, ridge regression and LASSO and from the results we identify 16 of the 36 features to be significant in predicting legal expense. We further use principal component analysis to reduce the dimensionality of the 16 features. Once we have the dimensions reduced, we apply an unsupervised k-means clustering to illustrate how banks cluster according to these features explained by the principal components. We cluster the BHCs in our sample into three clusters, and based on the significance of the principal component loading for each cluster, we are able to classify each cluster as having a high, moderate, or low risk culture. We denote high and moderate risk culture as a sound risk culture.

From our cluster analysis, we learn that 10-K reports of banks in Cluster 1 discuss more about uncertain strategy, decision and recruitment, and constraining recruitment. Cluster 2 represents 10-K reports that focus more on litigious and constraining leadership and control, negative strategy and control, negative and uncertain reward, positive portfolio and negative recruitment. Cluster 3 represents 10-K reports that discuss positive decisions and recruitment. Based on the properties of each cluster, we classify Cluster 3 as having a high risk culture, Cluster 1 a moderate risk culture and Cluster 2 a low risk culture. Out of the 5,760 10-K reports in our sample, 1,811 belong to a moderate risk culture class; 1,133 belong to a low risk culture class; and 2,726 belong to a high risk culture class. Hence, we observe that most 10-K reports show a high-risk culture.

Now to establish the effect of risk culture on performance, risk-taking incentives and governance, we need to address the concern for endogeneity in our study. We identify reverse causality as our main source of endogeneity, and a much better way to establish causality will be use an instrument. However, it is difficult to find a valid instrument in our study. Hence, we attempt to establish some form of causality in our study by using panel regression analysis and include firm and year fixed effects to address any omitted and unobservable characteristics within the firm and over time.

Using the results from our cluster analysis, we examine how the different classes of risk culture formed impact performance. We measure the banks' performance using the annual return on asset (ROA) and natural logarithm of Tobin's Q. First we run a univariate ANOVA analysis to determine the difference in means of the three clusters. The results show that there is a significant difference in mean for ROA and Tobin's Q. Next, we run panel regressions with firm and year fixed effects. We find that banks with a high and moderate risk culture have positive and significant ROA and Tobin's Q compared to banks with low risk culture. This result is expected and further confirms that our cluster analysis performed well. Hence, we conclude that a sound risk culture is beneficial to bank's performance.

Second, we test the effect of risk culture on governance and board characteristics. We examine the structure of the board of directors and find that banks with high and moderate risk culture do not impact the gender of the board or specifically the probability of a board having at least one female director. However banks with a moderate risk culture increase the diversity of the board relative to banks with low risk culture. Also the boards of banks with high and moderate risk cultures have more board meetings in a year compared to banks with a low risk culture, which shows that the boards are active and function accordingly. They also have a lower percentage of institutional ownerships relative to low risk cultured banks. We also find a negative but not significant decrease in the number of stocks issued to directors in high and moderate risk culture banks, and a negative and significant decrease in the number of options issued to directors decrease in moderate risk culture banks relative to low risk culture banks. This shows that banks with a sound risk culture reduce the risk-taking incentives of directors.

Last, we examine the impact of risk culture on the compensation structure of the executives of the banks. We define executives to include the CEOs and CFOs of the banks. We find that, the total compensation of executives in high and moderate risk culture firms is lesser compared to banks with low risk cultures. We also find that banks with high-risk culture increase the fraction of incentive pay and decrease the fraction of equity-based compensation for the executives compared to low risk culture banks. Our results suggest that banks with a sound risk culture avoid over paying executives and controls the risk-incentive packages given to executives.

Our paper offers a timely new contribution on identifying and measuring risk culture. First, it relates to previous studies that apply textual analysis to measure corporate risk culture of U.S. publicly traded firms (Fang et al., 2017) and risk culture of large European banks (Bianchi, Farina and Fiordelisi, 2016). We identify seven indicators of risk culture and build a two-dimensional dictionary of words, which provides a feasible and easy interpretation of risk culture aspects from textual sources. Second, we introduce a clustering technique, which enables us to cluster the risk culture of banks into three classes and examine the effect of different risk cultures on performance, risk-taking and governance characteristics.

Third our paper contributes to existing literature on the effect of risk culture on firm performance and corporate decisions. Previous studies show that corporate risk culture affect firm performance and value (Fang et al., 2017), corporate policies such as R&D intensity and acquisitions (Pan et al., 2016), banks' volatility (Collucia, 2017), and bank stability and loan portfolio quality (Bianchi et al. 2016). In our study, we show the effect of three different levels of risk cultures on performance, governance characteristics and compensation structure of banks. We provide empirical evidence that a poor risk culture has a negative impact on performance and provides incentives for excessive risk taking and increases poor governance.

Lastly, our study has some implications for monitoring and regulating risk culture within a bank. Our results provide summaries of features of the different risk culture groups, which can guide regulators to differentiate between risk cultures of banks. Also, although it is difficult to change culture due to entrenchment values and assumptions (McConnell, 2013), our findings can help to assess and determine how a bank's risk culture change over time.

The rest of the paper is organized as follows. In section 2, we discuss the risk culture framework used to build our risk culture dictionary for the textual analysis. In Section 3 we introduce our text data source and risk culture measure. We also briefly discuss the methodology in parsing data, extracting risk culture features from the 10-K reports. We present the supervised and unsupervised machine learning models for feature selection and clustering in Section 4, and present some descriptive statistics. In Section 5, we present and discuss the empirical results. Finally, we conclude with a summary of findings and discuss future research directions in Section 6.

## 2. The Risk Culture Framework

From the IIF's definition of risk culture, the behavior of individuals and groups within an organization form the mainstay of risk culture. Therefore, the study on risk culture should focus on the risk appetite and factors that affect the risk behavior of the employees, managers, executives, and the business units of a firm. McConnell (2013) introduced a Risk Culture Framework (RCF) with six drivers reflecting managers' values and behavior, management system, and employees' individual activities. This framework was proposed as a starting model to assist individual banks in changing their risk culture, which is difficult to achieve. By taking into consideration proposals by regulators and industry bodies regarding risk culture concept, the RCF proposes a rigorous risk culture framework, which expands on the treating customers fairly (TCF) concept developed by the Financial Services Authority (FSA). This framework is a thorough description of risk culture, and includes Geretto and Pauluzzo's (2015) emphasis on values, norms, and practices of the members of the organization and Sheedy et al.'s (2015) four common factors of risk climate: value, manager, proactive, and avoidance.

The six key drivers of the framework include: Leadership, Strategy, Decision making, Controls, Reward, and Recruitment, training and competence. ‘Leadership’ and ‘Strategy’ reflect the tone-at-the-top that describes managers' values and behavior. ‘Decision Making’, ‘Controls’ and ‘Reward’ illustrate the mechanism of the risk management and controlling system. ‘Recruitment, Training and Competence’ provides information of employees' activities. Therefore, this framework offers a top-to-down information collection framework for risk culture measurement. Based on McConnell's (2013) rigorous risk culture framework, our identification of risk culture will take into consideration all the detailed features proposed by the studies and the instruction for soundness of risk culture. Table 1 shows the proposed six key drivers of risk culture by RCF and provides some risk indicators for each driver. Now to enable us to have a complete overview of risk culture within a bank, we add a seventh key driver, ‘Portfolio’ to the framework to capture the bank’s risk appetite from the balance sheet.

## 3. Data and Risk Culture Measure

### 3.1 Text Data

To serve the objective of this research, we need a large textual data set that is consistent across all the banks we study. Most importantly, the textual data should contain content that describes risk culture aspects adequately and should span for a long enough time period so that sufficient observations can be fed into the supervised and/or unsupervised learning models. We rely on the 10-K report filed by U.S. public banks with the SEC as our source for textual data. Form 10-K is an annual report offering a comprehensive summary of the financial performance of a bank. The form consists of up to over twenty sections. Among them, Item 1A discusses risk factors that may affect the bank in the future and Item 7 illustrates management's discussion on the operation of the bank in detail and their analysis of reasons for the issues in operation. According to SEC, these filings are required to be written according to ``plain English rule" for the benefit of shareholders to read and understand (Loughran and McDonald, 2014).

The 10-K report has been used in many research studies as an important source of rich content, especially in the domain of readability and financial health. Li (2008) measures the readability of 10-K reports and finds that the firms with lower earnings are harder to read. Bodnaruk, Loughran and McDonald (2015) parse the 10-K report and find the frequency of constraining words predicts subsequent liquidity events better than traditional financial constraint indexes. Kaplan and Zingales (1997) find that greater sensitivities of a firm's investment cash flow are significantly related to less financial constraint by looking at 10-K reports.

For our sample of bank holding companies, we start with the universe of U.S. BHCs that reported the 'FR-Y9C' between 2000 and 2016. These BHCs have unique 'RSSD ID' numbers assigned to them. Identifying the corresponding 10-K report filed with the SEC creates a challenge since SEC identifies the BHCs using their central index key (CIK). To overcome this challenge, we follow Gupta et al. (2017) and link the Federal Reserve Bank of New York's PERMCO-RSSD dataset to the CRSP-COMPUSTAT merged data, which provides us with a PERMCO-CIK dataset link. We exclude from our sample large banks or banks considered as Too-Big-To-Fail banks since they enjoy some government safety nets[†]. We identify 572 banks as

[†] In a preliminary cluster analysis, which includes the large banks, we find that large banks always cluster together and have a low risk culture. However we are not able to clearly observe their effect on performance due to some benefits that they enjoy due to their size. Results can be made available upon request.

our final sample of BHCs. From the SEC EDGAR system we download 5,670 annual 10-K filings for all the banks in our sample.

### 3.2 Risk Culture Measure

Although cultural sociologists have made progress in clarifying what the concept of culture is, there is still a problem with how to measure culture to make it concrete (Ghanziani, 2009). As analyzed by Sheedy and Griffin (2014), consistent measurements based on existing data are hard to construct because of diverse, “messy” objects for which data must be obtained from different data resources. Alternatively, “constructing” data by conducting interviews or surveys also have many constraints. Surveys are subject to limitations due to low response rates and biases of the respondents, and it is difficult for a firm to have extensive and periodic interviews, especially for large firms. These restrict objective comparisons over time and across the business and firms.

Now since the beginning of the Internet era, huge amounts of documents, comments and discussions are available as text from the World Wide Web, and not only computer science experts have taken advantage of this textual data, but there has already been a large fraction of application of text mining in finance, economics, biology. Recent studies by cultural sociologists have proposed focusing on this massive informative social science data, to address and resolve some of the un-measurable concepts of culture (Lazer et al., 2009; King, 2011; Bail, 2014; Evans and Aceves, 2016). Bail (2014) argues that, “integration of in depth qualitative coding techniques pioneered by cultural sociologists and anthropologists can be leveraged to improve already powerful automated text analysis techniques produced by computer scientists, linguists, and political scientists.” Hence we follow previous studies in finance that use data mining tools to corporate culture and risk culture (Bianchi, Farina and Fiordelisi, 2016; Fang et al., 2017). We take advantage of the availability of annual 10-K reports for banks, which gives a comprehensive summary of the financial performance, as well as discussions on risk factors, and apply textual analysis tools to extract and analyze the unstructured information.

Before extracting features from the text, the textual data should be preprocessed into a uniform and clean version. Starting with the raw file submitted by each bank, this process is broken down into several steps.

- Cleaning Irrelevance: The 10-K report is a multi-objective SEC filing. Tables and figures are contained in the forms as a description for the financial condition of the bank. They are objective quantitative data without attitude or sentiment of the bank. Since this paper only focuses on how each aspect of risk culture is described in the 10-K reports, quantitative data are removed.
- Tokenization: Each file submitted by the banks is represented as a corpus in text mining. Each corpus is broken (tokenized) into words to apply the bag-of-words method. They are also tokenized into paragraphs simultaneously to analyze the topic at the paragraph level.
- Stop Words Removing: Not all words that are tokenized from the corpus are meaningful. Most common words in the English language like 'a, the, that, about, etc. are removed before further processing of the textual data.

After parsing and cleaning the text document, we create a two-dimensional dictionary: risk culture dictionary from the framework RCF (McConnell, 2013) and a sentiment dictionary (Loughran and McDonald, 2014). The risk culture dictionary has seven categories (or key drivers): Leadership, Strategy, Decision making, Controls, Reward, and Recruitment, training and competence. Using the RCF framework, we identify key risk words for each category of risk culture. After, we use the 'Synset' function found in 'Wordnet', an NLTK Corpus reader, to identify synonyms of all the words in our risk culture dictionary, including their hyponyms and hypernyms. We include in our dictionary, only synonyms that have the closest meaning to the key words. In total, our risk culture dictionary has 600 risk culture words. A distribution of words in each category is provided in Panel A of Table 2, and Table A.1 in the Appendix shows a list of the 600 risk culture words.

The sentiment dictionary we use is the Loughran and McDonald's Sentiment Dictionary. Only considering whether a specific category of risk culture or sentiment appears yields an inaccurate result. Prollochs et al. (2015) considered the negation issue in a sentiment of negative and positive words. We apply their idea of detecting negation to the five sentiment categories. If a sentiment word is accompanied by '*no*' or negation terms such as *rather, hardly, does not, have not, has not, will not, had not, never* etc., the sentiment intent is reversed, from positive to negative, and vice-versa. This reversal is relevant for positive and negative category of sentiments, and not as much for the other three sentiment categories. An uncertainty word generally has no polarity, thus, the uncertainty sentiment words appearing with negation are ignored without any re-categorization. Litigious and constraining sentiments are not affected by negations, as their appearance itself is indicative of their degree of importance.

We address this issue by using a sentiment analysis library in Python called TextBlob. TextBlob is used for processing textual data and has a sentiment analysis property called sentiment. We use this property returns the polarity of the word or sentence on a scale of -1 to +1, with -1 being very negative and +1 being very positive. For each positive and negative word, we identify the polarity of the sentence in which it is found to determine the sentiment as either negative or positive. The distribution of words in each sentiment category is shown in Panel B of Table 2.

After labeling each specific term in the sentiment and risk culture dictionary by its category, all the terms are treated as equally important in the machine-learning algorithm. Considering word frequency as the metric gives too much weight to the paragraph that mentions a term many times, we introduce a paragraph topic and sentiment model to extract the metrics at a paragraph level.

Since our risk culture dictionary and sentiment dictionary are built by expert guidance, we believe once a term is mentioned in a paragraph, the paragraph is labeled as relevant to that term's category of risk culture indicator or sentiment. Thus, we categorize each paragraph in a corpus into a unique two-dimensional category of sentiment and risk culture. We then count the paragraph frequencies by each two-dimension category. They are classified in each corpus. Since there are seven categories in the risk culture dictionary and five categories in the sentiment dictionary, a corpus has in total of 35 features, each showing the frequency of paragraphs that belong to each joint category.

### 3.3 Financial, Governance and Compensation and Legal Expense Data

We obtain firm performance and characteristics data from Compustat. We also obtain from the ExecuComp database, the executive, director and board characteristics and compensation data, including the number of employees, and volatility. For the purpose of our supervised learning technique, we obtain legal expense data from the 'FR-Y9C' reports. The legal expense data has some missing observation as well as firms reporting non-zero legal expense. An annual summary of the legal expense data is provided in Figure 1. This figure shows that more banks report legal expense in earlier years than in later years. However, a larger number of banks report positive legal expense after 2008. The reason might be that during the initial years (2002-2006), the banks were not informed clearly of norms and standards. After the 2008 financial crisis, behavioral norms are emphasized with banks, which lead to strict monitoring and non-zero legal expense. In addition, the mean and standard deviation of the legal expense also increase after 2008.

Panel A of Table 3 shows the summary statistics of all banks in our sample. Definitions of all variables are presented in Table A2 in the Appendix. All variables are winsorized at the 1% and 99% levels.

# 4 Feature Reduction and Selection

## 4.1 Supervised Learning

In a study by McNulty and Akhigbe (2015) they show that high legal expenses incurred by a bank predict weak risk culture to some extent. Hence, we utilize legal expense as a guidance measure in a supervised learning approach to inform us what features significantly uncover aspects of risk culture of a bank. We utilize a linear regression based supervised learning approach, as described by Equation 1 below:

$$LE = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \ldots\ldots\ldots\ldots\ldots\ldots + a_{35}x_{35} + \varepsilon \qquad (1)$$

Where $\boldsymbol{LE}$ denotes the legal expense of the bank, $\boldsymbol{x_i}$ denotes the paragraph frequency for a specific joint category in the bank's 10-K report. Since we have a large set of 35 independent variables, we develop a linear regression analysis for supervised learning to reduce the number of independent variables by detecting insignificance in the regression model.

We implement four different regression models for our supervised learning: Ordinary Least Square (OLS), Stepwise Regression, Ridge Regression and Least Absolute Shrinkage and Selection Operator (LASSO). Results for the supervised learning are presented in Table 4. Column 1 is an OLS regression and Column 2 is a backward-selection stepwise regression model that removes independent variables that are not significant at the 10% significance level. Similar to the OLS Model, the stepwise regression selects 14 features, which are significant at the 10% level. One shortfall of the OLS and stepwise regression is that, they are not suitable to handle large features or variables, and tend to over fit the data. Hence, we consider the ridge and LASSO regularization techniques.

The ridge and LASSO regression models penalize the magnitude of coefficients of the features and minimizes the errors between actual and predicted observations. Whiles the ridge regression adds a penalty equal to the square of the magnitude of the coefficients, the LASSO models adds a penalty equal to the absolute value of the magnitude of the coefficients. From Columns 3 and 4, we observe that the LASSO regression shrinks to 0 similar predictor variables that are found to be insignificant using the ridge regression. We compare both models and select 16 variables that are significant in both models.

The selected variables include 'Litigious Leadership', 'Constraining Leadership', 'Negative Strategy', 'Uncertain Strategy,' 'Positive Decision', 'Uncertain Decision', 'Negative Control', 'Litigious Control', 'Constraining Control', 'Positive Recruitment', 'Negative Recruitment', 'Constraining Recruitment', 'Uncertainty Recruitment', 'Negative Reward', 'Uncertainty Reward', and 'Positive Portfolio'. The selected features cover all risk culture and sentiment dimensions.

### 4.2 Unsupervised Learning

Using the 16 significant features obtained through the supervised learning with guidance using the legal expense variable, we apply an unsupervised k-means clustering to illustrate how banks cluster according to these features. However, working with 16 features can still be complex and create some errors, hence before we cluster based on these features, we perform a Principal Component Analysis (PCA) to allow us to extract a smaller set of features that still captures most of the useful information in the variables.

Figure 2 shows the proportion of variation explained by each of the 16 principal components extracted. Now to determine the number of components to include in our analysis, we use the Kaiser's Rule, which suggests that one includes the number of components that have

Eigen values greater than 1. Figure 3 shows a scree plot of the Eigen values after PCA. Using Kaiser's Rule and the plot means we have to select the first two components only. Together they both explain approximately 62% of variation in data. However, our aim is to explain at least 80% of the variation in the data. Hence, we select the first 6 principal components. Table 4 shows the varimax rotated component matrix, which helps us to interpret the correlation between the 6 principal components and the original data features. Interpretation of each component is summarized in Panel B of Table 5. We find that at least one of each feature is explained by a principal component.

We move on to determine clusters based on the principal components selected using an unsupervised k-means clustering analysis. The idea in k-means clustering is to partition ***n*** observations into ***k*** clusters according to each cluster's distance from each of the k means. An observation is categorized into a cluster whose mean is nearest to the observation. To determine the best value for k in the k-means clustering, we use the Elbow Method. The elbow method looks at the percentage of variance explained s a function of the number of clusters. The plot in Figure 3 shows the optimal value for k to be 3.

To define the features of the clusters based on the principal components, we estimate the mean of the principal component for each cluster. Panel B of Table 5 summarizes these results. Cluster 1 has the highest mean in Components 2 and 6, and the second highest in the remaining components, hence we describe the risk culture features of Cluster 1 to show an uncertain Strategy and Decision, and constraining and uncertain Recruitment. Now since Cluster 1 has the second highest mean for the remaining components (1, 3, 4, 5), it is clearly observed to be in the middle of Cluster 2 and Cluster 3. Hence we classify Cluster 1 to be our moderate risk culture class.

Next, Cluster 2 shows the highest mean value in Components 1, 4 and 5, and these components describe a risk culture that has a constraining Leadership and Control, Negative Strategy and Control, Negative and uncertain Reward, Litigious Leadership and Control, Negative Recruitment, and also Positive Recruitment and Portfolio. Aside from a positive portfolio, Cluster 2 portrays a low risk culture. In addition, Cluster 2 has the least mean in Component 3, which describes a risk culture which is strong in Positive Decision making and Recruitment. Hence, Cluster 2 is our low risk culture class.

Lastly, Cluster 3 has the highest mean in Component 3, which describes a positive Decision Making and Recruitment, and the lowest in Component 4, which describes litigious Leadership and Control. Hence, we interpret Cluster 3 as having a high risk culture. In the next section, we examine how each cluster or risk culture class impact the bank's performance, corporate governance and compensation structure.

## 5. Empirical Analysis and Results

### 5.1 Endogeneity Concerns

The definition and features used to measure risk culture raises an endogeneity concern. First, risk culture of a firm is formed or structured by the behavior and risk appetite of individuals and groups in the firm. Second, the risk culture of the firm is not a stand-alone culture and affects many aspects of the firm's decisions and outcomes. Hence we identify reverse causality to be the main source of endogeneity in our empirical analysis. For instance, the risk culture is likely to impact investment decisions made by the firm such as acquisitions, and simultaneously, acquisition decisions could affect the risk appetite or taking behavior of the firm. One effective way to address the reverse causality issue is to use an instrument. However, it is difficult to find an instrument that will satisfy both the relevance and endogenous conditions of validity. Hence, we use panel regression analysis and include firm and year fixed effects to address any omitted and unobservable characteristics within the firm.

## 5.1 Risk Culture and Bank Performance

Our first analysis looks at how risk culture affects the performance output of a bank. We measure performance using ROA and the natural logarithm of Tobin's Q. Table 7 provides a summary of the results. The variable of interest is the class of risk culture mainly high and moderate risk classes. The excluded class is the low risk culture class. Hence all results are interpreted relative to low risk culture banks. We control for firm characteristics such as bank size, leverage, board size, number of independent and firm and year fixed effects. Columns 1 and 3 include only year fixed effects whiles columns 2 and 4 include both year and firm fixed effects.

Using ROA as the dependent variable, we find that banks with high and moderate risk cultures experience a positive and statistical significant increase in ROA relative to banks with a low risk culture. High and moderate risk culture banks have approximately 9.62% and 8.7% increases in ROA respectively. The results is the same when the logarithm of Tobin's Q is used. High and moderate risk culture banks have approximately 0.54% and 0.76% increases in Tobin's Q respectively. Hence we can conclude that having a sound risk culture can improve the performance of a bank.

Our results add to existing research findings on corporate culture and performance. A study by Fang et al. (2017) on U.S. public firms find that around the 2008-2009 financial crises, corporate culture created value for firms. Also, Guiso et al. (2013) uses a unique survey data that assesses the workplace of firms, and find that corporate culture has a positive relation with firm performance.

## 5.2 Risk Culture and Governance Characteristics

Next we examine the effect of risk culture on governance and board characteristics of banks. First, we consider the diversity of the board and focus on gender and race. We measure gender of the board using a dummy variable, which is 1 if there is a female director on the board, and zero otherwise. Race is measured using the nationality mix of the directors on the board. This is also a dummy variable, which is 1 if the board has more than 50% of directors having different nationalities. We include these two board characteristics since existing literature on gender shows that female executives and directors are more risk averse compared to males. Also the race of the board is of interest since the cultural origin of the founders and leaders at the top can be used to describe the risk culture of a firm (Pan, Siegal and Wang, 2016). From Column 1 in Table 8 we find no statistical significance of risk culture on gender of the board. In Column 2 we find that banks with moderate risk culture are likely to have a high nationality mix board compared to banks with a low risk culture. The coefficient for high risk culture banks is positive but insignificant. Hence we show that there is some relationship between a sound risk culture and diversity.

Second we look at the effect of risk culture on the percentage of institutional ownership. Chung and Zhang (2011) in their study show that, the fraction of institutional ownership increases the quality of corporate governance; hence we expect that banks with low risk culture would benefit from or need more monitoring and governing. From Column 3 we find that banks with high and moderate risk culture decrease the percentage of institutional ownership compared to banks with a low risk culture. This negative and significant effect can be interpreted as follows. A sound risk culture can also imply a good risk governance, hence banks with high and moderate risk cultures would have less concern for issues related to governance. On the other hand, banks with a low risk culture are more likely to have issues with corporate and risk governance and as such are more likely to have a higher percentage of institutional owners to monitor them.

We also examine the effect of risk culture on the number of board meetings held in a year. We use board meetings to proxy for how active a corporate board is or how well the board operates. From the results in Column 4, we find that boards of banks with a sound risk culture meet more often than banks with a low risk culture. Hence we can imply that these boards are active and operate effectively.

Lastly, we examine whether the risk culture of a bank can affect the pay structure of directors of the boards. Specifically we are interested to know if risk culture can create a risk incentive for directors through their pay structure. We consider the number of stocks and number of options included in the non-employee director's pay. The effect of risk culture on the director's stocks ownership is negative and insignificant. However, banks with moderate risk culture decrease the number of options held by directors. We can imply that some amount of sound risk culture reduces the risk incentives created for the directors of the bank in their compensation package.

### 5.3 Risk Culture and Compensation

In our last empirical analysis, we test the effect of risk culture on the compensation structure of the executives of the banks to determine risk-taking incentives. Our sample of executives includes both CEOs and CFOs. Our first regressions examine the total pay of executives. Table 9 presents the results of our analysis. We find that banks with a high and moderate risk culture decrease the number of total compensation given to their executives relative to banks with a low risk culture. We interpret this finding in two ways. First banks with sound risk culture reduce the risk taking incentives for executives. Existing literature documents a positive relationship between executive compensation and (excessive) risk taking (Rajgopal and Shevlin, 2011; Bolton et al, 2015). Second, these banks do not overpay their executives compared to low risk banks, however they outperform the latter.

We also find that banks with a high risk culture decrease the executive's fraction of equity based pay compared to banks with low risk culture. Once again, this implies that sound risk culture reduces the risk taking incentives of executives. We also find that the effect of sound risk culture on the fraction of executives' incentive pay is not significant. However, in Column 4 we find that banks with a high risk culture increase the fraction of executives incentive pay, but this is a small increase of 0297% and at a significance of 10%. Hence we conclude that banks with a sound risk culture create some incentive for their executives, but at the same time, they ensure that they do not excessively overpay them.

## 6. Conclusion

In this paper, we use textual analysis and unsupervised cluster analysis to explore bank's culture. Focusing on the 5,670 10-K reports of 572 bank holding companies between 2000 and 2016, downloaded from the SEC, we study the sentiment of the banks on seven risk culture aspects and categorize the 10-K reports into three different risk culture classes: high, moderate and low risk culture.

After the process of supervised learning for feature selection and feature reduction under the guidance of legal expense, we finally determine sixteen features as input to the unsupervised learning for the larger data set, covering all seven risk culture aspects and five sentiment categories. Further feature extraction is implemented using a principal component analysis and 6 components are selected to explain the variation in features, and 3-means clustering is applied to the six principal components.

Our main results show that sound risk culture matter and has an effect on a bank's performance, governance and compensation structure of executives. Using the three risk culture classes, we find that banks with high and moderate risk culture outperform banks with a low risk culture. They also have less monitoring incentives available since they are more likely to have a good risk governance structure in place. A sound risk culture also reduces the risk-taking incentive of executives.

For future research, we suggest a more robust approach can be used to test and define causality for our study.

## References


Ashby, S. Guidance on Supervisory Interaction with Financial Institutions on Risk Culture Feedback on the FSBs Consultative Document. PhD thesis, London School of Economics

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

Bianchi, N., Carretta, A., Farina, V., Fiordelisi, F. Risk culture in banking: just words?. Working paper 2016

Bodnaruk, A., Loughran,T. and Bill McDonald. 2015. Using 10-k text to gauge nancial constraints. Journal of Financial and Quantitative Analysis, 50(04):623-646

Bolton, B., Mehran,H. and Shapiro, J., 2015. Executive Compensation and Risk Taking, *Review of Finance*,16(6), 2139–2181

Coluccia, D., Fontana, S., Graziano, E. A., Rossi, M., & Solimene, S. (2017). Does risk culture affect banks' volatility? The case of the G-SIBs. Corporate Ownership & Control, 15(1), 33-43.

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

Fang, Y., Fiordelisi, F., Hasan, I.,, and Leung W. S., 2017 Does Corporate Culture Matter for Firm Value? Evidence from the 2008-09 Financial Crisis. Working paper

FSA, 2007, 'Treating Customers Fairly - Culture', Financial Services Authority, London. July 2007

FSB, 2011, 'Policy measures to address Systemically Important Financial Institutions', Financial Stability Board, Basel, 4 November 2011

FSB. Guidance on supervisory interaction with nancial institutions on risk culture: A framework for assessing risk culture. Technical report, Financial Stability Board, 2014.

Geretto, E. F and Pauluzzo, R. 2015. Knowledge management and risk culture in the banking industry: Relations and problems. In European Conference on Knowledge Management, page 313. Academic Conferences International Limited

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

Griffin, B and Sheedy, E. 2014Empirical analysis of risk culture in financial institutions: Interim report. Technical report, Macquarie University, Centre for International Finance and Regulation, 2014.

Guiso, L., P. Sapienza, and L. Zingales, 2013, The Value of Corporate Culture, NBER Working Paper 19557.

Gupta, A., Simaan,M. and Zaki, M. When positive sentiment is not so positive: Textual analytics and bank failures. Available at SSRN 2773939, 2016.

IIF, 2009, 'Reform in the financial services industry: Strengthening Practices for a More Stable System', Institute of International Finance, Washington, December 2009

Kaplan, S. and Zingales, L. 1997. Do investment-cash flow sensitivities provide useful measures of  nancing constraints? The Quarterly Journal of Economics, pages 169-215

Kee H. Chung and Hao Zhang.Corporate Governance and Institutional Ownership. The Journal of Financial and Quantitative Analysis. Vol. 46, No. 1

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

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.-L., Brewer, D., … Van Alstyne, M. 2009. Computational social science. Science (New York, N.Y.), 323(5915), 721-3

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. 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. The Journal of Finance, 66(1), 35-65

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

McConnell, P., 2013. A risk culture framework for systemically important banks. Journal of Risk and Governance 3, 23-68

McNulty, J. and Akhigbe, A., 2014. Bank litigation, bank performance and operational risk: Evidence from the financial crisis. Bank Performance and Operational Risk

Pan, Y., Siegel, S., Wang, T.Y., 2015. Corporate Risk Culture. Working paper

Prollochs N., Feuerriegel, S. and Neumann, D. 2015. Enhancing sentiment analysis of  nancial news by detecting negation scopes. In System Sciences (HICSS), 2015 48th Hawaii International Conference on, pages 959{968. IEEE, 2015.

Rajgopal, S. and T. Shevlin, 2002, Empirical evidence on the relation between stock option compensation and risk taking, Journal of Accounting and Economics, 33, 145-171.

Sheedy E, Griffin B. Risk governance, structures, culture, and behavior: A view from the inside. Corp Govern Int Rev. 2018;26:4–22

**Figure 1: Summary of Legal Expense Report by Frequency**

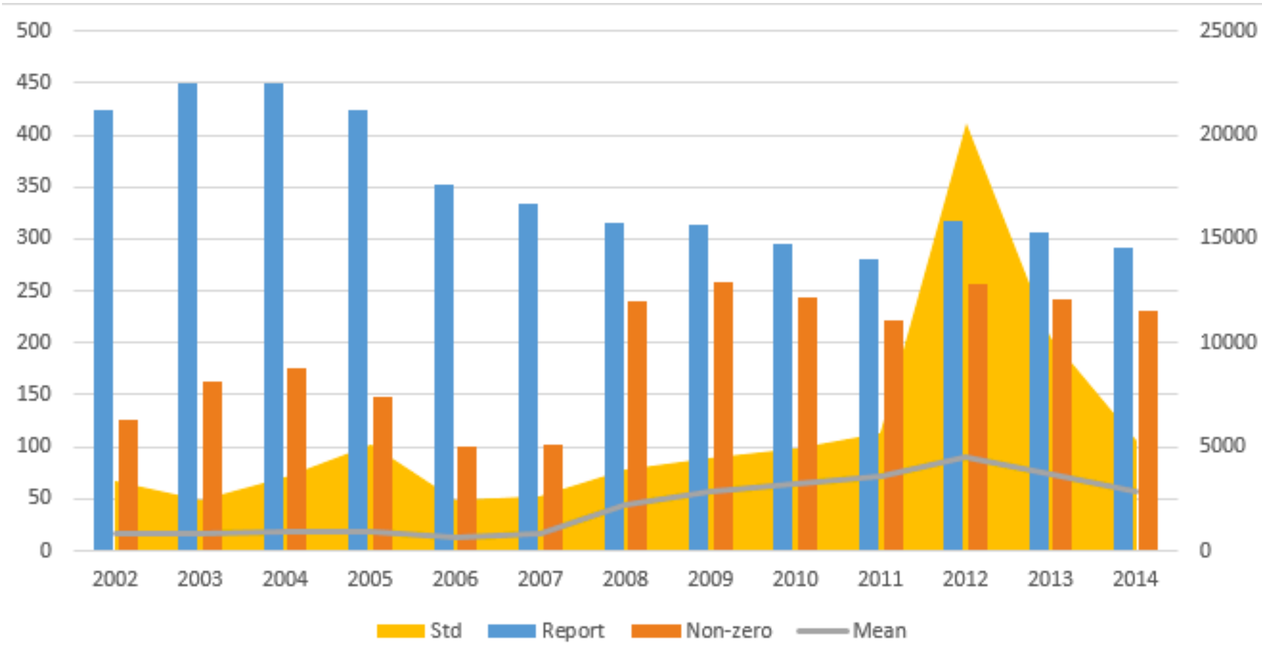



**Figure 2: Proportion of Variation Explained in Principal Component Analysis**

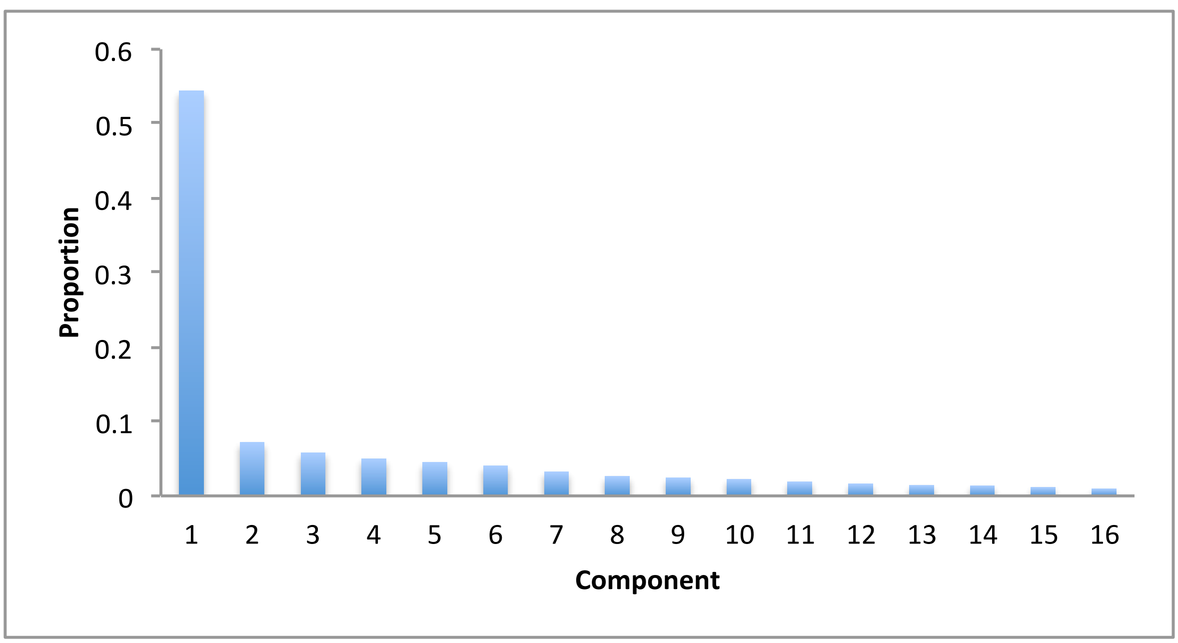


**Figure 3: Scree Plot of Eigen Values**

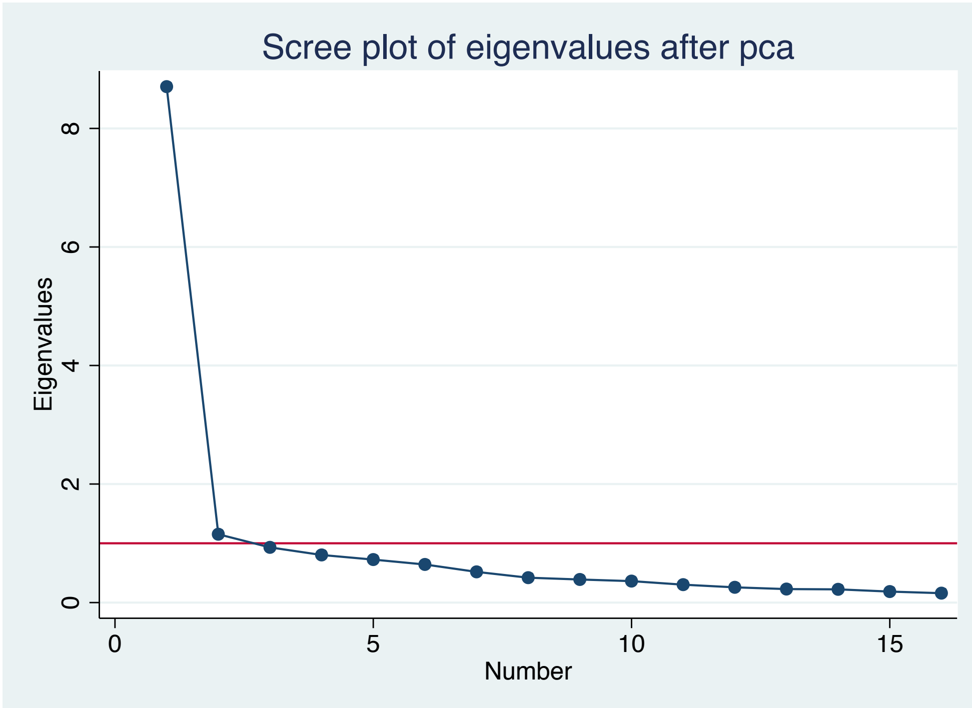
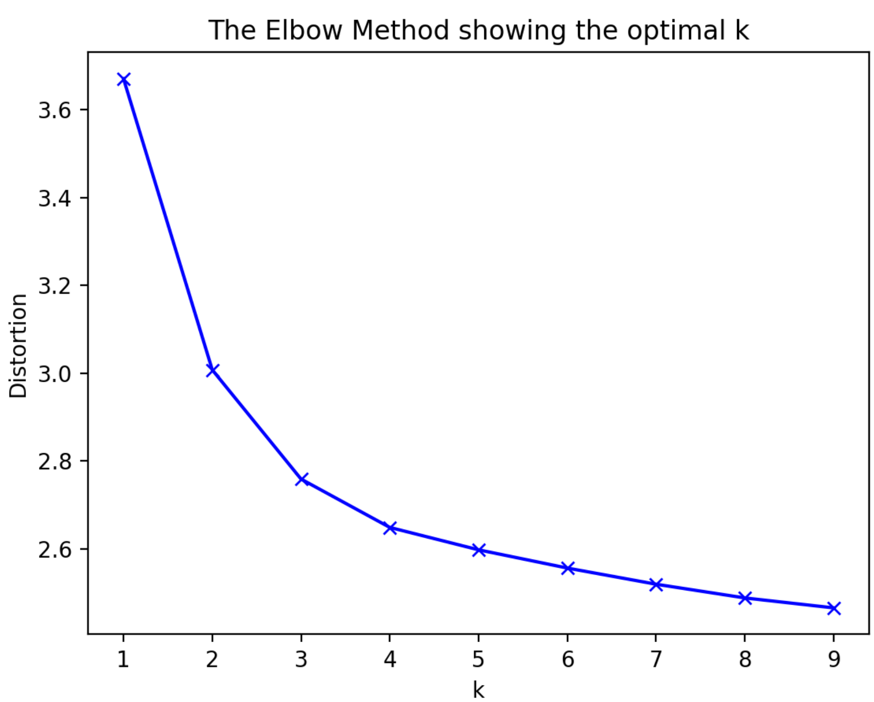



**Figure 4: The Elbow Method to detect the number of k-mean clusters**



**‡Table 1: Risk Culture Framework, *McConnell (2013)***

Key Drivers of Risk Culture and Indicators

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

‡ Table 3. Risk Culture Framework based on TCF model (FSA 2007); McConnell (2013)

## Table 2: Distribution of Words in Sentiment and Risk Culture Dictionary

Panel A shows the distribution of bag of words included in the risk culture dictionary used for text extraction. Panel B shows the distribution of words in the Loughran and McDonald (2013) sentiment dictionary.

Panel A: Risk Culture Dictionary

| **Risk Culture Key Driver** | Leadership | Strategy | Decision Making | Controls | Recruitment, Training and Competence | Reward | Portfolio |
|---|---|---|---|---|---|---|---|
| **Word Count** | 194 | 100 | 39 | 111 | 61 | 64 | 31 |

Panel B: Sentiment Dictionary

| **Sentiment** | Negative | Positive | Uncertainty | Litigious | Constraining |
|---|---|---|---|---|---|
| **Word Count** | 2355 | 354 | 297 | 903 | 184 |

## Table 3: Summary Statistics

Panel A shows the summary statistics of firm and executive compensation characteristic variables for the total number of observations in our sample. Panel B is the summary statistics of observations within the 3 Clusters: High, Moderate and Low risk culture. All variables are winsorized at the 1% and 99% levels.

**Panel A: Full Observation**

| Variables | Observations | Mean | Standard Deviation | Minimum | Median | Maximum |
|---|---|---|---|---|---|---|
| **Firm Characteristics** | | | | | | |
| Size | 5670 | 7.349 | 1.047 | 5.539 | 7.160 | 9.937 |
| Return on Asset | 5670 | 0.325 | 0.607 | -1.882 | 0.011 | 2.010 |
| Legal Expense | 5670 | 2001.034 | 3983.375 | 0 | 337 | 22822 |
| Leverage | 5670 | 0.127 | 0.086 | 0 | 0.111 | 0.423 |
| Return on Equity | 5670 | 7.528 | 15.780 | -99.502 | 9.576 | 26.554 |
| Number of Employees | 5670 | 1.568 | 1.266 | 0.114 | 1.123 | 7.023 |
| Tobin's Q | 5670 | 1.040 | 0.062 | 0.922 | 1.032 | 1.238 |
| Earning Per Share | 5670 | 1.268 | 1.506 | -6.360 | 1.370 | 4.230 |
| Volatility | 5670 | 0.261 | 0.077 | 0.124 | 0.252 | 0.524 |
| Sales | 5670 | 415.282 | 314.076 | 44.824 | 314.158 | 1484.036 |
| Board Size | 5670 | 10.893 | 3.154 | 5 | 10 | 21 |
| Female Director | 5670 | 0.787 | 0.410 | 0 | 1 | 1 |
| Independent Directors | 5670 | 8.468 | 2.716 | 3 | 8 | 16 |
| Institutional Ownership | 5670 | 0.277 | 0.221 | 0 | 0 | 1 |
| Nationality Mix | 5670 | 0.183 | 0.386 | 0 | 0 | 1 |
| Number of Meetings | 5670 | 9.239 | 4.027 | 3 | 9 | 23 |
| **Compensation Characteristics** | | | | | | |
| Executive Salary Pay | 5670 | 307.113 | 101.070 | 120.068 | 296.044 | 615 |
| Executive Bonus Pay | 5670 | 59.996 | 112.332 | 0 | 0 | 630 |
| Executive Other Pay | 5670 | 48.066 | 62.076 | 0 | 29.676 | 407.585 |
| Executive Shareholding | 5670 | 141.372 | 170.871 | 5.474 | 84.376 | 1010.273 |
| Executive Total Compensation | 5670 | 415.176 | 170.164 | 123.803 | 376.419 | 1472.765 |
| Fraction of Executive Incentive Pay | 5670 | 0.108 | 0.092 | 0 | 0.086 | 0.695 |
| Equity Fraction of Executive Pay | 5670 | 0.349 | 0.454 | 0.012 | 0.209 | 7.251 |
| Executive Age (Median) | 5670 | 54.616 | 4.767 | 43 | 55 | 66 |
| Director's Meeting Fee | 5670 | 0.969 | 0.930 | 0 | 1 | 6 |
| Number of Director's Stocks | 5670 | 0.168 | 0.379 | 0 | 0 | 1.912 |
| Number of Director's Options | 5670 | 7.495 | 13.931 | 0.672 | 3 | 80 |

**Panel B: Risk Culture Clusters**

| Variables | High N = 2726 Mean | High N = 2726 Standard Deviation | Moderate N = 1811 Mean | Moderate N = 1811 Standard Deviation | Low N = 1133 Mean | Low N = 1133 Standard Deviation |
|---|---|---|---|---|---|---|
| **Firm Characteristics** | | | | | | |
| Size | 7.360 | 0.958 | 6.883 | 0.897 | 8.060 | 1.066 |
| Return on Asset | 0.331 | 0.574 | 0.308 | 0.596 | 0.333 | 0.684 |
| Legal Expense | 1661.618 | 3475.840 | 857.085 | 2115.419 | 4151.212 | 5639.078 |
| Leverage | 0.126 | 0.082 | 0.130 | 0.093 | 0.122 | 0.081 |
| Return on Equity | 8.644 | 12.704 | 11.611 | 14.636 | 3.391 | 19.052 |
| Number of Employees | 1.534 | 1.218 | 1.401 | 1.295 | 1.720 | 1.297 |
| Tobin's Q | 1.039 | 0.061 | 1.053 | 0.065 | 1.024 | 0.056 |
| Earning Per Share | 1.394 | 1.292 | 1.622 | 1.263 | 0.871 | 1.807 |
| Volatility | 0.253 | 0.078 | 0.276 | 0.075 | 0.261 | 0.069 |
| Sales | 389.253 | 291.574 | 351.376 | 298.828 | 491.592 | 338.100 |
| Board Size | 11.107 | 3.163 | 10.275 | 3.196 | 11.159 | 2.969 |
| Female Director | 0.763 | 0.425 | 0.793 | 0.405 | 0.833 | 0.373 |
| Independent Directors | 8.695 | 2.736 | 7.775 | 2.641 | 8.801 | 2.602 |
| Institutional Ownership | 0.284 | 0.209 | 0.184 | 0.178 | 0.430 | 0.238 |
| Nationality Mix | 0.158 | 0.365 | 0.277 | 0.448 | 0.121 | 0.327 |
| Number of Meetings | 9.403 | 4.233 | 9.135 | 3.894 | 8.500 | 3.033 |
| **Compensation Characteristics** | | | | | | |
| Executive Salary Pay | 280.032 | 83.737 | 253.602 | 76.412 | 358.929 | 104.606 |
| Executive Bonus Pay | 48.815 | 79.650 | 99.258 | 152.561 | 56.875 | 121.327 |
| Executive Other Pay | 44.314 | 59.722 | 44.996 | 63.951 | 53.510 | 63.644 |
| Executive Shareholding | 124.632 | 156.455 | 153.962 | 155.133 | 152.780 | 184.228 |
| Executive Total Compensation | 373.160 | 136.007 | 397.856 | 192.272 | 469.314 | 180.586 |
| Fraction of Executive Incentive Pay | 0.110 | 0.093 | 0.104 | 0.094 | 0.107 | 0.089 |
| Equity Fraction of Executive Pay | 0.351 | 0.408 | 0.402 | 0.408 | 0.335 | 0.498 |
| Executive Age (Median) | 54.714 | 4.664 | 52.234 | 5.548 | 55.320 | 4.326 |
| Director's Meeting Fee | 0.915 | 0.693 | 1.086 | 1.201 | 0.783 | 0.850 |
| Number of Director's Stocks | 0.180 | 0.398 | 0.106 | 0.297 | 0.374 | 0.516 |
| Number of Director's Options | 7.031 | 16.551 | 8.468 | 10.481 | 6.826 | 8.713 |

## Table 4: Supervised Learning

This table reports the results of the supervised regressions with legal expense as a target variable. Column 1 is an ordinary least square regression, column 2 is a forward stepwise regression, column 3 is a ridge regression and column 4 is a Least Absolute Shrinkage and Selection Operator (LASSO). Numbers in parentheses are robust standard errors. Significance levels are as indicated; *** p<0.01, ** p<0.05, * p<0.1

| | Dependent Variable: Legal Expense | | | |
|---|---|---|---|---|
| | **OLS** | **Stepwise** | **Ridge** | **LASSO** |
| **Independent Variables** | **(1)** | **(2)** | **(3)** | **(4)** |
| Positive Leadership | 14.65* | 41.58*** | 405.95 | 268.76 |
| | (8.650) | (13.10) | (295.82) | (296.60) |
| Negative Leadership | 23.45 | | 121.92 | 36.30 |
| | (32.35) | | (184.68) | (185.16) |
| Uncertainty Leadership | 12.70 | | 172.27 | 0.00 |
| | (17.21) | | (310.49) | (311.30) |
| Litigious Leadership | 91.56* | 80.27* | 854.97*** | 753.58*** |
| | (49.64) | (44.24) | (197.81) | (198.33) |
| Constraining Leadership | -122.5*** | -118.3*** | -865.75*** | -610.05*** |
| | (31.77) | (36.07) | (206.15) | (206.69) |
| Positive Strategy | -4.240 | | -77.55 | 0.00 |
| | (11.66) | | (280.73) | (281.47) |
| Negative Strategy | 56.72** | 65.14*** | 394.54* | 361.72* |
| | (22.18) | (23.06) | (206.02) | (206.57) |
| Uncertainty Strategy | 118.4*** | 129.6*** | 1545.32*** | 1452.37*** |
| | (44.68) | (30.07) | (292.51) | (293.28) |
| Litigious Strategy | -0.442 | | -3.88 | 0.00 |
| | (19.03) | | (215.46) | (216.02) |
| Constraining Strategy | 38.68 | | 197.44 | 0.00 |
| | (29.50) | | (191.04) | (191.54) |
| Positive Decision | -112.5** | -64.23* | -322.66** | -190.23 |
| | (48.54) | (35.99) | (155.71) | (156.12) |
| Negative Decision | -173.0** | -183.7*** | -211.64 | -98.54 |
| | (79.82) | (69.14) | (131.85) | (132.19) |
| Uncertainty Decision | 99.56 | | 338.16* | 107.00 |
| | (63.30) | | (204.55) | (205.09) |
| Litigious Decision | -55.23 | | -148.86 | -17.51 |
| | (38.61) | | (137.68) | (138.04) |
| Constraining Decision | 261.9 | | 197.74 | 121.82 |
| | (171.4) | | (121.09) | (121.41) |
| Positive Control | 16.03 | | 191.25 | 0.00 |
| | (27.69) | | (257.06) | (257.74) |
| Negative Control | 43.06 | 57.83** | 329.64* | 270.49 |
| | (38.26) | (27.35) | (194.10) | (194.61) |
| Uncertainty Control | -19.42 | | -168.08 | 0.00 |
| | (33.64) | | (245.01) | (245.65) |
| Litigious Control | 68.38*** | 60.62*** | 897.58*** | 729.31*** |
| | (17.42) | (13.83) | (233.37) | (233.99) |
| Constraining Control | -77.15 | | -682.39*** | -381.10* |
| | (49.64) | | (213.00) | (213.56) |
| Positive Recruitment | -144.0*** | -92.80*** | -485.95*** | -440.15*** |
| | (52.73) | (29.56) | (156.58) | (156.99) |
| Negative Recruitment | -313.7*** | -344.8*** | -563.89*** | -496.26*** |
| | (88.19) | (104.1) | (140.71) | (141.08) |

| | | | | |
|---|---|---|---|---|
| Uncertainty Recruitment | -102.0** | | -349.52* | -71.40 |
| | (50.80) | | (197.50) | (198.02) |
| Litigious Recruitment | -6.166 | | -14.67 | 0.00 |
| | (73.17) | | (153.94) | (154.35) |
| Constraining Recruitment | -379.7*** | -401.0*** | -395.37*** | -373.66*** |
| | (121.3) | (139.1) | (128.90) | (129.24) |
| Negative Reward | 17.62 | | -1242.81*** | -884.55*** |
| | (19.58) | | (249.79) | (250.44) |
| Positive Reward | -99.81** | -62.03*** | 188.32 | 0.00 |
| | (48.90) | (22.13) | (229.06) | (229.66) |
| Uncertainty Reward | -75.03** | -59.13* | -644.89*** | -425.15** |
| | (32.06) | (32.89) | (211.17) | (211.73) |
| Litigious Reward | 46.89 | | 277.63 | 101.64 |
| | (71.11) | | (184.48) | (184.96) |
| Constraining Reward | -50.21 | | -280.93 | -77.45 |
| | (42.25) | | (186.66) | (187.15) |
| Positive Portfolio | 67.95 | | 1835.19*** | 1458.22*** |
| | (41.73) | | (329.14) | (330.01) |
| Negative Portfolio | -2.838 | | -36.27 | 0.00 |
| | (21.27) | | (264.11) | (264.80) |
| Uncertainty Portfolio | 8.933 | | 112.75 | 0.00 |
| | (25.77) | | (263.20) | (263.89) |
| Litigious Portfolio | -31.89 | | -189.06 | 0.00 |
| | (27.88) | | (174.43) | (174.89) |
| Constraining Portfolio | -2.328 | | -16.95 | 0.00 |
| | (22.69) | | (207.16) | (207.70) |
| Constant | -1,648*** | -2,076*** | 2228.41*** | 2228.41*** |
| | (297.7) | (572.1) | (114.84) | (115.14) |
| Observations | 4,011 | 4,011 | 4011 | 4011 |
| R-squared | 0.127 | 0.111 | 0.127 | 0.122 |

## Table 5: Unsupervised Learning

This table presents the results from the principal component analysis for feature extraction and Clustering Analysis. Panel A reports the results of a rotated component matrix using the Varimax method with Kaiser's normalization. Panel B summarizes the interpretation of the rotated component matrix in Panel A. Panel C summarizes the mean component for each cluster.

**Panel A: Rotated Component Matrix**

| Features | Component 1 | Component 2 | Component 3 | Component 4 | Component 5 | Component 6 | Unexplained |
|---|---|---|---|---|---|---|---|
| Litigious Leadership | -0.0511 | 0.0146 | 0.069 | **0.8231** | 0.0352 | -0.0509 | 0.09974 |
| Constraining Leadership | **0.2846** | 0.0061 | 0.1067 | 0.0888 | -0.0948 | 0.0954 | 0.2961 |
| Negative Strategy | **0.3632** | 0.0191 | -0.1244 | 0.0444 | 0.0974 | -0.0519 | 0.2855 |
| Uncertain Strategy | 0.2444 | **0.3213** | -0.0096 | -0.005 | -0.031 | -0.0383 | 0.1641 |
| Positive Decision | 0.0125 | -0.0197 | **0.7564** | 0.0651 | -0.1558 | -0.0915 | 0.1901 |
| Uncertain Decision | -0.0101 | **0.6692** | -0.0103 | 0.0156 | -0.0627 | -0.0346 | 0.1271 |
| Negative Control | **0.4557** | -0.07 | -0.1099 | -0.1142 | 0.0915 | -0.0665 | 0.2562 |
| Litigious Control | 0.2258 | -0.002 | -0.0974 | **0.4701** | -0.0458 | 0.0389 | 0.175 |
| Constraining Control | **0.3589** | -0.1082 | -0.1096 | 0.1678 | -0.051 | 0.1799 | 0.2467 |
| Positive Recruitment | -0.096 | -0.0026 | **0.5052** | 0.0121 | **0.4053** | 0.1896 | 0.2451 |
| Negative Recruitment | 0.0242 | 0.0046 | -0.0715 | 0.0216 | **0.8684** | -0.0478 | 0.1088 |
| Constraining Recruitment | -0.0027 | 0.01 | -0.0212 | -0.0293 | -0.0253 | **0.9385** | 0.05315 |
| Uncertainty Recruitment | -0.0414 | **0.6386** | -0.0191 | 0.012 | 0.0886 | 0.0762 | 0.1403 |
| Negative Reward | **0.3682** | -0.0327 | 0.12 | -0.1119 | 0.0849 | -0.0674 | 0.2158 |
| Uncertainty Reward | **0.2855** | 0.1449 | 0.2134 | -0.1625 | -0.0616 | -0.0524 | 0.2478 |
| Positive Portfolio | **0.3293** | -0.035 | 0.2032 | -0.0742 | -0.009 | 0.0639 | 0.189 |

**Panel B: Interpretation of Principal Components**

| Component | Interpretation based on Features | Highest mean by cluster | Risk Culture |
|---|---|---|---|
| 1 | Constraining Leadership and Control; Negative Strategy and Control; Negative and Uncertain Reward; Positive Portfolio | Cluster 2 | Low |
| 2 | Uncertain Strategy, Decision and Recruitment | Cluster 1 | Moderate |
| 3 | Positive Decision and Recruitment | Cluster 3 | High |
| 4 | Litigious Leadership and Control | Cluster 2 | Low |
| 5 | Positive and Negative Recruitment | Cluster 2 | Low |
| 6 | Constraining Recruitment | Cluster 1 | Moderate |

**Panel C: Mean of Principal Components by Clusters**

| Cluster | Component 1 | Component 2 | Component 3 | Component 4 | Component 5 | Component 6 |
|---|---|---|---|---|---|---|
| **1** | -3.251 | **0.015** | -0.043 | 0.008 | 0.191 | **0.012** |
| **2** | **4.237** | -0.016 | -0.051 | **0.021** | **0.208** | -0.021 |
| **3** | 0.399 | -0.003 | **0.050** | -0.014 | -0.213 | 0.001 |

## Table 6: ANOVA Analysis

This table summarizes the results for the ANOVA tests on the means of the three risk culture clusters. Variables are as defined in the variables table. Significance levels are as indicated; *** p<0.01, ** p<0.05, * p<0.1

| Variables | High | Moderate | Low | F-Statistic | p-value | Significance |
|---|---|---|---|---|---|---|
| Size | 7.290 | 6.766 | 7.989 | 342.550 | 0.000 | *** |
| Return on Asset | 0.202 | 0.179 | 0.232 | 3.850 | 0.021 | *** |
| Total Expense | 1252.613 | 504.028 | 3264.545 | 245.630 | 0.000 | *** |
| Leverage | 0.125 | 0.128 | 0.120 | 2.720 | 0.066 | * |
| Return on Equity | 1.912 | 1.802 | 1.317 | 2.110 | 0.121 | |
| # of Employees | 0.335 | 0.214 | 0.666 | 93.680 | 0.000 | *** |
| Tobin's Q | 0.518 | 0.529 | 0.509 | 0.510 | 0.601 | |
| Earning Per Share | 0.308 | 0.252 | 0.338 | 3.650 | 0.026 | *** |
| Volatility | 0.026 | 0.024 | 0.009 | 19.550 | 0.000 | *** |
| Sales | 86.104 | 54.521 | 191.343 | 132.690 | 0.000 | *** |
| Board Size | 7.966 | 5.776 | 7.722 | 86.770 | 0.000 | *** |
| Female Director | 0.763 | 0.793 | 0.833 | 12.150 | 0.000 | *** |
| Independent Directors | 6.236 | 4.371 | 6.090 | 101.730 | 0.000 | *** |
| Institutional Ownership | 0.221 | 0.148 | 0.302 | 173.880 | 0.000 | *** |
| Nationality Mis | 0.113 | 0.156 | 0.084 | 18.470 | 0.000 | *** |
| # of Meetings | 0.445 | 0.449 | 0.120 | 12.000 | 0.000 | *** |
| Executive Salary Pay | 49.411 | 23.946 | 135.588 | 306.330 | 0.000 | *** |
| Executive Bonus Pay | 8.613 | 9.372 | 21.485 | 24.530 | 0.000 | *** |
| Executive Other Pay | 7.819 | 4.249 | 20.214 | 88.460 | 0.000 | *** |
| Executive Shareholding | 15.728 | 7.226 | 55.422 | 134.790 | 0.000 | *** |
| Executive Total Compensation | 65.844 | 37.567 | 177.287 | 245.850 | 0.000 | *** |
| Executive Age (Median) | 9.112 | 4.182 | 20.898 | 240.290 | 0.000 | *** |
| Director's Meeting Fee | 0.044 | 0.054 | 0.013 | 8.130 | 0.000 | *** |
| # of Director's Stocks | 0.009 | 0.005 | 0.006 | 0.890 | 0.411 | |
| # of Director's Options | 0.162 | 0.182 | 0.078 | 0.810 | 0.445 | |
| Fraction of Executive Incentive Pay | 0.019 | 0.010 | 0.041 | 0.210 | 0.814 | |
| Equity Fraction of Executive Pay | 0.044 | 0.019 | 0.122 | 6.030 | 0.003 | *** |

## Table 7: Risk Culture and Firm Performance

This table presents a panel regression analysis of the effect of risk culture on the performance of bank holding companies. The dependent variable in Columns 1 and 2 is the return on asset (ROA), and in Columns 3 and 4 is the natural logarithm of Tobin's Q. All other variables are as defined in variable table A2. Numbers in parentheses are robust standard errors. Significance levels are as indicated; *** p<0.01, ** p<0.05, * p<0.1

| | Dependent Variables | | | |
|---|---|---|---|---|
| | ROA | | Log (Tobin's Q) | |
| **Independent Variables** | **(1)** | **(2)** | **(3)** | **(4)** |
| Moderate Risk Culture | 0.168*** | 0.0870*** | 0.0332*** | 0.00758** |
| | (0.0190) | (0.0215) | (0.00428) | (0.00319) |
| High Risk Culture | 0.133*** | 0.0962*** | 0.0201*** | 0.00535** |
| | (0.0151) | (0.0154) | (0.00324) | (0.00230) |
| Size | 0.00861 | 0.0111 | -0.0147*** | -1.32e-05 |
| | (0.00661) | (0.00709) | (0.00272) | (0.00253) |
| Board Size | 0.0147*** | 0.0112*** | 0.00366*** | 0.00213*** |
| | (0.00428) | (0.00418) | (0.000865) | (0.000610) |
| Independent Directors | -0.0201*** | -0.00656 | -0.00563*** | -0.00318*** |
| | (0.00539) | (0.00536) | (0.00109) | (0.000782) |
| Leverage | -0.222** | -0.0935 | -0.00202 | -0.0593*** |
| | (0.0908) | (0.0976) | (0.0187) | (0.0145) |
| Constant | 0.0551 | 0.104** | 0.131*** | 0.0269 |
| | (0.0514) | (0.0519) | (0.0214) | (0.0175) |
| Year Dummy | No | Yes | No | Yes |
| BHC Fixed Effect | Yes | Yes | Yes | Yes |
| Observations | 5,670 | 5,670 | 2,832 | 2,832 |
| R-squared | 0.023 | 0.096 | 0.107 | 0.574 |
| Number of BHC | 572 | 572 | 536 | 536 |

## Table 8: Risk Culture and Board Governance

This table presents a panel regression analysis of the effect of risk culture on the board characteristics and the governance of the bank holding companies. Female director is an indicator variable for female representation on a bank's board. Nationality mix is an indicator variable equal to 1 if the nationality mix of the board is 50% or more. Institutional ownership is the percentage of institutional investors for the bank. Number of Meetings is the number of meetings held by the board of directors in a year. Director's Stock and Director's Options are the number of awarded stocks and options respectively. All other variables are as defined in variable stable. Numbers in parentheses are robust standard errors. Significance levels are as indicated; *** p<0.01, ** p<0.05, * p<0.1

| | Dependent Variables | | | | | |
|---|---|---|---|---|---|---|
| | **Female Director** | **Nationality Mix** | **Institutional Ownership** | **Number of Meetings** | **Director's Stock** | **Director's Options** |
| **Independent Variables** | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| Moderate Risk Culture | -0.0120 | 0.0422*** | -0.0201** | 0.539*** | -0.00557 | -0.419*** |
| | (0.0196) | (0.0145) | (0.00827) | (0.169) | (0.00526) | (0.143) |
| High Risk Culture | -0.0133 | 0.00421 | -0.0138** | 0.300** | -0.00128 | -0.158 |
| | (0.0141) | (0.0105) | (0.00596) | (0.153) | (0.00379) | (0.103) |
| Size | -0.0134** | 0.00890* | 0.0205*** | 0.487*** | -0.00194 | 0.0420 |
| | (0.00647) | (0.00480) | (0.00274) | (0.131) | (0.00174) | (0.0474) |
| Board Size | -0.0381*** | 0.0170*** | -0.00659*** | 0.0904*** | 0.00600*** | -0.0125 |
| | (0.00381) | (0.00283) | (0.00163) | (0.0167) | (0.00102) | (0.0279) |
| Independent Directors | 0.0314*** | -0.00679* | 0.0123*** | -0.0688*** | -0.00631*** | 0.0297 |
| | (0.00488) | (0.00362) | (0.00207) | (0.0249) | (0.00131) | (0.0358) |
| Leverage | -0.0923 | -0.142** | -0.0945** | 0.869 | 0.0513** | 0.152 |
| | (0.0888) | (0.0659) | (0.0376) | (0.725) | (0.0239) | (0.651) |
| ROA | -0.0321** | 0.0214** | 0.00675 | -0.0277 | 0.0212*** | 0.533*** |
| | (0.0128) | (0.00948) | (0.00541) | (0.0930) | (0.00343) | (0.0936) |
| Tobin's Q | 0.00658 | -0.000162 | -0.000502 | 0.109** | -0.00275 | 0.0821 |
| | (0.00765) | (0.00567) | (0.00324) | (0.0479) | (0.00206) | (0.0561) |
| Female Director | | | 0.0139** | -0.243** | | |
| | | | (0.00594) | (0.102) | | |
| Meeting Fee | | | | 0.242*** | | |
| | | | | (0.0487) | | |
| Constant | 1.106*** | -0.0794** | -0.0247 | | 0.00766 | -0.0822 |
| | (0.0473) | (0.0351) | (0.0210) | | (0.0127) | (0.346) |
| Year Dummies | Yes | Yes | Yes | Yes | Yes | Yes |
| BHC Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 5,670 | 5,670 | 5,670 | 673 | 5,670 | 5,670 |
| R-squared | 0.157 | 0.127 | 0.465 | | 0.037 | 0.019 |
| Number of BHC | 572 | 572 | 572 | 60 | 572 | 572 |

## Table 9: Risk Culture and Executive Compensation

This table presents a panel regression analysis of the effect of risk culture on the executive compensation structure of bank holding companies. Executives include both CEO and CFOs. Total compensation is the sum of salary, bonus and other pay available to the executives. Fraction of Incentive Pay is 1 – (Salary + Bonus) / Total Compensation. Fraction of Equity-Based Pay is the ratio of the value of equity pay to total compensation each director receives in a year. All other variables are as defined in the variables table. Numbers in parentheses are robust standard errors. Significance levels are as indicated; *** p<0.01, ** p<0.05, * p<0.1

| Independent Variable | Dependent Variable | | | | | |
|---|---|---|---|---|---|---|
| | Total Compensation | | Fraction of Incentive Pay | | Fraction of Equity-based Pay | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Moderate Risk Culture | -21.75*** | -16.99*** | 0.000507 | 0.00342 | -0.00484 | 0.00106 |
| | (3.207) | (3.884) | (0.00207) | (0.00236) | (0.00900) | (0.0111) |
| High Risk Culture | -23.51*** | -18.45*** | 0.00218 | 0.00297* | -0.0119* | -0.0115 |
| | (2.613) | (2.812) | (0.00164) | (0.00171) | (0.00720) | (0.00804) |
| Size | 7.880*** | 5.285*** | 0.000513 | -0.000293 | 0.00671** | -0.00311 |
| | (1.068) | (1.287) | (0.000712) | (0.000783) | (0.00277) | (0.00368) |
| Board Size | 2.818*** | 1.970*** | -0.000292 | -0.000163 | 0.00329* | -0.000897 |
| | (0.727) | (0.760) | (0.000458) | (0.000462) | (0.00191) | (0.00217) |
| Independent Directors | -4.610*** | -3.332*** | 0.000238 | -0.000165 | -0.00580** | -0.000547 |
| | (0.914) | (0.970) | (0.000576) | (0.000590) | (0.00243) | (0.00277) |
| Leverage | 15.67 | 16.48 | 0.0196** | 0.0219** | 0.106*** | 0.184*** |
| | (14.21) | (17.65) | (0.00970) | (0.0107) | (0.0354) | (0.0504) |
| ROA | 35.90*** | 33.30*** | 0.00328** | 0.00355** | -0.0484*** | -0.0744*** |
| | (2.540) | (2.688) | (0.00159) | (0.00163) | (0.00696) | (0.00768) |
| Tobin's Q | -0.456 | -0.648 | 0.000365 | 0.000485 | -0.00249 | -0.00314 |
| | (1.526) | (1.518) | (0.000922) | (0.000923) | (0.00436) | (0.00434) |
| Median Age of Executive | 6.404*** | 5.843*** | 0.00162*** | 0.00157*** | 0.00512*** | 0.00310*** |
| | (0.0714) | (0.0895) | (5.36e-05) | (5.45e-05) | (0.000168) | (0.000256) |
| Volatility | 33.94** | -67.49*** | -0.0376*** | -0.0348*** | -0.138*** | -0.233*** |
| | (14.82) | (17.18) | (0.00995) | (0.0104) | (0.0413) | (0.0491) |
| Constant | -26.06*** | -17.40* | -0.00208 | -0.00367 | -0.0419* | 0.0359 |
| | (8.200) | (9.378) | (0.00546) | (0.00570) | (0.0220) | (0.0268) |
| Year Dummies | No | Yes | No | Yes | Yes | Yes |
| BHC fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 5,670 | 5,670 | 5,670 | 5,670 | 5,670 | 5,670 |
| R-squared | | 0.582 | 0.222 | 0.227 | | 0.134 |
| Number of BHCs | 572 | 572 | 572 | 572 | 572 | 572 |

## Appendix

### Table 1A: Risk Culture Dictionary

| Key Indicator | Bag of Words |
|---|---|
| Leadership | Ability, accentuate, accomplish, achieve, acknowledge, act, address, adherence, administration, advance, adviser, advisor, advocacy, aid, aim, anchor, appoint, appreciate, apprehend, apprise, apprize, aspire, assign, assistance, assurance, attain, authority, backbone, belief, board, care, ceo, certainty, cfo, chair, chairperson, change, charge, chief, clarify, command, commission, commit, committee, communicate, competence, complete, comprehend, consult, consummate, control, convey, corroborate, create, creativity, criticize, criticize, deal, decision, defend, delegate, demote, dependable, depute, designate, destine, determination, dictate, diligence, direct, directorate, discharge, discriminate, displace, distinguish, divide, downplay, effectuate, elevate, emphasize, empowerment, enact, encourage, end-all, enforcement, engage, enquire, enumerate, evince, excellence, executive, experience, expert, explain, express, fairness, firmness, foresight, format, goal, governance, guide, handle, head, honesty, identify, implement, indoctrinate, influence, inquire, inspire, instruct, integrity, intend, interact, interpret, judgement, judgment, justify, lead, leader, leadership, maintenance, management, mandate, master, misaddress, misdirect, misgovernment, misidentify, misrule, monitor, objective, observe, officialdom, optimize, optimize, order, organize, outcome, overact, oversee, perseverance, persistence, pertinacity, power, principles, professional, proficiency, promote, purpose, quality, re-address, re-emphasise, re-emphasize, reach, reassign, recognize, recognize, redirect, refinement, regard, relegate, reorganize, representative, resolution, resolve, resource, respect, risk, secern, senior, signal, solve, sovereignty, specify, stress, supervise, supervisor, support, sustainment, target, tenacity, tone, tone-from-the-top, tropicalize, underact, underscore, understand, upgrade, upkeep, values, vision, entrench |
| Strategy | Accomplishment, achieve, acquisition, action, active, administration, agenda, aid, align, alter, ameliorate, amend, appetite, approach, attain, business, buyout, change, complete, consult, contract, control, deliver, design, designate, develop, direction, discharge, effective, enforce, enterprise, execute, expand, forefront, formulation, framework, guidance, implement, improve, information, initiate, inquiry, instruction, investigation, know-how, maintenance, managing, manipulate, merger, method, misalign, models, modify, operation, organization, organize, orientation, outbalance, outcome, overbalance, performance, perspective, pioneer, plan, plot, policy, position, precedence, priority, procedure, process, product, program, project, redevelop, reformulate, reorganize, research, resource, restoration, restrain, restrict, revise, routine, rule , scheme, service, solution, strategy, structure, system, tactics, takeover, technique, undertaking, update, upkeep, venture, work, risk |
| Decision | Appointment, assignment, assurance, certainty, challenge, charge, choice, competent, confidence, control, convinced, decide, decision, designation, disclosure, effective, empowered, expert, exposure, feedback, guidance, independent, influence, information, inquiring, judgment, option, outcome, recommend, recorded, regulate, relevant, resolution, response, scenarios, selection, structured, transparent, risk |
| Control | Adhere, analyze, appoint, appraisal, assessment, assign, audit, authority, boundary, cap, certainty, challenge, change, charge, command, complaints, complete, compliance, confidence, confine, conformity, constrain, content, control, countersuit, criticism, decrease, define, delegate, delimit, delineate, demarcate, designate, dominance, effective, enforce, ensure, escalate, establish, evidence, examine, execute, experiment, explore, extent, extremity, finish, guarantee, honesty, identify, implement, increase, information, inspect, integrity, investigate, lawsuit, lessen, limits, litigation, mechanism, minify, modify, monitor, necessity, objection, obligate, obligation, observe, penalty, prerequisite, proactive, probe, proceedings, quantify, recognize, record, redefine, redouble, reexamine, regulation, regulatory, report, requirement, reset, restrict, review, scrutinize, situation, sue, testing, verify, whistleblow, fine, fraud, risk, protest, strike, unionize, crime, inhibit, mandate, prohibit, prevent, allegation, anticorruption, corruption, bail, breach, convict, felony |

| | |
|---|---|
| Recruitment | Ability, advancement, application, appoint , attrition, broadening, candidate, certification, chro, closedown , competence, competent, contract, course, development, drop-off, duty, education, employment, encourage, engage, enlist, enroll, feedback, growth, hiring, improvement, inspire, invest, job, knowledge, layoff, learn, occupation, overachievement, performance, position, procure, profession, proficiency, promote, proof, recruitment, role, rotation, scholarship, schooling, self-education, shutdown , skill, study, subcontract, support, training, tuition, underachievement, validation, work-study, workmanship, workplace, risk |
| Reward | Accredit, acknowledge, adjust, advancement, align, attainment, bonus, commission, compensation, consequence, credibility, cut, cutback, defrayment, discriminate, dividend, elimination, excellence, excessive, expel, fairly, fee, fine, fraud, gain, half-pay, honor, imposter, incentive, income, interest, law-breaking, layoff, misalign, offense, outcome, overachievement, overcompensation, payment, penalty, performance, premium, profit, progress, promotion, prosperity, punishment, quality, recognize, reduction, reimbursement, remunerate, remuneration, reorient, reputation, salary, share, stock, success, underachievement, wage, penalize, penalize, convict |
| Portfolio | Arrears, assets, collateral, credit, debt, deduction, default, delinquency, depreciation, derivatives, forfeiture, goal, guarantee, guaranty, impact, impairment, indebtedness, installment, instrument, investment, leverage, loan, loss, non-performing, nonpayment, overdue, principal, securitization, statement, target, write-off |

## Table A2: Description of Variables

This table presents a detailed description of the variables used in out analysis.

| Variable | Definition |
|---|---|
| Size | This is the size of the bank and is claculated as: the natural logaritm of total assets |
| Return on Asset | This is calculated as: Net Income/Total Assets |
| Legal Expense | This is calculated as: Total sum of legal expenses in a year |
| Leverage | This is calculted as: (long term debt + debt in current liabilities)/Total assets |
| Return on Equity | This is the Net Income Before Extraordinary Items and Discontinued Operations divided by Total Common Equity. This quotient is then multiplied by 100. Unit is perecnt |
| Number of Employees | The total employees as reported by the company. Unit is thousands |
| Tobin's Q | Tobin's Q is calculated a:s (Total Assets - Shareholder's Equity+Market Value of Equity)/Total Assets |
| Earning Per Share | The Earnings per Share (Primary) Excluding Extraordinary Items and Discontinued Operations. Unit is actual |
| Volatility | This is the volatility figure used in calculating Black-Scholes values for options. This is a standard deviation volatility calculated over 60 months. |
| Sales | The Net Annual Sales as reported by the company. Unit is millions |
| Board Size | The size of board of directors in a year |
| Female Director | This is an indicator variable equal 1 if there is at least one female director on the board in a year |
| Independent Directors | This is the number of independent directors on the board in a year |
| Institutional Ownership | Total institutional ownership, perecnt of shares outsatnding |
| Nationality Mix | This is the |
| Number of Meetings | The number of board meetings held during the indicated fiscal year. |
| Executive Salary Pay | This is calculated as: median dollar value of the base salary earned by the named executive officer during the fiscal year.Unit is thousands |
| Executive Bonus Pay | This is calculated as: median dollar value of a bonus earned by the named executive officer during the fiscal year. Unit os thousands |
| Executive Other Pay | This is calculated as:median dollar value of other annual compensation not properly categorized as salary or bonus. |
| Executive Shareholding | This is calculated as: Median shares owned by the executives, including options that are exercisable or will become exercisable within 60 days. Unit is thousands |
| Executive Total Compensation | This is calculted as: Salary + Bonus + Other Annual Pay |
| Fraction of Executive Incentive Pay | This is calculated as: 1 - (Salary + Bonus)/Total Compensation |
| Equity Fraction of Executive Pay | This is calculated as: Exective shareholding / Total Compensation |
| Executive Age (Median) | This is calculated as: Median age of executives as reported in the annual report. Unit is years |
| Director's Meeting Fee | This is the fee paid to each director for attending a meeting of the full board of directors. Unit is thousands |
| Number of Director's Stocks | The number of shares of stock (including restricted stock) that were granted to each non-employee director during the year. Unit is thousands |
| Number of Director's Options | The number of options which each non-employee director received during the year. Unit is thousands |