Journal of Economic Theory and Econometrics, Vol. 33, No. 2, Jun. 2022, 96–111

Can Lessons Learned from Past Economic Crises Help Predict Future Economic Crises?*

Chirok Han† Kwanho Shin‡

Abstract Based on experience of Korea's two recent economic crises in 1997 and 2008, we investigate if lessons learned from past economic crises can help predict future economic crises. Using the least absolute shrinkage and selection operator (lasso)-logit model, we find that the symptoms in the pre-1997 crisis period do not appear before the 2008 crisis. This finding suggests that the 2008 crisis could not be predicted using lessons learned from the precedent crisis. We also attempted to answer an even more hypothetical question of whether the first crisis could have been predicted if lessons from the second crisis had already leaned. Our findings suggest that the 1997 crisis could have been predicted using lessons learned from the 2008 crisis. Overall, our findings imply that factors that cause the future crisis encompass those of the past crisis, but it is difficult to predict a future crisis armed only with experience of a past crisis.

Keywords Economic Crises, Lasso, Lessons, Logit, Prediction.

JEL Classification E17, E32, G01

---

*The authors offer thanks to Professor Dongheon Kim for providing information on data. Research by Han was supported by Korea University (K1810141).

†Department of Economics, Korea University, E-mail: chirokhan@korea.ac.kr

‡Corresponding author. Department of Economics, Korea University, E-mail: khshin@korea.ac.kr

Received March 17, 2022, Revised June 21, 2022, Accepted Jun 26, 2022

1. INTRODUCTION

Is it possible to learn lessons from past economic crises? More specifically, can we learn from past economic crises, in order to reliably predict future economic crises? Several attempts have been made to predict economic crises, with one of the most well-known attempts being the so-called "early warning system," which, in recent times, has utilized state-of-the-art technologies, such as machine learning.[1]

However, some prominent policymakers and economists dispute the possibility of predicting future crises. For example, Geithner (2014), a former U.S. Secretary of the Treasury, wrote that "financial crises cannot be reliably predicted, so they cannot be reliably prevented. They are kind of like earthquakes..." Ben Bernanke, a former Chairman of the Federal Reserve, added, "What we didn't recognize immediately was the vulnerability of the system to a run of short-term funding. ... It was an old-fashioned run in new clothes."[2]

In this paper, based on experience of Korea's two recent economic crises in 1997 and 2008, we investigate why it is difficult to predict future economic crises using the lessons learned from past crises. Using the 1997 and 2008 crises as a test bed, we examine if it would have been possible to predict the second crisis using lessons learned from the first. We also try to answer an even more hypothetical question of whether the first crisis could have been predicted if lessons from the second crisis had already been learned. Our analyses rely on lasso, one of machine learning methodologies. Our findings strongly suggest the answer of 'no' to the first question and 'yes' to the second.

One caveat to note is that the purpose of our study is not to find causes of a crisis. Lasso can only pinpoint the features that best signal pre-crisis symptoms. In this sense, caution is warranted in interpreting our findings. However, our findings raise serious concerns about the effectiveness of early warning systems that are designed to rely on past experience to predict future crises.

The remainder of the paper is organized as follows: In the next section, we exposit two hypotheses and a methodology to test them; in Section 3, we explain the results; Section 4 concludes the paper.

---

[1] See Fouliard and Rey (2020) and Bluwstein *et al*. (2020), among others.

[2] See an article published by Columbia Business School (https://www8.gsb.columbia.edu/articles/chazen-global-insights/financial-system-will-survive-says-ben-bernanke).

2. METHODOLOGY

As explained in the introduction, we aim to answer questions related to two economic crises in Korea in 1997 and 2008. One question (Q1) asks whether we could predict the 2008 crisis using lessons learned from the 1997 crisis, with the other (Q2) asking whether the symptoms prevailing in the period preceding the 2008 crisis were also prevalent in the period prior to the 1997 crisis. The second question is a hypothetical one that asks whether we could have predicted the 1997 crisis if we had experienced the 2008 crisis beforehand and learned lessons from it.

We answer Q1 by first identifying the characterizing symptoms in the period prior to the 1997 crisis (hereafter 'Pre97C') and then by predicting the probability of the same symptoms being present in the period before the 2008 crisis (hereafter 'Pre08C'). Note that the targets are the periods prior to the crises rather than during them like in the literature on early warning systems (see, e.g., Park *et al*. (2017). If we observe high probabilities of those symptoms occurring in the Pre08C period, it means that the symptoms in Pre97C were also prevalent in Pre08C, which would, in turn, imply that we could have predicted the second crisis, had we learned lessons from the first crisis. On the other hand, if the symptoms characterizing Pre97C are not observed in Pre08C, it would imply that predicting the 2008 crisis was impossible even if we learned lessons from the first crisis; the 2008 crisis was characterized by unprecedented symptoms.

Q2 is answered using a similar strategy: symptoms prevalent in the Pre08C period are first identified and then the probability of these symptoms occurring in Pre97C is predicted. A high probability of symptoms from Pre08C being present in Pre97C is interpreted as a signal of the traits of the second crisis being already prevalent before the former crisis, therefore, we could have predicted the first crisis if lessons from the second crisis had already been learned.

Possible combinations of answers to Q1 and Q2 are summarized in Table 1.

We next explain how the symptoms are identified in Pre97C and Pre08C for Q1 and Q2, respectively. The beginning month of the 1997 crisis is December 1997 and Pre97C constitutes one year prior to it, i.e., December 1996 to November 1997, which is the test period in our analysis for Q1. The corresponding comparison period is chosen conservatively. As it is difficult to pin down exactly when the crisis began and ended, we omit a certain period of months before the beginning of and after the end of the test period. Specifically, we omit one year before the beginning of the crisis and one year after the end of the crisis. The comparison periods begin in January 1995 due to data availability for some variables and end in December 2001, sufficiently far from the 2003 credit card

Table 1: Summary of possible outcomes of analysis

<table>
<tr><td colspan="2" rowspan="2"></td><td colspan="2">Probability of symptoms of Pre08C occurring in Pre97C</td></tr>
<tr><td>High</td><td>Low</td></tr>
<tr><td rowspan="2">Probability of symptoms of Pre97C occurring in Pre08C</td><td>High</td><td>Two crises are similar and the Korean economy could learn lessons from either crisis to predict the other.</td><td>The second crisis has symptoms of both the first crisis and some new features and the Korean economy could predict the 2008 crisis by learning lessons from the 1997 crisis, but not the other way around.</td></tr>
<tr><td>Low</td><td>Symptoms of Pre97C are not observed in Pre08C, but symptoms of Pre08C were present in the Pre97C period. The Korean economy, by learning lessons from observed features of the 1997 crisis, could not predict the 2008 crisis. However, had the 2008 crisis occurred beforehand, the Korean economy would have been warned about the 1997 crisis by learning lessons from the 2008 crisis.</td><td>Two crises have distinct characteristics.</td></tr>
</table>

Note. Analyses in this paper support the shaded area for the Korean crises.

lending distress event. The results are robust when the terminal month is adjusted by ± 6 months. Similarly, the 2008 crisis is believed to have started in September 2008 and the test period (Pre08C) is one year prior to the beginning of the crisis (September 2007 to August 2008); the associated comparison periods are January 2004 to December 2005 and March 2010 to December 2014. Again, the

Table 2: Sample periods

Question to answer	Beginning of crisis	Duration of crisis	Pre-crisis (test) period	Comparison periods
Q1	12/97	$\sim$ 1 year	Pre97C (12/1996–11/1997)	01/1995–11/1995, 01/1999–12/2001
Q2	09/08	$\sim$ 1 year	Pre08C (09/2007–08/2008)	01/2004–12/2005, 03/2010–12/2014

Note: Data are collected for 1995–2014.

comparison periods are chosen not to be contaminated by the 2003 credit card lending distress, and the choice of the terminal month is hardly consequential. The results remain robust to changes in the specification of the periods. The test and comparison periods are summarized in Table 2.

Thirty-two macro-economic indicators, listed in Table A.1 in the appendix, are considered as candidate features that characterize the test periods. The target variable, denoted y, is the binary indicator that takes on the value 1 if the month is in Pre97C or Pre08C, depending on the question to answer. Let X denote the vector of the feature variables (i.e., the predictors).

With the binary target variable, we base our analysis on the logit regression. A logit model relates the logit (log-odds, where the odds are $\frac{p}{1-p}$) of the event to a linear combination of features. That is, the logit model is formulated as

$$\log\left[\frac{P(y=1|X)}{1-P(y=1|X)}\right]=\beta_0+X\beta.$$

This model is equivalently written as $y=I(\beta_0+X\beta+u>0)$, where the error term u is assumed to have the standard logistic distribution with the cumulative distribution function being $\Lambda(x)=e^x/(1+e^x)$. Logit models are typically fitted by the maximum likelihood estimation, which maximizes the logit log-likelihood

$$\log L=\sum_{i=1}^{n}\left[y_i(\beta_0+X_i\beta)-\log(1+e^{\beta_0+X_i\beta})\right]. \tag{1}$$

Considering the large number of candidates to explain the target variable y, we combine the logit regression and the least absolute shrinkage and selection operator (lasso; see Tibshinari (1996)) to select a ‘best’ model in terms of the leave-one-out cross validation criterion. Automatic model selection methods such as lasso are an effective tool for predictive analyses although they are often undesirable for causal analyses.

The resulting lasso-logit regression minimizes the penalized loss function

$$-\frac{1}{n}\log L(\beta_0,\beta)+\lambda\sum_{j=1}^{p}|\beta_j|,$$

where $\log L(\beta_0,\beta)$ is the logit log-likelihood in (1). (See Friedman and Tibshirani (2010); Hastie and Tay (2021).) It is remarkable that the estimator depends on the scales of the feature variables due to the presence of the lasso penalty. We follow the standard practice of normalizing the feature variables before estimation by dividing each by its sample standard deviation. The coefficients are however reverted to the original scale after estimation.

The 'tuning parameter' λ plays an important role in the model selection. A small λ gives a small penalty to the coefficients, vice versa. In one extreme, if $\lambda = 0$, the lasso-logit model reduces to the plain logit, and coefficients of all predictors will be estimated to have nonzero effects. In the other extreme, if $\lambda = \infty$, then the lasso-logit loss function is minimized by $\beta = 0$, thus leading to the intercept-only model. For postive and finite λ values, the lasso excludes unimportant variables from the model by estimating their coefficients to be zero, and only important variables survive. Large λ values are associated with fewer variables with nonzero coefficients and smaller magnitudes for the coefficient estimates.

The best lasso tuning parameter is chosen to ensure a good performance in prediction. For cross-sectional data, the leave-one-out cross-validation is often used, which is similar to the 'jackknife' methods in the statistical literature. This method fits the model using all the observations but one and then predicts the target value of the excluded and evaluate the prediction error. This leave-one-out prediction error is computed for every observation, and the leave-one-out cross-validation selects the λ value that minimizes the sum of squares of those prediction errors. Often the λ value associated with the smallest error turns out to be too small, leading to an overfitted model. To overcome this problem, the so-called 'one standard error' ('1se') rule is used, which chooses the largest λ value such that the leave-one-out cross-validation prediction error is within 1 standard error of the minimum. In our analysis in Section 3, we use the '1se rule', but the results change little when the 'minimum rule' is used instead.

For time series data with errors possibly serially correlated, cross-validation is less obvious especially when the sample size is not sufficiently large. There are some theoretical results on conditions for the leave-one-out cross validation

Bergmeir and Koo (2018), and alternative cross-validation methods are available for models with time series data. But considering the limited sample sizes in our applications and the purpose of cross validation as a means to choose a tuning parameter, we take the simpler approach of trying various tuning parameters and checking robustness.

Finally, for each of the chosen tuning parameters, the coefficients on the predictors are estimated using the full sample. The probability of the Pre97C symptoms appearing in every month of the whole sample period is then predicted by $\Lambda(\beta_0 + X_t\beta)$, where $\Lambda(x) = \exp(x)/[1+\exp(x)]$ is the standard logistic cumulative distribution function as defined before, β is the vector of the estimated coefficients, and β_0 is the intercept estimate. Models to compare Pre08C and the corresponding comparison periods are similarly estimated, and the probabilities are predicted for every month. The lasso-logit model is fitted using the R 'glmnet' package (see Friedman and Tibshirani (2010)), which uses a "proximal Newton" algorithm, which is an iteratively reweighted penalized least squares (see Hastie and Tay (2021)). The R source codes are provided in Table A.2 in the Appendix. See also James *et al*. (2013).

3. RESULTS

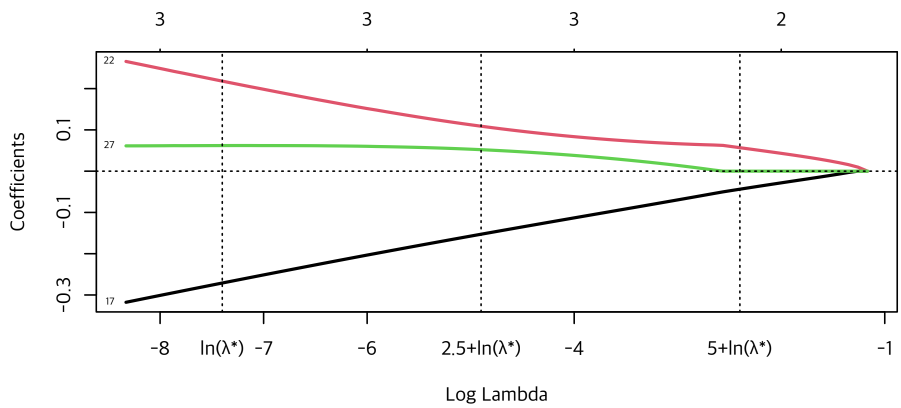
Data are collected from various sources including the Bank of Korea, the Bank for International Settlements (for the real effective exchange rates), Yahoo! Finance (for the Down Jones Index), and the Federal Reserve Bank of St. Louis (for the Federal Fund Rates, US term spread (10Y-3M), TED spread, and the Nikkei 255 index).

We first fit the model to explain the Pre97C period to answer the first question. The full lasso coefficient profiles corresponding to a range of tuning parameters are illustrated in Figure 1 against the logarithm of the tuning parameter, where λ^* is the estimated optimal value 0.00061. Three feature variables are found to have nonzero coefficients for the optimal tuning parameter. Figure 1 also presents two larger tuning parameters for comparison, which lead to coefficient estimates smaller in magnitude in absolute terms and thus better avoid overfitting.

The optimal (in terms of leave-one-out cross validation) lasso-logit model for Pre97C associated with the optimal tuning parameter is as follows.

The optimal model suggests that the best characteristic features before the 1997 crisis are rapid depletion of foreign reserves, higher manufacturing inventory index, and rapid growth of the Dow Jones index. While our model is not

Figure 1: Coefficient profile plot of coefficient paths for Pre97C



Note: λ^* is the tuning parameter value chosen by the 1se rule.

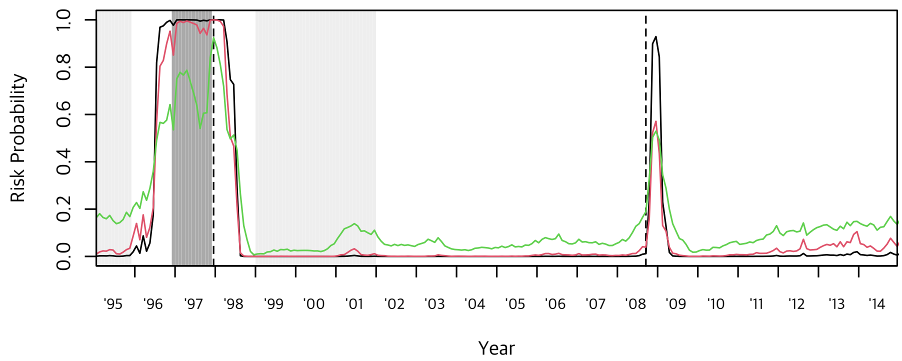
designed to find causal effects, the first two features are related to the fundamentals of the Korean economy. Before the 1997 Asian financial crisis, the Korean economy experienced a rapid decline in foreign exchange reserves, and manufacturers' inventories were increasing as large firms went bankrupt. The third feature is believed to be associated with capital flows. For example, foreign capital is more likely to be withdrawn from Korea when the stock market in the U.S is strong.

Table 3: Estimated coefficients in the optimal model for Pre97C

Variable name	Variable ID	Coefficient	Relative importance
Foreign reserves YOY	17	-0.2710	1.0000
Manufacturing inventories index	22	0.2182	0.6192
Dow Jones index YOY	27	0.0620	0.1401
Intercept		-24.8499	

Note: See Table A.1 for a full list of variables. The estimated lasso tuning parameter is 0.00061. 'Relative importance' is the absolute beta coefficient normalized to 1 for the largest, where a beta coefficient is the coefficient multiplied by the sample standard deviation of the corresponding explanatory variable.

Figure 2: Fitted and predicted probabilities for Q1



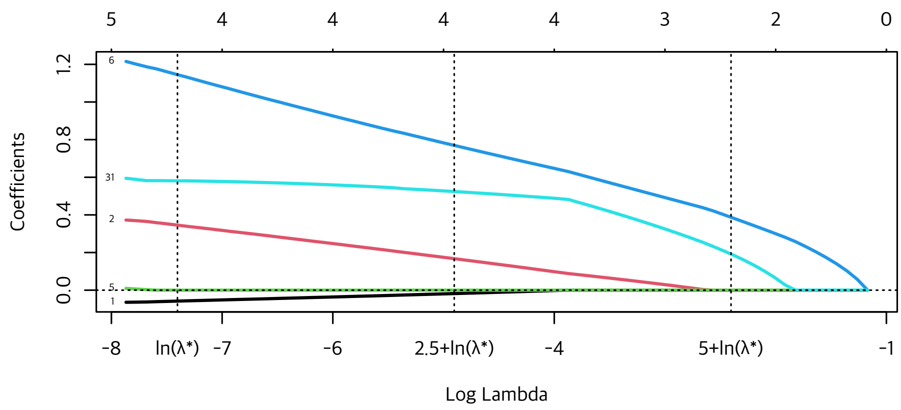
Note: The three lines correspond to the three lasso tuning parameters: the estimated optimal lambda ($\lambda^* = 0.00061$), $\exp(2.5)\lambda^*$, and $\exp(5)\lambda^*$. Smaller lambda values result in a better fit; larger values more shrinkage.

Figure 2 presents the predicted probabilities obtained using the results in Table 3 and two more sets of results with higher tuning parameters marked in Figure 1 to ensure avoiding overfitting. The area with dark vertical bars in Figure 2 indicates the test period (Pre97C) and the light vertical bars indicate the comparison months.

The feature variables characterizing Pre97C also appear *during* the 2008 crisis period (from September 2008), but not *before* the 2008 crisis. This confirms that the crisis in 2008 was a complete surprise to the Korean economy, and we conclude that it was not possible to predict the 2008 crisis using lessons learned from the precedent crisis. This result remains robust when larger tuning parameters are considered. We have also considered weighted lasso logit in order to address the class imbalance issue by giving quadruple weights to the treatment period observations, in which case two extra variables (the Nikkei Index YOY and the TED spread) are picked up, but the predicted probabilities are qualitatively the same.

Turning to the second question, the training sample period is now set to the 2004–2014 period with the test period being the Pre08C months and the comparison period the rest except the ‘gray’ period (see Table 2). The full lasso coefficient profiles are illustrated in Figure 3 against the logarithm of the tuning parameter, where λ^* is the estimated optimal value 0.000475. As before, λ^* and two extra tuning parameters, with larger magnitude, are considered for

Figure 3: Coefficient profile plot of coefficient paths for Pre08C



Note: λ^* is the tuning parameter value chosen by the 1se rule.

comparison and robustness check.

Four feature variables are chosen by the lasso-logit procedure for the optimal tuning parameter. The estimated coefficients are given in Table 4.

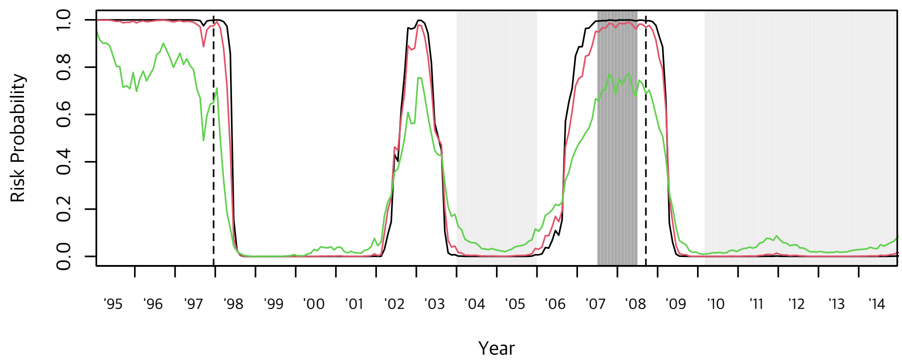
Note that the feature variables in Table 4 do not appear in Table 3. The most prominent features characterizing the Pre08C period include the growth rate of private credit, TED spread, the growth rate of M2 and the Federal funds rate.

Table 4: Estimated coefficient in optimal model for Pre08C

Variable name	Variable ID	Coefficient	Relative importance
Private creditYOY	6	0.9568	1.0000
TED spread	31	2.7760	0.2883
M2 YOY	2	0.3329	0.2536
Federal funds rate (1 year before)	29	-0.2528	0.0754
Intercept		-16.9919	

Note: See Table A.1 for a full list of variables. The estimated lasso tuning parameter is 0.000475 and the associated degrees of freedom are 4. 'Relative importance' is the absolute beta coefficient normalized to 1 for the largest, where a beta coefficient is the coefficient multiplied by the sample standard deviation of the corresponding explanatory variable.

Figure 4: Fitted and predicted probabilities for Q2



Note: The three lines are when the lasso tuning parameter is equal to the optimal lambda ($\lambda^* = 0.000475$), $\exp(2.5)\lambda^*$, and $\exp(5)\lambda^*$, respectively. Smaller lambda values result in a better fit; larger lambda values result in more shrinkage.

These new features, while not causal, are consistent with the findings by studies on factors contributing to the 2008 crisis. For example, increases in private credit (or M2) are an indication of financial imbalances that were one of the major causes of the 2008 crisis. The TED spread and the Federal funds rate are also sensible features because the crisis originated from advanced economy countries, especially the U.S.

The predicted probabilities obtained using the results in Table 4 and the extra results corresponding to the two larger tuning parameters marked in Figure 3 are plotted in Figure 4. As before, the dark vertical bars indicate the test period (Pre08C) and the light vertical bars indicate the comparison months. The results are robust when the pre-crisis comparison period is shortened (e.g., to Year 2013).

In contrast to Figure 1, the symptoms characterizing the Pre08C period (the dark shaded area in Figure 4) are also prevalent in Pre97C. This suggests that the Pre08C symptoms were already prevalent during the periods prior to the 1997 crisis. As such, if we had experienced the 2008 crisis (or a crisis homogenous to the 2008 crisis) beforehand and we had not adjusted our behavior in accordance, we would have been given some warnings before the outbreak of the 1997 crisis. The observed spike in the late 2002 and early 2003 period seems related with the credit card lending distress in 2003.

4. CONCLUSION

In this paper, based on the experience of Korea's two recent economic crises, we investigate if lessons learned from past economic crises are helpful in predicting future economic crises. Using a machine learning method, we select, among as many variables as possible, the most important variables that predict the occurrence of a crisis. One important drawback of the machine learning method is, however, that it does not uncover causal relationships. For this reason, we tried to refrain from economic interpretations as much as possible. At the same time, however, we also demonstrated how consistent our findings are with existing economic theories and empirical analyses of the two crises.[3]

We find that the symptoms in Pre97C do not appear before the 2008 crisis, which suggests that we could not predict the 2008 crisis using lessons learned from the precedent crisis. We also tried to answer an even more hypothetical question of whether it would have been possible to predict the first crisis if we had already learned lessons from the second crisis. Unlike what is suggested in the previous case, the symptoms characterizing Pre08C are also prevalent in Pre97C, suggesting that we could have predicted the 1997 crisis if we had learned lessons from the 2008 crisis. While it is not realistic to predict a previous crisis using the results derived from a future crisis, we conducted this experiment to find out to what extent variables useful in predicting future crises still played a role in predicting past ones, though not selected as key predictors. Overall, our findings suggest that factors that cause a future crisis encompass those of the past crises but it is difficult to predict a future crisis using only lessons learnt from past experience.

[3]We thank an anonymous referee who suggested that the role of the Chinese economy in the global economy had grown throughout the 2000s and the China factor may be crucial in accounting for the differences between the two crises. Explorations of this association are left for future research.

REFERENCES

Bergmeir, C, R. J. Hyndman, and B. Koo (2018). "A note on the validity of cross-validation for evaluating autoregressive time series prediction," *Computational Statistics and Data Analysis* 120, 70-83.

Bluwstein, K., M. Buckman, A. Joseph, M. Kang, S.Kapadia and O. Simsek (2020), "Credit Growth, the Yield Curve and Financial Crisis Prediction: Evidence from a Machine Learning Approach," *Staff Working Paper No. 848*, Bank of England.

Fouliard, J., M.J. Howell and H. Rey (2020). "Answering the Queen: Machine Learning and Financial Crises," *NBER Working Paper No. w28302*.

Friedman, J. H., T. Hastie, and R. Tibshirani (2010). "Regularization Paths for Generalized Linear Models via Coordinate Descent," *Journal of Statistical Software* 33(1), https://www.jstatsoft.org/article/view/v033i01.

Geithner, T. (2014). *Stress Test: Reflections on Financial Crises*, Crown, May 12, 2014.

Hastie, T., J. Qian, and K. Tay (2021). "An Introduction to glmnet," Vignette document of the R glmnet package."

James, G., D. Witten, T. Hastie, and R. Tibshirani (2013). *An Introduction to Statistical Learning: with Applications in R*, Springer.

Park, E., D. H. Kim, and K. Kim (2017). "The Study on the Development of the Financial Sector Early Warning System," *Journal of Economic Theory and Econometrics* 28(3), 36–67.

Tibshirani, R. (1996). "Regression Shrinkage and Selection via the Lasso," *Journal of Royal Statistical Society, Series B (Methodological)* 58(1), 267–288.

Table A.1: Macroeconomic indicators

No	Variable	Unit
1	M1 stock (end of month), YOY change rate	
2	M2 stock (end of month), YOY change rate	
3	Lf stock (end of month), YOY change rate	
4	Demand deposit turnover	
5	Demand deposit turnover, YOY change	
6	Loans (bank & nonbank), YOY change rate	
7	Currency stability security rate (1 yr)	%
8	Currency stability security rate (1 yr), YOY change	
9	Corporate bond yield (3 yr)	%
10	Corporate bond yield (3 yr), YOY change	
11	Producer price index, YOY change rate	
12	Consumer price index, YOY change rate	
13	House price index, YOY change rate	01/2019=100
14	Seoul apartment price index,YOY change rate	01/2019=100
15	Net commodity terms of trade, YOY change rate	
16	Income terms of trade, YOY change rate	
17	Foreign currency reserves, YOY change rate	
18	KRW/USD exchange rate, YOY change rate	
19	KRW/JPY exchange rate, YOY change rate	
20	Average manufacturing operation	%
21	Average manufacturing operation,YOY change rate	
22	Manufacture inventory index	2015=100
23	Manufacture inventory index,YOY change rate	
24	Construction order,YOY change rate	
25	KOSPI, YOY change rate	
26	Real effective exchange rate,YOY change rate	
27	Dow Jones index, YOY change rate	
28	US federal funds rate	
29	US federal funds rate, YOY change	
30	US term spread (10Y-3M)	
31	TED spread (3M-Libor)	
32	Nikkei Index, YOY change rate	

Note. Change rates are measured by log difference.

Table A.2: The R source codes

```
## Read data
z1 <- openxlsx::read.xlsx("train1set.xlsx")
z2 <- openxlsx::read.xlsx("train2set.xlsx")
zfull <- openxlsx::read.xlsx("fullset.xlsx")
zfull$yearmon <- as.yearmon(as.character(zfull$yearmon), "%Y%m")
x1train <- as.matrix(z1[, -(1:2)])
x2train <- as.matrix(z2[, -(1:2)])
xfull <- as.matrix(zfull[, -(1:2)])
y1train <- z1$y
y2train <- z2$y

library(glmnet)

## Direction 1: 97 -> 08

## Figure 1: Lasso coefficients profile
g1 <- glmnet(x1train, y1train, nfolds = nrow(x1train),
             family = "binomial", alpha = 1, grouped = FALSE)
plot(g1, xvar = "lambda")

## LOOCV
cv1 <- cv.glmnet(x1train, y1train, nfolds = nrow(x1train),
                 family = "binomial", alpha = 1, grouped = FALSE)

## Table 3: Coefficients
b1 <- as.matrix(coef(cv1, s = "lambda.1se"))
b1 <- b1[rownames(b1) != "(Intercept)" & b1 != 0, , drop = FALSE]
b1s <- b1*apply(x1train[, rownames(b1)], 2, sd)
print(cbind(b1,b1s/max(abs(b1s))))

## Prediction
lamb1 <- c(1, exp(2.5), exp(5))*cv1$lambda.1se
zfull$phat1a <- as.numeric(predict(g1,xfull,s=lamb1[1],type="r"))
zfull$phat1b <- as.numeric(predict(g1,xfull,s=lamb1[2],type="r"))
zfull$phat1c <- as.numeric(predict(g1,xfull,s=lamb1[3],type="r"))

## Figure 2
plot(phat1a~yearmon, data=zfull, type="l",ylab="Risk Probability")
lines(phat1b~yearmon, data=zfull, type="l", col=2)
lines(phat1c~yearmon, data=zfull, type="l", col=3)
```

```
## Direction 2: 08 -> 97

## Figure 3: Lasso coefficients profile
g2 <- glmnet(x2train, y2train, nfolds = nrow(x2train),
             family = "binomial", alpha = 1, grouped = FALSE)
plot(g2, xvar = "lambda")

## LOOCV
cv2 <- cv.glmnet(x2train, y2train, nfolds = nrow(x2train),
                 family = "binomial", alpha = 1, grouped = FALSE)

## Table 4: Coefficients
b2 <- as.matrix(coef(cv2, s = "lambda.1se"))
b2 <- b2[rownames(b2) != "(Intercept)" & b2 != 0, , drop = FALSE]
b2s <- b2*apply(x2train[, rownames(b2)], 2, sd)
print(cbind(b2,b2s/max(abs(b2s))))

## Prediction
lamb2 <- c(1, exp(2.5), exp(5))*cv2$lambda.1se
zfull$phat2a <- as.numeric(predict(g2,xfull,s=lamb2[1],type="r"))
zfull$phat2b <- as.numeric(predict(g2,xfull,s=lamb2[2],type="r"))
zfull$phat2c <- as.numeric(predict(g2,xfull,s=lamb2[3],type="r"))

## Figure 4
plot(phat2a~yearmon, data=zfull, type="l",ylab="Risk Probability")
lines(phat2b~yearmon, data=zfull, type="l", col=2)
lines(phat2c~yearmon, data=zfull, type="l", col=3)
```