JOURNAL OF APPLIED ECONOMICS AND BUSINESS

VOL. 6, ISSUE 4 – DECEMBER, 2018

Education and Novel Technology Research Association

Journal of Applied Economics and Business

VOL. 6, ISSUE 4 – DECEMBER, 2018

The Journal of Applied Economics and Business (JAEB – ISSN: 1857-8721) is an international peer-reviewed, open-access academic journal that publishes original research articles. It provides a forum for knowledge dissemination on broad spectrum of issues related to applied economics and business. The journal pays particular attention on contributions of high-quality and empirically oriented manuscripts supported by various quantitative and qualitative research methodologies. Among theoretical and applicative contributions, it favors those relevant to a broad international audience. Purely descriptive manuscripts, which do not contribute to journal's aims and objectives are not considered suitable.

JAEB provides a space for academics, researchers and professionals to share latest ideas. It fosters exchange of attitudes and approaches towards range of important economic and business topics. Articles published in the journal are clearly relevant to applied economics and business theory and practice and identify both a compelling practical issue and a strong theoretical framework for addressing it.

The journal provides immediate open-access to its content on the principle that makes research freely available to public thus supporting global exchange of knowledge.

JAEB is abstracted and indexed in: DOAJ, EZB, ZDB, Open J-Gate, Google Scholar, JournalITOCs, New Jour and UlrichsWeb.

Publisher

Education and Novel Technology Research Association

Web: www.aebjournal.org

E-mail: editorial@aebjournal.org

support@aebjournal.org

publisher@aebjournal.org

Editor-in-Chief

Editorial board

Content

Journal of Applied Economics and Business

OPEN ACCESS ARTICLE

DOES BOARD SIZE SPUR BANK PERFORMANCE IN NIGERIA? EVIDENCE FROM PANEL DATA ANALYSIS

Dosunmu Adebukola Ibitamuno*, Okey Onuchuku*, Alwell Nteegah*

*Institute of International Trade and Development, University of Port Harcourt

Abstract

This paper examines how board size affects the performance of Deposit Money Banks (DMBs) in Nigeria. To achieve the objectives of the study, data on return on equity (ROE) board size, executive and non-executive board members and interest rate margin (lending rate less deposit rate) of 15 DMBs were sourced from the NDIC and CBN records and analyzed using panel approach. The result of the analysis revealed that board size has negative implication on the performance of banks in Nigeria. This implies that banks are likely to perform better with reduced board size. Based on this finding, the paper concludes that, there is no significant relationship between board size and the performance of banks in Nigeria. It also suggests the reduction in board size and banks adherence to prudential guideline by the regulatory agencies as ways of improving their performance in Nigeria.

Key words: *Return on equity; Board size; Interest rate margin.*

INTRODUCTION

Corporate governance plays a critical role in the wellbeing of a firm because its determine the success or failure of an organization (Ogbechie, 2006: 6). Corporate governance therefore refers to the processes and structures by which the business and affairs of institutions are directed and managed, in order to improve long term, share holders' value by enhancing corporate performance and accountability, while taking into account the interest of other stakeholders (Jenkinson & Mayer, 1992). Corporate governance is therefore, about building credibility, ensuring transparency and accountability as well as maintaining an effective channel of information disclosure that will foster good corporate performance. The critical role of corporate governance in the performance of a firm could be deciphered in the work of Akingunola et al., (2013) which argued that the bitter experiences of Asian financial crisis of the 1990s underscore the importance of effective corporate governance procedures to the

survival of banks and the economy at large. The Asian financial crisis, they argued, demonstrated in no unmistakable terms that "even strong economies, lacking transparent control, responsible corporate boards and shareholder right can collapse quite quickly as investor's confidence collapse.

The banking industry in Nigeria in recent years has undergone major changes arising from the reforms from the monetary authorities which stimulates that banks increase their capital base (share) to a minimum level of twenty-five billion naira (₦25B), (Ogbeche, 2006: 1). This development led to merging among banks to meet up the monetary authority's guideline thereby reducing the number of banks from 89 banks to 25 banks as at 2006 till date. (Kama, 2006: 66). Consequently, merger and acquisition of banks increase the size of bank and also widen the span of control. This brought a serious challenge on corporate governance issue and some bank became illiquid due to their inability to manage challenges emanating from their internal and external environment. The global economic crisis and the decline in the value of investment collections of deposit money banks particularly in Nigeria are due to distorted credit management and this problem is associated with poor corporate governance.

One critical factor in corporate governance is board size. The agency theory argues that superior firm financial performance may be linked to smaller board size. Smaller boards are less likely to have difficulty in coordinating and communicating. Also, a smaller board is probably more effective at monitoring management's activities because it cannot be easily influenced by the CEO and thus smaller board size may cause better firm financial performance than larger board size due to the problem of span of control (Lipton & Lorsch, 1992; Jensen, 1993). The agency theory strongly emphasises the importance of smaller boards, whereas resource dependence approach is in favour of large boards. The resource dependence theory on the other hand, posited that boards with a large number of directors may be advantageous in reducing dependency on external resources because larger boards may provide greater opportunity for more environmental linkages than smaller boards (Pfeffer & Salancik, 1978; Goodstein et al., 1994). The Nigerian financial sector had recognised the critical role board size plays in the performance of banks and has periodically review its composition and size in order to enhance the efficiency of the banking sector. This study seeks to examine how board size affects the performance of banks in Nigeria. We continue our discussion by reviewing relevant literature that are related to the study, explain the methodology employed to achieve the objectives, provide the results, findings and finally, the concluding remarks and recommendations.

LITERATURE REVIEW

The agency theory states that better corporate governance should lead to higher stock prices or better long-term performance, because when managers are better supervised, agency costs are decreased (Albanese, et al., 1997). However, as Gompers, et al., (2003) suggest, the evidence of a positive association between corporate

governance and firm performance may be traced to the agency explanation. In connection with the relationship between corporate governance and firm performance, the most studied governance practices include board composition, size and shareholder activities.

Lipton and Lorsch (1992) and Jensen (1993) in their studies also confirmed that; limiting board size is believed to improve firm performance because the benefits by larger boards of increased monitoring are outweighed by the poorer communication and decision-making of larger groups. A large board is likely to be less effective in substantive discussion of major issues and to suffer from free-rider problems among directors in their supervision of management (Hermalin & Weisbach, 1999). Mak and Li (2001) conducted an empirical analysis of firms listed on the Stock Exchange of Singapore. They stated that the sign and significance of the relationship between board size and performance, is sensitive to the estimation method. They concluded that the board characteristics are endogenous and failing to take endogeneity into account may yield a significant relationship with performance, which in reality does not exist.

Mak and Kusnadi (2002) also asserted an inverse relationship between board size and firm value. Their observation is based on a comparative study done on the firms listed on Singapore Stock Exchange (SGX) and Kuala Lumpur Stock Exchange (KLSE). Board effect was found in both countries. They further supported Healey (2003) that large groups are less effective than small groups in decision-making. Dwedi and Jain (2002) conducted a study on 340 large, listed Indian firms for the period 1997-2001. This study found a weak positive relation between board size and performance of the firm.

Beiner, et al., (2003) conducted a study over companies listed on the Swiss Stock Exchange (SWX). Study did not find a significant relationship between board size and firm valuation, as measured by Tobin's Q. Authors suggested that Swiss firms, on average, choose their number of board members just optimally.

Mak and Yuanto (2003) echo the above findings in firms listed in Singapore and Malaysia when they found that firm valuation is highest when board has five directors. Bennedsen, et al., (2006) studied the relationship between board size and performance of 500 Danish firms. Their study also supported a negative relation between the two variables. Adams and Mehran (2002) accessed the relationship between banking firms' performance (represented by Tobin's Q) and board size and found a non-negative relationship between board size and Tobin's Q. They further argued that M and A activity and features of the bank holding company organizational form might make a larger board more desirable for these firms.

In Manas and Saravanan (2006) it was concluded that the absence of a relationship between board size and corporate governance exists in Indian banks. In Ghana, it has

been identified that small board sizes enhances the performance of MFIs (Coleman & Nicholas-Biekpe, 2006). While in a study conducted in Nigeria, Sanda et al., (2005) found that, firm performance is positively related with small size as opposed to large boards. In their study, Klein (1998), Booth and Deli (1999) and Anderson et al., (2004) tried to find out the relation between board size and ratio of debt to assets (book leverage). They presented a different result that firms with bigger boards have lower cost of debt. On contrary to the theory that larger boards are ineffective monitors, they stated that board plays an important advisory role that enables firms to gain access to low-cost debt. They observed that the board will be larger in firms with high leverage.

Klein (1998), Agrawal and Knoeber (1996), Adams and Ferreira (2003), and Adams and Mehran (2003) also tried to access the applicability of same board size for all classes of firms. Klein (1998) argued that the CEO's need for advice will increase with the complexity of the organization.

Baysinger and Butler (1985) found little evidence that corporate governance resolutions initiated by shareholders lead to better firm performance. Smith and Watts (1992), reported a positive performance effects for the Shareholder's activities of the California Public Employees' Retirement System. Huson et al., (2004) showed that financial institutions could be fairly effective in pushing target companies to take steps to comply with their corporate governance proposals. They also find that any short-term valuation effects resulting from activities are dependent on the specific type of governance issue targeted. Gillan (2006) find that shareholder proposals by individuals have small, positive announcement effects, while proposals by institutional investors have a small but significant negative effect on stock prices. Overall, the empirical literature on shareholder's activities in the United States seems to indicate that it has a negligible impact on corporate performance (Black et al., 2003).

In other studies, Frankel et al., (2002) showed a negative relationship between earnings and auditor's independence, but Ashbaugh et al., (2003) and Larcker and Richardson (2004) dispute their evidence arguing that the study dwelt more on intrinsic factors. Agrawal and Chadha (2005) find no relation between either audit committee independence nor the extent auditors provide non-audit services with the probability a firm restates its earnings.

Furthermore, several studies have examined the separation of CEO and chairman, positing that agency problems are higher when the same person holds both positions. Using a sample of 452 firms in the annual Forbes magazine rankings of the 500 largest U.S. public firms between 1984 and 1991, Yermack (1996) shows that firms are more valuable when the CEO and board chair positions are separate. Larcker (2004) also find out that CEO compensation is lower when the CEO and board chair positions are separate.

Most prior studies on corporate governance and performance make use of the market based performance measure and not accounting performance measures. In order to

cover the lapses in prior studies, this study will build on the studies by Brown and Caylor (2004), Sanda et al., (2005), Coleman and Nicholas-Biekpe (2006), to analyze the relationship between corporate governance and financial performance of banks in Nigeria. This study used the CBN code of best practice and also made use of the specific governance index as provided by the Institutional Shareholder Services and as adopted in Brown and Caylor (2004), to create a summary index of firm- specific governance i.e. "Gov- Score". This will be an improvement over the index as used in Gomper et al., (2003) (i.e. the GIM index), which focused only on anti- takeover measures.

The hallmark of banking is the observance of high degree of professionalism, transparency and accountability, which are essential for building strong public confidence. Due to the systemic distress witnessed in the nation's banking system and its unpleasant consequences on all stakeholders as a result of inadequacies in corporate governance of banks in recent times, series of initiatives had been taken by the nation's regulatory/supervisory authorities to encourage sound corporate governance in the system. Some of the initiatives included enhancing the legal framework; enhancing the surveillance activities of the financial system; strengthening the roles of internal and external auditors; developing of a code of best practices of corporate governance in the system; issuance of guidelines and circulars on matters such as pre-qualification for appointment to board and top management positions in banks, insider related credits, etc. While all the above-mentioned efforts are in the right direction, it is equally important to indicate some imperatives of good corporate governance for banks so as to ensure the safety and soundness of emerging bigger banks in the post-consolidation era with a view to enhancing public confidence in the nation's banking system.

METHODOLOGY

In order to achieve the objectives of this study, data on return on Return on Equity (ROE), board composition, board size and lending rate of 15 Deposit Money Banks in Nigeria were sourced from the records of the banks and the Central bank of Nigeria. The data was analyzed using the panel approach. The purposes of our analysis are: to examine the impact of board size on performance of the banking sector and to find out if the impact of board size on performance (ROE) varies across the banks in Nigeria or not.

This study employs a modified version of the econometric model of Miyajima et al., (2003) as adopted by Coleman and Nicholas-Biekpe (2006**).** These studies used the CBN code of best practice and also made use of the specific governance index as provided by the Institutional Shareholder Services. The CBN prudential guideline specifies the number of directors (executive and non-executive) that a given bank

should have. Deposit money banks are expected to comply with such guideline in order to enhance stability of the bank and ensure efficient management and performance. In this study, performance is measured by return on equity (ROE). The rationale for the use of this variable as a measure of performance is that banks in Nigeria are privately owned firms financed by individual/group of individuals whose interest is to maximize profit. Return of equity captures the proportion of profit earned that goes to the shareholders of the bank. The higher the proportion of profit to shareholders' fund, the more attractive a firm (bank) is to the shareholders and other potential investors.

Board size is critical for performance of a firm and the banking sector in particular. This is because the composition (size) of board is key in decision making and management of the business. Board size also has serious role in running cost of the bank while directly affect ROE/performance. Based on this illustration, the paper specifies a bank performance function thus:

$$ROE_{it} = f\left(BOS_{it},\ BOC_{it},\ BRN_{it},\ INTR_{it}\right) \quad \text{(I)}$$

In order to estimate the functional relationship between banks' performance and the components of board size using econometric technique, equation 1 could be expressed in mathematical form as follows:

$$ROE_{it} = \alpha_0 + \alpha_1 BOS_{it} + \alpha_2 BOC_{it} + \alpha_3 BRN_{it} + \alpha_4 INTR + U_{it} \quad \text{(II)}$$

Where; ROE_{it}= Return on Equity; BOS_{it} = Board size; BOC_{it}= Executive Board Composition;BRN_{it}= Non executive Board Composition; $INTR_{it}$= Bank charges proxy by interest rate margin and U_{it}= Error term.

RESULTS AND FINDINGS

TABLE 1. DESCRIPTIVE STATISTICS

Variable	Mean	Standard Deviation	Minimum	Maximum
Roe	0.56	0.91	0.02	12.34
Bos	13.20	3.28	6	22
Bed	4.57	1.98	1	11
Bnd	8.79	2.21	4	16
Intr	19.88	3.39	14.82	26.04

The descriptive statistics reported in Table 1 indicates that there exists a high deviation in the variables investigated in this study. The high degree of deviation in return of equity, board size, executive board composition, non-executive board size and interest rate margin show that the variables experienced very high level of fluctuation during the period under study. This wide fluctuation in the variables could be attributed to the continuous reforms that take place in the Nigeria's banking sector.

TABLE 2. PAIRWISE CORRELATION RESULT: ROE BOS BED BNDINTR, STAR(5)

Variable	Roe	Bos	Bed	Bnd	Intr
Roe	1.000				
Bos	-0.1528*	1.000			
Bed	-0.1020	0.6995*	1.000		
Bnd	-0.1255	0.7752*	0.2080*	1.000	
Intr	-0.1139	-0.0231	-0.0567	0.0214	1.000

The pairwise correlation result in Table 2 revealed a negative and weak correlation between board size, executive board composition, non-executive board composition, interest rate margin and return on equity. This implies that components of board size have very weak relationship with bank performance (ROE).

TABLE 3. PANEL RESULT

Random Effect (RE) Model				Fixed Effect (FE) Model		
Variable	Coefficient	T-statistic	Prob	Coefficient	T-statistic	Prob
Bos	-0.039	-0.63	0.531	-0.029	-0.43	0.665
Bed	-0.0038	-0.05	0.958	0.022	0.32	0.747
Bnd	-0.0053	-0.15	0.881	-0.023	-0.30	0.766
Intr	-0.032	-1.68	0.092	-0.037	-1.66	0.098
Cons	1.759	3.87	0.000	1.75	3.34	0.001
$R^2 = 0.037$				FE test F(14,191) = 0.83; F-Prob = 0.63; $R^2 = 0.032$		

The panel result reported in Table 3 indicates that board size is negatively and insignificantly related to return on equity both in the random and fixed effects models. This implies that an increase in board size (Bos) retarded bank performance proxy by return on equity. This result is in consonance with earlier studies by Lipton and Lorsch (1992) and Jensen (1993) which revealed that limiting board size improve firm performance because the benefits by larger boards of increased monitoring are outweighed by the poorer communication and decision-making of larger groups. Smaller board size tends to reduce cost and span of control which has direct bearing on the performance of a firm (bank).

Executive board size also has negative and insignificant relationship with return on equity in the random effect model but has positive impact on performance in the fixed effect model. This implies that reduction in the number on executive board members of banks increases their performance (return on equity). This result conforms to the work of Coleman and Nicholas-Biekpe, (2006) which shows that smaller executive relative to large members have direct impact on the performance of bank in Ghana. The fixed effect result implies that large executive board size is performance friendly. This result is in tandem with the works of Dwedi and Jain (2002) which found a weak positive relation between board size and performance of the firm in a study conducted on 340 large, listed Indian firms for the period 1997- 2001.

The negative and insignificant relationship between non-executive board member size and performance in both the random and fixed effects models also indicates that smaller number of non-executive board members improve the performance of deposit money banks in Nigeria. This result aligned with earlier study by Sanda et al., (2005) which found that, firm performance is positively related with small size as opposed to large boards. The Central bank and other financial regulatory agencies in Nigeria have in their policy reforms tried to stipulate smaller board members for the bank. This was aimed that curtailing cost and unhealthy rivalries that may arise from large board size.

Interest rate margin measured by the difference between lending rate and deposit rate also has negative relationship with bank performance. This implies that increase in bank charges reduced the performance of banks in Nigeria. From the theoretical view point higher interest rate is necessary for improve performance of bank since such charge has influence of the bottom line of banks. However, Trujillo-Ponce, (2010) found that macroeconomic and financial environment of low interest rates coupled with tense competition among banks could reduce the possibilities for banks to establish appropriate prices for their loans and deposits, thereby putting pressure on the cost of operation and negatively affecting banks' performance. It should be noted that Interest rate has serious implication on volume of liquidity banks give out as loan and the income earned from such loans.

The panel result further shows that the effect of board size on the performance of bank is the same across the banks in Nigeria. This implies that board size has similar impact on all the banks operating in Nigeria over the period of this study.

CONCLUDING REMARKS AND RECOMMENDATIONS

This paper examine how board size affects the performance of deposit money banks in Nigeria. The result of our investigation revealed that board size has negative implication on the performance of banks in Nigeria. This implies that banks are likely to perform better with reduced board size. This result is in tandem with earlier studies by Lipton and Lorsch (1992), Jensen (1993), Coleman and Nicholas-Biekpe, (2006) and Sanda et al., (2005). Hence the paper concludes that there is no significant relationship between board size and the performance of banks in Nigeria. Based on the finding, the paper suggests reduction in board size and banks adherence to prudential guideline by the regulatory agencies as ways of improving their performance in Nigeria.

REFERENCES

Adams, R. & D. Ferreira. (2003). A theory of friendly boards. *The Academy of Management Review*, 22(3), 609-611.

Adams, R. & H. Mehran. (2003). Is corporate governance different for bank holding companies?. *Federal Reserve Bank of New York, Economic Policy Review*, 9, 123-142.

Agrawal, A. & S. Chadha. (2005). Corporate governance and accounting scandals. *Journal of Law and Economics,* 10(4), 56-78.

Agrawal, A. & C. R. Knoeber. (1996). Firm performance and mechanism to control agency problems between managers and shareholders. *Journal of Financial and Quantitative Analysis,* 31, 377-397.

Albanese, R., M.T. Dacin & I. C. Harris. (1997). Agents as stewards. *The Academy of Management Review,* 22(3), 609-611.

Anderson, R., S. Mansi & D. Reeb. (2004). Board characteristics, accounting report integrity, and the cost of debt. *Journal of Accounting and Economics,* 37, 315-342.

Ashbaugh, H., R. Lafond & B. Mayhew. (2003). Do non-audit services compromise auditor independence? Further evidence. *Accounting Review,* 78, 611-639.

Baysinger, B. D. & H. N. Butler. (1985). Corporate governance and the board of directors: performance effects of changes in board composition. *Journal of Law, Economics, and Organization*, 1, 101-124.

Beiner, S., W. Drobetz., M. Schmid, & H. Zimmermann. (2003). An integrated framework of corporate governance and firm valuation evidence from Switzerland. University of Switzerland. *Finance Working Paper No. 34.*

Bennedsen, M., H. Kongsted & K. Nielsen. (2006). Board size effects in closely held corporations. *Working Papers 09-2004, Copenhagen Business School, Department of Economics.*

Booth, J. & D. Deli. (1999). On executives of financial institutions as outside directors. *Journal of Corporate Finance*, 5, 227-250.

Brown, L. D. & M. C. Caylor. (2004). Corporate governance and firm performance. Georgia State University *Working Paper 146-04.*

Coleman, A. & N. Nicholas-Biekpe. (2006). Does board and CEO matter for bank performance? A comparative analysis of Banks in Ghana. *Journal of Business Management,* 13, 46-59.

Diwedi, N. & A. Jain. (2002). Corporate governance and performance of Indian firms: the effect of board size and ownership. *Review of Economics and Statistics*, February, 104-107.

Frankel, R., Johnson, M. & K. Nelson. (2002). The relation between auditors' fees for non-audit services and earnings management. *Accounting Review,* 77, 71-95.

Gillan, S. L. (2006). Recent developments in corporate governance: An overview. *Journal of Corporate Finance,* 12, 381-402.

Gompers, P., Ishii, L. & A. Metick. (2003). Corporate governance and equity prices. *Quarterly Journal of Economics,* 118, 107-125.

Jenkinson, T. & C. Mayer. (1992). The assessment: corporate governance and corporate control, *Oxford Review of Economic Policy*, 8(3), 138-156.

Hermalin, B. E. & M. Weisbach, (1999). The effects of board composition and direct incentives on firm performance. *Financial Management,* 20, 101-112.

Huson, M., P. Malatesta & R. Parrino. (2004). Managerial succession and firm Performance. *Journal of Financial Economics,* 74, 237-265.

Jensen, M. (1993). The modern industrial revolution, exit, and the failure of internal control systems. *Journal of Finance,* 48, 831-880.

Kama, T. (2006). Event studies in economics and finance. *Journal of Economic Literature,* 32(3), 75-86.

Klein, E. (1998). Firm performance and board committee structure. *Journal of Law and Economics*, 41(2), 275- 293.

Larcker, D. & S. Richardson. (2004). Fees paid to audit firms, accrual choices, and corporate governance. *Journal of Accounting Research,* 42, 625-638.

Lipton, M. & J. W. Lorsch. (1992). A modest proposal for improved corporate governance. *The business lawyer*, 59-77.

Manas, M. & P. Saravanan. (2006). Does the board size really matter? An empirical investigation on the Indian banking sector. *Journal of Corporation Law,* 27(2), 231-243.

Mak, Y. T. & Y. Lin. (2001). Determinants of corporate ownership and Board structure: evidence from Singapore. *Journal of Corporate Finance,* 7, 235-256.

Mak, Y. T. & Y. Kusandi. (2002). Size really matters: further evidence on the negative relationship between board size and firm value. *Working Paper 575, May, National University of Singapore, NUS Business School.*

Mak, Y. T. & K. Yuanto. (2003). Board size really matters: Further evidence on the negative relationship between Board size and firm value. *Journal of Corporate Finance,* 5, 145-166.

Miyajima, H., Omi, Y. & N. Saito. (2003). Corporate governance and performance in twentieth century Japan. *Business and Economic History Vol. 1, pp. 1-26*

Ogbeche, G. (2006). Environmental dynamism, capital structure and performance: A theoretical integration and empirical test. *Strategic Management Journal,* 21, 1-15.

Pfeffer, J. & G. R. Salancik. (1978). *The external control of organisation: A resource dependency perspective.* New York: Harper and Roe.

Sanda, A. U., A. S . Mukaila & T. Garba. (2005). Corporate governance mechanisms and firm financial performance in Nigeria. *Final Report Presented to the Biannual Research Workshop of the AERC*, Nairobi, Kenya, 24-29.

Smith, Clifford W. & L. W. Ross. (1992). The investment opportunity set and corporate financing, dividend, and compensation policies. *Journal of Financial Economics*, 32, 263-292.

Yermack, D. (1996). Higher market valuation of companies with a small board of directors. *Journal of Financial Economics*, 40, 185-211.

Journal of Applied Economics and Business

OPEN ACCESS ARTICLE

CAN MACHINE LEARNING IMPROVE RECESSION PREDICTION ACCURACY?

Azhar Iqbal[1], Kyle Bowman[2]

[1]Wells Fargo Securities, LLC, Charlotte

[2]Wells Fargo Bank, N.A., Charlotte

Abstract

This paper proposes a framework to utilize machine learning and statistical data mining tools in the economics/financial world with the goal of more accurately predicting recessions. Decision makers have a vital interest in predicting future recessions in order to enact appropriate policy. Therefore, to help decision makers, we raise the question: Does machine learning and statistical data mining improve recession prediction accuracy? Our first method examined over 500,000 variables as potential predictor variables for recession forecasting. Furthermore, to obtain the final logit/probit model specification, we ran 30 million different models. The selected model was then utilized to generate recession probabilities. The second method is the random forest approach, a famous class of machine learning tools. The third approach we employ is known as gradient boosting, a technique that also belongs in the machine learning family. Moreover, we built an econometric model that utilizes the yield curve as a recession predictor and employ that model as a benchmark. To test a model's accuracy, we employ both in-sample and out-of-sample criteria. In our tests, the random forest approach outperforms all the other models (gradient boosting, statistical machine learning and the simple econometric model) in both the in-sample and out-of-sample situations. The gradient boosting model comes in second place, while the statistical data mining approach captures third. Furthermore, if we combine all four probabilities, then that method is still unable to beat the random forest's prediction accuracy. That is, the random forest approach, alone, is the best. Our analysis proposes that machine learning can improve recession prediction accuracy. Moreover, our models suggest a less than 5% chance of a recession during the next 12 months.

***Key Words**: Recession Prediction; Machine Learning; Statistical Data Mining; Logit/Probit; Yield Curve.*

INTRODUCTION

"Computers are useless. They can only give you answers." Pablo Picasso

One major challenge for decision making today is the abundance of information - as opposed to information scarcity - and how to efficiently utilize the available information to design effective policies. Our study provides a framework to utilize machine learning

and statistical data mining tools in economics/financial world. The proposed framework employs to predict recessions. Decision makers have a vital interest to predict future recessions accurately as different set of decisions are needed for a recession than those of an expansion phase of a business cycle. Therefore, to help decision makers, we raise a question whether machine learning and statistical data mining improve recession prediction accuracy?

Our first method to predict recession is the statistical machine learning (also known statistical data mining). The estimation method is logit/probit modeling approach. This method utilized over 500,000 variables as potential predictors. Furthermore, to obtain the final logit/probit model specification, we ran 30 million different models. Then the selected model is utilized to generate recession probabilities.

The second method is the random forest approach which is a famous class of machine learning tools. The random forest approach utilizes the same set of predictors which are utilized in the statistical machine learning method. The third approach we utilize is known as the gradient boosting and it is also belong to the machine learning family. We also build an econometric model which utilizes the yield curve as a predictor and that model is employed as a benchmark. Basically, our benchmark model rely on an economic/financial theory as the yield curve is a famous recession predictor. Other three approaches include hundreds of thousands of potential predictors and those methods do not utilize any prior economic/financial theories. Therefore, we raise question whether the machine learning tools beat (provide more accurate forecast) a simple econometric model (a model with only one predictor)?

To test a model's accuracy, we utilize both in-sample and out-of-sample criteria. The random forest (machine learning) approach outperform all other models which are the gradient boosting, the statistical machine learning and the simple econometric model in both the in-sample and out-of-sample criteria. The gradient boosting is at the second place and statistical machine learning capture the third position. Furthermore, if we combine all four probabilities then that method is unable to beat random forest's prediction accuracy. That is, the random forest approach is still the best. Therefore, our analysis suggests that machine learning can improve recession prediction accuracy. Our models suggest a less than 5% chance of a recession during the next 18 months.

To summing up our analysis, we would like to expand the above mentioned Picasso's quote by saying that it's up to an analyst to obtain either an accurate answer by utilizing computers (machine learning/big data) efficiently or end up having a useless answer by providing irrelevant input (more noises than signals) to the model. Therefore, the

accuracy/reliable answer may not depends on computers (machine learning/big data) but how one utilize those computers.

PREDICTING RECESSION IN THE BIG DATA AGE: SETTING THE STAGE

Accurately predicting recessions is crucial for decision makers who are tasked with designing bespoke policy responses. Every recession is unique in the sense that different recessions have varying drivers. For example, one of the major causes of the Great Recession was the housing sector, while the IT boom/bust was a major cause of the 2001 recession. Knowing what will cause the next recession is a trillion dollar question. However, finding the right set of predictor variables to forecast the next recession is challenging because of the changing nature of the economy. Likewise, including too many variables in a traditional econometric modeling approach creates issues, such as an over-fitting problem *(Please go to the end of the paper and see Note 1)*.

Machine learning tools, on the other hand, are capable of handling a very large set of variables while providing useful information to identify the target variable. Basically, in the machine learning approach, we are letting the data speak for themselves and predict recessions. The rationale is that recessions are the results of imbalances/shocks that must reveal themselves in certain sectors of the economy. By including information from various sectors of the economy, we can improve the prediction of those imbalances and corresponding recessions. One major challenge for today's modelers is the abundance of information, where noise in large data sets can prove distracting. This challenge is different than the traditional modeling process where too little information was the issue. In the following sections we provide a reliable framework to utilize machine learning tools and large dataset (big data) to generate accurate recession forecasts.

Statistical Machine Learning: Opening Doors, Finding Connections

Our first recession prediction method is statistical machine learning, which is sometimes referred to as statistical data mining. In statistical machine learning modeling, we can "train" machines to go through hundreds of thousands of potential predictors and select a handful of predictors (4 to 6 variables, for example). That is, machines will utilize some statistical criteria (forecast error for instance) to narrow down the large dataset of potential predictors to a more manageable variable-list. We asked machines to consider over 500,000 variables as potential predictors and return to us a combination of five variables that predict U.S. recessions accurately.

There are several major benefits of the statistical machine learning method over a traditional econometric model in which an analyst has a model with a set of predictors that are selected based on an economic/financial theory. First, economies evolve over time and so does the relationship between the variables of interest. Thereby, it would be

practical to re-evaluate existing relationships and, if needed, add/subtract variables to/from a model. Statistical data mining does not rely on any economic/financial theory but identifies relevant variables using statistical tools. Second, complex economic interactions between different sectors vary over time as well. Thus the question, "what will cause the next recession?" is a very difficult one to answer. Therefore, putting everything in the pot (statistical data mining) increases the chances of finding what is affecting the target variable (recession) in the recent periods.

Third, it is important to note that a combination of some factors may bring about a recession rather than weakness in a single sector. For example, a drop in equity prices (S&P 500 index) and house prices along with a rising unemployment rate may be more likely to pull the economy into a recession than, for example, weakness in the manufacturing sector alone. Statistical data mining would likely help an analyst explore deep and complex interactions between different sectors that are closely associated with the target variables.

An added benefit of using statistical data mining is that important connections between different sectors are often unknown to analysts. Statistical machine learning can help an analyst identify those obscure connections. A great illustration of such unknown connections between certain sectors of the economy is the financial crisis and the Great Recession. That is, the housing boom was initially thought to be a regional phenomenon that would not pose a serious risk to the national economy. The Federal Open Market Committee (FOMC) transcripts from this period show that at first the FOMC considered there to be isolated regional housing bubbles. Likewise, by 2006, the meeting transcripts show that Ben Bernanke, the Federal Reserve Board's Chairman at the time, discussed that falling home prices would not derail economic growth *(Please go to the end of the paper and see Note 2)*. Furthermore, the relationship between the housing market and financial sector was also underestimated and only appeared with the Lehman Brother's bankruptcy in September 2008. Statistical machine learning has the potential to uncover such complex connections by utilizing information across major sectors of the economy.

Therefore, the statistical machine learning can find such unknown connections by utilizing information from major sectors as well as from detailed sub-sector level variables. Basically, in statistical machine learning, we do not restraint a model to use only a handful variables instead we include as much as possible (essential all available) information in the modeling process which increases chances to find relevant information and hence boost accuracy.

Information Magic: How Does Statistical Machine Learning Work?

Here we outline our proposed framework to effectively utilize statistical machine learning to forecast recessions. It is important to note that here we are using recession prediction as a case study but our framework is flexible and would help an analyst to effectively utilize the statistical machine learning to obtain accurate forecasts for any variable of interest. The first step is to define the target variable (what we are forecasting?) which, in our case, is a recession. We utilize the National Bureau of Economic Research's (NBER) definition of recession dates to construct the dependent variable. The dependent variable is a dummy variable with a value of zero (the U.S. economy is not in a recession) and one (the U.S. economy is in a recession). The benefit of using a dummy variable as the target variable is that we can generate the probability of a recession for a certain period-ahead using predictor variables.

Before we look for predictors, we need to discuss the sample period of the study. We started our analysis from January 1972 (monthly dataset). There are some major reasons to pick 1972 as a starting year of the analysis. First, since our dependent variable is a dummy variable that includes recession (value equals one) and non-recession (value equal zero) periods, our sample period must include both recession and non-recession periods. There have been six recessions since 1972. Second, many variables go back to the early 1970s and, therefore, provide an opportunity to select a model's relevant predictors from a large dataset of potential predictors. As mentioned earlier, a large pool of potential predictors captures information from all major sectors of the economy, which provides an opportunity to detect obscure connections between different sectors, thereby improving forecast accuracy.

The final and most important reason for starting our analysis in 1972, is that it can provide enough observations in our modeling approaches to conduct both in-sample analysis and out-of-sample forecasting, helping us test the predictive power of all potential variables. That is, we utilize the 1972-1988 period for in-sample analysis and the 1989-2017 period is employed for out-of-sample forecasting purpose. What are in-sample analysis and out-of-sample forecasting? Why do we need to conduct in-sample and out-of-sample forecasting analysis?

When a model utilizes the complete available information to estimate a statistical association (or sometimes, statistical causality) between variables of interest, that process is known as an in-sample analysis. In machine learning, that process is called "trained" or "training period/sample." For example, we utilized the 1972-2017 period to reduce the potential pool of 500,000 predictors to a manageable size of predictors (we'll talk about the variable reduction process in the following section). Basically, we utilize the 1972-2017 period to examine which variables are statistically associated with recessions. The

out-of-sample process involves forecasting, and the model does not know (have information) about the actual outcome for the forecast-horizon at the time of forecasting. That is, we utilize the 1972-1988 period and ask the model to generate the probability of a recession during the next 12 months (forecast horizon is 12 months). The important point here is that the model does not *know* whether there is a recession during the next 12 months. Out-of-sample forecasting utilizes the available information to forecast a future period. Put simply, in- sample analysis utilizes the available information and provides a statistical relationship between the target variable and predictors for that sample period. Out-of-sample forecasting uses the discovered relationship between variables to predict the future values of the target variable *(Please go to the end of the paper and see Note 3)*.

Now we turn to the next question of why we need to conduct in-sample and out-of-sample analyses. The in-sample analysis is a very effective tool to reduce the large potential list of predictors (sometimes the list contains hundreds of thousands or millions of potential predictors) to a more manageable pool. There are a number of statistical tests available within the in-sample analysis, which helps analysts identify a handful of predictors from the larger pool.

The out-of-sample forecasting exercise, in our view, is the most important tool in selecting the final model and improving forecast accuracy. When we generate the probability of a recession in real time, we will not know whether there will be a recession in the next 12 months. This is essentially a simulated real time forecasting experimentation. There are two major benefits of the simulated real time out-of-sample forecasting experiment. First, a common issue with forecasting models selected with using only in-sample selection criteria is over-fitting. Typically, an over-fitted model performs well during the in-sample analysis but very badly during out-of-sample forecasting. A model selected based on the out-of-sample forecasting criterion would reduce the over-fitting problem and improve forecast accuracy significantly compared to a model that is selected using in-sample criteria. The second major benefit is that the simulated real time out-of-sample forecasting would help an analyst estimate a reliable potential risk to the forecast (such as an average forecast error).

TURNING COLORS INTO A PICTURE: SAMPLE PERIOD AND DATA REDUCTION STEPS

The starting year of our analysis is 1972, and we conduct an in-sample analysis using the 1972-1988 period and the out-of-sample simulation criterion utilizing the 1989-2017 era. According to the NBER, there are six recessions in the complete sample period of 1972-2017. Furthermore, those six recessions are evenly divided in the in-sample analysis

(three recessions in the 1972-1988 period) and in the out-of-sample forecasting period (three recessions in the 1989-2017 period). The three recessions of the 1989-2017 period contain different characteristics (different depth and duration, for example) such as the 2007-2009 recession, which is the deepest recession since the Great Depression and hence has been labeled the Great Recession. The 2001 recession, on the other hand, is one of the mildest recessions in the sample era while the 1990-1991 recession is widely considered a moderate (neither mild nor deep) recession.

The major benefit of this out-of-sample forecasting simulation is that we do not know whether the next recession will be mild, moderate or deep; historically, mild recessions are relatively difficult to predict. If a model can predict recessions of different depths in a simulation, then there is a decent probability that the model would repeat its accuracy in the future.

A Sea of Potential Predictors: The FRED Dataset

One major benefit of the advancement of the Internet is that large datasets are available in a ready-to-use format, often at no cost. One such dataset is available on the Federal Reserve Bank of St. Louis' website, commonly referred to as the FRED (Federal Reserve Economic Data) *(Please go to the end of the paper and see Note 4)*. There are more than 500,000 variables listed in FRED, collected from 86 different sources. For our analysis, we consider all the 500,000 variables as potential predictors and try to find reliable predictors from this dataset using statistical tools. As mentioned earlier, instead of picking a handful of predictors (a traditional modeling approach), we include everything in the pot to find useful predictors from over 500,000 variables (statistical data mining approach). By using all FRED data, one thing is certain, not all of the 500,000 variables are relevant to predicting recessions. Put differently, we are including lots of noises in the model in addition to useful signals. However, there are some major benefits, as discussed earlier, of using the entire FRED data. That is, we will be able to find some obscure/new connections between different sectors of the economy and those connections may improve recession prediction accuracy.

As discussed earlier, 1972 is the starting year of our analysis, but not all FRED data go back that far. Therefore, the pool of over 500,000 variables is easily reduced to 5,889 variables. Three very important points we want to stress here before we move forward and discuss data reduction steps. First, our propose framework (which is consists of in-sample and out-of-sample analysis) can effectively handle hundreds of thousands (millions) variables and observations. Furthermore, in the present case, due to the nature of the target variable (recessions are very less frequent as there are only 6 recessions in our sample period of over 45 years) the list of potential predictors reduce from over 500,000 to under 6,000. Second, if we move our analysis from 1972 to, let say, 1995 then

that may give an opportunity to analyze over 200,000 variables as potential predictors instead of under 6,000 variables in the present case. However, that analysis (1995 as starting year), will have a very serious issue and that issue would reduce recession prediction accuracy significantly. Why? There are only 2 recessions in the 1995-2017 period and that indicates we have to rely only on in-sample analysis and in-sample analysis is known for over-fitting problem. By the same token, 1990 as a starting years would add one more recession in the analysis (3 recessions in the 1990-2017 period) but again not enough observations/time span to conduct out-of-sample simulation. Once more, an analyst have to rely on in-sample analysis and potentially suffer with the over-fitting problem *(Please go to the end of the paper and see Note 5)*.

The third and final point we want to stress is that we must employ an appropriate simulated real time out-of-sample analysis to finalize a model. Sometimes, as in the present case, an analyst may face a tradeoff between longer time span vs. a large pool of potential predictors. That is, if utilize longer time period (with starting year of 1972) then potential pool reduce from over 500,000 variables to 5889 variables. On the other hand, a 200,000 plus list of potential predictors is attached with a short time span (from 1995 and that period contains only 2 recessions). Our analysis picks a longer time span (over 200,000 potential predictors list) with a decent size of potential predictors (5889 variables as potential predictors). A shorter time span, let's say 1995 as a starting year, will have only 2 recessions and that means our target variable will have lots of zeros (dummy variable with zero for no-recession and one for recession) and very few ones which naturally create a bias toward zeros. Therefore model will tend to predict very low probability of a recession and increases chances of a false negative scenarios (very high likelihood of missing a future recession).

Essentially, we provide information to a model through the target variable (recession and no-recession in the present case) and via predictors. If we have large pool of potential predictors then we are providing an opportunity to include information for predictors. However, if that model uses a shorter time span then we are not providing an appropriate information about the target variable and setting the model for a failure. Therefore, we need to balance and provide enough information about the target variable as well as predictors. That is, we are including 6 recessions by starting our analysis from 1972 which also provide appropriate bases for the in-sample analysis (3 recessions) as well as for the out-of-sample simulation (3 recessions of different characteristics). Furthermore, our pool of potential predictors consist of 5889 variables and that offers an opportunity to include variables from every major sector (also for sub-sector level) of the economy and thereby

information to find useful predictors and explore some possible connections between different sectors.

A Statistical Spell of Variables Reduction

The list of 5,889 potential predictors is large enough to conduct in-sample analysis and out-of-sample simulation. To obtain a more manageable set of predictors, we employ several statistical methods and utilize the complete sample period of 1972-2017. First, we run the Granger causality test between our target variable and each of the 5,889 variables *(Please go to the end of the paper and see Note 6)*. The Granger (1969) test is a precise method to find which variables are statistically useful to predict the target variable *(Please go to the end of the paper and see Note 7)*. For the Granger causality test, we set a 5% level of significance and keep all variables that produce the p-value of the Chi-square test less than or equal to 0.05 *(Please go to the end of the paper and see Note 8)*.

The next methods to reduce the number of variables is called Weight of Evidence (WoE) and Information Value (IV) *(Please go to the end of the paper and see Note 9)*. Both the WoE and IV are very good tools to find reliable predictors, particularly if dealing with a binary, dependent target variable (zero for no recession and one for recession). The WoE provides evidence of predictive power of a variable relative to the target variable. The IV method, on the other hand, helps to rank variables according to their predictive power (the Y-variable has a higher predictive power than the X-variable to forecast recession, for example).

The Granger causality test, WoE and IV methods help us reduce the list of 5,889 variables to a set of 1,563 potential predictors. However, 1,563 variables as potential predictors are a lot for the in-sample analysis and out-of-sample simulation. Therefore, we utilize economic theory and intuition to further narrow down the list of 1,563 variables. That is, we manually inspect these 1,563 variables and then categorize them to represent major sectors of the economy. For example, consider the category "current population survey." A few potential predictors in this category are civilian labor for men only, White, Black, 16-19 year old and so forth. Not all of these series make economic sense to predict recessions, thus we remove them.

Furthermore, we remove series that have statistical predictive power (the Granger causality test/WoE/IV suggested those series as predictors) but do not make intuitive sense to predict recessions such as the CPI of education, books and supplies for all urban consumers. With this manual procedure, we are able to reduce the set of potential predictors to 192 from 1,563 variables. Therefore, we utilize 192 potential predictors for the three competing models that are: (1) Logit/Probit (statistical data mining) model, (2) Random forest, and (3) The gradient boosting. Our benchmark model utilizes the yield curve as a predictor.

FINDING THE BEST SET OF PREDICTORS FOR RECESSION FORECASTING: DISCOVERING HIDDEN CONNECTIONS

We have now narrowed down the list of potential predictors to 192 variables. Next, we need to classify those 192 variables into categories. For example, we created the category "inflation" and put all inflation related variables (i.e. CPI and PCE deflator) in that category. Likewise, nonfarm payrolls and unemployment rate fall in the "employment" category and so on. We end up having 40 different categories. The 192 variables we have selected as potential predictors are individually statistically useful to predict recessions. Now we need to find the ideal combination of predictors that represent different sectors of the economy.

As we know, economies evolve over time and the strength of relations between different sectors of an economy also vary. Our approach will find a set of sectors that are statistically more accurate to predict recessions than any other set in our analysis. Basically, we utilize all possible combinations of the 192 variables and, by doing so, we explore the hidden connections between different sectors. Furthermore, including one variable from a category at a time avoids the potential multi-collinearity problem*(Please go to the end of the paper and see Note 10)*.

We Ran 30 Million Regressions

Here is the outline of our procedure to find the best set of predictors from the 40 different categories. We set a logit/probit modeling framework with eight predictors (nine variables in a model: one dependent variable and eight predictor variables). Moreover, we are interested in a distinct combination of the eight predictors, meaning we want eight predictors from eight different sectors. For example, we pick the unemployment rate as a predictor from the "employment" category and the next predictor comes from the "inflation" category (CPI for example), the S&P500 from "equity", 10-year Treasury yield from "interest rates" and housing starts from the "housing" category and so on. Therefore, eight predictors represent eight different sectors of the economy. In addition, we repeat the process by keeping the unemployment rate (to represent "employment") in the model but change the rest of the predictors of the model one by one. That is, we include eight predictors at a time and then replace predictors with others, but keep the total number of predictors to eight. Why do we do this?

This process tests the relationship of every combination of variables. For example, the unemployment rate will team up with each and every predictor of the rest of the 39 categories. Put differently, each category not only gets a chance to perform as a predictor

but also team up with other sectors to predict recessions. Therefore, we employ all possible combinations of these 40 categories and 192 variables and that process allows us to explore hidden connections between different sectors and improve recession prediction accuracy. The process is very time-intensive, taking several weeks of continuously running code. In total, we ran 30 million different models. We utilize the Schwarz information criterion (SIC) to narrow down 30 million models to a manageable list of models. We selected the top 1,500 models in this step using the SIC values (as we know a model with the lowest SIC value is the preferred one among competitors). The selected 1,500 models contain eight predictors in each model but all those models include distinct combinations of the eight predictors.

From the 1,500 different combinations of eight-predictors we need to select the final model (one model with eight predictors). Moreover, these 1,500 models were selected by using in-sample criterion, however, our objective is to forecast future recessions accurately (out-of-sample forecasting). Therefore, we utilize simulated real time out-of-sample forecast error as the criterion to find the best model among the 1,500 models.

Precisely, we utilize the 1972-1988 period to generate the probability of a recession during the next 12 months and then re-estimate the model using the 1972-1989: 1 period (include the next month in the estimation period) and again generate probability of a recession for the next 12 months. We iterated this process till we reached the last available data point, which is December 2016. The major benefit of this recursive forecasting is that we know the actual outcome (recession or no recession during the next 12 months), but we did not share that information with the model. This allows us to calculate the model's accuracy. We repeat this process for each of the 1,500 models and select the model with the highest accuracy. That is, we select the set of eight-predictors, which forecast recessions during the 1989-2017 (period for the simulated out-of-sample forecasting) more accurately than the rest of the 1,499 models. Basically, we ran over half a million (504,000) simulations to select final model. The selected logit/probit model is utilized to represent the statistical machine learning/data mining approach.

HAPPY HUNGER GAMES: AND MAY THE ODDS BE EVER IN YOUR FAVOR

Who Perform the Best?

The objective of this report is to find an approach/model that predicts recessions more accurately than other contenders. The first contestant, which is also the benchmark approach, is a probit model with the yield curve as the predictor. The second approach is the statistical machine learning/data mining and a logit/probit model with eight

predictors are utilized to represent the data mining approach. The random forest and gradient boosting methods are utilized to represent machine learning.

Before we introduce a statistical tool to evaluate a model's performance, we will discuss our precise objective about the target variable. That is, our target is to predict recessions accurately and our dependent variable is binary with zeros (non-recessionary periods) and ones (recessions). Furthermore, an accurate forecast from a model correctly predicts either a recession or a non-recessionary period in the forecast horizon. By the same token, an inaccurate forecast implies missing of a recession/non-recession. Precisely there are the following possibilities for a forecast:
(1) True positive: model correctly predicts recession;
(2) True negative: model accurately predicts non-recessionary period;
(3) False positive: model predicts a recession when there was no recession; and
(4) False negative: model predicts non-recession but there was a recession.

With this information, we can restate our objective: a forecast should be true positive and true negative and avoid both false negative and false positive.

In addition, adjusting the probability threshold for a recession directly influences the changes of false positives. For example, 60% or higher probability indicates a recession, otherwise no recession. That threshold helps reduce chances of false positives. However, a higher probability-threshold also poses the risk of missing a recession. On the other hand, a threshold using a lower probability (20% probability as a threshold, for instance) would lead to more false positives. With this discussion in mind, we can introduce our statistical method to evaluate forecasts of a model.

The Relative Operating Characteristic (ROC) Curve

The relative operating characteristics (ROC) curve is a helpful tool to evaluate a model's performance *(Please go to the end of the paper and see Note 11)*. The ROC curve helps to find an optimal threshold by plotting different thresholds' performances. Put differently, the ROC curve shows a plot of a true positive (correct forecast) against a false positive (false signal) of a given threshold. Essentially, the ROC curve depicts accuracy (true positive vs. false positive) of different thresholds and the threshold which produces the highest accuracy can be selected. That is, a threshold can be identified by the ROC curve which produces the maximum hit rate along with least false signals.

In addition, a further nuance of the ROC curve is known as the area under the curve (AUC). The ROC AUC, in the present case, is equal to the probability of predicting recessions accurately. That is, the ROC AUC values vary between zero and one and a

value close to one represents higher accuracy while a value near zero represents a useless model. Therefore, the ROC AUC will help us determine which model is the best among competitors. Furthermore, we will estimate the ROC and ROC AUC for both in-sample analysis (complete sample period) and out-of-sample forecasting simulation (1989-2017) for each of the four models to evaluate which model is the most accurate.

The Legends of Machine Learning: The Random Forest and the Gradient Boosting Approaches

Applications of machine learning techniques in economics/finance are a relatively new phenomenon *(Please go to the end of the paper and see Note 12)*. The basic logic behind machine learning techniques is to utilize the available information effectively to generate accurate predictions. That is, machine learning techniques allow us to transform a computationally-hard problem into a computationally-efficient solution. In contrast to the traditional econometric techniques, which worry about issues such as linear/non-linear, small/large samples and degree of freedom/more predictors than observations etc., machine learning techniques find a connection between the target variable and predictors and then utilize that information to form a prediction. Put differently, most machine learning techniques divide data into segments and then utilize some of those segments for estimation and others for validation.

The basic idea behind most machine learning techniques is that an algorithm sets a loss function (minimum forecast error, for example) and finds a combination of predictors that produce a minimum forecast error, on average, among competitors. Before we discuss our models, we need to clarify one more thing, which is the classification and regression problem. In machine learning, if the target variable is binary (or categorical), then it is called a classification-problem while for a continuous-target variable the term regression-problem is utilized. Since our target variable is binary, we are dealing with a classification problem.

The Random Forest Approach

The random forest is one of the more famous techniques of machine learning, and it is also our first model. Typically, the random forest approach produces accurate forecasts (both in-sample and out-of-sample) (Mullainathan & Jann, 2017). However, the random forest is a black box in the sense that there are no formal estimated parameters or explanations as to what variable has the highest predictive power. For example, in a traditional econometric model we estimate a coefficient that states an average relation between dependent and independent variables. However, in the case of a random forest, we do not have such coefficients. One major reason that the random forest is a black box is that the random forest is an ensemble technique that originates from the decision tree (or classification and regression tree, CART).

A tree, in simple words, successively chooses each predictor by splitting variables into two groups (partisans) and calculates the mean squared error (MSE). The tree splits at the point that minimizes MSE. The splitting process continues by further splitting each group into two new groups and calculates the MSE for each new group. Typically, in machine learning, these splitting points are called *nodes*. The splitting process continues until the stopping point is reached and the end point is labeled as *leaves*. A decision tree is simple to build and generates a very good in-sample fit but a horrible out-of-sample forecast. One major reason for bad out-of-sampling is that such trees are built using in-sample information and, typically, do not include out-of-sample forecasting.

Breiman (2001) improved the decision tree approach and his framework is known as the random forest *(Please go to the end of the paper and see Note 13)*. The basic logic behind the random forest approach is that instead of generating one tree, we can create many trees (number of trees can be in thousands or millions depending on the objective). Furthermore, if trees are independent and unbiased, then the average of those trees would be unbiased with a potentially small variance, and more likely to produce a better out-of-sample forecast. The averaging of different trees is called ensemble learning or the random forest approach. Essentially, averaging many models tends to provide better out-of-sample forecasts than a single model.

The Gradient Boosting Approach

The gradient boosting is also an ensemble approach and a very powerful machine learning tool for forecasting. The basic idea behind the gradient boosting is that a weak learner (in-accurate model) can be modified to become a better one (accurate model). A weak model can be defined as a model which produce a slightly better forecast than a random chance. Furthermore a tree (or decision tree) can be considered a weak learner as trees are infamous for very bad out-of-sample forecast. Friedman (1999) provides a formal framework to estimate a gradient boosting model *(Please go to the end of the paper and see Note 14)*.

Essentially, in the gradient boosting modeling approach, we set a loss function (minimum MSE, for example) and then add weak learners (trees for example) to optimize the loss function using a gradient descent process (Friedman, 1999). Put differently, the gradient boosting approach help us to form an accurate forecast using so many inaccurate predictions by setting a learning (additive) modeling process.

For the random forest and gradient boosting approaches, we utilize the set of 192 variables as potential predictors (as we have discussed those 192 variables are selected using statistical data mining). The logit/probit models represent statistical machine

learning, utilizing eight predictors. The benchmark probit model employs the yield curve as a predictor.

THE RESULTS: THE IN-SAMPLE AND OUT-OF-SAMPLE SIMULATIONS

As mentioned earlier, the ROC AUC is utilized to measure a model's performance. We estimated ROC AUC for all models and then compared them to select the best performing among the four models. The ROC curve along with an AUC for the random forest approach are plotted in Figure 1 (for in-sample analysis) and Figure 2 (for out-of-sample forecasts). A ROC curve, Figure 1, shows the plot of true positive rate (y-axis) against false positive rate (x-axis) at various threshold settings. The diagonal line (dotted line in Figure 1) is known as the line of no-discrimination, as an outcome on the line, point B for example, is almost as good as a random guess (probability of a true positive is equal to probability of a false positive). The area to the left of the diagonal line shows when the chance of a true positive rate is higher than the probability of a false positive rate at a given threshold. The left upper corner, point A for example, indicates the best possible prediction as it shows 100% accuracy. The right bottom corner, the corner closest to the point C, represents the worse possible prediction: a 100% chance of a false positive rate.

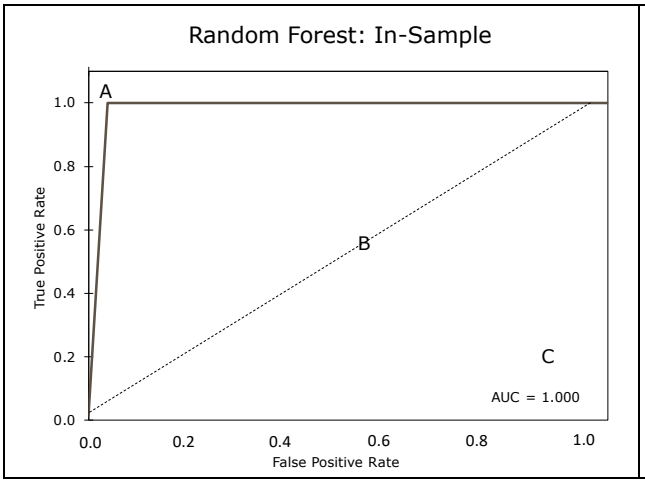

FIG 1. RANDOM FOREST: IN-SAMPLE

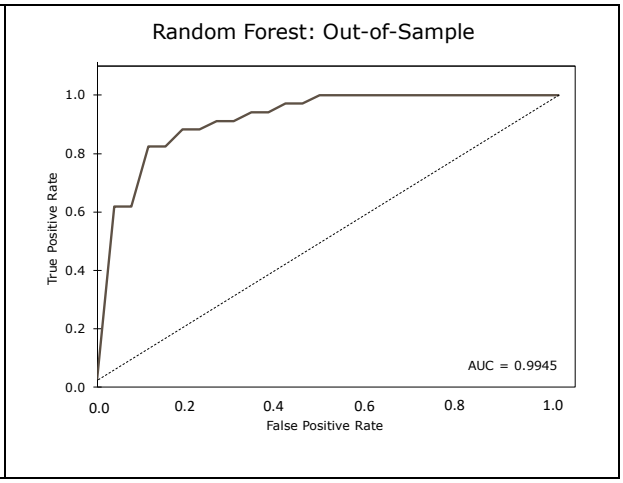

FIG 2. RANDOM FOREST: OUT-OF-SAMPLE

The random forest in-sample analysis that produces an ROC AUVC value of one indicates the best in-sample fit. It is not a surprise that the random forest approach tends to produce a great in-sample fit. The out-of-sample forecasting simulations prove that the random forest approach is able to predict all recessions (1990, 2001 and 2007-2009 recessions) without producing a false positive as the ROC AUC is very close to one (0.9945), Figure 2. The random forest approach performance is excellent in both in-sample and out-of-sample simulations.

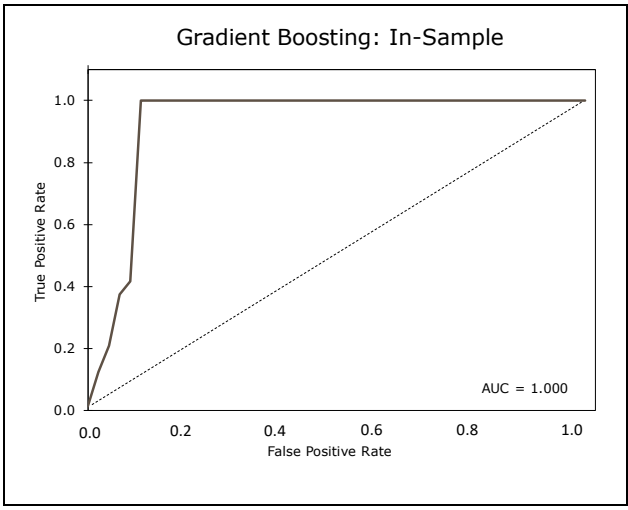

FIG 3. GRADIENT BOOSTING: IN-SAMPLE

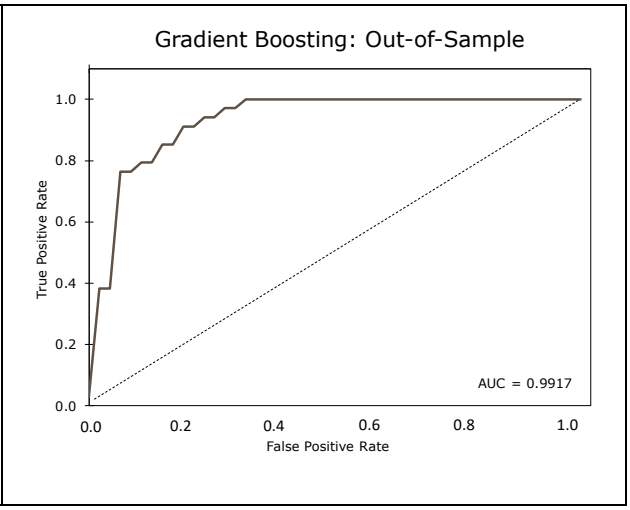

FIG 4. GRADIENT BOOSTING: OUT-OF-SAMPLE

The results based on the gradient boosting are shown in Figure 3 (in-sample) and Figure 4 (out-of-sample). The in-sample AUC value is 1 and 0.9917 for the out-of-sample simulations. That is, the in-sample performance of the gradient boosting is equal to the random forest in-sample accuracy, but the random forest performed slightly better than the gradient boosting in the out-of-sample forecasting. The statistical data mining (logit/probit) approach came in at the third position with the ROC AUC value of 0.9756 (in-sample) and 0.8746 (out-of-sample) (Figure 5 & Figure 6).

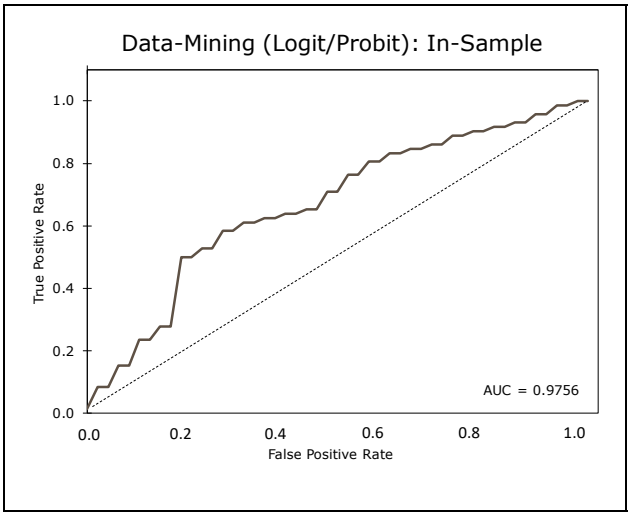

FIG 5. STATISTICAL DATA-MINING: IN-SAMPLE

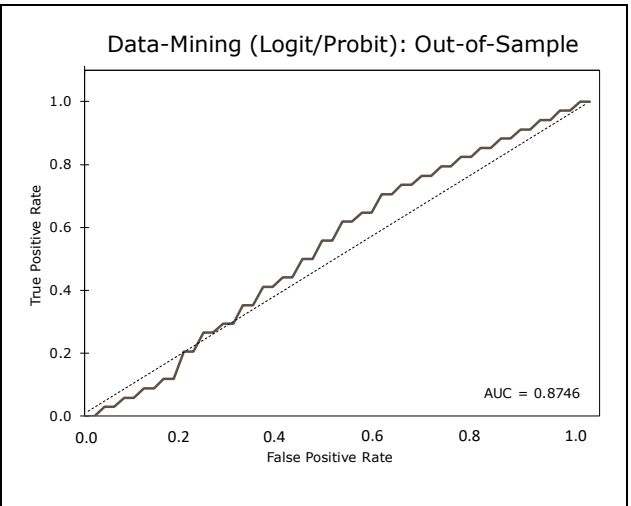

FIG 6. STATISTICAL DATA-MINING: OUT-OF-SAMPLE

The benchmark probit model produces 0.9560 (in-sample) and 0.8266 (out-of-sample) values for the ROC AUC, the worst performer in our analysis (Figure 7 and Figure 8). The

average of the all models are shown in Figure 9 and 10. The average of four models performed better than the benchmark and the data mining. However, the random forest and gradient boosting outperformed the average of all models. In addition, all four methods produce a very low probability (less than 5%) of a recession during the next 12 months.

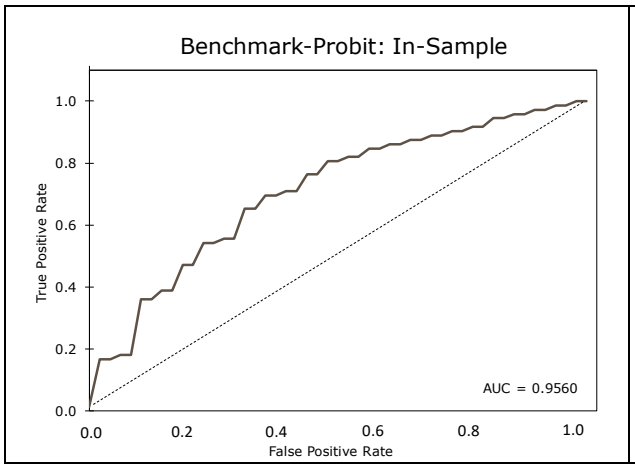

FIG 7. THE BENCHMARK-PROBIT: IN-SAMPLE

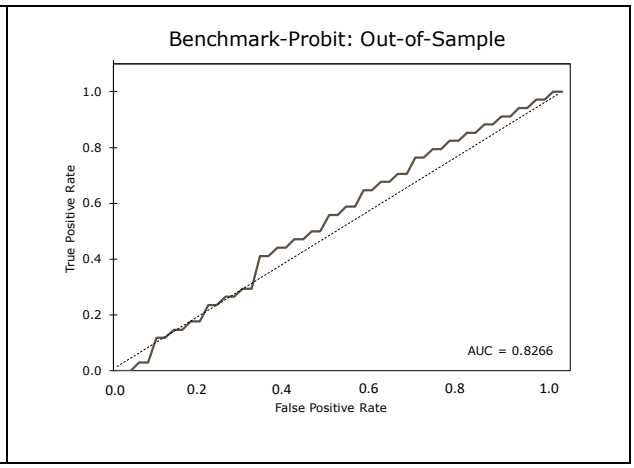

FIG 8. THE BENCHMARK-PROBIT: OUT-OF-SAMPLE

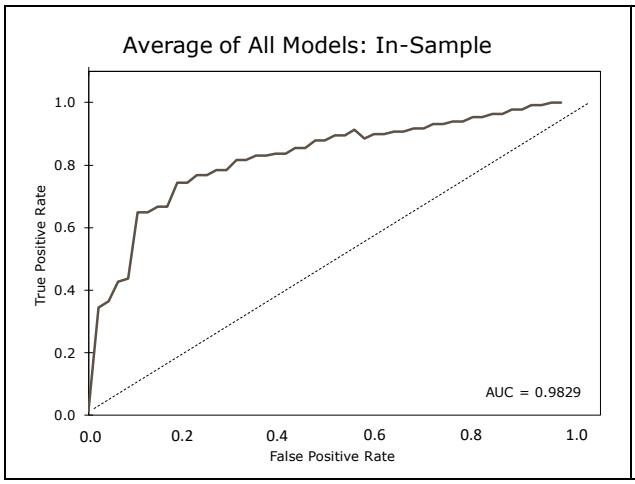

FIG 9. AVERAGE OF ALL MODELS: IN-SAMPLE

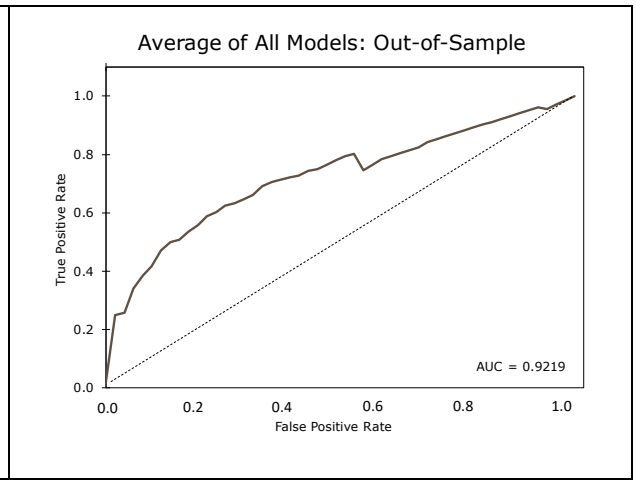

FIG 10. AVERAGE OF ALL MODELS: OUT-OF-SAMPLE

Summing up, the evolution of big data and machine learning techniques open doors to improve economic/financial variables' prediction. We believe that an effective modeling process can be dividing into two phases. The extraction of the useful information (signals vs. noises) is the first phase of an accurate modeling process and the second phase consists of utilizing that information efficiently (select appropriate estimation techniques). For example, in this analysis, we utilize the statistical data mining techniques to narrow down the FRED dataset (which contains more than 500,000 variables) to 192

variables. In other words, the data mining helps us to extract useful information (find signals and cancel noises).

In the estimation simulations, machine learning (the random forest and gradient boosting) techniques provided more accurate results using the same dataset than those of the logit/probit (statistical data mining) models. One major reason is that the logit/probit approach estimates an average relationship to predict an outcome. An average estimation process may limit the effectiveness of the modeling approach as relations between variables evolve overtime and the strength of the relationship fluctuates overtime as well. Machine learning techniques (both the random forest and gradient boosting) dig deeper and find useful statistical relationship between the target variable and predictors to generate forecasts. Therefore, both phases are necessary for accurate forecasting.

CONCLUDING REMARKS: IT'S NOT WHAT YOU HAVE, IT'S HOW YOU USE IT

The evolving nature of the economy forces decision makers to look for new tools to capture growing complexities in the economy to help them form effective policy. Our work proposes a new framework to generate accurate forecasts using a large set of predictors and machine learning tools. We stress that the extraction of useful information and the effective utilization of that information is crucial for accurate predictions.

The development of machine learning techniques along with large dataset availability opens doors to improving the predictive power of economic variables. We believe that an effective modeling process can be divided into two phases. The extraction of the useful information (signals vs. noises) first phase of an accurate modeling process. The second phase consists of utilizing that information efficiently.

Our analysis utilized the statistical data mining techniques to narrow down the FRED dataset (which contains more than 500,000 variables) to 192 variables. In the estimation simulations, machine learning techniques provided more accurate results using the same dataset than those of the logit/probit (statistical data mining) models. One major reason is that the logit/probit approach estimated an average relationship to predict an outcome. An average estimation process may limit the effectiveness of the modeling approach as relations between variables evolve over time, and the strength of the relationship fluctuates over time as well. Machine learning techniques (both the random forest and gradient boosting) dig deeper and find useful statistical relationship between the target

variable and predictors to generate forecasts. Therefore, both phases are necessary for accurate forecasting.

Notes:

[1]Typically, an over-fitted model shows very good in-sample fit but very bad out-of-sample forecasts. For more detail see, Silvia, J., Iqbal, A., et al. (2014). *Economic and Business Forecasting: Analyzing and Interpreting Econometric Results.* Wiley 2014.

[2]The FOMC releases its meetings transcripts with a five year lag and can be found here: https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm

[3]It is worth mentioning that sometimes in machine learning/other big data applications different terms (instead of in-sample and out-of-sample) are utilized such as training-sample or cross-validations etc. For more detail see, Hastie, T et al. (2008). *The Elements of Statistical Learning*. 2nd Edition, Springer. The basic logic behind all these procedures (analysis) is similar and that is to utilize some part of the available information (either time span, number of observations or both) to establish some statistical association/relationship and then utilize those relationships to forecast future events (unknown values/outcome).

[4]For more detail about the FRED dataset see: https://fred.stlouisfed.org/

[5]If we use 1990-2005 period for in-sample analysis and rest of the period for out-of-sample simulation then that will provide an opportunity to forecast only one recession and that recession is the Great Recession. That is not enough time span to test the real-time out-of-sample accuracy of a model.

[6]Note, for the Granger causality analysis, we utilize GDP growth rates as a target variable instead of a binary variable (dummy variable to represent recession and non-recession periods). As the Granger causality test assumes the target variables are continuous not binary variables.

[7]For more detail about the Granger causality test see, Granger, C.W.J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3).

[8]A p-value of less than or equal to 0.05 would reject the null hypothesis of no-causality and that indicates the variable in the model is a good predictor of the target variable.

[9]For more detail about WoE and IV see, Lin, Alex. (2013). Variable Reduction in SAS by using Weight of Evidence and Information Value. The full paper is available at: https://support.sas.com/resources/papers/proceedings13/095-2013.pdf

[10]In simple words, if two (or more) predictors of a model are highly correlated with each other than that issue is known as multi-collinearity. Typically, the multi-collinearity problem leads to an overfitting issue.

[11]For a detailed discussion about the ROC curve see, Lahiri, K., and J. G. Wang (2013). Evaluating Probability Forecasts for GDP Declines Using Alternative Methodologies. *International Journal of Forecasting*, 29, 175-190.

[12]For more details about machine learning applications in economics see Mullainathan, Sendhil and Jann Spiess. (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives*, 31(2).

[13]Breiman, Leo. (2001). Random Forests. Statistics Department, University of California, Berkeley, CA. The paper is available at: https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf

[14]Friedman, Jerome H. (1999). Greedy Function Approximation: A Gradient Boosting Machine. The full paper is available at: https://statweb.stanford.edu/~jhf/ftp/trebst.pdf

Journal of Applied Economics and Business

OPEN ACCESS ARTICLE

CORPORATE GOVERNANCE AND BANKING SECTOR PERFORMANCE IN NIGERIA

Dosunmu Adebukola Ibitamuno*, Okey Onuchuku*, Alwell Nteegah*

*Institute of International Trade and Development, University of Port Harcourt

Abstract

This study investigates the effect of corporate governance on the performance of banks in Nigeria over the period 2012-2016. To achieve the purpose of the study data on corporate governance proxy by board size, executive and non-executive board members, interest rate margin, profit level and Return on Asset (ROA) of 15 deposit banks in Nigeria were sourced from the CBN and the records of the banks and analysed using panel technique. The results indicated that none of the variables that represents corporate governance was significant in explaining changes in the performance of banks. This implies that corporate governance has less implication on the performance of banks in Nigeria. The result also shows that board size and non-executive board members have negative effect on ROA while executive board members have positive effect on the performance of banks over the period of this study. The implication of this result is that increase in executive members of a bank's board could improve the performance of the bank in Nigeria. Other variables like interest rate margin and profit level were also insignificant in explaining changes in the performance (ROA) of banks. The result further revealed that the effect of corporate governance on banks 'performance differs across the banks in Nigeria. Based on this result, the study recommends: an upward review of executive members of the board of banks and a periodic review of guidelines on the management of banks in order to enhance efficiency in management of banks and their performance.

Key words: *Return on Asset; Corporate governance; Board size; Interest rate margin and profit.*

INTRODUCTION

Financial scandals around the world and collapse of major corporate institutions in the USA, Europe such as Lehman Brothers, Merrill Lynch, American International Group (AIG), have brought to the forefront, the need for the practice of good corporate governance. Nigeria being a part of the global economy in the last two decades, has followed this development in the financial sector by reinforcing the need for greater concern for corporate governance in financial institutions in the country. According to

Shleifer and Vishny (1997), corporate governance means the ways in which suppliers of finance to corporations assure themselves of getting a return on their investment. Chow (1999) explained that the objectives of corporate governance are to ensure transparency, accountability, adequate disclosure and effectiveness of reporting systems. He asserted that the need for good corporate governance originated from what he termed expectation gap problem which arises when the behavior of companies falls short of shareholders and other stakeholders' expectations.

Nigeria has vibrant but challenging financial environment characterized by endemic systemic governance problems, capacity complains and defaulting in compliance and implementation of laws which has inhibited financial and economic growth. The global economic crisis and the decline in the value of investment of Deposit Money Banks (DMBs) banks particularly in Nigeria are due to distorted credit management and this problem is associated with poor corporate governance. Given the fury of activities that have affected the efforts of banks to comply with the various consolidation policies and the antecedents of some operators in the system, there are concerns on the need to strengthen corporate governance in banks.

Before the consolidation exercise in 2006, the banking industry had about 89 active players whose overall performance led to sagging of customers' confidence. The guiding laws and regulations which contain provisions that address the issue of corporate governance include the Company and Allied Matters Act (CAMA) of 1990, the Prudential Guidelines, the Statement of Accounting Standards (SAS 10), the Banks and Other Financial Institutions (BOFI) Act of 1991, Central Bank of Nigeria (CBN) Act of 1991, CBN Circulars, the Nigeria Deposit Insurance Corporation (NDIC) Act of 1988, and the Investment and Securities Act (ISA) of 1999. According to Sanusi (2010), the current banking crises in Nigeria, has been linked with governance malpractice within the consolidated banks which has therefore become a way of life in large parts of the sector. He further opined that corporate governance in many banks failed because boards ignored these practices for reasons including being misled by executive management, participating themselves in obtaining un-secured loans at the expense of depositors and not having the qualifications to enforce good governance on bank management. The boards of directors were further criticized for the decline in shareholders' wealth and corporate failure. They were said to have been in the spotlight for the fraud cases that had resulted in the failure of major corporations, such as Enron, WorldCom and Global Crossing.

The series of widely publicized cases of accounting improprieties recorded in the Nigerian banking industry in 2009 (for example, Oceanic Bank, Intercontinental Bank, Union Bank, Afri Bank, Fin Bank and Spring Bank) and even the recent sack of the

Chairman and Managing director of Skye Bank in July, 2016 for improper management of funds were related to the lack of vigilant oversight functions by the boards of directors, the board relinquishing control to corporate managers who pursue their own self-interests and the board being remiss in its accountability to stakeholders (Unadiale, 2010). Inan (2009) also confirmed that in some cases, these bank directors' equity ownership is low in other to avoid signing blank share transfer forms to transfer share ownership to the bank for debts owed banks. He further opined that the relevance of non- executive directors may be watered down if they are bought over, since, in any case, they are paid by the banks they are expected to oversee. From literature it can be deduced that corporate governance is influenced by board size, board composition, profitability, capital adequacy, asset base, policy shift, investment, liquidity ratio as well as inflation rate.

In Nigeria, few empirically feasible studies on corporate governance are available in literature, some of the available ones are: Sanda et al., (2005) Ogbechie (2006), Okike (2007), and Adegbite (2015) which all studied the corporate governance mechanisms and firms' performance. In order to address these paucity of facts on the effect of corporate governance on the performance of DMBs, this study examined the role of corporate governance in enhancing the financial performance of banks in Nigeria.

LITERATURE REVIEW

According to the Agency Theory, the need for corporate governance arises because of the separation of management and ownership in the modern corporation. This separation of ownership from control implies a loss of effective control by shareholders over managerial decisions. Partly as a result of this separation between the two parties, a system of corporate governance is implemented to assist in aligning the incentives of managers with those of shareholders. With the significant increase in equity holdings of investors, there has been an opportunity for a reversal of the separation of ownership and control problems because ownership is not so diffused. One consequence of the separation of ownership from management is that the day to today decision-making power that is, the power to make decision over the use of the capital supplied by the shareholders' rests with persons other than the shareholders themselves. The separation of ownership and control has given rise to an agency problem whereby there is the tendency for management to operate the firm in their own interests, rather than those of shareholders' (Jensen & Meckling, 1976; Fama & Jensen, 1983). This creates opportunities for managers to build illegitimate empires and, in the extreme, outright expropriation. These presumptuous agency theories are however predominantly invalid in developing

countries such as Nigeria. For instance, the aftermath of Nigeria independence from Britain in 1960 led to an indigenization programme with resulted in majority ownership by government, individuals and families in corporate Nigeria (Nmehielle & Nwauche, 2004) hence there is no single best institutional arrangement for organizing economic systems and corporate governance.

Agency theory supports the delegation and the concentration of control in the board of directors and use of compensation incentives. The board of directors' monitor agents through communication and reporting, review and audit and the implementation of codes and policies.

The Stakeholder theory by Sundaram and Inkpen (2004a) also suggest that "stakeholder theory attempts to address the question of which groups of stakeholder deserve and require management's attention". Shareholders play a key role in the provision of corporate governance. Small or diffuse shareholders exert corporate governance by directly voting on critical issues, such as mergers, liquidation, and fundamental changes in business strategy and indirectly by electing the boards of directors to represent their interests and oversee the myriad of managerial decisions in the banking sector, (CBN, 2015). Incentive contracts are a common mechanism for aligning the interests of managers with those of shareholders. The Board of directors may negotiate managerial compensation with a view to achieving particular results. Thus small shareholders may exert corporate governance directly through their voting rights and indirectly through the board of directors elected by them.

However, a variety of factors could prevent small shareholders from effectively exerting corporate control. There are large information asymmetries between managers and small shareholders as managers have enormous discretion over the flow of information. Also, small shareholders often lack the expertise to monitor managers accompanied by each investor's small stake, which could induce a free-rider problem. Stakeholder theory offers a framework for determining the structure and operation of the firm that is cognizant of the myriad participants who seek multiple and sometimes diverging goals.

Due to the vital role banks play in promoting economic growth and development, the conduct of their financial intermediation functions and the environment in which they operate remain particularly important. In recognition of this strategic importance of banks, knowing full well that the governance of any banking institution in Nigeria is statutorily placed in hands of board of directors, appointment and activities of bank directors in Nigeria are governed by laws and regulations, which presumably, the implementing bodies rigorously enforce. The financial crisis of 2008 has shown that the corporate governance of financial institutions has been an under highlighted area, as there were massive failures at major institutions in advanced countries. Corporate

governance in financial institutions has been identified to differ from that of corporations, but in which ways is not yet clear besides the important role of prudential regulations, given the special nature of banks. In this area, more work is needed for emerging markets as well, in part related to the role of banks in business groups. While there is some research on state ownership, corporate governance of banks in emerging markets is little analyzed. (Claessens &Yurtoglu, 2013).

According to Claessens and Yurtoglu (2013), the identified channels in which corporate governance affects corporations and countries include:

- The increased access to external financing by firms. This in turn can lead to greater investment, higher growth, and greater employment creation;
- Lowering of the cost of capital and associated higher firm valuation. This makes firms more attractive to investors, leading to growth and more employment;
- Better operational performance through better allocation of resources and better management. This creates wealth more generally;
- Good corporate governance can be associated with less financial crises, important, as highlighted recently again, given the large economic and social costs of crises; and
- Good corporate governance can mean generally better relationships with all stakeholders. This helps improve social and labor relationships and aspects such as environmental protection, and can help further reduce poverty and inequality.

All these channels matter for growth, employment, poverty, and well-being more generally. Empirical evidence using various techniques has documented these relationships at the level of the country, the sector, and the individual firm and from the investor perspectives. Klein et al., (2004: 32) examined the relationship between corporate governance and firm value by using the corporate governance index (CGI) and Tobin's Q, which measures the firm's value, the results concluded that corporate governance does matter in a firm value

According to Sanusi (2010) it was well known in the industry that since consolidation, some banks were engaging in unethical and potentially fraudulent business practices and the scope and depth of these activities were documented in recent CBN examinations. Governance malpractice within the consolidated banks has therefore become a way of life in large parts of the sector, enriching a few at the expense of many depositors and investors. Sanusi further opined that corporate governance in many banks failed because boards ignored these practices for reasons including being misled by executive management, participating themselves in obtaining un-secured loans at the expense of

depositors and not having the qualifications to enforce good governance on bank management. In addition, the audit process at all banks appeared not to have taken fully into account the rapid deterioration of the economy and hence of the need for aggressive provisioning against risk assets. As banks grew in size and complexity, bank board's often did not fulfill their functions and were lulled into a sense of well-being by the apparent year-over year growth in assets and profits. In hindsight, boards and executive management in some major banks were not equipped to run their institutions.

Eisenberg et al., (1998) studied 879 Finnish firms and found that companies with smaller boards had higher ROA, positing that the effect of board size may in part depends on the size and wellbeing of the firm. Spencer Stuart Board Index (2008) also indicated that globally, board size has been reducing over the years and that there is a continued quest towards smaller board size. However, other studies by (Druckeriv, 2002; Dalton et al., 1999; Kiel & Nicholson, 2003; Adams & Mehran, 2003; Anderson et al., 2004; Coles et al., 2008; Belkhir, 2009; Arslan et al., 2010; Chang & Duta, 2012), found that board size have a positive impact on the stock market performance of company. This implies that found that, large board size improves corporate performance through enhancing the ability of the company to establish external connection with the environment, providing on that way rare resources for company operations.

According to Caprio et al., (2007), and Andres and Vallelado, (2008) Size of board plays a critical role in the company's performance because it supervises the management and takes more human capital to advise management. Javid and Iqbal (2008) and Yasser et al., (2011) studies reported positive relationship between board size and firm performance. However, the findings of Yermack (1996); Eisenberg et al., (1998); Mak and Kusnadi (2004); and Andres et al., (2005) found is negative relationship between board size and firm efficiency.

In a related study by Tariq et al., (2014) found that Non-executive director participation in the board increases the performance of the decision and it also monitors the affair of corporation in a better way. According to the authors, the purpose of involvement in the board is to protect and increase the value of shareholder. Their participation in board brings new windows of universe (Tricker, 1984). They safeguard the interest of shareholder from the management. Empirical studies by (Weisbach, 1988; Prevost et al., 2002; Anderson & Reeb, 2004; Rebeiz & Salameh, 2006) reported a positive and significant relationship between outsider directors and firm performance. Notwithstanding, others studies by Baysinger and Butler (1985), Hermalin and Weisbach (1991), Agrawal and Knoeber (1996), and Yasser (2011) found a negative relationship between the outside directors and firm performance.

Kanwal and Nadeem (2013) studied the impact of macroeconomic factors on deposit money banks' performance in Pakistan over the period 2001 to 2011 using the Pooled Ordinary Least Square (POLS) method. The study used inflation rate, real gross domestic product (GDP) and real interest rate as explanatory variables while return on assets (ROA), return on equity (ROE) and equity multiplier (EM) were used as dependent variables. Their result indicated that real interest rate was significant and positively related to all the measures of profitability. GDP has negative nexus with ROE and EM only, while it is insignificant with ROA. The result also indicated that inflation rate was negatively related to the three measures of profitability. The study therefore concludes that macroeconomic variables have very strong impact on banks' performance in Pakistan.

Adeusi et al., (2014) examined factors that affect the profitability of 14 banks in Nigeria over the period 2000 to 2013. The study used ROA as a proxy for profitability and selected bank-specific, industry-specific and macroeconomic indicators. The findings of the study revealed that total loans to total assets, interest income to interest expenses, and GDP growth have the most significant effect on banks' profitability. The finding also revealed that capital adequacy, liquidity risk and inflation have no significant effect on banks 'performance in Nigeria.

Prowse (1997) argued that research on corporate governance applied to financial intermediaries especially banks, is indeed scarce. This shortage is confirmed in Oman (2001); Goswami (2001); Lin (2001); Malherbe and Segal (2001), and Arun and Turner (2002). They held a consensus that although the subject of corporate governance in developing economies has recently received a lot of attention in the literature, however, the corporate governance of banks in developing economies has been almost ignored by researchers. The idea was also shared by Caprio and Levine (2001). To the best of the researchers' knowledge, apart from the few studies by Caprio and Levine (2002), Peek and Rosengren (2000), Okoi and Ocheni (2000), Okike (2007), and Adegbite (2015) on corporate governance and bank performance, not much empirical studies have been carried out specifically on this subject especially in developing economies like Nigeria. A similar study carried out in Nigeria was by Sanda et al., (2005) where they looked at corporate governance and the financial performance of non-financial firms. It is on this premise that this study seeks to examine the impact of corporate governance on the performance on banks in Nigeria.

METHODOLOGY

To achieve the objectives of the study, data on Return on Asset (ROA), board composition, board size and interest rate margin and profit of 15 Deposit Money Banks selected for this study were sourced from the records of the banks and the Central bank of Nigeria from 2012-2016. The data was analyzed using the panel technique. The purposes of our analysis are: to examine the relationship between corporate governance and performance of the banking sector (ROA) and to find out if the impact of corporate governance on performance (ROA) varies across the banks in Nigeria or not.

This study employs a modified version of the econometric model adopted by Eisenberg et al., (1998) and Adeusi et al., (2014) in examining the effect of corporate governance on the performance of banks. It also used the CBN best practice rules and the specific governance index as provided by the Institutional Shareholder Services. The CBN prudential guideline specifies the number of directors (executive and non-executive) that constitute a board. Deposit money banks are expected to comply with such guideline in order to enhance efficient management and performance of the banks. In this study, performance is measured by return on Asset (ROA). The rationale for the use of this variable as a measure of performance is that banks in Nigeria are privately owned firms financed by individual/group of individuals whose interest is to maximize profit. Return of Asset captures how valuable the assets of banks are over time. The higher the quality of asset of a bank, the higher, its potentials to attract investors and growth

Corporate governance is critical for performance of a firm and the banking sector in particular. This is because the governance and management of a bank is key in decision making and growth of the business. Corporate governance also has serious implication on credit administration of the bank which directly affect performance. Based on this illustration, the study specifies a bank performance function thus:

$$ROA_{it} = f\left(BOS_{it},\ BED_{it},\ BND_{it},\ INTR_{it} + PROFT_{it}\right) \quad \text{(I)}$$

In order to estimate the functional relationship between corporate governance and banks' performance (ROA) using econometric technique, equation 1 is expressed in mathematical form as follows:

$$ROA_{it} = \beta_0 + \beta_1 BOS_{it} + \beta_2 BED_{it} + \beta_3 BND_{it} + \beta_4 INTR + \beta_5 PROFT_{it} + U_{it} \quad \text{(II)}$$

Where; ROA_{it}= Return on Asset; BOS_{it} = Board size; BED_{it}= Executive Board Composition; BND_{it}= Non executive Board Composition; $INTR_{it}$= Bank charges proxy by interest rate margin $PROFT_{it}$= Profit of the banks and U_{it}= Error term.

RESULTS AND FINDINGS

We started our analysis by examining the trend in board size in the selected 15 banks investigated. The trend in figure 1 shows that UBA, Union bank, first bank and Skye bank has the highest number of board size while Guaranty Trust bank, FCMB and Fidelity banks have the least number of board members. In nutshell, there appeared to be a common trend in board size among the banks operating in Nigeria. This may be the result of the guideline by CBN to all the banks on the maximum number of board.

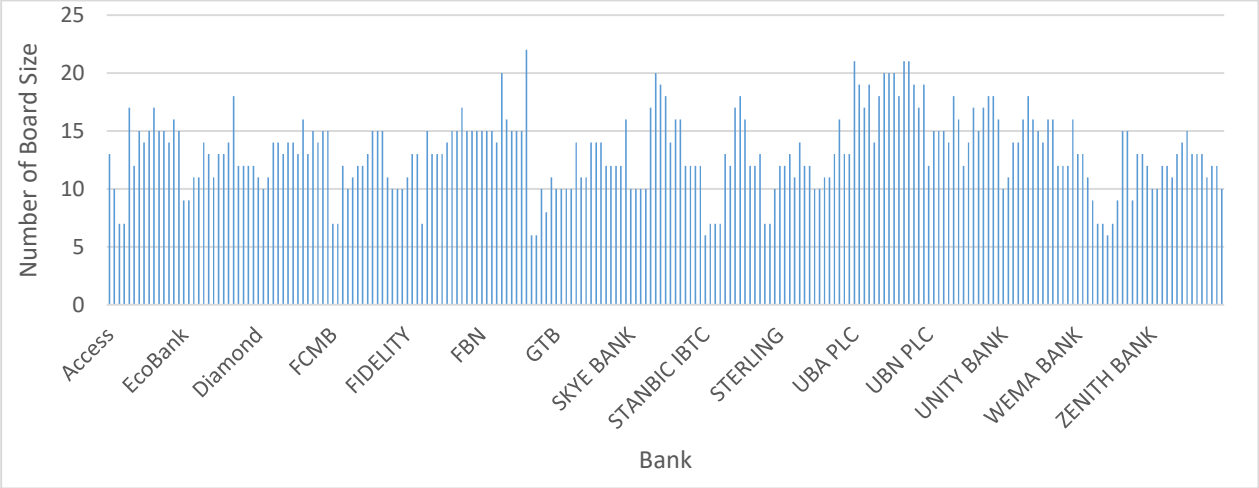

FIGURE 1. TREND IN BOARD SIZE IN THE 15 SELECTED BANKS IN NIGERIA 2002-2016

Figure 2 shows that Return on Asset (ROA) appeared to evenly distribute among the selected 15 banks. However, it was very high in Skye, UBA and Wema banks. This shows that ROA differ across the banks unlike the evidence/trend in board size.

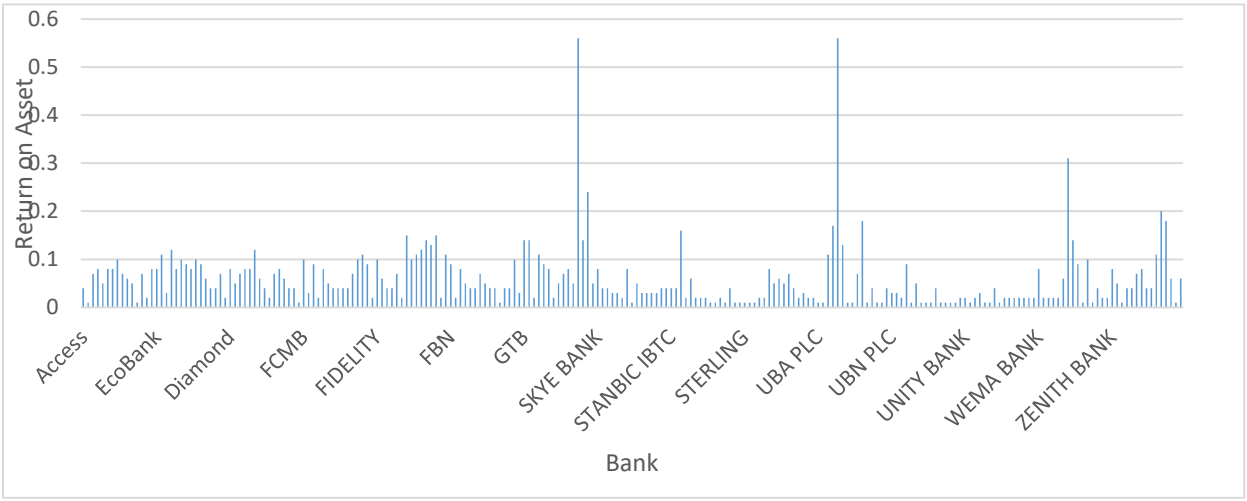

FIGURE 2. TREND IN RETURN ON ASSET IN THE 15 SELECTED BANKS IN NIGERIA 2002-2016

TABLE 1. DESCRIPTIVE STATISTICS

Variable	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**
ROA	0.594	0.066	0.01	0.56
Bos	13.20	3.282	6	22
Bed	4.57	1.981	1	11
Bnd	8.79	2.209	4	16
Intr	19.88	3.388	14.82	26.04
Proft	4335.09	27859.49	-286169.1	126836.8

The descriptive statistics reported in Table 1 revealed a high deviation in the mean of the variables investigated in this study. The high deviation from the mean in ROA, board size, executive board composition, non-executive board size, interest rate margin and profit level show that the variables experienced very high level of fluctuation during the period under study. This wide fluctuation in the variables could be attributed to the instability in the banking sector in Nigeria.

TABLE 2. PAIRWISE CORRELATION RESULT: ROE BOS BED BNDINTR,PROFTSTAR(6)

Variable	Roe	Bos	Bed	Bnd	Intr	Proft
ROA	1.000					
Bos	-0.112	1.000				
Bed	-0.003	0.6995*	1.000			
Bnd	-0.159*	0.7752*	0.2080*	1.000		
Intr	-0.050	-0.0231	-0.0567	0.0214	1.000	
Proft	0.182*	0.0421	0.0987	-0.0615	0.1136	1.000

The pairwise correlation result in Table 2 revealed a negative and weak correlation between board size, executive board composition, non-executive board composition, interest rate margin and ROA. This implies that components of corporate governance and interest rate margin have very weak relationship with bank performance (ROA). On the other hand, profit level has positive but weak correlation with return on asset. The correlation result also revealed that only board non-executive members and profit level are significant.

TABLE 3: PANEL RESULT OF CORPORATE GOVERNANCE AND RETURN ON ASSET (ROA)

Random Effect (RE) Model				Fixed Effect (FE) Model		
Variable	Coefficient	Z-statistic	Prob	Coefficient	T-statistic	Prob
Bos	-0.0027	-0.65	0.517	-0.0036	-0.81	0.416
Bed	0.0011	0.24	0.811	0.0008	0.18	0.857
Bnd	-0.0019	-0.42	0.675	-0.0015	-0.28	0.778
Intr	-0.0011	-0.80	0.426	-0.0015	-0.97	0.335
Proft	2.76e-07	1.76	0.079	2.14e-07	1.24	0.215
Cons	0.127	3.69	0.000	0.1449	4.25	0.000
$R^2 = 0.053$				FE test F(14,190) = 1.92; F-Prob = 0.03; $R^2 = 0.043$		

The panel result reported in Table 3 indicated that board size was negatively and insignificantly related to ROA in the random and fixed effects models. This implies that increases in board size retarded return on asset/performance in Nigerian banks. This result is in tandem with some earlier studies which found higher board size to be unfriendly with performance. Some of the earlier studies that found negative relationship between board size and performance are of Yermack (1996), Eisenberg et al., (1998), Mak and Kusnadi (2004), and Andres et al., (2005). Large Board size may lead to additional cost and reduce return on asset/ performance.

The result of the panel analysis also revealed that executive member of the board of the banks in Nigeria has positive but insignificant relationship with ROA both in the random and fixed effect models. This implies that increases in the executive members of the board of banks stimulated the performance of banks in Nigeria and vice versa. This result is in consonance with the studies of Druckeriv (2002), Dalton et al., (1999), Kiel & Nicholson, (2003), Adams & Mehran (2003), Anderson et al., (2004), Coles et al, (2008), Belkhir (2009), Arslan et al., (2010), Chang & Duta (2012), which found that large board size improves corporate performance through enhancing the ability of the company to establish external connection with the environment, providing on that way rare resources for company operations.

Non-executive board size appeared with a negative coefficient both in the random and fixed effect models. This implies that increases in non-executive members of a board have a negative implication on the performance of banks in Nigeria. This result agreed with the findings of: others studies by Baysinger and Butler (1985), Hermalin and Weisbach (1991), Agrawal and Knoeber (1996), and Yasser (2011), which reported a negative relationship between the outside directors and firm performance. It should be noted that non-executive board members are not directly involved in the day-to-day running and management of the banks. Hence increasing the numbers may offer higher cost burden on the banks.

Interest rate margin appeared with negative sign both in the random and fixed effect models. The implication of this result is that increase in charges impose on credit by banks have adverse effect on the performance (ROA) of banks in Nigeria. Increases in charges may drive away customers to other alternative sources of borrowing and funding for their businesses. This may affect the bottom line of banks and their performance. This result is not in tandem with earlier study by Kanwal & Nadeem (2013) which found a positive and significant relationship between interest rate and performance of banks.

Profit margin has a positive coefficient with return on asset (ROA) both in the random and fixed effect models. This implies that increase in profit level of banks increased ROA while a decrease in profit level diminished ROA. This result agrees with theoretical expectation. Firms/Banks ROA tend to improve as it profits margin rises.

The test for fixed effect indicates that the null hypothesis is rejected given the probability of the F-statistic. This implies that the effect of corporate governance on the performance of banks differ across the banks investigated in Nigeria. This return tends to agree with the trend analysis reported in Figure 2.

CONCLUSION AND RECOMMENDATIONS

The result of our investigation on the effect of corporate governance on the performance of banks indicated that none of the variables that represents corporate governance was significant in explaining changes in the performance of banks. This implies that corporate governance has less implication on the performance of banks in Nigeria. The result also shows that board size and non-executive board members have negative effect on ROA while executive board members have positive effect on the performance of banks over the period of this study. The implication of this result is that increase in executive members of a bank's board could improve the performance of the bank in Nigeria. Other variables like interest rate margin and profit level were also insignificant in explaining changes in the performance (ROA) of banks. The result further revealed that the effect of corporate governance on banks 'performance differs across the banks in Nigeria Based on this result, the study recommends: an upward review of executive members of the board of banks and a periodic review of guidelines on the management of banks in order to enhance efficiency in management of banks and their performance.

REFERENCES

Adegbite, E. (2015). Good corporate Governance in Nigeria: Antecedents, propositions and peculiarities. *Durham University Business School, Queens Campus Stockton, University Boulevard, Thornaby, Stockton-on-tees TS17 6BH, UK.*

Adams, R. & Mehran, H. (2003). Is corporate governance different for banking holding companies? *Economic policy review*, 9, 123-142.

Andres, P. & Vallelado, E. (2008). Corporate governance in banking: The role of the board of directors. *Journal of banking & finance*, 32(12), 2570-2580.

Adeusi, S. O., Kolapo, F. T., Aluko, A. O. (2014). Determinants of commercial banks' profitability panel evidence from Nigeria. *International Journal of Economics, Commerce and Management United Kingdom*, 2(12).

Agrawal, A. & Knoeber, C. R. (1996). Firm Performance and Mechanisms to Control Agency Problems Between Managers and Shareholders. *Journal of Financial and Quantitative Analysis*, 31, 377-397.

Anderson, R. C., Mansi, S. & Reeb, D. M. (2004). Board characteristics, accounting report integrity, and the cost of debt. *Journal of Accounting & Economics*, 37(3), 315-342.

Arslan, O., Karan, M. B. & Eksi, C. (2010). Board Structure and Corporate Performance. *Managing Global Transitions*, 8(1), 3-22.

Arun, T. G. & J. D. Turner. (2002). Corporate Governance of Banking Institutions in Developing Economies: The Indian Experience. Paper presented in the conference on 'Finance and Development' organized by IDPM, The University of Manchester, 23rd July.

Baysinger, B. D. & H. N. Butler. (1985). Corporate Governance and the Board of Directors: Performance Effects of Changes in Board Composition. *Journal of Law, Economics, and Organization*, 1, 101-124.

Chang, B. & Dutta, S. (2012). Dividends and Corporate Governance: Canadian Evidence. *The IUP Journal of Applied Finance*. 18(4), 5-30.

Coles, J. L., Daniel, N. D. & Naveen, L. (2008). Boards: Does one size fits all. *Journal of Financial Economics*, 87(2), 329-356.

Caprio, G., Laeven, L. & Levine, R. (2007). Governance and banks valuations. *Journal of Financial Intermediation*, 16, 584-617.

Capiro, G. Jr. & R. Levine. (2002). Corporate Governance of Banks: Concepts and International Observations. Paper Presented in the Global Corporate Governance Forum Research Network Meeting, April.5.

Central bank of Nigeria Annual report for 2006 'www.cbn.gov.ng'

Central bank of Nigeria Annual report for 2015 'www.cbn.gov.ng'

Chow, K. F. (1999). Effective Transparency and Disclosure Systems for Corporate Governance. Ppaper presented at the Asian Development Bank Institute Workshop on Corporate Governance, Manilla, 23 November.

Claessens S. B. & B. Yurtoglu. (2013). Corporate Governance in emerging markets: A survey Research Department, International Monetary Fund, University of Amsterdam, the Netherlands, CEPR and ECGIWHU–Otto Beisheim School of Management, Germany.

Dalton, D. R., Daily, C. M., Johnson, J. L. & Ellstrand, A. E. (1999). Number of directors and financial performance: A meta-analysis. *Academy of Management Journal*, 42(6), 674-686.

Druckeriv, P. (2002). Corporate Governance after Enron and WorldCom Applying Principles of Results-Based Governance. Insight Conference on Corporate Governance. Working Paper. 1-12.

Eisenberg, T., Sundgren, S. & Wells, M. T. (1998). Larger board size and decreasing firm value in small firms. *Journal of Financial Economics,* 48(1), 35-54.

Fama, E. F. & M. Jensen. (1983). Separation of Ownership and Control. *Journal of Law and Economics,* 26, 301-325.

Freixas, X., Parigi, B. & J. C Rochet. (2003). The Lender of Last Resort: A 21st Century Approach. *Working Paper No.298, European Central Bank.*

Goswami, O. (2001). The Tide Rises, Gradually: Corporate Governance in India, *OECD Development Centre Discussion Paper No 96.*

Iqbal, M. & Zaheer, A. (2011). Impact of corporate governance on financial performance of banks in Pakistan. *Interdisciplinary Journal of Contemporary Research in Business,* 2(12), 217-635.

Inan, W. (2009). Regulatory and Institutional Challenges of Corporate Governance in Nigeria Post Banking Consolidation. *Nigerian Economic Summit Group, Economic Indicators*, 12(2), 114-129.

Javid, A. Y. & Iqbal, R. (2008).Does Corporate Governance Affect a Firm's Performance? A Case Study Pakistani Market. *Journal of business and economics,* 1(1), 11-23.

Jensen, M. & W. Meckling. (1976). *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure, In: Putterman, L. (1986), The Economic Nature of the Firm.* Cambridge University Press.

Kanwal, S. & Nadeem, M. (2013). The impact of macroeconomic variables on the profitability of listed commercial banks in Pakistan. *European Journal of Business and Social Sciences,* 2(9), 186-201.

Kiel, G. C., & Nicholson, G. J. (2003). Board Composition and Corporate Performance: How the Australian Experience informs contrasting Theories of Corporate Governance: *An International Review,* 11(3), 189-205.

Klein, P., Shapiro D. & J. Young. (2004). Corporate governance family Ownership and firm value: *The Canadian evidence. Simon Fraser University working Paper.*

Koch, I. & S. Macdonald. (2000). Industrial Averages as Targets for Financial Ratios, *Journal of Accounting Research,* 4(2), 126-133.

Lin, C. (2001). Private Vices in Public Places: Challenges in Corporate Governance Development in China, *OECD Development Center, Discussion Paper 634-01.*

Mak, Y. T. & Kusnadi, Y. (2004). Size really matters: Further evidence on the negative relationship between board size and firm value. *Pacific-Basin Finance Journal*, 12(1), 1-18.

Malherbe, S. & N. Segal. (2001). Corporate Governance in South Africa, *OECD Development Centre Discussion Paper, Issue 44.*

Nmehielle, V. O. & E. S. Nwauche. (2004). External-Internal Standards in Corporate Governance in Nigeria. *Public Law and Legal Theory Working Paper NO. 115,* George Washington University Law School.

Ogbeche, G. (2006). Environmental dynamism, capital structure and performance: A theoretical integration and empirical test. *Strategic Management Journal,* 12, 1-15.

Okike, E. M. N. (2007). Corporate Governance in Nigeria, the status quo. *Journal compilation Blackwell publishing LTD 9600 Garsington road, Oxford, UK.*

Okoi, I. O. & Ocheni, S. (2014). Empirical study of the impact of corporate governance on the performance of financial institutions in Nigeria. *Journal of Good Governance and Sustainable Development in Africa,* 2(2), 57-73.

Oman, C. P. (2001). Corporate Governance and National Development, *OECD Development Center Technical Papers, Number 180, 362-388.*

Peek, J. & E. Rosengren. (2000). Implications of the Globalization of the Banking Sector: The Latin American Experience. *New England Economic Review,* Sept./Oct., 45-62.

Prevost, A. K., Rao, R. P. & Hossain, M. (2002). Determinants of board composition in New Zeland: a simultaneous equations approach. *Journal of Empirical Finance*, 9(4), 373-397.

Prowse, S. (1997). The Corporate Governance System in Banking: What do We Know? *BNL Quarterly Review*, March, 760-781.

Rebeiz, K. S. & Salameh, Z. (2006). Relationship between governance structure and financial performance in construction. *Journal of Management in Engineering*, 22(1), 20-26.

Sanda, A. U, A. S. Mukaila & T. Garba. (2005). *Corporate Governance Mechanisms and Firm Financial Performance in Nigeria*, Final Report Presented to the Biannual Research Workshop of the AERC, Nairobi, Kenya, 24-29.

Sanusi, L. S. (2010). The Nigerian Banking Industry: What Went Wrong and the Way Forward. *A Convocation Lecture Delivered at the Convocation Square, Bayero University, Kano, on Friday 26 February, 2010 to mark the Annual Convocation Ceremony of the University.*

Shleifer, A. & R. W. Vishny. (1997). A survey of corporate governance. *Journal of Finance,* 52, 737-783.

Sundaram, A. K. & A. C. Inkpen. (2004). The Corporate Objective Revisited. *Organization Science,* 15(3), 350-363.

Tariq, W, I. Ali, M. Ibrahim, M. Asim & N. U. Rehman. (2014). Theory and empirical evidence on corporate governance from banking sector of Pakistan. *Business and Economic Research,* 4(1), 163-174.

Tricker, R. I. (1984). Corporate Governance: Practices, procedures and powers in British companies and their boards of directors, Gower.

Unadiale, O. M. (2010). The Impact of Board Structure on Corporate Financial Performance in Nigeria. *International Journal of Business and Management,* 5(10), 155-166.

Weisbach, M. S. (1988). Outside Directors and CEO Turnover. *Journal of Financial Economics,* 20, 431-450.

Yermack, D. (1996). Higher market valuation of companies with a small board of directors. *Journal of Financial Economics, 40(2), 185-211.*

Yasser, Q. R. (2011). Corporate governance and performance: An Analysis of Pakistani listed firms. *Global Journal of Management and Business Research,* 11(10).

Yasser, Q. R. & Entebang, H., Mansor, S. (2011). Corporate governance and firm performance in Pakistan: The case of Karachi Stock Exchange (KSE)-30. *Journal of Economics and International Finance,* 3(8), 482-491.

Zingales, L. (1998). *Corporate Governance*, *the New Palgrave Dictionary of Economics and Law*. MacMillan, London.