International Journal of Forecasting 34 (2018) 440–455

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting bank failures and stress testing: A machine learning approach

Periklis Gogas, Theophilos Papadimitriou, Anna Agrapetidou *

Democritus University of Thrace, Department of Economics, Komotini 69100, Greece

ARTICLE INFO

Keywords:
Machine learning
Bank failures
Stress testing
Forecasting

ABSTRACT

This paper presents a forecasting model of bank failures based on machine-learning. The proposed methodology defines a linear decision boundary that separates the solvent banks from those that failed. This setup generates a novel alternative stress-testing tool. Our sample of 1443 U.S. banks includes all 481 banks that failed during the period 2007–2013. The set of explanatory variables is selected using a two-step feature selection procedure. The selected variables were then fed to a support vector machines forecasting model, through a training–testing learning process. The model exhibits a 99.22% overall forecasting accuracy and outperforms the well-established Ohlson's score.

1. Introduction

Historically, the banks' role of intermediation between surplus and deficit agents has been crucial for economic activity. This has remained true over the last few decades, despite the increased significance of capital markets and direct financing. Financial institutions are highly interconnected operationally, and the complex channels of interconnection increase the associated systemic risk. As a result, the issue of their financial health is always topical, and a prerequisite to maintaining stability in the economy. Usually, a banking crisis is transmitted swiftly to other sectors within the originating country or to other economies, triggering financial distress on an international scale. The financial crisis of 2007 is an example of the importance of this interconnection. Only 29 banks out of a total of more than 6000 failed in the U.S. during the seven years from 2000 to 2006. During the next seven years (2007–2013), though, bankruptcies increased by 17 times, reaching a total of 492 failed banks. At the same time, the crisis started spreading internationally. The global financial crisis that followed highlighted the need for a stricter and more efficient supervision of financial institutions, in addition to raising macro-prudential concerns. Stress-testing has proved to be a useful and popular tool for regulators internationally. In 2009, the Federal Reserve implemented the Supervisory Capital Assessment Program (SCAP), known as stress-testing, on the 19 largest bank-holding companies. Since 2011, stress tests have been being conducted as part of the Comprehensive Capital Assessment Review (CCAR) and the Dodd-Frank Act.

Various types of modeling techniques have been applied in the literature in an attempt to forecast bankruptcies. Among the most prominent techniques are: linear probability (Meyer & Pifer, 1970), multivariate discriminant analysis (MDA) (Altman, Haldeman, & Narayanan, 1977; Cox & Wang, 2014; Sinkey, 1975; Stuhr & Van Wicklen, 1974), probit and logit (Cole & Gunther, 1998; Cole & White, 2012; Espahbodi, 1991; Estrella, Park, & Peristiani, 2000; Hanweck, 1977; Martin, 1977; Ohlson, 1980; Thomson, 1991), and Cox proportional hazards models

* Corresponding author.
E-mail address: aagrapet@econ.duth.gr (A. Agrapetidou).

https://doi.org/10.1016/j.ijforecast.2018.01.009
0169-2070/

(Lane, Looney, & Wansley, 1986; Shumway, 2001; Whalen, 1991; Wheelock & Wilson, 2000). Recently, several studies have also investigated the use of machine learning based techniques for this purpose.[1]

Boyacioglu, Kara, and Baykan (2009) compared logistic regression, MDA, k-means cluster analysis (CA), support vector machine (SVM) and four neural network (NN) architectures in forecasting Turkish bank failures for the period 1997–2003. They used a very small sample with 21 failed banks, 44 solvent banks and 20 explanatory variables. Although NN yielded the best results, SVM outperformed the majority of the other techniques, with a forecasting accuracy of 90.90%. In a similar setup, Ecer (2013) also compared the performances of NN and SVM for forecasting bank failures for the period 1994–2001. The dataset consisted of a small sample of 34 Turkish banks and 36 financial ratios. Again, NN yielded the optimum out-of-sample result (97.06%). Along the same lines, Erdogan (2013) applied SVM to the forecasting of Turkish bank failures in the period 1997–2003. The dataset consisted of 42 Turkish commercial banks, of which 18 had failed and 24 were solvent. He achieved a forecasting accuracy of approximately 95% using 19 financial ratios.

In the relevant literature, most studies rely on accounting data that are augmented in some cases with macroeconomic and market-based variables (Agarwal & Taffler, 2008; Berger & Bouwman, 2013; Cole & Gunther, 1998; Curry, Elmer, & Fissel, 2007; Espahbodi, 1991; Kolari, Glennon, Shin, & Caputo, 2002; Männasoo & Mayes, 2009; Martin, 1977; Meyer & Pifer, 1970; Thomson, 1991). The most commonly used variables are those based on CAMELS[2] indicators. These include measures of capital adequacy, asset quality, management, earnings, liquidity and sensitivity to market risk. Regarding non-financial firms, the primary variables used for predicting bankruptcies are those based on the prediction models of Altman et al. (1977) and Ohlson (1980). The strong performance of Ohlson's model that researchers have demonstrated over time supports its effectiveness (Begley, Ming, & Watts, 1996; Grice & Dugan, 2003; Hillegeist, Keating, Cram, & Lundstedt, 2004; Karamzadeh, 2013).

Due to the large number of bank failures in the recent crisis in the U.S., numerous studies since have aimed to forecast the insolvency of financial institutions. Jordan, Rice, Sanchez, Walker, and Wort (2010) examined 225 failed banks and 225 solvent ones for the period 2007–2010. They performed discriminant analysis and achieved an out-of-sample forecasting accuracy of 78.10%. Mayes and Stremmel (2014) examined a large number of 16,188 U.S. banks from 1992 to 2012 with the aim of forecasting bank failures during the period 2008–2012. They employed a logit model and achieved an out-of-sample forecasting accuracy of 83%. Papadimitriou, Gogas, Plakandaras, and Mourmouris (2013) used a sample of 300 U.S. banks for forecasting U.S bank failures with SVM, and obtained an out-of-sample forecasting accuracy of 76.40% using only six input variables that refer to banks' efficiency, leverage and market appreciation in terms of goodwill and other intangibles. Iturriaga and Sanz (2015) compared NN and SVM for forecasting U.S. bank failures over the period 2002–2012. Their dataset consisted of 386 failed U.S. banks and 386 solvent ones. They found the NN to outperform the SVM for the short-term (one year) horizon, but the SVM to outperform the NN for the medium and long-term forecasting horizons (two and three years before failure). Their optimum one-year horizon model achieved a forecasting accuracy of 94.23% with NN. The two- and three-year horizon models produced forecasting accuracies of 86.54% and 82.69% via SVM. According to their study, the most important variables are provisions, risk concentration on the construction industry and equity support to loans. Cleary and Hebb (2016) examined 323 banks that failed in the U.S. over the period 2002–2011 and an equal sample of non-failed ones. They used discriminant analysis and variables related to bank capital, loan quality and profitability to forecast bank failures in out-of-sample data, and achieved a forecasting accuracy of 89.50%.

This paper presents an SVM-based methodology for forecasting the bankruptcy of U.S. financial institutions over the period 2007–2013 using financial data taken from the banks' publicly-available financial statements. The proposed approach includes a two-step feature selection process that is used to find the most relevant variables for the identification of soon-to-fail banks. These variables are then fed into an SVM model that has been optimized through a training and testing procedure. This study introduces three innovations. First, in contrast to the relevant literature, which uses one predetermined cut-off level in the SVM decision function, here we identify the optimum cut-off level out of several alternatives. Our results show that the optimum cut-off level is different from the standard one used in the literature. Second, even though the proposed model deals with a binary classification problem ("solvent" or "failed"), the corresponding sensitivity analysis offers a quantitative tool for measuring the confidence of our forecast. This can also be extended to an alternative stress-testing tool: for each explanatory variable (or a combination of them), we can measure the change that would be necessary in order to reclassify the bank from "solvent" to "failed" or vice versa. Thus, this procedure provides us with a sensitivity analysis of the resulting classification. Finally, the third innovation has to do with the sample size and is two-fold: (a) we attempt to construct a forecasting model for all U.S. bank failures for the period 2007–2013, and (b) we employ a realistic ratio of 1:10 of failed to solvent banks in the out-of-sample data, instead of the 1:1 that has been used in most previous studies.

[1] Foran extensive review of forecasting banks' bankruptcy via statistical and intelligent techniques, see Kumar and Ravi (2007), Demyanyk and Hasan (2010) and Chen, Ribeiro, and Chen (2016).

[2] CAMELS is a rating system that was introduced in 1979 by U.S. regulators for assessing the financial condition of banks by assigning ratings from 1 (strong) to 5 (weak).

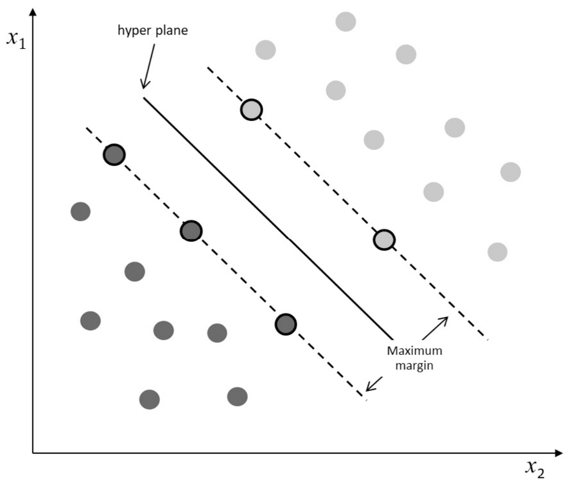

Fig. 1. Support vectors lie on the margin lines, which are represented by dashed lines and define the hyperplane. The hyperplane is represented by the continuous line and separates the data into two classes.

The rest of the paper is organized as follows: Section 2 presents the selected dataset. Section 3 provides a description of the proposed methodology. Section 4 analyzes the empirical forecasting results and presents our novel stress-testing method, along with various intervention scenarios that may be necessary in order to avoid a potential bankruptcy. Section 5 provides a discussion of some of the findings with respect to their supervisory and policy implications, as well as of possible future research based on this study. Finally, Section 6 concludes the paper.

2. The data

We collected data for all failed U.S. banks from 2007 to 2013. Our dataset includes all types of banking institutions, such as commercial banks, savings associations and national banks (as they are categorized by the FDIC). The period of interest coincides with the financial crisis, which led to a substantial number of bank failures. A total of 492 U.S. banks failed during the selected period, but we restrict our sample here to banks that have data available for a minimum of four years prior to their collapse. For example, if a bank failed in 2007, we collected data from its financial statements that pertained to the years 2006, 2005, 2004 and 2003. There were 11 banks that had been active for under four years prior to their collapse, and these were therefore excluded from our sample, leaving us with a total of 481 failed banks. Table 1 reports the total numbers of failed banks in the U.S. by year, as well the numbers that we included in our sample.

Next, for each failed bank in our sample we randomly added four-year windows of data for two solvent ones. Thus, our final dataset includes 481 failed and 962 solvent banks, giving a total of 1443. For each bank, four observations of 36 variables and financial ratios were collected from their financial statements over a period of four years, resulting in a total of 144 observations per bank. The initial 36 variables are reported in Table 2, with all of these variables being available publicly from the FDIC.

All of the balance sheet variables are expressed as ratios to total assets (except of course for total assets). Similarly, all of the income and expense variables are expressed as ratios of the total interest income.

3. Methodology

The support vector machine (Cortes & Vapnik, 1995) is a supervised machine learning methodology based on the statistical learning theory that was created for classification and regression tasks. This study employs SVM for classification in constructing a forecasting model for bank failures. The basic idea of the support vector classification is to separate the data into two classes by finding a hyperplane that maximizes the distance between the two classes. The hyperplane is defined by only a small number of data points, which are called support vectors (SV) and are derived through a minimization procedure.

The model includes two basic steps: training and testing. The majority of the dataset is used in the training process, where the separating hyperplane is defined. In the testing step, the model's generalization ability is evaluated by calculating the model's performance on the remaining subset, a small part of the dataset that was set aside during the first step. In what follows, we briefly present the mathematical derivation of the SVM theory.

3.1. Linearly separable case

In a linearly separable case, given a training dataset $\boldsymbol{x}_i \in \mathbb{R}^n$ and the corresponding label $y_i \in [+1, -1]$, $(i = 1, \ldots, N)$, a separator can be defined as

$$f(\boldsymbol{x}) = \mathbf{w}^T \boldsymbol{x} - b = 0, \tag{1}$$

satisfying the constraints:

$$\begin{aligned} \mathbf{x}_i^T \mathbf{w} - b > 0 \quad & \text{for } y_i = +1 \\ \mathbf{x}_i^T \mathbf{w} - b < 0 \quad & \text{for } y_i = -1, \end{aligned} \tag{2}$$

such that $y_i f(\mathbf{x_i}) > 0\ \forall_i$, where $\mathbf{w}$ is the parameter vector and b is the bias. The optimal hyperplane is defined as the decision boundary that classifies each data vector into the correct class and has the maximum margin from both classes. In Fig. 1, the SVs are represented by the pronounced contour, the margin lines by the dashed lines and the hyperplane by the continuous line.

The solution to the problem of finding the hyperplane can be determined using the Lagrange relaxation procedure in the following equation:

$$\min_{\mathbf{w},b} \max_{\mathbf{a}} \left(\frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^{N} a_i \left[y_i \left(\mathbf{w}^T \mathbf{x}_i - b \right) - 1 \right] \right), \tag{3}$$

where $\mathbf{a} = [a_1, \ldots, a_N]$ are the non-negative Lagrange multipliers. However, Eq. (3) is never used to estimate the solution. Instead, we always solve the dual problem, defined as:

$$\max_{\mathbf{a}} \left\{ \sum_{i=1}^{N} a_i - \sum_{j=1}^{N} \sum_{k=1}^{N} a_j a_k y_j y_k \mathbf{x}_j^T \mathbf{x}_k \right\}, \tag{4}$$

Table 1
Numbers of failed U.S. banks included in the dataset.

Year	Number of failed banks	Number of failed banks in our dataset
2007	3	2
2008	25	22
2009	140	135
2010	157	155
2011	92	92
2012	51	51
2013	24	24
Total	492	481

Table 2
The 36 initial variables collected for each bank.

Source	Variable	Description
A/L	TASSET	Total assets
A/L	CASH	Cash and due from depository institutions/total assets
A/L	NLL	Net loans and leases/total assets
A/L	LLA	Loan loss allowance/total assets
A/L	GOI	Goodwill and other intangibles/total assets
A/L	TD	Total deposits/total assets
A/L	IBD	Interest-bearing deposits/total assets
A/L	SD	Subordinated debt/total assets
A/L	AA	Average assets year-to-date/total assets
A/L	VL	Volatile liabilities/total assets
A/L	ULC	Unused loan commitments/total assets
A/L	T1CRC	Tier 1 (core) risk-based capital/total assets
A/L	T2RBC	Tier 2 risk-based capital/total assets
A/L	TUC	Total unused commitments/total assets
I/E	TIE	Total interest expense/total interest income
I/E	PLLL	Provision for loan and lease losses/total interest income
I/E	TNI	Total noninterest income/total interest income
I/E	TRACC	Trading account gains & fees/total interest income
I/E	ANI	Additional noninterest income/total interest income
I/E	TNE	Total noninterest expense/total interest income
I/E	SAL	Salaries and employee benefits/total interest income
I/E	PTNOI	Pre-tax net operating income/total interest income
I/E	SEC	Securities gains (losses)/total interest income
I/E	NIA	Net income attributable to bank/total interest income
I/E	DIVDS	Cash dividends/total interest income
I/E	NOI	Net operating income/total interest income
P/R	YOEA	Yield on earning assets
P/R	NIM	Net interest margin
P/R	ROA	Return on assets
P/R	ROE	Return on equity
P/R	ASSPE	Assets per employee ($ millions)
C/R	NLLTD	Net loans and leases to deposits
C/R	EQCTA	Equity capital to assets
C/R	LEV	Core capital (leverage) ratio
C/R	T1RBC	Tier 1 risk-based capital ratio
C/R	TRBCR	Total risk-based capital ratio

Note: The source of the variable is classified as A/L when it comes from assets and liabilities and I/E when it comes from the income and expense statement. P/R and C/R stand for the performance and condition ratios respectively.

subject to $\sum_{i=1}^{N} a_i y_i = 0$ and $0 \leq a_i, \forall_i$. The solution to Eq. (4) gives the location of the separating hyperplane, defined by:

$$\hat{\mathbf{w}} = \sum_{i=1}^{N} a_i y_i \mathbf{x}_i, \tag{5}$$

$$\hat{b} = \hat{\mathbf{w}}^T \mathbf{x}_i - y_i, i \in V, \tag{6}$$

where $V = \{i : 0 < a_i\}$ is the set of the support vector indices.

3.2. Non-linearly separable case

3.2.1. Error-tolerant SVM

The above methodology works for perfectly separable cases. However, datasets are often contaminated with noise and outliers. In that case, the position of the separating hyperplane could be affected by these un-fitted points. This problem is solved by using the error-tolerant model that was introduced by Cortes and Vapnik (1995). The basic idea involves the introduction of non-negative slack variables $\xi_i \geq 0, \forall_i$, controlled through a penalty parameter

C, for treating, erroneously classified data points. Eq. (4) is now defined as:

$$\min_{\mathbf{w},b,\boldsymbol{\xi}} \max_{\mathbf{a},\boldsymbol{\mu}} \left\{ \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i - \sum_{j=1}^{N} a_j\left[y_i\left(\mathbf{w}^T\mathbf{x}_j - b\right) - 1 + \xi_j\right] - \sum_{k=1}^{N}\mu_k\xi_k \right\}, \tag{7}$$

where ξ_i measures the distance of vector $\mathbf{x}_i$ from the hyperplane when classified erroneously, and $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_N]$ are the Lagrange multipliers.The optimal separating hyperplane is defined as:

$$\hat{\mathbf{w}} = \sum_{i=1}^{N} a_i y_i \mathbf{x}_i, \tag{8}$$

$$\hat{b} = \hat{\mathbf{w}}^T\mathbf{x}_i - y_i,\ i \in V, \tag{9}$$

where $V = \{i : 0 < a_i < C\}$ is the set of support vector indices.

3.2.2. Kernel methods

The underlying data generating processes of real life phenomena are often non-linear. Thus, linear classifiers underperform in these cases. In order to tackle this drawback, the SVM is coupled with the kernel functions. The so-called "kernel trick" follows the projection idea, presented in Fig. 2, while ensuring a minimum computational cost: the dataset is mapped in an inner product space, where the projection is performed using only dot products within the original space through "special" kernel functions, instead of computing the mapping of each data point explicitly. When the kernel function is non-linear, the SVM model produced is non-linear as well.

The solution to the dual problem with the projection of Eq. (4) now transforms to:

$$\max_{a} \sum_{i=1}^{N} a_i - \frac{1}{2}\sum_{j=1}^{N}\sum_{k=1}^{N} a_j a_k y_j y_k K\left(\mathbf{x}_j, \mathbf{x}_k\right), \tag{10}$$

under the constraints: $\sum_{i=1}^{N} a_i y_i = 0$ and $0 \le a_i \le C, \forall_i$, where $K\left(\mathbf{x}_j, \mathbf{x}_k\right)$ is the kernel function. Then, the rule for forecasting the label of a testing data point $\mathbf{x}$ is given as follows:

$$f(x) = sign\left\{\sum_{i=1}^{N} a_i y_i K\left(\mathbf{x}, \mathbf{x}_i\right) + b\right\}. \tag{11}$$

If $f(x) > 0$ (the cut-off level is zero), the data is classified as belonging to class $+1$; otherwise, it is in class -1. This study tested the shift of the separation hyperplane and tried alternative cut-off levels ranging from 1 to -1, in order to evaluate the performance of the modified model.

We performed this scheme for two different kernels: the linear kernel and the non-linear radial basis function (RBF). The mathematical representations of is the two are as follows:

- The linear:

$$K\left(\mathbf{x}_i, \mathbf{x}_j\right) = \mathbf{x}_i^T\mathbf{x}_j + c \tag{12}$$

- The radial basis function:

$$K\left(\mathbf{x}_i, \mathbf{x}_j\right) = \exp\left(-\gamma\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2\right), \tag{13}$$

with c and γ representing the hyperplane parameters that need to be optimized.

3.3. Over-fitting

It is possible that during the training step the model may "*learn*" to forecast accurately the outcome of the specific training data, instead of the general phenomenon under study. This situation is called *over-fitting*, and yields models with high levels of forecasting accuracy on the training samples, but significantly worse performances on the testing samples. We avoid the problem of over-fitting by training our model in a cross-validation framework: the training dataset is divided into n equal sized 'folds' (parts). Then, for each given pair of parameters c and γ, $n-1$ of them are used for training the model and the remaining part is used to evaluate the forecasting accuracy of the trained model. This procedure is repeated n times with the same set of parameters until all folds have passed through the testing procedure. The overall performance of the model with the given pair of parameters c and γ is evaluated by averaging its forecasting accuracy over the n testing folds. An example of a three-fold cross-validation scheme is represented in Fig. 3.

The empirical part of this study used a 10-fold cross-validation. Finally, the out-of-sample subset that was kept aside during training and did not participate in the cross-validation procedure was used to evaluate the model's generalization ability (out-of-sample forecasting).

3.4. Weights

Our unbalanced dataset contains twice as many solvent banks as failed ones. We restore the balance in the classification process by assigning inversely proportional weights to the misclassification cost of the two instances: one for the solvent cases and two for the insolvent ones. This means that the penalty for a misclassified insolvent bank in the minimization procedure inside the core of the SVM training is double that imposed on a misclassified solvent bank.

4. Empirical results

4.1. Feature selection

Our dataset consists of 36 variables for each bank, collected for a four-year time window, yielding a dataset of 144 observations. We identified the most informative variables for our forecasting model by employing a two-step feature selection procedure. The first step uses a local-learning-based feature selection scheme to filter the initial variable set, then the resulting set is refined scrupulously in an iteratively shrinking procedure.

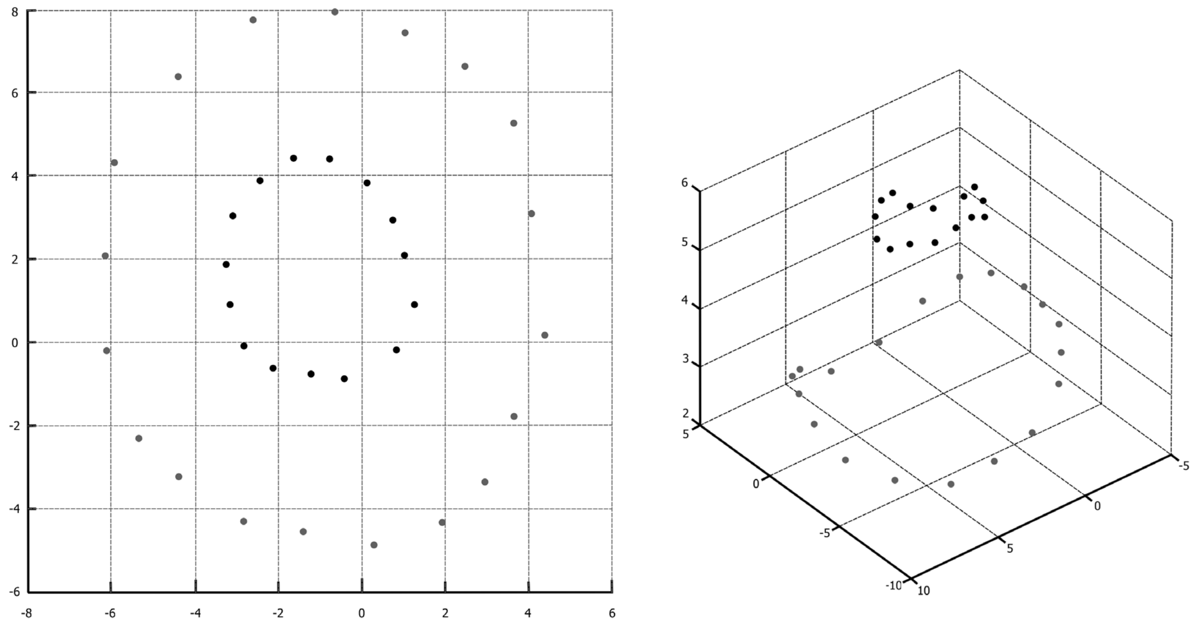

Fig. 2. The data cannot be separated by a hyperplane in the two-dimensional space (left). The use of a kernel function renders the data linearly separable in a higher dimensional feature space (right).

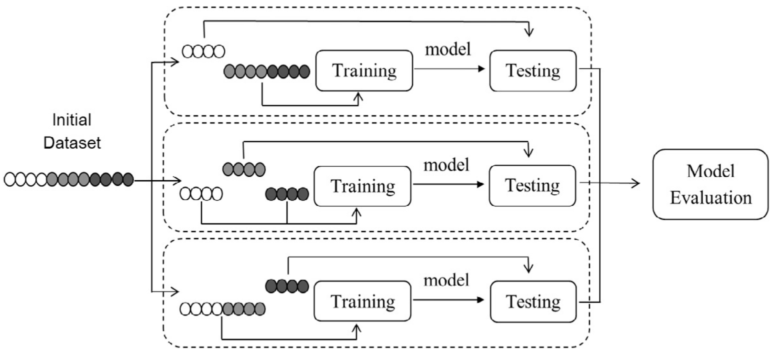

Fig. 3. A three-fold cross-validation example.

(a) The first step of our feature selection procedure is based on the local learning method for high dimensional data analysis that was proposed by Sun, Todorovic, and Goodison (2010).[3] The main goal of this algorithm is to simultaneously (a) decompose a non-linear problem into a set of locally linear ones through local learning and (b) identify the variables (features) that best represent the decomposed linear problems. The method calculates a relevance index between the explanatory variables and the dependent one, which in our case is the financial position of the bank (failed or solvent). Relevant variables accumulate positive weights, while irrelevant ones retain near-zero values. Only six out of the 144 observations in our dataset accumulated positive weights. According to their relevance index, these are:

(1) Tier 1 (core) risk-based capital/total assets (T1CRC $(t-1)$)

(2) Provision for loan and lease losses/total interest income (PLLL $(t-1)$)

(3) Loan loss allowance/total assets (LLA $(t-1)$)

(4) Total interest expense/total interest income (TIE $(t-1)$)

(5) Equity capital to assets (EQCTA $(t-1)$)

[3] For a detailed analysis, the interested reader can refer to the paper by Sun et al. (2010).

(6) Volatile liabilities/total assets (VL ($t-3$)).

As expected, the most relevant variables come from a bank's most recent financial statement, i.e., year $t-1$, with the exception of VL, which comes from year $t-3$.

(b) In the second step, the six selected variables were fed into the SVM-based forecasting model,[4] so as to refine the input set further using a shrinking procedure. At each iteration of this procedure, we compare the performance of the base variable set (the optimum set so far) with those of all sets generated by removing one variable from the base set. For example, if our base set includes n variables, we compare it with every possible set of $n-1$ variables created using the base set. If the base set does not perform the best, the $n-1$ sized set that yields the maximum improvement over the base variable set then replaces the base variable set in the next iteration. The procedure continues sequentially until no improvement can be achieved by shrinking the base set.

In our tests, the initial base set consists of the six variables identified in step (a) of our feature selection procedure. For every variable set that we tested, we exhaustively tried alternative cut-off levels for the decision function. Following this procedure, we ended up with an input set of only two explanatory variables for the linear kernel (T1CRC and TIE) and four explanatory variables with the RBF kernel (T1CRC, PLLL, TIE and EQCTA).

4.2. *Forecasting bank failures*

For the SVM approach, usually 70%–80% of the dataset is used for training the model and the rest is used for out-of-sample forecasting. Our study used a total of 1100 banks to train and cross-validate the model (≈76%), while the remaining 343 banks were set aside to be used to test the generalization ability of the optimum model to unknown data (out-of-sample forecasting). In both samples, the ratio of failed to solvent cases was maintained at 1:2.

When forecasting banking bankruptcy, two types of forecasting error may occur: (a) classifying a failed bank as solvent (Type II error) and (b) classifying a solvent bank as failed (Type I error). The optimum model will be the one that minimizes the sum over both types of forecast errors. The best forecasting result for the SVM-linear model is achieved when the cut-off is set to –0.7,[5] with an out-of-sample forecasting accuracy of 98.25%. Table 3 decomposes the out-of-sample forecasting results for solvent and insolvent banks. The forecasting accuracy for the solvent banks is 98.68% (225 out of 228) and that for the insolvent ones is 97.39% (112 out of 115). Thus, the model misclassifies only three banks in each case. With the RBF kernel, the best result is achieved when the cut-off level is set to –1, with an overall out-of-sample forecasting accuracy of 97.96%. The forecasting accuracy for the solvent banks is again 98.68% (225 out of 228), while that for the failed banks is 96.52% (111 out of 115). Thus, with the RBF kernel, three solvent banks were misclassified as insolvent and four insolvent banks were misclassified as solvent. Our optimum forecasting model is achieved using the linear kernel with an overall classification error of 1.75%; that for the RBF kernel is 2.04%. We will not address the RBF kernel model in the remainder of the paper, but instead will focus on the linear kernel, which produces slightly more accurate results. The two variables that provided the highest forecasting accuracy (98.25%) are: (1) Tier 1 (core) risk-based capital over total assets (T1CRC), and (2) total interest expense over total interest income (TIE). Both come from the most recent statement ($t-1$). T1CRC is a measure of capital adequacy, and is used as a cushion against unanticipated losses that might result from economic tribulations. Banks with higher T1CRC ratios are associated with lower probabilities of default. TIE is a ratio that represents the interest paid on any type of borrowing over the interest earned on any type of lending. It is evident that a rise in this ratio leads to a deterioration in the gross profit margin. Most importantly, though, with respect to the forecasting solvency that is our main concern here, it can be translated as an indicator of increased leverage (debt-to-own-funds) that produces a higher risk of bank failure. When a bank leverages an augmented balance sheet, the interest expense increases.

Moreover, we augmented our out-of-sample data with additional solvent banks (we have already included all failed banks), in order to test the robustness of our model in a more realistic situation, so that the ratio of failed to solvent banks is closer to that actually observed. Thus, the out-of-sample data of the augmented sample contains a ratio of 1:10 of failed to solvent banks. More specifically, the out-of-sample dataset consists of 115 failed banks (as previously) and 1164 solvent ones. The trained model is then used to forecast the financial position (solvent or failed) of the banks in the new out-of-sample dataset. The results are presented in Tables 4 and 5. This is a very interesting measure of the robustness of our empirical findings.

Table 3
Comparison of the out-of-sample forecasting results to true cases.

SVM-linear	True	
	Solvent	Insolvent
Classified solvent	225	3
Classified insolvent	3	112

Table 4
Out-of-sample forecasting accuracy of the SVM-linear augmented model.

Overall accuracy	Insolvent accuracy	Solvent accuracy
99.22%	97.39%	99.40%

Table 5
Performance of the proposed model on the out-of-sample augmented dataset.

SVM-linear	True	
	Solvent	Insolvent
Classified solvent	1157	3
Classified insolvent	7	112

[4] All variables were scaled to [−1, 1] before performing the tests.

[5] The results for different cut-off-levels of the SVM model with the two explanatory variables can be found in Appendix A.

The optimal forecasting model is still the one identified previously using the 1:2 ratio of failed to solvent banks in the training step – only the out-of-sample ratio is changed to 1:10.

The out-of-sample forecasting accuracy for the solvent banks is 99.40% (1157 out of 1164), while that for the insolvent banks is of course the same as before (97.37%). The overall out-of-sample forecasting accuracy of the model is 99.22% (1269 out of 1279). Thus, our optimum forecasting model produced a total of ten errors: seven solvent banks were misclassified as insolvent (false alarms) and three insolvent banks were misclassified as solvent.

As the number of misclassified insolvent banks is only three, we can take a closer look at these institutions and their idiosyncratic behaviors and characteristics. These are: (a) The First National Bank of Davis, Davis, Oklahoma; (b) The First State Bank, Camargo, Oklahoma; and (c) Glasgow Savings Bank, Glasgow, Missouri. The first two failed in 2011 and the third in 2012. Looking closely at the regulatory actions that took place at that period, we find that:

(a) The First National Bank of Davis was a small bank with \$90.2 million in total assets. There were no outstanding regulatory enforcement actions from the FDIC. According to the review by the Department of the Treasury, the bank was well capitalized prior to 2011. However, significant unrecognized losses were discovered during the 2011 examination that exceeded the bank's capital and allowance for loan and lease losses, and on March 11, 2011, the bank was acquired by The Pauls Valley National Bank, Pauls Valley, Oklahoma, with government assistance. The First National Bank of Davis was the 24th FDIC-insured institution to fail in 2011, and the second in Oklahoma.

(b) The first FDIC-insured institution to fail in 2011 in Oklahoma was The First State Bank. The bank received an enforcement action on 21 December, 2010, and very shortly after, on the 28 January, 2011, closed and was acquired by Bank 7, Oklahoma City, Oklahoma, with government assistance. This was an even smaller bank, with only \$43.5 million of total assets.

(c) Glasgow Savings Bank is an inactive institution as of July 13, 2012. At that time, the bank had approximately \$24.8 million in total assets. It was acquired by the Regional Missouri Bank with government assistance. The estimated cost to the FDIC was \$0.1 million.

Moreover, our optimum model misclassified seven solvent banks as insolvent, namely (i) the First Chatham Bank, Savannah, Georgia; (ii) the Americana Community Bank, Sleepy Eye, Minnesota; (iii) the First Cornerstone Bank, King of Prussia, Pennsylvania; (iv) the Heritage Bank, Hinesville, Georgia; (v) the Brand Banking Company, Lawrenceville, Georgia; (vi) the State Central Bank, Bonaparte, Iowa; and (vii) Eagle Valley Bank, National Association, St. Croix Falls, Wisconsin. In this case, the misclassification may be related to the fact that all of these banks received enforcement actions from either FDIC or OCC (Office of Comptroller of the Currency). In addition, the holding company of the Heritage Bank received financial help from the U.S. Treasury as part of the Capital Purchase Program under the Troubled Asset Relief Program (TARP). Thus, unknown to our model, these banks that it misclassified were the recipients of specific regulatory actions in attempts to prevent them from failing.

To summarize the above, our optimum model, with a 99.22% overall forecasting accuracy and only two explanatory variables, classified three insolvent banks as solvent and seven solvent banks as insolvent. In other words, the overall classification error of the model is 0.78% (10 out of 1279). According to the business (mergers and acquisitions) and regulatory (actions from the FDIC and the OCC) history uncovered for these banks, it is very probable that the model's forecasting accuracy would have been higher had it not been for external intervention.

4.3. Longer forecasting horizons

This section explores the proposed framework at longer forecasting horizons. We follow the same procedure as was described in Section 4.2: we use 1100 banks to train and cross-validate the model and the remaining 343 banks to test the model's generalization ability on the out-of-sample data for two- and three-year horizon forecasting. When forecasting bank failures for a two-year-ahead forecasting window, we use a reduced input set of six variables (see Section 4.1) that come from the banks' financial statements in year $t-2$, except for the variable volatile liabilities, which comes from the year $t-4$. In the same way, when forecasting bank failures for three-year horizon forecasting, we use the six variables that come from the banks' financial statements in year $t-3$, except for the variable volatile liabilities, which comes from the year $t-5$. The final input set is defined using the variable set shrinking methodology defined in Section 4.1.

Under this scheme, we end up with an input set of four explanatory variables for the two-year-ahead forecasting horizon:

(1) Tier 1 (core) risk-based capital/total assets ($t-2$),
(2) Provision for loan and lease losses/total interest income ($t-2$),
(3) Loan loss allowance/total assets ($t-2$) and
(4) Volatile liabilities/total assets ($t-4$).

The optimum model for the three-year-ahead forecasting horizon ($t-3$) employs three variables:

(1) Tier 1 (core) risk-based capital/total assets ($t-3$),
(2) Loan loss allowance/total assets ($t-3$) and
(3) Total interest expense/total interest income ($t-3$).

Table 6
Out-of-sample forecasting accuracy of the SVM-linear model – two – and three-year-ahead forecasting.

SVM-linear model	True	
	Solvent	Insolvent
Panel A: Two-year-ahead forecasting		
Classified solvent	202	18
Classified insolvent	26	97
Panel B: Three-year-ahead forecasting		
Classified solvent	185	27
Classified insolvent	43	88

Table 6 summarizes the forecasting accuracy of the SVM-linear model over the two- and three-year-ahead forecasting windows in Panels A and B respectively.

The overall two-year horizon forecasting accuracy of the model is 87.17%. The forecasting accuracy for the solvent banks is 88.60% (202 out of 228), while that for the failed banks is 84.35% (97 out of 115). The overall accuracy of the model for the three-year-ahead forecasting horizon is 79.59%. The forecasting accuracy for the solvent banks is 81.14% (185 out of 228), while that for the failed ones is 76.52% (88 out of 115). As expected, the forecasting accuracy declines as the forecasting horizon expands.

4.4. Comparison

This section compares our results to those that we get from the well-established Ohlson's score (Ohlson, 1980). Ohlson used a multi-factor financial formula to develop a statistical bankruptcy predictor that is often called the O-score and is calculated using the formula:

$$O = -1.32 - 0.407AS + 6.03LM - 1.43WCM + 0.0757ICR - 1.72DCLM - 2.37RoA - 1.83FtDR + 0.285DCRA - 0.521CiNI,$$

where

- AS is the adjusted size of the bank, calculated as
$$AS = \log \frac{Total\ Assets}{GNP\ price\ level\ index}, \text{ and}$$
$$GNP\ price\ level\ index = 100\frac{Nominal\ GNP}{Real\ GNP}$$
- LM is the leverage measure, calculated as
$$LM = \frac{Total\ Liabilities}{Total\ Assets}$$
- WCM is the working capital measure, calculated as
$$WCM = \frac{Working\ Capital}{Total\ Assets}$$
- ICR is the inverse current ratio, calculated as
$$ICR = \frac{Current\ Liabilities}{Current\ Assets}$$
- $DCLM$ is the discontinuity correction for leverage measure, calculated as
$$DCLM = \begin{cases} 1 \text{ if } Total\ Liabilities > Total\ Assets \\ \quad 0 \text{ otherwise} \end{cases}$$

Table 7
Performance of the O-score on the out-of-sample enhanced dataset.

	True	
	Solvent	Insolvent
Classified solvent	897	1
Classified insolvent	267	114

- RoA is the return on assets, calculated as
$$RoA = \frac{Net\ Income}{Total\ Assets}$$
- $FtDR$ is the funds to debt ratio, calculated as
$$FtDR = \frac{Funds\ from\ Operations}{Total\ Liabilities}$$
- $DCRA$ is the discontinuity correction for return on assets, calculated as
$$DCRA = \begin{cases} 1 \text{ if net loss for the last two years} \\ \quad 0 \text{ otherwise} \end{cases}$$
- $CiNI$ is the change in net income, calculated as $CiNI = \frac{NI_t - NI_{t-1}}{|NI_t| - |NI_{t-1}|}$.

Bankruptcy is forecasted using the probability $P = \frac{e^O}{1-e^O}$, where $P > 0.5$ means that the bank is predicted to fail and $P < 0.5$ describes a solvent bank.

The results on the application of the O-score to our dataset can be found in Table 7.

The O-score outperforms the SVM model in forecasting insolvency: it forecasts 114 of the 115 cases correctly, while the SVM model 112. However, this accuracy comes with a significantly higher number of false alarms (more than 26 times as high): the O-score misclassifies 267 solvent cases as insolvent, in comparison only 10 false alarms for the SVM model. The overall accuracy of the O-score is 79.05%, while the SVM model's overall accuracy is 99.22%.[6]

4.5. Stress testing tool

Thus far in Section 4, we have (a) identified the most relevant variables for forecasting bank failures and (b) created the optimum SVM-based forecasting model for forecasting the failure of U.S. financial institutions. The final variable set was selected using a two-step feature selection procedure. First, the local-learning feature selection scheme, introduced by Sun et al. (2010), identified the six forecasting variables that were "relevant" to insolvency out of the 144 initial variables. Then, the sequential shrinkage method introduced in Section 4.1 identified the two explanatory

[6] When we stressed the SVM-model by shifting the linear separator so as to minimize the misclassified failed banks (which is equivalent to choosing a different cut-off level during the training), we achieved the same performance as the O-score, while obtaining only 50 false alarms in the enhanced model (compared to the 267 of the O-score). These results can be found in Appendix B. We also performed logit and probit regression analyses in order to compare the results. The logit model produced the same accuracy as SVM-linear. The results are available upon request.

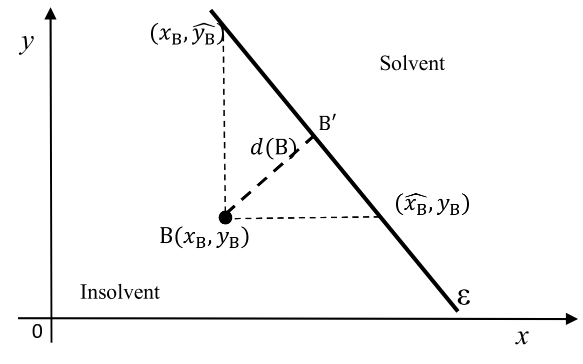

Fig. 4. Illustration of the stress-testing tool of a forecasted bank failure.

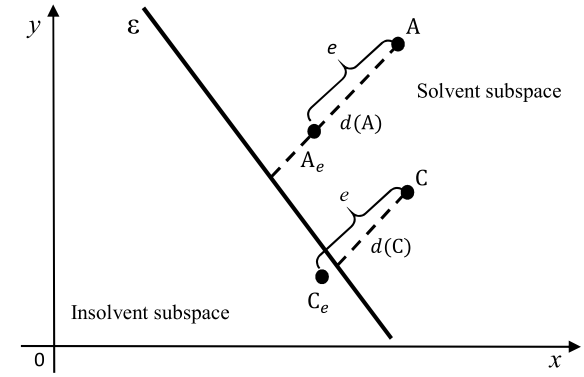

Fig. 5. Illustration of the stress-testing tool for a bank forecasted to be healthy.

variables of the optimum model. The optimum forecasting model is a linear decision boundary that separates the solvent from the failed banks with 99.22% accuracy using two explanatory variables (T1CRC and TIE), and is defined as:

$$\varepsilon : 23.45 \times \text{T1CRC} - 3.85 \times \text{TIE} + 10.55 = -0.7.$$

In what follows, we will perform a sensitivity analysis of our model in an attempt to answer two important questions: (i) given these forecasts, what kind of intervention is necessary to avoid a forecasted bank failure, and is this intervention quantifiable? (ii) how safe is a financial institution that is forecasted as healthy?

To answer these questions, we must first provide the geometric interpretation of our approach. The best forecasting model that we produced was equipped with the linear kernel fed with two explanatory variables. It is easy to visualize the classification space in this setup, as it has only two dimensions: the banking institutions are represented by points on the two-dimensional data space and the separating line is a function of the two explanatory variables. Fig. 4 provides an example of the stress-testing tool. The x-axis represents tier 1 (core) risk-based capital over total assets (T1CRC) and the y-axis the total interest expense over total interest income (TIE).

We depict the separating line $\varepsilon : \alpha x + \beta y + \gamma = 0$ that separates the two-dimensional space into two subspaces: solvent and insolvent banks, respectively. Thus, the separating line is the decision boundary between the banks that were forecasted healthy and those that were forecasted failed. It is trivial to calculate the analytical form of the separating line in the linear kernel case.[7]

Let us consider a bank (point B, with coordinates (x_B, y_B)) that lies in the insolvency subspace (so that the bank is forecasted to fail). The distance between point B and the separator is

$$d\,(B) = \frac{|\alpha x_B + \beta y_B + \gamma|}{\sqrt{\alpha^2 + \beta^2}}, \tag{14}$$

and corresponds to the distance between failure and solvency. We remind the reader here that the two axes represent real banking variables from the banks' annual publicly available financial statements scaled to $[-1, 1]$. Thus, any successful intervention in the financial status of the bank in question will result in the displacement of its projection to the other side of the separating line: the solvency subspace. The minimum possible changes[8] will result in a displacement by $d\,(B)$ perpendicularly to the separator (in the position $B' = (x_B - \alpha\frac{\alpha x_B+\beta y_B+\gamma}{\alpha^2+\beta^2}, y_B - \beta\frac{\alpha x_B+\beta y_B+\gamma}{\alpha^2+\beta^2}))$[9]; this is the closest "solvent point" to B, and the minimum changes can be calculated as the difference between the coordinates of point B and $B'(\alpha\frac{\alpha x_B+\beta y_B+\gamma}{\alpha^2+\beta^2}, \beta\frac{\alpha x_B+\beta y_B+\gamma}{\alpha^2+\beta^2})$.

If we restrict changes to occur only along the x-axis, the second coordinate y_B will remain constant and only the values x_B can be altered. In this scenario, the first point with a solvent forecast is the point $(\widehat{x_B}, y_B)$ that lies in the separator ε. This means that the following is true: $\alpha\widehat{x_B} + \beta y_B + \gamma = 0$ and $\widehat{x_B} = -(\beta y_B + \gamma)/\alpha$. Similarly, if we consider changes only on the y axis, bank B will be solvent if $\widehat{y_B} = -(\alpha x_B + \gamma)/\beta$.

The proposed methodology is able not only to forecast the solvency or insolvency of a banking institution, but also to quantify the minimum changes in the financial status of the bank that will change its forecasted position under various scenarios.

In Fig. 5, banks A and C are both forecasted as healthy, though $d\,(A) > d\,(C)$. We will use a simple example to show that this means that bank A is safer than bank C. Let

7 In fact, $\boldsymbol{w}$ from Eq. (8) is $[\alpha, \beta]$, and γ is calculated in Eq. (9).

8 This is the minimum distance because it was taken perpendicular to the separator.

9 The calculations are trivial but outside the scope of this paper.

us assume that the external financial environment changes in a way that results in the displacement of their projection perpendicularly to the separator by e at points A_e and C_e.

- If $d(C) > e$, nothing happens to either bank.
- If $d(A) > e > d(C)$ (see Fig. 5), bank C fails, while bank A remains healthy.
- If $e > d(A)$, both banks fail.

Thus, two points can be obtained from this example:

(a) We see that our forecast indicates that bank A is more robust to changes than bank C; the distance between any point and the separator reflects its robustness or the confidence we have in our forecast.

(b) This example outlines a simple stress testing tool that allows the regulator to apply different scenarios for the financial variables of the bank, find the exact position of the new bank projection, and see whether the altered bank is forecasted to fail or survive. Moreover, the distance between the actual position and the separating line provides the regulator with an estimate of the maximum changes that the bank can withstand without its classification changing.

In our case, the exact values of $[\alpha, \beta, \gamma]$ are [23.45, −3.85, 10.55] respectively, and the separating line is defined as

$$\varepsilon : 23.45 \cdot \mathrm{T1CRC} - 3.85 \cdot \mathrm{TIE} + 10.55 = -0.7. \qquad (15)$$

Given the two variables T1CRC and TIE of a bank $B = \left(\mathrm{T1CRC}_{B,} \mathrm{TIE}_{B,}\right)$, we can calculate the financial position of the bank and its distance $d(B)$ from the separating line as

$$d(B) = \frac{|23.45 \cdot T1RBC_B - 3.85 \cdot TIE_B + 11.25|}{\sqrt{23.45^2 + (-3.85)^2}}. \qquad (16)$$

Fig. 6 illustrates the position of each out-of-sample bank on the solvent and insolvent sub-spaces, as defined by the dashed line. The dashed line represents the decision boundary defined in the training process of the SVM-linear model. The deep gray points represent banks that are forecasted as solvent, the light gray points banks that are forecasted as insolvent, and the black points misclassified banks.

In our empirical results, the distances from the separating line of the banks that were forecasted as solvent range from 0.001 to 2.379, while those of the banks that were forecasted as insolvent range from 0.004 to 0.249. For each bank, whether forecasted as solvent or failed, the proposed stress-testing tool quantifies the exact amount of Tier 1 (core) risk-based capital (or total assets or total interest expense or total interest income) that banking institutions should maintain in order to avoid failure and remain solvent in the following year.

4.6. Simulation and scenarios

What follows is a simulation of the proposed stress-testing tool for an out-of-sample bank that failed in 2011.

Table 8
Balance sheet and income and expense statement of bank A ($thousands).

Assets and liabilities	December 2010
Tier 1 (core) risk-based capital	52,508
Total assets	2,153,690
Income and expense statement	
Total interest expense	18,776
Total interest income	84,660

Table 9
Critical values in order for bank A to move to the subspace of solvent banks.

Variable	Value	Critical	Margin
T1CRC	0.024	0.040	40.0%
TIE	0.222	0.216	2.8%

Table 8 presents the financial statements of bank A[10] as of December 2010.

I. Outcome: Failure in year $t + 1$

The Tier 1 (core) risk-based capital over total assets (T1CRC) is equal to 0.024 and the total interest expense over total interest income (TIE) is 0.222. The scaled values[11] for T1CRC and TIE are −0.640 and −0.711 respectively. According to the decision boundary defined in the training process of the linear model Eq. (15), the bank is forecasted to be insolvent in the following year. The distance of the bank from the decision boundary is 0.043, calculated using Eq. (16).

II. Change of classification

Given the distance, we calculate the minimum changes that will result in the reclassification of bank A's projection to the solvency subspace. Table 9 shows both the initial values and the critical ones in order for the bank to change classification. These critical values show that the bank should maintain a T1CRC greater than 0.040 and a TIE equal to or less than 0.216 in order to change class and become solvent.

III. Marginal scenarios

As expected, the bank should increase its capital and decrease its interest expense ratio. However, the real question is exactly how much capital is needed or how much the bank should downsize its expenses in order to avoid failure in the next year.

(a) Changes in one variable

If we restrict changes to only T1CRC (TIE remains constant), then the bank should maintain a T1CRC greater than 0.041; in other words, its Tier 1 (core) risk-based capital should be over $87,486 or its total assets should be less than $1,292,615 in order for the bank to be displaced to the subspace of solvent banks. Similarly, if we restrict changes

[10] This is the statement of a real failed bank, though we do not name the bank.

[11] The scaled values are calculated using the min and max of the training data.

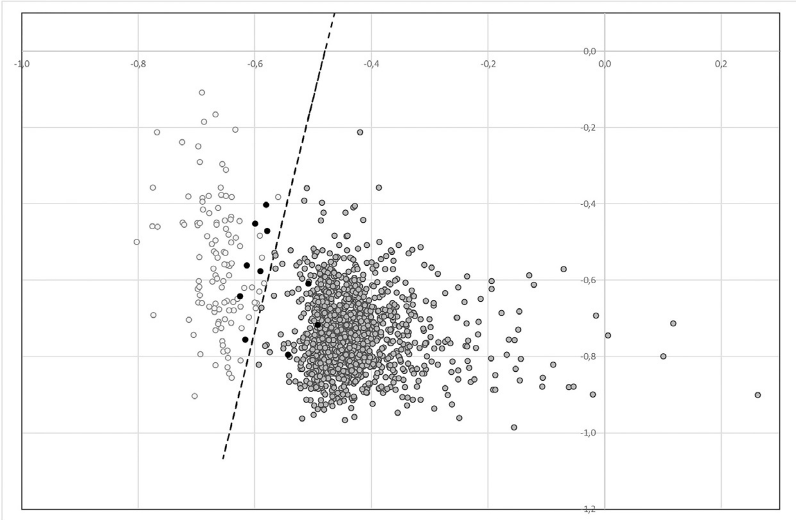

Fig. 6. Graphical representation of the out-of-sample forecasting: the dashed line represents the decision boundary obtained in the training process of the SVM-linear model, the light and deep gray dots the banks that are forecasted as insolvent and solvent respectively, and the black dots banks that were misclassified.

to only TIE (T1CRC remains stable), then the bank should maintain a TIE less than 0.017; or, equivalently, either its total interest expense should be less than $1398 or its total interest income should be more than $1,136,295.

(b) Changes in both variables

Table 10 presents the minimum-change alternative scenarios that will reclassify the bank to the solvent subspace.

According to Scenario 1, the bank should increase its Tier 1 (core) risk-based capital from $52,508 to at least $86,568, while at the same time decreasing its expenses from $18,776 to $18,320 or less (while total assets and total interest income are held fixed). In this way, T1CRC will be equal to 0.040 and TIE will be equal to 0.216, and the bank will be on the decision boundary, so that any increase in capital over $86,568 or any decrease in the total interest expense under $18,320 will result in the reclassification of the bank to the subspace of solvent banks. Scenario 2 prescribes increasing the capital and total interest income to over $86,568 and $86,768 respectively. Similarly, Scenarios 3 and 4 require the bank to downsize in terms of total assets to at least $1,306,324 and either decrease the total interest expense to less than $18,320 or increase the total interest income to over $86,768.

With respect to the banks forecasted as solvent, the proposed method can be used as a tool for assessing their resilience by providing a quantitative framework of their safety margins in terms of T1CRC and TIE; in other words, a safety margin regarding how much Tier 1 (core) risk-based capital or total assets a bank may maintain, without affecting its financial position in the following year. Moreover, the method provides an acceptable level of TIE that will maintain the bank in the solvent subspace. Specifically, one can calculate a safety margin of an increase in total interest expenses and a decrease in total interest income.

The proposed stress-testing tool can be employed by the supervisory authorities for making decisions on individual banking institutions and the financial system. It can also be used for assessing the default risk of a bank by both the bank's risk management division and investors making financial and investment decisions. It is noteworthy that the proposed stress testing model was created without any a priori information regarding the underlying data.

5. Discussion

The forecasting and stress-testing framework developed above has several implications regarding the separator boundary, the supervisory policy at the microprudential level and future research. We provide a brief discussion of these issues here.

Table 11 summarizes the forecasts of our SVM methodology vs. the *O*-score.

As we can see, in the case of the true solvent banks, the SVM approach results in 0.6% false alarms (7 out of 1164), i.e., true solvent banks that were predicted as failed. On the other hand, the percentage of false alarms produced by the *O*-score is 22.9% (267 out of 1164). In the case of the true insolvent banks, the respective percentages are 2.6% (3 out of 115) and 0.8% (1 out of 115). Considering the criticism in the literature (see for example the work of Danielsson, James, Valenzuela, & Zer, 2016a,b) on the possible negative effects of a greater strictness of the supervisory microprudential regulations, this is an important improvement to

Table 10
Scenarios of minimum changes to bank A's financial statements in order to avoid failure in year $t + 1$ (the stress test variable is marked in bold).

Assets and liabilities	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Tier 1 (core) risk-based capital	**86,568**	**86,568**	52,508	52,508
Total assets	2,153,690	2,153,690	**1,306,324**	**1,306,324**
Income and expense statement				
Total interest expense	**18,320**	18,776	**18,320**	18,776
Total interest income	84,660	**86,768**	84,660	**86,768**

the overall efficiency of prediction in terms of the costs associated with the regulatory requirements and actions.

This raises an important issue with respect to banking supervision: how strict should microprudential (bank-specific) regulation be? Initially, the consensus was that false positives, also called false alarms, were not a major issue, and were treated rather as a nuisance that was necessary for efficient and successful supervision. This is generally evident in Basel III. There are strict rules on capital adequacy, leverage and liquidity for each individual bank. Recently, though, there has been some relevant criticism, as some researchers have stated (Repullo & Suarez, 2012; Danielsson et al. 2016a,b) that in some cases "too-safe" restrictions in terms of microprudential regulations may actually increase the risk of default. There is concern that strict prudential rules at the micro (individual bank specific) level may increase both the systemic risk and the risk of contagion in the event of a mild (otherwise) downturn in economic activity. For example, in the case of a stress scenario with falling prices, financial institutions may start to sell some of the assets on their balance sheet in an attempt to meet the microprudential requirements of the supervising agent. At the micro level this is wise, but at the macroeconomic, system-wide level, if all banks start selling assets, this fuels the fall in prices further, and may start a vicious cycle and possibly even a sector-wide or economy-wide crisis. The forecasting approach that we follow in this study may help to relax these concerns by reducing the number of false positive predictions.

The three banks that are misclassified as solvent using the SVM model are very small institutions in terms of their total assets, with \$90, \$43.5 and \$24.8 million respectively. Thus, the associated cost to the FDIC was also minimal; e.g., the cost to the FDIC of the Glasgow Savings Bank failure was \$0.1 million. Moreover, the largest of the three, the First National Bank of Davis, with assets \$90 million, was misclassified, but it was discovered later that there were significant losses that were not reported in its financial statements. This bank was mis-predicted as solvent by the O-score as well.

An interesting and natural future extension of this framework would be to apply this methodology to European banks supervised by the ECB. However, this may prove to be more difficult, as there is no central data depository that provides data on the European banks' publicly available financial statements in a consistent and comparable format as we see with the FDIC. Moreover, the usual policy of the ECB and its national central bank branches is to recapitalize (bail-in) financial institutions that are in distress, without letting them fail. Thus, it may prove almost impossible to compile a similar solvent/failed dataset for Europe. One possible way to remedy this issue might be to forecast whether a recapitalization (bail-in) or bail-out intervention will be necessary in the future.

The forecasting model in this study employs only variables that are publicly available through the banks' published financial statements, as was mentioned in the data section. Both bank officials and the supervising authorities have access to data and information that is available to them instantly and in much more detail and depth which may improve the accuracy of this methodology further, especially for longer forecasting horizons.

6. Conclusion

The early identification of potentially failing banks is essential to both regulators and bank managers. This is evident from the microprudential regulations that have been put into effect by various bank supervising authorities worldwide over the last decade. This paper employs a support vector machines framework from the field of machine learning to construct a forecasting model for bank failures, and we compare it to the well-established Ohlson bankruptcy forecasting model. Next, based on the optimum SVM forecasting model, we propose a new form of stress-testing that informs the interested parties about necessary actions that a specific bank should take in order to avoid failure in the near future. We collected data on 144 variables from the publicly-available financial statements of 1443 U.S. banks. Our sample includes all 481 banks that failed during the period under study, along with 962 solvent ones, so that a ratio of 1:2 of failed to solvent banks is established. We selected the most informative variables by following a two-step feature selection procedure, then fed the selected variables into an SVM model through a training-testing learning process. The proposed methodology defines a linear decision boundary that separates the solvent from the failed banks with 98.25% accuracy, employing only two explanatory variables:

$$23.45 \times \text{T1CRC} - 3.85 \times \text{TIE} + 10.55 = -0.7.$$ [12]

[12] The variables are normalized to $[-1, 1]$.

Table 11
Performance comparison of the *O*-score and the SVM-linear model.

	True solvent		True insolvent	
	SVM	*O*-Score	SVM	*O*-Score
Classified solvent	1157	897	**3**	**1**
Classified insolvent	**7**	**267**	112	114
Total	1164	1164	115	115

Moreover, we tested the robustness of our model by augmenting our out-of-sample data with additional solvent banks so as to establish a more realistic ratio of 1:10 for failed to solvent banks. Again, the SVM trained model exhibits a significant ability to forecast the out-of-sample data consistently, with the overall accuracy of the model being 99.22%.

In addition, when used as a stress test, this approach offers a simple geometric interpretation of the results, providing the ability to measure the distance to default for institutions that were predicted to be solvent. Thus, banking institutions not only receive a binary classification as "solvent" or "insolvent", but also can be ranked according to their distance or safety margin from insolvency.

Apart from the obvious utility of the approach for private market participants in the banking sector, such as share-holders, depositors, borrowers, etc., these results are also significant for the regulators and supervising authorities (usually central banks) in terms of microprudential (bank-specific) supervision. This approach can raise red flags for financial institutions that are predicted to fail in the near future, thus allowing prompt action to be taken to monitor these institutions more closely and intervene early on so as to avoid insolvency or contagion, or at least to minimize its cost to the regulator (and taxpayers) and the economy as a whole.

Acknowledgments

The authors would like to thank Professor Nejat Seyhun for his valuable comments and suggestions that improved the manuscript. We also thank the participants in the seminar series at the Ross School of Business at the University of Michigan, the Department of Economics at Temple University, the BETA of the University of Strasbourg, and the External Developments and Macro-Financial Linkages Divisions of the European Central Bank for constructive comments and suggestions on this paper. This research has been partly co-financed by the European Union – European Social Fund (ESF) – and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) – Research Funding Program: THALES. Investing in knowledge society through the European Social Fund (MIS 380292).

Appendix A. Different cut-off levels of the SVM model with two explanatory variables achieved using the dataset with 1443 samples

See Table A.1.

Table A.1
Different cut-off levels of the SVM model with two explanatory variables achieved using the dataset with 1,443 samples.

Cut-off	Overall accuracy	Insolvent accuracy	Solvent accuracy
0	96.79%	97.39%	96.49%
0.1	95.92%	97.39%	95.17%
0.2	95.63%	97.39%	94.74%
0.3	95.63%	97.39%	94.74%
0.4	95.33%	97.39%	94.30%
0.5	95.33%	97.39%	94.30%
0.6	95.04%	97.39%	93.86%
0.7	94.75%	97.39%	93.42%
0.8	94.75%	97.39%	93.42%
0.9	94.75%	98.26%	92.98%
1	94.75%	99.13%	92.54%
−0.1	97.38%	97.39%	97.37%
−0.2	97.38%	97.39%	97.37%
−0.3	97.38%	97.39%	97.37%
−0.4	97.38%	97.39%	97.37%
−0.5	97.38%	97.39%	97.37%
−0.6	97.96%	97.39%	98.24%
−0.7	**98.25%**	97.39%	98.68%
−0.8	97.96%	96.52%	98.68%
−0.9	97.38%	94.78%	98.68%
−1	96.79%	92.17%	99.12%

Appendix B. Performance of the shifted SVM model on the out-of-sample enhanced dataset

The shifted linear separator of the SVM model (cut-off = 1) achieved the same forecasting accuracy of 114 cases out of 115 as the *O*-score, while keeping a lower level of false alarms (50 false alarm cases versus 267). The overall accuracy of this model is 96% (see Table B.2).

Table B.2
Performance of the shifted SVM model on the out-of-sample enhanced dataset.

SVM-linear	TRUE	
	Solvent	Insolvent
Classified solvent	1114	1
Classified insolvent	50	114

References

Agarwal, V., & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking and Finance*, *32*(8), 1541–1551.

Altman, E. I., Haldeman, R. G., & Narayanan, P. (1977). ZETA analysis: a new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance*, *1*(1), 29–54.

Begley, J., Ming, J., & Watts, S. (1996). Bankruptcy classification errors in the 1980s: an empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies*, *1*(4), 267–284.

Berger, A. N., & Bouwman, C. H. (2013). How does capital affect bank performance during financial crises?. *Journal of Financial Economics*, *109*(1), 146–176.

Boyacioglu, M. A., Kara, Y., & Baykan, Ö. K. (2009). Predicting bank financialfailures using neural networks, support vector machines and multivariatestatistical methods: a comparative analysis in the sample of savings depositinsurance fund (SDIF) transferred banks in Turkey. *Expert Systems with Applications*, *36*(2), 3355–3366.

Chen, N., Ribeiro, B., & Chen, A. (2016). Financial credit risk assessment: a recent review. *Artificial Intelligence Review*, *45*(1), 1–23.

Cleary, S., & Hebb, G. (2016). An efficient and functional model for predicting bank distress: in and out of sample evidence. *Journal of Banking and Finance*, *64*, 101–111.

Cole, R., & Gunther, J. (1998). Predicting bank failures: A comparison of on- and off-site monitoring systems. *Journal of Financial Services Research*, *13*(2), 102–117.

Cole, R. A., & White, L. J. (2012). Déjà vu all over again: the causes of US commercial bank failures this time around. *Journal of Financial Services Research*, *42*(1–2), 5–29.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297.

Cox, R. A., & Wang, G. W. Y. (2014). Predicting the us bank failure: a discriminant analysis. *Economic Analysis and Policy*, *44*(2), 202–211.

Curry, T. J., Elmer, P. J., & Fissel, G. S. (2007). Equity market data, bank failures and market efficiency. *Journal of Economics and Business*, *59*, 536–559.

Danielsson, J., James, K., Valenzuela, M., & Zer, I. (2016a). Can we prove a bank guilty of creating systemic risk? a minority report. *Money, Credit and Banking*, *48*, 795–812.

Danielsson, J., James, K., Valenzuela, M., & Zer, I. (2016b). Model risk of risk models. *Journal of Financial Stability*, *23*, 79–91.

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

Ecer, F. (2013). Comparing the bank failure prediction performance of neural networks and support vector machines: the Turkish case. *Ekonomska Istraživanja*, *26*(3), 81–98.

Erdogan, B. E. (2013). Prediction of bankruptcy using support vector machines: an application to bank bankruptcy. *Journal of Statistical Computation and Simulation*, *83*(8), 1543–1555.

Espahbodi, P. (1991). Identification of problem banks and binary choice models. *Journal of Banking and Finance*, *15*(1), 53–71.

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

Grice, J. S., & Dugan, M. T. (2003). Re-estimations of the Zmijewski and Ohlson bankruptcy prediction models. *Advances in Accounting*, *20*, 77–93.

Hanweck, G. A. (1977). Predicting bank failure (No. 19). Board of Governors of the Federal Reserve System (US).

Hillegeist, S. A., Keating, E. K., Cram, D. P., & Lundstedt, K. G. (2004). Assessing the probability of bankruptcy. *Review of Accounting Studies*, *9*(1), 5–34.

Iturriaga, F. J. L., & Sanz, I. P. (2015). Bankruptcy visualization and prediction using neural networks: a study of US commercial banks. *Expert Systems with Applications*, *42*(6), 2857–2869.

Jordan, D. J., Rice, D., Sanchez, J., Walker, C., & Wort, D. H. (2010). *Predicting bank failures: Evidence from 2007 to 2010. Working paper*.

Karamzadeh, M. S. (2013). Application and comparison of Altman and Ohlson models to predict bankruptcy of companies. *Research Journal of Applied Sciences, Engineering and Technology*, *5*(6), 2007–2011.

Kolari, J., Glennon, D., Shin, H., & Caputo, M. (2002). Predicting large us commercial bank failures. *Journal of Economics and Business*, *54*(4), 361–387.

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.

Lane, W. R., Looney, S. W., & Wansley, J. W. (1986). An application of the cox proportional hazards model to bank failure. *Journal of Banking and Finance*, *10*(4), 511–531.

Männasoo, K., & Mayes, D. G. (2009). Explaining bank distress in eastern european transition economies. *Journal of Banking and Finance*, *33*(2), 244–253.

Martin, D. (1977). Early warning of bank failure: a logit regression approach. *Journal of Banking and Finance*, *1*(3), 249–276.

Mayes, D. G., & Stremmel, H. (2014). The effectiveness of capital adequacy measures in predicting bank distress. SUERF Studies.

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

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, *18*, 109–131.

Papadimitriou, T., Gogas, P., Plakandaras, V., & Mourmouris, J. C. (2013). Forecasting the insolvency of US banks using support vector machines (SVM) based on local learning feature selection. *International Journal of Computational Economics and Econometrics*, *3*, 83–90.

Repullo, R., & Suarez, J. (2012). The procyclical effects of bank capital regulation. *The Review of Financial Studies*, *26*(2), 452–490.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business*, *74*(1), 101–124.

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

Stuhr, D. P., & Van Wicklen, R. (1974). Rating the financial condition of banks: A statistical approach to aid bank supervision. Federal Reserve Bank of NewYork.

Sun, Y., Todorovic, S., & Goodison, S. (2010). Local-learning-based feature selection for high-dimensional data analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(9), 1610–1626.

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

Whalen, G. (1991). A proportional hazards model of bank failure: an examination ofits usefulness as an early warning tool. *Federal Reserve Bank of Cleveland Economic Review*, *27*(1), 21–31.

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.

Periklis Gogas is an Associate Professor at the Democritus University of Thrace and was recently a Visiting Scholar at the Ross School of Business, University of Michigan. His research interests are in macroeconomics, banking and finance, forecasting, complex networks and machine learning. He has presented invited seminars at the ECB, University of Michigan, Temple University, and others. He has also published in the *Journal of Money, Credit and Banking*, *Journal of Banking and Finance*, *International Finance*, *Journal of Forecasting*, *Physica A*, *Computational Economics*, *Economic Modelling*, *Open Economies Review*, and *Energy Economics*, among others.

Theophilos Papadimitriou is an Associate Professor at the Department of Economics, Democritus University of Thrace. He has been an invited speaker at ECB, the University of Michigan, Temple University, and the University of Strasbourg. His research interests include machine learning and complex networks. He has co-authored more than 40 research papers in leading and prestigious journals.

Anna Agrapetidou is a Ph.D. student at the Democritus University of Thrace in the department of Economics. Her research interests focus on machine learning and forecasting methods, with a specific emphasis on the banking sector.