Optim Lett (2010) 4:543–558
DOI 10.1007/s11590-010-0182-1

ORIGINAL PAPER

# Modelling banking sector stability with multicriteria approaches

**Chrysovalantis Gaganis · Fotios Pasiouras · Michael Doumpos · Constantin Zopounidis**

Received: 17 December 2009 / Accepted: 21 February 2010 / Published online: 9 March 2010


**Abstract** Banking crises can be damaging for the economy, and as the recent experience has shown, nowadays they can spread rapidly across the globe with contagious effects. Therefore, the assessment of the stability of a county's banking sector is important for regulators, depositors, investors and the general public. In the present study, we propose the development of classification models that assign the banking sectors of various countries in three classes, labelled "low stability", "medium stability", and "high stability". The models are developed using three multicriteria decision aid techniques, which are well-suited to ordinal classification problems. We use a sample of 114 banking sectors (i.e., countries), and a set of criteria that includes indicators of the macroeconomic, institutional and regulatory environment, as well as basic characteristics of the banking and financial sector. The models are developed and tested using a tenfold cross-validation approach and they are benchmarked against models developed with discriminant analysis and logistic regression.

**Keywords** Banking · Multicriteria decision aid · Risk · Stability

C. Gaganis · F. Pasiouras · M. Doumpos · C. Zopounidis (✉)
Financial Engineering Laboratory,
Department of Production Engineering and Management,
Technical University of Crete, University Campus, 73100 Chania, Greece
e-mail: kostas@dpem.tuc.gr

C. Gaganis
Department of Economics, University of Crete, Crete, Greece

F. Pasiouras
Centre for Governance and Regulation,
University of Bath School of Management, Bath, UK

## 1 Introduction

Banking crises can be damaging for the economy, and as the recent experience has shown, nowadays they can spread rapidly across the globe with contagious effects. In particular, after a relatively stable period between the second World War and the early 1970s, several countries experienced a banking crisis over the last 30 years. Caprio and Klingebiel [8] provide information on 117 systemic banking crises that occurred in 93 countries and 51 borderline and smaller banking crises in 45 countries since the late 1970s. Laeven and Valencia [23] also provide details as well as management strategies for 42 systemic banking crises from 37 countries between 1970 and 2007. These crises have both direct and indirect costs for the economy. First, as documented in [8] the costs for restructuring and recapitalisation can reach 10–20% and occasionally 40–55% of GDP (e.g., Argentina, Indonesia). Second, the crises have adverse effects on the efficient operation of the market economy due to the central role of banks as financial intermediates. Such adverse developments result in reduction in investment and consumption, increases in unemployment, and disturb the flow of credit to individuals and firms, thus causing an overall economic slowdown. This makes the assessment of the stability of the banking sectors of particular importance for regulators, depositors, investors and the general public.

Therefore, a number of studies examine the determinants of systemic banking crises or develop early warning models to predict the crises (e.g. [9,11,25]). However, there are a number of problems associated with these studies. First, they focus on the 1980s and the 1990s, when we experienced the majority of banking crises, and their results may not be applicable to the current financial environment. Second, they concentrate on emerging market economies due to the higher frequency of crises in these economies in the past [6]. Yet, the current crisis started from developed countries like the US and the UK. In addition, there are differences in the dates attributed to the banking crises [6], thus making their empirical modelling problematic. Third, dating is also problematic when there are successions of crises episodes as later crises can be extensions or re-emergences of previous financial distress rather than individual events [7,9]. Fourth, the binary classification of the banking sectors, in the ones that experienced a crisis and those that did not experienced a crisis, reduces the usefulness of the models.

The present paper has two overall aims. The first is the development of a classification model for the assessment of the stability of the banking sectors. We rely on the Economist Intelligent Unit (EIU) "Banking Sector Risk Ratings" rather than banking crises, and we examine the classification of a country's banking sector in three classes, labelled "low stability", "medium stability", and "high stability".[1] To

[1] In the case of banking sector risk, EIU assesses whether payment problems are likely to occur within the banking sector of a country (e.g., due to non-performing loans or rapid growth of bank credit to the private sector). Furthermore, political, economic policy, economic structure and liquidity risk factors including regulatory reforms and the involvement of the government in the banking sector are taken into account in the risk assessments. In our study we work under the assumption that higher banking sector risk, as reflected in the EIU ratings, corresponds to lower banking stability and vice versa. This assumption is consistent with the one used in recent studies that assess the soundness of individual banks on the basis of credit ratings
Footnote 1 continued

the best of our knowledge this is the fist study that follows this approach. First, this allows us to avoid some of the aforementioned problems, such as the dating of the crisis or the investigation of crises originating mainly from developing countries. Second, this approach can provide additional information compared to a binary setting, since it allow us to monitor the banking sectors as they deteriorate from the "high stability" to the "low stability" class. The second aim of the study is to examine the comparative ability of alternative multicriteria decision aid techniques. We focus on three techniques, namely UTilities Additives DIScriminantes (UTADIS), Multi-group Hierarchical DIScrimination (MHDIS), and ELECTRE TRI. These methods are well-suited to ordinal problems, like the one that we examine, but they implement different modelling forms (e.g., value functions and outranking relations). Thus, their consideration in this study enables the investigation of the generalizing power of different multicriteria models in evaluating the stability of a country's banking sector. We also use discriminant analysis and logistic regression for benchmarking purposes. We use a sample of 114 banking sectors (corresponding to an equal number of countries) and a set of criteria that includes indicators of the macroeconomic, institutional and regulatory environment, together with variables that consider the basic characteristics of the banking and financial sector. To ensure the proper estimation and validation of the models we follow a tenfold cross-validation approach.

The rest of the paper is as follows. Section 2 discusses the sample and the variables used in the study, while Sect. 3 outlines the classification techniques. Section 4 discusses the empirical results, and Sect. 5 concludes the study.

## 2 Sample and variables

### 2.1 Sample

We start by considering the 120 banking sectors which were assigned an EIU "Banking Sector Risk Rating" during 2008. The EIU ratings classify these sectors in 8 risk groups, ranging from C to AA. However, as our purpose is not to explain or replicate the ratings, but rather to use them as the basis for the development of a general model, we group the banking sectors in three broad classes.[2] The first class includes banking sectors with ratings A and AA, the second class includes those banking sectors with ratings B, BB and BBB, and the third class consists of banking sectors with ratings C,

(e.g., [13,21]). However, by looking at the banking sector as a whole we implicitly account for the probability that all the banks in the system experience large losses simultaneously.

2 We are not interested in replicating all the ratings of EIU for two reasons. First, the classification of the banking sectors in three general classes allows us to avoid potential problems that could arise during the estimation and validation of the models due to the small number of observations falling in some of the original EIU groups. Second, this approach allows us to avoid (at least to some extent) problems associated with the timely adjustment of ratings. For instance, a delay in a downgrade from AA to A or from BBB to B would have no impact in assessing the overall stability of a banking sector as we do. Furthermore small errors of judgment in the assignment of ratings such as rating an A banking sector as AA would also had no impact on our model. Obviously, large errors of judgment could make a difference but we have no reason to believe that EIU would classify let us say a B banking sector as A and vice versa.

**Table 1** Within-stability group geographical composition (%) of the sample

| | High stability | Medium stability | Low stability | Overall |
|---|---|---|---|---|
| Africa | 5.0 | 25.3 | 33.3 | 22.8 |
| Northern and Western Europe | 55.0 | 5.1 | 0.0 | 13.2 |
| Western Asia and Middle East | 0.0 | 15.2 | 6.7 | 11.4 |
| South-eastern Asia and Oceania | 15.0 | 8.9 | 13.3 | 10.5 |
| Central America | 0.0 | 10.1 | 20.0 | 9.6 |
| Southern America | 5.0 | 8.9 | 13.3 | 8.8 |
| Eastern Europe | 0.0 | 11.4 | 0.0 | 7.9 |
| Southern Europe | 10.0 | 8.9 | 0.0 | 7.9 |
| Central Asia | 0.0 | 6.3 | 13.3 | 6.1 |
| North America | 10.0 | 0.0 | 0.0 | 1.8 |
| All sample | 100 | 100 | 100 | 100 |

CC, CCC. Thus, Class 1 includes "high stability" banking sectors, Class 2 includes "medium stability" banking sectors, and Class 3 includes "low stability" sectors.

Data for end-2007 for the macroeconomic and institutional environment, as well as basic characteristics of the banking sectors, all at the country/sector level, were collected from the following sources: (i) the deposit insurance database developed by Demirguc-Kunt et al. [15], (ii) the financial structure database developed by Beck et al. [3,4], (iii) heritage foundation, and (iv) World development indicators (WDI).

After excluding the banking sectors from 6 countries due to missing data for the selected criteria, the final sample consists of 114 banking sectors. The distribution in the three classes is as follows: 20 (Class 1), 79 (Class 2), and 15 (Class 3). Table 1 presents the within-stability class geographical percentage composition of the sample. In particular, the figures in Table 1 show that 5% of the banking sectors in the sample that are classified as "high stability" involve countries from Africa, 55% of them involve countries from Northern and Western Europe, etc.

## 2.2 Criteria of banking stability

We use a total of 11 criteria falling in four general categories: (i) regulations, (ii) other banking and financial sector attributes, (iii) institutional environment, and (iv) macroeconomic conditions. These criteria and the corresponding data sources are presented in Table 2 and discussed below.

### *2.2.1 Regulations*

Theoretical and empirical evidence suggests that banking regulations such as entry into the banking industry, restrictions on activities, etc., as well as state ownership of banks can influence the stability of a country's banking sector see [2]. As in [16,27] we use an overall indicator of the relative openness of each country's banking and financial

**Table 2** Definition and sources of criteria

| | Calculation-description | Source |
|---|---|---|
| Regulations | | |
| BFREG | Index of banking and financial regulatory freedom. Higher scores indicate higher freedom | Heritage foundation |
| DEPINS | Dummy variable taking the value of 1 if there is an explicit deposit insurance scheme and 0 otherwise | [15] |
| Other banking and financial industry attributes | | |
| BLIQ | Average ratio of bank credit to bank deposits in the banking sector | [3,4] |
| BCONC | Concentration in the banking industry (% of assets held by three largest banks) | [3,4] |
| BROA | Average return on assets in the banking industry | [3,4] |
| CRGDP | Domestic credit to private sector /GDP | World development indicators |
| Institutional environment | | |
| PRIGHTS | Index of property rights. Higher figures indicate more secured property rights | Heritage foundation |
| CORRUPT | Index of corruption higher figures indicate lower corruption | Heritage foundation |
| GDPCAP | GDP per capita ($US) in constant prices | World development indicators |
| Macroeconomic conditions | | |
| GDPGR | GDP growth (%) | World development indicators |
| INFL | Inflation rate (%) | World development indicators |

system (BFREG), taken by Heritage Foundation. This index indicates the extent of restrictions on financial services, central bank independence, government ownership of banks, the difficulty of opening and operating domestic and foreign financial firms, and government influence on the allocation of credit. Higher scores indicate higher freedom (i.e., less restrictions) in the banking and financial sector.

Deposit insurance is another regulatory tool used in many countries as a way to avoid bank runs. However, deposit insurance schemes may encourage excessive risk-taking behaviour [12]. The main reason is that depositors will have no incentives to monitor bank managers, who can take on riskier investments under the assumption that depositors are protected in the event of a failure. Demirguc-Kunt and Detragiache [12] provide evidence that an explicit deposit insurance scheme, in the absence of strong banking regulations tends to increase the probability of banking crises. Barth et al. [2]

also report a positive relationship between deposit insurance "generosity" and the likelihood of a crisis. Therefore, we use a dummy variable indicating whether an explicit deposit insurance scheme has been adopted (DEPINS = 1) or not (DEPINS = 0).

### *2.2.2 Other banking and financial sector attributes*

As the recent crisis revealed, liquidity can become a very important problem for banks especially when there is reluctance for interbank borrowings and depositors demand a higher rate for their savings. To assess the liquidity of the banking sector we use the average ratio of bank credit to bank deposits (BLIQ), which shows the percentage of deposits that is tied up in loans. Therefore, higher ratios may indicate that the banking sector has fewer funds available to meet a sudden recall of its funding.

The literature suggests that increased competition decreases bank charter value and induces bank managers to increase risk [22]. Cross-country evidence by De Nicolo et al. [10], showed that during 1993–2000 highly concentrated banking markets faced higher levels of systemic risk compared to less concentrated ones, and this relationship strengthened between 1997 and 2000. In contrast, Beck et al. [5] report that more concentrated national systems are subject to a lower probability of systemic banking crises. As a rough measure of competition, we use the percentage of assets held by the three largest commercial banks relative to the total assets of the commercial banking sector within a country (BCONC).

Results from bank-level studies indicate that profitability is negatively related to the probability of failure, e.g., [24,31]. Therefore, we use the average return on assets in the banking industry (BROA) under the assumption that higher ROA will result in a more stable sector.

Finally, as discussed in [11], financial liberalization may increase banking sector fragility due to increased opportunities for excessive risk-taking and fraud. Therefore, following [9,11] among others, we use the ratio of domestic credit to private sector over GDP (CRGDP) to proxy for financial liberalization.

### *2.2.3 Institutional environment*

The stability of the banking sector may also be affected by the country's institutional environment, which can also mitigate the adverse effects of deposit insurance. For example, Barth et al. [2] find that better-developed private property rights and greater political openness mitigate the negative association of moral hazard and bank fragility. Demirguc-Kunt and Detragiache [12] and Demirguc-Kunt and Kane [14] also conclude that a sound legal system with proper enforcement of rules reduces the adverse effects of deposit insurance on bank risk-taking.

In the present study, we use three indicators of the level of the development of institutional environment. The first is an index of the protection of property rights (PRIGHTS) taken by Heritage Foundation. This index indicates the ability to accumulate private property, secured by clear and fully enforced laws. This index ranges between 0 (private property is outlawed and all property belongs to the state) and 100 (private property is guaranteed by the government). The second is the Heritage index of corruption (CORRUPT), which reveals the degree of corruption in the business environment, including levels of governmental, legal, judicial, and administrative corruption. It ranges between 0 and 100, with higher values indicating lower corruption. Finally, as in past studies, we use the GDP per capita as a general indicator of institutional development (e.g., [16]).

### 2.2.4 Macroeconomic conditions

Several studies document a relationship between real GDP growth and the probability or hazard rate of banking crisis (e.g., [9,11,20,26]). As Davis and Karim [9] mention, GDP growth not only reduces non-performing loans, but it can also delays banking crises due to pro-cyclicality. Following these studies, we use the real GDP growth (GDPGR) as an overall indicator of economic growth. Finally, we use the annual inflation rate (INFL), since past studies show that it can affect the stability of the banking sector (e.g., [9,11]).

### 2.2.5 Descriptive statistics

Table 3 presents some descriptive statistics for the above criteria while distinguishing between the three classes of observations. The Kruskal–Wallis non-parametric test indicates that there are statistically significant differences between the three groups in all the cases. These univariate results show that the development of the financial sector as measured by the credits to private sector (% GDP) and regulatory freedom in the financial industry as measured by the Heritage banking and financial index, improve the stability of a country's banking sector. Consistent with our expectations, inflation increases considerably as we move from countries with banking sectors of high stability to the ones with low stability. GDP growth appears to be lower (on average) in countries with high stability banking sectors compared to the ones with medium and low stability sectors. This is consistent with the findings of Tornell et al. [30] who mention that fastest growing countries are typically those that have experienced boom-bust cycles. We also observe important differences in all the indicators of institutional development. For example, the average GDP per capita of countries with banking sectors of high stability is around four times higher than that of countries with banking sectors of medium stability. Similarly, the values of the properly rights and corruption indices of countries with high stability are approximately two and three times higher than the ones of countries with medium stability and low stability banking sectors, respectively. Contrary to our expectations, the average ratio of bank credit to bank deposits is higher in the "high stability" banking sectors, indicating higher percentage of deposits being tied up in loans. However, that could also simply mean that banks in these banking sectors manage to transform more efficiently a higher percentage of their deposits into loans, which is considered to be the core financial intermediation activity in banking. The differences on average bank concentration, return on assets, and the existence of deposit insurance, are statistically significant only at the 10% level.

**Table 3** Descriptive statistics

| | High stability | | | Medium stability | | | Low stability | | | Kruskal–Wallis |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | Median | SD | Mean | Median | SD | Mean | Median | SD | |
| Deposit insurance | 0.85 | 1.00 | 0.49 | 0.62 | 1.00 | 0.49 | 0.53 | 1.00 | 0.52 | 4.67* |
| Credit to private sector/GDP | 133.26 | 129.84 | 45.27 | 54.30 | 40.40 | 45.27 | 31.11 | 25.67 | 21.33 | 33.90*** |
| GDP growth | 3.55 | 3.25 | 3.30 | 6.54 | 6.20 | 3.30 | 6.00 | 6.90 | 4.41 | 19.29*** |
| GDP per capita | 26971.63 | 27732.77 | 6500.55 | 5796.14 | 2867.75 | 6500.55 | 1375.54 | 783.03 | 1482.34 | 48.90*** |
| Inflation | 2.13 | 2.01 | 6.63 | 7.34 | 5.58 | 6.63 | 25.04 | 8.24 | 59.18 | 25.65*** |
| Banking and financial index | 68.00 | 70.00 | 16.55 | 53.42 | 50.00 | 16.55 | 38.46 | 38.46 | 19.58 | 19.47*** |
| Property rights index | 83.50 | 90.00 | 18.31 | 44.43 | 40.00 | 18.31 | 27.69 | 30.00 | 10.79 | 51.12*** |
| Corruption index | 79.55 | 82.50 | 16.42 | 38.97 | 34.00 | 16.42 | 26.08 | 26.00 | 4.22 | 50.35*** |
| Bank credit/bank deposits | 1.30 | 1.20 | 0.44 | 0.98 | 0.91 | 0.44 | 0.82 | 0.83 | 0.25 | 7.35** |
| Bank concentration | 75.65 | 0.78 | 0.18 | 66.85 | 0.66 | 0.18 | 75.66 | 0.79 | 0.19 | 5.79* |
| ROA | 1.14 | 0.01 | 0.02 | 1.72 | 0.01 | 0.02 | 2.33 | 0.02 | 0.04 | 5.63* |

The Kruskal–Wallis test indicates whether there are statistically significant differences between the three groups
*** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level

## 3 Multicriteria classification techniques

The problem considered in this case study falls within the multicriteria classification problematic, which, in general involves, the assignment of a finite set of alternatives $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ to a set of $q$ ordered classes $C_1 \succ C_2 \succ \cdots \succ C_q$. Each alternative is described by $m$ criteria and consequently it can be considered as a multivariate vector $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{im})$, where $x_{ij}$ is the description of alternative $i$ on criterion $j$. In the present study, the alternatives involve the 114 banking sectors, the criteria correspond to the 11 variables discussed in Sect. 2.2, and there are three ordered classes.

### 3.1 UTilities Additives DIScriminantes

The UTADIS method develops an additive value function, which is used to score the banking sectors and decide upon their classification. The value function has the following general form:

$$U(\mathbf{x}) = \sum_{j=1}^{m} w_j u'_j(x_j) \in [0, 1]$$

where $w_j$ is the weight of criterion $j$ (the criteria weights sum up to 1) and $u'_j(x_j)$ is the corresponding marginal value function normalized between 0 and 1. The marginal value functions provide a mechanism for decomposing the aggregate result (global value) in terms of individual assessments on the criteria level. To avoid the estimation of both the criteria weights and the marginal value functions, it is possible to use the transformation $u_j(x_j) = w_i u'_j(x_j)$. Since $u'_j(x_j)$ is normalized between 0 and 1, it is obvious that $u_j(x_j)$ ranges in $[0, w_i]$. In this way, the additive value function is simplified to the following form, which provides an aggregate score $U(\mathbf{x})$ for each banking sector along all criteria:

$$U(\mathbf{x}) = \sum_{j=1}^{m} u_j(x_j) \in [0, 1]$$

To classify the banking sectors it is necessary to estimate the thresholds $0 \leq t_{q-1} < \cdots < t_2 < t_1 \leq 1$ that distinguish the class. Comparing the value utilities with the thresholds, the classification is achieved as follows:

$$\left.\begin{array}{ll} U(\mathbf{x}) \geq t_1 & \Rightarrow \mathbf{x} \in C_1 \\ \ldots & \ldots \\ t_k \leq U(\mathbf{x}) < t_{k-1} & \Rightarrow \mathbf{x} \in C_k \\ \ldots & \ldots \\ U(\mathbf{x}) < t_{q-1} & \Rightarrow \mathbf{x} \in C_q \end{array}\right\}$$

The estimation of the additive value function and the cut-off thresholds is performed through linear programming techniques. The objective of the method is to develop the additive value model so that the above classification rules can reproduce the predetermined grouping of the banking sectors as accurately as possible. Therefore, a linear programming formulation is employed to minimize the sum of all violations of the above classification rules for all the observations in the training sample. A detailed description of the mathematical programming formulation can be found in [18].

### 3.2 Multi-group hierarchical DIScrimination

In contrast to UTADIS, MHDIS distinguishes the classes progressively, starting by discriminating the first class from all the others, and then proceeds to the discrimination between the alternatives belonging into the other classes. To accomplish this task, instead of developing a single additive value function that describes all alternatives (as in UTADIS), two additive value functions are developed in each one of the $q-1$ steps, where $q$ is the number of classes. At each stage $k$, an additive value function $U_k(\mathbf{x})$ is used to describe the alternatives of class $C_k$, while a second function $U_{\sim k}(\mathbf{x})$ is developed to describe the alternatives that belong in the classes $C_{k+1}, \ldots, C_q$.

The classification of an observation $\mathbf{x}$, is performed as follows:

If $U_1(\mathbf{x}) \geq U_{\sim 1}(\mathbf{x})$ then $\mathbf{x}$ belongs in $C_1$

else if $U_2(\mathbf{x}) \geq U_{\sim 2}(\mathbf{x})$ then $\mathbf{x}$ belongs in $C_2$

.......

The estimation of the weights of the criteria in the value functions as well as the marginal value functions, is accomplished through mathematical programming techniques. More specifically, at each stage of the hierarchical discrimination procedure, two linear programs and a mixed-integer one are solved to estimate optimally the two required functions and minimize the classification error. Further details of the mathematical programming formulations used in MHDIS can be found in [34].

### 3.3 ELECTRE TRI

The ELECTRE TRI method implements the outranking relations approach of multicriteria decision aiding [29]. The outranking relation is a binary relation that enables the assessment of the outranking degree of an alternative $\mathbf{x}_i$ over an alternative $\mathbf{x}_j$. The outranking relation allows to conclude that $\mathbf{x}_i$ outranks $\mathbf{x}_j$ if there are enough arguments to confirm that $\mathbf{x}_i$ is at least as good as $\mathbf{x}_j$ (concordance), while there is no essential reason to refute this statement (discordance).

Within the context of classification problems, the outranking relation is used to estimate the outranking degree of an alternative $\mathbf{x}_i$ over a reference profile $\mathbf{r}_k$, which distinguishes the classes $C_k$ and $C_{k+1}$. Each reference profile $\mathbf{r}_k$ is defined as a vector of individual profiles for each criterion, i.e., $\mathbf{r}_k = (r_{k1}, r_{k2}, \ldots r_{km})$.

In order to determine whether an alternative $\mathbf{x}_i$ outranks a reference profile $\mathbf{r}_k$, all paired comparisons $(x_{ij})$, $(r_{kj})$ and $(r_{kj}, x_{ij})$ should be performed for each criterion $j$. The former comparison enables the assessment of the strength $\sigma(\mathbf{x}_i, \mathbf{r}_k)$ of the affirmation "alternative $\mathbf{x}_i$ is at least as good as profile $\mathbf{r}_k$", while the latter comparison leads to the assessment of the strength $\sigma(\mathbf{r}_k, \mathbf{x}_i)$ of the affirmation "profile $\mathbf{r}_k$ is at least as good as alternative $\mathbf{x}_i$". An alternative $\mathbf{x}_i$ is preferred to a profile $\mathbf{r}_k (\mathbf{x}_i \,\mathrm{P}\, \mathbf{r}_k)$ if

$\sigma(\mathbf{x}_i, \mathbf{r}_k) \geq \lambda$ and $\sigma(\mathbf{r}_k, \mathbf{x}_i) < \lambda$ ($\lambda$ is a pre-specified cut-off point). If $\sigma(\mathbf{x}_i, \mathbf{r}_k) \geq \lambda$ and $\sigma(\mathbf{r}_k, \mathbf{x}_i) \geq \lambda$, then $\mathbf{x}_i$ and $\mathbf{r}_k$ are considered as indifferent ($\mathbf{x}_i \, \mathrm{I} \, \mathbf{r}_k$). Finally, if $\sigma(\mathbf{x}_i, \mathbf{r}_k) < \lambda$ and $\sigma(\mathbf{r}_k, \mathbf{x}_i) < \lambda$, then $\mathbf{x}_i$ and $\mathbf{r}_k$ are considered incomparable ($\mathbf{x}_i \, \mathrm{R} \, \mathbf{r}_k$). The estimation of the credibility index $\sigma(\mathbf{x}_i, \mathbf{r}_k)$ is performed in two stages. The first stage involves the concordance test, which considers the criteria for which $\mathbf{x}_i$ is at least as good as $\mathbf{r}_k$. The second stage considers the veto conditions, which may arise if $\mathbf{x}_i$ is significantly worse than $\mathbf{r}_k$ in some criteria. The details of this process can be found in [29].

Once the outranking relation is constructed, its exploitation to classify the alternatives is performed through heuristic assignment procedures. For instance, ELECTRE TRI employs two assignment procedures, the pessimistic and the optimistic one. Assuming a classification problem with $q$ classes, in the pessimistic assignment, each alternative $\mathbf{x}_i$ is compared successively to the profiles $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{q-1}$. Let $\mathbf{r}_k$ be the first profile such that $\sigma(\mathbf{x}_i, \mathbf{r}_k) \geq \lambda$. Then, $\mathbf{x}_i$ is assigned to group $C_k$ (if there is no profile such that $\sigma(\mathbf{x}_i, \mathbf{r}_k) \geq \lambda$, then $\mathbf{x}_i$ is assigned to group $C_q$). In the optimistic assignment each alternative $\mathbf{x}_i$ is compared successively to the profiles $\mathbf{r}_{q-1}, \mathbf{r}_{q-2}, \ldots, \mathbf{r}_1$. Let $\mathbf{r}_k$ be the first profile such that $\mathbf{r}_k \, \mathrm{P} \, \mathbf{x}_i$. Then, $\mathbf{x}_i$ is assigned to group $C_{k+1}$ (if the there is no profile satisfying the above condition, then $\mathbf{x}_i$ is assigned to group $C_1$).

The differences between the two procedures arise in the presence of the incomparability relation. For example, in a two-group case an alternative that is incomparable to the profile $\mathbf{r}_1$ will be assigned to group $C_1$ with the optimistic procedure and to group $C_2$ with the pessimistic procedure. Thus, the differences between the two assignment rules facilitate the identification of alternatives with special characteristics, which make the comparison of the alternatives to the profiles difficult.

In this study we employ the pessimistic assignment procedures and all the parameters of the ELECTRE TRI model (weights of the criteria, preference, indifference and veto thresholds, as well as the $\lambda$ cut-off point) are estimated using the evolutionary optimization approach, which has been recently proposed by Doumpos et al. [17].

## 4 Results

Table 4 presents the average weights (in %) of the criteria along all replications over the tenfold cross-validation analysis.[3] The banking-financial regulatory environment (BFREG) and corruption (CORRUPT) are the two most important criteria in the UTADIS model, accounting together for around 40%. Property rights (PRIGHTS) and liquidity (BLIQ) are of medium importance with weights around 12% each, while the weights of the rest of the criteria range between 0.21% (DEPINS) and 9.78% (INFL). Turning to the ELECTRE TRI model, we observe that the weights of the criteria are quite more balanced ranging between 3.54% (DEPINS) and 15.88% (GDPCAP). The

[3] As mentioned earlier, we adopt a tenfold cross validation approach to develop and evaluate the models. The full sample of the 114 banking sectors is randomly split into 10 mutually exclusive sub-samples (i.e., non-overlapping folds of approximately equal size). Then, ten models are developed in turn, using nine folds for training and leaving one fold out each time for validation. The average error rate over all the ten replications is the cross-validated error rate.

**Table 4** Weights of criteria (averages over 10 replications, in %)

| | UTADIS | ELECTRE TRI | MHDIS | | | |
|---|---|---|---|---|---|---|
| | | | $U_1$ | $U_{\sim1}$ | $U_2$ | $U_{\sim2}$ |
| DEPINS | 0.21 | 3.54 | 0.00 | 0.00 | 1.26 | 0.68 |
| CRGDP | 2.37 | 6.33 | 20.47 | 26.87 | 14.98 | 13.61 |
| GDPGR | 6.91 | 11.46 | 3.85 | 0.36 | 13.36 | 17.41 |
| GDPCAP | 4.78 | 15.88 | 18.32 | 16.67 | 5.78 | 8.07 |
| INFL | 9.78 | 11.66 | 9.77 | 13.43 | 5.75 | 7.36 |
| BFREG | 20.27 | 8.30 | 0.01 | 0.01 | 18.67 | 17.43 |
| PRIGHTS | 12.61 | 11.47 | 8.75 | 6.34 | 6.77 | 3.08 |
| CORRUPT | 20.17 | 9.36 | 19.35 | 20.91 | 7.28 | 9.46 |
| BLIQ | 12.36 | 5.66 | 10.63 | 7.76 | 7.27 | 5.86 |
| BCONC | 5.19 | 8.33 | 1.90 | 2.29 | 16.23 | 12.20 |
| BROA | 5.34 | 8.02 | 6.95 | 5.36 | 2.63 | 4.83 |

interpretation of the weights is more complicated in the case of MHDIS due to the multiple functions that are developed. The most important criteria in the first set of the utility functions $U_1$ (i.e., "high stability") and $U_{\sim1}$ (i.e., "medium" and "low" stability) are the domestic credit to private sector over GDP (CRGDP), GDP per capita (GDPCAP), and corruption (CORRUPT). However, in the second set of utility functions other criteria become important. In particular, in $U_2$ which characterizes the "medium stability" banking sectors and $U_{\sim2}$ which characterizes the "low stability" banking sectors, the four most important criteria are: bank concentration (BCONC), banking-financial regulatory environment (BFREG), domestic credit to private sector over GDP (CRGDP), and GDP growth (GDPGR).

Overall, it appears that deposit insurance is the less important criterion in all models, while the institutional environment is a good predictor (on an aggregate basis) of the stability of the banking sectors. Yet, there is no general agreement in the models as for the importance of the criteria. While there is no particular reason for that, such differences have been observed in past studies as well (e.g., [1,19,28]). One possible explanation is that although all methods attempt to classify correctly as many observations as possible, they consider different ways of processing the same information in the dataset. Another explanation is that, while UTADIS and ELECTRE develop only one function characterizing all banking sectors, MHDIS develops four functions that correspond to difference classes. It should also be noted, while the weights in the value functions developed with UTADIS and MHDIS represent tradeoffs, the weights in ELECTRE TRI represented the strength of the criteria in a weighted voted process. As discussed in [28], whether the weights attributed by one method are intuitively more appealing than those selected by another method is a matter of subjective judgment.

Table 5 presents the average classification results obtained over the ten replications. Since the classification accuracies in the training sample can be upward biased, we focus on the ones achieved in the validation sample. Panel B shows that these accuracies are quite satisfactory being 79.81% (ELECTRE TRI), 78.83% (UTADIS), and

**Table 5** Classification results (averages over tenfold cross-validation)

| | High stability (%) | Medium stability (%) | Low stability (%) | Average (%) |
|---|---|---|---|---|
| Panel A: Estimation | | | | |
| ELECTRE TRI | 93.83 | 78.56 | 98.35 | 90.25 |
| UTADIS | 95.53 | 65.26 | 97.08 | 85.96 |
| MHDIS | 100.00 | 99.86 | 100.00 | 99.95 |
| DA | 90.51 | 63.47 | 81.81 | 78.60 |
| LR | 80.93 | 95.90 | 26.31 | 67.71 |
| Panel B: Validation | | | | |
| ELECTRE TRI | 82.41 | 73.69 | 83.33 | 79.81 |
| UTADIS | 87.96 | 65.21 | 83.33 | 78.83 |
| MHDIS | 78.70 | 79.34 | 68.75 | 75.60 |
| DA | 76.67 | 62.98 | 62.50 | 67.38 |
| LR | 85.19 | 93.64 | 22.22 | 67.02 |

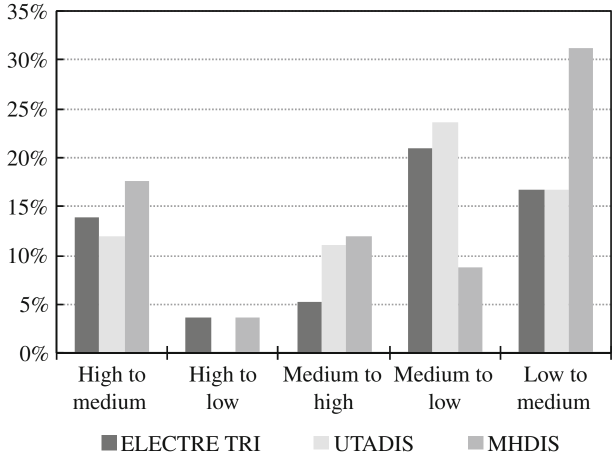


**Fig. 1** Analysis of classification errors for the three multicriteria methods

75.60% (MHDIS). Of particular importance is that ELECTRE TRI and UTADIS perform very well in identifying banking sectors that belong in Class 3, which bear the highest risk. MHDIS on the other hand achieves the highest accuracy in classifying banking sectors in Class 2. This is also a difficult task since the characteristics of these banking sectors may overlap with the ones belonging in the lower band of Class 1 and/or the upper band of Class 3.

Details on all types of classification error for the three multicriteria methods are given in Fig. 1. None of the methods led to a significant overestimation error (i.e., classification of a low stability sector to the high stability class). On the other hand, ELECTRE TRI and MHDIS have a very small major underestimation error (3.7%) involving a misclassification from the high stability class to the low stability one.

As a benchmark to the three MCDA methods, we also develop two models using discriminant analysis (DA) and logistic regression (LR). The average classification accuracies of DA and LR are considerably lower than the ones of the three MCDA methods. Despite being quite similar on average (i.e., DA: 67.38%; LR: 67.02%), these accuracies are achieved in a different way. While the performance of DA is balanced among the three classes, LR classifies very poorly banking sectors belonging in class 3, with an accuracy as low as 22.22%.

A closer look at the two models with the lowest misclassification errors in the validation sample indicates the following. First, 62% of the misclassification errors of UTADIS involve downgrades. In particular, Greece, Portugal, and Botswana are downgraded from the class of high stability to the one of medium stability. Countries downgraded from medium to low stability come mostly from Asia (India, Bangladesh, Indonesia, etc.), Africa (Algeria, Egypt, Nigeria, etc.), Latin-Southern America (Guatemala, Paraguay, Bolivia) and Eastern Europe (Russia, Romania). Only 38% of the errors involve upgrades. Most of them refer to upgrades of banking sectors from the medium stability class to the high stability one,[4] while there are also three misclassifications involving upgrades from the low stability class to the medium stability one (i.e., Sudan, Jamaica, Syria).

We observe a similar picture in the case of ELECTRE TRI, with downgrades accounting for 70% of the misclassification errors. In particular, the model downgrades four banking sectors from the high stability class to the medium stability one (Germany, USA, Singapore and Switzerland) and one country (Portugal) to the low stability group. As far as downgrades from the medium to the low stability class are concerned, they mostly involve countries from Western Asia (Azerbaijan, United Arab Emirates, Turkey), South-Eastern Asia (Indonesia, Malaysia), Eastern Europe (Moldova, Slovakia), Africa (Tunisia, Gabon, Equatorial Guinea), and Central-South America (Guatemala, Argentina, Uruguay). The misclassification errors of the model that are due to upgrades account for 30%, involving one upgrade from the medium to high stability class (Tanzania) and seven upgrades from low to medium stability class (Kenya, Honduras, Sudan, Uzbekistan, Nicaragua, Syria, and Vietnam).

## 5 Conclusions

The recent financial crisis that started in the US and the UK, and spread across the globe, highlighted the importance of early warning models to assess the stability of the countries' banking sectors. Using a sample of 114 countries, and a set of eleven variables, we developed three multicriteria decision aid models to classify banking sectors as "high stability", "medium stability", and "low stability". These models were capable in classifying correctly between 75.60 and 79.81% of the observations in the validation sample. In comparison, models developed with discriminant analysis and logistic regression achieved accuracies around 67%. The models developed in the

[4] These are mostly countries from Oceania (New Zealand, Australia), Europe (Spain, UK, Slovenia, Czech Republic), Israel, and South-eastern Asia (Korea, Malaysia).

present study could be useful in assessing and monitoring the soundness of a country's banking sector.

From a practical perspective, providing a crystal-clear suggestion on which method should be employed is a difficult if not impossible task (for theoretical results on this issue see [32,33]). The additive model of UTADIS is easy to build (from a computational point of view), use, and understand. On the other hand, it is more difficult to fit the multiple value functions of MHDIS or the outranking model of ELECTRE TRI to a set of training data. Their interpretation is also more involved. Nevertheless, they enable the modelling of some problem characteristics, which cannot be taken into consideration with an additive model. For instance, the multiple function approach of MHDIS enables the handling of situations where a criterion provides important information for the description of a specific class of countries, whereas its importance for other groups is limited. In a somewhat similar way, the outranking model of ELECTRE TRI provides the possibility to introduce veto in the analysis. In that respect, the final choice of the most appropriate method depends on its predictive power (in a specific problem domain), the insight information that it provides, and the way it fits the general evaluation scheme that the decision maker/analyst wants to implement. The results of this study indicate that in the context of evaluating the stability of a country's banking sector, multicriteria techniques can address in a satisfactory way, at least the first two of the above issues.

One of the limitations of our study is the development of models that classify the banking sectors in three groups, rather than the actual rating risk groups used by EIU. While this approach has certain advantages, it also comes with the disadvantage of a potentially less informative model. The small number of countries with banking sectors falling in certain EIU risk groups, did not allow us to explore this issue further. We hope that future research will improve upon that. The present study could also be extended towards various directions falling outside the scope of this paper. First, further research could compare additional classification techniques. Second, the study could be extended by exploring the development of integrated models to combine the classification estimations of the individual modes.

## References


1. Barnes, P.: The identification of U.K. takeover targets using published historical cost accounting data. Some empirical evidence comparing logit with linear discriminant analysis and raw financial ratios with industry-relative ratios. Int. Rev. Financ. Anal. **9**(2), 147–162 (2000)
2. Barth, J.R., Caprio, G. Jr., Levine, R.: Bank regulation and supervision: what works best? J. Financ. Intermed. **13**, 205–248 (2004)
3. Beck, T., Demirguç-Kunt, A., Levine, R.: A new database on financial development and structure. World Bank Econ. Rev. **14**, 597–605 (2000)
4. Beck, T., Demirguç-Kunt, A., Levine, R.: A new database on financial development and structure (1960–2007), September 2007 update, World Bank (2007)
5. Beck, T., Dermirguc-Kunt, A., Levine, R.: Bank concentration, competition and crises: first results. J. Banking Finance **30**, 1581–1603 (2006)
6. Bell, J., Pain, D.: Leading indicator models of banking crises—a critical review. Financial stability review, December, pp. 113–129 (2000)
7. Caprio, G., Jr., Klingebiel, D.: Bank insolvencies: cross-country experience. World Bank Policy Research Working Paper 1620, July (1996)

8. Caprio, G., Klingebiel, D.: Episodes of systemic and borderline financial crises, January, World Bank (2003). Available at http://go.worldbank.org/5DYGICS7B0
9. Davis, E.P., Karim, D.: Comparing early warning systems for banking crises. J. Financ. Stab. **4**, 89–120 (2008)
10. De Nicolo, G., Bartholomew, P., Zaman, J., Zephirin, M.: Bank consolidation, internationalization, and conglomeration: trends and implications for financial risk. Financ. Markets Inst. Instrum. **13**, 173–217 (2004)
11. Demirguc-Kunt, A., Detragiache, E.: The determinants of banking crises in developing and developed countries. IMF Staff Pap. **45**, 81–109 (1998)
12. Demirguc-Kunt, A., Detragiache, E.: Does deposit insurance increase banking system stability? An empirical investigation. J. Monet. Econ. **49**, 1373–1406 (2002)
13. Demirgüç-Kunt, A., Detragiache, E., Tressel, T.: Banking on the principles: compliance with basel core principles and bank soundness. J. Financ. Intermed. **17**, 511–542 (2008)
14. Demirguc-Kunt, A., Kane, E.J.: Deposit insurance around the world: where does it work? J. Econ. Perspect. **16**, 175–195 (2002)
15. Demirguc-Kunt, A., Karacaovali, B., Laeven, L.: Deposit Insurance around the World: A Comprehensive Database. World Bank Policy Research Working Paper 3628, June (2005)
16. Demirguc-Kunt, A., Laeven, L., Levine, R.: Regulations, market structure, institutions and the cost of financial intermediation. J. Money Credit Banking **36**, 593–622 (2004)
17. Doumpos, M., Marinakis, Y., Marinaki, M., Zopounidis, C.: An evolutionary approach to construction of outranking models for multicriteria classification: The case of the ELECTRE TRI method. Eur. J. Oper. Res. **199**, 496–505 (2009)
18. Doumpos, M., Zopounidis, C.: Developing sorting models using preference disaggregation analysis: An experimental investigation. Eur. J. Oper. Res. **154**, 585–598 (2004)
19. Espahbodi, H., Espahbodi, P.: Binary choice models for corporate takeover. J. Banking Finance **27**, 549–574 (2003)
20. Evrensel, A.Y.: Banking crisis and financial structure: a survival-time analysis. Int. Rev. Econ. Finance **17**, 589–602 (2008)
21. Ioannidis, C., Pasiouras, F., Zopounidis, C.: Assessing bank soundness with classification techniques. Omega (2010, in press). doi:10.1016/j.omega.2009.10.009
22. Keeley, M.C.: Deposit insurance, risk, and market power in banking. Am. Econ. Rev. **5**, 1183–1200 (1990)
23. Laeven, L., Valencia, F.: Systemic Banking Crises: A New Database. IMF Working Paper 08/224, October (2008)
24. Lanine, G., Vander Vennet, R.: Failure prediction in the Russian bank sector with logit and trait recognition models. Exp. Syst. Appl. **30**, 463–478 (2006)
25. Moshirian, F., Wu, Q.: Banking industry volatility and banking crises. Int. Financ. Markets/Institutions Money **19**(2), 351–370 (2009)
26. Noy, I.: Financial liberalization, prudential supervision, and the onset of banking crises. Emerging Markets Rev. **5**, 341–359 (2004)
27. Pasiouras, F., Gaganis, C., Doumpos, M.: A multicriteria discrimination approach for the credit ratings of Asian banks. Ann. Finance **3**, 351–367 (2007)
28. Pasiouras, F., Tanna, S., Zopounidis, C.: The identification of acquisition targets in the EU banking industry: An application of multicriteria approaches. Int. Rev. Financ. Anal. **16**, 262–281 (2007)
29. Roy, B., Bouyssou, D.: Aide Multicritère à la Décision: Méthodes et Cas. Economica, Paris (1993)
30. Tornell, A., Westermann, F., Martinez, L.: The Positive Link Between Financial Liberalization, Growth and Crises. NBER Working Paper No. 10293, February (2004)
31. Wheelock, D.C., Wilson, P.W.: Why do banks disappear? The determinants of U.S. bank failures and acquisitions. Rev. Econ. Stat. **82**, 127–138 (2000)
32. Wolpert, D.H.: The lack of a priori distinctions between learning algorithms. Neural Comput. **8**(7), 1341–1390 (1996)
33. Wolpert, D.H., Macready, W.G.: No free lunch theorems for optimization. IEEE Trans. Evol. Comput. **1**, 67–82 (1997)
34. Zopounidis, C., Doumpos, M.: Multi-group discrimination using multi-criteria analysis: illustrations from the field of finance. Eur. J. Oper. Res. **139**, 371–389 (2002)