

## Article

# Managerial Self-Attribution Bias and Banks' Future Performance: Evidence from Emerging Economies

Javid Iqbal

Department of Management Sciences, COMSATS University Islamabad, Islamabad 44000, Pakistan;  
javidiqbal@comsats.edu.pk or ijavid100@gmail.com

Received: 16 March 2019; Accepted: 22 April 2019; Published: 25 April 2019

**Abstract:** The objective of the study was to predict the future performance of banks based on the contextual information provided in annual reports. The European Central Bank has observed that performance prediction models in earlier studies mainly rely on quantitative financial data, which are insufficient for the comprehensive assessment of banks' performance. There is a need to incorporate the qualitative information along with numerical data for better prediction. In this context, this study employed the attribution theory for understanding the contextual information of behavioral biases of management towards the expected outcomes. The sample consisted of 58 banks of 16 emerging economies, and the period covered from 2007–2015. Unsupervised hierarchical clustering was performed to identify the latent groups of banks within the data. For performance prediction, system GMM was employed, because it helped to deal with the endogeneity and heterogeneity problems. The results of the study were consistent with the attribution theory that management took credit for favorable expected outcomes and distanced from bad outcomes. An important policy implication of the study is that the prevalence of self-attribution bias of management in annual reports provides an additional source of information for the regulators to identify the banks at risks and take preventive measures to avoid the expected cost of failure. It can also help investors, and gives analysts a better tool for a comprehensive analysis of the profitability of prospective investments.

**Keywords:** performance predictions; attribution theory; hierarchical clustering; system GMM; endogeneity

## 1. Introduction

Banks are important financial intermediaries within the financial system, and play a significant role in the economic growth of a country by mobilizing the savings of depositors and making them available to the borrowers for productive ventures (Jasevičienė et al. 2013). Bank failures led to contagious effects, as witnessed in the global banking crisis of 2007–2008 that affected many sectors of major economies of the world (Edey 2009; Laeven and Valencia 2010; Fethi and Pasiouras 2010; Jasevičienė et al. 2013). Therefore, prediction of bank performance is important, because bank failures cause vulnerabilities to the financial system (De Haan and Vlahu 2016). The Federal Deposit Insurance Corporation of the US reported that almost 400 commercial banks were bankrupted in the US during 2008 to 2012. Furthermore, 10 large banking groups survived with the help of government bailout packages. The total estimated cost of the financial crisis that the US economy had to face ranged between \$10 trillion to \$20 trillion (Seamans 2013).

In this context, a large body of literature has emerged to deal with investigating bank future performance from traditional econometric models to advanced machine learning methods (see Horváthová and Mokrišová (2018) for more details). However, these models were exclusively built on quantitative financial data (Gaganis et al. 2006). In addition, the focus of the earlier studies was to

classify the banks into two groups, ‘failed’ and ‘non-failed’ (Lin et al. 2009). In response to that, the European Central Bank advised using both the business data as well as qualitative information for the extensive and comprehensive assessment of bank’s performance (European Central Bank 2010). This was earlier mentioned by the Securities and Exchange Commission (1987) of the US that numerical presentations and brief notes were insufficient for investors to assess expected profitability of banks. It is already evident from the research on non-financial firms that qualitative information not only helps in understanding the financial statement, but also contains incremental information about future financial performance (Li 2008, 2010b). Therefore, Craig et al. (2013) proposed that qualitative data should also be considered for examining the firm’s future performance. This view was also endorsed by Beattie (2014) who asserted that narratives could be used to assess the future financial performance of firms. Thus, qualitative data seems a valuable additional information source to be used for supplementing the quantitative financial data provided in corporate financial statements.

Researchers in non-financial firms have already incorporated qualitative data to empirically testing the hypothesis of interest, including the relationship between stock market returns and media discussions (Tetlock 2007), the relationship between firm performance and complexity of corporate textual documents (Li 2008; You and Zhang 2009), press releases and future financial performance (Davis et al. 2012), the language of CEO letter to shareholders and firm performance (Craig et al. 2013), speculative forecast of investment and behavior of investors (Mushinada and Veluri 2018), and the relationship between self-attribution bias of managers and firms’ future performance (Amernic and Craig 2006).

Self-attribution bias (SAB) of management based on attribution theory<sup>1</sup> has become a well-recognized qualitative text analysis technique in corporate finance research. It is a tendency of individuals taking credit for successful outcomes, but blaming unsuccessful outcomes to external factors (Clarke 1988; Daniel et al. 2001; Gervais and Odean 2001; Malmendier and Tate 2008; Billett and Qian 2008; Libby and Rennekamp 2012).

The prevalence of self-attribution bias is likely in annual reports, press releases, conference calls, and the President’s letter to shareholders, because managers try to attribute favorable outcomes to their abilities to get rewards, and unfavorable outcomes to external factors to avoid penalties. Accounting literature has widely analyzed CEO’s self-attribution bias from different perspectives, especially linking it with the firm’s future financial performance. Keusch et al. (2012) argued that self-attribution behavior of the managers could alert the misleading information related to financial performance, especially during bad times of firms.

Research on non-financial firms has already started to investigate the possibilities of using self-attribution bias of managers for testing hypothesis of interest. However, the banking sector has not benefited from this line of research. De Haan and Vlahu (2016) maintained that banks are different from non-financial firms in many important ways: (a) banks have a unique role of creating credit for economy, (b) they are highly regulated due to contagious effects, and (c) banks are highly leveraged. Likewise, banks act as delegated monitors of-resources, share risk of the economy, and ensure that firms efficiently use the resources extended to them in terms of loans (Allen and Carletti 2012). Therefore, the studies conducted on non-financial firms cannot be generalized to the banking sector.

In addition, Demirgüç-kunt et al. (2012) explained that a substantial body of economic theory also emphasized the comparative importance of banks, and concluded that banks are important in early stages of economic development of a country. The authors suggested that in the transition stage of economies, banks have a comparative advantage in financing collateralized, and standardized endeavors than those of other financial institutions. Fink et al. (2005) also suggest that banks are more important in the majority of the emerging economies. Similarly, Djalilov and Piesse (2016) also argue that when the economies are in the early transition stage, the share of banks towards whole financial system and GDP growth are relatively high.

The ‘immeasurably important’ place of banks as central of economic activities in the emerging economies underscore the merit for studies that exclusively investigate the bank future financial

<sup>1</sup> Attribution theory has been discussed in detail in literature review section.

performance. To the best of my knowledge, there is no study that has used self-attribution bias of management focusing on the emerging economies to predict the future performance of banks. This study attempts to fill this gap.

This study poses the research question, does the self-attribution of management in annual reports offer incremental predictive power, over and above the models based on the traditional quantitative financial data alone?

The study contributes in two important ways. First, it contributes toward a comprehensive performance prediction models of banks in which contextual information using self-attribution bias of management is leveraged with the traditional numerical data to predict the future performance of banks. This could help potential, as well as existing, investors with a comprehensive information about future profitability. Second, management of banks have the advantage of superior information, and are aware that regulators mostly focus on few financial ratios, which can be distorted (Gandhi et al. 2019). This study helps in reducing the information asymmetry between the management and shareholders about expected future performance by getting signals with the help of self-attribution of management in annual reports of the banks (Principal-Agent model).

The remaining part of the paper has been organized as follows: the next section shed lights on existing literature, taking up each issue under scope of attribution theory. Section 3 provides the research methodology that describes the sample selection, how managerial attribution bias was measured, and describes the econometric models. Section 4 explores the data with the help of descriptive statistics, correlation matrix and scatter plot matrix. Section 5 presents the exploratory data analysis with agglomerative hierarchal cluster analysis to find the latent groups within the data. Section 6 provides the estimates of the model that leads to the discussion of results. Section 7 concludes the research findings.

## 2. Literature Review

Canbas et al. (2005) have explained that it is important to predict the future performance of banks, so that the regulators could take timely actions to mitigate the disastrous outcomes resulting from banking crisis. In this context, a large body of the literature emerged that dealt with investigating the future performance of banks. Methods developed earlier for manufacturing firms were also adopted in financial firms with some modifications. The methods included an analysis of financial ratios, discriminant analysis ((Beaver 1966; Altman 1968), and logistic models, data envelopment analysis or DEA, (Fethi and Pasiouras 2010), and machine learning algorithms, particularly, the neural networks (Ravi Kumar and Ravi 2007). These methods provide the foundation for researchers to predict the future financial performance of banks from more than two decades (Board et al. 2003). In addition to the earlier models, stress test which has become more widespread after the financial crisis of 2007–2008 that is performed under hypothetical bad economic conditions to observe that a bank has required capital to bear the impact of adverse conditions (Petrella and Resti 2013). However, this test is performed by management, and based on hypothetical situations which may be strategically manage by the managers to show a better financial condition of the bank (principals-agent problem).

Gaganis et al. (2006) criticized earlier developed models on the basis that these models exclusively relied on quantitative financial data. The European Central Bank (2010) also observed that researchers were developing models based only on the financial data for predicting the bank's performance. However, there is a large quantity of unstructured qualitative information that should also be used in conjunction with numerical data for a comprehensive performance prediction model. In this perspective, Smith and Taffler (2000) augmented that qualitative sections of annual report provide “nearly twice the quantity of information as do the basic financial statements”. Moreover, the qualitative information could also help in understanding financial statement, and signals about future financial performance (Li 2008, 2010b). This was earlier suggested by Craig et al. (2013) that the textual data should also be considered in examining whether firms are healthy or at risk.

Earlier studies on non-financial firms have empirically evidenced that future performance could be predicted using qualitative information in corporate documents (Li 2008, 2010a; Craig et al. 2013).

Thus, the qualitative data seems to be a valuable additional information source to supplement the financial data available in corporate financial statements. Research in the manufacturing sector has already begun to explore the narratives in textual disclosures in many ways, including complexity of text, sentiment analysis, and self-attribution bias. In this study, self-attribution bias of management is used to predict the future financial performance of banks.

Managerial attribution bias can be understood in the context of attribution theory (Bloomfield 2008; Hooghiemstra 2001; Clatworthy and Jones 2003). Attribution theory was first proposed by Heider in 1958, but Weiner and colleagues developed a theoretical framework that has become a major research paradigm in social psychology (Sparck Jones 1972; Li 2008; Baginski et al. 2000; Clatworthy and Jones 2003; Merkl-davies et al. 2014; Leary and Kowalski 1990; Moosa and Ramiah 2017). The theory explains that achievers believe success was attained due to their abilities, and efforts (internal attribution), whereas, failure would be considered due to bad luck, and was not their fault (external attribution). Therefore, Bloomfield (2008) claims that attribution theory in psychology predicts that using internal attribution or external attribution in textual disclosures is an import indication of a firm's future performance.

Literature has witnessed that attribution of management provides involuntary signals about the expected outcomes of firms in the textual sections of annual reports, especially, in the CEO's letter to the shareholders, and management discussion and analysis (Aerts 1994; Clatworthy and Jones 2003; Merkl-davies et al. 2014). Amernic and Craig (2006) *provided that when the outcomes of firms were good, the CEO took credit, and attributed to internal factors (their ability, skills, vision and foresightedness)*. In contrast, if the outcomes were unfavorable, they made external attributions (bad economic, and market situations) in an attempt to distance from bad performance, so that they would not be held personally accountable.

Accounting literature has widely analyzed management's self-attribution bias from different perspectives. For instance, Craig et al. (2013) analyzed the bankrupted Indian firm 'Satyam', and found that there was a shift in first-person singular pronouns to first-person plural pronouns in the CEO's letter to shareholders. It also showed blame shifting signals about the bad outcomes of the firm. Similarly, Clatworthy and Jones (2003) examined the chairman's letter to shareholders of the 50 top and bottom UK companies, and found that management took credit for good news, and blamed the external environment for bad outcomes. Likewise, Li (2010) analyzed the managements' tendency towards self-serving attribution with the help of computational linguistic technique in management discussions and analysis (MD&A), and found that the inclination of attributional words by management was positively related to firms' future performance. More recently, Lehmberg and Tangpong (2018) examined how top management communicate the bad or good firm performance to the stakeholders. The authors found a significant and positive relationship between subsequent good performance and internal attribution, while bad performance was attributed to external factors.

Researchers also used self-attribution biases from many other aspects other than firm performance. For instance, Adam et al. (2015) analyzed whether the past speculative forecast of managers increased more investments by testing the data of 92 North American gold mining firms over a period of 1989 to 1999. The results of the study demonstrated that bad forecast did not reduce investments, because managers attributed successful outcomes to their abilities, while losses were attributed to bad luck or some other external factors.

From the investors' perspective, Mushinada and Veluri (2018) analyzed the behavior of investors for their investment decisions using behavioral biases. The authors used 1290 stocks traded in the Bombay Stock Exchange of the India during 2004 to 2012 examining whether the self-attribution bias existed among investors about their earnings' forecasts. The results of the study revealed that when the forecast was accurate, the investors took credit, while the wrong forecast was attributed to external factors, especially the excessive volatility in the stocks. Likewise, Chen et al. (2016) investigated the relationship between attribution of managers for their earnings forecast, and investors' utilization of that attribution. The study found that investors followed the internal attribution of management for their investments. In another study, Asay et al. (2018) investigated the attribution with the help of personal pronouns of CEO for the prediction of winning law suits or

good/bad news of CEO. Results clearly indicated that participants' likelihood of winning law suits or good performance was associated with the use of more personal pronouns.

### 3. Methodology

#### 3.1. Data

The author selected all the emerging economies which were reported in IMF emerging economies list. Specialized banks were excluded to keep the sample homogeneous. Annual reports were downloaded from the websites of banks, and the financial data were obtained from Bureau van Dijk (BvD)<sup>2</sup>. Another condition to the sample selection was the availability of CEO letters to shareholders, and management discussion and analysis in annual reports. Moreover, banks were who published annual reports in their national language excluded from the sample, because this could create reliability problems in the construction of self-attribution bias indices.

After accounting for all changes, the data consisted of 58 banks from 16 emerging economies, and the period covered 2007 to 2015. The list of banks is provided in Appendix A.

#### 3.2. Variables and Their Definitions

##### 3.2.1. Bank Performance Indicator

*Return on Average Equity (ROAE):* Return on average equity was taken as a performance indicator of banks. It was calculated as the net income divided by average total equity. ROAE is an internal performance measure of shareholder value and has been widely used for performance prediction of banks (Aerts 2001; Petria et al. 2015; Beccalli et al. 2015; Yao et al. 2018; Akhisar et al. 2015). It is a fundamental ratio that tells the investors, how effectively management uses their money. It proposes a direct assessment of the financial return of shareholders' investment. This ratio shows whether the management is growing banks' value at an acceptable rate.

##### 3.2.2. Performance Determinants

*Self-Serving Attribution Bias (CEO letter to shareholders):* In the corporate context, more use of first, and second-person pronouns is an indication of taking credit for good outcomes, compared to third-person pronouns. Researchers have widely analyzed corporate narrative to explore the relationship between CEO's self-attribution bias in letters to the shareholders and the firm's future financial performance (Clatworthy and Jones 2003; Amernic and Craig 2006; Craig et al. 2013; Lehmberg and Tangpong 2018). In this study, a positive relationship is expected between self-attribution bias of CEO in letter to shareholders and bank's future performance.

*Self-Serving Attribution Bias (Management Discussions and Analysis):* This section of annual report explains the firm's overall performance, challenges face by the management, internal and external risks involved in the operations, and indications about the future prospects. A similar methodology was adopted to calculate the self-attribution bias of management in management discussions and analysis as described in CEO letter to shareholders. Earlier studies have evidenced the positive relation between self-attribution of management in management discussions and analysis and firm's future financial performance (Li 2010a; Lehmberg and Tangpong 2018; Aerts 2001, 2005). Similarly, a relationship is also expected in the bank's financial performance.

*Total Assets:* In this study, total assets represent the size of bank with the absolute values in million US dollars. There was huge variation in the assets of the banks in the emerging economies. Thus, the logarithm of total assets (Log\_Assets) was employed as a proxy for bank size. The size of assets could provide higher profit up to a certain level, thereafter, the profitability could be lowered as compared to small banks. Thus, the relationship between total assets and banks performance could be negative, because percentage of profit does not increase with the equivalent proportion of assets (Trujillo-Ponce 2013; Panta 2018; Terraza 2015; Shehzad et al. 2013). Some other studies have also

<sup>2</sup> Bureau van Dijk (BvD) is a Moody's Analytics company that has database for business information.

shown an insignificant relationship of total assets and firms' performance (Athanasoglou et al. 2008; Petria et al. 2015).

*Assets Growth Ratio:* The assets growth ratio was calculated by current year assets minus last year assets divided by last year assets. Assets growth ratio indicates the percentage increase or decrease over the prior year. The increase in profitability of the bank depends upon the increase in quality of assets. Higher quality assets would increase the profitability of the banks. The earlier studies suggested that as the assets increased, profitability of the bank were also increased (Mathuva 2009; Ahamed 2017; Bougatef 2017; Yao et al. 2018). Thus, the relations between assets growth and banks future performance is expected to be positive.

*Non-Performing Loans to Gross Loans Ratio:* Non-performing loans to gross loans was calculated as total non-performing loans divided by gross loans of the bank. The loans are classified as nonperforming when the borrower defaults or declares bankruptcy. It measures the effectiveness of a bank in receiving repayments on its loans. The higher the ratio, lower the profitability of the banks (Trujillo-Ponce 2013; Panta 2018; Petria et al. 2015).

*Tier1Capital Ratio:* Tier1capital ratio was calculated as total tier1capital<sup>3</sup> divided by total risk-weighted assets<sup>4</sup>. The tier1capital ratio measures a bank's core capital. In 2015, under Basel III, the minimum tier 1 capital ratio was 6 percent. The regulators use this ratio to determine, whether a bank is well capitalized, undercapitalized or adequately capitalized relative to the minimum requirements. Since, net income is spread over increased equity, the relationship between performance of banks and tier1capital is expected to be negative (Stovrag 2017).

*Loans to Asset Ratio:* Loans to asset ratio was calculated as total loans held by bank's borrowers divided by total assets of the bank. The loans included cash deposits at other banks, financial assets, securities, advances to the borrowers. This ratio indicates to what extent assets are devoted to loans. Literature has shown that the relationship between loans to assets ratio and bank performance could be either positive or negative. Higher the ratio, lower the liquidity position of bank, and may face a higher risk of failure (Goddard et al. 2013). In addition, a bank holding more liquid assets (lower loan to asset ratio) may suffer from lower profitability (Demirgüç-Kunt and Huizinga 1999; Yao et al. 2018). On the other hand, higher loans to borrowers could provide more interest income to the banks (Trujillo-Ponce 2013).

*GDP Growth Rate:* GDP growth is the rate of change in gross domestic products of a country. Positive relationship between bank's performance and GDP growth rate is expected, because increase in economic activities would lead to demand for bank's transactions, resulting higher profitability of the banks (Petria et al. 2015; Trujillo-Ponce 2013; Yao et al. 2018; Athanasoglou et al. 2008). Other studies showed that there was no relationship between GDP growth and banks profitability, and was more country specific (Shehzad et al. 2013; Djalilov and Piesse 2016).

*Interest Rate Spread:* Interest rate spread is the country level variable that refers to the difference between the borrowing and the lending rates of banks. Higher spread shows that banks are advancing loans at a higher premium. Thus, the relationship between interest rate spread and bank's performance is expected to be negative.

*Exchange Rate (1 \$ equals local currency):* The exchange rate is the price of one country's currency in terms of foreign currency. The financial data for this study was obtained from Bureau van Dijk (BvD), which was available in US dollar. Exchange rate has indirect impact on the profitability of the banks. There is a tendency among some countries to keep local currency weaker to stimulates exports. Because the export transactions are executed through banking channels, it increases the non-interest income of the banks (Medura 2006; Pagratis et al. 2014). Thus, a positive relationship is expected between exchange rate and return on average equity.

The list of variables and their definitions are provided in Table 1.

<sup>3</sup> Tier1capital consists of shareholders' common equity and retained earnings.

<sup>4</sup> Risk-weighted assets are calculated by adjusting category of asset class for risk to determine a bank's real exposure to the potential losses.

**Table 1.** Variables and Definitions.

Variable Name	Definition	Expected Signs
ROAE	Ratio of net profit to the average equity one year ahead.	
SAB_CEO	Self-Attribution bias measured from CEO letter to shareholders.	+
SAB_MD&A	Self-Attribution bias measured from management's discussions and analysis.	+
Assets (Log)	Log of total assets as a proxy of banks size.	+/-
Assets Growth (%)	Increase in total assets from last year.	+/-
NPL/GLoans (%)	Ratio of non-performing loans to gross loans.	-
Tier1Capital (%)	Ratio of core capital to risk adjusted assets.	-
Loans/Assets (%)	Ratio of loans to total assets.	+/-
Interest Spread (%)	Interest rate lending minus borrowing of banks.	-
GDP Growth (%)	GDP growth rate of an economy.	+
Exchange Rate (equals 1\$)	Exchange rate of an economy where 1 US dollar equal to local currency.	+

### 3.3. Measurement of Managerial Self-Attribution Bias

Text preprocessing is an exhaustive process, especially when someone needs to analyze unstructured qualitative information. The purpose of preprocessing is to convert qualitative text in a usable form to get insight for further analysis. There are different software programs which have built-in functions that convert text into predefined objects. Many of those software programs do not provide customized options to the users to process text according to specific requirements. For instance, two software programs are commonly used in literature; (i) Diction, and (ii) Linguistic Inquiry and Word Count (LIWC) that contain predefined built-in functions for text analysis. Our text analysis required customized functions, because annual reports of selected sample countries did not follow any standard pattern to measure the managerial self-attribution bias. Therefore, an open source R software is used for measurement of attribution bias indices, because it provides a rich selection of text preprocessing packages.

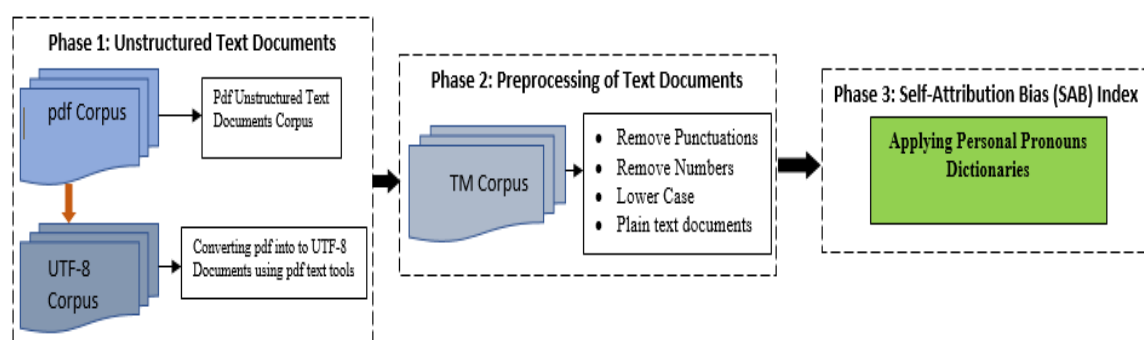
I also used SAS software for agglomerative hierarchical clustering, and estimation of system GMM.

After downloading the annual reports from the websites of the banks, two sections namely, CEO's letter to shareholders, and the management discussions and analysis (MD&A)<sup>5</sup> were extracted.

At an initial stage, all the documents were converted from pdf to plain, UTF-8 encoded text using "pdftools" of R, and were stored in a "Corpus". A corpus is considered to be a "library" of all the original documents that have been converted to plain, UTF-8 encoded text. Figure 1 exhibits the whole process for the measurement of self-attribution bias indices. Self-attribution bias was constructed by first and second-person pronouns minus third-person pronouns. To construct the self-attribution bias from the CEO's letter to shareholders (CEO SAB), two separate dictionaries of first and second-person pronouns, and third-person pronouns were constructed. These two dictionaries were employed to obtain the score by matching with the corpus of CEO letter to shareholders. Each term of pronoun was counted as many times it appeared in the document, which reflected the stress made by management to take credit of good performance and vice versa.

A similar process was adopted to obtain the score of first and second-person pronouns and third-person pronouns of management discussions and analysis (MD&A SAB).

<sup>5</sup> Some countries use the term Directors' reports to shareholders, or financial and operational review, instead of management discussions and analysis.



**Figure 1.** Construction of self-attribution bias (SAB) Indices using Natural Language Processing (NLP).

### 3.4. Econometric Model Using System GMM

For performance prediction, three models were estimated. First, prediction of banks performance using only two self-attribution bias (SAB) indices. Second, the model was estimated with the help of SAB indices along with bank level quantitative financial variables. Third model was estimated with full set of SAB indices, bank level variables, and macroeconomic indicators. The notion behind estimating three models was to observe the predictive power of SAB indices, over and above the models that were based on quantitative financial data alone.

Dynamic panel models are linear regression models that consist of individual effects, yield individual-level errors, and overall model residual errors. It allows for dependent variables to depend on its own value from its previous time, thus making the model dynamic. The following Equation (1) is specified for model 1.

$$ROAE_{i,t+1} = \beta_1 ROAE_{i,t} + \beta_2 CEO\_SAB_{i,t} + \beta_3 MD\&A\_SAB_{i,t} + v_i + \varepsilon_{i,t} \quad (I = 1, \dots, N, t = 1, \dots, T) \quad (1)$$

where  $ROAE_{it}$  is the return on average equity one year ahead taken as a performance indicator of banks.  $CEO\_SAB$  is the self-attribution bias calculated from CEO letter to shareholders, and  $MD\&A\_SAB$  is the self-attribution bias calculated from management discussions and analysis. The  $v_i$  is individual effects, and the  $\varepsilon_{it}$  observation-level regression errors.

The second model is shown in Equation (2):

$$ROAE_{i,t+1} = \beta_1 ROAE_{i,t} + \beta_2 CEO\_SAB_{i,t} + \beta_3 MD\&A\_SAB_{i,t} + \beta_4 Assets(log)_{i,t} + \beta_5 Assets\_Growth_{i,t} + \beta_6 NPL/GLoans_{i,t} + \beta_7 Tier1Capital_{i,t} + \beta_8 Loans/Assets_{i,t} + v_i + \varepsilon_{i,t} \quad (2)$$

The second model consisted of SAB indices, bank level quantitative financial variables that includes log of assets, assets growth ratio, ratio of non-performing loans to gross loans, tier1capital ratio, and loans to assets ratio.

The third model is shown in Equation (3):

$$ROAE_{i,t+1} = \beta_1 ROAE_{i,t} + \beta_2 CEO\_SAB_{i,t} + \beta_3 MD\&A\_SAB_{i,t} + \beta_4 Assets(log)_{i,t} + \beta_5 Assets\_Growth_{i,t} + \beta_6 NPL/GLoans_{i,t} + \beta_7 Tier1Capital_{i,t} + \beta_8 Loans/Assets_{i,t} + \beta_9 Interest\_Spread_{i,t} + \beta_{10} GDP\_Growth_{i,t} + \beta_{11} Exchange\_Rate_{i,t} + v_i + \varepsilon_{i,t} \quad (3)$$

The third model consists of a full set of SAB indices, bank level quantitative financial variables, and macroeconomic indicators that includes interest rate spread, GDP growth, and exchange rate.

### Diagnostic Checks

Initially, the model was estimated using fixed effects, and random effects, and observed the individual effects using F-test. The Hausman test was estimated for model's selection either to use the fixed effects or the random effects. The problem with the fixed effects and the random effects are that the error term being correlated with the lagged dependent variable created an endogeneity problem, and even the error term is not autocorrelated (Greene 1996, p. 536). Both models are



presented in Appendix C. Thus, the regressors are said to be endogenous when random errors are correlated to the regressors. In fact, endogeneity is a major methodological concern in many areas of research in corporate finance (Abdallah et al. 2015). If endogeneity is present in the model, then the statistical inference from the analysis may be biased (Abdallah et al. 2015). Endogeneity may occur due to omitted variables, which may result in the error term being correlated with the explanatory variables. Alternatively, the endogeneity may be of the dynamic type, whereby the past realizations of the dependent variable influence current realizations of one or more of the explanatory variables. Finally, endogeneity can be of the simultaneous type, where the contemporaneous realizations of both the dependent variable and the explanatory variables affect each other (Abdallah et al. 2015; Roberts and Whited 2012).

The model in this study potentially faces two types of endogeneity issues, i.e., omitted variable bias, and the past realization of dependent variable in terms of earning persistence. For example, SAB indices were developed from the textual data to capture the private information of the management. Nevertheless, these indices may not necessarily present a perfect proxy of private information, and is affected by the agency problem. This is referred to as endogeneity of omitted variable bias. Moreover, lagged dependent variable is included in the right-hand side of the model, making it dynamic. The reason to include the lagged dependent variable is the performance persistence of the banks, which was the continuity of the current earnings affected by the magnitude of the accruals. The higher persistent earnings are accompanied with more ability to maintain the current earnings (Lipe 1990). Hence, failure to address the endogeneity may lead to poor statistical inference (Abdallah et al. 2015).

Some of researchers have mentioned that there should be a theoretical reasoning for considering the regressor as endogenous. However, for a robustness check, the Durbin-Wo-Hausman test was used to test whether the theoretical reasoning justify the empirical reason of endogeneity problem.

Similarly, the model could also face heteroscedasticity problem. The White test was used for detecting the heteroscedasticity of residuals, and the null hypothesis is that the variances for the errors are homoscedastic, where  $H_0 : \sigma_i^2 = \sigma^2$  for all  $i$ . In presence of the heteroscedasticity, estimates are still unbiased, but become inefficient. However, the standard errors of estimates are wrong, leading to incorrect inferences (White 1980).

Arellano and Bover (1995) and Blundell and Bond (1998) proposed system GMM that could help deal with the endogeneity problems (Abdallah et al. 2015), and heterogeneity problems (Baum et al. 2003). System GMM consists of two sets of equations, each with its own internal instruments, i.e., the first set is levels equations, and second set is the difference equations (Zheng et al. 2017).

A generalization of the linear regression model is an autoregressive (AR) model. The AR test was conducted to identify the serial correlation in residual. The null hypothesis is that there is no serial correlation within residuals.

The validity of the instruments was confirmed via a the Sargan test that checks for overidentifying restrictions, and it is asymptotically distributed as a  $\chi^2(n)$  with  $n$  degrees of freedom under the null hypothesis that the instrument set is appropriate for the data at hand.

In addition, robust standard errors were calculated, which are more reliable than conventional standard errors (Gutierrez and El-khattabi 2017).

The estimates of all models were conducted with the help of SAS 'proc panel' procedure.

#### 4. Descriptive Statistics

Table 2 provides the descriptive statistics of 58 banks of 16 emerging economies. Furthermore, detailed descriptive statistics were calculated at country level, so that insight could be obtained in depth about how one country's statistics were different from another. Country level statistics are shown in Appendix B. The Mean of ROAE was 16 percent, which was close to median, whereas the lowest value was −27 percent, and the highest was 41 percent, and the standard deviation was 6.75. It was found that the highest ROAE belonged to the NDB bank of Sri Lanka in 2012, because in that year NDB bank became the first investment bank. It made divestment of AVIVA insurance, and made new investment with A/A corporation. Further, economic growth, and the post war period in Sri

Lanka also helped to increase profitability. Lowest ROAE −27.80 belonged to an Askari bank of Pakistan due to huge non-performing loans, and the bank written-off non-banking assets.

**Table 2.** Descriptive Statistics of SAB, Financial, and Macroeconomic Variables.

Variable	Mean	Median	Minimum	Maximum	Std. Dev.
ROAE_Lead1 (%)	15.75	15.43	−27.8	41.02	6.18
CEO SAB	37.84	29.00	−139	181.00	42.63
MD&A SAB	72.78	14.00	−191	1128.00	149.64
Total Assets (million dollars)	187,224.9	32,380.45	142.81	3421,363.00	486,273.00
Assets Growth (%)	10.01	10.38	−22.32	52.85	10.76
NPL to Gross Loans (%)	4.24	3.15	0.01	21.89	3.70
Tier1Capital (%)	11.95	11.67	4.02	26.7	3.29
Loans to Assets (%)	58.95	60.54	16.21	83.06	11.38
Interest Rate Spread (%)	5.09	4.81	0.15	35.59	3.86
GDP Growth (%)	5.16	5.23	−6.56	15.24	3.19
Exchange Rate (equal 1\$)	756.88	42.34	0.27	13,389.41	2612.74

SAB of CEO ranged between −139 to 181 that means there was polarity of managerial attribution, because in some of the banks, the CEO used more first and second-person pronouns, and in other cases, the CEO used more third-person pronouns. SAB of MD&A also showed a greater dispersion, and ranged between −191 to 1128. The highest values of SAB of MD&A related to those banks who had published annual reports of more than 700–1000 pages, meaning that a large number of pages were also allocated to MD&A section.

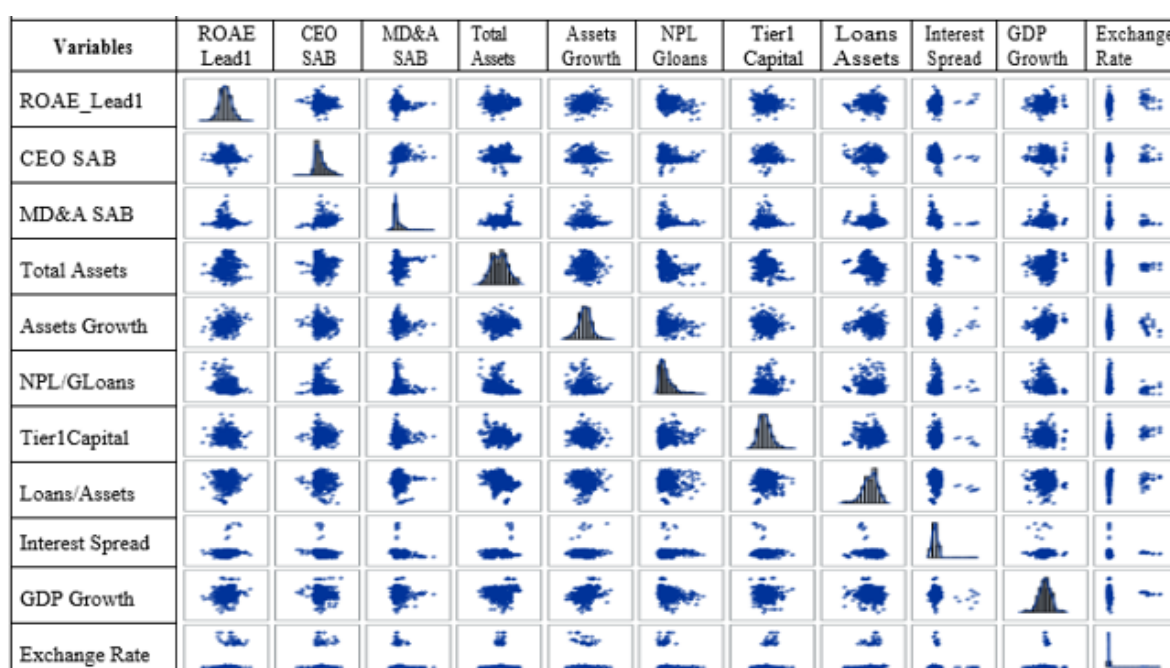
Financial variables also presented greater variability in the data. Absolute values were reported in millions of US dollars converted from local currency on the last day of the financial year of the respective country. The mean value of total assets was 187,225 US dollars, and ranged between 143 million US dollars to 3,421,363 million US dollars, having a standard deviation of 486,273. Largest value belonged to the ICBC, because it was one of the largest banks in the China, whereas, lowest assets were reported by the PABC bank of the Sri Lanka. It was important to notice that the Chinese banks were much larger than any other banks in selected samples. The lowest value of assets growth was −22% belonged to the Bank Alfalah of the Pakistan in 2008, which was year of financial crisis, and 56% belonged to the ThanaChart bank of the Thailand. The PABC bank of the Sri Lanka has the highest non-performing loans with the value of 27%, whereas, an average value was 4% close to median.

In the macroeconomic indicators, the Do-Brazil bank had the highest interest rate spread ranged between 30–35 percent. The high interest rate spread was due to the Brazil's large public debt, and debt services. Another reason of high interest rate spread was the history of default that's the government had to pay a high default risk premium to attract foreign capital. Moreover, the exchange rate was lowest in the Indonesia for 1 dollar equals 13,389 Indonesian Rupiah. The GDP Growth rate was lowest in Hungary in 2009 with −6.56, and highest in Singapore with 15.24 in 2010.

The author provides the correlation matrix in the Table 3 showing that there was no multicollinearity between the variables. Figure 2 provides the scatter plot matrix of all the SAB indices, quantitative financial variables, and macroeconomic indicators.

**Table 3.** Correlation Matrix of SAB, Financial and Macro-Economic variables.

Variables	ROAE Lead1	ROAE	CEO SAB	MD&A SAB	Log Assets	Assets Growth	NPL/ Gloans	Tier1 Capital	Loans/ Assets	Interest Spread	GDP Growth
ROAE	0.69										
CEO SAB	−0.11	−0.10									
MD&A SAB	−0.13	−0.14	0.38								
Assets (Log)	0.04	0.04	0.04	0.19							
Assets Growth	0.26	0.29	−0.10	−0.06	0.00						
NPL/GLOANS	−0.24	−0.31	−0.07	−0.12	−0.38	−0.19					
Tier1Capital	−0.10	0.00	0.05	−0.05	−0.17	−0.10	0.13				
Loans/Assets	−0.07	−0.03	−0.08	−0.03	−0.25	−0.08	−0.05	0.01			
Interest Spread	0.12	0.13	−0.11	−0.20	0.04	0.01	0.07	−0.04	−0.15		
GDP Growth	0.15	0.14	−0.03	−0.03	0.16	0.34	−0.30	−0.10	−0.12	−0.13	
Exchange Rate	0.25	0.25	0.07	−0.04	0.01	−0.07	−0.08	0.23	0.06	0.00	0.03

**Figure 2.** Scatter Plot Matrix of Independent and Dependent Variables.

## 5. Agglomerative Hierarchical Clustering Analysis

Cluster analysis is an important form of exploratory data analysis that tries to explore hidden groups within the data. It is an unsupervised machine learning method, which is widely used for classification problems, and exploring the data based on some similarity of features<sup>6</sup> through a structured pattern (Wu et al. 2009; Myatt and Johnson 2014). As a result, observations in the same cluster are more analogous than those in other clusters (Hsu et al. 2007). More importantly, clustering helps to discover latent natural group, and categorize data into a hierarchical set of clusters organized in a tree structure (Loewenstein et al. 2008).

There are different clustering techniques, including k-mean clustering, k-mode clustering, but the underlying study adopted the hierarchical clustering, because it is mainly used for small data sets, where each observation forms hierarchy. SAS ‘proc cluster’ procedure was used for agglomerative hierarchical cluster analysis.

To calculate the similarities between observations, it is required to normalize all the variables, so that the distance between these observations are computed to prevent disproportionate weights, and biases. The data used in this study consisted of total assets reported in million in the dollar, and other financial variables are in percentage. Therefore, the whole dataset was normalized using the following formula:

<sup>6</sup> Features in a sense of variables, which is named in machine learning.

$$z_n = \frac{x_n - \bar{x}}{\sigma_X}$$

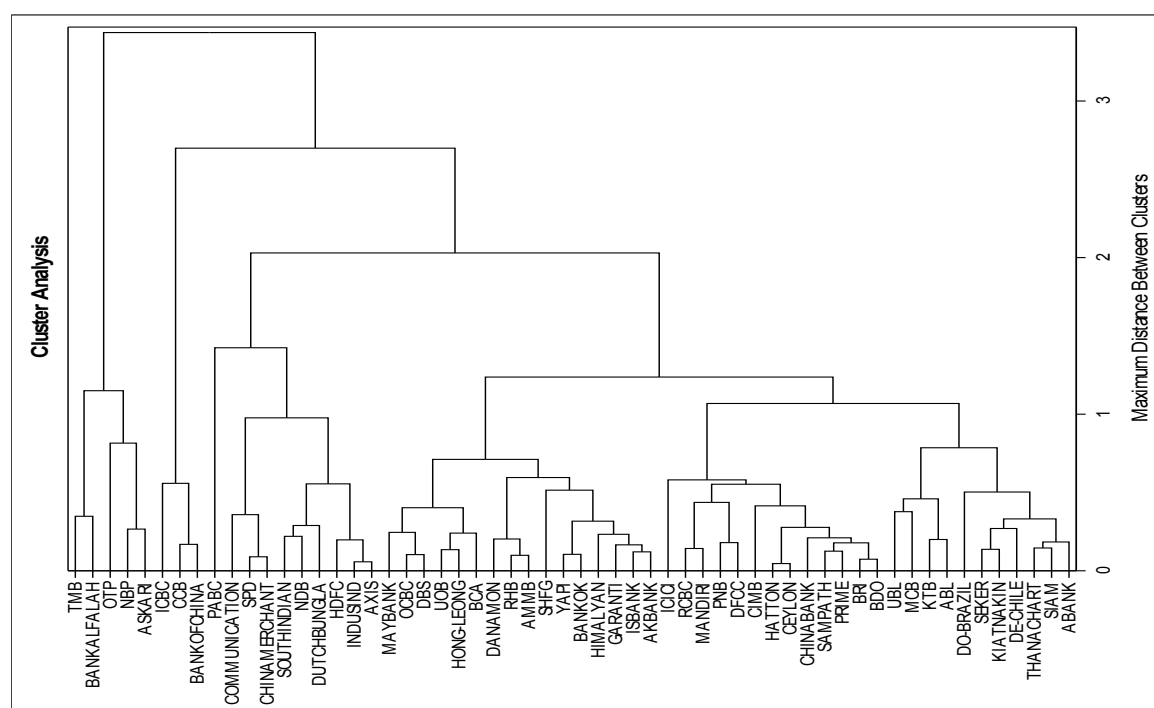
where  $z_n$  is the standardized value of observation  $n$ ,  $x_n$  is the original value of observation  $n$ ,  $\bar{x}$  and  $\sigma_X$  are the mean, and standard deviation of the variable  $X$ . The sample data consisted of a panel of 58 banks, and the time period covers from 2007–2015. Directly making clusters of the panel data could match the observation of one bank' year with another bank. This could distort the whole data, and understanding of clusters. Therefore, the mean of each bank was taken, where each bank represented a single observation. Moreover, the data consisted of 11 variables that could also create problems for differentiating variables for making the clusters. Principal component analysis was performed to identify the variables that were loaded on the first factor (see Appendix C). Hence, four variables: assets growth, total assets, NPL to gross loans, and GDP growth, were loaded on factor-1 that were taken to form clusters.

### 5.1. Measurement of Distance between Observations

The distance between observations was calculated using Euclidean distance matrix (Myatt and Johnson 2014). It is shown as follows.

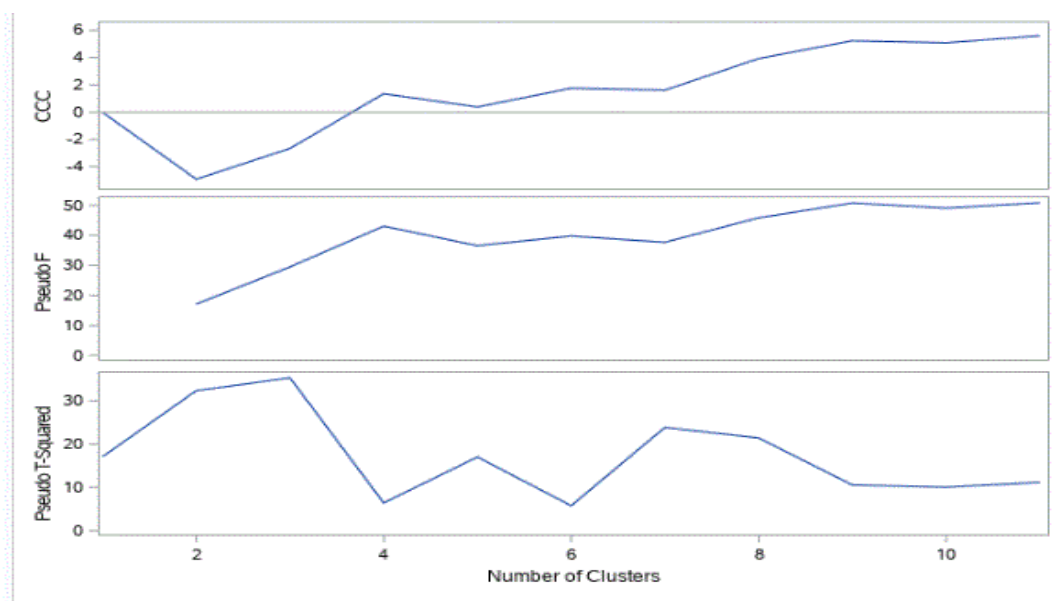
$$d = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

where  $d$  is the distance between  $p$ , and  $q$  the observations of  $n$  variables. Thus, the Euclidean method calculates the distance of each combination of all the observations. Then, the linkage rule creates clusters by comparing the distance between the clusters, and observations. This process continues until a diagram called dendrogram is created, which explains how the observations are connected to each other based on similarities of observations. The dendrogram is shown in Figure 3.



**Figure 3.** Dendrogram of Hierarchical Clustering.

To determine the optimum number of clusters, cubic clustering criterion (CCC), (b) pseudo T statistic, and (c) pseudo F statistics were used. The highest peak of CCC is considered as the optimum number of clusters. In our case, the peak of the plot at 4 CCC provided 9 optimal number of clusters, as shown in Figure 4.



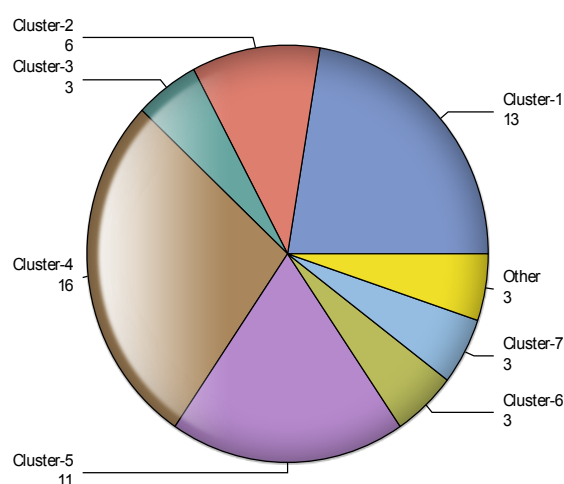
**Figure 4.** Criteria for Selecting number of Clusters.

Pseudo  $F$  statistic (PSF) also supplemented the selection of 9 clusters by showing the highest peak (middle graph). In pseudo  $t^2$  statistic, observing the graph from right to left, until the first height gave the indication of about 9 as the number of clusters. Table 4 shows the number of banks in each cluster, percent frequency, and cumulative percent frequency.

Figures 5 and 6 show the distribution of clusters in pie chart, and bar chart respectively.

**Table 4.** Frequency of Banks in Each Cluster.

Cluster	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Cluster-1	13	22.41	12	68.97
Cluster-2	6	10.34	19	84.48
Cluster-3	3	5.17	22	74.14
Cluster-4	16	27.59	38	46.55
Cluster-5	11	18.97	49	18.97
Cluster-6	3	5.17	52	89.66
Cluster-7	3	5.17	55	98.28
Cluster-8	2	3.45	57	93.10
Cluster-9	1	1.72	58	100.00



**Figure 5.** Pie Chart of Clusters.

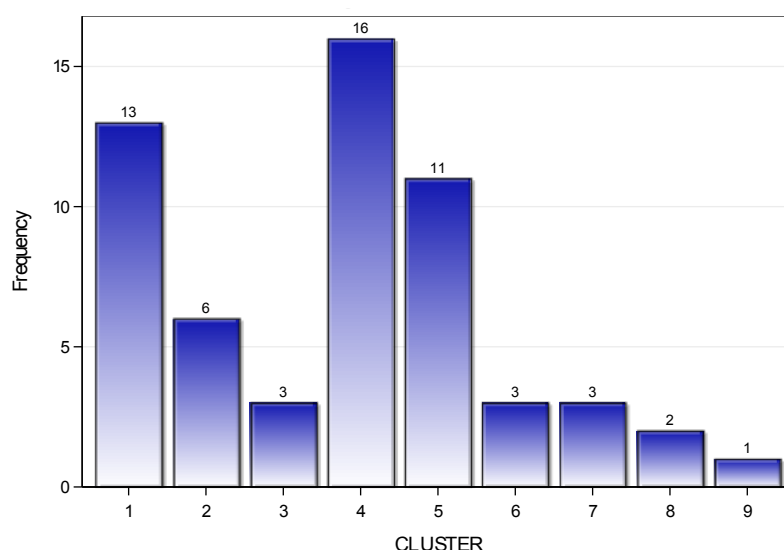


Figure 6. Bar Chart of Clusters.

### 5.2. Cluster Profile

To profile the clusters, descriptive statistics were calculated on the unstandardized dataset based on the clusters' distribution of banks to explore the reasons of similarities, and differences between clusters.

It helps to understand the patterns between clusters, and the rationale behind how these clusters were formed. Statistics included mean, median, standard deviation, skewness, and kurtosis shown in the Table 5.

Table 5. Profile.

Clusters	Statistics	Frequency	Assets Growth	GDP Growth	NPL/Loans	Total Assets
Sample	Mean		10.01	5.16	4.24	187,224.90
	Median		10.38	5.23	3.15	32,380.45
	CV		107.48	61.78	87.33	259.73
	Skewness		0.07	−0.21	1.90	4.24
	Kurtosis		0.99	1.31	4.04	19.20
Cluster-1	Mean	13	11.17	5.82	4.58	45,945.75
	Median		11.38	5.64	4.43	8027.08
	CV		13.91	9.84	25.34	190.72
	Skewness		−0.78	0.89	1.19	2.86
	Kurtosis		−0.17	2.28	0.96	8.67
Cluster-2	Mean	6	15.71	6.83	1.78	21,990.68
	Median		16.01	7.14	1.64	8783.72
	CV		12.10	8.06	38.05	121.11
	Skewness		0.03	−0.91	1.51	1.00
	Kurtosis		−0.37	−1.88	2.52	−1.29
Cluster-3	Mean	3	17.92	9.26	1.07	543,580.65
	Median		18.40	9.26	0.99	481,710.87
	CV		9.02	0.00	22.67	27.68
	Skewness		−1.23		1.31	1.54
Cluster-4	Mean	16	6.92	4.43	2.40	105,430.02
	Median		7.32	4.85	2.77	82,865.12
	CV		30.69	19.15	44.75	78.14
	Skewness		−0.24	−0.21	−0.25	0.75
	Kurtosis		−0.81	−1.50	−1.07	−0.50

Cluster-5	Mean	11	9.66	3.30	5.69	58,061.85
	Median		9.97	3.45	5.51	8502.77
	CV		20.99	8.23	29.26	213.44
	Skewness		−0.71	−1.08	0.86	3.15
	Kurtosis		1.05	1.26	0.17	10.16
Cluster-6	Mean	3	13.59	9.26	1.74	2,028,139.68
	Median		13.59	9.26	1.57	1,922,427.58
	CV		4.41	0.00	23.52	15.89
	Skewness		0.02		1.56	1.32
Cluster-7	Mean	3	3.68	2.48	13.79	20,778.74
	Median		4.28	3.45	13.95	13,033.19
	CV		79.03	67.51	5.47	106.01
	Skewness		−0.89	−1.73	−0.94	1.39
Cluster-8	Mean	2	0.20	3.30	7.82	13,180.11
	Median		0.20	3.30	7.82	13,180.11
	CV		391.08	6.34	22.38	83.45
Cluster-9	Mean	1	18.14	6.10	9.44	387.31
	Median		18.14	6.10	9.44	387.31

Observing the mean statistic of four variables used in cluster analysis, the mean value of assets growth was 10 percent, and GDP growth at 5.16 percent. However, Natural Language Processing (NLP) to gross loans were only 4.24 percent. These values were also close to median, meaning that the distribution of these variables was symmetrical. These statistics were compared with cluster statistics.

Table 5 shows that cluster-1 included those banks which had a smaller mean value, and median as compared to the whole dataset, except total assets. It included 13 banks: four banks of Philippines, four banks of the Sri Lanka, two banks of Indonesia, and one bank each from the India, Bangladesh, and the Malaysia. Cluster-2 consisted those banks which had high assets growth, banks belonged to those countries with a better GDP growth rate, less NPL, and low averaged total assets as compared to whole sample statistics. This cluster consisted 6 banks: four Indian banks, and one bank each from the Sri Lanka, and Bangladesh. Cluster-3 comprised of banks with highest assets growth, and banks belonged to those countries with the highest GDP growth rate, lowest NPL, and largest banks in the selected sample. It was interesting to notice that all the three banks in cluster-3 belonged to the China.

Cluster-4 included 16 banks where five from Turkey, three each from Malaysia and Singapore, and one bank each from Thailand, Indonesia, and Nepal. Banks included in this cluster had lower assets growth which was also manifested in countries with lower GDP growth.

Banks included in cluster-4 had low averaged NPL, and total assets as compared to whole sample statistics. Cluster-5 carried those banks having low averaged assets, GDP growth, and low NPL, as compared to total sample statistics, which was also supplemented by other statistics. Banks included in this cluster were relatively small banks. Three Chinese banks also constructed the cluster-6 without including any other bank from emerging economies. Cluster-3 and cluster-6 carried the Chinese banks, but the major differentiators between both were assets growth, and the size of banks. Cluster-3 banks were small sized, but the highest assets growth as compared to the banks of cluster-6.

Cluster-7 included banks which had the highest non-performing loans with second lowest assets growth. The data also revealed that the Askari bank and NBP bank from the Pakistan, and the OTP bank from Hungary had persistent high non-performing loans. In cluster-8, the major differentiator was the assets growth, having the lowest mean value and highest coefficient of variation. Finally, cluster-9 carried only one bank that did not merge in any other cluster. The reason being that the separated bank was the highest mean value of assets growth of 18.14 percent, and smallest bank in terms of total assets. The list of banks in each cluster along with their belonging to the country are shown in Table 6.

**Table 6.** List of Banks in Each Cluster.

Clusters	Country	Name of Banks in Each Cluster				
1	Sri Lanka	Ceylon	Hatton	Sampath	DFCC	
	Philippines	China Bank	PNB	RCBC	BDO	
	Indonesia	BRI	Mandiri			
	India	ICICI				
	Bangladesh	Prime				
2	Malaysia	CIMB				
	India	AXIS	Indusind	HDFC	Southindian	
	Sri Lanka	NDB				
3	Bangladesh	Dutch Bungla				
	China	China Merchant	Communication	SPD		
4	Turkey	AKBank	BCA	Garanti	ISbank	Yapi
	Malaysia	AMMB	Hong-Leong	Maybank		
	Thailand	Bankok				
	Indonesia	Danamon				
	Singapore	DBS	OCBC	UOB		
	Nepal	Himalayan				
5	Turkey	ABank	Seker			
	Pakistan	ABL	MCB	UBL		
	Chile	De-Chile				
	Brazil	Do-Brazil				
6	Thailand	Kiatnakin	KTB	Siam	Thanachart	
	China	Bankofchina	CCB	ICBC		
7	Pakistan	Askari	NBP			
	Hungary	OTP				
8	Pakistan	BankAlfalah				
	Thailand	TMB				
9	Sri Lanka	PABC				

## 6. Econometric Analysis

### 6.1. Pre-Diagnostics Checks

Initially, fixed effects and random effects were estimated. The results of both models, fit statistics and diagnostics tests are reported in Appendix D. The Hausman test was used to select the fixed effects or random effects model. Based on the results, we reject the null hypothesis, and suggested to employ fixed effects model. In addition, F-test was also performed to observe the bank effects in the model, and the null hypothesis is that bank effects are all zero. Based on the results, we rejected the null hypothesis at 1 percent ( $p$ -value  $< 0.01$ ), and suggest that bank effects exist in the data. In both models, fixed effects and random effects, CEO SAB and MD&A SAB indices were insignificant. However, assets growth was significant at 1 percent, NPL to gross loans and tier1capital ratios, and exchange rate from macroeconomic indicator were negative and significant at 5 percent. Assets growth was positive and significant at 5 percent.

Although, the Hausman test suggested for fixed effects model, but the problem with the fixed effects was that the error term was correlated with the lagged dependent variable, creating an endogeneity problem (Greene 1996, p. 536). Moreover, the model could also face the heterogeneity problem. The White test was used for detecting the heteroscedasticity of residuals. Based on the results, the null hypothesis was rejected at 1 percent significance ( $p < 0.01$ ), and suggested that the variances of error were heteroscedastic<sup>7</sup>. Thus, the results produced could be biased, and lead to poor statistical inferences.

In addition, self-attribution bias indices were also theoretically endogenous, which were further tested using *Durbin-Wu-Hausman test*, known as the Hausman specification test. The null hypothesis is that there is no measurement error. The results of the test rejected the null hypothesis at 5 percent ( $p < 0.026$ ), and suggested for use as the instrumental variable<sup>8</sup>.

<sup>7</sup> White Test for Heteroscedasticity, 182.6 ( $p < 0.01$ ).

<sup>8</sup> Hausman's Specification test, OLS efficient under Ho, 2SLS Consistent under H1, 14.74 ( $p < 0.026$ ).



To deal with the endogeneity and heterogeneity problem, the author used the GMM system, developed for dynamic panel models by Arellano and Bover (1995), and Blundell and Bond (1998). System GMM helps to deal endogeneity problem using instrumental variables. It also deals the heteroskedasticity using the orthogonality conditions to allow for efficient estimation in the presence of heteroskedasticity, even of unknown form (Baum et al. 2003).

Table 7 shows the results of model-1 (Equation (1)) of system GMM, parameter estimates, fit statistics such as means squared error (MSE), Root MSE, and r-squared.

Table 7. Model-1.

Variables	Signs	Estimates	t-Value	SE
ROAE	+	0.3775 ***	23.97	0.0157
CEO SAB	+	0.0829 ***	9.07	0.0091
MD&A SAB	+	0.0806 ***	4.01	0.0201
<b>Fit Statistics</b>				
RMSE		0.8897		
MSE		0.7915		
SSE		317.39		
<b>Diagnostic Tests</b>				
AR(m) Test Lag 1 (Statistic)		2.47 **		
$p >  Statistic $		(0.0136)		
AR(m) Test Lag 2 (Statistic)		0.40		
$p >  Statistic $		(0.6918)		
Sargan Test (Statistic)		51.01		
$p > ChiSq$		(0.4695)		

\*\*\*/\*\*/\* means significance at 0.01, 0.05, and 0.10 level respectively. Return on average equity (ROAE) is the dependent variable, explanatory variables included one-year lagged SAB of CEO and SAB of MD&A Indices.  $p$ -value in parenthesis.

## 6.2. Post Diagnostics Checks

Post diagnostic checks were conducted to determine the robustness of the empirical results. The Sargan test provided the test of overidentification restrictions. Therefore, the null hypothesis was rejected, and suggest that the instrument set is appropriate for the data ( $p$ -value < 0.4695). Similarly, the AR test was conducted to identify the serial correlation in residual, and the null hypothesis that there is no serial correlation, and the results suggested no serial correlation at the second-order ( $p$ -value < 0.6918).

Observing the model-1, lagged ROAE was positive and statistically significant at 1 percent ( $p$ -value < 0.01). The relationship between CEO SAB and MD&A SAB with ROAE was positive and statistically significant at 1 percent ( $p$ -value < 0.01). This suggested that these variables could be used for the prediction of banks performance at least one-year prior to the announcement of the results. These results were consistent with the attribution theory that management took credit for expected favorable outcomes, and distance itself when expected outcomes were poor (Bloomfield 2008).

Table 8 provides the results of model-2 which is a combination of SAB indices, and five financial ratios. Fit statistics showed little improvement in MSE, RMSE, and SSE as compared to model-1. In addition, there was no serial correlation within the residuals at second-order ( $p < 0.9051$ ). The Sargen test also showed that the instruments set was appropriate ( $p$ -value < 5275).

Table 8. Model-2.

Variables	Sign	Estimates	t-Value	SE
ROAE	+	0.2980 ***	12.01	0.0248
CEO SAB	+	0.0354	1.55	0.0229
MD&A SAB	+	0.1474 ***	3.00	0.0492
Total Assets (log)	-	0.5283 ***	4.51	0.1173
Assets Growth (%)	+	0.1104 ***	3.99	0.0277
NPL/Gross Loans (%)	-	0.2398 ***	2.85	0.0840
Tier1Capital (%)	-	0.4379 ***	6.21	0.0705
Loans/Assets (%)	-	0.3260 ***	5.16	0.0631
<b>Fit Statistics</b>				
RMSE		0.8437		
MSE		0.7119		
SSE		281.91		
<b>Diagnostic Test</b>				
AR(m) Test Lag 1 (Statistic)		-2.50 ***		
$p >  Statistic $		(0.0125)		
AR(m) Test Lag 2 (Statistic)		0.12		
$p >  Statistic $		(0.9051)		
Sargan Test (Statistic)		48.65		
$p > ChiSq$		(0.5275)		

\*\*\*/\*\*/\* means significance at 0.01, 0.05, and 0.10 level, respectively. Return on average equity (ROAE) is the dependent variable, explanatory variables included one-year lagged SAB, and financial variables. P-value in parenthesis.

Observing the Table 8, MD&A SAB was positive, and highly significant at 1 percent ( $p$ -value < 0.01) without major change in coefficients. It means that the self-attribution bias of management provided additional information, over and above the traditional bank level quantitative variables. However, CEO SAB became insignificant by adding the financial ratios in the model-2. This suggested that CEO managerial attribution provided indications about future performance, when it was used without financial variables. All the bank level financial variables were significant at 1 percent ( $p$ -value < 0.01). For example, the size of banks in terms of total assets was negative and statistically significant at 1 percent. Non-performing loans to Gross Loans showed a negative relationship with future ROAE and was significant at 1 percent. Likewise, the tier1Capital ratio, and Loans to assets ratio were negative and significant at 1 percent. Nevertheless, assets growth was positive and statistically significant at 1 percent ( $p$ -value < 0.01).

Finally, Table 9 demonstrates the results of model-3 using the full set of SAB indices, quantitative financial variables, and macroeconomic indicators. Fit statistics were further improved in MSE, RMSE, and SSE from previous both model-1 and model-2. The AR test provided no serial correlation at the second order within the residual ( $p$ -value < 0.9422), and the Sargen test demonstrated that instruments used in the model were appropriate ( $p$ -value < 0.3189).

Adding macroeconomic indicators within the model, CEO attribution bias index became positive and significant at 5 percent ( $p$ -value < 0.05). Management discussions and analysis persistently remained positive and significant at 1 percent ( $p$  < 0.001). Adding the macroeconomic indicators did not make any change in the bank level financial variables as compared to model-2. Some of the coefficient values were changed, but all the financial variables remained significant at 1 percent ( $p$ -value < 0.01). However, all the macroeconomic indicators were insignificant in model-3.

Table 9. Model-3.

Variables	Sign	Estimates	t-Value	SE
ROAE	+	0.2284 ***	6.58	0.0347
CEO SAB	+	0.0580 **	2.21	0.0262
MD&A SAB	+	0.1360 ***	3.18	0.0428
Assets (log)	−	0.3877 **	2.37	0.1634
Assets Growth (%)	+	0.1466 ***	4.41	0.0332
NPL/Gross Loans (%)	−	0.2259 ***	2.83	0.0798
Tier1Capital (%)	−	0.5513 ***	7.85	0.0703
Loans/Assets (%)	−	0.3782 ***	4.02	0.0941
Interest Rate Spread (%)	+	0.1737	1.49	0.1164
GDP Growth (%)	−	0.0031	0.08	0.0383
Exchange Rate (%)	−	0.1829	0.88	0.2081
<b>Fit Statistics</b>				
RMSE		0.8390		
MSE		0.7040		
SSE		276.66		
<b>Diagnostic Test</b>				
AR(m) Test Lag 1 (Statistic)		−2.41 ***		
$p >  Statistic $		0.0159		
AR(m) Test Lag 2 (Statistic)		−0.07		
$p >  Statistic $		0.9422		
Sargan Test (Statistic)		51.01		
$p > ChiSq$		0.3189		

\*\*\*/\*\*/\* means significance at 0.01, 0.05, and 0.10 level, respectively. Return on average equity (ROAE) is the dependent variable, explanatory variables included one-year lagged SAB, and financial variables.

### 6.3. Discussions

Observing model-1 and model-3, the results suggested that managerial self-attribution bias in CEO letter to shareholders predicted the future ROAE, either alone or in combination of quantitative financial variables and macroeconomic indicators. This suggests that CEO of banks use more first and second-person pronouns in the letter to shareholders to take credit for expected positive outcomes. Thus, it also suggests that the self-attribution bias of CEO provides additional predictive power, over and above the traditional quantitative financial data that researchers have been using from many decades. The results produced in this study were consistent with prior studies conducted on non-financial firms (Clatworthy and Jones 2003; Amernic and Craig 2006; Craig et al. 2013; Lehmberg and Tangpong 2018).

Similarly, self-attribution bias of management discussions and analysis also had predictive power, over and above quantitative financial and macroeconomic indicators. This index was positive and statistically significant in all the three models. Management discussions and analysis is a comprehensive document, where management discusses about overall bank performance, challenges faced by management, future plans, and expected outcomes. Thus, the results were supported by previous studies who had evidenced the positive relationship between self-attribution of management and firm's future financial performance (Li 2010b; Lehmberg and Tangpong 2018; Aerts 2001, 2005). The results of both indices were also consistent with the attribution theory that managerial attribution bias provided contextual information for understanding the behavioral bias of management that could be linked to the expected outcomes of the banks (Bloomfield 2008; Hooghiemstra 2001; Clatworthy and Jones 2003).

In financial variables, the size of the banks in terms of assets was negative and statistically significant at 1 percent and 5 percent in model-2, and model-3 respectively, meaning that as the size of banks increased, future ROAE of the bank decreased. Thus, the size of banks in terms of assets

could stipulate higher profit up to a certain level, thereafter, the profitability could be lower. Moreover, this also shows higher unproductive, non-interest earning assets in the balance sheet of the banks. Contrary to the study of Athanasoglou et al. (2008) and Petria et al. (2015), the results of both models were support by earlier studies that there was an inverse relationship between total assets and ratio of future ROAE, because the percentage of profit did not increase with the same proportion of assets (Trujillo-Ponce 2013; Panta 2018; Terraza 2015; Shehzad et al. 2013). Likewise, the ratio of asset growth was positive and significant at 1 percent in model-2 and model-3. This is directly proportional to the banks' profitability, i.e., as the assets of the bank increases, banks have more loanable investments to earn profits. Positive relationship also shows that banks have higher interest earning assets. Thus, an increase in customer borrowing means an increase in the interest income of the banks. These results were in line with the earlier studies that an annual increase in assets, profitability of the banks also increased (Mathuva 2009; Ahamed 2017; Bougatef 2017; Yao et al. 2018).

Non-performing loans to gross loans was also negative and the significant at 1 percent in model-2 and model-3 that means profitability decreases with an increase in non-performing loans, either used with self-attribution bias of management or in combination of macroeconomic indicators. Earlier discussed in the methodology section that loans were classified as non-performing, if the borrower defaulted or declared bankruptcy. Our results were supported by previous studies that higher the ratio, lowered the ROAE of the banks (Trujillo-Ponce 2013; Panta 2018; Petria et al. 2015).

The relationship between the performance of banks and Tier1capital was negative, because return on average equity was calculated by dividing the net income over equity. Thus, net income is spread over increased equity. The coefficient of tier1capital was highest in all the variables used in models. The results supported the view that the higher the ratio, the lower the future performance of banks (Stovrag 2017). Finally, the loans to assets ratio also showed a negative and significant relationship with future ROAE. This ratio indicates to what extent assets were devoted to loans as opposed to other assets, including cash, securities, and equipment. The results presented in this study are consistent with Goddard et al. (2013) that higher the ratios, lower the liquidity position of the bank, and may face a higher risk of failure.

In economic indicators, it was expected that there would be a positive relationship between banks future performance in terms of ROAE and GDP growth rate. However, the author failed to find any relationship as witnessed by (Shehzad et al. 2013; Djalilov and Piesse 2016). The reason of insignificance would be the direction of dependent and independent variables. It was argued in the introduction that banks promote economic growth. However, the present study considered GDP growth as determinant of future ROAE. Similarly, the relationship between interest rate spread and future banks performance was also insignificant. Finally, contrary to the studies of (Medura 2006; Pagratis et al. 2014), this study failed to find any relationship between future ROAE and exchange rate.

Based on the analysis, the study answers the research question that self-attribution bias of management in annual reports of banks provided additional predictive power, over and above the quantitative financial variables. The results provided in this study were more robust, because models were estimated in a way that reduced the endogeneity and heterogeneity problems along with valid instruments.

## 7. Conclusions

Banks are important financial intermediaries within the financial system because they help to promote the economic growth of a country. Nevertheless, the banking sector crisis had a history for their role in the financial turmoil, especially in the 2007–2008 sub-prime crisis. Such a banking crisis had not only reduced the industrial production, entrepreneurial innovation, trade, but also had knock-on effects to the rest of the world. Therefore, the prediction of bank performance is important for regulators to take pre-emptive actions to avoid huge losses. However, developed models for performance prediction of banks were only based on quantitative financial data.

In this research, self-serving attribution bias, which is a text analysis technique based on attribution theory, was used for a contextual understanding of managerial behavioral bias towards the outcomes of banks. The notion behind self-attribution was that management uses more first and second-person pronouns as compared to third person pronouns in annual reports if they anticipate better future performance.

The sample consisted of 58 banks of 16 emerging economies for a period from 2007–2015. For exploratory data analysis, hierarchical clustering from unsupervised machine learning was performed to detect latent groups within the data. It was observed that some of the banks joined the clusters based on asset growth, other banks formed clusters due to high NPL. GDP growth also worked as a differentiator for grouping the banks into the cluster. Finally, the size of the banks in terms of assets distinguished the small banks with medium and large banks.

To predict the future performance of banks, system GMM proposed by Blundell and Bond (1998) was used to estimate the models. System GMM helps to deal with the endogeneity and heterogeneity problems within the data. The results of the study have shown that there existed a strong relationship between managerial attribution bias and the future performance of banks in emerging economies. The results were consistent with the attribution theory, which predicted that managers took credit for good outcomes and distanced from bad outcomes. Therefore, the study concludes that self-attribution bias of management signals about the future performance of banks, over and above the quantitative financial data provided in financial statements.

### 7.1. Policy Implications

*Regulators:* Any technique that could even marginally improve the ability of regulatory authorities to make an assessment of overall bank performance would be beneficial, because supervisors may intervene in a timely manner to avoid bank failure (Gandhi et al. 2019). The findings of the study have shown that self-attribution bias of management in annual reports provides incremental information, over and above the quantitative data provided in financial statements of banks. Such information could be used as indications of early warnings, and help the regulatory authorities of emerging economies to differentiate the bad performing banks. As a result, the banks could be supervised more efficiently, and take preventive, and corrective measures to avoid huge losses to the investors, and government.

*Investors and Analysts:* The findings also show that the contextual information of management in emerging economies' banks can help reduce information asymmetry between shareholders and management (principal-agent theory). It can provide the existing as well as potential investors with a better tool for a comprehensive assessment of banks profitability for prospective investments.

*Researchers:* Use of textual information from the textual information of banks is a relatively unexplored area where researchers may yield rich insights for testing further hypotheses of interest.

### 7.2. Limitation of the Study

Sample data was relatively small, because the author had to follow certain criteria for inclusion of banks in the sample. For example, annual reports of banks must include CEO letter to shareholders, and management discussions and analysis (MD&A). The CEO letters were mainly available in the annual reports; however, MD&A was not regulatory requirement in emerging economies. In addition, the banks in emerging economies were either state owned or family owned or became public limited only recently. Therefore, most of the banks did not have these two sections. These issues reduced the sample size of banks.

Cultural differences of countries may hold individual effects, while using first-person pronouns, and second person pronouns by management in annual reports. For instance, it might be the convention in some sample countries that management uses more plural pronouns in these two sections.

Likewise, English was not the first-language in annual reports of banks in emerging economies. Therefore, use of pronouns, while writing annual reports may hold some implications for construction of self-attribution bias indices.

Finally, the literature on textual analysis was mainly focused on developed economies. Most of the banks in emerging economies are either state owned or family owned. Thus, management might not necessarily take credit for good performance, and bad performance attributed to external factors, due to restricted power of management over the board of directors and shareholders.

**Author Contributions:** All the work related to problem identification, analytical solution, interpretation and writing was done by J.I. This article is from PhD dissertation.

**Funding:** This research received no external funding.

**Acknowledgments:** Author is very grateful to two anonymous referees for their helpful suggestion. Author is responsible for any remaining errors.

**Conflicts of Interest:** The author declares no conflict of interest.

## Appendix A. List of Banks in Sample

Country	Bank Name	Country	Bank Name
Bangladesh	Dutch Bungla Bank	Singapore	DBS
	Prime Bank		OCBC
Brazil	Do-Brazil Bank		UOB
Chile	De-Chile	Sri Lanka	Ceylon Bank
China	Bank of China		DFCC Bank
	CCB		Hatton Bank
	China Merchant Bank		NDB Bank
	Communication bank		PABC Bank
	ICBC		Sampath Bank
	SPD Bank		Bangkok Bank
Hungary	OTP Bank	Thailand	Kiatnakin
India	Axis Bank		KTBank
	HDFC		Siam Bank
	ICICI Bank		Thana Chart Bank
	Indusind Bank		TMB Bank
	South Indian Bank	Turkey	ABank
Indonesia	BCA Bank		AKBank
	BRI Bank		Garanti Bank
	Danamon Bank		ISBank
	Mandiri Bank		Seker Bank
Korea	SHFG Bank		Yapi
Malaysia	AMMB Bank		
	CIMB Bank		
	Hong-Leong		
	May Bank		
Nepal	RHB Bank		
	Himalayan Bank		
Pakistan	ABL Bank		
	Askari Bank		
	Bank Alfalah		
	MCB		
	NBP		
	UBL		
Philippines	BDO Bank		
	China Bank		
	PNB		
	RCBC Bank		

## Appendix B. Descriptive Statistics by Country

Country	N Obs	Variable	Mean	Median	Minimum	Maximum	CV	Std Dev
Bangladesh	18	ROAE_LEAD1	19.75	19.71	8.45	35.33	4.40	8.77
		CEO SAB	28.94	28.50	8.00	57.00	48.26	3.97
		MD&A SAB	23.72	25.00	−17.00	58.00	93.34	22.14
		Total Assets	2084.23	2058.43	719.95	3288.66	41.15	857.72
		Assets Growth	14.11	15.39	−4.85	26.42	60.42	8.53
		NPL/GLoans	3.19	2.82	1.15	7.67	57.97	1.85
		Tier1Capital	8.64	9.44	4.65	10.95	21.20	1.83
		Loans/Assets	66.57	66.19	59.63	77.61	8.39	5.58
		Interest Spread	4.24	4.48	1.87	5.64	29.24	1.24
		GDP Growth	6.14	6.06	5.05	7.06	9.37	0.58
Brazil	9	Exchange Rate	73.98	74.15	68.60	81.86	6.66	4.93
		ROAE_LEAD1	20.72	18.41	15.44	29.61	25.26	5.23
		CEO SAB	34.89	39.00	12.00	52.00	41.52	14.49
		MD&A SAB	17.00	4.00	3.00	67.00	142.58	24.24
		Total Assets	426,310.40	481,190.04	200,340.00	554,609.90	30.44	129,776.38
		Assets Growth	9.70	7.61	−12.33	46.31	169.32	16.43
		NPL/GLoans	4.48	4.17	2.30	8.61	51.63	2.31
		Tier1Capital	8.83	8.84	5.83	11.39	24.74	2.18
		Loans/Assets	45.15	43.13	40.38	50.90	9.24	4.17
		Interest Spread	29.97	31.34	19.58	35.59	18.87	5.66
Chile	9	GDP Growth	2.69	3.00	−3.77	7.53	129.89	3.49
		Exchange Rate	2.11	1.95	1.67	3.33	23.72	0.50
		ROAE_LEAD1	22.50	22.48	18.61	27.11	12.11	2.72
		CEO SAB	53.89	57.00	26.00	82.00	30.79	16.59
		MD&A SAB	121.44	133.00	45.00	182.00	43.98	53.41
		Total Assets	29,227.61	21,741.00	12,583.00	49,514.80	49.27	14,399.67
		Assets Growth	12.81	12.77	−8.78	38.98	132.54	16.98
		NPL/GLoans	5.02	3.53	1.34	9.92	62.91	3.16
		Tier1Capital	10.03	8.92	7.25	13.32	24.38	2.45
		Loans/Assets	79.00	79.13	75.29	83.06	3.08	2.44
China	54	Interest Spread	3.91	4.09	1.91	5.77	29.88	1.17
		GDP Growth	3.63	3.98	−1.04	5.84	63.14	2.29
		Exchange Rate	533.99	522.46	483.67	654.12	10.17	54.31
		ROAE_LEAD1	19.31	19.84	11.71	27.90	17.44	3.37
		CEO SAB	17.87	8.00	−3.00	123.00	124.97	22.33
		MD&A SAB	14.00	2.00	−19.00	100.00	209.31	29.30
		Total Assets	1,285,860.17	996,974.44	153,270.17	3,421,363.16	72.79	936,017.03
		Assets Growth	15.75	16.78	1.55	33.35	40.78	6.42
		NPL/GLoans	1.40	1.13	0.44	3.17	51.05	0.72
		Tier1Capital	9.81	10.08	4.02	13.48	18.03	1.77
Hungary	9	Loans/Assets	52.78	53.27	44.52	59.58	6.38	3.37
		Interest Spread	3.03	3.06	2.85	3.33	4.44	0.13
		GDP Growth	9.26	9.40	6.92	14.23	23.10	2.14
		Exchange Rate	6.61	6.46	6.14	7.61	6.90	0.46
		ROAE_LEAD1	9.07	8.36	−7.37	26.25	101.55	9.22
		CEO SAB	38.56	41.00	18.00	70.00	52.12	20.09
		MD&A SAB	−5.22	−6.00	−23.00	9.00	−192.63	10.06
		Total Assets	45,631.94	46,948.80	37,396.11	51,701.50	9.34	4264.13
		Assets Growth	0.52	3.97	−13.69	16.13	2071.80	10.75
		NPL/GLoans	13.95	15.72	6.90	19.57	37.99	5.30
India	45	Tier1Capital	13.12	13.30	10.03	17.40	19.07	2.50
		Loans/Assets	71.27	72.06	59.93	79.49	8.95	6.38
		Interest Spread	2.80	2.67	0.26	5.21	49.83	1.39
		GDP Growth	0.55	0.89	−6.56	4.05	571.95	3.13
		Exchange Rate	214.20	207.94	172.11	279.33	14.62	31.32
		ROAE_LEAD1	15.83	17.11	6.24	21.60	24.29	3.85
		CEO SAB	−1.09	6.00	−139.00	71.00	−4428.10	48.31
		MD&A SAB	19.07	−3.00	−191.00	323.00	651.83	124.29
		Total Assets	49,115.58	33,227.81	3115.59	138,506.86	89.84	44,125.07
		Assets Growth	15.13	13.89	−7.37	38.70	73.36	11.10

		NPL/GLoans	1.88	1.36	0.81	5.65	63.26	1.19
		Tier1Capital	11.23	11.54	6.70	14.92	16.73	1.88
		Loans/Assets	58.18	58.54	47.33	68.01	9.22	5.37
		Interest Spread	6.93	7.10	4.40	8.40	17.72	1.23
		GDP Growth	7.19	7.18	3.89	10.26	24.90	1.79
		Exchange Rate	51.29	48.41	41.35	64.15	14.89	7.64
<b>Indonesia</b>	36	ROAE_LEAD1	21.81	22.54	7.39	36.44	33.88	7.39
		CEO SAB	49.44	33.00	−1.00	160.00	97.38	48.15
		MD&A SAB	55.28	7.50	−86.00	360.00	182.63	100.95
		Total Assets	37,175.34	38,368.90	9492.49	68,826.61	51.07	18,986.62
		Assets Growth	8.01	8.38	−15.63	33.35	148.74	11.92
		NPL/GLoans	2.90	2.97	0.01	8.64	68.02	1.97
		Tier1Capital	14.58	14.53	10.11	18.40	14.52	2.12
		Loans/Assets	60.76	63.18	38.44	74.55	14.84	9.02
		Interest Spread	5.26	5.39	3.85	6.24	13.85	0.73
		GDP Growth	5.64	6.01	4.63	6.35	11.28	0.64
		Exchange Rate	10,243.69	9698.96	8770.43	13,389.41	14.13	7.81
<b>Korea</b>	9	ROAE_LEAD1	10.28	10.93	7.13	16.77	28.54	2.94
		CEO SAB	30.89	27.00	8.00	48.00	42.26	13.05
		MD&A SAB	272.00	284.00	3.00	566.00	72.98	198.51
		Total Assets	262,576.90	250,079.69	221,996.06	316,025.25	14.25	7424.60
		Assets Growth	3.69	4.08	−3.30	11.01	107.78	3.97
		NPL/GLoans	1.34	1.33	0.77	1.76	25.62	0.34
		Tier1Capital	9.43	8.56	8.21	11.40	15.07	1.42
		Loans/Assets	66.95	66.88	65.78	68.04	0.97	0.65
		Interest Spread	2.97	2.60	2.00	5.00	36.11	1.07
		GDP Growth	3.37	2.90	0.71	6.50	50.98	1.72
		Exchange Rate	0.28	0.28	0.27	0.30	3.06	0.01
<b>Malaysia</b>	45	ROAE_LEAD1	14.35	14.46	7.26	23.29	21.31	8.97
		CEO SAB	76.73	73.00	10.00	155.00	47.57	36.50
		MD&A SAB	330.82	213.00	2.00	1128.00	84.61	279.91
		Total Assets	131,320.86	105,154.00	23,156.10	489,782.86	87.28	114,615.41
		Assets Growth	7.49	8.29	−12.80	35.21	118.19	8.86
		NPL/GLoans	2.99	2.82	0.83	8.84	53.39	1.60
		Tier1Capital	9.46	8.90	5.11	14.47	25.54	2.42
		Loans/Assets	52.40	62.19	16.21	68.70	33.20	17.40
		Interest Spread	2.24	2.00	1.45	3.24	29.63	0.66
		GDP Growth	4.85	5.29	−2.53	9.43	63.84	3.09
		Exchange Rate	3.33	3.27	3.06	3.91	7.59	0.25
<b>Nepal</b>	9	ROAE_Lead1	21.07	22.13	5.18	26.49	20.76	4.37
		CEO SAB	15.22	11.00	7	47.00	105.20	16.01
		MD&A SAB	25.56	17.00	8.00	74.00	83.84	21.42
		Total Assets	19.99	7.78	494.33	818.68	17.47	108.29
		Assets Growth	35	9.30	−6.06	23.29	130.47	9.59
		NPL/GLoans	.79	2.89	1.74	.65	26.91	0.75
		Tier1Capital	9.14	9.03	8.68	9.64	3.89	0.36
		Loans/Assets	63.99	6.16	53.09	70.54	9.11	5.83
		Interest Spread	6.58	7.00	4.38	8.00	20.14	1.33
		GDP Growth	4.43	4.53	2.73	6.10	25.89	1.15
		Exchange Rate	13	77.57	66.42	102.41	15.77	12.95
<b>Pakistan</b>	54	ROAE_Lead1	16.00	17.69	−7.80	27.32	47.49	7.60
		CEO SAB	23.52	15.00	−94.00	120.00	175.78	41.34
		MD&A SAB	27.81	14.00	−54.00	275.00	204.05	56.76
		Total Assets	7404.19	6655.89	2605.97	16,324.48	46.45	3438.88
		Assets Growth	5.36	7.86	−22.32	22.67	213.15	11.43
		NPL/GLoans	9.62	8.64	2.68	18.47	39.48	3.80
		Tier1Capital	12.70	11.91	5.91	21.01	34.77	4.42
		Loans/Assets	48.92	49.81	32.69	72.65	19.34	9.46
		Interest Spread	6.18	5.81	4.30	.30	22.08	1.36
		GDP Growth	3.45	3.51	1.61	4.83	35.49	1.22
		Exchange Rate	87.03	86.34	60.74	102.77	15.94	13.87
<b>Philippines</b>	36	ROAE_Lead1	1.34	11.90	3.76	18.72	32.19	3.65
		CEO SAB	37.44	34.50	−10.00	98.00	69.98	26.20



		MD&A SAB	77.00	72.50	−35.00	269.00	84.63	65.16
		Total Assets	13,040.79	9076.78	4243.65	43,066.06	79.86	10,414.92
		Assets Growth	11.11	9.78	−2.23	42.09	76.53	8.50
		NPL/GLoans	5.09	4.97	1.21	16.06	64.77	3.30
		Tier1Capital	13.23	12.92	8.31	17.43	16.64	2.20
		Loans/Assets	49.34	49.85	26.82	62.98	19.16	9.45
		Interest Spread	4.19	4.26	2.52	5.83	21.47	0.90
		GDP Growth	5.45	6.22	1.15	7.63	36.37	1.98
		Exchange Rate	44.57	44.40	42.23	47.68	3.79	1.69
<b>Singapore</b>	27	ROAE_Lead1	11.57	11.59	6.72	15.76	16	1.85
		CEO SAB	109.44	105	15	181	33.36	36.51
		MD&A SAB	214.52	208	−1	568	64.3	137.93
		Total Assets	2,13,055.8	2,13,544.7	1,21,154.1	3,33,509.4	31.28	66,637.83
		Assets Growth	9.22	9.18	−10.03	22.26	83.98	7.74
		NPL/GLoans	1.37	1.33	0.61	2.91	38.12	0.52
		Tier1Capital	13.55	13.8	8.86	16.6	14.12	1.91
		Loans/Assets	54.64	54.06	41.68	65.62	13.06	7.13
		Interest Spread	5.12	5.18	4.8	5.24	2.81	0.14
		GDP Growth	5.04	3.67	−0.6	15.24	90.08	4.54
		Exchange Rate	1.35	1.36	1.25	1.51	6.86	0.09
<b>Sri Lanka</b>	54	ROAE_Lead1	16.16	16.12	2.75	41.02	36.19	5.85
		CEO SAB	55.35	45.5	0	145	70.09	38.8
		MD&A SAB	81.07	46	−30	396	129.44	104.94
		Total Assets	2072.02	1639.06	142.81	6123.62	74.18	1537.06
		Assets Growth	12.84	12.66	−0.92	32.93	58.96	7.57
		NPL/GLoans	5.23	4.44	1.2	21.89	73.33	3.83
		Tier1Capital	13.87	12.91	7.23	26.7	33.02	4.58
		Loans/Assets	66.19	67.28	51.4	78.3	8.84	5.85
		Interest Spread	3.31	3.11	0.15	7.54	73.41	2.43
		GDP Growth	6.1	5.95	3.4	9.14	32.98	2.01
		Exchange Rate	120.07	114.94	108.33	135.86	8.36	10.04
<b>Thailand</b>	54	ROAE_Lead1	12.72	12.31	1.15	21.73	33.14	4.22
		CEO SAB	11.81	8.5	−7	48	122.12	14.43
		MD&A SAB	22.8	0	−82	251	322.22	73.45
		Total Assets	39,542.5	33,997.39	2586.96	84,614.34	67.76	26,794.25
		Assets Growth	7.91	5.83	−13.05	52.85	151.74	12
		NPL/GLoans	5.48	4.5	2.3	16.64	61.28	3.36
		Tier1Capital	11.8	11.25	7.5	15.83	17.91	2.11
		Loans/Assets	69.94	70.07	55.33	79.14	6.65	4.65
		Interest Spread	4.62	4.64	4.08	5.15	7.6	0.35
		GDP Growth	3.15	2.7	−0.74	7.51	88.83	2.8
		Exchange Rate	32.54	32.48	30.49	34.52	4.72	1.54
<b>Turkey</b>	54	ROAE_LEAD1	14.46	14.95	−2.92	26.53	39.94	5.77
		CEO SAB	42.24	33	4	134	73.43	31.02
		MD&A SAB	−9.85	−6	−130	82	−407.5	40.15
		Total Assets	56,503.58	67,250.35	2253.28	1,17,689.8	69.99	39,547.92
		Assets Growth	8.32	5.1	−8.62	45.14	148.34	12.34
		NPL/GLoans	3.53	3.4	1.17	7.31	40.37	1.43
		Tier1Capital	13.5	12.9	9.52	19.89	17.56	2.37
		Loans/Assets	62.6	63.12	43.57	78.95	13.9	8.7
		Interest Spread	6.42	6.2	5	8.3	15.29	0.98
		GDP Growth	3.53	3.97	−4.83	9.16	113.46	4
		Exchange Rate	1.77	1.67	1.3	2.72	24.56	0.43

## Appendix C. Principal Component Analysis for Clustering

Variables	Standardized Scoring Coefficients			
	Factor 1	Factor 2	Factor 3	Factor 4
GDP Growth	0.29403	0.01737	−0.04448	0.00861
Assets Growth	0.22054	−0.08433	0.00653	−0.08169
Total Assets	0.17464	−0.00418	−0.13046	−0.04752
ROAE	0.22968	−0.14158	0.31266	−0.13597
NPL/Gross Loans	−0.28315	−0.10349	−0.08393	−0.09274
CEO SAB	−0.02437	0.33124	0.04707	−0.01532
MD&A SAB	0.00450	0.36246	−0.07034	0.08208
Exchange Rate	0.01936	0.00823	0.34408	0.02595
Tier1Capital	−0.07521	0.00212	0.19086	0.01110
Loans/Assets	−0.04309	−0.03043	0.05416	0.30824
Interest Rate Spread	−0.03976	−0.16266	0.06292	−0.29118

## Appendix D. Fixed Effects and Random Effects Models

Variable	Model 1	Model 2
	Fixed Effects	Random Effects
ROAE	0.233 (0.046)	0.4315 *** (0.0409)
CEO SAB	−0.0054 (0.0479)	−0.0443 (0.0412)
MD&A SAB	0.0546 (0.0502)	−0.0021 (0.0425)
Assets (log)	−0.9233 *** (0.2041)	−0.0918 (0.0588)
Assets Growth (%)	0.0678 ** (0.0341)	0.0650 * (0.0337)
NPL/Gross Loans (%)	−0.1420 ** (0.0623)	−0.0670 (0.0484)
Tier1Capital (%)	−0.0858 ** (0.0529)	−0.1405 *** (0.0415)
Loans/Assets (%)	−0.0937 (0.069)	−0.1198 * (0.0472)
Interest Rate Spread (%)	0.0244 (0.0846)	0.0497 (0.0494)
GDP Growth (%)	−0.0454 (0.036)	−0.0059 (0.0352)
Exchange Rate (%)	−0.4464 ** (0.2003)	0.1550 *** (0.0551)
<b>Fit Statistics</b>		
Cross Sections	58	58
Time Series Length	9	9
MSE	0.3939	0.4231
Root MSE	0.6276	0.6504
R-Square	0.6575	0.2974
<b>Diagnostics Tests</b>		
F-Test for Fixed Effects ( $p > F$ )	3.32 (0.001)	
Hausman Test for Random Effects ( $p > m$ )		114.66 (0.001)
Variance Component for Cross Sections		0.11634
Variance Component for Error		0.393914

\*\*\*/\*\*/\* means significance at 0.01, 0.05, and 0.10 level, respectively. Return on average equity (ROAE) is the dependent variable. Standard error in Parenthesis.

## References

- Abdallah, Wissam, Marc Goergen, and Noel O’Sullivan. 2015. Endogeneity: How Failure to Correct for It Can Cause Wrong Inferences and Some Remedies. *British Journal of Management* 26: 791–804. doi:10.1111/1467-8551.12113.
- Adam, Tim R., Chitru S. Fernando, and Evgenia Golubeva. 2015. Managerial Overconfidence and Corporate Risk Management. *Journal of Banking and Finance* 60: 195–208. doi:10.1016/j.jbankfin.2015.07.013.
- Aerts, Walter. 1994. On the Use of Accounting Logic as an Explanatory Category in Narrative Accounting Disclosures. *Accounting, Organizations and Society* 19: 337–53. doi:10.1016/0361-3682(94)90001-9.
- Aerts, Walter. 2001. Inertia in the Attributional Content of Annual Accounting Narratives. *European Accounting Review* 10: 3–32. doi:10.1080/09638180122562.
- Aerts, Walter. 2005. Picking up the Pieces: Impression Management in the Retrospective Attributional Farming of Accounting Outcomes. *Accounting, Organizations and Society* 30: 493–517.
- Ahamed, M. Mostak. 2017. Asset Quality, Non-Interest Income, and Bank Profitability: Evidence from Indian Banks. *Economic Modelling* 63: 1–14. doi:10.1016/j.econmod.2017.01.016.
- Akhisar, İlyas, K. Batu Tunay, and Necla Tunay. 2015. The Effects of Innovations on Bank Performance: The Case of Electronic Banking Services. *Procedia—Social and Behavioral Sciences* 195: 369–75. doi:10.1016/j.sbspro.2015.06.336.
- Allen, Franklin, and Elena Carletti. 2012. *The Roles of Banks in Financial Systems: The Oxford Handbook of Banking*. Oxford: Oxford University Press.
- Altman, Edward I. 1968. The Prediction of Corporate Bankruptcy: A Discriminant Analysis. *The Journal of Finance* 23: 193–94.
- Amernic, Joel, and Russell Craig. 2006. *CEO Speak: The Language of Corporate Leadership*. Montreal: McGill-Queen’s University Press.
- Arellano, Manuel, and Olympia Bover. 1995. Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics* 68: 29–51.
- Asay, H. Scott, Robert Libby, and Kristina M. Rennekamp. 2018. Do Features That Associate Managers with a Message Magnify Investors’ Reactions to Narrative Disclosures? *Accounting, Organizations and Society* 68–69: 1–14. doi:10.1016/j.aos.2018.02.003.
- Athanasoglou, Panayiotis P., Sophocles N. Brissimis, and Matthaios D. Delis. 2008. Bank-Specific, Industry-Specific and Macroeconomic Determinants of Bank Profitability. *Journal of International Financial Markets, Institutions and Money* 18: 121–36. doi:10.1016/j.intfin.2006.07.001.
- Baginski, Stephen P., John M. Hassell, and William A. Hillison. 2000. Voluntary Causal Disclosures: Tendencies and Capital Market Reaction. *Review of Quantitative Finance and Accounting* 15: 371–89. doi:10.1023/A:1012002608615.
- Baum, Christopher F., Mark E. Schaffer, and Steven Stillman. 2003. Instrumental Variables and GMM: Estimation and Testing. *The Stata Journal* 3: 1–31.
- Beattie, Vivien. 2014. Accounting Narratives and the Narrative Turn in Accounting Research: Issues, Theory, Methodology, Methods and a Research Framework. *The British Accounting Review* 46: 111–34. doi:10.1016/j.bar.2014.05.001.
- Beaver, William H. 1966. Financial Ratios as Predictors of Failure. *Journal of Accounting Research* 4: 71–111. doi:10.2307/2490173.
- Beccalli, Elena, Mario Anolli, and Giuliana Borello. 2015. Are European Banks Too Big? Evidence on Economies of Scale. *Journal of Banking & Finance* 58: 232–46. doi:10.1016/j.jbankfin.2015.04.014.
- Billett, Matthew T, and Yiming Qian. 2008. Are Overconfident CEOs Born or Made? Evidence of Self- Attribution Bias from Frequent Acquirers. *Management Science* 54: 1037–51. doi:10.1287/mnsc.1070.0830.
- Bloomfield, Robert. 2008. Discussion of ‘Annual Report Readability, Current Earnings, and Earnings Persistence.’ *Journal of Accounting and Economics* 45: 248–52. doi:10.1016/j.jacceco.2008.04.002.
- Blundell, Richard, and Stephen Bond. 1998. Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics* 87: 115–43. doi:10.1016/S0304-4076(98)00009-8.
- Board, John, Charles Sutcliffe, and William T Ziemba. 2003. Applying Operations Research Techniques to Financial Markets. *Interfaces* 33: 12–24.

- Bougatef, Khemaies. 2017. Determinants of Bank Profitability in Tunisia: Does Corruption Matter? *Journal of Money Laundering Control* 20: 70–78. doi:10.1108/jmlc-10-2015-0044.
- Canbas, Serpil, Altan Cabuk, and Suleyman Bilgin Kilic. 2005. Prediction of Commercial Bank Failure via Multivariate Statistical Analysis of Financial Structures: The Turkish Case. *European Journal of Operational Research* 166: 528–46. doi:10.1016/j.ejor.2004.03.023.
- Chen, Wei, Jun Han, and Hun Tong Tan. 2016. Investor Reactions to Management Earnings Guidance Attributions: The Effects of News Valence, Attribution Locus, and Outcome Controllability. *Accounting, Organizations and Society* 55: 83–95. doi:10.1016/j.aos.2016.10.002.
- Clarke, Robert L. 1988. *Bank Failure: An Evaluation of the Factors Contributing to the Failure of National Banks*. Washington, DC: Comptroller of the Currency, Administrator of National Banks.
- Clatworthy, Mark, and Michael John Jones. 2003. Financial Reporting of Good and Bad News: Evidence from Accounting Narratives. *Accounting and Business Research* 33: 171–85. doi:10.1080/00014788.2003.9729645.
- Craig, Russell, Tony Mortensen, and Shefali Iyer. 2013. Exploring Top Management Language for Signals of Possible Deception: The Words of Satyam's Chair Ramalinga Raju. *Journal of Business Ethics* 113: 333–47. doi:10.1007/s10551-012-1307-5.
- Daniel, Kent D, David Hirshleifer, and Avanidhar Subrahmanyam. 2001. Overconfidence, Arbitrage, and Equilibrium Asset Pricing. *The Journal of Finance* 56: 921–65.
- Davis, Angela K., Jeremy M. Piger, and Lisa M. Sedor. 2012. Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language. *Contemporary Accounting Research* 29: 845–68. doi:10.1111/j.1911-3846.2011.01130.x.
- De Haan, Jakob, and Razvan Vlahu. 2016. Corporate Governance of Banks: A Survey. *Journal of Economic Surveys* 30: 228–77. doi:10.1111/joes.12101.
- Demirgüç-Kunt, Ash, and Harry Huizinga. 1999. Determinants of Commercial Bank Interest Margins and Profitability: Some International Evidence. *The World Bank Economic Review* 13: 379–408. doi:10.1093/wber/13.2.379.
- Demirgüç-kunt, Asli, Erik Feyen, and Ross Levine. 2012. *The Evolving Importance of Banks and Securities Markets*. National Bureau of Economic Research Working Paper Series 18004. Cambridge: National Bureau of Economic Research.
- Djalilov, Khurshid, and Jenifer Piesse. 2016. Determinants of Bank Profitability in Transition Countries: What Matters Most? *Research in International Business and Finance* 38: 69–82. doi:10.1016/j.ribaf.2016.03.015.
- Edey, Malcolm. 2009. The Global Financial Crisis and Its Effects. *Economic Papers* 28: 186–95. doi:10.1111/j.1759-3441.2009.00032.x.
- European Central Bank. 2010. *Beyond ROE — How to Measure Bank Performance*. Frankfurt: European Central Bank.
- Fethi, Duygun Meryem, and Fotios Pasiouras. 2010. Assessing Bank Efficiency and Performance with Operational Research and Artificial Intelligence Techniques: A Survey. *European Journal of Operational Research* 204: 189–98. doi:10.1016/j.ejor.2009.08.003.
- Fink, Gerhard, Peter Haiss, and Mantler Hans Christian. 2005. *The Finance-Growth Nexus: Market Economies vs. Transition Countries*. EI Working Papers/Europainstitut, 64. Vienna: Europainstitut, WU Vienna University of Economics and Business.
- Gaganis, Chrysovalantis, Fotios Pasiouras, and Constantin Zopounidis. 2006. A Multicriteria Decision Framework for Measuring Banks' Soundness around the World. *Journal of Multi-Criteria Decision Analysis* 14: 103–11. doi:10.1002/mcda.405.
- Gandhi, Priyank, Tim Loughran, and Bill McDonald. 2019. Using Annual Report Sentiment as a Proxy for Financial Distress in U.S. Banks. *Journal of Behavioral Finance* 7560: 1–13. doi:10.2139/ssrn.2905225.
- Gervais, Simon, and Terrance Odean. 2001. Learning to Be Overconfident. *The Review of Financial Studies* 14: 1–27. doi:10.1093/rfs/14.1.1.
- Goddard, John, Hong Liu, and Phil Molyneux. 2013. Do Bank Profits Converge? *European Financial Management* 19: 345–65. doi:10.1111/j.1468-036X.2010.00578.x.
- Greene, William H. 1996. *Econometric Analysis*. Internatio. New York: Pearson Education, Inc.
- Gutierrez, Roberto G., and A. Rachid El-khattabi. 2017. *Using a Dynamic Panel Estimator to Model Change in Panel Data*. SAS Paper Series. Cary: SAS Institute.
- Hooghiemstra, Roggy. 2001. Cultural Differences in Self-Serving Behaviour in Accounting Narratives. Paper present at the APIRA Conference 2001, Adelaide, Australia, July 15–17.

- Horváthová, Jarmila, and Martina Mokrišová. 2018. Risk of Bankruptcy, Its Determinants and Models. *Risks* 6: 117. doi:10.3390/risks6040117.
- Hsu, Chung-Chian, Chin-Long Chen, and Yu-Wei Su. 2007. Hierarchical Clustering of Mixed Data Based on Distance Hierarchy. *Information Sciences* 177: 4474–92. doi:10.1016/j.ins.2007.05.003.
- Jasevičienė, Filomena, Bronius Povilaitis, and Simona Vidzbelytė. 2013. Commercial Banks Performance 2008–2012. *Business, Management and Education* 11: 189–208. doi:10.3846/bme.2013.11.
- Keusch, Thomas, Laury H. H. Bollen, and Harold F. D. Hassink. 2012. Self-Serving Bias in Annual Report Narratives: An Empirical Analysis of the Impact of Economic Crises. *European Accounting Review* 21: 1–26. doi:10.1080/09638180.2011.641729.
- Laeven, Luc, and Fabian Valencia. 2010. *Resolution of Banking Crises: The Good, the Bad, and the Ugly*. IMF Working Paper WP/10/146. Washington, DC: International Monetary Fund.
- Leary, Mark R., and Robin M Kowalski. 1990. Impression Management: A Literature Review and Two-Component Model. *Psychological Bulletin* 107: 34–47.
- Lehmberg, Derek, and Chanchai Tangpong. 2018. Do Top Management Performance Attribution Patterns Matter to Subsequent Organizational Outcomes? A Two-Country Study of Attribution in Economic Crisis. *Journal of Management & Organization* 1–20. doi:10.1017/jmo.2018.32.
- Li, Feng. 2008. Annual Report Readability, Current Earnings, and Earnings Persistence. *Journal of Accounting and Economics* 45: 221–47. doi:10.1016/j.jacceco.2008.02.003.
- Li, Feng. 2010a. Managers' Self-Serving Attribution Bias and Corporate Financial Policies. SSRN. doi:10.2139/ssrn.1639005.
- Li, Feng. 2010b. Textual Analysis of Corporate Disclosure: A Survey of Literature. *Journal of Accounting Literature* 29: 143–65.
- Libby, Robert, and Kristina Rennekamp. 2012. Self-Serving Attribution Bias, Overconfidence, and the Issuance of Management Forecasts. *Journal of Accounting Research* 50: 197–231. doi:10.1111/j.1475-679X.2011.00430.x.
- Lin, Shih Wei, Yeou Ren Shiue, Shih Chi Chen, and Hui Miao Cheng. 2009. Applying Enhanced Data Mining Approaches in Predicting Bank Performance: A Case of Taiwanese Commercial Banks. *Expert Systems with Applications* 36: 11543–51. doi:10.1016/j.eswa.2009.03.029.
- Lipe, Robert. 1990. The Relation Between Stock Returns and Accounting Earnings Given Alternative Formation. *The Accounting Review* 65: 49–71.
- Loewenstein, Yaniv, Elon Portugaly, Menachem Fromer, and Michal Linial. 2008. Efficient Algorithms for Accurate Hierarchical Clustering of Huge Datasets: Tackling the Entire Protein Space. *Bioinformatics* 24: 41–49. doi:10.1093/bioinformatics/btn174.
- Malmendier, Ulrike, and Geoffrey Tate. 2008. Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction. *Journal of Financial Economics* 89: 20–43. doi:10.1016/j.jfineco.2007.07.002.
- Mathuva, D. M. 2009. Capital Adequacy, Cost Income Ratio and the Performance of Commercial Banks: The Kenyan Scenario. *The International Journal of Applied Economics & Finance* 3: 35–47.
- Medura, Jeff. 2006. *International Financial Management*, 6th ed. Mason: Thomson South-Western.
- Merkel-davies, Doris M., Niamh M. Brennan, and Petros Vourvachis. 2014. Text Analysis Methodologies in Corporate Narrative Reporting Research. Paper presented at the 23rd CSEAR International Colloquium, St Andrews, UK, August 26–29.
- Moosa, Imad A., and Vikash Ramiah. 2017. Overconfidence and Self-Serving Bias. In *The Financial Consequences of Behavioural Biases: An Analysis of Bias in Corporate Finance and Financial Planning*. Basingstoke: Palgrave Macmillan, pp. 45–69. doi:10.1007/978-3-319-69389-7\_3.
- Mushinada, Venkata Narasimha Chary, and Venkata Subrahmanya Sarma Veluri. 2018. Investors Overconfidence Behaviour at Bombay Stock Exchange. *International Journal of Managerial Finance* 14: 613–32. doi:10.1108/IJMF-05-2017-0093.
- Myatt, Glenn J., and Wayne P. Johnson. 2014. *Making Sense of Data: A Practical Guide to Exploratory Data Analysis and Data Mining*, 2nd ed. West Hoboken: John Wiley & Sons, Inc.
- Pagratis, Spyros, Eleni Karakatsani, and Helen Louri. 2014. *Bank Leverage and Return on Equity Targeting: Intrinsic Procyclicality of Short-Term Choices*. Athens: Bank of Greece. ISSN 1109-6691.
- Panta, Bishop. 2018. Non-Performing Loans & Bank Profitability: Study of Joint Venture Banks in Nepal. *International Journal of Sciences: Basic and Applied Research* 42: 151–65.
- Petrella, Giovanni, and Andrea Resti. 2013. Supervisors as Information Producers: Do Stress Tests Reduce Bank Opaqueness? *Journal of Banking & Finance* 37: 5406–20. doi:10.1016/j.jbankfin.2013.01.005.

- Petria, Nicolae, Bogdan Capraru, and Iulian Ihnatov. 2015. Determinants of Banks' Profitability: Evidence from EU 27 Banking Systems. *Procedia Economics and Finance* 20: 518–24. doi:10.1016/s2212-5671(15)00104-5.
- Ravi Kumar, P., and V. Ravi. 2007. Bankruptcy Prediction in Banks and Firms via Statistical and Intelligent Techniques—A Review. *European Journal of Operational Research* 180: 1–28. doi:10.1016/j.ejor.2006.08.043.
- Roberts, Michael R., and Toni M. Whited. 2012. Endogeneity in Empirical Corporate Finance. In *Handbook of the Economics of Finance*. Amsterdam: Elsevier, pp. 493–572. doi:10.1016/B978-0-44-453594-8.00007-0.
- Seamans, Michael A. 2013. When Gauging Bank Capital Adequacy, Simplicity Beats Complexity. *Economic Letter* 8: 1–4.
- Shehzad, C. T., J. De Haan, and B. Scholtens. 2013. The Relationship Between Size, Growth and Profitability of Commercial Banks. *Applied Economics* 45: 1751–65. doi:10.1080/00036846.2011.637896.
- Smith, Malcolm, and Richard J. Taffler. 2000. The Chairman's Statement—A Content Analysis of Discretionary Narrative Disclosures. *Accounting, Auditing & Accountability Journal* 13: 624–47. doi:10.1108/09513570010353738.
- Sparck Jones, Karen. 1972. A Statistical Interpretation of Term Specificity and Its Application in Retrieval. *Journal of Documentation* 28: 11–21.
- Stovrag, Arijan. 2017. *Capital Requirements and Bank Failure: A Comparison between the Large Swedish Banks and Niche Banks*. Jönköping: Jönköping University.
- Terraza, Virginie. 2015. The Effect of Bank Size on Risk Ratios: Implications of Banks' Performance. *Procedia Economics and Finance* 30: 903–9. doi:10.1016/s2212-5671(15)01340-4.
- Tetlock, Paul C. 2007. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *Journal of Finance* 62: 1139–68. doi:10.1111/j.1540-6261.2007.01232.x.
- Trujillo-Ponce, Antonio. 2013. What Determines the Profitability of Banks? Evidence from Spain. *Accounting and Finance* 53: 561–86. doi:10.1111/j.1467-629X.2011.00466.x.
- White, Halbert. 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *The Economic Journal* 48: 817–38. doi:10.2307/2223855.
- Wu, Junjie, Hui Xiong, and Jian Chen. 2009. Towards Understanding Hierarchical Clustering: A Data Distribution Perspective. *Neurocomputing* 72: 2319–30. doi:10.1016/j.neucom.2008.12.011.
- Yao, Hongxing, Muhammad Haris, and Gulzara Tariq. 2018. Profitability Determinants of Financial Institutions: Evidence from Banks in Pakistan. *International Journal of Financial Studies* 6: 53. doi:10.3390/ijfs6020053.
- You, Haifeng, and Xiao Jun Zhang. 2009. Financial Reporting Complexity and Investor Underreaction to 10-k Information. *Review of Accounting Studies* 14: 559–86. doi:10.1007/s11142-008-9083-2.
- Zheng, Changjun, Mohammed Rahman, Munni Begum, and Badar Ashraf. 2017. Capital Regulation, the Cost of Financial Intermediation and Bank Profitability: Evidence from Bangladesh. *Journal of Risk and Financial Management* 10: 9. doi:10.3390/jrfm10020009.

