Predicting US Banks Bankruptcy: Logit Versus Canonical Discriminant Analysis

Zeineb Affes[1] · **Rania Hentati-Kaffel[1]** (ORCID)

Abstract In this paper, we use random subspace method to compare the classification and prediction of both canonical discriminant analysis and logistic regression models with and without misclassification costs. They have been applied to a large panel of US banks over the period 2008–2013. Results show that model's accuracy have improved in case of more appropriate cut-off value C^*_{ROC} that maximizes the overall correct classification rate under the ROC curve. We also have tested the newly H-measure of classification performance and provided results for different parameters of misclassification costs. Our main conclusions are: (1) The logit model outperforms the CDA one in terms of correct classification rate by using usual cut-off parameters, (2) C^*_{ROC} improves the accuracy of classification in both CDA and logit regression, (3) H-measure and ROC curve validation improve the quality of the model by minimizing the error of misclassification of bankrupt banks. Moreover, it emphasizes better prediction of banks failure because it delivers in average the highest error type II.

Keywords Bankruptcy prediction · Early-warning system · Canonical discriminant analysis · Logistic regression · Principal component analysis · CAMELS · ROC curve · H-measure · Cost of misclassification

✉ Rania Hentati-Kaffel
rania.kaffel@univ-paris1.fr

Zeineb Affes
zeineb.affes@univ-paris1.fr

[1] Centre d'Economie de la Sorbonne, Maison des Sciences Economiques, Université Paris1 Panthéon-Sorbonne, 106-112 Boulevard de l'Hôpital, 75013 Paris, France

1 Introduction

The financial crisis starting in 2007 is considered as the first real crisis of excess financial complexity. It has illustrated the degree of the existing inter-connectivity between banks and financial institutions and highlighted the phenomenon of contagion that might exist in the interbank market. Since then, a swarming literature has been developed on the subject of quantification, prediction and control of systemic risk.

One of the methods proposed to prevent contagion of bank failures is to assess the bank failure rate. This approach helps to establish an early warning model of bank difficulties. Thus, interactions between solvency and refinancing risk can identify banks which have the most difficulties to refinance and therefore it would be perceived as risky by the other institutions. This flagging process would limit counterparty risk and warn the financial authorities of a pending liquidity risk in case of default of these banks.

The financial literature is rich of methods and models which aim to identify institutions whose financial situations appear alarming and call supervisors to act. In this study, we use random subspace method to compare the classification and prediction of both Canonical Discriminant Analysis (CDA) and Logistic Regression (LR) model with and without misclassification costs, applied to a large panel of US banks over the period 2008–2013.

Main questions raised in our paper are how to aggregate classifications results, what EWS should be proposed to regulators and also how Area Under Curve and H-measure (Hand 2009) perform classification problem?

The specificity of our study lies in the extensive list of used financial ratios (solvency ratios, quality of assets, cash or liquidity...) and the depth of sample (US banks:1224 per years) including large ans small bank. The choice of the period is justified by the high number of failed banks which have taken place there. Also, we show in this paper how the clustering quality is improved by using a more appropriate cut-off. By contrast, studies in the literature wrongly use the probability "0.5" as a cut-off value in the logit and "0" in the CDA model. Moreover, the majority of papers discussing the same topics did not consider the cost of misclassifying an active or a defaulted bank. Our main contribution here is the use of the misclassification cost to enhance the conformity between obtained results and reality.

The empirical literature distinguishes two methods: parametric and non-parametric validation. Beaver (1966) was the pioneer in using a statistical model for predicting bankruptcy. The approach is to select from thirty financial ratios those which are the most effective indicators of financial failures. The study concludes that the (Cash flow/total debt) ratio is the best forecasting indicator.

Altman (1968) tested Multiple Discriminant Analysis (MDA) to analyze 70 companies, first by identifying the best five significant explanatory variables from a list of 22 ratios and then by applying the (MDA) to calculate a Z-Altman score for each company. This score was almost accurate in predicting bankruptcy one year ahead. This model was then subsequently improved in Altman and Narayanan (1997) by proposing the zeta model that includes seven variables and classified correctly 96% of companies one year before bankruptcy and 70% five years before bankruptcy.

Since then, the use of discriminant analysis has grown through the different published studies (Bilderbeek 1979; Ohlson 1980; Altman 1984; Zopounidis and Dimitras 1993...). The vast majority of studies achieved after 1980 used the logit models to overcome the drawbacks of the DA method (Zavgren 1985; Lau 1987; Tennyson et al. 1990...). The logit analysis fits linear logistic regression model by the method of maximum likelihood. The dependent variable(the probability of default) gets the value "1" for bankrupted banks and "0" for healthy banks.

Numerous comparative studies were carried out (Keasey and Watson 1991; Dimitras et al. 1996; Altman and Narayanan 1997; Wong et al. 1997; Olmeda and Fernández 1997; Adya and Collopy 1998; O'leary 1998; Zhang et al. 1998; Vellido et al. 1999; Coakley and Brown 2000; Aziz and Dar 2004; Balcaen and Ooghe 2006; Balcaen et al. 2004; Kumar and Ravi 2007). However, the supremacy of one method over another remains subject to various controversies because of the heterogeneity of used data in the validation (database, number of points in the data, sample selection, validation methods for forecasting, the number and the nature of explanatory variables tested in the models (financial, qualitative...).

The aim of this paper is twofold: descriptive and predictive. Descriptive is to be understood as a detailed analysis of used models inputs from financial and statistics point of view. Thus, we proceed by describing and analyzing key financial ratios of the active and non-active banks for the entire period from 2008 to 2013.

We combined two parametric models (Canonical Discriminant Analysis and Logit) with the descriptive Principal Component Analysis model (PCA) to construct an Early Warning System (EWS).

First, (PCA) reduced the dimension size of data and insure an uncorrelated blend of variables framework. Then, factor scores were estimated for each bank. These scores were used to estimate (CDA) and Logit models.

One among the important results of this paper is to have compared several methods to calculate the theoretical value of the probability of default that will serve as threshold to split the bank universe into two set : failed and healthy.

The paper consists of five sections. After the introduction, an overview of the existing literature dealing with the bank failure prediction is given. Section 2 describes used data and the methodology. In Sect. 3 we implement Principal component Analysis on our data. Section 4 provides the empirical results with and without misclassification costs. Final section contains concluding remarks.

2 Description of the Methodology and the Variables

This section focuses on the data gathered for the estimation of our models. We begin with the description of the collected data and the variables selection process. Next, we present the financial and economic ratios, then we provide some descriptive statistics and correlation analysis.

We examine whether one can enhance bankruptcy prediction accuracy by a careful examination of the functional relationship between explanatory variables and the probability of bankruptcy.

Table 1 Data analysis

Year	Number of failed banks	Number of non-failed bank	% of failed banks
2013	15	835	1.76
2012	33	850	3.74
2011	72	883	7.54
2010	122	955	11.33
2009	131	1077	10.84
2008	37	1205	2.98

2.1 Data Description and Methodology

We proceed to the constitution of our database of US banks from mainly two sources: "BankScope" and FDIC. It covers the period 2008–2013. Statistics shows that the period from 2008 to 2013 was marked by a wave of bank failures in the United States: more than 450 bank failures.

After data reprocessing, sampled banks were split into two categories: active banks and non-active banks. Non-active banks are those which have been declared as bankrupted by the Federal Deposit Insurance Corporation (FDIC). The information on the identity and the bank's balance sheet data are obtained from the FDIC website. Indeed, all US banks must report their financial statements in the Uniform Bank Performance Report. Some treatments have been applied to our sample to allow homogeneity between banks. A bank that has been declared bankrupted in the first quarter of year "N" will be reclassified and considered as bankrupted in late "N−1".

Banks declared bankrupt by the FDIC after 01/04/N and for which there is no information for the current year, will be considered as inactive for the year "N". For banks that will bankrupt later and for which data are available on 31/12/N, will be considered as active for "N". Financial variables of active banks were retrieved from the database "BankScope". We notice that, data were available for only 928 banks each year in the period 2008–2013.

On this basis, the number of failed banks was reduced to 410 over the entire period. Table 1 gives more details as such constituted database.

To investigate robust performance of classification and prediction models, we split data in to Testing Set (TES) and Training Set (TRS). In fact, it was proven that classification tends to favor active banks (AB) which represent the majority class.

Thus, a random sampling was used to avoid the selection bias due to the concentration of specific bank-year samples in Training Set (TRS). (TES) represent 20% of data and were randomly selected.

Otherwise, one of the main questions raised in this study is how to improve parametric model accuracy by using the best cut-off value to classify banks.

Studies in the literature wrongly use the probability value $P* =$ "0.5" as cut-off value in the LR and a score target $D* =$ "0" in the CDA model. We show in this paper how the clustering quality is improved by using a more appropriate cut-off (a cut-off value that maximizes the overall correct classification rate obtained under the ROC curve).

Indeed, the asymmetrical composition of used samples, due to the low number of failed banks relative to total number of banks, requires the use of a cut-off based on the ROC curve equal to C*.

We propose to compare accuracy and sensitivity of classification under $P*=0.5$ and C^*_{Logit} for Logistic Regression and D* according to the formula (9) versus C^*_{CDA} for CDA model.

After, presenting the results of variables selection under the Principal Component Analysis (PCA), the procedure for each model is the same and summarized as follow:

1. *Step 1* Score calculation
2. *Step 2* Clustering for both (TRS) and (TES). We establish the confusion matrix based on :
   (a) D* Vs C^*_{CDA}
   (b) P*=0.5 Vs C^*_{Logit}
3. *Step 3* Evaluating performance of models with and without cost [correct classification rate, Sensitivity, Specificity, AUC and H-measure (Hand 2009)].

2.2 Variables: Review of the Literature

Federal regulators developed the numerical CAMEL rating system in the early 1970s to help structuring their examination process. This rating is based on the capital adequacy, asset quality, management quality, earnings ability, and liquidity position ratios. Capital adequacy evaluates the quality of a bank's capital. Asset quality measures the level of risk of a bank's assets. This reflects the quality and the diversity of the credit risk and the ability of the bank to repay issued loans. Management quality is a measure of the quality of a bank's officers and the efficiency of its management structure. Earnings ability reflects the performance of banks and the stability of its earnings stream. Liquidity measures the ability of banks to meet unforeseen and unexpected deposit outflow in the short time. In February 1997, a sixth component sensitivity to market risk was added to the CAMEL rating system.

A very abundant literature has tried to identify the most significant variables of the bank financial health of banks. According to Sinkey (1975), the quality of bank assets is the most significant ratio. Assets composition, loans characteristics, capital adequacy, source and use of income, efficiency and profitability are also discriminant variables. Poor asset quality and low capital ratios were the two characteristics of banks most consistently associated with banking problems during the 1970s (Sinkey 1978). Avery and Hanweck (1984), Barth et al. (1985) and Benston (1985) conclude that the proxies of loans portfolio composition and quality, capital ratio and the source of income are significant. Thomson (1991) demonstrate that the probability that a bank will fail is a function of variables related to its solvency, including capital adequacy, asset quality, management quality, and the relative liquidity of the portfolio. Martin (1977) found that the capital asset ratio, and the loans portfolio's composition to total assets ratio have a high level of significance. Pantalone et al. (1987) proposed a model including most of CAMEL proxies: profitability, management's efficiency, leverage, diversification and economic environment. Their results confirm the main cause of default was bad credit risk management. The model of Barr and Siems (1994) includes CAMEL

proxies and efficiency scores as management's quality proxies and a proxy of the economic conditions. The six variables selected for their failure-prediction models are equity/total loans (C), non performing loans/total assets (A), DEA efficiency score (M), net income/total assets (E) and large dollar deposits/total assets (L).

Our main objective in this study is to provide an accurate bank failure model based on the significant fragility factors. In line with the literature, we maintain the most commonly used financial ratios which can forecast potential failures (Beaver 1966; Altman 1968; Thomson 1991; Kolari et al. 1996; Jagtiani et al. 2003; Dabos and Sosa-Escudero 2004; Lanine and Vander Vennet 2006).

3 Principal Component Analysis

3.1 Statistic Description of Variables

We include in our analysis four categories of variables: (1) two measures of capital adequacy. They assesses the financial strength of a bank and determine its capacity to cover time liabilities and other risk such as credit risk, market risk, operational risk and others. The most popular proxies for capturing capital adequacy in previous literature were total equity divided to total assets or to total loans ratios. (2) Assets quality measures are considered through the data construction process. These variables have a crucial role in the assessment of the current condition and the financial capacity in the future. We employ four variables related to asset quality (NPLTA, NPLGL, LLRTA, and LLRGL). For NPLTA and NPLGL variables, we add the proxy loans not accruing to loans over 90 days late divided by total assets (non-performing loans/total assets). (3) Bank profitability which is assessed through two ratios. The first ratio is the net profit as a share of total assets. Like the second measure, it is the net profit as a share of total shareholders' equity. Both measures are positively related to the financial performance of the bank and negatively related to the later failure (Hassan Al-Tamimi and Charif 2011). (4) The liquidity level of the bank is assessed through employing two ratios. The first one is total liquid assets to total assets. This indicates the ability of the bank to cover its liabilities. The second ratio, used to estimate liquidity, is total liquid assets as a share of total deposits. This ratio depicts the capacity of the bank to cover unanticipated deposit withdrawal. The ratio of liquid assets to short term liabilities is the last ratio to determine the liquidity. The explanatory variables are shown below:

Categories CAMEL	Variables	Definition
Capital adequacy	EQTA	Total equity/total assets
	EQTL	Total equity/total loans
Assets quality	NPLTA	Non performing loans/total assets
	NPLGL	Non performing loans/gross loans
	LLRTA	Loan loss reserves/total assets
	LLRGL	Loan loss reserves/gross loans
Earnings ability	ROA	Net income/total assets
	ROE	Net income/total equity
Liquidity	TLTD	Total loans/total customer deposits
	TDTA	Total customer deposits/total assets

Table 2 presents the means of the ten financial ratios for the two groups [Active Bank (AB) and Default Bank (DB)], and significance tests for the equality of group means for each ratio.

First, according to capital adequacy ratios which are measures of how much capital is used to support the banks' risk assets. (EQTA) ratios for (DB) are on average very low. A low ratio means a significant leverage of these banks. This makes banks less resistant to shocks. Thus, the higher (EQTA) value is; the lower the probability of default will be. As banks trend toward failure, their equity position is likely to decrease, thus a negative relationship is expected between total loans and failure. The same conclusions emanate from the (EQTL) ratio analysis.

According to the asset quality ratio, we note that (NPLTA) ratios for (DB) are very high and disparate for the period spanning between 2008 and 2013. The immediate consequence of large amount of non-performing loans(NPL) is bank failure. In fact, according to our data, the economic environment has pushed up (NPL), consequently ratios (NPLTA), (NPLGL), (LLRTA) and (LLRGL) increased. Banks with a high (NPL) amount tend to carry out internal consolidation to improve the asset quality rather than distributing credit and then be obliged to raise provision for loan loss. For example, (NPLTA) ratio fell by 3.95% for (DB) and by 1.16% for (AB). We note that low value of loan portfolio signals the potential existence of an important vulnerability in the financial system (17.75% in 2008 and 17.83% in 2010). (LLRTA) and (LLRGL) provide an useful indication for analysts willing to evaluate the stability of bank's lending base. The higher the ratio, the poorer the quality of the loan portfolio is (3.56% for (LLRTA) ratio and 5.15% for (LLRGL) ratio for the (DB) in 2011).

Finally (TLTD) and (TDTA) liquidity ratios are often used by policy makers to assess the lending practices of banks and draw some statistics. If the ratio is too high, it means that banks might not have enough liquidity to cover any unforeseen fund requirements; if the ratio is too low, banks may not be earning as much as they could be. Table 2 exhibits high values for(DB)in average (for example (TDTA) are 86.72 and 91.12% for 2008 and 2009). These high ratios mean that those banks are relying on borrowed funds.

Table 3 is used to analyze the correlation coefficients between the different explanatory variables and the dependent variable (probability of default). We note the significance at 1 and 5% for all the variables that we have retained in our study. Moreover, most of the coefficients have the expected signs. For example, a negative correlation is confirmed for (EQTA) and (EQTL) for all years. Indeed, an increase in the value of the two ratios (a high level of equity) will reduce the probability of default of the bank (a positive effect on the bank's survival).

Table 4, presents the correlation matrix of ratios. Here, it can be seen that most of the ratios are correlated to each other. When we look more closely to the correlation matrix, we can distinguish the following aspects: there is a strong correlation (over 90%) between the pairs of variables (NPLTA)/(NPLGL) and (LLRTA)/(LLRGL). This result is generalized over the whole period and it indicates a strong link between them and that one of the variables can be substituted by another. Also, the Asset quality component (AQ) which groups (NPLTA), (NPLGL), (LLRTA) and (LLRGL) variables is negatively correlated with the return on assets (ROA). The variables (EQTA) and (EQTL), which are a proxy of capital adequacy, are negatively correlated with proxies

Table 2 Descriptive statistics of the ten variables

Variables	Bankrupt (37)		NonBankrupt (1205)		All (1242)		Fisher's F statistical test	
	Average (%)	SD	Average (%)	SD	Average (%)	SD	Value	Proba
2008								
EQTA	3.74	0.028	10.31	0.049	10.11	0.050	65.2521	0
EQTL	5.43	0.043	16.50	0.179	16.17	0.177	14.1981	0.000172
NPLTA	13.34	0.079	2.63	0.038	2.95	0.044	259.1017	0
NPLGL	17.75	0.099	3.57	0.051	3.99	0.058	255.0764	0
LLRTA	3.27	0.017	1.23	0.010	1.29	0.011	144.5561	0
LLRGL	4.39	0.021	1.70	0.014	1.78	0.015	134.5807	0
ROA	−5.53	0.035	−0.52	0.023	−0.66	0.025	160.7106	0
ROE	−155.24	4.400	−9.41	0.399	−13.75	0.882	106.5167	0
TLTD	83.08	0.147	91.10	0.236	90.86	0.234	4.2442	0.039593
TDTA	86.72	0.075	78.50	0.124	78.74	0.124	16.0524	0.000065

Variables	Bankrupt (131)		NonBankrupt (1077)		All (1208)		Fisher's F statistical test	
	Average (%)	SD	Average (%)	SD	Average (%)	SD	Value	Proba
2009								
EQTA	1.49	0.035	9.71	0.044	8.81	0.050	416.9754	0
EQTL	2.29	0.057	16.01	0.156	14.52	0.154	100.1981	0
NPLTA	11.70	0.078	3.49	0.043	4.38	0.055	338.3866	0
NPLGL	16.14	0.108	4.92	0.059	6.13	0.075	336.0824	0
LLRTA	3.44	0.023	1.47	0.011	1.68	0.014	270.4139	0
LLRGL	4.74	0.032	2.10	0.015	2.39	0.020	260.8247	0
ROA	−6.28	0.049	−0.91	0.026	−1.49	0.034	386.4597	0
ROE	−1060.78	81.646	15.58	12.607	−101.15	29.511	15.7272	0.000077
TLTD	75.78	0.150	82.64	0.203	81.90	0.199	14.0123	0.00019
TDTA	91.12	0.084	80.96	0.118	82.06	0.119	92.1347	0

Table 2 continued

Variables	Bankrupt (122)		NonBankrupt (955)		All (1077)		Fisher's F statistical test	
	Average (%)	SD	Average (%)	SD	Average (%)	SD	Value	Proba
2010								
EQTA	1.32	0.032	10.04	0.043	9.05	0.050	464.2639	0
EQTL	1.91	0.045	17.13	0.159	15.40	0.158	109.9660	0
NPLTA	12.95	0.068	3.11	0.037	4.23	0.052	596.9111	0
NPLGL	17.83	0.087	4.57	0.053	6.07	0.072	562.5029	0
LLRTA	3.46	0.018	1.45	0.011	1.68	0.013	309.2754	0
LLRGL	4.75	0.022	2.16	0.015	2.45	0.018	274.0009	0
ROA	−3.51	0.031	−0.31	0.022	−0.67	0.025	207.7540	0
ROE	−306.91	16.474	−21.04	1.235	−53.42	5.718	27.7129	0
TLTD	75.51	0.113	79.53	0.185	79.07	0.179	5.4807	0.01941
TDTA	91.84	0.069	81.73	0.118	82.88	0.118	85.3704	0

Variables	Bankrupt (72)		NonBankrupt (883)		All (955)		Fisher's F statistical test	
	Average (%)	SD	Average (%)	SD	Average (%)	SD	Value	Proba
2011								
EQTA	1.05	0.016	10.76	0.043	10.02	0.049	358.7365	0
EQTL	1.62	0.025	18.94	0.166	17.63	0.166	78.2484	0
NPLTA	12.01	0.064	2.54	0.030	3.25	0.042	518.3910	0
NPLGL	17.33	0.086	3.87	0.045	4.89	0.061	497.7271	0
LLRTA	3.56	0.018	1.31	0.009	1.48	0.011	365.3116	0
LLRGL	5.15	0.025	2.01	0.013	2.25	0.016	346.5782	0
ROA	−2.82	0.026	0.21	0.016	−0.02	0.018	222.0039	0
ROE	−104.06	8.068	−5.28	0.615	−12.73	2.294	12.4922	0.000428
TLTD	70.81	0.098	77.37	0.191	76.88	0.186	8.3528	0.003938
TDTA	92.71	0.056	81.50	0.119	82.35	0.119	62.6818	0

Table 2 continued

Variables	Bankrupt (33)		NonBankrupt (850)		All (883)		Fisher's F statistical test	
	Average (%)	SD	Average (%)	SD	Average (%)	SD	Value	Proba
2012								
EQTA	0.51	0.025	11.13	0.042	10.73	0.046	205.2486	0
EQTL	0.86	0.041	20.32	0.217	19.59	0.216	26.6014	0
NPLTA	10.23	0.039	1.94	0.023	2.25	0.029	379.4552	0
NPLGL	16.38	0.068	3.04	0.034	3.54	0.044	432.8428	0
LLRTA	3.03	0.013	1.16	0.007	1.23	0.008	231.6479	0
LLRGL	4.75	0.021	1.84	0.010	1.95	0.012	225.4214	0
ROA	−2.90	0.031	0.59	0.010	0.46	0.014	278.7442	0
ROE	−1016.60	33.699	4.64	0.420	−33.53	6.718	79.9908	0
TLTD	64.77	0.109	76.20	0.196	75.77	0.195	11.0998	0.000899
TDTA	94.71	0.052	81.65	0.117	82.14	0.117	41.0856	0

Variables	Bankrupt (15)		NonBankrupt (835)		All (850)		Fisher's F Statistical test	
	Average (%)	SD	Average (%)	SD	Average (%)	SD	Value	Proba
2013								
EQTA	1.13	0.019	11.34	0.047	11.16	0.048	71.8599	0
EQTL	1.87	0.031	19.84	0.181	19.52	0.181	14.8360	0.000126
NPLTA	9.39	0.074	1.47	0.018	1.61	0.023	222.3304	0
NPLGL	14.01	0.103	2.32	0.028	2.52	0.035	211.6835	0
LLRTA	2.96	0.015	1.05	0.006	1.09	0.006	151.7121	0
LLRGL	4.39	0.020	1.67	0.010	1.72	0.011	109.6913	0
ROA	−1.91	0.020	0.69	0.010	0.65	0.011	95.5193	0
ROE	−1076.54	39.628	6.98	0.288	−12.14	5.293	66.5167	0
TLTD	66.43	0.068	77.38	0.200	77.19	0.198	4.5111	0.033965
TDTA	94.65	0.067	81.50	0.117	81.73	0.118	18.7148	0.000017

Table 3 Correlation coefficients between the explanatory variables and the dependent variable

Variables	2008	2009	2010	2011	2012	2013
EQTA	−0.224∗	−0.507∗	−0.549∗	−0.523∗	−0.435∗	−0.279∗
EQTL	−0.106∗	−0.277∗	−0.305∗	−0.275∗	−0.171∗	−0.131∗
NPLTA	0.416∗	0.468∗	0.598∗	0.594∗	0.549∗	0.456∗
NPLGL	0.413∗	0.467∗	0.586∗	0.586∗	0.574∗	0.447∗
LLRTA	0.323∗	0.428∗	0.473∗	0.526∗	0.456∗	0.390∗
LLRGL	0.313∗	0.422∗	0.451∗	0.516∗	0.451∗	0.338∗
ROA	−0.339∗	−0.493∗	−0.402∗	−0.435∗	−0.490∗	−0.318∗
ROE	−0.281∗	−0.113∗	−0.159∗	−0.114∗	−0.289∗	−0.270∗
TLTD	−0.058∗∗	−0.107∗	−0.071∗∗	−0.093∗	−0.112∗	−0.073∗∗
TDTA	0.113∗	0.266∗	0.271∗	0.248∗	0.211∗	0.147∗

∗ Significant at 1%
∗∗ Significant at 5%

of asset quality. There is a negative correlation between asset quality and profitability of the bank.

3.2 PCA's Results

In this section we present the results of variables selection under the Principal Component Analysis (PCA). The aim is to extract the most important information from the data and to compress the data dimension by keeping only the most important ratios to explain the changes in financial conditions of banks.

Several tests are provided as following (Jolliffe 2002):

1. Bartlett's test to validate the assumption of equality of variances. In this sense, if the test statistic is larger than the critical value, we reject the null hypotheses at the 5% significant level (Table 5). Thus, the sample correlation matrix did not come from a population where the correlation matrix is an identity matrix.
2. Kaiser-Meyer-Olkin (KMO) to test if the variables have enough in common to warrant a factor analysis. In this test (KMO) retain only components with eigenvalues greater than one. Eigenvalues, also called characteristic roots are presented in Table 5.

In addition to theses tests, we perform (PCA) by analysing Factor Loading which are correlation coefficients between the financial variables and factors. Finally, we determine the (PCA) scores.

Before getting to the description of (PCA), we first analyse the correlation matrix. Then after centering and standardizing each ten variables, we determine the optimal number of principal component analysis.

The starting point is the correlation matrix. Table 4 presents the degree of dependence between the initial ten variables. It can be easily seen that variables are correlated. This means that the information they convey have some degree of redundancy. To per-

Table 4 Correlation matrix

Variables	EQTA	EQTL	NPLTA	NPLGL	LLRTA	LLRGL	ROA	ROE	TLTD	TDTA
2008										
EQTA	1									
EQTL	0.628	1								
NPLTA	−0.288	−0.191	1							
NPLGL	−0.287	−0.144	0.978	1						
LLRTA	−0.248	−0.209	0.693	0.665	1					
LLRGL	−0.209	−0.064	0.636	0.655	0.917	1				
ROA	0.134	0.065	−0.561	−0.566	−0.614	−0.58	1			
ROE	0.165	0.082	−0.311	−0.325	−0.304	−0.302	0.449	1		
TLTD	0.04	−0.342	−0.03	−0.098	0.072	−0.098	0.064	0.072	1	
TDTA	−0.264	−0.173	0.196	0.175	0.144	0.112	−0.129	−0.1	−0.408	1
2009										
EQTA	1									
EQTL	0.641	1								
NPLTA	−0.561	−0.354	1							
NPLGL	−0.553	−0.333	0.98	1						
LLRTA	−0.506	−0.351	0.645	0.61	1					
LLRGL	−0.475	−0.274	0.6	0.6	0.965	1				
ROA	0.592	0.343	−0.651	−0.653	−0.71	−0.717	1			
ROE	0.05	0.028	−0.128	−0.091	−0.079	−0.054	0.051	1		
TLTD	0.105	−0.291	−0.044	−0.119	0.002	−0.134	0.153	−0.014	1	
TDTA	−0.406	−0.268	0.283	0.266	0.231	0.198	−0.322	−0.037	−0.294	1

Table 4 continued

Variables	EQTA	EQTL	NPLTA	NPLGL	LLRTA	LLRGL	ROA	ROE	TLTD	TDTA
2010										
EQTA	1									
EQTL	0.618	1								
NPLTA	−0.588	−0.372	1							
NPLGL	−0.585	−0.351	0.98	1						
LLRTA	−0.518	−0.372	0.664	0.631	1					
LLRGL	−0.485	−0.29	0.621	0.626	0.962	1				
ROA	0.564	0.329	−0.632	−0.635	−0.657	−0.651	1			
ROE	0.145	0.081	−0.12	−0.128	−0.105	−0.108	0.197	1		
TLTD	0.113	−0.343	0.003	−0.086	0.076	−0.091	0.116	0.012	1	
TDTA	−0.407	−0.253	0.266	0.254	0.218	0.192	−0.281	−0.034	−0.278	1
2011										
EQTA	1									
EQTL	0.626	1								
NPLTA	−0.522	−0.328	1							
NPLGL	−0.503	−0.282	0.977	1						
LLRTA	−0.489	−0.368	0.707	0.67	1					
LLRGL	−0.444	−0.277	0.665	0.672	0.955	1				
ROA	0.529	0.292	−0.612	−0.621	−0.617	−0.613	1			
ROE	0.111	0.061	−0.103	−0.125	−0.101	−0.135	0.211	1		
TLTD	0.128	−0.362	0.015	−0.08	0.081	−0.11	0.116	0.062	1	
TDTA	−0.398	−0.229	0.162	0.154	0.161	0.144	−0.217	−0.035	−0.278	1

Table 4 continued

Variables	EQTA	EQTL	NPLTA	NPLGL	LLRTA	LLRGL	ROA	ROE	TLTD	TDTA
2012										
EQTA	1									
EQTL	0.454	1								
NPLTA	−0.382	−0.217	1							
NPLGL	−0.359	−0.161	0.971	1						
LLRTA	−0.334	−0.276	0.603	0.57	1					
LLRGL	−0.221	−0.107	0.543	0.585	0.903	1				
ROA	0.447	0.161	−0.505	−0.511	−0.35	−0.35	1			
ROE	0.131	0.051	−0.22	−0.199	−0.087	−0.071	0.235	1		
TLTD	0.076	−0.397	0.043	−0.088	0.134	−0.17	0.181	0.015	1	
TDTA	−0.351	−0.141	0.088	0.096	0.099	0.082	−0.245	−0.079	−0.264	1
2013										
EQTA	1									
EQTL	0.538	1								
NPLTA	−0.16	−0.131	1							
NPLGL	−0.13	−0.023	0.969	1						
LLRTA	−0.148	−0.218	0.534	0.497	1					
LLRGL	0.029	0.056	0.441	0.494	0.861	1				
ROA	0.299	0.109	−0.265	−0.268	−0.119	−0.146	1			
ROE	0.079	0.036	−0.051	−0.055	−0.133	−0.131	0.126	1		
TLTD	0.077	−0.381	0.07	−0.062	0.156	−0.195	0.176	0.028	1	
TDTA	−0.302	−0.156	0.013	0.017	0.05	0.043	−0.131	−0.042	−0.282	1

Table 5 Results of Bartlett's test of sphericity and KMO

Bartlett's test	2008	2009	2010	2011	2012	2013
CHISQ	10,373.3	12,135.77	10,903.66	9434.528	7092.255	6092.29
df	45	45	45	45	45	45
p value	0	0	0	0	0	0
KMO	0.6332	0.6474	0.6372	0.6014	0.5476	0.4495

form this finding of correlation, we present in Table 5 Bartlett's test of Sphericity. Bartlett test compares the correlation matrix with a matrix of zero correlation . A zero p value is obtained over all the period from 2008 to 2013. Thus, we perform a valid factor analysis.

Table 6 shows the estimated factors and their eigenvalue. In 2008, we retain the first three factors. These factors explain 72.04% of the total variation of the financial conditions of banks. The most important factor (F1) explains 41.70% of the total variance of the selected financial ratios. The both next ones (F2) and (F3) respectively explain 16.13 and 14.21% of the total variance.

Under the same decision rule of (KMO) measure and based on the results of the Eigenvalue's factors of 2009, the first four factors accounted for 81.66% of the total variation of the financial conditions of banks divided up as follows 46.99% for (F1), 13.24% for (F2), 11.43% for (F3) and 10% for (F4).

In 2010 and 2011, the distribution between the 3 mains factors were close to 2008 one. Finally, years 2012 and 2013 displayed some similarities, we should go up to the 4th factor to totalize more than 70% of total variance.

We then consider and evaluate Factor's loadings (see Table 7). In this case, the contribution ratios in the main components vary between 0 and 1 in absolute values. If a variable contributes more than 0.5 in a specific factor, it will be retained.

For 2008, variables that explain better the first factor (F1) are (NPLTA), (NPLGL), (LLRTA), (LLRGL) and (ROA). (F1) refers to both assets quality and return on assets components. The component loadings stress the importance of the variable for the component. For example asset quality loading values are negatives. Thus, an increase in the value of these ratios will result in a lower score factor F1. So, the increase in these ratios will result in a lower score factor (F1). So, the increase in these ratios will decline the asset quality. This implies, subsequently, an increase in the probability of default of the bank. The (ROA) ratio has a positive loading, which means that an increase in its value will increase the (F1) score. We find the same results for 2009, 2010 and 2011 for (F1) factor. For 2012 and 2013, (F1) groups only variables of asset quality. For 2013, all the ratios are positively related to factor (F1).

(F2) groups ratios of Capital Adequacy (EQTA) and (EQTL) only for 2008 and 2013. In 2008, loadings for these two variables are negatives. An increase in the value of these ratios will reduce the score of the factor (F2). For 2013, the Capital Adequacy loadings are positives. An increase in the value of these two ratios will increase the value of the score of the Capital Adequacy factor and reduce the probability of default.

Table 6 Eigenvalues of the factors

Factors	Eigenvalues	Variance (%)	Cumulative (%)	Factors	Eigenvalues	Variance (%)	Cumulative (%)
2008				*2009*			
F1	4.170	41.70	41.70	F1	4.699	46.99	46.99
F2	1.613	16.13	57.83	F2	1.324	13.24	60.23
F3	1.421	14.21	72.04	F3	1.143	11.43	71.66
F4	0.883	8.83	80.87	F4	1.000	10.00	81.66
F5	0.625	6.25	87.11	F5	0.731	7.31	88.97
F6	0.559	5.59	92.70	F6	0.521	5.21	94.18
F7	0.416	4.16	96.87	F7	0.317	3.17	97.35
F8	0.240	2.40	99.26	F8	0.228	2.28	99.63
F9	0.061	0.61	99.87	F9	0.029	0.29	99.93
F10	0.013	0.13	100.00	F10	0.007	0.07	100.00
2010				*2011*			
F1	4.699	46.99	46.99	F1	4.562	45.62	45.62
F2	1.354	13.54	60.53	F2	1.387	13.87	59.50
F3	1.118	11.18	71.71	F3	1.259	12.59	72.08
F4	0.985	9.85	81.56	F4	0.982	9.82	81.90
F5	0.692	6.92	88.47	F5	0.616	6.16	88.06
F6	0.520	5.20	93.68	F6	0.532	5.32	93.39
F7	0.376	3.76	97.44	F7	0.420	4.20	97.58
F8	0.223	2.23	99.67	F8	0.206	2.06	99.64
F9	0.026	0.26	99.93	F9	0.028	0.28	99.92
F10	0.007	0.07	100.00	F10	0.008	0.08	100.00

Table 6 continued

Factors	Eigenvalues	Variance (%)	Cumulative (%)	Factors	Eigenvalues	Variance (%)	Cumulative (%)
2012				*2013*			
F1	3.832	38.32	38.32	F1	3.087	30.87	30.87
F2	1.499	14.99	53.31	F2	1.736	17.36	48.23
F3	1.396	13.96	67.27	F3	1.506	15.06	63.29
F4	1.033	10.33	77.60	F4	1.052	10.52	73.81
F5	0.758	7.58	85.18	F5	0.965	9.65	83.46
F6	0.579	5.79	90.97	F6	0.721	7.21	90.67
F7	0.487	4.87	95.84	F7	0.564	5.64	96.31
F8	0.363	3.63	99.46	F8	0.300	3.00	99.32
F9	0.042	0.42	99.89	F9	0.055	0.55	99.87
F10	0.011	0.11	100.00	F10	0.013	0.13	100.00

In 2009 this factor is influenced by the ratios (TLTD) and (TDTA). Loadings are negatives and an increase of the value will accentuate the probability of default.

In 2010, 2011, 2012 and 2013, we retain also the second factor (F2) which depends on the ratio (TLTD), the Liquidity component, and the variable (EQTL), a proxy of the Capital Adequacy. (TLTD) ratio has a positive loading, which means that an increase in its value will increase the score of the factor. (TLTD) ratio is considered as a good proxy of short term viability and a low value means that there is no optimal reallocation of resources.

For 2009, 2010 and 2011 the factor (F3) is composed by capital adequacy variables and (TDTA) variable.

In 2009, 2012 and 2013, a fourth factor (F4) is considered. It firstly added the ratio (ROE) of the list of depending variables, among two ratios of asset quality.

Finally, we determine the factor score coefficient matrix for each bank. According to the Table 8 which describes the factor score coefficients, we calculate factor scores for each bank using the formula below:

$$F_{bi} = \sum u_{ij} z_{bj} \tag{1}$$

where

- F_{bi}: the estimated factor i for bank b.
- z_{bj}: the standardized value of the jth ratios for a bank b.
- u_{ij}: the factor score coefficient for the ith factor and the jth ratios.
- These scores (F_{ai}) were used as independent variables in estimating the discriminant and the Logit model.

4 Empirical Results

4.1 Score's Results by Model

4.1.1 Canonical Discriminant Analysis

In this section, we provide Canonical Discriminant Analysis (CDA) to conduct and recuperate an early warning system indicators of failed banks. In this sense, we propose to describe the relationships among the two groups of banks (bankrupt or not) based on a set of discriminating variables.

The canonical discriminant function is expressed as follow:

$$D_{bi} = \beta_0 + \beta_1 F_{b1} + \beta_2 F_{b2} + ... + \beta_i F_{bi} \tag{2}$$

where

- D_{bi}: the value (score) on the canonical discriminant function for bank b.
- F_{bi}: represents factors validated in (PCA) section.

Each sampling group of bank has a single composite canonical score, and the group centroids indicates the most typical location of a bank from a particular group.

Table 7 Factors loading

2008 variables	F1	F2	F3
EQTA		-0.834	
EQTL		-0.918	
NPLTA	-0.867		
NPLGL	-0.871		
LLRTA	-0.885		
LLRGL	-0.874		
ROA	0.778		
ROE			
TLTD			0.872
TDTA			-0.779

2009 Variables	F1	F2	F3	F4
EQTA			-0.726	
EQTL			-0.833	
NPLTA	-0.816			
NPLGL	-0.813			
LLRTA	-0.909			
LLRGL	-0.916			
ROA	0.800			
ROE				-0.984
TLTD		-0.911		
TDTA		0.582	0.630	

2010 variables	F1	F2	F3
EQTA	0.528		0.682
EQTL		-0.672	0.590
NPLTA	-0.845		
NPLGL	-0.843		
LLRTA	-0.903		
LLRGL	-0.907		
ROA	0.775		
ROE			
TLTD		0.901	
TDTA			-0.815

2011 Variables	F1	F2	F3
EQTA			0.755
EQTL		-0.558	0.711
NPLTA	-0.890		
NPLGL	-0.890		
LLRTA	-0.887		
LLRGL	-0.888		
ROA	0.735		
ROE			
TLTD		0.923	
TDTA			-0.758

Table 7 continued

2012 Variables	F1	F2	F3	F4	2013 Variables	F1	F2	F3	F4
EQTA			0.740		EQTA		0.554	0.658	
EQTL		0.798			EQTL		0.891		
NPLTA	−0.817				NPLTA	0.943			
NPLGL	−0.828				NPLGL	0.945			
LLRTA	−0.889				LLRTA	0.614			0.670
LLRGL	−0.902				LLRGL	0.584			0.692
ROA					ROA			0.570	
ROE				0.866	ROE				
TLTD		−0.848			TLTD		−0.688	0.590	
TDTA			−0.811		TDTA			−0.697	−0.650

Table 8 Factor scores coefficient matrix

Variables	EQTA	EQTL	NPLTA	NPLGL	LLRTA	LLRGL	ROA	ROE	TLTD	TDTA
2008										
F1	−0.091	−0.130	−0.430	−0.437	−0.464	−0.473	0.424	0.237	−0.024	0.098
F2	−0.642	−0.723	0.030	−0.005	−0.002	−0.107	0.122	0.018	0.212	0.245
F3	0.156	−0.142	0.008	−0.024	0.131	0.044	−0.013	0.065	0.728	−0.648
2009										
F1	−0.047	−0.223	−0.407	−0.414	−0.584	−0.623	0.405	−0.055	0.011	0.337
F2	−0.054	0.317	−0.022	0.025	−0.122	−0.033	−0.074	0.008	−0.796	0.513
F3	−0.553	−0.747	−0.010	−0.032	−0.197	−0.282	0.012	0.035	0.149	0.599
F4	0.028	0.038	0.119	0.084	−0.068	−0.099	0.062	−0.984	0.023	0.035
2010										
F1	−0.011	−0.205	−0.434	−0.440	−0.562	−0.600	0.393	0.069	0.025	0.315
F2	−0.096	−0.568	0.010	−0.051	0.051	−0.083	0.094	0.002	0.784	−0.226
F3	0.517	0.515	0.032	0.028	0.239	0.276	0.009	0.034	0.212	−0.791
2011										
F1	−0.035	−0.153	−0.492	−0.505	−0.493	−0.520	0.346	0.095	−0.012	0.318
F2	−0.004	−0.476	0.026	−0.048	0.090	−0.053	0.114	0.145	0.784	−0.334
F3	0.577	0.599	0.104	0.130	0.110	0.167	0.046	−0.002	0.056	−0.707
2012										
F1	−0.071	−0.092	−0.415	−0.436	−0.586	−0.638	0.090	−0.206	0.023	0.194
F2	0.194	0.643	−0.036	0.063	−0.101	0.147	−0.104	0.008	−0.705	0.140
F3	0.573	0.320	0.115	0.100	0.064	0.082	0.259	−0.115	0.300	−0.723
F4	0.007	−0.078	−0.240	−0.213	0.276	0.303	0.279	0.866	−0.059	0.144

Table 8 continued

Variables	EQTA	EQTL	NPLTA	NPLGL	LLRTA	LLRGL	ROA	ROE	TLTD	TDTA
2013										
F1	0.018	0.067	0.670	0.673	0.192	0.180	−0.226	0.349	0.036	−0.185
F2	0.386	0.685	−0.039	0.074	−0.133	0.179	−0.034	0.029	−0.581	0.026
F3	0.479	0.090	0.010	−0.030	0.155	0.072	0.436	0.077	0.514	−0.547
F4	0.052	−0.047	−0.243	−0.236	0.506	0.532	0.186	−0.674	−0.027	0.158

Discriminant analysis assumes the normality of the underlying structure of the data for each group. The proposed procedure is as follow:

1. Estimate the D-score of each bank via the Eq. (2).
2. Calculate the cut-off score.
3. Classify banks according to the optimal cut-off.

Table 9 report statistics for CDA model. Statistics show that discriminant function does not allow easy identification of banks status in 2008 and 2013. Indeed, eigenvalue's for 2008 and 2013 are relatively low. Furthermore, statistics of Wilk's lambda, which corresponds to the total variance in the discriminant scores not explained by differences among the groups, confirm this result. Wilk's values for 2008 and 2013 are respectively 83.27, and 76.9% .

However, for 2012, a high eigenvalue (0.9115) shows the ability of discriminant function to differentiate (AB) and (DB). The model explains only 47.68% (Square canonical R) of the variance between the two classes. The square canonical correlation coefficient testifies the weak association between the discriminant scores and the set of independent variables (among 34%).

The linear combination of the factors scores provide for each bank a D-score as below:

$$D_{score2008} = 0.956F_1 - 0.337F_2 + 0.207F_3$$
$$D_{score2009} = 0.893F_1 - 0.261F_2 - 0.598F_3 - 0.222F_4$$
$$D_{score2010} = 0.930F_1 + 0.607F_3$$
$$D_{score2011} = 0.957F_1 + 0.162F_2 + 0.543F_3$$
$$D_{score2012} = 0.896F_1 + 0.555F_3 + 0.585F_4$$
$$D_{score2013} = -0.828F_1 + 0.201F_2 + 0.583F_3 - 0.367F_4 \tag{3}$$

Tables 10 and 11 display standardized canonical discriminant functions and correlations between predictor variables.

Over all there are no surprises in score's factors. In 2008 for example, F1 and F3 where positively related to the D-score's bank. Clearly, good asset quality improves profitability. In addition, a sufficient level of liquidity helps bank to be able to perform its score and its ranking so it promotes in (AB) group. In 2009, according to the structure of the correlation matrix (see Table 11) we retain all the 4 Factors. F2, F3 and F4 are negatively related to the score. This means that the rise of these later will reduce the bank's score. Indeed, a low level of liquidity coupled with a low level of funds and low profitability reduces the bank's score.

4.1.2 Logit Regression

In this section, we propose the validation of the logit model which is considered as one of the most commonly applied parametric failure prediction models in both the academic literature, as well as in the banking regulation and supervision. Logit model is based on a binomial regression and is based on the estimation of the probability of

Table 9 Statistics of the estimated CDA model

	2008	2009	2010	2011	2012	2013
Eigenvalue	0.2009	0.5202	0.5466	0.7229	0.9115	0.3004
Canonical R	0.4090	0.5850	0.5945	0.6478	0.6905	0.4806
Wilks lambda	0.8327	0.6578	0.6466	0.5804	0.5232	0.7690
CHI-2	181.3229	402.9032	374.3427	413.7364	454.7975	177.5493
df	3	4	3	3	4	4
p value	0	0	0	0	0	0
Sq canonical corr.	0.1673	0.3422	0.3534	0.4196	0.4768	0.2310

Table 10 Canonical discriminant function

Standardized coefficients	2008	2009	2010	2011	2012	2013
F1	0.956*	0.893*	0.930*	0.957*	0.896*	−0.828*
F2	−0.337*	−0.261*	−0.061	0.162*	0.003	0.201*
F3	0.207*	−0.598*	0.607*	0.543*	0.555*	0.583*
F4	–	−0.222*	–	–	0.585*	−0.367*

* Significant variables according to Factor structure matrix correlation

failure $P(Z)$. This probability is defined as a linear function of a vector of covariates Z_i and a vector of regression coefficients β_i:

$$Z_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \tag{4}$$

In this study, the logistic regression model used the dependent variable Y_i which takes a value of 1 ($Y = 1$) when a failure occurred in a predefined period following the date at which the financial statement data are determined. If not, Y_i, takes on a value of 0 ($Y = 0$) when no failure occurred.

The relationship between the dependent variable and the predictor variables is expressed as follows:

$$\begin{cases} P(Y = 1) = P(Z) = \frac{1}{(1 + \exp^{(-Z)})} \\ P(Y = 0) = 1 - P(Z) = \frac{1}{(1 + \exp^{(Z)})} \end{cases} \tag{5}$$

In our analysis we consider the factors determined from the (PCA) as explanatory variables. After estimating the coefficients of the Logit model, we obtain the score of each bank Z_a:

$$Z_a = \beta_0 + \sum_{i=1}^{n} \beta_i F_i \tag{6}$$

Subsequently, we determine the probability of default of each bank.

The estimated probability of default allows the reallocation of each bank to a specific risk class. Subsequently, a threshold P^* is set to enable segregation between banks

Table 11 Factor structure matrix—correlations

		F1	F2	F3	F4
2008	Total	0.933	−0.277	0.177	–
	Within	0.921	−0.254	0.162	–
2009	Total	0.816	−0.214	−0.489	−0.183
	Within	0.754	−0.175	−0.414	−0.149
2010	Total	0.858	−0.059	0.492	–
	Within	0.802	−0.047	0.414	–
2011	Total	0.895	0.113	0.428	–
	Within	0.837	0.087	0.340	–
2012	Total	0.788	−0.018	0.453	0.448
	Within	0.680	−0.013	0.345	0.341
2013	Total	−0.786	0.173	0.502	−0.274
	Within	−0.744	0.153	0.453	−0.242

and the allocation of these to one of two classes. If the estimated default probability is greater than $P*$ then the bank will be considered bankrupted. Conversely, if the estimated default probability is lower than $P*$, it will be considered active. Most previous work consider a bank as defaulted when its default probability default probability is greater than or equal to 0.5.

$$Z_{score2008} = -6.560 - 0.651F_1 + 3.155F_2$$
$$Z_{score2009} = -4.394 - 0.697F_1 + 2.232F_3 + 0.307F_4$$
$$Z_{score2010} = -4.804 - 0.870F_1 + 1.811F_2 - 2.313F_3$$
$$Z_{score2011} = -6.901 - 1.100F_1 + 1.037F_2 - 2.858F_3$$
$$Z_{score2012} = -7.870 - 0.847F_1 - 2.236F_3 - 0.866F_4$$
$$Z_{score2013} = -8.559 + 0.379F_1 - 3.584F_2 - 1.428F_3 \tag{7}$$

From Tables 12 and 13, we can confirm that for 2008, the level of R^2 is relatively low (48%) and only F1 and F2 are significant. For 2009, only F1, F3 and F4 are significant with a value of R^2 equal to 51.67%. For 2010 and 2011, all factors are significant. For 2012, the model is very satisfactory with R^2 value close to 79.83%. For the last year, the quality of the regression is good. All of F1, F2 and F3 are significant at the 1% level.

For all the years, except for 2013, the component of the Asset Quality (F1) is negatively related to the score of the bank. A good Asset Quality (an increase of the F1) will reduce the probability of default of the bank. For 2013, the positive effect of the F1 on the score means that the rise of this factor (bad asset quality) will penalize the bank with a high probability of default.

For 2012, all the significant factors are negatively related to the score of the bank. This means that an improvement in the asset quality, a better profitability, a high level of equity and a sufficient level of liquidity will reduce the probability of default.

Table 12 Significance tests of factors

Attribute	Coef.	SD	Wald	Signif.
2008				
Constant	−6.560*	0.682	92.661	0
F1	−0.651*	0.091	50.953	0
F2	3.155*	0.544	33.683	0
F3	−0.31	0.17	3.35	0.067
2009				
Constant	−4.394*	0.337	170.486	0
F1	−0.697*	0.074	87.692	0
F2	0.031	0.228	0.018	0.893
F3	2.232*	0.268	69.281	0
F4	0.307**	0.156	3.873	0.049
2010				
Constant	−4.804*	0.406	140.244	0
F1	−0.870*	0.087	100.349	0
F2	1.811*	0.318	32.425	0
F3	−2.313*	0.257	81.197	0
2011				
Constant	−6.901*	0.899	58.932	0
F1	−1.100*	0.156	49.814	0
F2	1.037**	0.423	6.015	0.014
F3	−2.858*	0.457	39.168	0
2012				
Constant	−7.870*	1.396	31.799	0
F1	−0.847*	0.164	26.742	0
F2	−0.407	0.694	0.345	0.557
F3	−2.236*	0.566	15.613	0
F4	−0.866*	0.265	10.677	0.001
2013				
Constant	−8.559*	1.672	26.215	0
F1	0.379*	0.134	7.991	0.005
F2	−3.584*	1.18	9.234	0.002
F3	−1.428*	0.354	16.285	0
F4	0.083	0.18	0.212	0.646

* Significant at 1%
** Significant at 5%

Table 13 Statistical tests of logit models

Model	2008	2009	2010	2011	2012	2013
Chi-2	123.8316	325.3991	333.9415	272.4553	171.7757	66.5756
d.f.	3	4	3	3	4	4
P (>Chi-2)	0	0	0	0	0	0
McFadden's R	0.4854	0.5167	0.5583	0.6803	0.7713	0.6387

Table 14 Confusion matrix

	Predicted class	
	Bankruptcy Y = 1	Non-bankruptcy Y = 0
Actual class		
Bankruptcy Y = 1	Bankrupt banks correctly classified	Bankrupt banks incorrectly classified (Type I error)
Non-bankruptcy Y = 0	Non-bankrupt banks classified as bankrupted ones (Type II error)	Non-bankrupt banks correctly classified

For 2010 and 2011, the factor F3 was negatively related to the Z-score. In fact, a high level of equity decreases the probability of default of the bank. To sum up, all variables (ratios) have the expected signs.

4.2 The Predictive Performance of Models According to the Cut-Off Value

To evaluate the prediction performance of CDA and LR models, we consider type I/II error rates. The later can be measured by a confusion matrix shown in Table 14. This matrix summarizes the correct and incorrect classifications that models produced from our dataset. Rows and columns of the confusion matrix correspond to the true and predicted classes respectively. The type I error is the error of not rejecting a null hypothesis when the alternative hypothesis is the true state of nature. This latter concern the prediction error of the classifier which incorrectly classifies the bankrupted bank into non-bankrupted bank. Thus, type II error presents the rate of prediction errors of a classifier to incorrectly classify healthy bank into bankrupted bank. As consequence, a natural criterion for judging the performance of a classifier is the probability for making a misclassification error.

We consider an early warning model as good when it delivers a low probability of committing type I error and avoid classifying a failed bank into the group of healthy banks.

In this paper, we carry out also the predictive ability of classification methods via the construction of so-called Receiver Operating Characteristics (ROC) curves. (ROC) curve shows the relation between specificity[1] and sensitivity[2] of the given test or detector for all allowable values of threshold (cut-off).

We propose to improve the performance of the clustering by using both Area under the ROC curve (AUC) and H-measure. We note that the Hand ratio (Hand 2009) deals to avoid drawback of AUC.

---

[1] The specificity represents the number of the actives banks classified in the group of the actives banks.

[2] The sensitivity represents the percentage of the bankrupt banks correctly classified.

4.2.1 Prediction Accuracy of the Traditional D* Versus the Minimization of Errors (C^*_{CDA}) Cut-Off

For the prediction accuracy of the (CDA) model, we follow two approaches to select the best cut-off score.

In the first one, we calculate D*.

In the available literature and according to Canbas et al. (2005), the default cut-off value in the two class classifiers is approximately equal to zero and computed by the equation below:

$$Cut-Off = \frac{(N_1 D_1 + N_0 D_0)}{(N_1 + N_0)} \tag{8}$$

where

- N_1: number of bankrupted bank.
- D_1: average score for bankrupted bank.
- N_0: number of non-bankrupted bank.
- D_0: average score for non-bankrupted bank.

However, if the two classes are asymmetric and have different sizes, the optimal cutting score for a discriminant function is the weighted average of the group centroids (Hair et al. 2010). The formula for calculating the critical score between the two groups is:

$$Cut - Off = \frac{(N_A Z_B + N_B Z_A)}{(N_A + N_B)} \tag{9}$$

where Z_A and Z_B are the centroids for group A and B and N_A and N_B are the number of banks in each group. This formula is adopted in our paper for the (CDA) analysis.

The second methodology to select the optimal cut-off score is based on the Receiver operating characteristics (ROC curve) graphs.

Then, we classify banks in failed or healthy group according to the comparison between D_{score} and the cut-off score (D* or C^*_{CDA}):

- If $D_{score} >$ cut-off , the bank is classified in the (AB) group.
- If $D_{score} <$ cut-off , the bank is classified in the (DB) group.

Table 15 presents the classification results obtained by using the cut-off "D*". In the (TRS) of 2008, we note that 21.43% of the (DB) and 7.04% of the (AB) banks are misclassified. We obtain the same result for the (TES) of 2008 with a type I error around 22% and a type II error about 7%.

In 2009, we obtain relatively low results in terms of correct classification rate[3] [88.61% for (TRS) and 85.12% for (TES)]. Indeed, for (TRS), the CDA model classifies correctly 76.29% of the (DB) and 89.99% of the (AB). However, for the (TES), the model classifies 44.12% of the (DB) in the group of (AB).

The results of 2010 are much better in terms of type I error: 13.68% of (AB) in (TRS) and only 3.7% of (DB) in (TES) are misclassified by the CDA model.

For (TRS) in 2011, 2012 and 2013, we observe an improvement of the correct classification rate (resp. 93.72, 98.44 and 98.24%) and a decrease of the type II error

---

[3] The correct classification rate corresponds to the number of all the banks correctly classified.

Table 15 Classification results with $P^* = 0.5$ and D^*

	2008		2009		2010		2011		2012		2013	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
LOGIT												
Sensitivity (%)	28.57	66.67	55.67	29.41	67.37	74.07	67.86	81.25	73.08	57.14	40.00	80.00
Type I error (%)	71.43	33.33	44.33	70.59	32.63	25.93	32.14	18.75	26.92	42.86	60.00	20.00
Specificity (%)	99.79	99.58	98.27	98.08	97.26	96.28	98.02	98.29	99.26	98.82	99.55	100.00
Type II error (%)	0.21	0.42	1.73	1.92	2.74	3.72	1.98	1.71	0.74	1.18	0.45	0.00
Error rate (%)	2.21	1.61	6.00	11.57	6.03	6.51	4.19	3.14	1.7	2.82	1.32	0.59
Correct classification (%)	97.79	98.39	94.00	88.43	93.97	93.49	95.81	96.86	98.3	97.18	98.68	99.41
CDA												
D*	−2.55		−1.92		−1.84		−2.78		−4.69		−4.41	
Sensitivity (%)	78.57	77.78	76.29	55.88	86.32	96.3	82.14	81.25	73.08	57.14	60.00	100.00
Type I error (%)	21.43	22.22	23.71	44.12	13.68	3.70	17.86	18.75	26.92	42.86	40.00	0.00
Specificity (%)	92.96	92.89	89.99	89.9	90.61	88.83	94.63	93.71	99.41	98.24	98.81	100.00
Type II error (%)	7.04	7.11	10.01	10.1	9.39	11.17	5.37	6.29	0.59	1.76	1.19	0.00
Error rate (%)	7.44	7.66	11.39	14.88	9.86	10.23	6.28	7.33	1.56	3.39	1.76	0.00
Correct classification (%)	92.56	92.34	88.61	85.12	90.14	89.77	93.72	92.67	98.44	96.61	98.24	100.00

(resp. 5.37, 0.56 and 1.19%). However, the discriminant function delivers a high type I error that reaches 40% in 2013.

For (TES) in 2012, 42.86% of (DB) was classified as (AB) (type I error). On the opposite, the CDA model in 2013 was able to classify correctly all the banks (sensitivity and specificity equal to 100%).

Table 16 displays results according to the best cut-off obtained under the ROC curve. We remind that optimal critical point corresponds to the value which minimizes both the type I error [(DB) classified in (AB)] and the type II error [(AB) classified in (DB)]. It is also the value which makes it possible to maximize the sensitivity [correctly classified (DB)] and the specificity [(AB) correctly assigned to the group of (AB)].

Results show an improvement of correct classification rate. For (TES) in 2008, we observe that the discriminating function was able to classify correctly 97.18% of the banks in their appropriate groups (versus 92.34% by using D*). By raising the cut off (passing from -1.84 to -3.35), 94.42% (against 89.77%) of the banks in the (TES) in 2010 were correctly classified.

We note that CDA model achieved a 100% of correct classification (TES) in 2013 by using optimal ROC cut-off.

When we use C^*_{CDA}, we observe that the CDA model has failed in classifying the bankrupt banks correctly. We note for (TES) a high level of type I error (66.67, 76.47, 29.63, 31.25 and 42.86% for respectively 2008, 2009, 2010, 2011 and 2012). This result is obvious. Actually when we rise the cut-off value we will penalize the (DB) with a score lower than the optimal C^*_{CDA}.

To sum up, with the canonical discriminant analysis achieved with the D* (Table 17), we obtain similar results on average over the period 2008-2013 for (TRS) and (TES). Indeed, results show that the model classifies correctly 93.62% for (TRS) and 92.75% for (TES). Moreover, 94.40% of (AB) in (TRS) and 93.93% of (AB) were correctly classified (type II error: 5.6% for (TRS) and 6.07% for (TES)). However, we note a high level of type I error, which means that the model was failed to classify correctly 23.93% of the (DB) in (TRS) and 21.94% of (DB) in (TES).

The result of the CDA when using the C^*_{CDA} show that the model obtain on average for the (TES) a correct classification rate about 94.97%. Moreover, we note a low type II error [on average 1.67% in the (TES)]. By against, the CDA model was failed to classify correctly on average 41.15% in the (TES) and 52.75% in the (TRS) of the (DB).

By analyzing the period (T1) which cover 2008, 2009 and 2010 and (T2) spanning between 2011 and 2013, we obtain different results. Indeed, the first period was characterized by a relatively lower correct classification rate [90.44% in (TRS) and 89.08% in (TES) Vs 96.80% in (TRS) and 96.43% in (TES)]. This result is explained by the fact that the high level of the type I error was added with a relatively high type II error. In other words, the model of the "first period" misclassify a large number of (AB) [8.81% in (TRS) and 9.46% in (TES) versus 2.38% in (TRS) and 2.68% in (TES)]. The same tendency was observed by analyzing the CDA model with a cut-off of C^*_{CDA}.

Table 16 Classification results with ROC curve C^*_{Logist} C^*_{CDA}

LOGIT	2008		2009		2010		2011		2012		2013	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
C^*_{Logist}	0.435	0.513	0.499	0.222	0.623	0.618	0.523	0.476	0.385	0.200	0.764	0.552
Sensitivity (%)	35.71	66.67	56.7	61.76	57.89	70.37	66.07	87.5	84.62	85.71	40.00	80.00
Type I error (%)	64.29	33.33	43.3	38.24	42.11	29.63	33.93	12.5	15.38	14.29	60.00	20.00
Specificity (%)	99.79	99.58	98.27	93.27	98.83	98.4	98.31	98.29	99.12	98.24	100.00	100.00
Type II error (%)	0.21	0.42	1.73	6.73	1.17	1.6	1.69	1.71	0.88	1.76	0.00	0.00
Error rate (%)	2.01	1.61	5.9	11.16	5.68	5.12	4.06	2.62	1.42	2.26	0.88	0.59
Correct classification (%)	97.99	98.39	94.1	88.84	94.32	94.88	95.94	97.38	98.58	97.74	99.12	99.41
CDA												
C^*_{CDA}	−6.85	−8.56	−4.50	−3.26	−3.85	−3.35	−3.91	−3.48	−4.68	−6.56	−4.72	−2.73
Sensitivity (%)	21.43	33.33	26.8	23.53	37.89	70.37	64.29	68.75	73.08	57.14	60.00	100.00
Type I error (%)	78.57	66.67	73.2	76.47	62.11	29.63	35.71	31.25	26.92	42.86	40.00	0.00
Specificity (%)	99.48	99.58	99.31	97.12	97.26	97.87	97.46	96	99.41	99.41	99.25	100.00
Type II error (%)	0.52	0.42	0.69	2.88	2.74	2.13	2.54	4.00	0.59	0.59	0.75	0.00
Error rate (%)	2.72	2.82	7.97	13.22	9.28	5.58	4.97	6.28	1.56	2.26	1.32	0.00
Correct classification (%)	97.28	97.18	92.03	86.78	90.72	94.42	95.03	93.72	98.44	97.74	98.68	100.00

Table 17 Results in average

	2008–2013				2008–2010				2011–2013			
	LR	LR*	CDA	CDA*	LR	LR*	CDA	CDA*	LR	LR*	CDA	CDA*
Training sample												
Sensitivity (%)	55.42	56.83	76.07	47.25	50.54	50.1	80.39	28.71	60.31	63.56	71.74	65.79
Type I error (%)	44.58	43.17	23.93	52.75	49.46	49.9	19.61	71.29	39.69	36.44	28.26	34.21
Specificity (%)	98.69	99.05	94.4	98.7	98.44	98.96	91.19	98.68	98.95	99.14	97.62	98.71
Type II error (%)	1.31	0.95	5.6	1.3	1.56	1.04	8.81	1.32	1.05	0.86	2.38	1.29
Error rate (%)	3.58	3.33	6.38	4.64	4.75	4.53	9.56	6.66	2.4	2.12	3.2	2.62
Correct classification (%)	96.42	96.67	93.62	95.36	95.25	95.47	90.44	93.34	97.6	97.88	96.8	97.38
Testing sample												
Sensitivity (%)	64.76	75.34	78.06	58.85	56.72	66.27	76.65	42.41	72.8	84.4	79.46	75.3
Type I error (%)	35.24	24.66	21.94	41.15	43.28	33.73	23.35	57.59	27.2	15.6	20.54	24.7
Specificity (%)	98.51	97.96	93.93	98.33	97.98	97.09	90.54	98.19	99.04	98.84	97.32	98.47
Type II error (%)	1.49	2.04	6.07	1.67	2.02	2.91	9.46	1.81	0.96	1.16	2.68	1.53
Error rate (%)	4.37	3.89	7.25	5.03	6.56	5.96	10.92	7.21	2.18	1.82	3.57	2.85
Correct classification (%)	95.63	96.11	92.75	94.97	93.44	94.04	89.08	92.79	97.82	98.18	96.43	97.15

LR* Results of the logistic regression with C^*_{ROC}
CDA* Results of the canonical discriminant analysis with C^*_{ROC}

4.2.2 Prediction Accuracy of the Traditional P*= 0.5 Versus the Minimization of Errors (C^*_{Logit}) Cut-Off

For the prediction accuracy of the Logit model, we firstly compared the probability of default obtained from the scoring function with the $P^* = 0.5$.

After, we use the best cut-off point which minimizes the overall error (sum of type I and II errors).

- If Probability of default $<P^*$ or C^*_{Logit}, the bank is classified in (AB) group.
- If Probability of default $>P^*$ or C^*_{Logit}, the bank is classified in (DB) group.

Table 15 display results based on a probability of default "P^*" equal to 0.5. For 2008, we obtain a high correct classification rate. But we note a high level of type I error which means that the logistic regression misclassifies 71.43% (DB) in (TRS) and 33.33% of (DB) in (TES).

For (TES) in 2009, the model achieves its low level with a rate of correct classification around 88.43%. From 2010 to 2013, we observe an improvement in terms of correct classification rate (passing from 93.49 to 99.41% in the (TES))and a decrease of the type II error (3.72, 1.71, 1.18 and 0% respectively in 2010, 2011, 2012 and 2012). In contrast, the logit model was not able to classify correctly (DB). We note a high type I error in the (TES) about 33.33, 70.59, 25.93, 18.75, 42.86 and 20% respectively in 2008, 2009, 2010, 2011, 2012 and 2013).

Table 16 summarizes results with C^*_{Logit} in (TRS) and (TES). For all the years, we note a slight improvement in the correct classification rate. But we note a decrease in the type I error.

For (TES) in 2009, results with a probability of default equal to 0.222 calculated from the ROC curve allow to reduce the type I error by 32.35% (38.24% against 70.59% with a probability of default of 0.5). We also obtain a high type II error with a rate around 6.73% (against 1.92%). In 2012, we note approximately the same trend of decline in the type I error. A lower probability of default (0.2) allows the Logit model to classify correctly 85.71% of (DB) (against 57.14% of (DB) with $P^* = 0.5$). The results of 2011 show also a decrease in the type I error (12.5% against 18.75%) and an improvement of the correct classification rate (97.38 versus 96.86% for $P^* = 0.5$).

According to the results in Table 17 we observe that the Logistics Regression obtained with a probability of default equal to 0.5 delivers a satisfactory overall result in terms of correct classification rate [on average 96.42% for the (TRS) and 95.63% for the (TES)]. We also note that the model obtains a lower type II error with on average value of 1.31% in (TRS) and 1.49% for (TES). However, we obtain a higher type I error [44.58% for the (TRS) and 35.24% for the (TES)].

The results for (TES) of the Logistic Regression, using a C^*_{Logit}, show an improvement of the correct classification rate (on average 96.11% versus 95.63% for LR with P^*) and a decrease of the type I error (on average 24.66 versus 35.24% for LR with P^*). But we note a slight increase in the type II error (on average 2.04 versus 1.49% for LR with P^*).

When we consider the two periods T1 (2008–2010) and T2 (2011–2013), results for T2 are better in term of correct classification regardless of the choice of the cut-off and model. Also, error type I and II still the lowest in (TES) in the second period.

Conclusion

To conclude, we can summarize the results as follow:

1. The Logit model with $P*$ outperforms the CDA model with D* in terms of correct classification rate. Indeed, on average in the (TRS) we report a correct classification rate about 96.42% for the Logit versus 93.62% for the CDA. For (TES), the Logistic model classifies correctly 95.63% of the banks. We record only 92.75% CDA.
2. In terms of the type II error, the Logit is more efficient than the CDA model with an average type II error for (TRS) about 1.31% and 1.49% for (TES) [against 5.60% for (TRS) and 6.07% for (TES)]. By contrast, type I error for the CDA model outperforms the Logit.
3. C_{ROC}^* improves the accuracy of classification in both Logit and CDA . For example, an average rate results of (TES) equal to 96.11% for the LR and 94.97% for the CDA model (against 95.63 and 92.75%). This result confirm the supremacy of the Logistic Regression.
4. The Logit model with C_{logit}^* allows to reduce the type I error (on average 24.66 versus 35.24% for Logit with $P*$). However, with C_{CDA}^* the classifier has failed to classify correctly 41.15% of the DB (against a type I error of 21.94% for the CDA with D*).

The Table 18 exhibits the *P* value of a Student's *t* test for difference between the correct classification rates achieved with the optimal cut-off C_{ROC}^*. The results of the *P* value show that these differences are, in average, significant and that the Logit and the CDA models do not have a statistical identical accuracy.

4.3 Performance of Predictability of Models with Cost

4.3.1 Cost of Misclassification

To assess model's performance we introduce the ratio of misclassification costs. In bankruptcy forecasting literature, the cost of a type I error is largely greater than the cost of type II error (Balcaen and Ooghe 2006). This means that misclassifying a (DB) is more costly than misclassifying (AB).

We consider several cost scenarios. Lets $Cost_{II}$ the value of the numerator of the misclassification ratio which corresponds to the cost of misclassification of (AB): i.e C_{II}: cost of type II error. This value will be constant and equal to 1. We define also $Cost_I$ as the misclassifying (DB) cost (C_I: cost of type I error). We test several parameters: 2, 5, 10, 20, 30, 40 and 50 (Etheridge 2015; Gepp and Kumar 2015; Jardin 2015, 2016; Jardin and Séverin 2012; Frydman et al. 1985).

Tables 19 and 20 summarize percentage of misclassified banks (type I error, type II error and total error) by each model with different penalty's parameters. The main important results of (TES) are as follow:

1. The total error rate increases when the misclassification costs of the bankrupt banks increase. Indeed, for a low parameter of $Cost_I = 2$ Logit model achieved an error rate varying between 0.59 and 11.16%, and the CDA achieved an error

Table 18 Student's t tests for differences between correct classification rates achieved with C^*_{ROC}

Training sample				Testing sample			
Model	LOGIT	CDA	P value	Model	LOGIT	CDA	P value
2008	97.99	97.28	0.002	2008	98.39	97.18	0.004
2009	94.1	92.03	0.007	2009	88.84	86.78	0.007
2010	94.32	90.72	0.012	2010	94.88	94.42	0.002
2011	95.94	95.03	0.003	2011	97.38	93.72	0.012
2012	98.58	98.44	0.000	2012	97.74	97.74	0.000
2013	99.12	98.68	0.001	2013	99.41	100.00	0.002

Table 19 Results of CDA with cost

C_I	Training sample						Testing sample					
	2008	2009	2010	2011	2012	2013	2008	2009	2010	2011	2012	2013
CDA												
2												
C^*_{CDA}	−4.19	−2.34	−1.86	−3.42	−3.68	−4.72	−7.62	−1.63	−2.38	−2.57	−3.09	−2.73
Type I (%)	39.29	31.96	13.68	26.79	11.54	40.00	55.56	29.41	11.11	12.50	0.00	0.00
Type II (%)	2.48	7.02	9.00	3.39	1.76	0.75	0.84	12.50	6.91	6.29	4.12	0.00
Total error (%)	3.52	9.52	9.51	5.10	2.12	1.32	2.82	14.88	7.44	6.81	3.95	0.00
5												
C^*_{CDA}	−4.19	−1.74	−1.12	−1.66	−3.33	−3.54	−3.15	−0.25	−1.92	−2.12	−3.09	−2.73
Type I (%)	39.29	20.62	5.26	0.00	7.69	20.00	22.22	5.88	3.70	6.25	0.00	0.00
Type II (%)	2.48	10.93	14.08	8.62	2.35	1.49	3.77	22.12	10.11	8.57	4.12	0.00
Total error (%)	3.52	11.90	13.11	7.98	2.55	1.76	4.44	19.83	9.30	8.38	3.95	0.00
10												
C^*_{CDA}	−2.30	−0.07	−1.12	−1.66	−2.90	−3.54	−1.43	−0.04	−0.77	−1.24	−3.09	−2.73
Type I (%)	14.29	2.06	5.26	0.00	3.85	20.00	0.00	2.94	0.00	0.00	0.00	0.00
Type II (%)	8.18	25.43	14.08	8.62	3.24	1.49	12.13	25.96	14.36	12.00	4.12	0.00
Total error (%)	8.35	23.08	13.11	7.98	3.26	1.76	11.69	22.73	12.56	10.99	3.95	0.00

Table 19 continued

C_I	Training sample						Testing sample					
	2008	2009	2010	2011	2012	2013	2008	2009	2010	2011	2012	2013
20												
C^*_{CDA}	−1.97	−0.07	−0.82	−1.66	−2.43	−2.37	−1.43	−0.04	−0.77	−1.24	−3.09	−2.73
Type I (%)	10.71	2.06	4.21	0.00	0.00	10.00	0.00	2.94	0.00	0.00	0.00	0.00
Type II (%)	9.32	25.43	16.04	8.62	4.85	4.33	12.13	25.96	14.36	12.00	4.12	0.00
Total error (%)	9.36	23.08	14.73	7.98	4.67	4.41	11.69	22.73	12.56	10.99	3.95	0.00
30												
C^*_{CDA}	−1.97	−0.07	−0.82	−1.66	−2.43	−1.51	−1.43	−0.04	−0.77	−1.24	−3.09	−2.73
Type I (%)	10.71	2.06	4.21	0.00	0.00	0.00	0.00	2.94	0.00	0.00	0.00	0.00
Type II (%)	9.32	25.43	16.04	8.62	4.85	8.21	12.13	25.96	14.36	12.00	4.12	0.00
Total error (%)	9.36	23.08	14.73	7.98	4.67	8.09	11.69	22.73	12.56	10.99	3.95	0.00
40												
C^*_{CDA}	−1.97	−0.07	−0.33	−1.66	−2.43	−1.51	−1.43	−0.04	−0.77	−1.24	−3.09	−2.73
Type I (%)	10.71	2.06	3.16	0.00	0.00	0.00	0.00	2.94	0.00	0.00	0.00	0.00
Type II (%)	9.32	25.43	19.95	8.62	4.85	8.21	12.13	25.96	14.36	12.00	4.12	0.00
Total error (%)	9.36	23.08	14.73	7.98	4.67	8.09	11.69	22.73	14.36	10.99	3.95	0.00
50												
C^*_{CDA}	−1.97	−0.07	−0.33	−1.66	−2.43	−1.51	−1.43	−0.04	−0.77	−1.24	−3.09	−2.73
Type I (%)	10.71	2.06	3.16	0.00	0.00	0.00	0.00	2.94	0.00	0.00	0.00	0.00
Type II (%)	9.32	25.43	19.95	8.62	4.85	8.21	12.13	25.96	14.36	12.00	4.12	0.00
Total error (%)	9.36	23.08	14.73	7.98	4.67	8.09	11.69	22.73	12.56	10.99	3.95	0.00

Table 20 Results of logit with cost

C_I	Training sample						Testing sample					
	2008	2009	2010	2011	2012	2013	2008	2009	2010	2011	2012	2013
LOGIT												
2												
C^*_{Logit}	0.33	0.44	0.33	0.27	0.22	0.76	0.51	0.22	0.36	0.47	0.20	0.55
Type I (%)	53.57	37.11	21.05	3.57	7.69	60.00	33.33	38.24	3.70	6.25	14.29	20.00
Type II (%)	0.62	2.65	4.69	4.52	1.32	0.00	0.42	6.73	6.38	2.29	1.76	0.00
Total error (%)	2.11	6.11	6.50	4.45	1.56	0.88	1.61	11.16	6.05	2.62	2.26	0.59
5												
C^*_{Logit}	0.13	0.16	0.19	0.19	0.16	0.09	0.21	0.05	0.36	0.25	0.07	0.10
Type I (%)	25.00	15.46	10.53	1.79	3.85	10.00	11.11	8.82	3.70	0.00	0.00	0.00
Type II (%)	3.11	9.09	7.69	4.80	1.76	1.94	2.93	22.60	6.38	4.57	4.12	1.21
Total error (%)	3.72	9.73	8.00	4.58	1.84	2.06	3.23	20.66	6.05	4.19	3.95	1.18
10												
C^*_{Logit}	0.13	0.07	0.07	0.19	0.16	0.09	0.21	0.05	0.23	0.25	0.07	0.10
Type I (%)	25.00	7.22	4.21	1.79	3.85	10.00	11.11	8.82	0.00	0.00	0.00	0.00
Type II (%)	3.11	15.77	13.04	4.80	1.76	1.94	2.93	22.60	10.11	4.57	4.12	1.21
Total error (%)	3.72	14.91	12.06	4.58	1.84	2.06	3.23	20.66	8.84	4.19	3.95	1.18
20												
C^*_{Logit}	0.04	0.05	0.07	0.06	0.06	0.02	0.21	0.02	0.23	0.25	0.07	0.10
Type I (%)	10.71	6.19	4.21	0.00	0.00	0.00	11.11	0.00	0.00	0.00	0.00	0.00
Type II (%)	7.76	18.76	13.04	6.36	3.38	4.33	2.93	40.87	10.11	4.57	4.12	1.21
Total error (%)	7.85	17.49	12.06	5.89	3.26	4.26	3.23	35.12	8.84	4.19	3.95	1.18

Table 20 continued

C_I	Training sample						Testing sample					
	2008	2009	2010	2011	2012	2013	2008	2009	2010	2011	2012	2013
30												
C^*_{Logit}	0.04	0.03	0.07	0.06	0.06	0.02	0.21	0.02	0.23	0.25	0.07	0.10
Type I (%)	10.71	3.09	4.21	0.00	0.00	0.00	11.11	0.00	0.00	0.00	0.00	0.00
Type II (%)	7.76	24.86	13.04	6.36	3.38	4.33	2.93	40.87	10.11	4.57	4.12	1.21
Total error (%)	7.85	22.67	12.06	5.89	3.26	4.26	3.23	35.12	8.84	4.19	3.95	1.18
40												
C^*_{Logit}	0.04	0.03	0.07	0.06	0.06	0.02	0.21	0.02	0.23	0.25	0.07	0.10
Type I (%)	10.71	3.09	4.21	0.00	0.00	0.00	11.11	0.00	0.00	0.00	0.00	0.00
Type II (%)	7.76	24.86	13.04	6.36	3.38	4.33	2.93	40.87	10.11	4.57	4.12	1.21
Total error (%)	7.85	22.67	12.06	5.89	3.26	4.26	3.23	35.12	8.84	4.19	3.95	1.18
50												
C^*_{Logit}	0.02	0.07	0.07	0.06	0.06	0.02	0.21	0.02	0.23	0.25	0.07	0.10
Type I (%)	7.14	4.21	4.21	0.00	0.00	0.00	11.11	0.00	0.00	0.00	0.00	0.00
Type II (%)	12.42	13.04	13.04	6.36	3.38	4.33	2.93	40.87	10.11	4.57	4.12	1.21
Total error (%)	12.27	12.06	12.06	5.89	3.26	4.26	3.23	35.12	8.84	4.19	3.95	1.18

rate ranging from 0 to 14.88%. For a parameter equal to 50, the Logit provide an error rate varying between 1.18 and 35.12% and the CDA provide an error rate in [0, 22.73%] interval.

2. For a $Cost_I = 10 * Cost_{II}$, the CDA was able to classify correctly all the bankrupt banks (error type I equal to zero). We note that in 2008, when $Cost_I = 5 * Cost_{II}$ with a cut-off value of 0.21, the Logit model was not able to classify correctly 11.11% of the (DB) banks. Moreover, for all the others years, we achieve a type I error equal to zero in Logistic model

3. When the parameter is low all type II errors are also low. For example, in Logit when $Cost_I = 2 * Cost_{II}$, we note an average type II error about 2.93 versus 5.11% in CDA. In contrast, for a high level $Cost_I$, we note an increase in the type II error.

4. When we increase $Cost_I$, we note an improvement in the results. For a cost parameter equal to 20, the Logit model classifies correctly all the bankrupt banks for each year (except for 2008). With a parameter equal to 10, the CDA model classifies correctly all the bankrupt banks (except for 2009). We can conclude, that when we raise the parameter, we can reduce the type I error.

4.3.2 H-Measure

We used the area under the receiver operating characteristic curve (AUC) for the estimation of the performance of the models.

Hand (2009) proposed H-measure as a coherent alternative measure to the AUC. It seeks to quantify the relative severity of one type of error over the other (Jardin 2016; Fitzpatrick and Mues 2016). A high value of H-measure is associated with a better performance (similar to the AUC).

The Table 21 presents the results of the AUC (ROC curve) and the H-measure proposed by Hand (2009). We note that the results achieved by the H-measure are consistent with those obtained with the AUC. In terms of supremacy between CDA and Logit, both performance measures converge to the same results: Logit is more efficient than the CDA.

4.4 Bankruptcy Prediction as a Classification Problem

In this paper we proved that the choice of the cut-off is crucial. If it is low, type I error will increase and as a result crises will be more accurately detected. However, if the type II error increases the number of false alarms will increase but at the same time, the number of false alarms will increase (i.e. the type II error) generating good control in terms of economic policy.

In this section we analyze the impact of letting $Cost_I = 20 * Cost_{II}$ on type II error in order to validate the accuracy of models with misclassification cost (Tables 22, 23).

In testing set, Logit regression failed to classify correctly 7, 85, 19, 8, 7 and 2 actives banks respectively for the years 2008, 2009, 2010, 2011, 2012 and 2013. But an out-of-sample monitoring year by year proves that 100, 54.12, 89.47, 87.5 and 85.71% of

Table 21 Performance of models: AUC and H-measure

	2008		2009		2010		2011		2012		2013	
	Logit	CDA	Logit	CDA	Logit	CDA	Logit	CDA	Logit	CDA	Logit	CDA
Training sample												
AUC	0.946	0.922	0.947	0.929	0.950	0.931	0.983	0.973	0.995	0.992	0.990	0.982
H	0.725	0.684	0.664	0.589	0.738	0.682	0.881	0.821	0.933	0.902	0.907	0.839
AUCH	0.958	0.938	0.953	0.935	0.960	0.940	0.985	0.976	0.996	0.993	0.993	0.987
Gini	0.892	0.843	0.895	0.858	0.900	0.861	0.966	0.945	0.989	0.983	0.980	0.964
KS	0.815	0.800	0.770	0.725	0.828	0.807	0.936	0.914	0.966	0.952	0.957	0.918
MWL	0.005	0.006	0.021	0.025	0.017	0.019	0.004	0.006	0.001	0.002	0.001	0.001
Testing sample												
AUC	0.933	0.965	0.899	0.886	0.976	0.968	0.989	0.963	0.987	0.981	0.998	1
H	0.811	0.765	0.537	0.528	0.834	0.774	0.911	0.775	0.911	0.910	0.973	1
AUCH	0.964	0.974	0.916	0.903	0.983	0.975	0.993	0.970	0.991	0.989	0.999	1
Gini	0.865	0.930	0.799	0.772	0.951	0.936	0.977	0.925	0.973	0.961	0.995	1
KS	0.860	0.879	0.686	0.720	0.899	0.862	0.954	0.880	0.959	0.959	0.988	1
MWL	0.005	0.004	0.038	0.034	0.011	0.015	0.004	0.009	0.002	0.002	0.000	0.000

AUC The area under the ROC curve

H The H-measure

AUCH The area under the convex Hull of the ROC curve

Gini The Gini coefficient

KS The Kolmogorov–Smirnoff statistic

MWL The minimum cost-weighted error rate

Table 22 Analysis of type II error achieved with the logistic regression $C_I = 20$

	2008	2009	2010	2011	2012	2013
Logit/testing sample						
AB classified as DB	7	85	19	8	7	2
AB really bankrupt	7	46	17	7	6	0
AB really bankrupt (%)	100.00	54.12	89.47	87.50	85.71	0.00

	2009	2010	2011	2012	2013	2014	AB really bankrupt
2008	6						7
2009		21	15	8	2		46
2010			10	6	1		17
2011				5	2		7
2012					5	1	6

	N	N + 1	N + 2	N + 3	N + 4
2008		85.71	32.61	14.29	4.35
2009		45.65	35.29	17.39	
2010		58.82	28.57	5.88	
2011		71.43	16.67		
2012		83.33			
Average		68.99	32.16	12.52	4.35

Table 23 Analysis of type II error achieved with the canonical discriminant analysis with $C_I=20$

	2008	2009	2010	2011	2012	2013
CDA/testing sample						
AB classified as DB	29	54	27	21	7	0
AB really bankrupt	29	48	21	10	5	0
AB really bankrupt (%)	100.00	88.89	77.78	47.62	71.43	–

	2009	2010	2011	2012	2013	2014	2015	AB really bankrupt
2008	16	7	3	2		1		29
2009		21	14	9	3		1	48
2010			11	8	2			21
2011				7	3			10
2012					5			5
2013								0

N	N + 1	N + 2	N + 3	N + 4	N + 5	N + 6
2008	55.17	24.14	10.34	6.90		3.45
2009	43.75	29.17	18.75	6.25		2.08
2010	52.38	38.10	9.52			
2011	70.00	30.00				
2012	100.00					
Average	64.26	30.35	12.87	6.25		2.77

the misclassified active banks will go bankrupt in the following years (respectively in 2008, 2009, 2010, 2011 and 2012).

Thus, for all banks that the model identified as (DB), we check over the following years if they will or won't fail. For example, in 2009, among the 46 (AB) classified as a (DB), 21, 15, 8 and 2 banks will go bankrupt respectively in 2010, 2011, 2012 and 2013. Therefore, the logit model was able to detect the failure of banks 1 year, 2 years, 3 years and 4 years before bankruptcy occurrence.

The logit model could predict the failure on average 68.99, 32.16, 12.52 and 4.35% respectively one year, two years, three years and four years before the bankruptcy.

In CDA model, results show that among the 29, 54, 27, 21 and 7 misclassified (AB), 29, 48, 21, 10 and 5 respectively for 2008, 2009, 2010, 2011 and 2012 will really go bankrupt. For example, in 2008, among the 29 misclassified (AB), 16, 7, 3, 2 and 1 banks will fail respectively in 2009, 2010, 2011, 2012 and 2014. The CDA model was able to predict the failure of banks 6 years before the bankruptcy.

The results of 2009 show that the CDA model detects the failure of 43.75, 29.17, 18.75, 6.25 and 2.08% of the misclassified actives banks that will really bankrupt respectively one year, two years, three years, four years and six years before the fall. In 2012, all the misclassified actives banks (5 banks) will go bankrupt the next year (2013).

5 Conclusion

In this paper, we show that the Logistic Regression (LR) and Canonical Discriminant Analysis (CDA) can predict banks failure with accuracy. Main models inputs are CAMEL's Variables of a large panel of US banks over the period from 2008 to 2013.

First, Principal Component Analysis (PCA) was performed to compress the data dimension by keeping only the most important ratio combinations. The results show the importance of the Asset Quality, the Capital Adequacy and the Liquidity as an indicators of the financial conditions of the bank.

We use random subspace method to compare the classification and the prediction accuracy of CDA and LR models, with and without misclassification cost. We compared different cut-off point formulas to provide and evaluate classification accordingly.

Our results confirm, first that the more accurate the critical probability of default value is, the more accurate is the sensitivity of the model. In this sense, comparative results over the entire period prove that correct classification was improved with the C^*_{ROC} for both Logit and CDA model.

The first finding proves that sensitivity of classification was improved and in average Logit model outperforms CDA (in the (TES)75.34 versus 58.85%).

The second finding concerns the supremacy of the ROC curve validation with regard to the quality of the Logit model by minimizing the error of misclassification of (DB) (in the (TES), the type I error is on average 24.66 versus 35.24% with a probability of default $P^* = 0.5$).

Third, in term of correct classification we show that the Logistic Regression is better by using C^*_{Logit} (in the (TES) 96.11% against 95.63% with $P^* = 0.5$).

Then, we evaluate the Logit and the CDA model by introducing a misclassification cost (cost for type I error higher than the cost of type II error). The results show that with a high misclassification cost of type I error we obtain a high level of error rate. Indeed, the classifiers misclassify a large number of (AB) meaning that we obtain a high type II error. By against, the models are able to classify correctly all the (DB) (error type I equal to zero).

Finally, models were used to pick-up early warning signals. Moreover, The combination of the two models allows a better information about the future prospect of banks. Indeed, ROC curve validation emphasizes better prediction of bank failure because it delivers, in average, the highest error type II 2.04%. This means that the model classifies some solvable banks in bankrupt group. Consequently, we can conclude that the Logit was able to predict the failure of banks. It gives good signal on banks, which would failed one or two year later.

Overall, the study reveals also that our choice resulting from combinations of ten financial ratios which represent Capital adequacy, Assets quality, Earnings ability and Liquidity are obvious determinants to predict bankruptcy.

Our results can be used for several purposes. For instance, regulators and banks can predict problems in order to avoid financial distress which can lead to bankruptcy. Our methodological framework helps to construct an Early Warning System that can be used by supervisory authorities to detect banks close to failure state.

Finally, further extensions can take account of non-parametric methods (Trait Recognition Model, Intelligence techniques such as induction of classification trees and Neural Networks methods).

References

Adya, M., & Collopy, F. (1998). How effective are neural networks at forecasting and prediction? A review and evaluation. *Journal of Forecasting*, *17*, 481–495.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, *23*(4), 589–609.

Altman, E. I. (1984). A further empirical investigation of the bankruptcy cost question. *The Journal of Finance*, *39*(4), 1067–1089.

Altman, E. I., & Narayanan, P. (1997). An international survey of business failure classification models. *Financial Markets, Institutions & Instruments*, *6*(2), 1–57.

Avery, R. B., Hanweck, G. A., et al. (1984). *A dynamic analysis of bank failures*. Board of Governors of the Federal Reserve System (US): Technical report.

Aziz, M. A., & Dar, H. A. (2004). *Predicting corporate financial distress: Whither do we stand?*. Department of Economics: Loughborough University.

Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems. *The British Accounting Review*, *38*(1), 63–93.

Balcaen, S., Ooghe, H., et al. (2004). Alternative methodologies in studies on business failure: do they produce better results than the classical statistical methods? *Vlerick Leuven Gent Management School Working Papers Series*, (16).

Barr, R. S., Siems, T. F., et al. (1994). *Predicting bank failure using dea to quantify management quality*. Federal Reserve Bank of Dallas: Technical report.

Barth, J. R., Brumbaugh, R. D., Sauerhaft, D., Wang, G. H., et al. (1985). Thrift institution failures: Causes and policy issues. In *Federal Reserve Bank of Chicago Proceedings*, p 68.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, *4*, 71–111.

Benston, G. J. (1985). *An analysis of the causes of savings and loan association failures*. Salomon Brothers Center for the Study of Financial Institutions: Graduate School of Business Administration, New York University.

Bilderbeek, J. (1979). Empirical-study of the predictive ability of financial ratios in the netherlands. *Zeitschrift fur Betriebswirtschaft*, *49*(5), 388–407.

Canbas, S., Cabuk, A., & Kilic, S. B. (2005). Prediction of commercial bank failure via multivariate statistical analysis of financial structures: The Turkish case. *European Journal of Operational Research*, *166*(2), 528–546.

Coakley, J. R., & Brown, C. E. (2000). Artificial neural networks in accounting and finance: Modeling issues. *International Journal of Intelligent Systems in Accounting, Finance & Management*, *9*(2), 119–144.

Dabos, M., & Sosa-Escudero, W. (2004). Explaining and predicting bank failure using duration models: The case of argentina after the mexican crisis. *Revista de Análisis Económico*, *19*(1), 31–49.

Dimitras, A. I., Zanakis, S. H., & Zopounidis, C. (1996). A survey of business failures with an emphasis on prediction methods and industrial applications. *European Journal of Operational Research*, *90*(3), 487–513.

Du Jardin, P. (2015). Bankruptcy prediction using terminal failure processes. *European Journal of Operational Research*, *242*(1), 286–303.

Du Jardin, P. (2016). A two-stage classification technique for bankruptcy prediction. *European Journal of Operational Research*, *254*(1), 236–252.

Du Jardin, P., & Séverin, E. (2012). Forecasting financial failure using a kohonen map: A comparative study to improve model stability over time. *European Journal of Operational Research*, *221*(2), 378–396.

Etheridge, H. (2015). Minimizing the costs of using models to assess the financial health of banks. *International Journal of Business and Social Research*, *5*(11), 09–18.

Fitzpatrick, T., & Mues, C. (2016). An empirical comparison of classification algorithms for mortgage default prediction: Evidence from a distressed mortgage market. *European Journal of Operational Research*, *249*(2), 427–439.

Frydman, H., Altman, E. I., & KAO, D.-L. (1985). Introducing recursive partitioning for financial classification: The case of financial distress. *The Journal of Finance*, *40*(1), 269–291.

Gepp, A., & Kumar, K. (2015). Predicting financial distress: A comparison of survival analysis and decision tree techniques. *Procedia Computer Science*, *54*, 396–404.

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis: A global perspective*. Upper Saddle River, NJ: Pearson.

Hand, D. J. (2009). Measuring classifier performance: A coherent alternative to the area under the roc curve. *Machine Learning*, *77*(1), 103–123.

Hassan Al-Tamimi, H. A., & Charif, H. (2011). Multiple approaches in performance assessment of UAE commercial banks. *International Journal of Islamic and Middle Eastern Finance and Management*, *4*(1), 74–82.

Jagtiani, J., Kolari, J., Lemieux, C., Shin, H., et al. (2003). Early warning models for bank supervision: Simpler could be better. *Economic Perspectives-Federal Reserve Bank of Chicago*, *27*(3), 49–59.

Jolliffe, I. (2002). *Principal component analysis*. New York: Wiley.

Keasey, K., & Watson, R. (1991). Financial distress prediction models: A review of their usefulness. *British Journal of Management*, *2*(2), 89–102.

Kolari, J., Caputo, M., & Wagner, D. (1996). Trait recognition: An alternative approach to early warning systems in commercial banking. *Journal of Business Finance & Accounting*, *23*(9–10), 1415–1434.

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.

Lanine, G., & Vander Vennet, R. (2006). Failure prediction in the russian bank sector with logit and trait recognition models. *Expert Systems with Applications*, *30*(3), 463–478.

Lau, A. H.-L. (1987). A five-state financial distress prediction model. *Journal of Accounting Research, 25*(1), 127–138.

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

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*(1), 109–131.

O'leary, D. E. (1998). Using neural networks to predict corporate failure. *International Journal of Intelligent Systems in Accounting, Finance & Management*, *7*(3), 187–197.

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

Pantalone, C. C., Platt, M. B., et al. (1987). Predicting commercial bank failure since deregulation. *New England Economic Review, 15*, 37–47.

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

Sinkey, J. F. (1978). Identifying "problem" banks: How do the banking authorities measure a bank's risk exposure? *Journal of Money, Credit and Banking*, *10*(2), 184–193.

Tennyson, B. M., Ingram, R. W., & Dugan, M. T. (1990). Assessing the information content of narrative disclosures in explaining bankruptcy. *Journal of Business Finance & Accounting*, *17*(3), 391–410.

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

Vellido, A., Lisboa, P. J., & Vaughan, J. (1999). Neural networks in business: A survey of applications (1992–1998). *Expert Systems with Applications*, *17*(1), 51–70.

Wong, B. K., Bodnovich, T. A., & Selvi, Y. (1997). Neural network applications in business: A review and analysis of the literature (1988–1995). *Decision Support Systems*, *19*(4), 301–320.

Zavgren, C. V. (1985). Assessing the vulnerability to failure of american industrial firms: A logistic analysis. *Journal of Business Finance & Accounting*, *12*(1), 19–45.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, *14*(1), 35–62.

Zopounidis, C., & Dimitras, A. (1993). The forecasting of business failure: Overview of methods and new avenues. *Applied Stochastic Models and Data Analysis, World Scientific Publications, London*.