A new approach to Early Warning Systems for smaller European banks

Despo Malikkidou*
EUROPEAN BANKING AUTHORITY
Regulation
Credit, Market and Operational Risk
Despo.Malikkidou@eba.europa.eu

Michael Bräuning
EUROPEAN CENTRAL BANK
Directorate General Micro-Prudential Supervision III
Analysis and Methodological Support
Michael.Brauning@ecb.europa.eu

Stefano Scalone
EUROPEAN CENTRAL BANK
Directorate General Micro-Prudential Supervision III
Analysis and Methodological Support
Stefano.Scalone@ecb.europa.eu

ABSTRACT

The paper presents an early warning model for predicting distress events tailored for smaller European banks. The underlying approach considers bank distress events based on the Bank Recovery and Resolution Directive. Consequently, the sample of distress events is significantly expanded compared to conventional approaches and allows for a better calibration of the model. The resulting early warning system is modelled using decision-tree techniques and a binary classification tree is grown. The tree comprises bank-specific, banking-sector and macro-financial explanatory variables. The results are quite promising and indicate that the model could provide added value in the supervision of smaller institutions.

Keywords

Early Warning System; Bank Distress Prediction; Machine Learning in Supervision

1. INTRODUCTION

Existing early warning models are usually based on conventional modeling techniques, such as multivariate logit models, and are calibrated using only a very small number of distress events. The proposed approach in defining bank distress enlarges the sample size of distress events and therefore improves the learning of the model. We propose to classify banks as distressed based on the triggering of the Bank Recovery and Resolution Directive's early interventions measures and on its criteria for categorizing banks as failing or likely to fail. More specifically, banks which breach or are close to breaching the minimum capital requirements are considered as distressed. This is a novel approach to the best of our knowledge. Since the definition does not constitute the final stage of a bank's failure, the system will predict the pre-failure stage early enough to allow supervisors to adopt preemptive measures to tackle the situation. Moreover, the paper makes use of decision tree models, often applied in machine learning for classification problems (Mitchell, 1997), to construct a flexible and interpretable signaling tool for supervisors. The proposed early warning system (EWS) is able to predict individual bank distress events and identify which are the key explanatory variables for distress.

The framework is applied to a unique dataset of more than 3,000 small European banks, the so called Less Significant Institutions (LSIs). The EWS for LSIs (LSI-EWS) is built using three sets of variables:

- bank-specific,
- banking-sector variables and
- country-level macro-financial indicators.

The remainder of this paper is organized as follows. In the next section, we give a brief overview of related work. In Section 3, we present our approach to EWS for smaller banks. The corresponding results for a sample of banks are presented in Section 4 prior to concluding the paper in Section 5.

2. RELATED WORK

The paper encompasses two strands of literature which are the base for the system: that on early warning systems (modeling) for bank distress, and the latter on machine learning techniques, in particular classification algorithms.

The research for predicting bank distress events has received a lot of attention by academics and policymakers. One of the first studies on the topic dates back to 1975, when Sinkey employed Altman's (1968) multiple discriminant analysis technique to predict bank distress. He found that asset composition, loan characteristics, capital adequacy, sources of revenues, efficiency and profitability are good discriminators between distressed and non-distressed banks.

Subsequent studies have focused on the six components of the so called CAMELS ratings system to identify problematic banks. These components are

- capital adequacy,
- asset quality,
- management,
- earnings,
- liquidity and
- sensitivity to market risk.

Thomson (1992) and Cole and Gunther (1998) estimate logit and probit models to show that vulnerability indicators covering the CAMELS dimension are good predictors of bank failure. Other studies, including Flannery (1998), Gonzalez-Hermosillo (1999) and Betz et al. (2013), complemented the CAMELS-motivated variables with both macroeconomic and market price-based indicators and found increased predictive power.

* This work was prepared while the author was working at the European Central Bank (ECB) in the Directorate General Micro-Prudential Supervision III, Analysis and Methodological Support.

NOTE: This paper should not be reported as representing the views of the ECB. The views expressed are those of the authors and do not necessarily reflect those of the ECB.

Interest in predicting bank failures and banking crises has re-emerged after the global financial crisis of 2007. Jin et al. (2011, 2013) found that complementing traditional CAMELS indicators with data on banks' internal controls on risk-taking and audit quality variables improves the predictive rate. Moreover, Cole and White (2012) found that measures of commercial real estate investments are relevant for predicting bank distress.

(Among) The most established methods for constructing an EWS are the simple binary choice models (Logit and Probit). Alternative methods also used with high degree of predictability include multiple discriminant analysis (Sinkey, 1975; Jordan et al., 2010), multiple regression analysis (Meyer and Pifer, 1970) and neural networks (Loopez-Iturriaga et al., 2010, Ng et al., 2008).

From a methodological perspective, the paper is linked to decision tree learning in the general field of machine learning, a methodology which has already been successfully applied to various tasks such as assessing credit risk (Mitchell, 1997). Kohavi and Provost (1998) describe machine learning as the exploration of the study and construction of such algorithms that can learn from and make predictions on data. Two popular algorithms for the induction of decision trees in the context of classification are Classification And Regression Trees (CART) and C4.5 (which has recently been superseded by C5.0, Kohavi and Quinlan, 2002).

Established methods which base the analysis on traditional statistical techniques require pre-specifications of the relationship between explanatory variables and the output, i.e. the distress event (Lee, 2008). On the other hand, there are several machine learning concepts which are not restricted by this requirement. Artificial neural networks (ANN), for example, do not need any specification of this relationship and several research studies regarding distress prediction for banks have already been performed (among others Altman et al. (1994), Boritz and Kennedy (1995) or Atiya (2001)). Altman et al. (1994) concluded that neural networks are interesting tools for distress diagnosis. However, the interpretation of the model's operating logic is not as convenient as in traditional statistical tools. Sometimes ANN are even considered as "black boxes" and, therefore, not capable of clearly presenting the relationship between explanatory and dependent variables.

In contrast to ANN, decision trees offer a more transparent representation of the model. Frydman et al. started applying decision trees to predict business failure in 1985, and found that their Recursive Partitioning Algorithm (RPA) outperforms discriminant analysis. Joos et al. (1998) applied decision trees in credit classification and compared them with logit analysis based on three different data sets comprising a full set of financial variables, a reduced one, and one containing only qualitative variables. It turned out that the logit analysis is a superior predictor for the full data set. However, the decision tree outperformed the logit analysis on the reduced set and for qualitative variables (Gepp et al., 2010). Additional studies mostly support decision trees as being superior to discriminant and logit analyses for predicting business failures (Gepp and Kumar, 2015).

3. APPROACH

This section describes first the data and variables' definitions used in the analysis. Second, the methodological framework for the LSI-EWS is presented, which consists of three blocks: data pre-processing, model estimation and model evaluation. In general, we follow the CRISP-DM[1] methodology when developing the LSI-EWS, to ensure a structured and robust approach.

3.1 Data

The study uses a unique dataset which covers more than 3,000 small banks in the euro area and spans over the period 2014Q4 – 2016Q1. While we focus on small banks, the recent global financial crisis has made it clear how small institutions are still relevant for financial stability due to the high interconnectedness observed in the euro area banking system[2].

Actual bank failures in the euro area are quite rare, making the calibration of an EWS challenging. This paper introduces an approach for identifying bank distress based on the Bank Recovery and Resolution Directive (BRRD), which is a novel approach to the best of our knowledge. As a result, there are approximately 350 distress observations throughout a sample of only six quarters. These events were extracted according to the following criteria defining the dependent variable. We consider a bank to be in distress when

- it is deemed to be failing or is likely to fail within the meaning of Article 32 BRRD. For categorizing a bank as failing or likely to fail, indicators assessing whether a bank has breached the minimum capital requirements or capital buffers are constructed;
- it meets the conditions for early intervention pursuant to Article 27 of BRRD. The triggers used to meet the conditions of early interventions consist of indicators for assessing if a bank is close to breaching minimum capital requirements;
- it is placed under special administration and/or is appointed of a temporary administrator pursuant to Article 29 of the BRRD;
- there is a rapid and significant deterioration of its financial situation according to Article 96 of Framework Regulation. This is based on expert judgement by national central banks and in-house qualitative data;
- one of the four types of conventional bank distress events proposed by Betz et al (2013) (i.e. bankruptcies, liquidations, state interventions and forced mergers) is observed.

The explanatory variables, on the other hand, aim at capturing a bank's vulnerability to potential distress and the environment in which it operates. The variables are categorized in three groups: bank-specific indicators[3], banking-sector indicators and country-level macro-financial indicators.

[1] CRISP-DM stands for cross-industry process for data mining. The CRISP-DM methodology provides a structured approach to planning a data mining project.

[2] While for individual LSIs the impact on the domestic financial system might be limited, a substantial large number of them are organized in sectors. Together, they can have a non-negligible impact on the domestic financial system.

[3] The terms explanatory variable and indicator are used as synonyms.

First, the bank-specific indicators are constructed based on balance sheet and income statement data. These explanatory variables capture a bank's capital adequacy and account for the dimensions of business model and profitability, credit risk, market risk, interest rate risk in the banking book, operational risk and risks to liquidity and funding.

Second, banking-sector indicators aim to capture the banking sector structure, lending and leverage, profitability, liquidity and asset quality at the country level.

Finally, the country-specific macro-financial indicators capture the macroeconomic and financial environment of each euro area country.

The resulting total number of explanatory variables is quite high, comprising more than 3,000 variables at a first step. The majority of the data is available at quarterly frequency apart from balance-sheet, income statement, and country-specific banking sector data, which are only available bi-annually. Thus, we use the same information in two consecutive quarters. To account for publication lags, we lag all the explanatory variables by one quarter. Other data quality issues related to cross-data comparability due to different accounting standards in different countries are partially addressed by adjusting the affected variables accordingly.

Prior to modelling, the data is split into two separate sets to learn (train) and validate (test) the model independently, rather than just working with all of the data. 75% of the data is assigned to the training and 25% to the test partition. The out-of-sample validation ensures more reliable results in case of predictive modelling.

3.2 Methodology

Data pre-processing steps are required to ensure that unreliable and noisy data as well as irrelevant and redundant information is eliminated prior to the modelling phase. As such, the final training dataset used for the analysis is of high quality, thus increasing the efficiency and performance of the final model. We here summarize the main steps taken in this direction:

- We start with the data cleaning, aiming at removing incomplete or non-informative data. The step includes the removal of banks where the majority of relevant data is missing; the removal of explanatory variables for which the majority of values is missing or where the variance is nearly zero. This step already significantly reduces the number of indicators to less than 500.
- The following step is to transform the data in a way to ensure consistency and comparability across banks. Cross-data comparability issues arise mainly because of different accounting standards across euro area countries, and we overcome the issue by adjusting the data according to a conversion technique. Moreover, variables are normalized through the creation of ratios, and aggregated to produce consolidated vulnerability indicators. We refer to ratios usually applied in related literature such as return on assets/equity or non-performing loans ratio to increase comparability of the banks in scope.
- We remove explanatory variables which are highly correlated with each other, using a correlation threshold of 0.9.

We select the final set of indicators based on their ability to predict distress: Variables are ranked according to their importance, which is captured by the individual Area under the Receiver Operating Characteristic curve (AUC) for each indicator; using this technique, we select the top 100 variables in terms of predictive performance. The LSI-EWS is based on a decision tree model leveraging from its simplicity, good performance and transparency. It allows supervisors to interpret the output tree and understand which indicators affect bank distress, thus minimizing the risk of creating a "black box" model. Moreover, it can easily accommodate unbalanced panels and missing data (Mitchell, 1997), which is a common issue in the early warning literature.

A decision tree is a classification technique commonly used in machine learning. The tree recursively identifies the indicators and their respective thresholds which best split the sample in the pre-determined classes (i.e. distress and no distress). The final output is a tree, illustrating a set of if-then rules (decision nodes) to reach a final decision on the classification (leaf nodes). In practice, for each bank, the classification starts from the root decision node, and based on predictors values you move along the tree until you reach a leaf node, classifying if the bank is in distress or not.

We use Quinlan's C5.0 algorithm[4] to build the classification tree model. The C5.0 algorithm is fast, has high accuracy and efficiency and produces lower error rates compared to alternative algorithms. It became quite popular after ranking first in the Top 10 Algorithms in Data Mining in 2008 (Wu et al., 2008). The algorithm is also able to handle noise and missing data as well as remove unhelpful attributes. The training of the model requires the selection of a number of specification options.

- The choice of the prediction horizon is an important element of the model and depends on the objective of the supervisor. Given that the scope of this EWS is short-term (<1 year) and taking into account the data limitations in the time dimension, we consider predicting distress within the next quarter as a starting point. By considering pre-default events as target variable, we ensure that the system has a forward looking manner.
- We implement asymmetric misclassification costs when assessing the performance of explanatory variables. In other words, we consider Type I errors (miss a distress events) twice as costly as Type II errors (issue a false alarm). In principle, this assumes that when a policymaker is faced with a tradeoff of issuing more false alarms or missing a distress event, he or she takes a conservative stance and prefers the former.

Alessi and Detken (2014) argue that decision trees are not particularly robust when additional predictors or observations are included in the sample. To overcome this problem, they use the Random Forest method in order to pre-select the key indicators on which the final decision tree will be constructed. In this way, they ensure that the variables selected are the truly most important ones, for both in- and out- of sample exercises.

We follow a similar logic and employ boosting techniques (Freund et al., 1999) in identifying the most important variables to

[4] The R language and environment was used to learn the decision tree applying the package 'C50' by Kuhn et al. (2015).

build the decision tree. Boosting is a technique for generating and combining multiple classifiers to improve the predictive accuracy of the model. Instead of using a single tree, n separate decision trees (trials) are grown and combined to make predictions. The error rate of the boosted classifier is often substantially lower than the single tree. However, presentation and communication of the boosted classifier is hard since there is not one preferred tree among the trials. Therefore, we rely on the boosting in order to rank variables by importance and select the top 20 variables to go into the final model. We complement this list with variables that are considered relevant for identifying bank distress based on expert judgement.

The final decision tree is assessed based on its performance and complexity. For this purpose, we make use of both qualitative and quantitative evaluation measures. In particular, for evaluating the performance of the model, we rely on the area under the Receiver-Operating-Characteristic curve (AUC) and Cohen's kappa statistic, both standard measures of accuracy in the early warning system literature (e.g. Peltonen at al. 2015). These performance measures have the advantage of being robust to imbalance classes, where the number of distress events is much lower than the number of non-distress ones.

4. RESULTS

The resulting tree[5] consists of 19 nodes, covers 12 distinct explanatory variables and is presented in Figure 1. The indicator in the parent node is profitability, adjusted for different accounting standards across Single Supervisory Mechanism jurisdictions. The indicator splits the sample to profitable (right branch) and loss making (left branch) banks. The subsequent paths within the tree contain the following variables

- non-performing loans ratio
- non-performing loans coverage ratio
- deficit-to-GDP ratio
- GDP growth
- liquidity coverage ratio (LCR)
- leverage ratio
- equity exposures
- exposures in default
- market risk captured by
    - sum of trading financial assets and financial liabilities held for trading over total assets
    - net gains on financial assets and liabilities held for trading over total operating income
- membership in an institutional protection scheme (IPS)

The tree nicely illustrates interactions between different variables. For profitable banks, one can notice that if the non-performing loans ratio is high, the coverage ratio is considered. Consequently, the coverage ratio can compensate if sufficient allowances on non-performing loans are created also taking into account the economic environment of the country (as captured by the deficit-to-GDP ratio). On the other hand, for banks with low non-performing loans (lying in the 1st quantile of the distribution), market risk becomes relevant in capturing distress. In particular, banks relying on income from market activities are more sensitive to distress, in particular in countries with fragile economic fundamentals.

For loss-making banks, the variables explaining distress depend on whether the institution lies in the extreme of the distribution or not. For extremely heavy losing banks, the tree suggests to look into the GDP growth of the country where the bank is located. If the growth is low, then the model suggests that the bank is in distress, while vice versa, if the growth is high, then the model suggests looking into the bank's equity position. When the bank has relatively high equity exposures in combination with a weak LCR ratio, then the membership in an IPS determines whether a bank will experience distress. IPS members are in fact less vulnerable to distress, as the scheme protects them from financial difficulties. For moderately losing banks, high deficit and low leverage can explain whether a bank is in distress or not.

The in-sample predictive performance of the LSI-EWS is very high (see Table 1):

- The true positive rate is 0.89, while the false positive rate (Type II error) and false negative rate (Type I error) equal 0.01 and 0.11, respectively.
- The AUC is 0.95 and much higher than the 0.5 value of a random classifier and very close to the 1 value of the perfect classifier. The Cohen's kappa statistic is also high, at a value of 0.89.

The performance of the model is also assessed on the testing set, covering 25% of the initial sample. The out-of-sample performance of the model remains very high with an AUC of 0.92 and a Cohen's kappa statistics of 0.80. Type I and II errors are comparable to the in-sample error rates, remaining at adequately low levels.

Table 1. Validation results

Measures	In-sample (train)	Out-of-sample (test)
Type I error rate	0.01	0.03
Type II error rate	0.11	0.10
AUC	0.95	0.92
Cohen's Kappa	0.89	0.80

Logistic regression has been one of the most popular approaches in the literature to predict distress. Thus, it is worth comparing the performance of the resulting tree against a logit model. We estimate a benchmark logit model, also with a forecast horizon of 1 quarter. The logit model performs also well and showed an AUC of 0.90, Type I and II error rate of 0.01 and 0.02, respectively,[6] but reacted much more sensitively to missing values compared to the decision tree.

[5] Due to data confidentiality reasons the thresholds for the variable splits cannot be disclosed.

[6] An out of-sample-validation showed similar results.

5. CONCLUSION

This paper presents an early warning model for predicting future bank distress cases in the European banking system. We introduce a novel approach in defining distress based on the basis of the Bank Recovery and Resolution Directive, which overcomes the problem faced in the literature of limited actual distress events. As such, the calibration of the EWS is significantly improved and policymakers are able to predict pre-default events early enough to take preemptive measures. The early warning system is built using decision tree modelling, a method which has the benefits of transparency and good performance, this confirmed by the promising results presented in Section 4.

Future work will cover the enrichment of the data and the extension of the prediction horizon to up to six months. Once further data is available an extensive back-testing shall assess the robustness of the LSI-EWS and its prediction results over time. Moreover we plan to develop a new Early Warning System which could also identify the severity of the bank distress. This can be achieved by using a multi-class target model, which will classify distress across three categories: Mild, Moderate and Severe Distress. Such classification is important for supervisors in order to efficiently allocate resources and improve financial surveillance.

Acknowledgements We thank *Caroline Liesegang* for her significant contribution to the development of this work. We also thank the reviewers for the helpful comments.

REFERENCES

Alessi L., and Detken C. (2014). *Identifying excessive credit growth and leverage*. ECB Working Paper No. 1723.

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

Altman, E. I., Marco, G., and Varetto F. (1994). *Corporate Distress Diagnosis: Comparisons Using Linear Discriminate Analysis and Neural Networks*. Journal of Banking and Finance, 18, 505–529.

Atiya, A. (2001). *Bankruptcy prediction for credit risk using neural networks: A survey and new results*. Transactions on Neural Networks 12, 4 (July 2001), 929-935.

Betz, F., Opricǎ, S., Peltonen, T. A., and Sarlin, P. (2014). *Predicting distress in European banks*. Journal of Banking & Finance, 45, 225-241.

Boritz, J. E., and D. B. Kennedy (1995). *Effectiveness of Neural Network Types for Prediction of Business Failure*. Expert Systems with Applications, 9(4), 95–112.

Cole, R. A., and J. Gunther (1998). *Predicting bank failures: A comparison of on- and off- site monitoring systems*. Journal of Financial Services Research 13(2), 103-117.

Cole, R. A. and L. J. White (2012). *Deja vu all over again: The causes of U.S. commercial bank failures this time around*. Journal of Financial Services Research 42(1-2), 5-29.

Flannery, M. J. (1998). *Using Market Information in Prudential Bank Supervision: A Review of the U.S. Empirical Evidence*. Journal of Money, Credit and Banking 30, no. 3: 273-305.

Freund, Y., Schapire, R., and Abe, N. (1999). *A short introduction to boosting*. Journal-Japanese Society For Artificial Intelligence, 14(771-780), 1612.

Frydman, H., Altman, EI., and Kao, DL. (1985). *Introduction Recursive Partitioning for Financial Classification: The Case of Financial Distress*. Journal of Finance, vol. 40(1), pp. 269–291.

Gepp, A., Kumar, K., and Bhattacharya, S. (2010). *Business failure prediction using decision trees*. Journal of forecasting, 29(6), 536-555.

Gepp, G., and Kumar, K. (2015). *Predicting financial distress: A comparison of survival analysis and decision tree techniques*. 11th International Conference on Data Mining and Warehousing.

González-Hermosillo, M. B. (1999). *Determinants of ex-ante banking system distress: A macro-micro empirical exploration of some recent episodes*. International Monetary Fund.

Jin, J., K. Kanagaretnam, and G. Lobo (2011). *Ability of accounting and audit quality variables to predict bank failure during the financial crisis*. Journal of Banking & Finance 35(11), 2811-2819.

Jin, J., K. Kanagaretnam, G. Lobo, and R. Mathieu (2013). *Impact of FDICIA internal controls on bank risk taking*. Journal of Banking & Finance 37(2), 614-624.

Joos, P., Vanhoof, K., Ooghe, H., and Sierens, N. (1998). *Credit classification: A comparison of LOGIT models and decision trees*. Proceedings notes of the Workshop on Application of Machine Learning and Data Mining in Finance, 59-72.

Kohavi, R., and Quinlan, R. (2002). *Decision Tree Discovery, in Handbook of Data Mining and Knowledge Discovery*. Klosgen & Zytkow Editors, Chapter 16.1.3, pages 267-276, Oxford University Press.

Kuhn M., Weston S., Coulter N., and Culp M. (2015). *C5.0 Decision Trees and Rule-Based Models*.

Lee, W. (2008). *An empirical comparison of bankruptcy models: Evidence from Taiwan*. Retrieved December, 6, 2009.

Meyer, P. A., and Pifer, H. W. (1970). *Prediction of bank failures. The Journal of Finance*, 25(4), 853-868.

Mitchell, T. (1997). *Machine Learning* (1 ed.). McGraw-Hill, Inc., New York, NY, USA.

Ng, G. S., Quek, C., and Jiang, H. (2008). *FCMAC-EWS: A bank failure early warning system based on a novel localized pattern learning and semantically associative fuzzy neural network*. Expert Systems with Applications, 34(2), 989-1003.

Peltonen, T. A., Piloiu, A., and Sarlin, P. (2015). *Network linkages to predict bank distress*. Available at SSRN 2579584 (2015).

Provost, F., and Kohavi, R. (1998). Glossary of *terms*. Machine learning, 30(2), 271–274

Sinkey, J. F. (1975). *A multivariate statistical analysis of the characteristics of problem banks*. The Journal of Finance, 30(1), 21-36.

Thomson, J. B. (1991). *Predicting bank failures in the 1980s*. Economic Review-Federal Reserve Bank of Cleveland, 27(1), 9.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., … Zhou, Z. H. (2008). *Top 10 algorithms in data mining*. Knowledge and information systems, 14(1), 1-37.

Figure 1. LSI-EWS decision tree