The 2022 International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/5/20220097

Prediction and Analysis of Financial Crisis

Yingjie Fu[1,a,*],Yunxiang Gao[2,b], Zhanhao Xu[3,c], Jingxuan Gu[4,d]

[1]Department of Arts, University of Waterloo, Waterloo, N2N 3G1, Canada
[2]Beijing National Day School, Beijing, 100871, China
[3]Shenzhen College of International Education, Shenzhen, 518000, China
[4]Shanghai American School, Shanghai, 201107, China
a. g22fu@uwaterloo.ca, b. gaoyx0126@126.com,
c. s20184.gu@stu.scie.com.cn, d. kalson01px2023@saschina.org
**corresponding author*

Abstract: Financial crisis forecasting (FCP) plays a crucial role in economic phenomena. An accurate forecast of the number and likelihood of failures indicates the growth and strength of a country's economy. Traditionally, several effective FCP methods have been proposed. On the other hand, classification performance, prediction accuracy, and data legitimacy are not good enough for practical applications. In addition, many developed methods perform well for some specific datasets but do not apply to different datasets. Therefore, there is a need to develop an effective prediction model to obtain better classification performance and to adapt to other datasets. In this paper, we improve the data characteristics of the existing methods, including introducing time series variables, macroeconomic indicators interaction terms, etc. Finally, this paper attempts to predict financial crises using logistic regression models. The analysis of the results ensures that the proposed FCP model outperforms other classification models based on different metrics and explores the essential factors affecting financial crises.

Keywords: financial crisis, logistic regression, macroeconomic indicator.

1. Introduction

Finance is indispensable in society, yet finance can also be a terror. Between 1929 and 1933, the most significant financial crisis originated in America, resulting in nearly 11,000 American banks failing. The unemployment rate in America surged to more than 25%. Almost 300,000 companies in the U.S., U.K., France, and Germany have closed. The financial crisis caused by subprime loans in the U.S. from 2007 to 2008 caused a trillion-dollar loss in global subprime mortgages and dealt a massive blow to the worldwide economy, spreading to all sectors and lasting for several years. As the financial sector grows, the threat of financial crises gradually expands. The risk of financial problems expands as the financial industry grows in size and complexity [1]. World Bank estimates that the global economy will contract by 5.2% in 2020, which will be the deepest recession since the Second World War. With the spread of the epidemic, global financial markets have experienced significant volatility, and countries' financial markets have been hit to varying degrees. The global economy is facing a new round of crises. Aspired by the current situation, in this paper, two questions will be discussed, can we predict the financial crisis, and what are the main reasons for the financial crisis?

© 2023 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/5/20220097

2. Literature Review

This paper begins with a short review of the literature regarding the method of forecasting financial crises. Current studies use various ways with different variables and models to extrapolate the probability of a financial crisis in a country and whether there will be a financial crisis or not. A series of recent studies have indicated that the linear regression model is appropriate and quickly forecasts the financial crisis. However, the Logistic regression model will be much more convenient when a dichotomy exists [2]. Another technique used by the previous studies is MLP. MPL has demonstrated high accuracy in forecasting financial time series, macroeconomic factors, and market trends. By this method, [3] identified a significant reduction in investment risk because of the low correlations with developed countries' financial markets.[4] construct an analytic network process model for financial crisis forecasting. It provides a judgmental predictive structure to evaluate a certain number of heuristics when data is difficult to predict statistically promptly. The relationship between risk and return in the financial market has also been analyzed on target to predict the financial crisis.[5] employed several methods, including standard deviation, semi-deviation beta, and downside beta. And they concluded two parameter asset pricing model might help predict the crisis. A lot of investigation has also been carried out to determine the main reasons behind the financial crisis using modeling. [6] constructed a very and fascinating statistical model base on a panel of LIBRO-OIS (London Interbank Offered Rate and overnight index swap) spreads (including variation across banks, currencies, and terms) and bank CDS (credit default swap) rates. Their results show that liquidity and credit rates are significant factors affecting financial crisis. Especially when considering in terms of 1 to 3 months, liquidity becomes increasingly essential.

3. Methodology

3.1. Data

3.1.1. Data Are Complementary

Since a few of the economic data haven't been recorded in the bibliography of this work, especially for the data during the two world wars, Kernel and Nearest Neighbour Imputation (KNN Imputation) was used to fill the empty cells. The purpose of KNN imputation is to find K nearest neighbors for missing data (incomplete instances) from all complete models (without missing values) in a given data set. Then, if the target feature (or attribute) is definite, the missing data is filled with the data with the highest frequency among neighbors, which is called the majority rule. If the target feature is numerical, the lost information is filled with the average of the neighbors, which is called the mean rule. Compared with its competitors, KNN is relatively more straightforward, more easy-understanding, and more accurate, which makes it very popular in real data processing applications, such as surveys conducted at Statistics Canada, the US Bureau of Labor Statistics, and the US Census Bureau [7].

3.1.2. Data Transformation

To make the features of data more salient and make it easier to find the trends within these figures, four new variables were formed by transforming the existing data. The first one is dlrgdp, which was made according to the change in the log of real GDP of two consecutive years. This variable can show the growth rate of real GDP annually. The second one is dlpc. Similarly, this one shows the change in the log of CPI in two consecutive years, indicating the annual CPI inflation of each country. The third one, dlcred, shows the difference in the record of total bank assets in two consecutive years. [8] constructed a few models and proved that from 1860 to 1913, bank assets could be influenced

DOI: 10.54254/2754-1169/5/20220097

dramatically during a financial crisis. The fourth variable was loans, which shows the change in the log of loans to private sectors in two consecutive years.

3.2. Models

3.2.1. Generalized Linear Model (GLM)

As mentioned above, this paper aims to discuss whether financial crises are predictable and determine factors that have a more significant impact in the formation of financial problems. To achieve these targets, a class of generalized linear models was employed to analyze the economic data of 14 countries. Linear models customarily embody both systematic and random (error) components [9]. The unexpected part of a GLM is the probability distribution of the response variable. And the systematic component included the exploratory variables that combine to form the linear predictor of the GLM.

3.2.2. Logistic Regression Model

The logistic regression model has become a widely used and accepted method of analyzing of binary outcome variables [10]. The popularity stems from the availability of computer languages (in this paper, R) and software in computers that can be used to construct logistic regression models. Logistic regression is the combination of linear regression and a logistic function which can be simplified as $y = w * x + b$. In this function, w and b are unknown variables, and y is a dummy variable. A standard logistic function is $f(x) = \frac{1}{1+\exp(-x)}$, and in the logistic process, thex becomes the linear regression output. When the production of the linear regression part is greater than 1, the result of the sigmoid function will be greater than 0.5, and we will give a dummy variable a value of 1, indicating that there is a financial crisis in this year. Similarly, when the output of the linear regression part is less than one, the production of this sigmoid function will become less than 0.5, so the dummy variable will be given a value of 0, indicating that there is no financial crisis in this year. This model is constructed by using R.

4. Data Analysis

4.1. Data Processing

The original dataset contained data starting from the year 1870 up until 2008 and collected data from 14 countries. It had indicators of the economy such as bank assets, broad money, little money, loans, and GDP. The data also contained the ratio between all the different variables within each country to show the economic trends.

Though containing data from 1870 to 2008, a significant portion is missing from our dataset. After analyzing the data, we noticed that the majority of years where multiple columns are missing are pre-1900. Therefore, we completely removed all years before the year 1900. For the rest of the years, with missing data, we used the k-nearest neighbor's algorithm to fill in the mixing cells.

After inspection of the dataset, we could see a clear difference in trends both before and after World War 2.

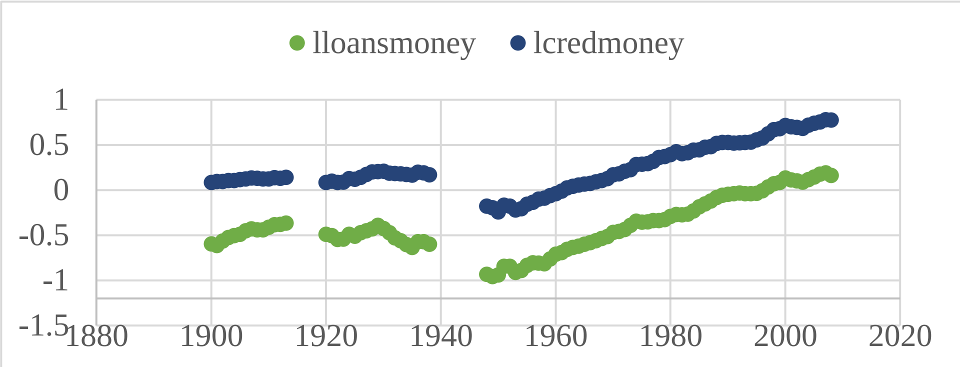

Figure 1: The ratio of loans to money and bank assets to money before and after World War 2.

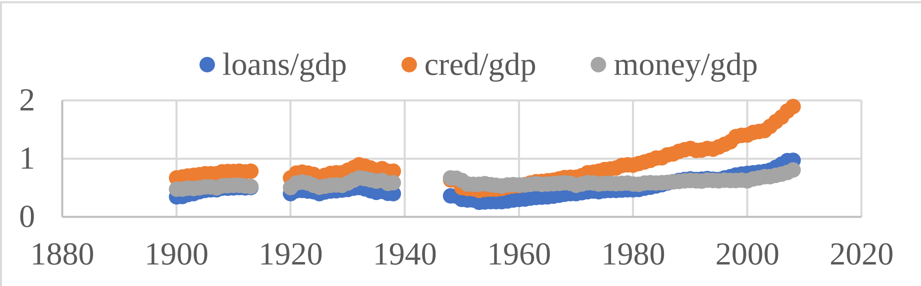

Figure 2: The trend of loans to GDP ratio, bank assets to GDP ratio, and broad money to GDP ratio before and after World War 2.

Since the war that took place between 1938-1945 ushered in a new financial era around the world. The economic and financial structure changed markedly during the pre-war and post-war financial periods. Before and after the Second World War, very different monetary and regulatory frameworks, macroeconomic policies, etc., played a different role. The economy started growing significantly faster, and increased growth in the global economy is visible. To counteract these change effects and better study the problem, we divided the data set into two parts based on time, before and after World War II. To exclude the results of the war on the financial system and to make the forecasts more general, we further exclude data from countries between 1938 and 1945.

We analyzed the data based on prewar and postwar, but we explored the different countries separately.

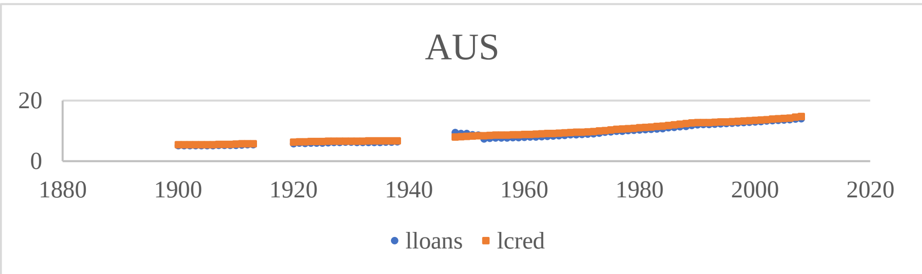

Figure 3: The trend of loans and total bank assets in before and after the war in Australia.

DOI: 10.54254/2754-1169/5/20220097

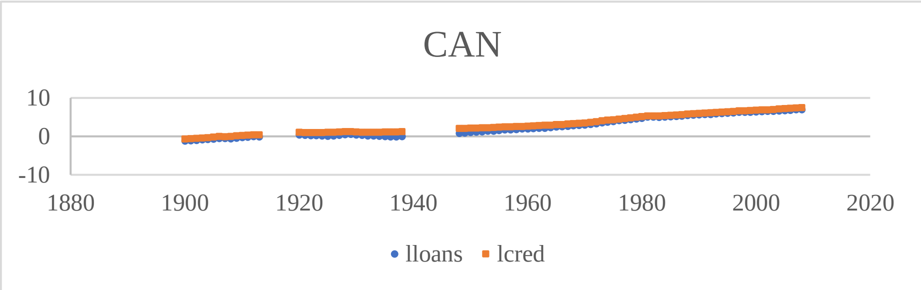

Figure 4: The trend of loans and total bank assets in before and after the war in Canada.

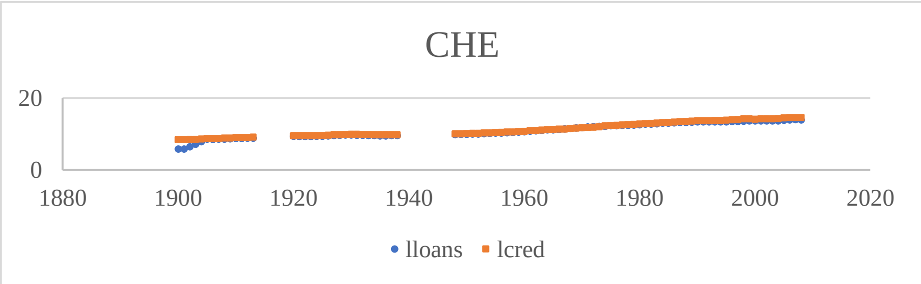

Figure 5: The trend of loans and total bank assets in before and after the war in Chile.

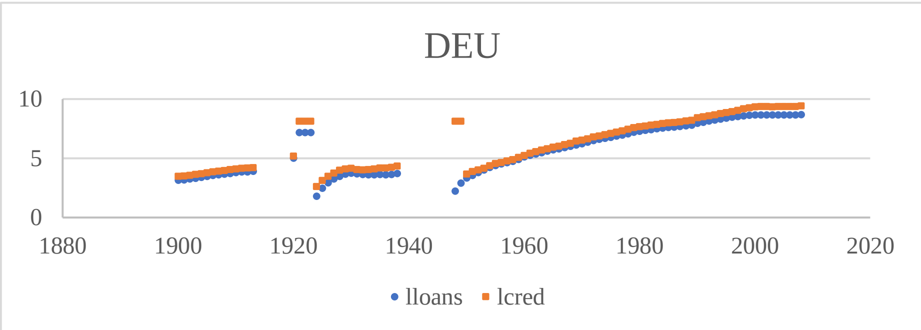

Figure 6: The trend of loans and total bank assets in before and after the war in Deutschland.

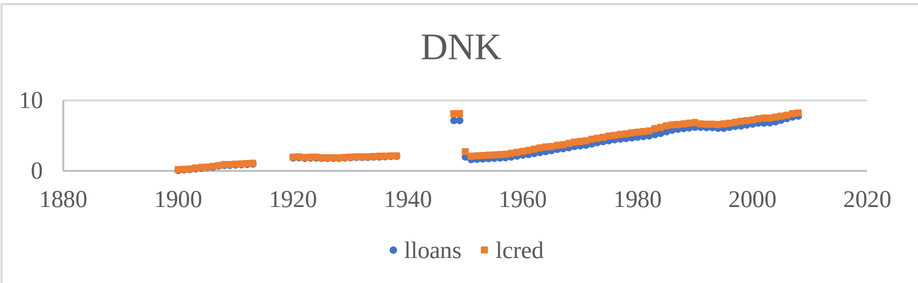

Figure 7: The trend of loans and total bank assets in before and after the war in Denmark.

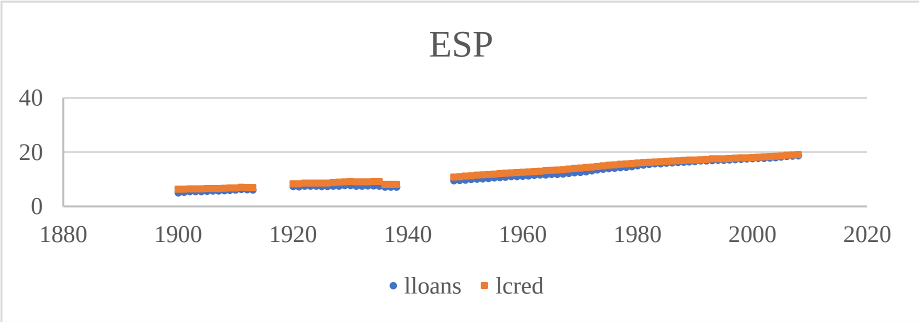

Figure 8: The trend of loans and total bank assets in before and after the war in Australia.

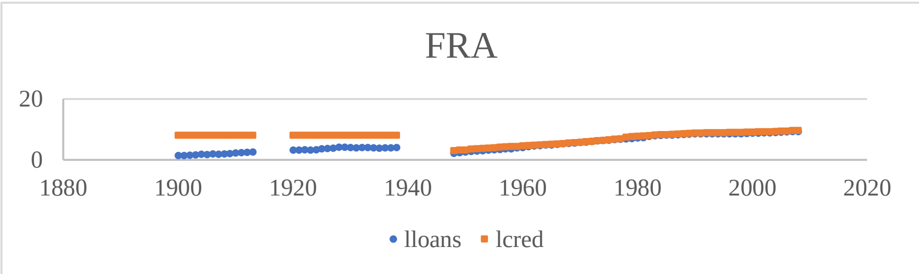

Figure 9: The trend of loans and total bank assets in before and after the war in Spain.

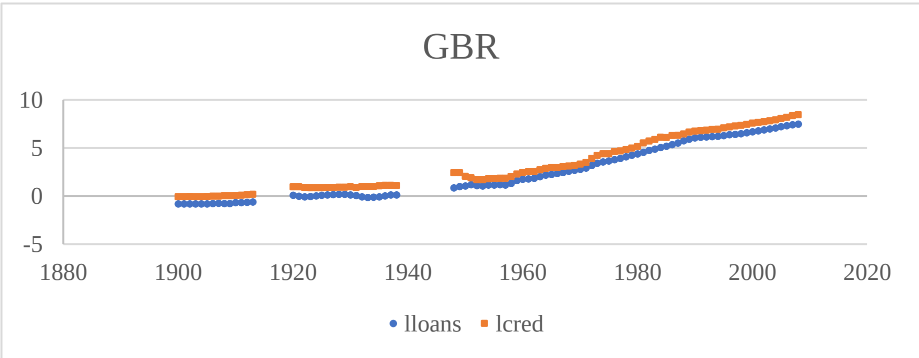

Figure 10: The trend of loans and total bank assets in before and after war in England.

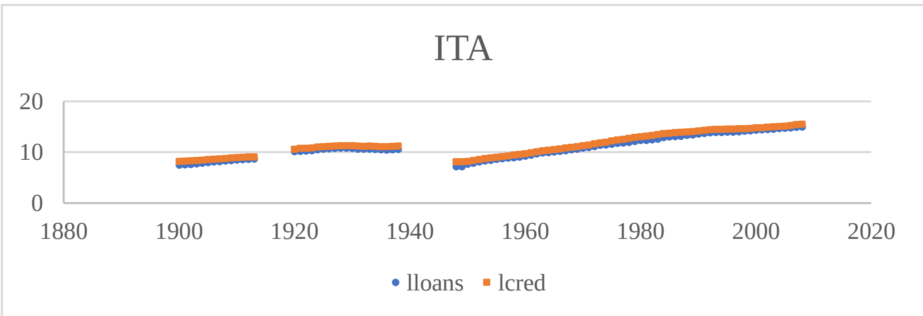

Figure 11: The trend of loans and total bank assets in before and after the war in Italia.

DOI: 10.54254/2754-1169/5/20220097

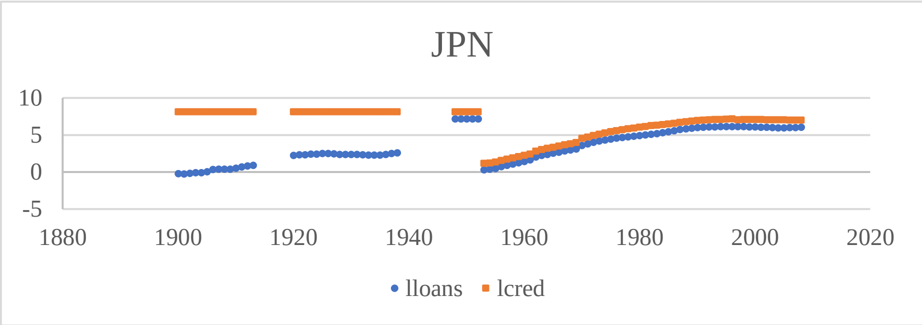

Figure 12: The trend of loans and total bank assets in before and after war in Japan.

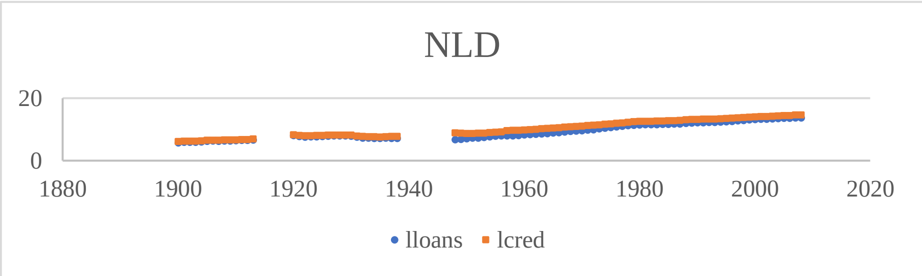

Figure 13: The trend of loans and total bank assets in before and after war in Dutch.

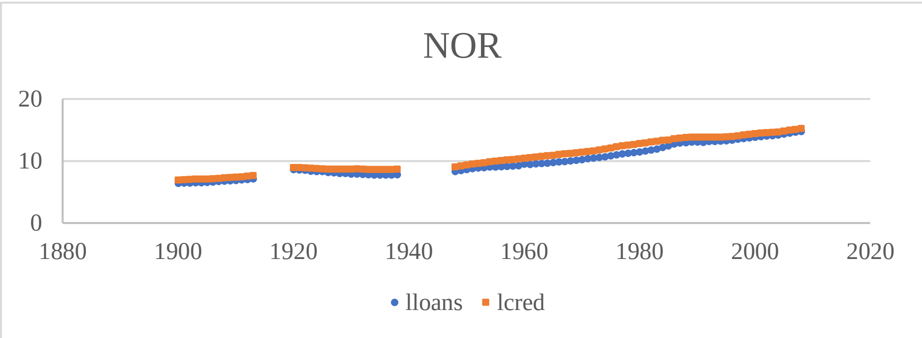

Figure 14: The trend of loans and total bank assets in before and after war in Norway.

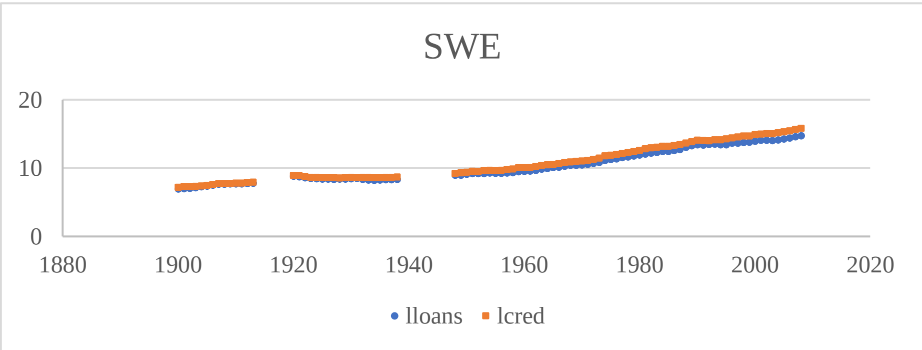

Figure 15: The trend of loans and total bank assets in before and after war in Sweden.

DOI: 10.54254/2754-1169/5/20220097

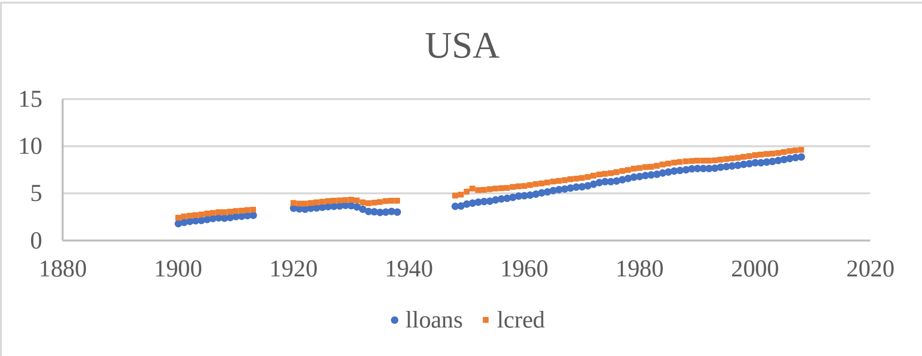

Figure 16: The trend of loans and total bank assets before and after the war in America.

The difference in the trends from before and after the war is significantly larger than the difference between the individual countries. Therefore, we only separated them into pre- and post-war, and if we further break the data down, we will not have enough data to create an effective model to predict the crisis.

We then separated the dataset into training and testing sets. Since we are analyzing pre- and post-war separately, we need two of each test. Our data is time-series data from 1900 to 2008, but since it is a time series, we couldn't just randomly pick points since the data would be discontinued. Therefore, we selected the first 75% of the years as the training set, used to train and construct the model, and the rest of the years as the testing set, used to see the effectiveness of our model.

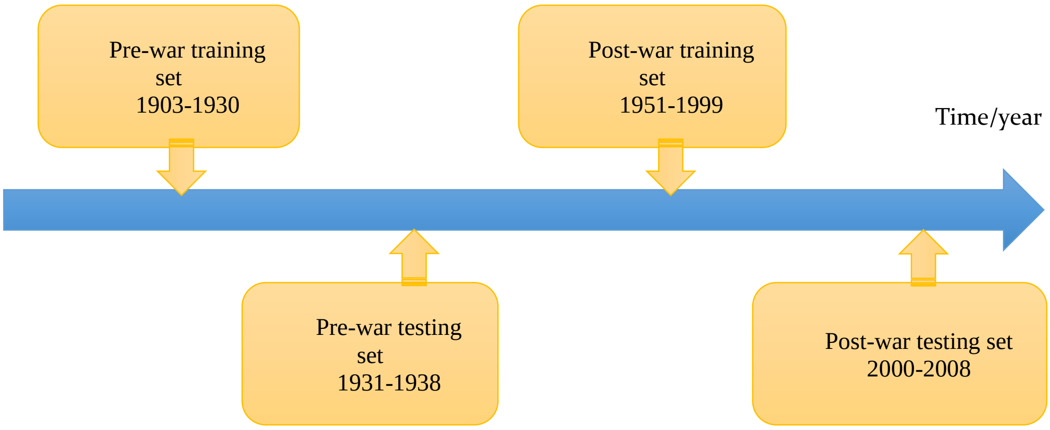

Figure 17: Time ordering of training and test set data.

Also, we introduce four-time lag terms, which are dlrgdpt, dlpct, dlcredt, dlloanst . These variables represent the growth rate of real GDP, CPI inflation, the growth rate in bank assets, and growth in loans in t years before, respectively, and they are defined as:

$$dlrgdp_t = lrgdp_t - lrgdp_{(t-1)}$$

$$dlpc_t = lpc_t - lpc_{(t-1)}$$

$$dlcred_t = lcred_t - lcred_{(t-1)}$$

$$dlloans_t = lloans_t - lloans_{(t-1)}$$

4.2. Prediction Model

We choose a logistic regression model since our goal is to determine the probability of a financial crisis by using macroeconomic indicators to predict whether a financial crisis will ultimately occur.

To make the most appropriate use of the variables in our dataset and the time-lagged items, we have introduced. Therefore, we use the backward stepwise regression technique to automatically build a model as a reference.

4.2.1. Pre-war Model.

First, we distinguish between the concepts of lloansmoney, lcredmoney, and loansmoney, credmoney. Since the first set of variables is the logarithmic take of the second set of variables and has the same economic meaning, we bring it into the model and perform the AIC and Anova tests, respectively.

Table 1: AIC and Anova test result for a model with different variables.

	AIC	Pr (> Chi)
Model with *lloansmoney, lcredmoney*	185.47	0.601
Model with *loansmoney, ceremony*	184.04	0.743
Full Model	188.45	

Based on the results of the tests, we finally dropped the second set of variables.

Because our variables are taken from macroeconomic indicators, they are, to some extent, relevant. Firstly, we performed a collinearity test for the remaining variables.

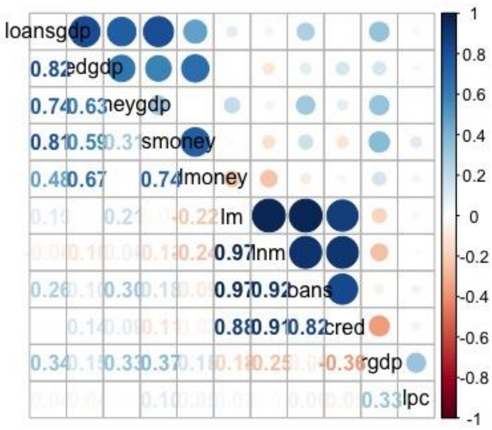

Figure 18: The correlation coefficient between each variable in the pre-war model.

According to the correlation test, the correlation coefficient between loansgdp and loan money is 0.75, and between ceremony and credgdp is 0.7. To eliminate the error caused by the covariance problem, we can, among the two groups of loansgdp, credgdp, and credmoney, loans money, Delete one group of variables. At the same time, the share of credit to cash better reflects the change of leverage compared to the percentage of credit to GDP, which is crucial for such a financial crisis, so we deleted the first group. The chi-square and AIC tests support our view that the model after deletion has better results.

DOI: 10.54254/2754-1169/5/20220097

It is clear from the historical data and literature studies that systemic financial crises tend to occur before rapid credit expansion, and excessive credit growth implies an increase in non-performing bank loans, which can lead to a collapse of credit if some part of the social credit system is damaged. In this case, the financial system can be shown to disruption.

Therefore, lm is a better explanation and predictor of the financial crisis in the economic sense than lnm. Based on their correlation coefficients and chi-square tests, we delete lnm and keep lm.

From the historical data and studies, it is clear that the occurrence of a financial crisis cannot be judged by the quantity of credit alone, which varies significantly from country to country and from period to period, and we tend to focus on the change of ratio. We can conclude that the AIC of the model decreases after excluding the two variables, lloans and lured, and the chi-square test is significant.

In addition, we find that the coefficients for lrgdp, ceremony, loansmoney, and lpc are not significantly different from zero ($P > 0.05$) when we regress our model with all remaining variables so far. We also use the AIC test and Chisq-test with the model without these two terms vs. the model with them, pointing all to removing them from the model.

After that, we introduced the time series term based on historical studies and cardinality tests, and we chose to retain dlcred and dlloans3.

Finally, we have decided to look into interaction terms between all variables. Investigating them through t-test, F-test, and the influence on the sum of squares, we found that the interaction between credgdp and dlloans and loansgdp increased the model's predictive power.

4.2.2. Post-War Model

Similarly, we distinguish between the concepts of loan money, lcredmoney, and loan money, ceremony. Since the first set of variables is the logarithmic take of the second set of variables and has the same economic meaning, we bring it into the model and perform the AIC and Anova tests, respectively.

Table 2: AIC and Anova test results for the model with different variables.

	AIC	Pr (> Chi)
Model with lloansmoney, lcredmoney	110.28	0.45
Model with loansmoney, ceremony	108.7	0.99
Full Model	112.69	

Based on the results of the tests, we finally dropped the second set of variables. Furthermore, we perform a correlation test for the post-war data set.

DOI: 10.54254/2754-1169/5/20220097

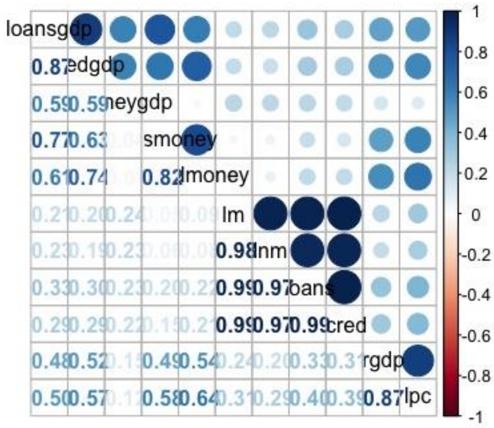

Figure 19: The correlation coefficient between each variable in the post-war model.

As a result of the post-war period, financial styles have changed considerably. The average growth rate of broad money was much smaller than that of loans and assets after World War II, and the critical leverage ratio, the money ratio, increased nearly 20 times after World War II. Both loans and investments began a strong long-term upward trend relative to broad money. Therefore, we start with the selection of variables based on characteristic importance. Based on the postwar economy and finance characteristics, we chose and lcred, lpc, lrgdp, and lloans for initial model building.

The results of the chi-square test and the AIC test support our elimination of these two variables.

In the post-war period, domestic credit was used to finance the investment activities of domestic firms. Therefore, domestic credit growth was significant in driving a country's economic growth, but with it came higher financial risk due to regulation, policy, and other factors, so we introduced dlcred3, loans, which is the growth rate of credit three years ago, indicating changes in financial risk and participates in the forecasting of the model. Since we know that a credit boom is marked by a rise in both the price index and the quantity of credit and that a credit boom may be one of the triggers of a financial crisis, we chose variables representing the CPI index and GDP index, which is dlrgdp, dlrgdp2, dlpc2. And ANOVA tests and AIC tests were used to select the variables to participate in the model. The test result pointed in favor of adding these variables.

4.3. Model Check

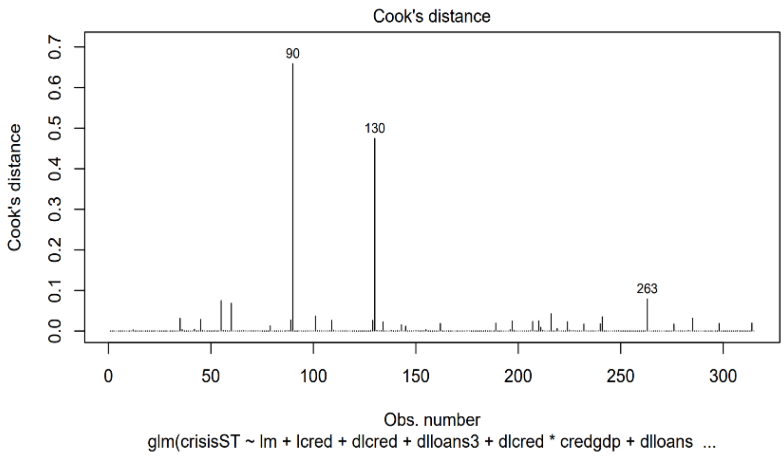

Figure 20: Cook’s distance.

DOI: 10.54254/2754-1169/5/20220097

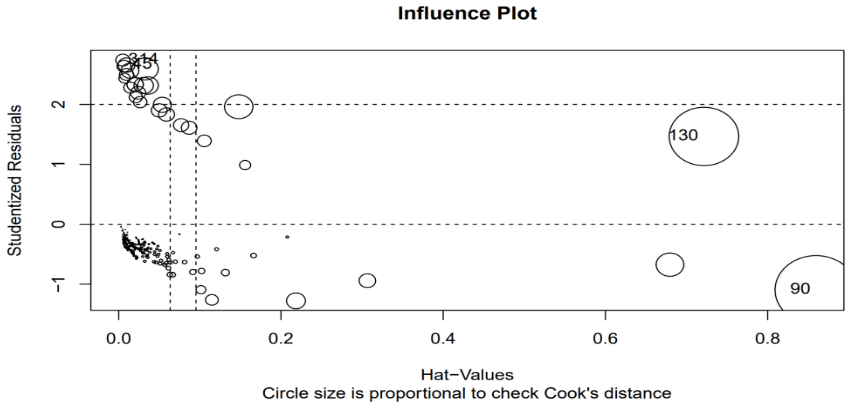

Figure 21: Influence Plot.

In the model check, we tried to find some unusual values that might appear in the data and used them to eliminate or reduce their impact on the model.

To find high leverage points, strong influence points, and outliers present in the data, two plots are used to plot the approximate distribution of data in the dataset, one for detecting the cook's distance and one for the influence of the data. In the cook's distance plot, two lines stand out as having a much higher cook's length than the rest. These two lines correspond to Obs. numbers 130 and 90, and their cook's distance is around 0.65 and 0.5, respectively.

As shown in the influence plot, the larger the circle, the larger the Cook's distance, so it can be seen that 130 and 90 correspond to large Cook's spaces, which leads to two points that can be considered solid influence points and can have a significant impact on the estimation of the model parameters.

Meanwhile, they are also high-leverage points. This is because they have a Hat-value too large or too small and will pull the fitted model straight toward itself. We can see that 90 and 130 correspond to Hat values greater than 0.7, while most points have values below 0.1. There are also two outliers in the graph, 314 and 45, with large residuals greater than 0.2.

After looking at the information in rows 130 and 90, we find that the values for dlcred and loans in this country for this year far exceed the data for other years and other countries. Since these variables and their interaction terms significantly affect the results, these two points become high-leverage points and influential solid points.

However, we do not intend to remove these points because the data for each country are accurate, and they are not significantly recording errors, so we have to consider them. Part of the reason these points are influential solid points and high leverage points is that the overall sample is so tiny that the issues with insufficient data significantly affect the overall model results, so we can address these points in the future by adding variables or increasing the amount of data.

DOI: 10.54254/2754-1169/5/20220097

5. Model Result

Table 3: Pre-war model result.

	Estimate Std.	Error	Z-value	Pr(>|z|)
(Intercept)	-2.6834	0.7378	-3.637	***
lm	-0.7155	0.2915	-2.455	*
lcred	0.7431	0.2984	2.49	*
dlcred	18.4823	5.1256	3.606	***
dlloans3	1.0984	0.6159	1.783	.
credgdp	-2.8339	2.0165	-1.405	
dlloans	-15.9152	4.8877	-3.256	**
loansgdp	2.9657	2.2095	1.461	
dlcred:credgdp	-11.3942	4.0915	-2.785	**
dloans:loansgdp	9.9703	4.2421	2.35	*

5.1. Post-war Model Result

After filtering the variables, we identified the appropriate ones and calculated their slopes and intercepts.

Table 4: Post-war model result.

	Estimate Std.	Error	Z-value	Pr(>|z|)
(Intercept)	-23. 4737	7.2796	-3.226	***
lnm	-0.8770	0.3999	-2.193	*
lcred	0.8560	0.4304	1.989	*
lpc	3.0815	1.4006	2.200	*
dlrgdp	-24.4098	16.4467	-1.484	
dlrgdp2	48.7996	22.4958	2.169	*
dlpc2	25.3947	13.3619	1.900	.
dlcred3	8.9617	3.7388	2.397	*
dloans3	7.0874	3.9857	1.791	.

The summary of the final pre-war model is shown in the figure. According to the final logistic regression model, the occurrence of financial crises was negatively correlated with the interaction terms of log of broad money, bank assets/gdp, log bank loans growth rate, log bank assets growth rate, and bank assets/gdp, and positively correlated with the other variables. The growth rate of real log gdp and the growth rate of log of the CPI price level is highly significant. The growth rate of log bank loans and assets is highly effective as they have significant coefficients.

Nevertheless， the post-war model shows a very different message. After the end of World War II, the occurrence of financial crises is negatively correlated with the log of little money and growth rate of log real GDP, while positively correlated with the record of CPI price level and its growth rate, log bank assets, and its growth rate, and the growth rate of log bank loans. Notably, the growth rate of real log GDP and the growth rate of log of the CPI price level is highly significant.

DOI: 10.54254/2754-1169/5/20220097

5.2. Model Performance

Two methods are used here to determine the model's performance: AIC and R-squared.

AIC is the Akaike information criterion, a measure of the goodness of fit of a statistical model that weighs the complexity of the estimated model against the goodness of fit of this model to the data. Since AIC is a model designed to best explain the data but contains the least number of free parameters, and a considerable AIC value represents overfitting, a smaller AIC value represents a good fit and is not overly complex. According to the table, the training set for the pre-war dataset has an AIC value of 168.2, for post-war dataset has an AIC value of 90.24. At the same time, AIC of the testing set is 46.8 for pre-war and 62.12 for post-war.

Table 5: AIC Value.

	Training set	Testing set
Pre-war	168.2	46.8
Post-war	90.24	62.16

Table 6: R-square.

	Training set	Testing set
Pre-war	0.0017	0.0474
Post-war	0.0064	0.0317

R-squared represents the proportion of the total variation in the dependent variable that can be explained by the independent variable through the regression relationship. It shows the fit between the estimated values and the corresponding actual data. According to the table, the training set for the pre-war dataset has an R- squared value of 0.0017; the post-war dataset has an R- squared value of 0.0064. At the same time, the R-squared value of the testing set is 0.0474 for pre-war and 0.0317 for post-war. Also, the low R-squared values indicate the high generalizability of the model.

Here, we can see that the value of R-squared in the test set is larger than the value of R-squared in the training set, while the value of AIC is smaller than the value of AIC in the training set, so we can conclude that the model performs better in the test set than in the training set.

5.3. Predicting Result

To successfully predict the probability of the final financial crisis occurring to determine whether this time would occur in the current year, I used the prediction function in RStudio to predict the pre-and post-war situation based on a model built from the training data set. The test data was put into the pre-war and post-war models constructed from the training set to predict the probability of the final financial crisis occurring. From the output, there were very few numbers above 0.1, meaning that in most years, the likelihood of the problem occurring was less than 10%, a value that was not sufficient for me to make a yes or no judgment, as in normal circumstances these values would represent an unlikely occurrence. However, there were still some years when it did occur. I then summarized the two figures, which showed that the pre-war mean was around 7% and below 10%, while the post-war mean was just over 2%, about a third of the pre-war value, and the pre-war and post-war maxima were 15.9% and 7.2% respectively. Here, we can see that the probability of a financial crisis is much greater in wartime than in post-war due to the instability of the various data. Since the likelihood of a financial crisis is much lower than 50%, in this case, we cannot predict a financial crisis because we cannot choose a reasonable probability to say that we expect the occurrence of a financial crisis.

DOI: 10.54254/2754-1169/5/20220097

6. Conclusion

In summary, I am aware of several problems and limitations with our model. Firstly, the dataset size is not large enough as we only used 30 years before and 50 years after the war for 14 countries, which led to bias in the model and poor model performance when we ran the tests. In addition, the crisis is almost impossible to predict because it is a small probability event, which means that the probability of it happening is very small in a vast dataset (it can even occur when the chance is 1% if there are many years when it is 1%). As I said earlier, it is difficult to tell the probability of it happening because our models are not sophisticated enough to summarise all the useful variables and information to manage it.

To improve our model and its performance, we can increase the size of the dataset, add more years and countries, attenuate the impact of strong and high leverage points, and build and fit the model. At the same time, we could look for additional influential variables that might help to provide the model better. Also, find more efficient ways to do more complex models to predict the crisis better.

Reference

[1] *Schularick, M., & Taylor, A. M. (2009). Credit booms gone bust: Monetary policy, leverage cycles and financial crises, 1870-2008. National Bureau of Economic Research.*

[2] *Lee, K., Lee, H. and Kim, S. (2007) Factors Influencing the Adoption Behavior of Mobile Banking: A South Korean Perspective. Journal of Internet Banking & Commerce, 12, 1-9.*

[3] *Harvey, C. (1994). Predictable risk and returns in emerging markets. https://doi.org/10.3386/w4621*

[4] *Niemira, & Saaty, T. L. (2004). An Analytic Network Process model for financial-crisis forecasting. International Journal of Forecasting, 20(4), 573–587. https://doi.org/10.1016/j.ijforecast.2003.09.013*

[5] *Sehgal, S., & Pandey, A. (2018). Predicting financial crisis by examining risk-return rela tionship. Theoretical Economics Letters, 08(01), 48–71. https://doi.org/10.4236/tel.2018.81003*

[6] *Gefang, D., Koop, G., & Potter, S. M. (2010). Understanding liquidity and credit risks in the financial crisis. Journal of Empirical Finance, 18(5), 903–914. https://doi.org/10.1016/j.jempfin.2011.07.006*

[7] *Chen, & Shao, J. (2000). Nearest Neighbor Imputation for Survey Data. Journal of Official Statistics, 16(2), 113–.*

[8] *Baker, M., & Collins, M. (1999). Financial crises and structural change in English Com mercial Bank assets, 1860–1913. Explorations in Economic History, 36(4), 428–444. https://doi.org/10.1006/exeh.1999.0727*

[9] *McCullagh, P., & Nelder, J. A. (2000). Generalized linear models. Champman and Hall/CRC.*

[10] *Hsomer, D. W., Hosmer, T., Le Cessie, S., & Lemeshow, S. (1997). A comparison of g oodness-of-fit tests for the logistic regression model. Statistics in Medicine, 16(9), 965–980. https://doi.org/10.1002/(sici)1097-0258(19970515)16:9<965::aid-sim509>3.0.co;2-o*