Comput Econ (2018) 52:275–297
https://doi.org/10.1007/s10614-017-9676-6

Financial Soundness Prediction Using a Multi-classification Model: Evidence from Current Financial Crisis in OECD Banks

D. Fernández-Arias[1] · M. López-Martín[1] · T. Montero-Romero[3] · F. Martínez-Estudillo[2] · F. Fernández-Navarro[2]

Accepted: 9 March 2017 / Published online: 17 March 2017

Abstract The paper aims to develop an early warning model that separates previously rated banks (337 Fitch-rated banks from OECD) into three classes, based on their financial health and using a one-year window. The early warning system is based on a classification model which estimates the Fitch ratings using Bankscope bank-specific data, regulatory and macroeconomic data as input variables. The authors propose a "hybridization technique" that combines the Extreme learning machine and the Synthetic Minority Over-sampling Technique. Due to the imbalanced nature of the problem, the authors apply an oversampling technique on the data aiming to improve the classification results on the minority groups. The methodology proposed outperforms other existing classification techniques used to predict bank solvency. It proved essential in improving average accuracy and especially the performance of the minority groups.

✉ D. Fernández-Arias
davidfas@outlook.es

M. López-Martín
mclopez@uloyola.es

T. Montero-Romero
tmontero@uloyola.es

F. Martínez-Estudillo
fjmestud@uloyola.es

F. Fernández-Navarro
i22fenaf@uco.es; fafernandez@uloyola.es

1 Department of Economics, Universidad Loyola Andalucia, Córdoba, Spain

2 Deparment of Quantitative Methods, Universidad Loyola Andalucia, Córdoba, Spain

3 Department of Financial Economics and Accounting, Universidad Loyola Andalucia, Córdoba, Spain

Keywords Bankruptcy prediction · Artificial neural networks · Extreme learning machine · Early warning models · Computational intelligence · Financial crisis

1 Introduction

The banking sector is considered to be one of the most supervised sectors in the world due to the impact that the bankruptcy of banking institutions can have on the economy (Chortareas et al. 2012). Over the last three decades, banking crises have become more frequent and severe both in emerging markets and in industrial countries. These crises have resulted in a strong contraction in production and substantial fiscal and financial losses for the economies concerned. During recent years we have realized that the collapse of even massive non-financial firms is less likely to endanger an entire economy than bank insolvencies; and that the latter can either result in systemic crises with adverse consequences for the economy as a whole, or as experienced recently (US and Europe), they can have a rapid impact worldwide (see, e.g., Rose and Spiegel 2012).

Thus, there exists a need for predictive techniques to provide an early warning system to regulatory agencies and other stakeholders regarding distressed or failing institutions. Once distressed banks are identified, regulator intervention may be able to prevent ultimate failure, or at least minimize the impact of such failure. Developing an early warning model that divides already rated banks into three groups, based on their financial health and using a one-year window is the intention of this study. While our model is based on the Fitch ratings, in assessing the soundness[1] of the banks, we do not attempt to replicate all the ratings of Fitch.[2] We therefore follow the approach of Gaganis et al. (2006) and classify banks into three general classes. The first class includes very strong and strong institutions; the second class contains adequate institutions, while banks with weaknesses or serious problems are placed into the third class. By focusing only on banks that have not failed, and then distinguishing between these three groups, the model can be useful in reducing the costs of bank failure, either by minimizing these costs to the public or by taking action to prevent failure.

The prediction of bankruptcy for banking firms is an area that has been extensively researched using classical statistical techniques, such as discriminant, logit or probit (Martin 1977). However, all these methods have the disadvantage that they make

[1] The concept of soundness is commonly used to denote, for example, an ability to withstand adverse events. Solvency is reflected in the positive net worth of a bank, as measured by the difference between the assets and liabilities (excluding capital and reserves) in its balance sheet. The likelihood of remaining solvent will depend, inter alia, on banks' being profitable, well managed, and sufficiently well capitalized to withstand adverse events. In a dynamic and competitive market economy, efficiency and profitability are linked, and their interaction will indicate the prospects for future solvency. Inefficient banks will make losses and will eventually become insolvent and illiquid. Undercapitalized banks, that is, those with low net worth, will be fragile in the sense of being more prone to collapse when faced with a destabilizing shock, such as a major policy change, a sharp asset price adjustment, financial sector liberalization, or a natural disaster (Lindgren et al. 1996).

[2] Fitch evaluates the current banks' solvency however our intention is to develop an early waning model in which we predict the solvency in a future time.

assumptions about the model or the data distribution (independence of multiple features, etc.) that are not usually satisfied. For example, most parametric models require that the input variables be linearly separable. That is, no two input variables should have similar influence on the dependent variables. When financial ratios and aggregate account balances are used as inputs, this requirement can easily be violated.

So, in order to avoid the drawbacks of the classic statistical models, it has recently been suggested in the field of economics that soft-computing techniques should be used (Chen and Cheng 2012), such as neural networks or evolutionary algorithms (Fernández-Navarro et al. 2011). Artificial neural networks (ANNs) have been used as an alternative to these traditional techniques, and have gained popularity over the past two decades. One of the most important advantages of an ANN is that it can automatically approximate any non-linear mathematical function. ANNs are particularly useful when statistical assumptions are not valid, or when the relationship between variables is not known, or is complex and difficult to handle statistically. Moreover, ANNs are more robust with respect to outliers than classic parametric models. Conversely, the main disadvantages are to find the optimal framework for the ANN model, and the corresponding training required to estimate the neural network weights (Huang et al. 2012).

This paper uses a recent methodology named extreme learning machine (ELM) as a learning algorithm for single-hidden-feed-forward neural networks (Huang et al. 2006a) . ELM appears to overcome the drawbacks mentioned above concerning the ANN model, offering the advantages of a faster learning speed, better generalization performance and ease of implementation. The imbalanced nature of the problem lead us to applying an oversampling technique on the data aiming to improve the classification results on the minority classes (the extreme ones, classes 1 and 3). Specifically, synthetic patterns were generated on those classes as suggested in Chawla et al. (2002). The suitability and robustness of the proposed methodology is examined and tested through a set of variables (bank specific, regulatory and macroeconomic). It has been necessary to resolve a complex problem as financial soundness prediction is.

The main empirical finding of this paper is to introduce a new hybrid methodology to the banking field, even though it has been extensively used in other research fields. The model proposed is able to resolving the limitations of the classical statistical techniques. The use of the oversampling technique implemented on the minority classes (specially interesting in the problem) significantly improved the classification performance on those classes. In addition, this methodology outperforms other existing classification techniques used for bank solvency prediction. At the same time, our proposal could be used as an early warning system which distinguishes three classes of banks based on financial health unlike traditional methods which only consider two classes. This approach has rarely been used in the field of banking prediction. It provides more information and may therefore reduce the cost of banks' bailouts and thereby minimize the fiscal cost (to the taxpayer) or place banks in a better position to take action to prevent bankruptcies (Thomson 1991).

The rest of the paper is structured as follows: the next section provides a brief literature review. Section 3 describes the data. Section 3 describes the methodology applied. Section 5 presents the experimental framework. The results are showed in Sect. 6, and finally Sect. 7 concludes the research.

2 Literature Review

The challenge of forecasting insolvency and assessing bank soundness is an important and widely studied topic (for a complete review of this literature and methodology employed see, e.g., Demyanyk and Hasan 2010; Gramlich et al. 2010; Kumar and Ravi 2007). Current empirical research can be divided into two strands. The first strand involves studies that examine banks and try to classify them into two categories, distinguishing between solvent and insolvent banks, and taking into account the influence of explanatory variables on banks' solvency. One limitation of the studies in this strand, in so far as the extensive literature reveals it, is that they have concentrated the classification of banks into only two groups, those that have failed and those that have not. Obviously, the simple classification of banks as either "good" or "bad" reduces the usefulness of a model and the information it provides. Also, they have focused on the determinants of banks failures rather than on a correct out-of-sample classification of the banks. Therefore, during the last few years, several studies have tried to develop models capable of classifying banks into more than two categories, thus giving rise to the second strand of literature. Following this last approach, and in order to resolve the imbalanced problem, we propose a new hybrid method.

There are several interesting reviews related to the first strand. For example, the U.S. financial crisis of the late eighties and early nineties served as the sample for the work of Thomson (1991). He analyzed 16 independent variables based on the CAMEL ratings. The model used was logistic regression, and obtained consistent 93% prediction results for failed banks between 6 and 12 months in advance.

Olmeda and Fernández (1997) compared the accuracy of parametric and nonparametric classifiers on the problem of predicting bankruptcy. They used a database from 1977 to 1985 when the Spanish banking system suffered a dramatic crisis, affecting 52 percent of the 110 banks that were operative at the beginning of this period. An Artificial Neural Network is found to be superior to both classical and recently developed statistical classifiers. Also they found that when one combines two or more methods by simple voting, the predictions are generally more accurate than the one obtained by applying any single method.

Among the most important analyses of banking crises is that conducted by Demirgüç-Kunt and Detragiache (1998). Their work analyzed the factors associated with the emergence of banking crises throughout 65 developed and developing countries during the period 1980–1994. Regarding the dependent variables, the authors set eight indicators to identify the fragility of the banking system, all of which were macroeconomic in nature. They used a multivariate logit model, concluding that banking crises tend to emerge when there are weak macroeconomic conditions, low GDP growth, high real interest rates, high inflation, strong growth of bank credit in the past, and a large cash flow ratio of international reserves.

González-Hermosillo (1999) analyzed the contribution of microeconomic and macroeconomic factors in five episodes of banking crises in the Southeast of U.S. (1986–1992), the U.S. Northeast (1991–1992), California (1992–1993), Mexico (1994–1995) and Colombia (1982–1987). This author found that low levels of capital ratios and coverage of nonperforming loans were the main indicators of bank stress, indicating high probability of bank failure in the short term. Porath (2006) used a new database, which contained information on the financial stress of all savings banks and credit cooperatives in Germany during the period 1995–2002. He covered the effect of capitalization, profitability, credit risk, market risk and the macroeconomic environment (regional macroeconomic data) to study their influence on the probability of failure. Testing the importance of the microeconomic and macroeconomic variables, he decided to use a discrete risk model, and concluded that the information provided by regional macroeconomic variables was a significant element in forecasting banks' defaults. Classification trees were used in banking crisis analysis by Duttagupta and Cashin (2008) for a large sample of 50 emerging and developing countries during the years 1990–2005, covering the geographical area of Asia, Africa , Europe, South and Central America and the Caribbean. In explaining the crises a long list of independent variables was considered, covering four blocks: macroeconomic fundamentals, external liquidity, monetary conditions and financial soundness indicators. According to their results, high inflation, low profitability and liquidity tensions were all related to the precipitation of banking crises.

Čihák and Poghosyan (2009) composed a database of 5,708 banks for the EU-25 during the period 1997–2008. The empirical technique used to calculate the probability of failure was multivariate logistics. Among their findings was that banks that paid higher remuneration for their deposits were more likely to collapse. They also showed that a higher concentration of banks in a market was associated with a greater probability of failure in the financial system. Gutiérrez et al. (2010) tried to explain the existence of systemic financial crises, linking these events to banking crises. Their explanatory variables attempted to cover the main banking risks: credit risk, interest rate risk, currency risk and liquidity risk. The authors used a novel model for the detection and prediction of crises, based on the hybridization of a logit regression model with product unit (PU), with a neural network and radial basis function (RBF) neural network. The new model they developed for the analysis of 79 countries obtained robust results, classifying the dependent variables in crisis or no-crisis situations and increasing the efficiency relative to other models they compared theirs with.

To complete the first strand, DeYoung and Torna (2013) contributed by studying the failures of hundreds of U.S. commercial banks during the financial crisis lasting from 2008 to 2010, using non-traditional banking activities as explanatory variables estimated with a multi-period logit model. They found that non-interest income from stakeholder activities, such as investment banking, insurance underwriting and venture capital increased the probability of bank failure. They also discovered that Fee-for-Service income from non-traditional activities, such as insurance sales, loan servicing and securities brokerage actually reduced the probability of bank failure during the crisis. Regarding the second strand, we highlight the research by Gaganis et al. (2006), who used a sample of 894 banks from 79 globally distributed countries to develop a multiclass decision model where banks were classified into groups based on their creditworthiness. The method used was Discriminant Utilities Additives (UTADIS). Their results indicated that asset quality, capitalization and the markets in which banks operate were the most important criteria in their classification. The method used (UTADIS) correctly classified 68.91% of the banks in the validation sample, and was more accurate and efficient than the other models that were compared with it (discriminant analysis and ordinal logistic regression).

Given the current context of crisis, which increases the importance of early warning models to assess the solvency of banks, Ioannidis et al. (2010) presented a study to classify banks into three groups. They used a database of 944 banks from 78 countries, using the ratings in late 2008 and the bank specific variables in 2007. Regarding variables, the authors relied on the CAMEL model (for financial variables), using in addition other qualitative variables and the regulation characteristics of the countries where the banks were operating. The results showed a satisfactory level of precision for the UTADIS model in that it correctly classified 78.45% of the banks in the holdout sample. Working within the same context of crisis, Moore (2011) analyzed the determinants required to classify banks according to their level of financial health, using different thresholds of net income as a proxy for soundness. The model developed was a decision tree, providing 78.8% correct classification, and increasing the robustness of the model to 84.7% when possible classification groups were reduced from four to three. Finally, Table 1 presents for all papers, the details such as the number of groups in which the classification of banks is divided, techniques used in the study and the typology of inputs used (microeconomic/macroeconomic variables).

3 Variables Involved in Financial Soundness Prediction

Table 2 presents descriptive statistics of the variables used in the model as criteria of banks' soundness. Credit agencies, auditors and bank regulators tend to evaluate banks' performance on the basis of the CAMEL model. We followed the same approach and selected financial variables that proxy for the five CAMEL dimensions, but included financial efficiency and size. All these bank-specific data were obtained from the Bankscope database. Being aware of the importance of incorporating regulatory variables, we constructed indices, obtained from the World Bank (WB) database, on bank regulations and supervision (Barth et al. 2001a, b, 2008). Also, we add country-level variables, which are obtained from the World Development Indicators, the Heritage Foundation and the 2008 update of the WB database on financial development and structure (Beck et al. 2000).

3.1 Bank-Specific Ratios

Capital adequacy measures have been found to be significant predictor of bank failures in a number of studies (Bongini et al. 2001; Estrella et al. 2000). The Basel Committee on Banking Supervision has emphasized the significance of capital adequacy, which has become the main subject of regulation in all Basel Accords in the form of capital requirement ratios. Due to this importance as emphasized in earlier literature, we use four different ways to evaluate the equity funding of a balance sheet as an indicator of capital strength: (1) the equity to total assets (C1) ratio that measures the amount of protection offered to the bank by its equity; (2) the equity to net loans (C2) ratio, which measures the equity available to absorb losses on the bank's loan portfolio; (3) the ratio of equity to customer and short term funding (C3), measuring the amount of permanent funding (i.e. equity) with respect to customer deposits and short term funding; and (4) equity to liabilities ratio (C4), which provides another way of looking

Table 1 Details of each paper

Studies	Level	Methodology	Independent variables
Thomson (1991)	B	Logistic regression	Variables based on the CAMEL ratings (microeconomic).
Olmeda and Fernández (1997)	B	ANNs	Microeconomic variables
Demirgüç-Kunt and Detragiache (1998)	B	Multivariate logit model	Macroeconomic variables
González-Hermosillo (1999)	B	Fixed-effects logit model	Microeconomic and macroeconomic factors
Porath (2006)	B	Discrete risk model	Microeconomic and macroeconomic (regional) variables
Čihák and Poghosyan (2009)	B	Multivariate logistic	Microeconomic and macroeconomic factors
Gutiérrez et al. (2010)	B	ANNs	Microeconomic variables
DeYoung and Torna (2013)	B	Multi-period logit model	Microeconomic variables
Gaganis et al. (2006)	M	UTADIS	Microeconomic and macroeconomic factors
Ioannidis et al. (2010)	M	UTADIS	Microeconomic and macroeconomic data
Moore (2011)	M	Decision trees	Microeconomic and macroeconomic input variables

B binary, *M* multiclass or multigroup

at the equity funding of the balance sheet. Table 2 shows how the mean values of these ratios decrease from the strongest group to the weakest one. This shows the importance of these ratios in evaluating banking solvency. The higher these ratios are, the more protection there is.

When a banking crisis emerges, factors such as profitability and asset quality show a decline in their values. Related to this, both operating costs and interest margins reflect notorious deterioration at the time of a crisis. Therefore, as a result of these problems with their margins, financial institutions are motivated to increase their operating efficiency (Demirgüç-Kunt et al. 2006). Through these statements, we justify the incorporation of asset quality, earnings and financial efficiency ratios into our model.

Thus, in order to measure asset quality, we use a loan loss provision to net interest revenue ratio (A1), which is the relationship between provisions in the profit and loss

Table 2 Descriptive statistics

	Group 1			Group 2			Group 3		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Bank ratios									
C1	13.00	6.60	17.53	11.96	8.17	14.29	8.67	5.36	11.24
C2	31.74	12.06	97.20	28.48	14.50	60.21	18.42	11.06	23.46
C3	43.01	10.64	125.21	25.77	11.25	73.44	20.14	7.92	59.93
C4	29.03	7.15	69.99	20.22	9.32	50.43	16.03	5.81	48.38
A1	20.35	18.20	21.29	23.32	22.44	41.78	37.82	30.51	53.14
M1	10.07	10.07	22.79	13.36	8.20	48.80	32.41	11.67	85.19
E1	2.32	0.92	6.03	0.90	0.49	4.99	0.14	0.16	3.08
E2	3.47	1.74	4.45	3.18	2.46	3.41	2.63	2.14	2.21
E3	2.86	1.16	5.29	1.56	0.84	3.55	0.87	0.74	1.54
E4	18.28	12.29	48.73	3.61	5.72	19.08	−5.03	3.36	50.72
E5	3.24	1.62	6.08	1.45	1.27	1.92	1.34	0.79	2.45
L1	104.24	84.88	85.05	90.41	86.64	64.63	74.55	73.65	35.28
L2	27.00	16.30	39.32	22.11	13.53	24.23	21.44	18.29	17.56
EF1	3.62	3.60	2.10	3.67	3.57	2.01	5.81	5.26	3.64
EF2	2.78	2.28	2.53	2.71	2.71	1.83	2.71	2.49	2.51
EF3	1.94	1.92	1.66	1.81	1.81	1.20	1.74	1.74	2.03
S1	7.79	7.98	0.91	7.46	7.49	0.81	7.31	7.38	0.91
Regulatory									
R1	5.05	5.00	1.72	5.18	5.00	1.30	4.65	5.00	1.64
R2	6.44	6.00	0.79	6.16	6.00	0.48	6.11	6.00	0.84
R3	11.25	13.00	2.36	11.15	12.00	2.24	10.80	11.00	1.99
R4	9.16	10.00	2.48	9.80	11.00	2.32	8.65	8.00	2.56
Country-level									
MAC1	77.21	81.00	5.93	74.42	73.00	7.42	71.12	81.00	8.05
MAC2	168.33	191.25	39.19	156.21	191.25	49.25	129.15	108.74	64.90
MAC3	0.26	−0.02	1.45	0.22	−0.02	1.31	0.42	0.66	1.54
MAC4	120.39	112.49	48.65	111.20	93.34	43.74	109.69	93.34	42.74
MAC5	59.94	60.85	24.04	53.56	42.43	20.35	62.79	64.75	20.57

account, and the interest income over the same period. Ideally, this ratio should be as low as possible and, in a well-run bank, if the lender is at higher risk, this should be reflected by higher interest margins. If the ratio deteriorates, this means that risk is not being properly remunerated by the margins. Looking again at Table 2, we can observe the increase of this variable through the three groups. In our case, earnings and profitability are represented by the following variables: (1) returns to average total assets (E1), which is perhaps the most important single ratio in the operational performance of banks as it looks at the returns generated from the assets financed by the bank; (2) non-interest expenses to average assets (E2) gives a measure of the costs of the bank's performance with respect to the assets invested; (3) income from other operations to the average assets (E3) ratio, which indicates to what extent fees and other income represent a greater percentage of the bank's earnings (as long as this is not volatile trading income, it can be seen as a lower risk form of income); (4) return on average equity (E4) is a measure of the return on shareholder funds; and (5) recurring earning power (E5) provides a measure of after-tax profits, adding back provision for bad debts as a percentage of total assets. In fact, this is a return-on-assets performance measurement without deducting provisions. Regarding financial efficiency, three variables were studied: (1) interest expenses to average interest-bearing liabilities (EF1); (2) net interest income to average earning assets (EF2), the higher this figure, the cheaper the funding, or the higher the margin the bank is commanding; and, (3) net interest income less loan impairment charges to average earning assets (EF3), which is a financial efficiency ratio before impairment charges.

Another important bank aspect that influences the likelihood of bankruptcy is the bank's liquidity position. According to Friedman and Schwartz (2008), the 1930's economic deceleration caused by previous bank failures was not due to banking problems of insolvency but rather to an unexpected illiquidity in the financial institutions. Therefore, we consider two ratios related to liquidity to demonstrate whether there is an influence on bank soundness in the actual financial crisis context: (1) net loans to deposit and short-term funding measures (L1) both highlight the association between comparatively illiquid assets (loans) and moderately stable funding sources (deposits and other short term funding), thus showing the extent to which the bank has lent its deposits in illiquid form; (2) liquid assets divided by depositors and short term funding (L2), which is a deposit run-off ratio and looks at the percentage of customers and short term funds that could be met if they were suddenly withdrawn. A possible description of management quality could be the competence of the board of directors and management in recognizing, measuring and managing the risks of a financial institution's activities to guarantee its safe, sound, and efficient operation in accordance with applicable laws and regulations. We use growth of gross loan (M1) as a measure of management. Foos et al. (2010) have examined the relationship between loan growth and loan losses, the interest income of banks, and bank solvency. As a result, the ratio of loan growth is shown to have a positive and highly significant influence on ensuing loan losses; nonstandard loan growth leads to a decline in the relative interest income of banks, and finally, abnormal loan growth leads to lower capital ratios, indicating a decrease in bank creditworthiness. Accordingly, monitoring loan growth of institutions may be beneficial for banking supervisors and deposit insurers. Finally, casual empiricism suggests that small banks may be more likely to fail (Wheelock and Wilson 2000). So, in line with previous studies, a logarithm of total assets is used as a measure of size (S1).

3.2 Regulatory Variables

This subsection contributes some variables from theoretical and empirical studies that examine the impact of Basel II type regulations on aspects related to bank performance, especially related to their soundness. The three Pillars of Basel II (i.e., capital adequacy requirements, official supervisory power, and market discipline mechanisms), as well as restrictions on bank activities, on cost and the profit efficiency of banks, are examined under the following indices[3] (Barth et al. 2001a, b, 2008) : the first one (R1), is an index of capital requirements that accounts for both initial and overall capital stringency. Initial capital stringency indicates whether certain funds may be used initially to capitalize a bank and whether they are officially verified. Overall capital stringency indicates whether risk elements and value losses are considered while calculating regulatory capital.

The second Pillar of Basel II emphasizes that supervisors will evaluate the activities and risk profiles of individual banks to determine whether those organizations should hold higher levels of capital than the minimum requirements that Pillar 1 would specify, and to see whether there is any need for remedial action. Therefore, the second index (R2) is constructed, which reveals the ability of the supervisory agencies to take specific actions in banking decisions against the bank management and directors, shareholders, and bank auditors to prevent and correct problems. The third index (R3) that represents Basel II's third Pillar, leverages the ability of market discipline to motivate prudent management by enhancing the degree of transparency in banks' public reporting to shareholders and customers. Thus, we use R3, which reflects the degree to which banks are forced to disclose accurate information to the public (e.g., disclosure of off-balance sheet items, risk management procedures, etc.) and whether there are incentives to increase market discipline, such as subordinated debt and an absence of deposit insurance schemes. Finally, the fourth (R4) index is elaborated to determine whether securities, insurance, real estate activities, and ownership of non-financial firms are unrestricted, permitted, restricted, or prohibited.

3.3 Other Macroeconomic Variables

A number of additional independent variables are incorporated that include a set of macroeconomic and banking sector ratios. The role that macroeconomic factors can play in determining the soundness of individual banks has been examined by González-Hermosillo et al. (1997). It seems that bank fragility is determined not only by bank-specific ratios but also by the condition of the banking sector as a whole, as well as by the macroeconomic environment in which the banks operate.

We use the Heritage Foundation indicator of the protection of property rights, which is an assessment of the capacity of individuals to accumulate private property, secured by clear laws that are fully enforced by the state (MAC1). It also assesses the likelihood of private property being expropriated, and analyzes the independence of the judiciary, the existence of corruption within the judiciary and the ability of individuals and businesses to enforce contracts. It can take values between 0 and 100. The more certain the legal protection of property is, the higher a country's score; similarly, the greater the chances of government expropriation of property, the lower a country's score. Normally, excessive risk-taking and fraud occur in countries with high financial

[3] In the Appendix we explain how all these indexes-variables, which are proxies of regulatory aspects, are constructed.

liberalization, which assumes an increase in banking sector fragility (Demirgüç-Kunt and Detragiache 1998). As an alternative for financial liberalization, we use the ratio of domestic credit to the private sector over GDP (%) (MAC2).

Trying to understand economic growth, and its positive influence on decreasing non-performing loans and the banking crisis (Davis and Karim 2008), we use real annual GDP growth (%) (MAC3). Low liquidity in the banking sector appears to be linked to a susceptibility to crisis (Demirguc-Kunt and Detragiache 1997). Accordingly, the ratio of bank credit to bank deposits is used (MAC4) as a proxy for banking sector liquidity.

As a final factor, and a controversial point, we attempt to look at a measure of concentration. Illustrating this controversy, Liang (1989) found that bank risk and market concentration are positively related. In contrast, Beck et al. (2006) found that a systemic banking crisis would be less likely to occur in more concentrated national systems. Hence, we use the percentage of assets held by the three largest commercial banks in relation to the total assets of the commercial banking sector in the country to measure concentration (MAC5).

4 Methodology

4.1 Extreme Learning Machine

In recent times, computational intelligence techniques have been employed for a wide range of applications, such as marketing and sales, manufacturing, finance, risk assessment, etc. The application of artificial neural networks (ANNs) to finance issues started in the 1990s, gaining popularity as an alternative to traditional statistical techniques due to their flexibility and their non-parametric nature (Tam and Kiang 1990). However, beyond doubt, the learning speed of forward neural networks is generally far slower than required, and it has shown poor computational scalability. A new learning algorithm called the Extreme Learning Machine (ELM) Huang et al. (2006a) appears to overcome these drawbacks. The ELM scheme proposed by Huang et al. (2006a) overcomes the problems associated to gradient-based methods by randomly assigning weights to the input layers and analytically computing the weights for the output layer using a simple generalized inverse operation. This finding has recently attracted the attention of computational intelligence and the machine learning field, in both theory and application.

Suppose that there are N training patterns $(\mathbf{x}_n, \mathbf{y}_n)$, $n = \{1, 2, \ldots, N\}$, where $\mathbf{x}_n = (x_{n1}, x_{n2}, \ldots, x_{nK})$ and $\mathbf{y}_n = (y_{n1}, y_{n2}, \ldots, y_{nJ})$ are the n-th input pattern and its target, respectively (assuming a classification problem with J classes and K input variables). Let's note as $\mathbf{v}_s = (v_{s1}, v_{s2}, \ldots, v_{sK})$ the weight vector connecting the input nodes to the s-th basis function, for $s = \{1, 2, \ldots, S\}$ and with $\boldsymbol{\beta}^j = (\beta_1^j, \ldots, \beta_S^j)$ the weight vector connecting the basis functions to the j-th output node for $j = \{1, \ldots, J\}$. During the training process, ELM determines the parameters $\boldsymbol{\beta}^j$, for all j values, by minimizing the following error function:

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^S \times \mathbb{R}^J} \left(\|\mathbf{H}\boldsymbol{\beta} - \mathbf{Y}\|^2,\ \|\boldsymbol{\beta}\| \right) \tag{1}$$

where $\|\cdot\|$ is the L2 norm, $\mathbf{H}$ is the hidden layer output matrix of the SLFN:

$$\begin{aligned}\mathbf{H} &= (\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_S) \\ &= \begin{pmatrix} \phi_1(\mathbf{x}_1; \mathbf{v}_1) & \ldots & \phi_S(\mathbf{x}_1; \mathbf{v}_S) \\ \ldots & \ldots & \ldots \\ \phi_1(\mathbf{x}_N; \mathbf{v}_1) & \ldots & \phi_S(\mathbf{x}_N; \mathbf{v}_S) \end{pmatrix} \in \mathbb{R}^N \times \mathbb{R}^S \qquad (2)\\ \mathbf{Y} &= (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N)^T \in \mathbb{R}^N \times \mathbb{R}^J, \qquad (3)\end{aligned}$$

and

$$\boldsymbol{\beta} = (\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, \ldots, \boldsymbol{\beta}^J) \in \mathbb{R}^S \times \mathbb{R}^J. \qquad (4)$$

where $\phi(\mathbf{x}_n; \mathbf{v}_s)$ is the activation function of the n-th pattern for the s-th basis function. The ELM algorithm starts choosing the activation function $\phi(\mathbf{x}, \mathbf{v})$ (generally the sigmoidal one) and the number of basis functions S. In the first step, random values are assigned to the input weight vectors $\mathbf{v}_s$. Thus, the problem of minimizing the error reduces to solve:

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{Y}. \qquad (5)$$

The output weights, $\boldsymbol{\beta}$, are estimated as:

$$\hat{\boldsymbol{\beta}} = \mathbf{H}^{\dagger}\mathbf{Y}, \qquad (6)$$

where

$$\mathbf{H}^{\dagger} = \begin{cases} \mathbf{H}^T \left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^T\right)^{-1} & \text{for } N < S, \\ \left(\frac{\mathbf{I}}{C} + \mathbf{H}^T\mathbf{H}\right)^{-1} \mathbf{H}^T & \text{otherwise,} \end{cases} \qquad (7)$$

and $C \in \mathbb{R}$ is a user-specified parameter that promotes generalization performance.

The output function of the ELM classifier is defined as (just for the case $N < S$)

$$\begin{aligned} f(\mathbf{x}) &= \mathbf{h}(\mathbf{x})\hat{\boldsymbol{\beta}} \\ &= \mathbf{h}(\mathbf{x})\mathbf{H}^T \left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^T\right)^{-1} \mathbf{Y}, \qquad (8)\end{aligned}$$

where $\mathbf{h}(\mathbf{x})$ is a mapping function which corresponds to the basis functions output in the ANN literature or it is unknown to users in the kernel machines literature. Therefore, the output function can be kernelized (obtaining in this way the so-called kernelized extreme learning machine, K-ELM), as suggested in Huang et al. (2012), as

$$f(\mathbf{x}) = \mathbf{K}(\mathbf{x})^T \left(\frac{\mathbf{I}}{C} + \boldsymbol{\Omega}_{\mathrm{ELM}}\right)^{-1} \mathbf{Y}, \qquad (9)$$

where $\mathbf{K}(\mathbf{x}) : \mathbb{R}^K \rightarrow \mathbb{R}^N$ is the vector of kernel functions

$$\mathbf{K}(\mathbf{x})^T = (K(\mathbf{x}, \mathbf{x}_1), \ldots, K(\mathbf{x}, \mathbf{x}_N)).$$

The Gaussian kernel function implemented in this study is

$$K(\mathbf{x}, \mathbf{x}_i) = \exp(-k||\mathbf{x} - \mathbf{x}_i||^2), \quad i = \{1, \ldots, N\}, \tag{10}$$

where $k \in \mathbb{R}$ is the kernel parameter. The kernel matrix $\boldsymbol{\Omega}_{\text{ELM}} = [\Omega_{i,j}]_{i,j=1,\ldots,N}$ is defined as:

$$\Omega_{i,j} = K(\mathbf{x}_i, \mathbf{x}_j). \tag{11}$$

4.2 SMOTE: Synthetic Minority Over-Sampling Technique

The success of machine learning algorithms is classically evaluated using predictive accuracy. Even so, the performance of machine learning algorithms is limited when the data is imbalanced and/or the costs of different errors vary markedly Akbani et al. (2004). The SMOTE algorithm oversamples the minority class by taking each minority class sample and introducing synthetic examples along the line segments joining any or all of the k minority class' nearest neighbors. Depending on the amount of over-sampling required, neighbors are randomly chosen from the k nearest neighbors. Synthetic samples are generated as follows: take the difference between the feature vector (sample) under consideration and its nearest neighbor. Multiply this difference by a random number between 0 and 1, and add it to the feature vector under consideration. This results in the selection of a random point along the line segment between two specific features. This approach effectively forces the decision region of a minority class to become more general (Chawla et al. 2002; Fernández-Navarro et al. 2011). In this study, k was set to five and the number of instances for classes 1 and 3 were doubled (with synthetic patterns) in the training set (the minority classes).

5 Experimental Framework

5.1 Data Description

In the present study, we use the Fitch Individual bank ratings, which are based on an A to F scale and represent Fitch's opinion on the financial strength of a bank independently. This is complemented by a support rating that expresses the likelihood and sources of external support in the event that a financial institution undergoes hardship. As mentioned earlier, the purpose of the present study is not to explain or replicate the ratings of Fitch, but rather to use them as the basis for the development of a general model to assess the soundness of banks. We, therefore, classify the banks into three broad groups. The first consists of banks with ratings A and B, the second of banks with rating C, and the third includes banks rated D, E and F. Hence, banks in Group 1 can be considered as very strong or strong banks, banks in Group 2 can be considered adequate banks, and those in Group 3 can be considered as banks with weaknesses or serious problems.

An advantage in using ratings as an indicator of bank soundness rather than individual measures, such as non-performing loans, coverage ratio or Z-scores, is the consensus of the analysts, which indicates that rating agencies are believed by the market to have superior information, and in addition, to support a comprehensive measure of the ability of a bank to meet its obligations to depositors and other creditors[4] (Demirgüç-Kunt et al. 2008). Another benefit of Fitch's ratings is that they evaluate bank solvency independently of the safety net, so that cross-country differences in the safety net, which are difficult to observe and measure, should not affect the results. A limitation of using Fitch's ratings is that the sample is restricted to larger banks, as smaller banks are not rated. Thus, our investigation will not address the impact of financial health on smaller banks. As these banks are not likely to be of systemic importance, this limitation should be relatively minor.

Our dataset consists of 337 banks from 28 countries with data and Fitch individual bank ratings available in the Bankscope database.[5] The ratings were obtained at the end of 2009, while the bank specific characteristics correspond to the end of 2008. The distribution of banks into the three groups is as follows: 64 (Group 1), 179 (Group 2), and 94 (Group 3). Table 3 presents the definitions of the Fitch individual bank ratings, together with the coding used in the present study and the distribution of banks into the three groups. Table 4 presents the number of banks in the sample by country and group.

For all the methods, we adopted a 10-fold cross validation in the experimentation in order to assess how the results generalize. The procedure was repeated 10 times resulting in a total of 100 training and test splits. The same partitions were used for all the methods. Furthermore, we did a simple linear rescaling of the input variables over the interval $[-2, 2]$, with X_i^* being the transformed variables.

5.2 Evaluation Metrics

Two metrics were used to assess the performance of each method in the generalization set:

– *Accuracy rate* (*Acc*) It is the number of successful hits (correct classifications) relative to the total number of classifications. It is the most widely used metric in the Machine Learning community.

[4] In support of this view, Sironi (2003) finds that credit ratings outperform the balance sheet variables in predicting spreads on bank subordinated notes and debentures in Europe. Other studies have shown that changes in credit ratings cause changes in equity prices of banks in the United States (Billett et al. 1998; Schweitzer et al. 1992) and in Europe (Gropp and Richards 2001), indicating that rating agencies are believed by the market to have superior information.

[5] This is a comprehensive, global database containing information on public and private banks commercialized by the Bureau van Dijk Group. The database is updated to December 6, 2012 (version number 1349).

Table 3 Definitions of Fitch's bank individual ratings and bank soundness

Fitch rating	Code	Sample	Fitch definition
A	Group 1	64	A very strong bank. Characteristics may include outstanding profitability and balance sheet integrity, franchise, management, operating environment or prospects
B	Group 1		A strong bank. There are no major concerns regarding the bank. Characteristics may include strong profitability and balance sheet integrity, franchise, management, operating environment or prospects
C	Group 2	179	An adequate bank which, however, possesses one or more troublesome aspects. There may be some concerns regarding its profitability and balance sheet integrity, franchise, management, operating environment or prospects
D	Group 3	94	A bank, which has weaknesses of internal and/or external origin. There are concerns regarding its profitability, substance and resilience, balance sheet integrity, franchise, management, operating environment or prospects. Banks in emerging markets are necessarily faced with a greater number of potential deficiencies of external origin
E	Group 3		A bank with very serious problems, which either requires or is likely to require external support
F	Group 3		A bank that has either defaulted or, in Fitch's opinion, would have defaulted if it had not received external support. Examples of such support include state or local government support; (deposit) insurance funds; acquisition by some other corporate entity or an injection of new funds from its shareholders or equivalent

Fitch also uses the following intermediate assignments among the major five categories: A/B, B/C, C/D, and D/E

– *Minimun Sensitivity* (MS) The accuracy for the worst classified class. This measure has been recently used in Machine Learning to evaluate the performance of a classifier in imbalanced classification environments. For example, accuracy Acc and MS are optimized through an evolutionary process in Fernández-Navarro et al. (2011), while, the optimization is carried out by a Pareto-based multiobjective optimization methodology based on a memetic evolutionary algorithm in Caballero et al. (2010).

Table 4 Distribution of banks by country and soundness group

	G-1	G-2	G-3		G-1	G-2	G-3
Austria	–	1	3	Italy	2	18	1
Australia	7	4	1	Japan	1	11	4
Belgium	–	–	3	Korea (Rep. of)	–	9	2
Switzerland	1	6	–	Luxembourg	–	–	1
Chile	–	1	–	Mexico	–	–	4
Czech Republic	–	3	–	Netherlands	1	4	2
Germany	–	7	14	Norway	2	5	–
Denmark	1	1	–	New Zealand	4	–	1
Spain	6	6	2	Poland	–	2	4
Finland	2	–	–	Portugal	–	2	3
France	3	9	3	Sweden	3	2	–
United kingdom	6	14	7	Turkey	–	5	12
Greece	–	–	5	USA	25	68	16
Hungary	–	–	1				
Ireland	–	1	5	Total	64	179	94

5.3 ELM Algorithms Selected and Parameter Selection

The proposed method is compared to the following ELM approaches:

– The Extreme learning machine (ELM) (Huang et al. 2006a) with the sigmoidal basis function. In this method, the number of basis functions has been selected by a nested five fold cross-validation procedure over the training set, i.e., once the lowest cross-validation error alternative was obtained, it was applied to the complete training set and test results were extracted. The criteria for selecting the best configuration was the Acc metric. The values considered in the cross-validation procedure were $\{10, 20, \ldots, 100\}$.
– The optimally pruned extreme learning machine (OP-ELM) (Miche et al. 2010). The OP-ELM extends the original ELM algorithm wrapping it around with a methodology that prunes the hidden neurons. Again, the sigmoidal non-linear transformation was the one considered in this study. The number of basis functions in the OP-ELM algorithm is set at the beginning to 100.
– The incremental extreme learning machine (I-ELM) (Huang et al. 2006b). This algorithm proposes a procedure that increases the network architecture adding random nodes till the residual error is bigger than a threshold.
– The pruned extreme learning machine (P-ELM) (Rong et al. 2008) uses statistical methods to measure the relevance of basis functions (which are ranked using the chi-square and information gain statistical techniques). Irrelevant nodes are pruned by considering their relevance to the class labels. The optimal number of hidden nodes is estimated considering the performance of the classifier on a (stratified)

validation set constructed with the 25% of the training set patterns as suggested by the authors.
– The principal component analysis extreme learning machine (PCA-ELM) (Castaño et al. 2013). The parameters of the hidden layer are estimated using the information retrieved from principal components analysis and, therefore, the number of basis functions is calculated deterministically (according to the accumulated variance expressed into the eigenvectors).

The optimal two hyperparameter values for the method proposed (the K-ELM model) were selected using a nested five fold cross-validation over the training set ($k \in \{10^{-3}, 10^{-2}, \ldots, 10^{3}\}$ and $C \in \{10^{-3}, 10^{-2}, \ldots, 10^{3}\}$). These methods have been selected for comparison as they share the same foundations as the method proposed and furthermore they are some of the best performing algorithms of recent literature on ELM. Some of these models (with or without some slight modifications) have also been tested before in bankruptcy detection problem (Yu et al. 2014; Zhao et al. 2016).

6 Results

Table 5 shows the overall generalization results obtained with the different techniques tested. As can be seen in Table 5, the K-ELM method achieved the best results in all the cases considered. In the first part of Table 5, models were tested using all the bank-specific variables. The K-ELM yielded 72.44% of accuracy using only this set of variables (and a minimum sensitivity of 55.14%). Thus, we can affirm that the variables proposed are able to adequately resolve the challenge of predicting bank classification into the three groups. In addition, we achieved low levels of extreme misclassification of banks. So, extensive lists of variables that reflect the main features of banks (i.e. capital, assets quality, management, earnings, liquidity, financial efficiency and size) seem to be good predictors of bank soundness.

The classification results, after the incorporation of some regulatory variables and other country specific characteristics, are displayed in the second part of Table 5. At this point, the model is in a position to distinguish the banks in class 3 more accurately[6] (with an improvement of 10.79% points). This means that the macroeconomic variables have helped to differentiate, above all, the banks that have weaknesses or serious problems. Furthermore, the overall performance has improved to achieve 77.31% accuracy.

Last part of Table 5 presents the results of the model when all the input variables are considered and the Synthetic Minority Over-sampling Technique (SMOTE) technique is applied on the minority classes. The results showed that the weakest previous performance (MS) has improved almost 3% points if compared with the previous case. Consequently, we achieved the results expected, confirming the advantages of implementing the novel hybrid model (combining ELM and SMOTE) for this type of problem.

Finally, each pair of algorithms was compared by means of the Wilcoxon test (as the normality assumption was rejected) (Demšar 2006). A level of significance

[6] The class which reports the minimum sensitivity value.

Table 5 Generalization results of the Acc and MS of the ELM methods along with the p values of the Wilcoxon rank sum test

	Acc	MS	p value$_{Acc}$	p value$_{MS}$
Using only bank-specific variables				
ELM	$66.87_{2.21}$	$50.20_{4.56}$	1.8E-4*	1.8E-4*
OP-ELM	$62.65_{3.17}$	$45.86_{4.53}$	1.8E-4*	1.8E-4*
I-ELM	$67.23_{2.20}$	$48.99_{5.23}$	1.8E-4*	1.8E-4*
P-ELM	$67.19_{2.29}$	$47.11_{5.03}$	1.8E-4*	1.8E-4*
PCA-ELM	$\mathit{71.92_{2.32}}$	$\mathit{52.07_{4.04}}$	0.0781	0.0257*
K-ELM	$\mathbf{72.44_{2.45}}$	$\mathbf{55.14_{5.04}}$	–	–
Using all variables				
ELM	$\mathit{76.53_{1.90}}$	$\mathit{54.20_{6.12}}$	0.1276	0.0455*
OP-ELM	$71.65_{2.42}$	$57.21_{5.34}$	1.8E-4*	1.8E-4*
I-ELM	$72.23_{1.73}$	$59.13_{5.04}$	1.8E-4*	1.8E-4*
P-ELM	$71.99_{1.45}$	$59.05_{5.85}$	1.8E-4*	1.8E-4*
PCA-ELM	$75.22_{1.32}$	$63.47_{6.54}$	0.0318*	0.0210*
K-ELM	$\mathbf{77.31_{2.05}}$	$\mathbf{65.93_{3.91}}$	–	–
Including SMOTE on the training set				
ELM	$78.32_{2.88}$	$66.95_{5.89}$	0.0253*	0.0371*
OP-ELM	$73.41_{2.01}$	$60.83_{6.09}$	1.8E-4*	1.8E-4*
I-ELM	$75.94_{3.34}$	$64.26_{5.11}$	1.8E-4*	1.8E-4*
P-ELM	$76.03_{2.29}$	$63.87_{4.28}$	1.8E-4*	1.8E-4*
PCA-ELM	$\mathit{78.68_{1.89}}$	$\mathit{67.59_{4.34}}$	0.0123*	0.0785
K-ELM	$\mathbf{80.05_{1.43}}$	$\mathbf{68.21_{2.12}}$	–	–

The best result is in bold face and the second one in italics

* The null hypothesis that results provided by the comparison method the results of K-ELM are samples with equal medians is rejected

of $\alpha = 0.05$ was considered, and the corresponding correction for the number of comparisons was also included. The control method was the K-ELM method, since it got the best mean ranking in Acc and MS. As shown in Table 5, the K-ELM method achieved significantly better results than the remaining methods, except PCA-ELM in the MS metric when the whole set of variables was considered and SMOTE was applied on the training set.

7 Conclusions

The cost of the financial crisis which peaked in 2008 was felt by all the members of the international banking system. Six years of deep banking turbulence have revealed the critical importance of these institutions to the overall economy. Therefore the need to develop early warning models that will help to maintain a healthy banking system and avoid similar problems in the future has become a priority. Using a sample of 337 banks from 28 countries, we developed a quantitative model and examined their accuracy in classifying banks into three groups. We presented a recent methodology called Extreme learning machine (ELM) combine with an oversampling technique (SMOTE). This hybrid method resolves the drawbacks related to parametric models, as mentioned in the introduction, which are easily violated when financial ratios and aggregate account balances are used as input. In addition, it overcomes the main disadvantages of needing to find the optimal framework of the ANN model and the corresponding training required to estimate the neural network weights, and finally, solve the imbalanced nature of the problem. The application of the hybrid model to the evaluation of bank soundness in OECD countries resulted in a 80.05% correct classification. The model proposed thus provides a useful tool for meeting the aims of the "prudential" supervision of banks as it presents a low ratio of misclassification.

The results and comparison with other classification models confirmed the suitability and robustness of the hybrid methodology proposed, as it presents better performance rates than all other methods tested. We believe that the model developed in the present study could be beneficial in measuring the solvency of banks, and in helping to identify potential deterioration when financial institutions move from a "strong group" to a "weaker group" . Also, due to the easy implementation and accessibility of the data used, supervisors, investors and analysts could all use this model.

Moreover, the methodology allows for the addition of multiple variables that have no correlation problems. As a result, we added different perspectives on the CAMEL variables as well as information regarding financial efficiency and size. This additional information proved essential in improving average accuracy and especially the performance of the minority groups. Thus, the advantages provided by the methodology in incorporating more information constitute a significant gain. Once regulatory and macroeconomic variables were included, the model recorded an important improvement in classification accuracy of Group 3 banks, and an improved overall accuracy.

Future research could be extended in various ways: first, different criteria could be used to rate banks into three groups, other than the ratings of Fitch, thereby spreading the analysis to more banks. A second possible direction would be to do a sensitivity analysis of the variables employed in the model. Third, it would be interesting to study the temporary effects of this type of model by making a longitudinal study instead of a cross-sectional one.

Acknowledgements The research work of F. Martínez-Estudillo and F. Fernández-Navarro was partially supported by the TIN2014-54583-C2-1-R project of the Spanish Ministry of Economy and Competitiveness (MINECO).

Appendix A: Definition of Regulatory Variables

See Table 6.

Table 6 Definition of regulatory variables

Var.	Category	Description
R1	Capital requirements index	This variable is determined by adding 1 if the answer is yes to questions 1–6 and 0 otherwise, while the opposite occurs in the case of questions 7 and 8 (i.e. yes=0, no=1). (1) Is the minimum required capital asset ratio risk-weighted in line with Basel guidelines? (2) Does the ratio vary with market risk? (3–5) Before minimum capital adequacy is determined, which of the following are deducted from the book value of capital: (a) market value of loan losses not realized in accounting books? (b) unrealized losses in securities portfolios? (c) unrealized foreign exchange losses? (6) Are the sources of funds to be used as capital verified by the regulatory/supervisory authorities? (7) Can the initial or subsequent injections of capital be done with assets other than cash or government securities? (8) Can initial disbursement of capital be done with borrowed funds? Higher values indicate higher capital stringency
R2	Official disciplinary power index	This variable is determined by adding 1 if the answer is yes and 0 otherwise, for each one of the following fourteen questions: (1) Does the supervisory agency have the right to meet with external auditors to discuss their report without the approval of the bank? (2) Are auditors required by law to communicate directly to the supervisory agency any presumed involvement of bank directors or senior managers in illicit activities, fraud, or insider abuse? (3) Can supervisors take legal action against external auditors for negligence? (4) Can the supervisory authorities force a bank to change its internal organizational structure? (5) Are off-balance sheet items disclosed to supervisors? (6) Can the supervisory agency order the bank's directors or management to constitute provisions to cover actual or potential losses? (7) Can the supervisory agency suspend a director's decision to distribute dividends? (8) Can the supervisory agency suspend a director's decision to distribute bonuses? (9) Can the supervisory agency suspend a director's decision to distribute management fees? (10) Can the supervisory agency supersede bank shareholder rights and declare banks insolvent? (11) Does banking law allow a supervisory agency or any other government agency (other than court) to suspend some or all ownership rights of a problem bank? (12) Regarding bank restructuring and reorganization, can the supervisory agency or any other government agency (other than a court) supersede shareholder rights? (13) Regarding bank restructuring and reorganization, can a supervisory agency or any other government agency (other than a court) remove and replace management? (14) Regarding bank restructuring and reorganization, can a supervisory agency or any other government agency (other than a court) remove and replace directors? Higher values indicate more powerful supervisors

Table 6 continued

Var.	Category	Description
R3	Market discipline index	This variable is determined by adding 1 if the answer is yes to questions 1–7 and 0 otherwise, while the opposite occurs in the case of questions 8 and 9 (i.e. yes=0, no=1). (1) Is subordinated debt allowable (or required) as part of capital? (2) Are financial institutions required to produce consolidated accounts covering all bank and any non-bank financial subsidiaries? (3) Are off- balance sheet items disclosed to public? (4) Must banks disclose their risk management procedures to public? (5) Are directors legally liable for erroneous/misleading information? (6) Do regulations require credit ratings for commercial banks? (7) Is an external audit by certified/licensed auditor a compulsory obligation for banks? (8) Does accrued, though unpaid interest/principal enter the income statement while a loan is non-performing? (9) Is there an explicit deposit insurance protection system? Higher values indicate a regulatory framework that promotes market discipline
R4	Restrictions on banks activities index	The score for this variable is determined on the basis of the level of regulatory restrictiveness for bank participation in: (1) securities activities (2) insurance activities (3) real estate activities (4) bank ownership of non-financial firms. These activities can be unrestricted, permitted, restricted or prohibited, which are assigned the values of 1, 2, 3 or 4 respectively. We use an overall index by calculating the average value over the four categories. Higher values indicate higher restrictions

Source of the variables: Bank Regulation and Supervision World database

References

Akbani, R., Kwek, S., & Japkowicz, N. (2004). Applying support vector machines to imbalanced datasets. In *European conference on machine learning*, Springer, pp. 39–50.

Barth, J.R., Caprio, G., & Levine, R. (2001). The regulation and supervision of banks around the world—A new database, Policy Research Working Paper Series 2588, The World Bank.

Barth, J. R., Caprio, G., & Levine, R. (2001). *The regulation and supervision of banks around the world: A new database* (Vol. 2588). Washington: World Bank Publications.

Barth, J. R., Caprio, G, Jr., & Levine, R. (2008). Bank regulations are changing: For better or worse? *Comparative Economic Studies*, *50*(4), 537–563.

Beck, T., Demirgüç-Kunt, A., & Levine, R. (2000). A new database on the structure and development of the financial sector. *The World Bank Economic Review*, *14*(3), 597–605.

Beck, T., Demirgüç-Kunt, A., & Levine, R. (2006). Bank concentration, competition, and crises: First results. *Journal of Banking & Finance*, *30*(5), 1581–1603.

Billett, M. T., Garfinkel, J. A., & O'Neal, E. S. (1998). The cost of market versus regulatory discipline in banking. *Journal of Financial Economics*, *48*(3), 333–358.

Bongini, P., Claessens, S., & Ferri, G. (2001). The political economy of distress in east asian financial institutions. *Journal of Financial Services Research*, *19*(1), 5–25.

Caballero, J. C. F., Martínez, F. J., Hervás, C., & Gutiérrez, P. A. (2010). Sensitivity versus accuracy in multiclass problems using memetic pareto evolutionary neural networks. *IEEE Transactions on Neural Networks*, *21*(5), 750–770.

Castaño, A., Fernández-Navarro, F., & Hervás-Martínez, C. (2013). Pca-elm: A robust and pruned extreme learning machine approach based on principal component analysis. *Neural Processing Letters*, *37*(3), 377–392.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.

Chen, Y.-S., & Cheng, C.-H. (2012). A soft-computing based rough sets classifier for classifying ipo returns in the financial markets. *Applied Soft Computing*, *12*(1), 462–475.

Chortareas, G. E., Girardone, C., & Ventouri, A. (2012). Bank supervision, regulation, and efficiency: Evidence from the european union. *Journal of Financial Stability*, *8*(4), 292–302.

Čihák, M., & Poghosyan, T. (2009). Distress in European banks: An analysis based on a new data set, Technical report, IMF Working Paper.

Davis, E. P., & Karim, D. (2008). Comparing early warning systems for banking crises. *Journal of Financial Stability*, *4*, 89–120.

Demirgüç-Kunt, A., & Detragiache, E. (1998). The determinants of banking crises in developing and developed countries. *IMF Staff Papers*, *45*(1), 81–109.

Demirguc-Kunt, A., Detragiache, E., et al. (1997). The determinants of banking crises: evidence from industrial and developing countries, Technical report. The World Bank.

Demirgüç-Kunt, A., Detragiache, E., & Gupta, P. (2006). Inside the crisis: An empirical analysis of banking systems in distress. *Journal of International Money and Finance*, *25*(5), 702–718.

Demirgüç-Kunt, A., Detragiache, E., & Tressel, T. (2008). Banking on the principles: Compliance with basel core principles and bank soundness. *Journal of Financial Intermediation*, *17*(4), 511–542.

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, *7*, 1–30.

Demyanyk, Y., & Hasan, I. (2010). Financial crises and bank failures: A review of prediction methods. *Omega*, *38*(5), 315–324.

DeYoung, R., & Torna, G. (2013). Nontraditional banking activities and bank failures during the financial crisis. *Journal of Financial Intermediation*, *22*(3), 397–421.

Duttagupta, R., & Cashin, M.P. (2008). *The anatomy of banking crises*. IMF Working Paper, International Monetary Fund, *93*, 1–37.

Estrella, A., Park, S., & Peristiani, S. (2000). Capital ratios as predictors of bank failure. *Economic Policy Review*, *6*(2), 33–52.

Fernández-Navarro, F., Hervás-Martínez, C., & Gutiérrez, P. A. (2011). A dynamic over-sampling procedure based on sensitivity for multi-class problems. *Pattern Recognition*, *44*(8), 1821–1833.

Foos, D., Norden, L., & Weber, M. (2010). Loan growth and riskiness of banks. *Journal of Banking & Finance*, *34*(12), 2929–2940.

Friedman, M., & Schwartz, A. J. (2008). *A monetary history of the United States, 1867–1960* (2nd ed.). Princeton: Princeton University Press.

Gaganis, C., Pasiouras, F., & Zopounidis, C. (2006). A multicriteria decision framework for measuring banks' soundness around the world. *Journal of Multi-Criteria Decision Analysis*, *14*(1–3), 103–111.

González-Hermosillo, B., Pazarbaşioğlu, C., & Billings, R. (1997). Determinants of banking system fragility: A case study of mexico. *Staff Papers*, *44*(3), 295–314.

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

Gramlich, D., Miller, G., Oet, M. V., & Ong, S. J. (2010). Early warning systems for systemic banking risk: Critical review and modeling implications. *Banks and Bank Systems*, *5*, 199–211.

Gropp, R., & Richards, A. J. (2001). Rating agency actions and the pricing of debt and equity of european banks: What can we infer about private sector monitoring of bank soundness? *Economic Notes*, *30*(3), 373–398.

Gutiérrez, P., Segovia-Vargas, M., Salcedo-Sanz, S., Hervás-Martínez, C., Sanchis, A., Portilla-Figueras, J., et al. (2010). ybridizing logistic regression with product unit and rbf networks for accurate detection and prediction of banking crises. *Omega*, *38*(5), 333–344.

Huang, G. B., Zhu, Q., & Siew, C. (2006a). Extreme learning machine: Theory and applications. *Neurocomputing*, *70*(1–3), 489–501.

Huang, G.-B., Chen, L., & Siew, C. K. (2006). Universal approximation using incremental constructive feedforward networks with random hidden nodes. *IEEE Transactions on Neural Networks*, *17*(4), 879–892.

Huang, G.-B., Zhou, H., Ding, X., & Zhang, R. (2012). Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *42*(2), 513–529.

Ioannidis, C., Pasiouras, F., & Zopounidis, C. (2010). Assessing bank soundness with classification techniques. *Omega*, *38*(5), 345–357.

Kumar, P. R., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques-a review. *European Journal of Operational Research*, *180*(1), 1–28.

Liang, N. (1989). Bank profits, risk, and local market concentration. *Journal of Economics and Business*, *41*(4), 297–305.

Lindgren, C.-J., Garcia, G. G., & Saal, M. I. (1996). *Bank soundness and macroeconomic policy*. Washington: International Monetary Fund.

Martin, D. (1977). Early warning of bank failure: A logit regression approach. *Journal of banking & finance*, *1*(3), 249–276.

Miche, Y., Sorjamaa, A., Bas, P., Simula, O., Jutten, C., & Lendasse, A. (2010). OP-ELM: Optimally pruned extreme learning machine. *IEEE Transactions on Neural Networks*, *21*(1), 158–162.

Moore, J. (2011). Bank performance prediction during the'great recession'of 2008-'09: A pattern-recognition approach. *Academy of Banking Studies Journal*, *10*(2), 87.

Olmeda, I., & Fernández, E. (1997). Hybrid classifiers for financial multicriteria decision making: The case of bankruptcy prediction. *Computational Economics*, *10*(4), 317–335.

Porath, D. (2006). Estimating probabilities of default for german savings banks and credit cooperatives. *Schmalenbach Business Review*, *58*, 214–233.

Rong, H.-J., Ong, Y.-S., Tan, A.-H., & Zhu, Z. (2008). A fast pruned-extreme learning machine for classification problem. *Neurocomputing*, *72*(1–3), 359–366.

Rose, A. K., & Spiegel, M. M. (2012). Cross-country causes and consequences of the 2008 crisis: Early warning. *Japan and the World Economy*, *24*(1), 1–16.

Schweitzer, R., Szewczyk, S. H., & Varma, R. (1992). Bond rating agencies and their role in bank market discipline. *Journal of Financial Services Research*, *6*(3), 249–263.

Sironi, A. (2003). Testing for market discipline in the european banking industry: Evidence from subordinated debt issues. *Journal of Money, Credit, and Banking*, *35*(3), 443–472.

Tam, K. Y., & Kiang, M. (1990). Predicting bank failures: A neural network approach. *Applied Artificial Intelligence an International Journal*, *4*(4), 265–282.

Thomson, J. (1991). Predicting bank failures in the 1980s. *Economic Review*, *27*(1), 9–20.

Wheelock, D. C., & Wilson, P. W. (2000). Why do banks disappear? The determinants of us bank failures and acquisitions. *Review of Economics and Statistics*, *82*(1), 127–138.

Yu, Q., Miche, Y., Séverin, E., & Lendasse, A. (2014). Bankruptcy prediction using extreme learning machine and financial expertise. *Neurocomputing*, *128*, 296–302.

Zhao, D., Huang, C., Wei, Y., Yu, F., Wang, M., & Chen, H. (2016). An effective computational model for bankruptcy prediction using kernel extreme learning machine approach. *Computational Economics*, *49*(2), 325–341.