LUND UNIVERSITY
School of Economics and Management

Predicting Bank Insolvency with Random Forest Classification

Brynjar Harðarson

Michael Vuono

A thesis submitted in June 2019 for the degree of

Master's in Finance

Supervisor: Anders Vilhelmsson

Abstract

The aim of this paper is to evaluate a machine learning technique, Random Forest, to predict default rates for banks in the United States. This study extends the findings of a Random Forest model first introduced by Petropoulos et al. (2017) by extending their model by evaluating a longer sample period and adding macroeconomic variables to analyze how current market conditions impact the prediction of default rates of U.S. Banks from 1994-2016. Petropoulous et al. (2017) evaluated multiple traditional and artificial intelligence models to find that Random Forest produced the best results utilizing quarterly data from the FDIC from 2008-2014. Numerous studies have suggested that the financial condition of banks is purely determined by bank specific variables. Our empirical results confirm that theory as a bank default prediction model utilizing Random Forest classification performed worse with the addition of macroeconomic variables when compared to a model based purely on bank specific variables.

Keywords: Random Forest, Forecasting Defaults, CAMELS, Macroeconomic Variables

Acknowledgements

We want to gratefully acknowledge the inputs and feedback from our supervisor Anders Vilhelmsson and Dinh-Vinh Vo throughout the process.

Table of Contents

List of Figures

List of Tables

1 Introduction

Concern and instability in the financial markets has traditionally triggered two main consequences in the banking industry, a higher percentage of bank insolvency and the introduction of regulators. Regulators must leverage limited resources to minimize the probability of default or overall impact of future bank defaults. To improve efficiency and effectiveness of their actions, regulators continuously improve Early Warning Systems (EWS) to better predict banks at risk of default. One such approach that has gained more attention in empirical studies over recent years is a model based on Random Forest, a non-parametric machine learning technique. The goal of this study is to evaluate the Random Forest approach to predict future bank defaults using both bank specific and macroeconomic variables.

We build upon a recent study by Petropoulos et al. (2017) by examining their robust Random Forest model after the addition of macroeconomic variables. Additionally, we evaluate the model over an extended time period to better account for shifting overall market environments. To our knowledge, no literature has evaluated the combination of a Random Forest model based on bank specific and macroeconomic variables. In the following sections, we illustrate the overall success in past studies of the Random Forest method compared to other approaches and the significance of macroeconomic variables in bank default prediction models. By testing the predictive power of a Random Forest model including direct measures of sensitivity to the overall market, our analysis will introduce a bank failure prediction model that will benefit the current research and regulators ability regarding predicting bank defaults.

One such regulating body that evaluates banks and their probability of default is the Federal Deposit Insurance Corporation (FDIC) in the United States of America. The FDIC was created with the passing of the Banking Act of 1933 after the stock market crash in 1929, which caused approximately 9,000 banks to suspend operation, and the closure of 4,000 banks in early 1933. Since, the FDIC has continuously improved policies and procedures to preserve and promote the public's confidence in the U.S. financial system. One such improvement was the introduction of EWS in the late 1970s. The purpose of EWS is to identify and then target limited resources to banks at risk of default to better maintain stability in financial markets and potentially reduce the expected cost of a bank failure. Most bank predicting models (including ours) have been related to the FDIC's original EWS model from the 1970's.

Figure 1 below shows the bank failures from the FDIC database over time. Clearly, bank defaults in the United States have been concentrated around three significant time periods. The first in the late 1930s, the second (and most significant) was the banking crisis of 1985-1992, and the last being the Great Financial Crisis starting in 2008. Each period indicated limitations to regulators current procedures, which ultimately led to changes to the regulatory environment. Given the systematic risk in the market during these prominent periods, Figure 1-1 supports our theory that the combination of the overall health of the economy and bank specific variables plays a key role in predicting future bank defaults.

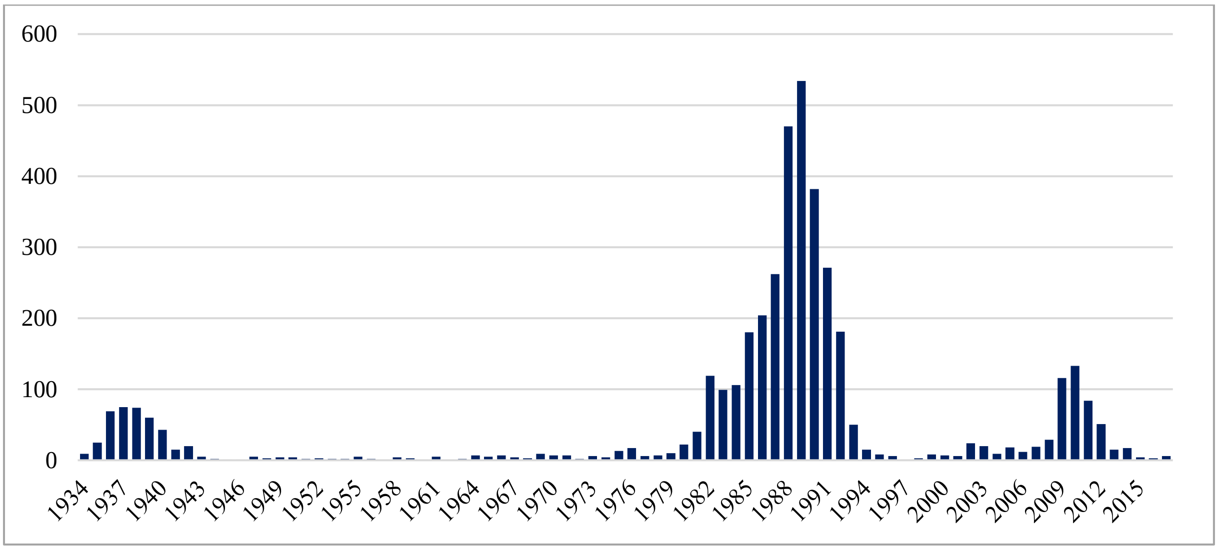

Figure 1-1: United States Bank Failures over Time (1934-2017)

Globally, other methods to evaluate banks have been established such as Financial Soundness Indicators (FSI) by the International Monetary Fund or the Basel Committee on Banking Supervision (BCBS). BCBS was established at the end of 1974 following instability in international currency and banking markets. The main goal of the Committee is to enhance financial stability by improving the quality of banking supervision worldwide. The Basel Capital Accord was published in 1988 to create consistent standards across all countries. The main focus of the Basel Capital Accord was a bank's capital adequacy, as the Committee feared the deterioration of capital ratios at the time. Similar to the FDIC, the Basel Committee has adjusted its policies to better reflect the true measurement of risk. Basel II was released in 2004 to account for the changing environments in the banking industry. During the Great Financial Crisis, Basel II proved to be an inaccurate measure of banks at risk of default considering there were a large number of banks that were in a worse position than expected. Therefore, Basel III was introduced to update the regulation once again. Basel III added a leverage ratio and liquidity requirements that was previously not accounted for in Basel I and Basel II. Along with the traditional bank specific variables, macroeconomic elements were added to the regulation to better account for systematic across the industry.

Given a multitude of factors including the variety of regulators, types of banks, dynamic financial markets, and the relative rarity of bank defaults, much analysis has been centered on what type of method provides the best estimator for a troubled bank. To assist authorities in monitoring and assessing the economic health of financial institutions, statistical methods continue to be leveraged to guide supervisory authorities. In Petropoulos et al. (2017), the authors evaluated the performance of multiple bank insolvency prediction methods to find Random Forest as the best performing model. Random Forest, due to its ability to handle numerous underlying features in a dataset, produces high quality prediction rates for classification questions. In our analysis, we utilized the *randomForest* package available in *R*, a free programming language used for statistical calculations. The aim of our paper is to extend the model introduced by Petropoulos et al. (2017) by implementing a Random Forest model including bank specific *and* macroeconomic variables to calculate the models ability to predict bank defaults given an extensive time period to capture multiple business cycles.

The remainder of this paper is organized as follows. We first discuss and evaluate previous literature describing the techniques and success of different bank failure prediction methods. Based on our literature review findings, we further analyze studies that include sensitivity to the overall market through the addition of macroeconomic variables. Section 3 focuses on the method of Random Forests, its previous applications, and main benefits. Section 4 and 5 details our data, data transformations and model developments needed to optimize each model. Section 6 discusses our validation measures and results. We then discuss the main empirical findings. Section 8 contains our conclusion and finishing thoughts.

2 Bank Failure Prediction Models – Literature Review

2.1 Application of CAMEL(S) Framework

Supervisory authorities are mainly concerned with bank specific weaknesses that drive banks into insolvency. Most models are evaluated based on the CAMEL(S) framework including the FDIC's EWS model introduced in the 1970s. The initial model evaluated different financial and accounting ratios that fell into 5 categories: capital adequacy (C), asset quality (A), management competence and expertise (M), earnings ability and strength (E), and liquidity (L). The model then scored firms 1-5 based on the relevant ratios. 1 indicated a stable (good) firm and 5 being a major concern that may require supervisory intervention. From each individual score, an aggregate score would be calculated and utilized by supervisory authorities. Leveraging the CAMEL(S) framework, many subsequent models have benefitted by following a similar framework pioneered by the FDIC.

Many studies and models have found strong empirical success following the CAMEL framework by focusing on individual bank-level data as the driver of bank failure predication models. For example, literature by Berger & Bouwman (2013) and Vazquez & Federico (2015) suggested that the financial condition of individual banks was the key driver in distinguishing performance during the Great Financial Crisis. Cole and Wu (2009) used a simple time-varying hazard model and a simple static probit model to provide empirical evidence in support of supervisory authorities' use of CAMELS ratios from bank financial data.

2.2 Machine Learning to Predict Bank Failures

Over recent decades, more literature and empirical evidence has been published regarding different model approaches to efficiency analyze and identify banks in distress. Demyank and Hasan (2010) evaluated and summarize the different prediction methods available and tested. Model types range from traditional approaches, such as discriminant analysis, first introduced by Altman (1968), and logit/probit regressions, introduced by Ohlson (1980), to more advanced models based on artificial intelligence such as Neutral Networks, Support Vector Machines, and Random Forests.

Petropoulos et al. (2017), from the Bank of Greece, fill a gap in academic research by testing all the above model types simultaneously on a common dataset. The authors test the traditional and machine-learning techniques to produce a model that best distinguishes characteristics of U.S. banks that failed or received financial assistance from 2008-2014 based on quarterly data from the FDIC. Their empirical evidence finds that the method of Random Forest, a machine learning technique, demonstrates superior predictive performance compared to traditional bank failure methods and other advanced machine learning techniques.

Messai and Gallali (2015) compare early warning systems for European Banks from 2007-2011. Based on their data, they test discriminant analysis, logistic regression, and artificial intelligence. The artificial intelligence method they test is Neutral Network, which they find to be the best performing model to predict future bank distress. Along with the strong theoretical backing and Petropoulos et al. (2017) empirical evidence, this provides additionally evidence to the predicting power of artificial intelligence models in forecasting bank failures. Similar to Petropoulos et al. (2017), this study leverages only bank specific details with CAMEL indicators to predict bank defaults.

2.3 Systematic Risk in Bank Failure Prediction Models

In the 1990s, the FDIC added a sixth category to the CAMEL framework: sensitivity to market risk (S). This additional category can account for a bank's sensitivity to systematic risk across the financial industry. Although not a direct exposure to overall market risk, Flannery (1998) was one of the first to add market information to the EWS approach by using market expectations from stock prices, volatility, and bond spreads. He found that bond ratings are a good proxy for bank condition and equity market volatility is a good predictor of bank holding company performance. However, many banks are not publicly traded which may limit the effectiveness of using market-based information sources. Therefore, other macroeconomic variables such as GDP growth, inflation, inter-bank interest rates, or exchange rates may better capture economic pressure or shocks that could trigger bank failures.

Other studies have included macroeconomic variables to predict bank defaults. Kaminsky and Reinhart (1996) evaluate different macroeconomic variables as leading indicators of a banking crisis. Low GDP growth, decline in stock prices, and high real interest rates were shown to perform as the best signals. A concern regarding macroeconomic variables is the distinction between being a leading or lagging indicator to a banking crisis. Their research evaluated and indicated that a decline in GDP growth tends to precede the onset of a banking crisis by about 8 months.

Demirguc-Kunt and Detragiache (1997) evaluated systemic banking crises across developed and developing countries in 1980-1994. Their empirical evidence found common factors across banking crises such as low GDP growth and high inflation. Additionally, they find that systemic banking sector problems are linked to excessively high real interest rates. This environment negatively impacts a bank's balance sheet as they are forced to increase the interest paid to depositors while the asset side of the bank's balance sheet consists of long-term loans that cannot adjust to the new environment. The authors argue that even if the banks can pass on the higher rates to borrowers that will increase the rate of non-performing loans.

Arena (2008) evaluated bank failures in Latin American and East Asia during the 1990's to evaluate how bank specific data impacted default rates and how systematic macroeconomic and liquidity shocks impacted the banking sector. The results explain that bank specific data partly explain bank defaults while systematic macroeconomic and liquidity shocks destabilized all banks in terms of their financials. To account for macroeconomic conditions, they evaluated the banking system liquidity to capture potential contagion effects, volatility in real effective exchange rates, and economic activity measured by GDP growth.

By looking at banks in the United States, the following two research papers found macroeconomic variables to be significant when predicting bank defaults. Cole and White (2012) found macroeconomic variables such as real-estate loans and mortgages to be leading indicators when examining US commercial bank defaults in 2009. Mayes and Stremmel (2014) incorporate the CAMELS framework to consider both bank specific variables and macroeconomic conditions in a logistic regression and a discriminant analysis based on Cox's proportional hazard model. The authors included real GDP to the FDIC insured banks over the time period of 1992 to 2012. GDP, along with other CAMELS variables, were significant in the logit model tests and was one of the few coefficients significant in the Cox proportional hazard model.

3 Random Forest

3.1 *Theory and Main Benefits*

Random Forest's is a machine learning technique used to model classification problems given a collection (or forest) of binomial trees based on a training set of data. Breiman (2001) first introduced the Random Forest framework. Random Forest includes an efficient calculation algorithm to better analyze big datasets with many input variables without correlation restrictions and can recognize non-linear relationships in the data.

Kartasheva and Traskin (2011) effectively lay out the main benefits of Random Forest compared to other methods. First, Random Forest is more insensitive to tuning parameters of the models as shown by Breiman (2001). Random Forest also reduces the need to sub-sample or shrink the data by effectively evaluating the unbalanced and non-linearity nature of large datasets (like bank defaults as there is a much larger percentage of solvent banks for each given observation period). Lastly, Random Forest presents a variable importance plot, which ranks each variable's "importance" to the model's predictive power. The model evaluates the change in prediction accuracy if the variable was removed. Therefore, highly ranked variables will significantly decrease the predictive accuracy of the model when removed. We display and discuss the importance plots in our model in Section 5.2 – Model Development.

An example of a variable importance plot is shown below in Figure 3-1. Based on this simplified example from the DnI Institute (2015), the variables are ranked based on the *Mean Decrease Accuracy* and *Mean Decrease Gini*. The *Mean Decrease Accuracy* measures the decrease in prediction accuracy if the variable was removed from the model. The *Mean Decrease Gini* is a measure of the decrease in a variable's total Gini (or node) impurity. Gini impurity is the probability of an observation being incorrectly classified at a given node based on the training dataset of the model. Since Gini impurity is evaluated at each node, the importance plot considers the average Gini impurity across all the decision trees generated. As shown below in Figure 3-1, the variable labeled *duration* is highly significant based on the *Mean Decrease Accuracy* and *Mean Decrease Gini*.

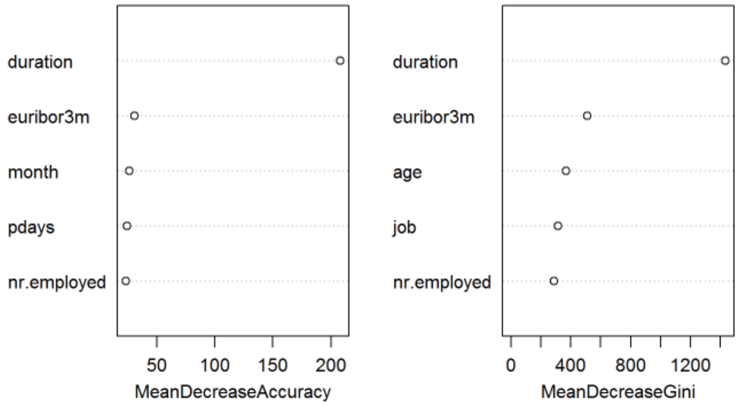

Figure 3-1: Variable Importance Plot Example

The method of Random Forest is to test each new observation characteristic as a vector through the created decision trees to predict the observation's classification. The decision trees are generated through a bootstrap sample of the train data. Based on this process, the binomial trees are created where each split creates the "branches" and the terminal nodes represent the final "leaves" in the Random Forest. Each "branch" of the decision represents a true/false classification question for the observation that continues to the final "leaf." Then, the out-of-sample (or Out-Of-Bag (OOB) sample) data, that was not included in the bootstrap sample, is classified based on the trees. As shown in the simplified example in Figure 3-1 above, if the OOB data is based on *N* number of decision trees, each observation in the sample will then be tested *N* times. Given the characteristics of the OOB observations, a vector will be created that estimates the classification for each tree. Random Forest then uses a majority vote of the predicted values to determine how the observation should be classified.

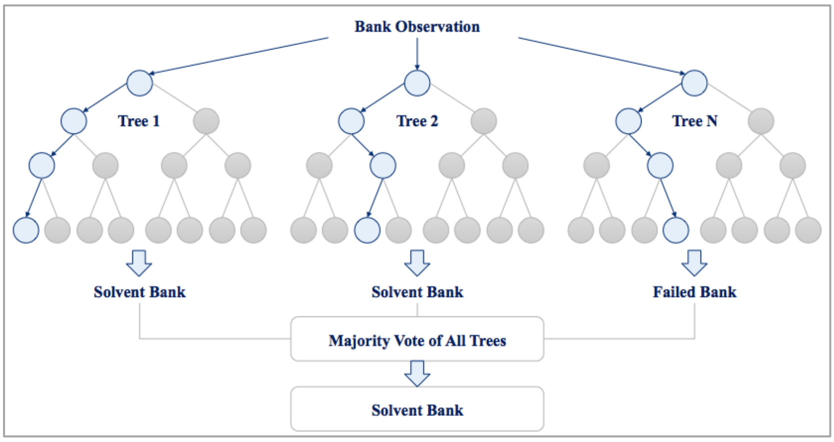

Figure 3-2 Random Forest Simplified Example

The Random Forest method has unique features that may influence the classification of the OOB sample and the variable importance. One such feature is *mtry*, which classifies the number of variables available for splitting at each tree node. Research by Cutler et al. (2007) and Strobl et al. (2008) show conflicting evidence based on *mtry*'s impact on their model, but Breiman (2001) suggests *mtry* equals either $(0.5)\sqrt{\mathrm{m}}, \sqrt{\mathrm{m}},$ or $(2)\sqrt{\mathrm{m}}$ with m equaling the number predictive variables in the model. Another feature is *ntree*, which identifies the number of trees grown to create the forest. The preference is to have $ntree = \infty$ as a larger number of trees increases the stability of the model. However, this is not realistic, as it would increase the memory needed and run time to test the model. If *ntree* is too low, this could negatively impact the correlation between trees and overall strength of the trees. Therefore, common values for *ntree* range from 50 – 500 as there is a diminishing value to each number of trees added. Although best practice is to tune the models based on these features, a benefit of Random Forest compared to other methods is the insensitivity to these types of measures.

3.2 Random Forest Applications in Finance

Until recently, there has been limited use of Random Forest in practice. Denil et al. (2013) tried to narrow the gap between the strong theoretical foundation detailed above and the limited applications utilizing Random Forest. The study displayed strong consistency in the results of Random Forest.

Building upon their analysis, the method of Random Forest has gained momentum as it has been applied to generate successful predictions from a variety of areas ranging from medicine to sports. Gurm et al. (2014) predicted transfusion risk regarding the impact of bleeding avoidance strategies during surgery while Lock and Nettleton (2014) apply a Random Forest model to predict the outcome of individual plays and ongoing win probability for National Football League (NFL) games. The increasing use of Random Forest as a predictive tool has been driven by its strong results while containing distinctive benefits.

Random Forest has received increased attention in the finance industry to improve prediction accuracy in a variety of topics. One such area that has benefitted from Random Forest has been the benefits of automated trading in equity trading. Booth et al. (2014) produce an automated trading system based on a Random Forest method to improve profitability and stability of trading seasonality events. They also find that Random Forest produced superior results in terms of profitability and prediction accuracy compared to other techniques. Khaidem et al.

(2016) use Random Forest to predict the returns in stock market prices to limit overall risk in an investment.

Random Forest has also been used to assess credit risk in individual firms. Yeh et al. (2016) used a hybrid Random Forest method with a rough set theory approach to increase the accuracy of going-concern predictions. Their empirical evidence displays the best classification rate and lowest Type I and II errors. Wu et al. (2016) found that applying a Random Forest approach to predict credit ratings was the most effective.

4 Data & Data Transformations

4.1 Bank Specific Data

To replicate and extend the Petropoulos et al. (2017) model using their Random Forest method, we collected our data from the same source, the Federal Deposit Insurance Corporation (FDIC). The FDIC maintains data on all FDIC insured banks and thrift institutions in the United States dating back to its inception in 1934 and bank specific data back to 1992. Additionally, the FDIC details all banks that either failed or required assistance transactions. Therefore, we define a bank failure (or "Bad Event") in our data as any bank that failed or received assistance to be consistent with the FDIC definition. In our model, a bank's previous reporting prior to a "Bad Event" will be indicated by 1 and 0 if not. We gathered data from the database on all available firms for the model to compare the indicators of banks that remained solvent and banks that did not. We first started with quarterly data from FDIC for the sample 2008-2014 to replicate Petropoulos et al. (2017) with the goal of then extending the data to incorporate multiple business cycles.

We first identified 8,602 unique banks and evaluated 210,213 observations based on quarterly data over the sample period 2008-2014. As shown in Figure 4-1, the number of observed banks decreases linearly over time as banks fail or consolidate in the industry. A total of 445 banks were impacted by failures or assistance transactions resulting in 5.2% of unique banks failing over the sample. Over 50% of the bank failures during the period were concentrated to 2009-2010. The sample period accounts for approximately 11% of the total failed banks dating back to the creation of the FDIC in 1934 but is the second largest period impacted behind bank failures behind only the banking crisis during 1985-1992.

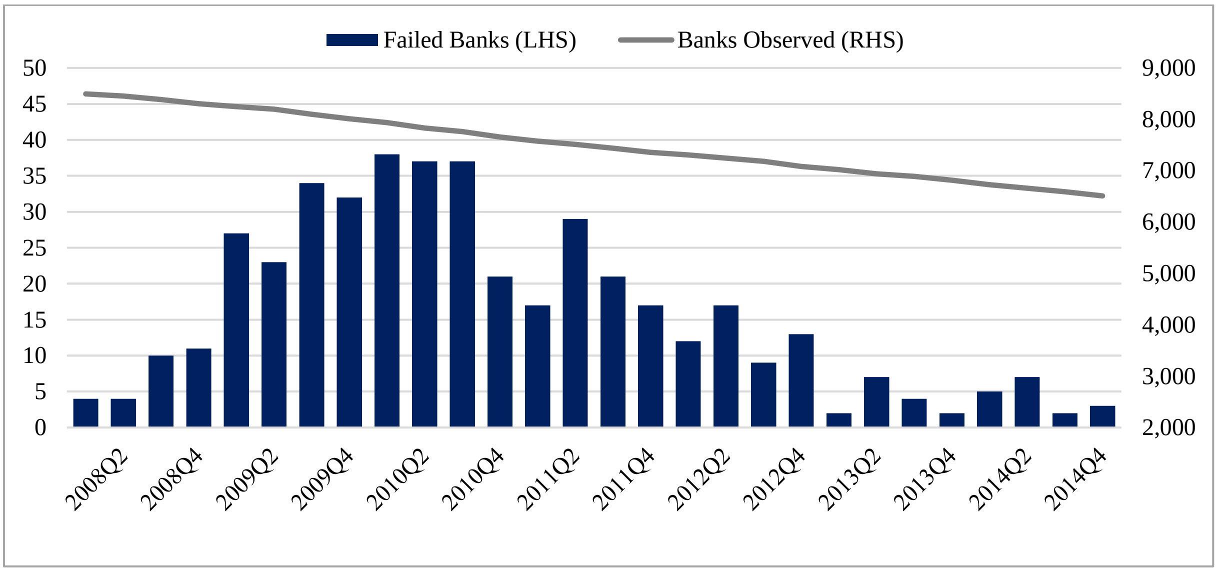

Figure 4-1: Total Banks and Failure: 2008-2014 (Quarterly Observations)

Given our goal of extending Petropoulos et al. (2017) and incorporating the impact of macroeconomic variables, we therefore increased our sample period from 2008-2014 to 1994-2016. One aspect we considered when deciding on a proper sample size was the FDIC database only includes annual reporting from 1992 to 2000 and quarterly thereafter. Therefore, we had to evaluate the cost/ benefit between having more observations using the quarterly data from 2000 or a longer time sample using annual data. We chose the latter as our main goal is to evaluate bank defaults over multiple time periods (and business cycles). This decision also resulted in the addition of more "Bad Event" observations compared to the bank defaults from 2000 to present.

Using annual observations from 1992 to 2016, we observed 16,504 unique banks and a total of 232,188 total observations. As shown in Figure 4-2, the number of observations decreases over time similarly to Figure 4-1 above. Considering the total data sample, we account for 662 "Bad Events" and 599 in our final observation range of 1994-2016. The concentration to defaults during the Great Financial Crisis as shown in Figure 4-1 is emphasized when compared to the previous 15-year period.

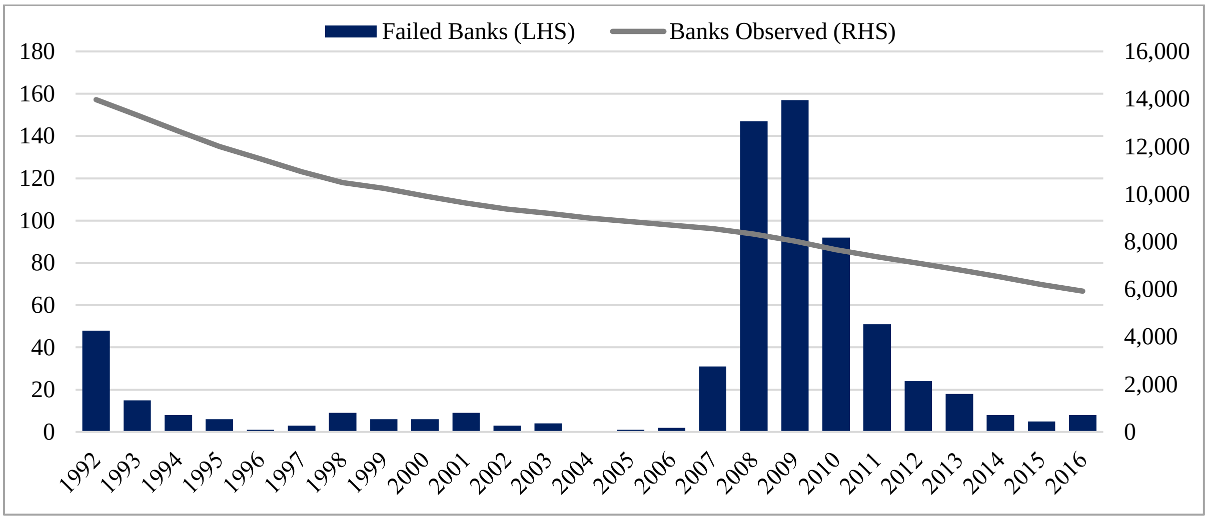

Figure 4-2: Total Banks and Failure: 1992-2016 (Annual Observations)

Each bank accounted for in the FDIC database is assigned a regulator. The regulators include the FDIC, Federal Reserve Board (FRB), Office of the Comptroller of the Currency (OCC), or Office of Thrift Supervision (OTS). The FRB is responsible for members of the Federal Reserve System while the FDIC regulates all remaining state-charted bank that are not members of the Federal Reserve System. The OCC regulates nationally chartered commercial banks. The OTS was responsible for federally chartered thrifts but was merged into the OCC on July 21, 2011. In our sample, the largest allocation of banks and "Bad Events" fell under the supervision of the FDIC at 55.1% and 59.5%, respectively. Of the 9,088 FDIC regulated banks observed over our sample period, 4.3% of the banks failed or received assistance. This was higher than FRB and OCC/OTS banks, which saw 3.1% and 3.7% of their respective banks, defined as "Bad Events."

An interesting aspect of the data is the large percentage of defaults allocated to a relatively small number of states in the United States over the sample period. Four states in particular, Georgia, California, Florida, and Illinois, account for over 50% of the bank defaults during 1992-2016. This is despite these four states only making up 20.9% of all unique banks. Each state saw their defaults be concentrated to the Great Financial Crisis.

4.2 Bank Specific Variables

The independent variables used in Petropoulos et al. (2017) are based on the CAMELS (i.e. Capital, Asset Quality, Management, Earnings, Liquidity, and Sensitivity to market risk) framework. In total, 44 bank specific variables from the FDIC database were tested and transformed to create their model. Petropoulos et al. (2017) started with a total of 660 variables, which compromises transformations such as different lag lengths and first differenced variables. By evaluating the correlation with the dependent variable, a cross-correlation analysis, LASSO process, and the Random Forest variable importance, the authors reduced the independent variables to 23 variables used in the Random Forest model. The last step in the process was to evaluate the importance of each individual variable in the Random Forest model, which is a key benefit of the approach and distinguishes Random Forest from other machine learning techniques.

The book value of equity is included with three different transformations in the model. One of the transformations includes a one-quarter lag and is therefore excluded in our model as we utilize annual observations. Table 4-1 (below) displays the 22 independent variables that remained. Loss Allowance to Loan is accounted for 3 different times in the model after considering the importance of transformations. This, along with Average Equity, causes the CAMELS Category "Asset Quality" to be the most prevalent category in the model. The "Earnings" category is the second most common considering the inclusion of measures such as Return on Equity (ROE), Return on Assets (ROA), Retained Earning to Average Equity (RE_EQ), Net Interest Margin (NIM) and Cost of Funding Earning Assets (CFEA). Liquidity is only accounted for once by the Net Loan & Leases to Core Deposits (NLOAN_CDEP).

Table 4-1: Independent Variables Added to Random Forest Model

Independent Variable	Following Transformation	CAMELS Category
Average Equity	d_log_equity_lag1_DFS	Asset Quality
Average Equity	PCT_log_equity_lag1_DFS	Asset Quality
Capital Adequacy Ratio	CAR	Capital Adequacy
Core Capital (Leverage) Ratio	LEV	Capital Adequacy
Core Capital (Leverage) Ratio	d_LEV_lag1	Capital Adequacy
Cost of Funding Earning Assets	CFEA	Earnings
Efficiency Ratio	EFF_DFS	Management Capability
Loss Allowance to Loan	d_Loss_Loan_lag1_DFS	Asset Quality
Loss Allowance to Loan	Loss_Loan_DFS	Asset Quality
Loss Allowance to Loan	Loss_Loan	Asset Quality
Loss Allowance to Noncurrent Loans	Loss_NPL	Asset Quality
Net Interest Margin	NIM	Earnings
Net Loan & Leases to Core Deposits	NLOAN_CDEP	Liquidity
Net Operating Income to Assets	NOI_ASS	Management Capability
Net Operating Income to Assets	d_NOI_ASS_lag2	Management Capability
Noncurrent Assets plus Other Real Estate Owned to Assets	d_NCASS_ORE_lag1	Asset Quality
Noncurrent Loans to Loans	NPL	Asset Quality
Retained Earnings to Average Equity	RE_EQ	Earnings
Retained Earnings to Average Equity	d_RE_EQ_lag1	Earnings
Return on Assets	d_ROA_lag1	Earnings
Return on Equity	ROE_DFS	Earnings
Return on Equity	ROE	Earnings

From our sample of annual observation from 1994 to 2016, Table 4-2 shows the resulting descriptive table of the variables shown in Table 4-1. The table represents the final statistics of the bank specific data as the FDIC data required modifications to correct for issues such as observations with a large percentage of blanks and account for outliers by windorize the 1% and 99% quantile. DFS (or distance from sector) transformation accounts for the difference between the mean value for the bank's sector (as determined by the FDIC) and the bank's value each year.

As shown in the descriptive statistics table below, when comparing the mean and standard deviations of the good banks and distressed banks, there is a significant difference. For example, the mean for Capital Adequacy Ratio (CAR) for good firms is 18.98 while bad firms are almost half at 9.57. This significance is present across the bank specific variables to a varying degree. This makes intuitive sense as this high level review reinforces Petropoulos et al. (2017) key findings that the bank specific variables are significance. As expected, the good observations and whole sample are highly similar regarding the statistics considering the banks that did not experience a default or assistance transactions dominates the sample as there are only 599 Bank Bad observations.

Table 4-2: Descriptive Statistics Table of Bank Specific Variables 1994-2016

	Whole Sample (N=204,496)		**Good Observations (N=203,897)**		**Distressed Banks (N=599)**	
Variable	**Mean**	**St. Dev**	**Mean**	**St. Dev**	**Mean**	**St. Dev**
d_log_equity_lag1_DFS	9.37	1.31	9.36	1.31	9.53	1.60
PCT_log_equity_lag1_DFS	0.00	0.01	0.00	0.01	-0.02	0.01
CAR	18.96	11.66	18.98	11.67	9.57	2.80
LEV	10.82	5.23	10.83	5.23	5.65	1.67
d_LEV_lag1	-0.19	1.88	-0.18	1.87	-3.64	2.92
CFEA	2.46	1.36	2.46	1.36	2.65	1.16
EFF_DFS	0.00	23.61	-0.15	23.32	49.38	52.47
d_Loss_Loan_lag1_DFS	0.00	0.37	0.00	0.37	0.57	0.93
Loss_Loan_DFS	0.00	0.79	-0.01	0.78	1.81	1.43
Loss_Loan	1.47	0.86	1.46	0.85	3.53	1.52
Loss_NPL	589.53	1788.29	591.11	1790.67	49.69	74.55
NIM	4.08	1.00	4.09	1.00	3.11	1.14
NLOAN_CDEP	88.60	34.28	88.56	34.27	102.00	37.09
NOI_ASS	0.82	1.06	0.83	1.03	-3.01	1.85
d_NOI_ASS_lag2	0.04	0.92	0.04	0.90	-2.08	1.99
d_NCASS_ORE_lag1	0.03	0.88	0.02	0.86	2.53	1.99
NPL	1.30	1.81	1.28	1.76	8.36	3.21
RE_EQ	3.67	8.94	3.78	8.67	-34.18	15.42
d_RE_EQ_lag1	-0.39	8.82	-0.32	8.66	-25.55	20.04
d_ROA_lag1	0.02	0.73	0.02	0.72	-1.38	1.87
ROE_DFS	0.00	8.88	0.09	8.70	-30.99	14.24
ROE	8.71	9.55	8.82	9.30	-29.11	15.42

4.3 Macroeconomic Variables

Given the model explained above, all the CAMELS categories are accounted for except for the "Sensitivity to market risks (S)." Petropoulos et al. (2017) only evaluated one variable, Asset Fair Value, categorized as sensitivity to market risk. This underscores the fact that a bank's sensitivity to the overall market has been relatively neglected compared to the other CAMELS categories. Avkiran and Cai (2012) explain how this omission is due to the inability of capturing the relationship from accounting and financial data. Therefore, to better account for systematic risk in the financial system, we evaluated the addition of macroeconomic variables to the model.

Based on the literature review of the addition of macroeconomic variables and data availability, we evaluated the following 22 macroeconomic variables in the United States listed in Table 4-3. The macroeconomic variable data was mainly sourced from the Federal Reserve Economic Database (FRED) with the Gross Domestic Product (GDP) levels sourced from the Bureau of Economic Analysis (BEA).

Table 4-3: Macroeconomic Variables Evaluated

Macroeconomic Variable	Variable
Baa-10 Spread	Baa.10Spread
Bond Yield Spread: 30-1 Year	X30.1YrSpread
Bond Yield Spread: 10-1 Year	X10.1YrSpread
CBOE Volatility Index	VIX
Chicago Fed National Activity Index	CFNAI
Chicago Fed National Financial Conditions Index	NFCI
Consumer Credit	CC
Consumer Debt Service Ratio	CD.SR
Consumer Price Index	CPI
Crude Oil Prices: West Texas Intermediate (WTI)	OIL
Effective Federal Funds Rate	FEDFUNDS
Financial Obligations Ratio	FD.SR
Household Debt Service Ratio	TD.SR
Inflation Rate	INFL
Leading Index for United States	LEADING
Mortgage Debt Service Ratio	MD.SR
Real Gross Domestic Product Percent Change	GDP
Real Trade Weighted U.S. Dollar Index	RealTrade
St. Louis Fed Financial Stress Index	FinStress
TED Spread	TED
Unemployment Rate	EMPLOY
Wilshire 5000 Total Market Full Cap Index	WILL5000

We included a wide variety of traditional variables such as the level of GDP, Inflation, VIX, or Unemployment Rate as well and non-traditional economic variables such as the Chicago Fed National Activity Index (CFNAI) or Leading Index for the United States. These non-traditional economic variables test an aggregation of economic indicators that have been shown to be leading indicators for the economy in the United States. For example, the CFNAI is a weighted average of 85 monthly indicators of national economic activity in the U.S. that provides early indications of business cycle changes and inflationary pressures. We believe that the combination of the traditional and aggregated economic indicators will provide a robust analysis of the macroeconomic environment's impact on default probabilities.

Table 4-4: Descriptive Statistics Table of Macroeconomic Variables (1992-2016)

Macroeconomic Variable	Mean	Median	Min.	Max.	St. Dev.
Baa-10 Spread	2.44	2.33	1.30	5.62	0.89
Bond Yield Spread: 30-1 Year	2.00	2.32	-0.19	4.16	1.40
Bond Yield Spread: 10-1 Year	1.52	1.62	-0.29	3.38	1.16
CBOE Volatility Index	19.48	18.31	11.56	40.00	6.79
Chicago Fed National Activity Index	-0.24	0.15	-4.76	0.72	1.09
Chicago Fed National Fin. Conditions Index	-0.35	-0.53	-0.87	2.27	0.62
Consumer Credit	14.49	14.60	13.60	15.11	0.44
Consumer Debt Service Ratio	5.65	6.71	4.66	6.71	0.56
Consumer Price Index	193.61	191.60	142.80	243.78	32.41
Crude Oil Prices: West Texas Intermediate	46.37	34.31	12.52	100.27	29.41
Effective Federal Funds Rate	2.62	2.28	0.07	5.98	2.29
Financial Obligations Ratio	16.50	16.57	14.90	18.13	0.89
Household Debt Service Ratio	11.37	11.29	9.84	13.22	1.07
Inflation Rate	0.02	0.03	0.00	0.04	0.01
Leading Index for United States	1.24	1.54	-2.40	1.96	0.90
Mortgage Debt Service Ratio	5.72	5.63	4.39	7.21	0.73
Real Gross Domestic Product Percent Change	4.73	5.00	-7.20	9.30	3.23
Real Trade Weighted U.S. Dollar Index	94.63	93.98	83.27	112.19	7.97
St. Louis Fed Financial Stress Index	0.15	0.02	-1.23	3.37	1.04
TED Spread	0.52	0.47	0.17	1.41	0.34
Unemployment Rate	5.97	5.60	4.00	9.80	1.58
Wilshire 5000 Total Market Full Cap Index	45.12	41.58	12.13	100.27	29.41

To better understand the macroeconomic variables added and the potential overlap between variables, we evaluated the correlation matrix (shown in Appendix – Table A-1). Highlighted in the table is any correlation between variables greater than 0.8. The variables most highly correlated with others are the Wilshire 5000 Total Market Index and the Debt-Service Ratios. The Wilshire 5000 Index, a proxy for the overall health and level of the equity markets in the United States, is highly correlated with GDP log, CPI, and Consumer Credit (CC). Additionally, the debt service ratio measures, Total (TD.SR), Mortgage (MD.SR), Consumer (CD.SR), and Financial Obligations (FD.SR) are all highly correlated with each other. The only relationship with a correlation less than 0.7 is Mortgage and Consumer at 0.35. Non-traditional macro-economic variables including the Leading Index, CFNAI, and NFCI all have a correlation greater than 0.7 with GDP. Since these measures are aggregate variables of economic health, this is not surprising considering GDP is an input in the calculations.

Along with the raw level of the macro variables, we also added lagged values and first differences (measured as proportional change) of different lengths to the model. The model evaluates one, two, three, and four-year lagged values and first differences. Lagged values of macroeconomic variables that show strong trend (shown in Appendix – Figure A-1), such as consumer credit and consumer price index, are excluded and are therefore only included as proportional change, first differenced values. Transformed macroeconomic variables with invalid reference are excluded (see final transformations applied in Appendix – Table A-2). The goal of adding the transformations is to have the model incorporate any trends in the overall market that may be underlying the market conditions. For example, if the unemployment rate trended down, this would indicate an improved economy, leading to less bank defaults.

4.4 Sample Selection

As mentioned previously, the FDIC database includes annual data dating back to 1992. One of the bank specific variables includes a two-year lagged value and therefore, our total model sample will be based on observations from 1994 to accurately present the lags in the model. Since there are no defaults in FDIC data in 2017 and 2018 these two years are not included in our sample. Therefore, our total sample period is from 1994 to 2016. Although one of Random Forest's main benefits is its ability to handle large, unbalanced datasets, we configured our Total Sample of bank observations into the below subsamples. This allows us to take a similar approach as Petropoulos et al. (2017) while allowing the dataset to be easily implemented into other prediction model methods at a later date. Figure 4-3 below shows how we divided our total sample to train and validate our Random Forest model.

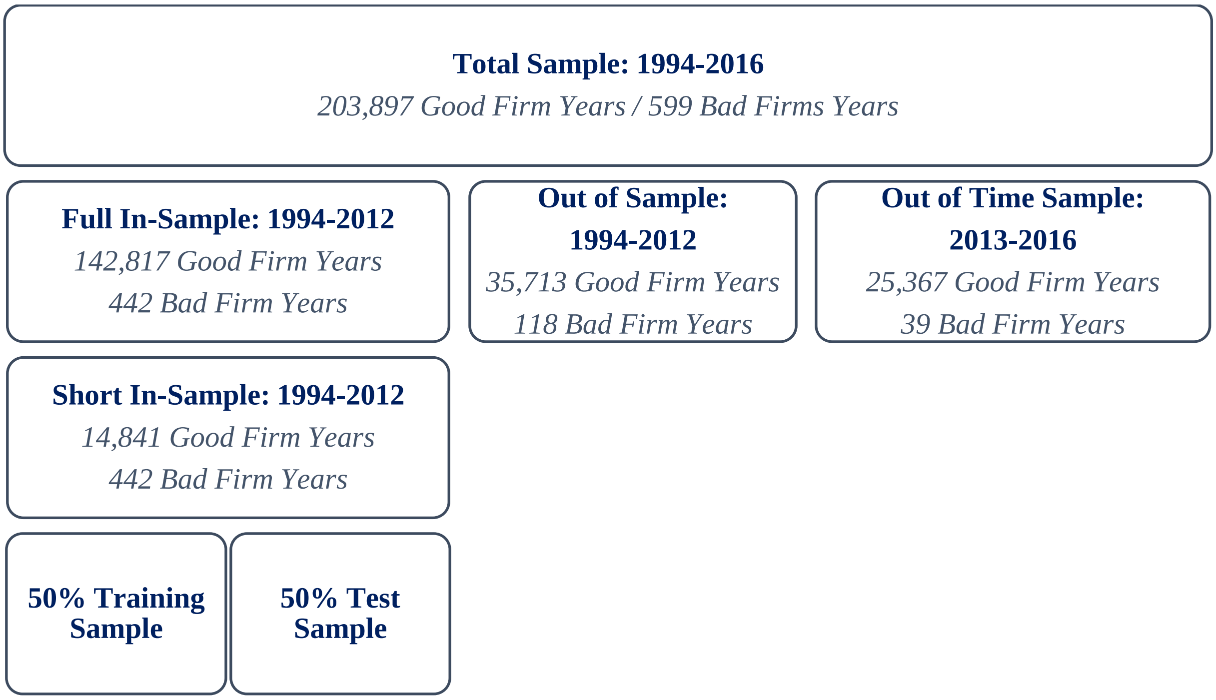

Figure 4-3: Model Samples

From the total sample, we first divided it into three samples. We split all yearly bank observations from 1994-2012 by an 80/20% to create our Full In-Sample and Out of Sample with the remaining observations from 2013-2016 as our Out of Time Sample. Given our Full In-Sample is highly concentrated with good firm year observations, we randomly selected 10% of the good firm years with all 442 bad firm year observations. The main reason to perform this step is to verify that the good firms don't suppress the "Bad Events" in the model. We then split the data 50/50% into a Training and Validation Sample as a robustness check of our Random Forest model.

The total sample includes 599 Bad Firm Years (or observations). The bad observations are based on all banks impacted by default or assistance transactions from the FDIC database over the sample. The number of bad events is in line with Mayes and Stremmel (2014) who evaluated bank defaults from 1992 to 2012 despite Petropoulos et al. (2017) displaying a higher number of bad events in their dataset.

5 Models & Model Development

5.1 Models Tested

The goal of our research is to effectively predict the solvency of banks by binary classification from a Random Forest approach. Ultimately, we are evaluating if the addition of macroeconomic variables improve the predictive power of identifying bad bank observations using a Random Forest model. Therefore, we are comparing the prediction accuracy between models with only bank specific variables and bank specific plus macroeconomic variables. We developed, trained, and validated the models to evaluate which of the models performed best during the sample period and out of time periods.

5.2 Model Development

To effectively develop our models, we analyzed multiple factors such as the variable importance, *ntree*, and *mtry* as detailed in Section 3.1. Analyzing the variable importance is key to model optimization to avoid generalization between variables. Breiman (2001) explains Random Forest models perform better given a lower generalization error. Generalization is measured by determining the model variables correlations and overall strength. Therefore, a lower generalization error is defined as a model that includes variables with a lower correlation between each other and higher strength. By removing correlated and unimportant variables, the inter-tree correlation improves resulting in a stronger predictive ability.

As shown in Figure 5-1 and 5-2 below, we first evaluated the variable importance plots for the bank specific only model and the bank specific model plus all macroeconomic variables. Once accounting for all the variations of macroeconomic and the bank specific variables, there was a total of 185 variables tested in the combined model.

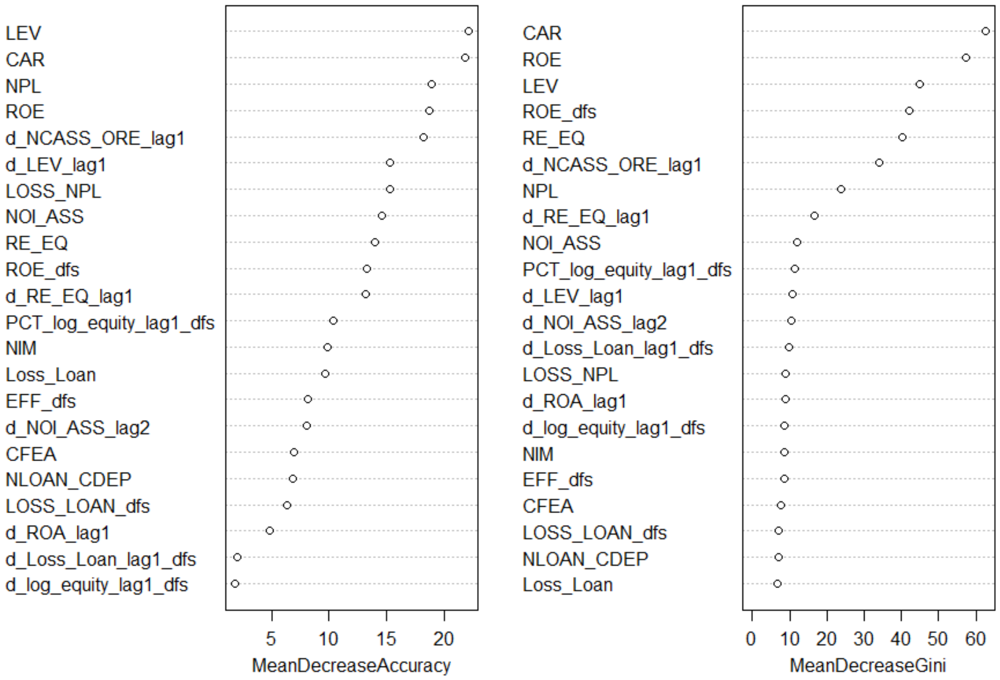

Figure 5-1: Variable Importance Plot (Bank Specific Only)

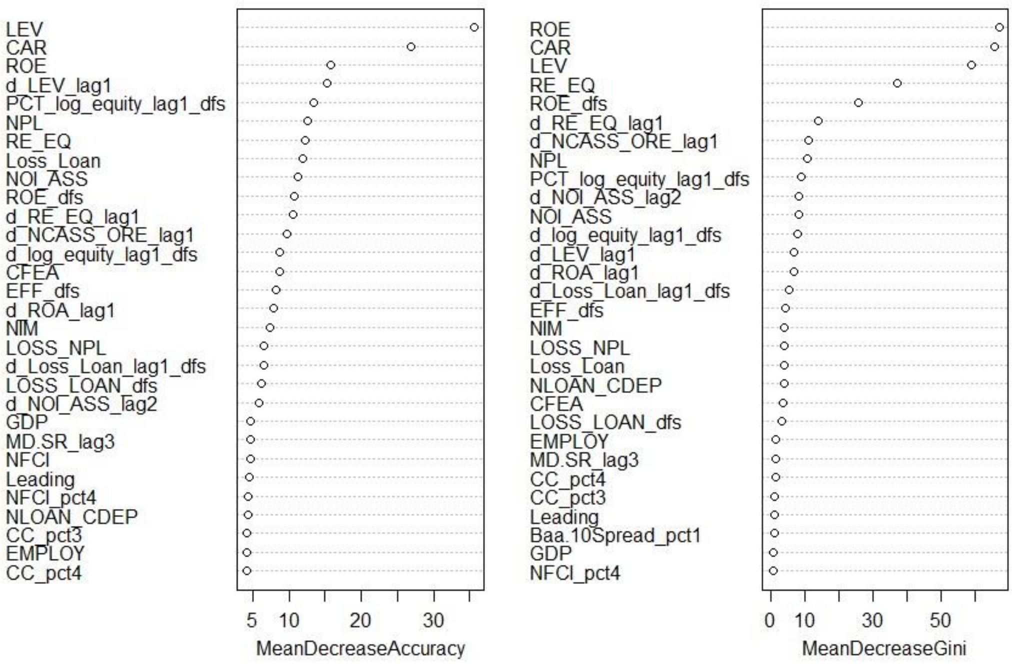

Figure 5-2: Variable Importance Plot (Bank Specific + All Macro)

Petropoulous et al. (2017) evaluated the bank specific variable importance plot resulting in the final bank specific variables we have implemented in our model. Therefore, we will focus on the importance of macroeconomic variables. At first glance, the importance plot of the combined model (Figure 5-2) shows that the bank specific variables are driving the prediction model based on importance. The first macroeconomic variable to appear on Mean Decrease Accuracy plot is the GDP level. The macroeconomic variables are shown to have little importance when compared to the bank specific variables as the highest important macroeconomic variables (GDP, MD.SR_lag3, NFCI, and Leading) are the only greater than a one bank specific variable, Net Loan & Leases to Core Deposits (NLOAN_CDEP). However, with the large number of lags and first differenced variables, it is highly possible that the macroeconomic variables are being generalized in the current model. Therefore, we evaluated removing the least important variables according to the importance plot while considering the correlation matrix discussed in Section 4.3.

To remove the unimportant variables in the model, we focused our attention on the left column of the importance plot in Figure 5-2 as Genuer et al. (2010) explains that the Mean Decrease Accuracy is the most prevalent used score of importance when using Random Forest. By using the Mean Decrease Accuracy, Genuer et al. (2010) evaluated the sensitivity of variable importance based on variable selection and other factors. One such variable selection process detailed in Díaz-Uriarte and Alvarez de Andrés (2006) is based on recursive elimination of variables. While consider the above importance plot, we also implemented a similar approach by evaluating the impact of removing 20% of the less important variables and re-running our model with the decreased number of variables. By continuing this process and evaluating the model's performance, we arrived at our "Final Model". Figure 5-3 shows the importance plot of the Final Model, which includes all the bank specific variables plus a revised list of macroeconomic variables.

Based on our analysis GDP, Leading Index, Mortgage Debt Service Ratio lagged three years (MD.SR_lag3) and NFCI were the remaining macroeconomic variables in the model. Each macroeconomic variable is shown as being moderately important based on the plot. Although not shown as a most important variable, MD.SR_lag3 is the tenth most important variable, which is shown as drastic difference in importance compared the model with all macroeconomic variables included in Figure 5-2. GDP and NFCI are slightly behind

MD.SR_lag3. The last macroeconomic variable remaining, Leading, is shown as one the least important variables but is still shown as an improvement over four bank specific variables.

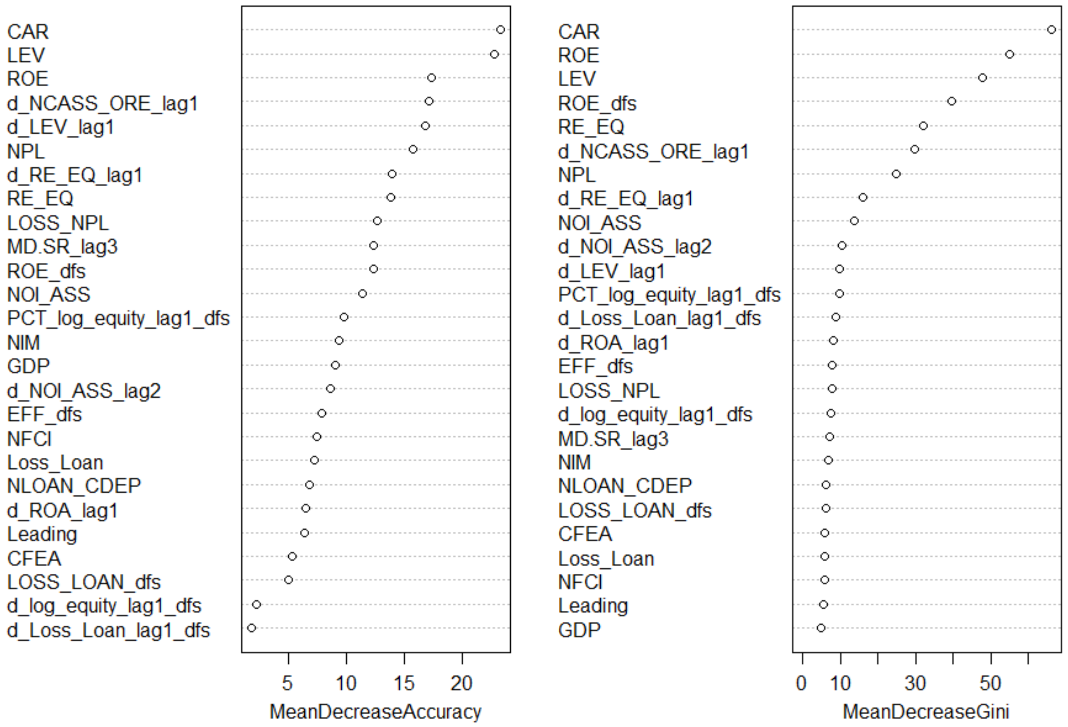

Figure 5-3: Variable Importance Plot (Final Model)

One of the other factors Genuer et al. (2010) evaluated was the number of trees (*ntree*) included in the model. Although impactful, they found the selection of *ntree* to be less sensitive than variable selection. To evaluate the accurate number of trees (*ntree)* in our model, we examined the below error rate relative to the number of trees from R (Figures 5-4 and 5-5). For both models, the overall out-of-bag error (black lines) and the error for good observations (red lines) drop significantly with a limited number of trees. Afterwards, the error rates decrease marginally as more trees are added. The class' error for bad observations (green line) stabilizes after 300 trees. As discussed earlier in Section 3.1, it is optimal to have an infinite number of trees in the model. Genuer et al. (2010) showed stability at *ntree* of 2,000 based on their data while we chose the number of trees for our model to be 650 after evaluating the error rates and downsides of increasing the number of trees such as the running time of the model.

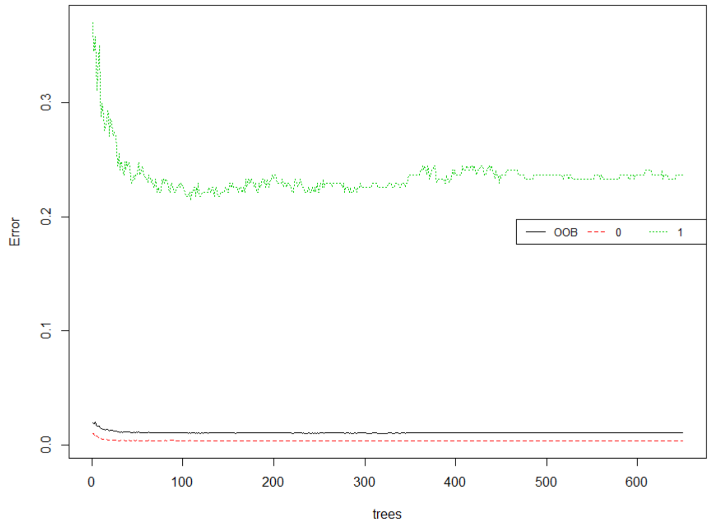

Figure 5-4: Random Forest Error Relative to Number of Trees (Bank Specific Only)

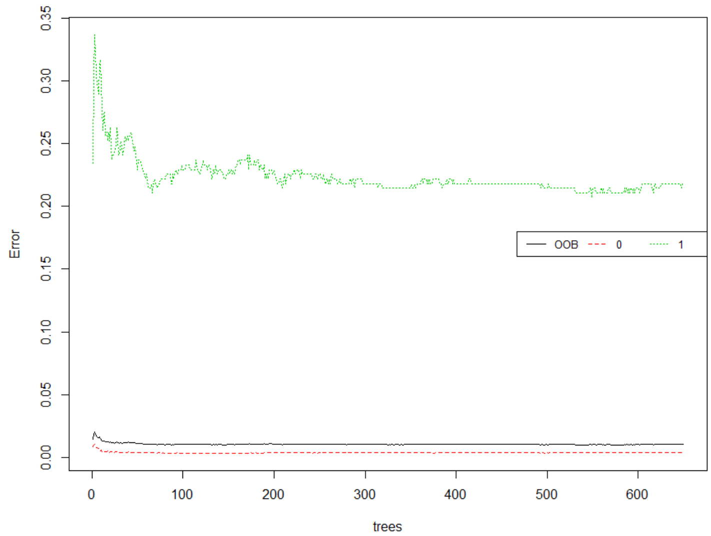

Figure 5-5: Random Forest Error Relative to Number of Trees (Bank Specific + All Macro)

Lastly, we assessed the appropriate value of number of variables randomly sampled as candidates at each split (*mtry*) for all three models. By using the *tuneRF* function in R, we optimized the *mtry* for each model. The function calculates the optimal *mtry* based on the minimum OOB error. The function calculated an *mtry* of 4 for the bank specific model, 52 for the model with bank specific variables and all macro variables, and 5 for the Final Model. The difference in the *mtry* in each model is explained by the large changes in variables in each model and is line with Breiman (2001) suggested range for each model. The development of OOB error by value of *mtry* for bank specific model and the bank specific and all macro variables model are shown in Figures 5-6 and 5-7 below.

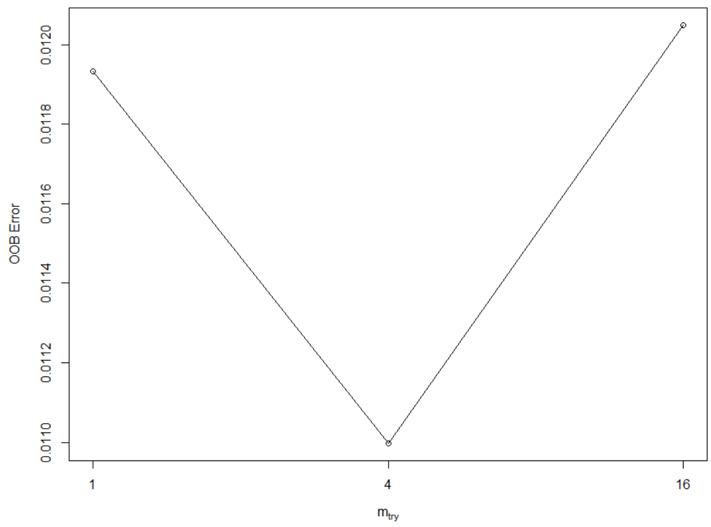

Figure 5-6: Random Forest Error Relative to Number of Trees (Bank Specific Only)

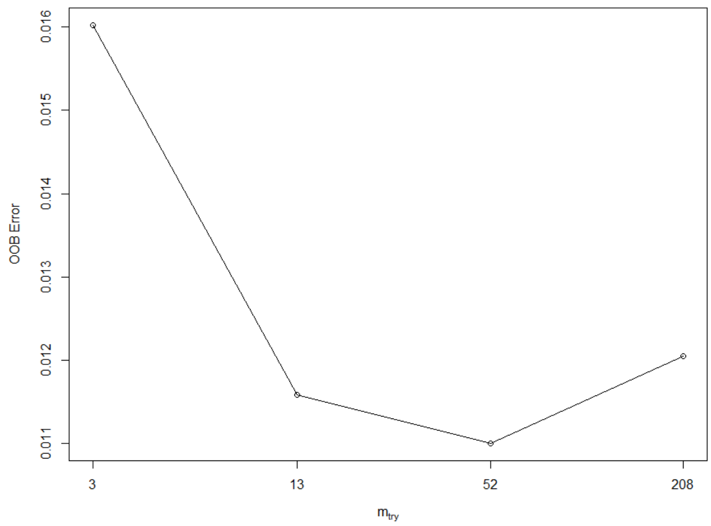

Figure 5-7: Random Forest Error Relative to Number of Trees (Bank Specific + All Macro)

6 Validation

6.1 Measures

To validate the performance of the constructed model of bank specific and macroeconomic variables compared to the bank specific only model, we focus the validation measures and benchmarks towards the classification accuracy of the model as explained by Bekkar et al. (2013). Meaning how often each model accurately predicts if a bank is solvent (good bank observation) or insolvent (bad bank observation). Therefore, we will evaluate multiple performance validation measures based on the rates of sensitivity and specificity, and the relationship between the measures to calculate other performances measures such as the ROC curve, Negative Likelihood Ratio, and others explained below.

To measure the overall effectiveness of the model, we can evaluate the model's ability to accurately classify bank observations through a confusion matrix. There are four classifications based on each bank observation – True Positive, False Positive, True Negative, and False Negative – as shown in Table 6-1 below.

Table 6-1: Confusion Matrix

	Actual Positive	**Actual Negative**
Predicted Positive	True Positive	False Positive
Predicted Negative	False Negative	True Negative

- True Positive = good banks observations *correctly* identified as good banks observations
- False Negative = good banks observations *misclassified* as bad banks observations
- True Negative = bad banks observations *correctly* identified as bad banks observations
- False Positive = bad banks observations *misclassified* as good banks observations

Using the above confusion matrix structure and classifications, the overall accuracy of model performance is based on rate of True Positive and True Negative classifications as shown below.

$$Accuracy = \left(\frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative}\right)$$

$$Error\ Rate = (1 - Accuracy)$$

Since our dataset is heavily influenced by good bank observations and our goal is to effectively predict bad bank observations, we can further split the model accuracy into sensitivity and specificity. Sensitivity measures the rate of solvent (good) banks being correctly identified while specificity measures the rate of insolvent (bad) banks being correctly identified. Therefore, sensitivity is measure of the true positive compared to all actual good banks while specificity measures true negatives compared to all actual bad banks available.

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive}$$

Many of the following performance measure compare the rates of sensitivity and specificity. The first measure examined is Area Under the Curve (AUC), which allows us to plot the Receiver Operating Characteristic (ROC) curve. The ROC curve utilizes the sensitivity and specificity measures explained above. The sensitivity, or true positive rate, is plotted on the y-axis while the 1-specificity, or false positive rate, is plotted on the x-axis. The curve visually displays the overall accuracy of the test. A prediction model with 100% accuracy passes through the upper-left corner of the diagram. Therefore, the closer the curve moves towards the upper-left corner, the better the overall accuracy of the model. The Area Under Curve (AUC) summarizes the position and curvature of the ROC curve. AUC will range from 0.5 to 1 with values of 0.5-0.6 defined as poor model performance and values of 0.9-1 indicating an excellent model performance.

The geometric mean, or G-Mean, measures the product of sensitivity and specificity. A higher G-Mean indicates a better performing model as the metric balances the classification performance of good and bad bank predicts.

$$G = \sqrt{sensitivity * specificty}$$

The Negative Likelihood Ratio (LR) is the ratio between the probability of a false negative and the probability of a true negative. LR is a key metric in our analysis as it measures the performance of observations relating to bad banks, which is what we are ultimately trying to predict in the model.

$$LR- = \left(\frac{1 - sensitivity}{specificity}\right)$$

Youden's Y measures the linear transformation of the mean sensitivity and specificity. A higher Youden's Y indicates a better avoidance in misclassifying banks.

$$Youden's\ Y = sensitivity - (1 - specificity)$$

Discriminant Power (DP) summarizes interaction between sensitivity and specificity. A value for DP greater than 3 indicates a model that differentiates well between good and bad observations.

$$DP = \frac{\sqrt{3}}{\pi}\left[\log\left(\frac{sensitivity}{1 - sensitivity}\right) + \log\left(\frac{specificity}{1 - specificity}\right)\right]$$

Balanced Accuracy (BA) is a simple average of sensitivity and specificity. This measure helps identify the prediction power between classifiers (good and bad bank observations). Since solvent banks make up a much larger percent of the sample, it may signal a stronger performing model just based on its ability to predict good banks. By averaging the two, the measure better accounts for the smaller classification of insolvent banks. We also evaluated weighting the Balanced Accuracy 75%/25% providing two other performance measures Weighted Balance Accuracy 1 (WBA1 – 75% specificity / 25% sensitivity) and Weighted Balance Accuracy 2 (WBA2 – 25% specificity / 75% sensitivity)

6.2 Validation Results

Based on the performance measures discussed above, we evaluated the performance of the three models tested, Bank Specific only, Bank Specific plus All Macroeconomic Variables, and Bank Specific plus Targeted Macroeconomic Variables (or Final Model). Tables 6-2-7 disclose the performance of each model based on the different samples tested.

Table 6-2: Confusion Matrices for Test Sample

	Bank Specific		Bank + All Macros		Final Model	
	Actual Positive	**Actual Negative**	**Actual Positive**	**Actual Negative**	**Actual Positive**	**Actual Negative**
Predicted Positive	7,136	74	7,137	74	7,138	73
Predicted Negative	25	137	24	137	23	138

Table 6-3: Confusion Matrices for Out of Sample

	Bank Specific		Bank + All Macros		Final Model	
	Actual Positive	**Actual Negative**	**Actual Positive**	**Actual Negative**	**Actual Positive**	**Actual Negative**
Predicted Positive	35,608	32	35,600	35	35,600	35
Predicted Negative	105	86	113	83	113	83

Table 6-4: Confusion Matrices for Out of Time

	Bank Specific		Bank + All Macros		Final Model	
	Actual Positive	**Actual Negative**	**Actual Positive**	**Actual Negative**	**Actual Positive**	**Actual Negative**
Predicted Positive	25,342	18	25,364	35	25,363	35
Predicted Negative	25	21	3	4	4	4

Table 6-5: Test Sample Performance Measures

Measure	Goal	Bank Specific	Bank + All Macros	Final Model
Accuracy	1	0.9866	0.9867	0.9870
Sensitivity	1	0.9965	0.9966	0.9968
Specificity	1	0.6493	0.6493	0.654
AUC	> 0.8	0.9177	0.9203	0.9235
G-mean	Max	0.8044	0.8044	0.8074
LR	0	0.0054	0.0052	0.0049
Discriminant power	>3	3.4554	3.4715	3.5164
Balanced Accuracy	1	0.8229	0.8230	0.8254
Youden's γ	1	0.6458	0.6459	0.6508
WBA1	0.5	0.4549	0.4549	0.4556
WBA2	0.5	0.3681	0.3681	0.3699

Table 6-6: Out of Sample Performance Measures

Measure	Goal	Bank Specific	Bank + All Macros	Final Model
Accuracy	1	0.9962	0.9959	0.9959
Sensitivity	1	0.9971	0.9968	0.9968
Specificity	1	0.7288	0.7034	0.7034
AUC	> 0.8	0.7274	0.7112	0.7347
G-mean	Max	0.8525	0.8373	0.8373
LR	0	0.0040	0.0045	0.0045
Discriminant power	>3	3.7649	3.6415	3.6415
Balanced Accuracy	1	0.8630	0.8501	0.8501
Youden's γ	1	0.7259	0.7002	0.7002
WBA1	0.5	0.4650	0.4617	0.4617
WBA2	0.5	0.3979	0.3884	0.3884

Table 6-7: Out of Time Performance Measures

Measure	Goal	Bank Specific	Bank + All Macros	Final Model
Accuracy	1	0.9983	0.9985	0.9866
Sensitivity	1	0.9990	0.9999	0.9998
Specificity	1	0.5385	0.1026	0.1026
AUC	> 0.8	0.7279	0.7850	0.7576
G-mean	Max	0.7335	0.3203	0.3203
LR	0	0.0019	0.0010	0.0019
Discriminant power	>3	3.8930	3.8822	3.5000
Balanced Accuracy	1	0.7688	0.5513	0.5512
Youden's γ	1	0.5375	0.1025	0.1024
WBA1	0.5	0.4419	0.3878	0.3878
WBA2	0.5	0.3268	0.1635	0.1635

Based on the above tables, the models show varying results depending on the sample. The accuracy measured in all models is never lower than 98.66% and is the highest in the out of sample period across all three models. Comparing results for the test data, all three models perform almost identically as shown in the confusion matrix for the test data (in Table 6-2). However, the Final Model performs slightly better than the Bank Specific and Bank Specific plus All Macroeconomic Variables. This is also confirmed by the validation measures in Table 6-5. The Final Model performs best across all validation measures with the most significant difference being the AUC measurement (0.9235 compared to 0.9177 for the Bank Specific Model.) This makes theoretical sense considering the Final Model includes the most important macroeconomic variables based on the sample as discussed previous.

When evaluating the validation results outside the test sample, the results are significantly different. The Bank Specific Model outperforms the models with the addition of macroeconomic variables in both the out of sample and out of time. Tables 6-3 and 6-6 display the Bank Specific Model stronger performance across all measures for out of sample, as the models with macroeconomic variables produce a larger number of false negatives, or solvent banks that were predicted to be insolvent.

When first evaluating the out of time sample in Tables 6-4 and 6-7, the Bank Specific plus All Macroeconomic Variables Model shows the highest accuracy and AUC. However, this higher performance is purely based on the stronger prediction of solvent banks during the sample. Both models with macroeconomic variables perform very poorly when classifying bad banks as they both predict almost all 25,406 banks as solvent banks and of the banks that were classified as insolvent, about half were wrong. This poor performance in predicting bad banks is detailed in the low specificity score. Whereas the Bank Specific Model performs meaningfully better when predicting bad bank observations as shown by specificity as well as the LR and Balance Accuracy measures which better account for accuracy across good and bad bank observations.

Although the test sample results indicated the Final Model performed best, the out of sample and out of time strongly suggests that the Bank Specific Model outperforms the models with macroeconomic variables. Based on the consistency of importance across the models of bank specific variables, such as leverage ratio (LEV), capital adequacy ratio (CAR) and return on equity (ROE), these results are not overly surprising. But the level of outperformance between the models is noteworthy.

The ROC curves displayed below in Figure 6-1 display many of the key findings we discussed based on above validation tables. First, across all three models, the ROC curve for the test data is almost identical. The ROC curves show very high performing models, as the AUC measure is greater than 0.9, pulling the curve to the top left corner. Although the Final Model performed slightly better, this confirms the limited impact of the additional macroeconomic variables impact on the model. The Final Model and All Variables Model displayed the best performing ROC curve in the out of sample and out of time, respectively. Based on the previous discussion, this can be misleading as the Bank Specific Model performs better across other measures as it more accurately predicts the bad bank observations.

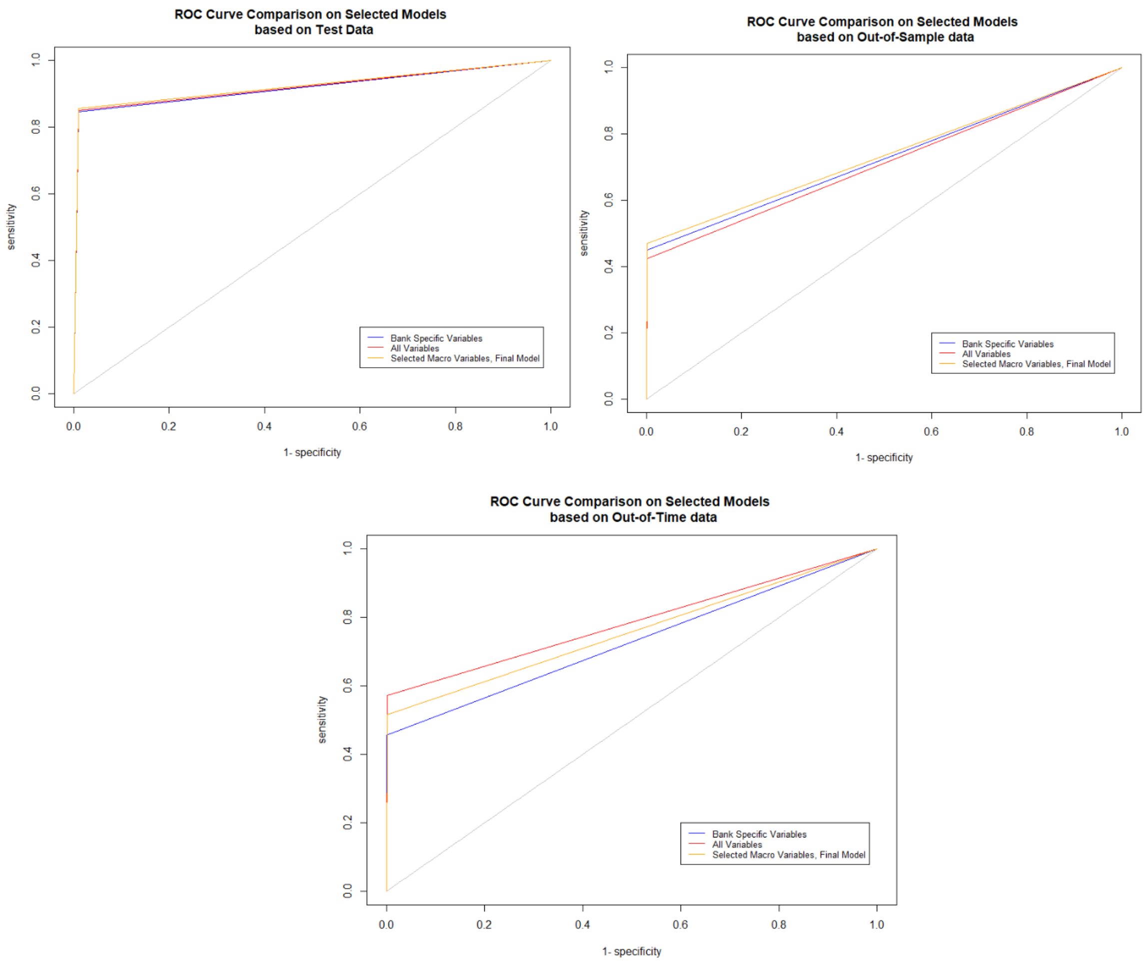

Figure 6-1: ROC Curve Performance between Models and Sample Periods

6.3 Variable Importance

To complement the variable importance plots based on Mean Decrease Accuracy, the below Forest Floor Main Effect plots of the Final Model's structure describe the variable importance effect with the dependent variable as shown in Figures 6-2 and 6-3. Welling et al. (2016) first introduced the graphical measure to show the relationship between variable values (x-axis) and the out-of-bag cross-validated feature contribution (y-axis). The cross-validation feature expresses the change of the predicted probability due to changes of the variable value. Therefore, a flat line shows a weaker contribution by the independent variable. Highly important variables such as return on equity (ROE), leverage ratio (LEV), and capital adequacy ratio (CAR) show more variation and non-linearity in their plots and a higher R-squared values, which measures the goodness-of-fit when visualizing the variable effect as the main effect. Less important variables like Loss Allowance to Noncurrent Loans (LOSS_NPL) are shown as relatively flat and have little contribution according to the R-squared. The plots also quantify the interactions of different regressors with the most important variable (CAR) by using color gradient. The span is from red to green to blue, where low interactions are colored red and high values in blue.

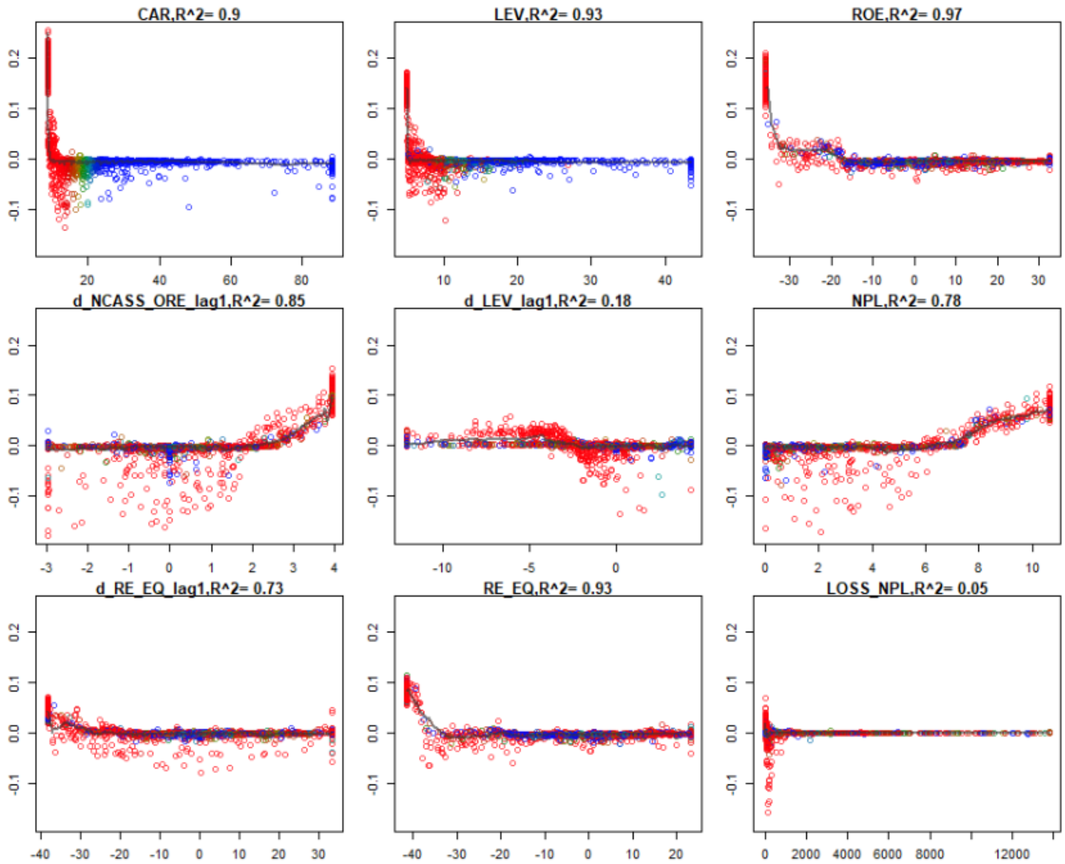

Figure 6-2: Variable Importance Effect - Bank Specific Variables from Final Model

In Figure 6-3, we evaluated the plots of the macroeconomic variables included in the Final Model. First, you can see that the number of data points is significantly less compared to the above banks specific graphs since we had thousands of bank observations each year included in the model. NFCI and Leading have similar R-squared values placing them as more important macroeconomic variables than GDP or the Mortgage Debt Service Ratio – lagged 3 years (MD.SR_lag3). It is noteworthy that both NFCI and the Leading Index are aggregate economic variables evaluating multiple relationships known to be leading indicators in the United States. NFCI has the largest concentration of plots towards the right of the chart while Leading has a similar plot on the left. This difference is related to NFCI having an inverse relationship with the market, meaning the NFCI is higher during poor market conditions. Therefore, regarding both cases, the macroeconomic variables show higher importance during poor economic conditions, which supports our initial expectations. GDP and MD.SR_lag3 display a similar relationship but have more dispersion across the data points, which decreases the overall significance as shown by the low R-squared terms.

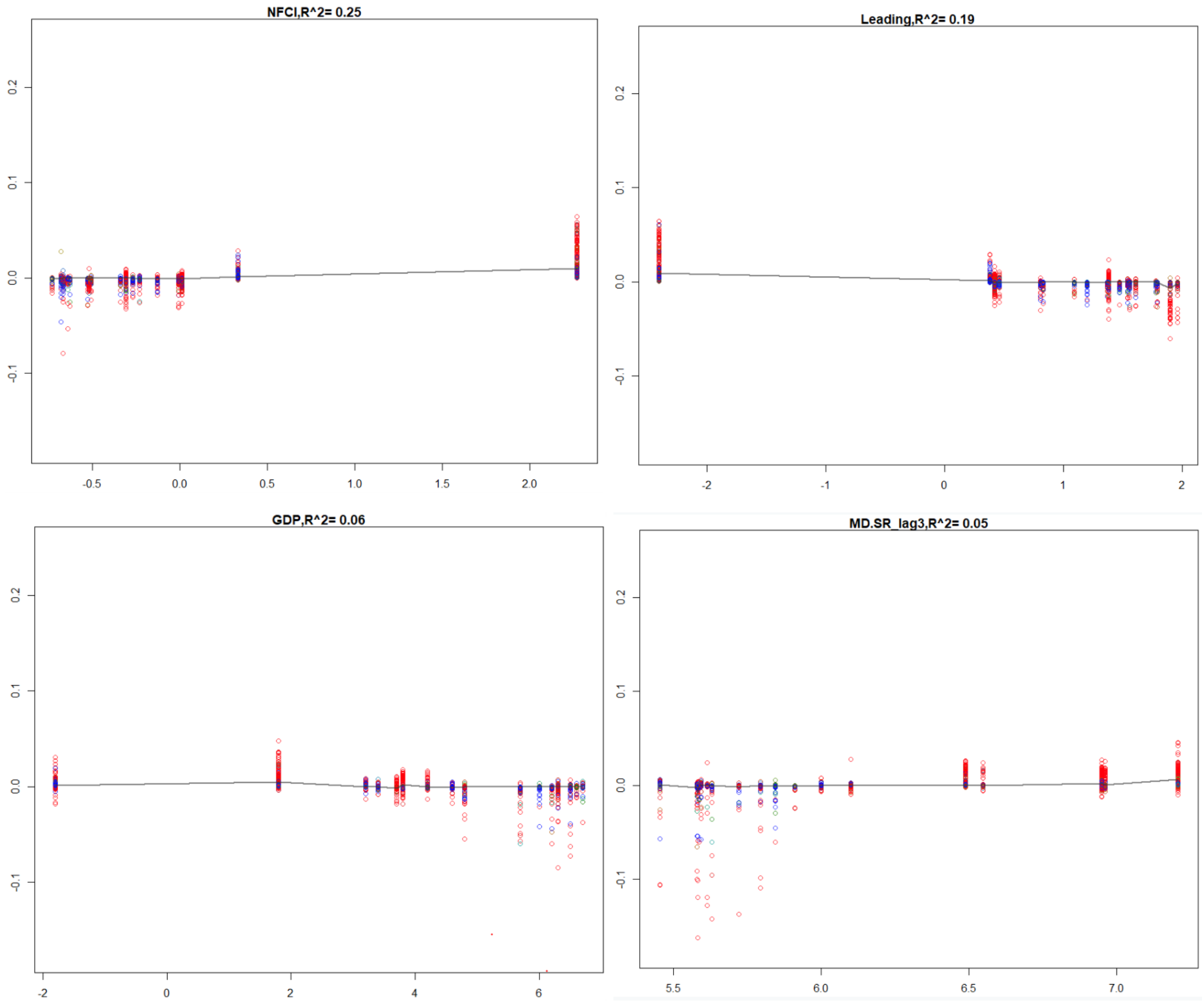

Figure 6-3: Variable Importance Effect – Macroeconomic Variables from Final Model

The most important macroeconomic variable based on the above importance plots is the NFCI, or Chicago Fed's National Financial Conditions Index. The Federal Reserve Bank of Chicago publishes the NFCI weekly to capture financial conditions in the United States by considering three sub-indexes measuring risk, credit, and leverage to capture the economic and financial conditions of money markets, debt and equity markets, and the banking systems. Risk is measured by the volatility and funding risk in the financial sector, credit measures the credit conditions, and leverage captures debt and equity levels. Given the construction of the index and the bank specific variables that have shown to be most significant, it easy to understand why the NFCI was the most significant macroeconomic variable as shown by the variable importance effect.

The Leading Index was also important in the Final Model. The Federal Reserve Bank of Philadelphia publishes the Leading Index monthly by includes multiple leading economic variables in the United States such as state-level housing permits, state initial unemployment insurance claims, delivery times from the Institute for Supply Management (ISM) manufacturing survey, and the interest rate spread between 10-year Treasury bond and 3-month Treasury bill. Many of these relationships were included in our sample of macroeconomic variables however the relationship between the variables in one constructed index proved to be more significant than any single variable.

The remaining two macroeconomic variables added to the bank specific model was the level of GDP and MD.SR_lag3. The additions of these two variables are consistent with previous empirical studies. Mayes and Stremmel (2014) found GDP to be significant in their empirical research, which also evaluated CAMELS indicators using FDIC data. Cole and White (2012) found that real-estate loans and mortgages were a leading indicator to bank defaults during the Great Financial Crisis. As our evaluation period includes this time, the inclusion of the Mortgage Debt Service Ratio is aligned with their findings.

The variety of inputs and relationships between the macroeconomic variables described above adds to the overall robustness of the model as shown in the comparison between the Final Model and the All Variable Model. Table 6-8 shows the correlation between the macroeconomic variables added. Many of the relationships actually show negative correlations. It is worth mentioning that the levels of GDP and Leading Index are fairly correlated at 0.758 and NFCI and Leading are fairly negatively correlated at –0.881.

Table 6-8: Correlation Matrix of Final Macroeconomic Variables

	GDP	Leading	NFCI	MD.SR
GDP	1			
Leading	0.758	1		
NFCI	-0.705	-0.881	1	
MD.SR	-0.095	-0.497	0.518	1

6.4 Models Performance as Early Warning System (EWS)

Petropoulos et al. (2017) mentioned that probability of default (PD) models could be useful for banking supervisors to identify vulnerable institutions and work in a sense as EWS. They define the limit of high-risk banks when the probability of default is larger than 25%. In Tables 6-9-11, we can see how our models based on Random Forest would have classified bad banks as either low or high risk based on the latest annual financial statements prior to the default.

Table 6-9: Low and High Risk Classification of Defaulted Banks in Test Sample

	Bank Specific	Bank + All Macros	Final Model
	Actual Negative	**Actual Negative**	**Actual Negative**
Low Risk	15.2%	19.4%	15.2%
High Risk	84.8%	80.6%	84.8%

Table 6-10: Low and High Risk Classification of Defaulted Banks in Out of Sample

	Bank Specific	Bank + All Macros	Final Model
	Actual Negative	**Actual Negative**	**Actual Negative**
Low Risk	13.6%	20.3%	12.7%
High Risk	86.4%	79.7%	87.3%

Table 6-11: Low and High Risk Classification of Defaulted Banks in Out of Time

	Bank Specific	Bank + All Macros	Final Model
	Actual Negative	**Actual Negative**	**Actual Negative**
Low Risk	25.6%	30.8%	28.2%
High Risk	74.4%	69.2%	71.8%

A strong EWS identifies high risk and vulnerable institutions timely to improve decision-making and actions taken by regulators. Our models identify 69.2% to 87.3% of the failed banks as high risk across the models and samples. The Final Model and Bank Specific Model perform identically during the test sample. The Final Model performs better in the out of sample while the Bank Specific model identifies the bad banks slightly better in the out of time. Based on these results, the Bank Specific Model more consistently classifies bad bank observations accurately prior to default across samples. This aligns with the validation measures indicating the stronger performance of the Bank Specific Model.

7 Empirical Results of Macroeconomic Variables Effect

The purpose of this section is to evaluate the overall effect of the macroeconomic variables' addition to the Random Forest Model first introduced by Petropoulos et al. (2017). The concentration of historical defaults in the FDIC data indicates a direct correlation between a weaker economic environment and an increased number of defaults. Based on our research and subsequent models, the overall effect of macroeconomic variables has shown negative results as the additional macroeconomic variables decreased the predictive power during the out of sample and out of time periods. By decreasing the number of macroeconomic variables to the most important variables (GDP, Leading Index, NFCI, and Mortgage Debt Service Ratio – lagged 3 years), the model is shown to have highly similar validation results compared to the model with all macroeconomic variables included.

One of the most relevant measurements explaining the difference in performance between the models is the specificity measure. The specificity measures the model's ability to predict true negatives (our main goal), which is insolvent or "bad bank" observations in our research. In our analysis, the Bank Specific Model shows improvement over the other two models when classifying bad bank observations during the out of sample and a significant improvement during the out of time period. Additionally, the consistency of the EWS confirms these results indicating the Bank Specific Model as being performing better.

This underscores previous literature that describes bank specific variables based on the CAMELS framework as driving bank default prediction models. Along with the validation measures, we discovered similar results considering the consistency of importance of bank specific variables across each model. Multiple views indicate the significance of the bank specific variables including the Variable Importance Plots, Variable Importance Effects, and R-squared values of explaining the classification of bank defaults.

Based on the muted impact of the macroeconomic variables, we hypothesis multiple responses that may explain the limited improvement in predictive power. One such effect questioned is that the current market environment may already be accounted for in the bank-specific data. For example, since banks are pro-cyclical, if macroeconomic variables indicate a weaker economy, banks will likely perform worse since there will be less activity, leading to lower commissions and less quality in the loan book, which will ultimately increase the probability of loan provisions or impairments. Therefore, a poor economy indirectly impacts the banking industry by forcing banks to adjust or display weaker financial statistics.

Figure 7-1 extends this hypothesis further. Figure 7-1 plots the GDP and Leading Index level over the average return on equity (ROE) and leverage (LEV) of all banks evaluated each year. The figure shows a correlation between the macroeconomic variables and the bank specific variables. ROE explicitly shows this relationship as ROE across all firms' significantly decreases during time of low GDP and Leading Index levels such as 2001-2002 and 2008-2009. LEV follows a similar pattern but with less dispersion compared to ROE.

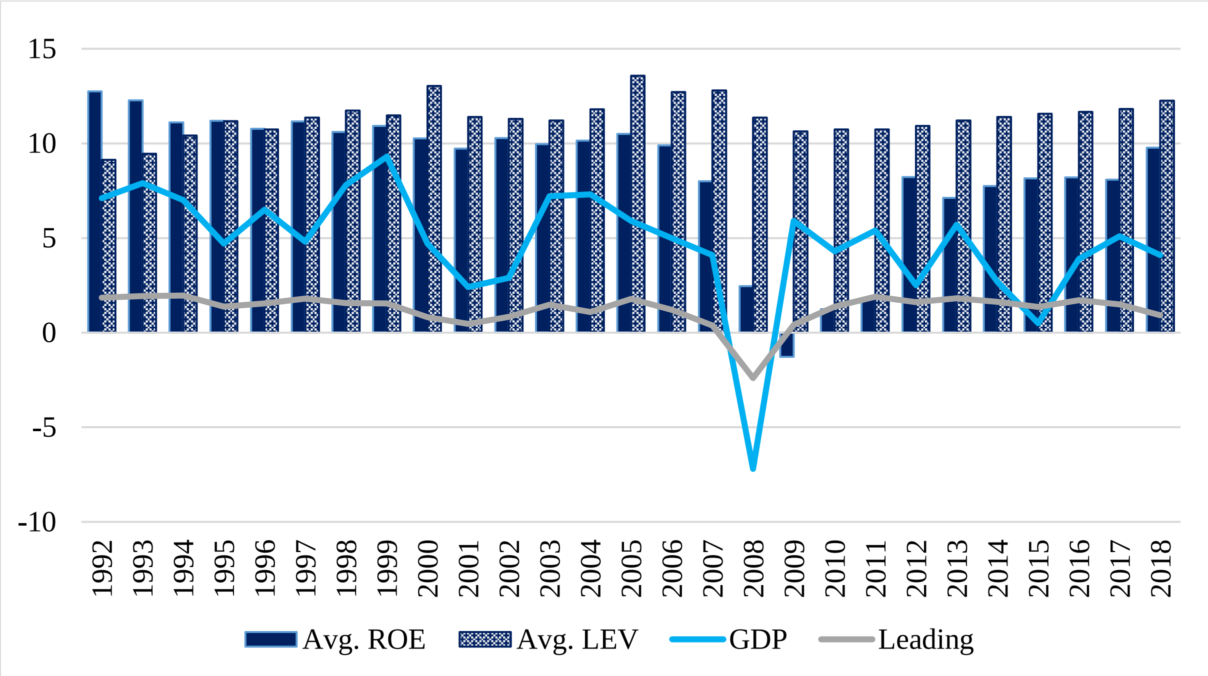

Figure 7-1: Relationship between Bank Specific and Macroeconomic Variables Overtime

Another hypothesis we consider is each economic downturn or recession in the United States has been driven by different factors. The Great Financial Crisis was impacted by real estate and real estate loans indicating the importance of the Mortgage Debt Service Ratio in our Final Model. Since our study evaluates defaults over a 20+ year period, multiple macroeconomic relationships impact bank defaults. Given each economic downturn is driven by different factors, this may help explain the decreased predictive power of the Final Model which includes macroeconomic variables. This also adds to the importance of the aggregated variables such the Leading Index and NFCI since they combine many leading indicators in the United States.

As the Federal Reserve refines these aggregation measures with increased samples, these measurements may increase in importance and accuracy to better indicate economic strength or weaknesses.

Based on our results, regulators such as the FDIC would be prudent to continue to focus on the CAMELS indicators to best align resources to banks who default or are on the verge of default. By using a Random Forest classification method based on FDIC-insured banks in the United States, the addition of macroeconomic variables has proven to have a limited impact on the predictive power when evaluated over an extended period of bank defaults. However, given the level of change in the banking industry, regulators and researchers should continue to search for key variables and relationships that drive bank defaults to better stabilize the industry prior to the next financial downturn.

8 Conclusion

When predicting bank defaults in the United States from 1994-2016 using a Random Forest classification method, our empirical results indicate that the addition of macroeconomic variables reduce the predictive power compared to a model explained purely by bank specific variables based on the CAMELS approach. The main goal was to evaluate an extended period of banking defaults in the U.S. to build a model that would improve default predictions during an out of time sample and be a strong predictor of bank defaults across future banking environments. However, the addition of macroeconomic variables decreased the performance during the out of sample and out of time period of 2013-2016.

Future research would be prudent to evaluate the model's performance in predicting bank defaults during the next market downturn in the United States as the current out of time sample is characterized by a low default frequency. A higher frequency of bank defaults, and a poorer macroeconomic environment, will likely provide a true indication of the overall performance of the model. However, since the Bank Specific Model performed significantly better than the macroeconomic models, this may have a limited effect.

It may also be advised to evaluate the most important macroeconomic variables during a shorter time period. Therefore, the macroeconomic variables chosen may be more relevant for the specific sample. However, every recession or banking crisis has been driven by different factors, making a macroeconomic variable important during one period, and display limited importance in the next. This provides additional support to the usage of aggregate macroeconomic variables like the Leading Index or NFCI. However, by using aggregate macroeconomic variables that were constructed based on leading indicators for the United States of America, the transferability of our model to other regions or countries is limited. Therefore, our Final Model would be inappropriate to predict bank defaults in Europe or other areas even if future research showed improved results.

Our model utilized annual data, which in general displayed worse performance compared to the bank specific model of Petropoulos et al. (2017). Their model based on quarterly data from a concentrated time period implies that decreasing the frequency from quarterly to annual data creates a loss in valuable information. This suggests that market environments and a bank's financing situation can change quickly. It is worth mentioning that when using annual data, the difference between the last bank reporting and a failure can be more than 12 months as some of the banks failed in the first months of the year had not issued statements for the preceding year.

References

Altman, E.I. (1968). “Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy”. *The Journal of Finance 23, Pages 589-609*.

Arena, M. (2008). “Bank Failures and Bank Fundamentals: A Comparative Analysis of Latin American and East Asia During the Nineties Using Bank-Level Data”. *Journal of Banking & Finance 32, Pages 299-310*.

Avkiran, Necmi and Cai, Cynthia. (2012). “Predicting bank financial distress prior to crises”. *The University of Queensland, Australia* [online]. Available at: https://www.researchgate.net/publication/228492072_Predicting_bank_financial_distress_prior_to_crises

Basel Committee on Banking Supervision. (2004). “Basel II: International Convergence of Capital Measurement and Capital Standards (A Revised Framework)”. *Bank for International Settlements* [online]. Available at: https://www.bis.org/publ/bcbs107.pdf

Basel Committee on Banking Supervision. (2011). “Basel III: A global regulatory framework for more resilient banks and banking systems”. *Bank for International Settlements* [online]. Available at: https://www.bis.org/publ/bcbs189.pdf

Bekkar, Mohamed and Djema, Hassiba and Alitouche, T.A.. (2013). “Evaluation measures for models assessment over imbalanced data sets.” *Journal of Information Engineering and Applications. 3. 27-38.*

Berger, Allen N., and Bouwman, Christa HS. (2013). “How does capital affect bank performance during financial crises?”. *Journal of Financial Economics. Pages 146-176*.

BIS. (2018). “*History of the Basel Committee”* [online]. Available at: https://www.bis.org/bcbs/history.htm?m=3%7C14%7C573%7C76

Booth, Ash. Gerding, Enrico. and McGroarty, Frank. (2014). “Automated trading with performance weighted random forests and seasonality”. *Expert Systems with Applications, Vol. 41, Issue 8, Pages 3651-3661*.

Breiman, Leo. (2001). “Random forests”. *Machine Learning 45:5-32*.

Cole, Rebel A. and White, Lawrence. (2012).” Déjà Vu All Over Again: The Causes of U.S. Commercial Bank Failures This Time Around”. *Journal of Financial Services Research, Vol. 42, Issue 1, Pages 5-29.*

Cole, Rebel A. and Wu, Qiongbing. (2009). “Is hazard or probit more accurate in predicting financial distress? Evidence from U.S. bank failures”. *22nd Australasian Finance and Banking Conference.*

Cutler, D.R., Edwards Jr., T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., and Lawler, J.J. (2007). Random Forests for Classification in Ecology. *Ecology Vol. 88, Pages 2783-2792*.

Díaz-Uriarte, R. and Alvarez de Andrés, S. (2006) "Gene selection and classification of microarray data using random forest". *BMC Bioinformatics. 7, 3.*

Demirguc-Kunt, Asil. and Detragiache, Enrica. (1997). "The determinants of banking crises: evidence from industrial and developing countries". *IMF Research Working Paper*.

Demyanyk, Yuliya. and Hasan, Iftekhar. (2010). "Financial Crisis and Bank Failures: A Review of Prediction Methods". *Omega 38.5, Pages 315-324*.

Denil, Misha. Matheson, David. and Freitas, Nando De. (2013). "Consistency of Online Random Forests". *ICML (3) 28, Pages 1256-1264*.

DnI Institute. (2015). "*Random Forest Using R: Step by Step Tutorial*" [online]. Available at: http://dni-institute.in/blogs/random-forest-using-r-step-by-step-tutorial/

Federal Deposit Insurance Corporation. (1984). "*Federal Deposit Insurance Corporation: The First Fifty Years*". FDIC.

Federal Deposit Insurance Corporation. (2017). "*Who is the FDIC?*" [online]. Available at: https://www.fdic.gov/about/learn/symbol/index.html

Federal Reserve Bank of Chicago. (2019). "*Chicago Fed National Activity Index (CFNAI)* Current Data" [online]. Available at: https://www.chicagofed.org/research/data/cfnai/current-data

Federal Reserve Bank of Chicago, Chicago Fed National Financial Conditions Index [NFCI], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/NFCI, May 19, 2019.

Federal Reserve Bank of Philadelphia, Leading Index for the United States [USSLIND], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/USSLIND, May 19, 2019.

Flannery, Mark. (1998). "Using Market Information in Prudential Bank Supervision: A Review of the U.S. Empirical Evidence". *Journal of Money, Credit and Banking, Vol. 30, Issue 3, Pages 273-305.*

Genuer, Robin. Poggi, Jean-Michel, Tuleau-Malot, Christine. (2010). "Variable selection using Random Forests." *Pattern Recognition Letters, Elsevier, 31 (14), Pages 2225-2236.*

Gurm, HS. Kooiman, Judith. Lalonde, Thomas. Grines, Cindy. Share, David. and Seth, Milan. (2014). "A Random Forest Based Risk Model for Reliable and Accurate Prediction of Receipt of Transfusion in Patients Undergoing Percutaneous Coronary Intervention". *Journal PLoS ONE 9(5): e96385.*

Kaminsky, Graciela L. and Reinhart, Carmen M. (1996). "The twin crises: the causes of banking and balance-of-payments problems". *International Finance Discussion, Pages 544.* Board of Governs of the Federal Reserve System (U.S.).

Kartasheva, A. V. and Traskin, M. (2011). "Insurers' Insolvency Prediction using Random Forest Classification". *Journal of Risk and Insurance, revised and resubmitted*.

Khaidem, Luckyson. Saha, Snehanshu. and Roy Dey, Sudeepa. (2016). "Predicting the direction of stock market prices using random forest". *Applied Mathematical Finance.*

Lock, Dennis. and Nettleton, Dan. (2014). "Using random forests to estimate win probability before each play of an NFL game". *University of Iowa* [online]. Available at: http://homepage.divms.uiowa.edu/~dzimmer/sports-statistics/nettletonandlock.pdf

Mayes, David G. and Stremmel, Hanno. (2014). "The Effectiveness of Capital Adequacy Measures in Predicting Bank Distress". *Financial Markets & Corporate Governance Conference* [online]. Available at SSRN: https://ssrn.com/abstract=2191861 or http://dx.doi.org/10.2139/ssrn.2191861

Messai, Ahlem Selma. and Gallali, Mohamed Imen. (2015). "Financial Leading Indicators of Banking Distress: A Micro Prudential Approach-Evidence from Europe". *Asian Social Science 11.21, Page 78.*

Ohlson, James A. (1980). "Financial ratios and the probabilistic prediction of bankruptcy". *Journal of Accounting Research, Pages 109-131.*

Petropoulos, Anastasios. Siakoulis, Vasilis. Stavroulakis, Evangelos. and Vlachogiannakis, Nikolaos E. (2017). "Predicting bank insolvencies using machine learning techniques". *Bank of Greece* [online]. Available at: https://pdfs.semanticscholar.org/f059/6623add1f92af34045fefd4bbadd16a99841.pdf?_ga=2.262563983.224226644.1557822881-577554370.1557822881

Strobl, S., Boulesteix, A.-L., Kneib, T., Augustin, T., and Zeileis, A. (2008). "Conditional Variable Importance for Random Forests. *BMC Bioinformatics 9:307.*

Vazquez, Franciso. and Federico, Pablo. (2015). "Bank funding structures and risk: Evidence from the global financial crisis". *Journal of Banking & Finance. Pages 1-14.*

Welling, Søren and Refsgaard, Hanne and Brockhoff, Per and Clemmensen, Line. (2016). "Forest Floor Visualizations of Random Forests." *Technical University of Denmark. arXiv:1605.09196.*

Wu, Hsu-Che. and Wu, Yu-Ting. (2016). "Evaluating credit rating prediction by using the KMV model and random forest". *Kybernetes, Vol. 45, Issue 10, Pages 1637-1651.*

Yeh, Ching-Chiang. Chi, Der-Jang. and Lin, Yi-Rong. (2014). "Going-concern prediction using hybrid random forests and rough set approach". *Information Sciences, Vol. 254, Pages 98-110.*

Appendix

Table A-1: Correlation Matrix of All Macroeconomic Variables Evaluated

	GDP	GDP_log	CC	TED	RealTrade	CPI	Infl	VIX	Leading	CFNAI	NFCI	FinStress
GDP	1											
GDP_log	-0.392	1										
CC	-0.417	**0.994**	1									
TED	-0.224	-0.220	-0.229	1								
RealTrade	-0.094	-0.102	-0.039	-0.070	1							
CPI	-0.386	**0.991**	**0.976**	-0.238	-0.176	1						
Infl	0.565	-0.265	-0.260	0.203	-0.051	-0.268	1					
VIX	-0.547	0.003	0.049	0.374	0.309	-0.028	-0.233	1				
Leading	0.758	-0.151	-0.193	-0.424	-0.145	-0.110	0.308	-0.708	1			
CFNAI	0.760	-0.138	-0.153	-0.435	0.064	-0.132	0.419	-0.459	0.799	1		
NFCI	-0.705	0.071	0.093	0.616	-0.045	0.050	-0.237	0.776	-0.881	-0.745	1	
FinStress	-0.343	-0.536	-0.504	0.711	0.229	-0.565	0.008	0.694	-0.663	-0.527	0.746	1
WIL5000												
Baa.10 Spread												
10.1Yr Spread												
Oil												
TD.SR												
MD.SR												
CD.SR												
FD.SR												
EMPLOY												
FED FUNDS												

	FED FUNDS	EMPLOY	FD.SR	CD.SR	MD.SR	TD.SR	Oil	10.1Yr Spread	Baa.10 Spread	WILL 5000
GDP	0.442	-0.252	-0.159	-0.061	-0.095	-0.098	-0.102	-0.177	-0.836	-0.175
GDP_log	-0.693	0.128	-0.313	-0.129	-0.336	-0.308	0.692	0.181	0.346	**0.876**
CC	-0.690	0.107	-0.246	-0.031	-0.307	-0.241	0.668	0.192	0.371	**0.858**
TED	0.450	-0.225	0.416	0.093	0.429	0.359	-0.123	-0.500	0.233	-0.248
RealTrade	0.213	-0.608	0.293	0.733	-0.117	0.253	-0.600	-0.188	0.056	0.050
CPI	-0.719	0.186	-0.396	-0.235	-0.391	-0.398	0.716	0.210	0.333	**0.882**
Infl	0.487	-0.172	0.243	0.180	0.257	0.273	0.150	-0.341	-0.441	-0.250
VIX	-0.096	0.133	0.368	0.318	0.239	0.324	-0.126	0.068	0.738	-0.160
Leading	0.203	-0.234	-0.585	-0.358	-0.497	-0.533	-0.002	-0.149	-0.765	0.098
CFNAI	0.138	-0.166	-0.269	-0.062	-0.256	-0.218	-0.010	-0.005	-0.655	0.006
NFCI	-0.178	0.356	0.492	0.134	0.518	0.445	0.093	0.113	**0.821**	-0.214
FinStress	0.399	-0.022	0.654	0.357	0.581	0.595	-0.460	-0.163	0.428	-0.625
WIL5000	-0.473	-0.213	-0.524	-0.156	-0.650	-0.553	0.425	-0.034	0.051	1
Baa.10 Spread	-0.553	0.418	0.190	0.065	0.182	0.165	0.193	0.358	1	
10.1Yr Spread	-0.778	0.714	-0.081	-0.145	0.004	-0.063	0.276	1		
Oil	-0.545	0.514	-0.176	-0.353	0.059	-0.120	1			
TD.SR	0.330	-0.040	**0.982**	0.721	**0.901**	1				
MD.SR	0.276	0.251	**0.877**	0.350	1					
CD.SR	0.272	-0.489	0.721	1						
FD.SR	0.351	-0.067	1							
EMPLOY	-0.604	1								
FED FUNDS	1									

Table A-2: List of Transformation to Macro Variables

		Lagged Values				First Difference			
Variable	**Current**	**1 yr**	**2 yr**	**3 yr**	**4 yr**	**1 yr**	**2 yr**	**3 yr**	**4 yr**
Baa.10Spread	X	X	X	X	X	X	X	X	X
X30.1YrSpread									
X10.1YrSpread	X	X	X	X	X	X	X	X	X
VIX	X	X	X	X	X	X	X	X	X
CFNAI	X	X	X	X	X	X	X	X	X
NFCI	X	X	X	X	X	X	X	X	X
CC						X	X	X	X
CD.SR	X	X	X	X	X	X	X	X	X
CPI						X	X	X	X
OIL						X	X	X	X
FEDFUNDS	X	X	X	X	X	X	X	X	X
FD.SR	X	X	X	X	X	X	X	X	X
TD.SR	X	X	X	X	X	X	X	X	X
INFL	X	X	X	X	X	X	X	X	X
LEADING	X	X	X	X	X	X	X	X	X
MD.SR	X	X	X	X	X	X	X	X	X
GDP	X	X	X	X	X	X	X	X	X
RealTrade	X	X				X	X	X	X
FinStress	X	X	X	X	X	X	X	X	X
TED	X	X	X	X	X	X	X	X	X
EMPLOY	X	X	X	X	X	X	X	X	X
WILL5000						X	X	X	X

"X" indicates inclusion in Model

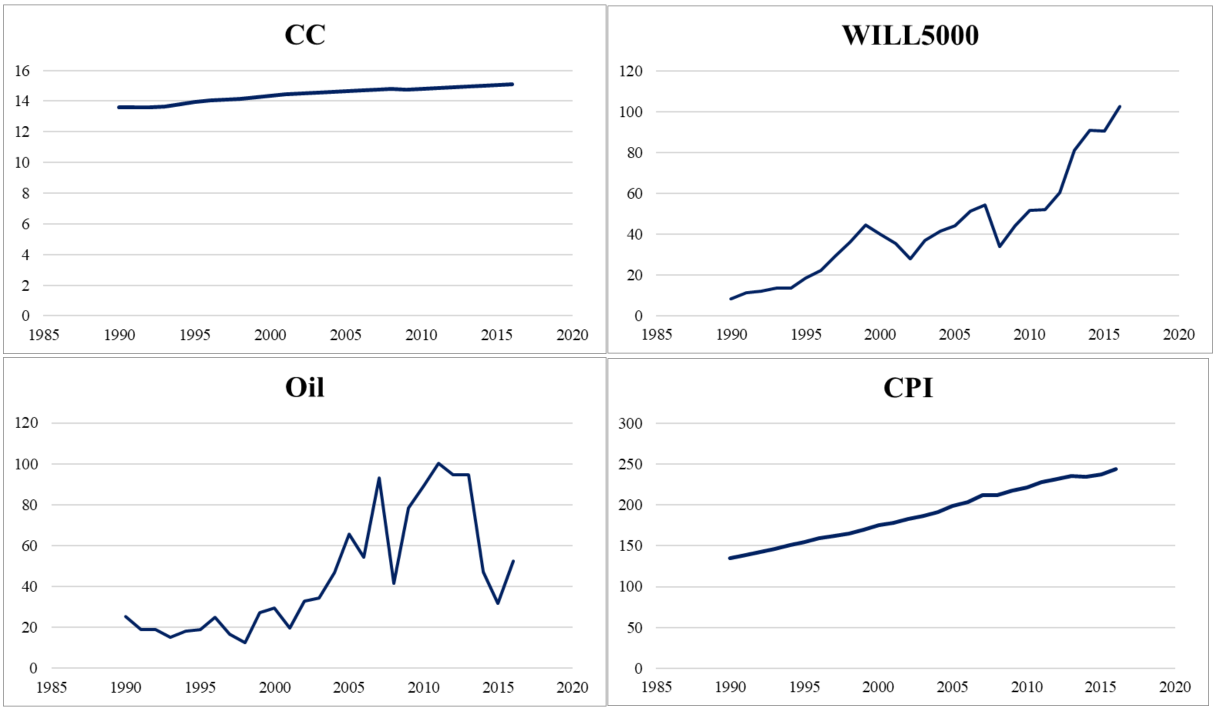

Figure A-1: Trend Figures for Macro Variables Transformation Not Included