



Małgorzata Iwanicz-Drozdowska 

Warsaw School of Economics, Collegium of Management and Finance, Institute of Finance
Financial System Department, miwani@sgh.waw.pl

Aneta Ptak-Chmielewska 

Warsaw School of Economics, Collegium of Economic Analysis, Institute of Statistics and Demography
Applied Statistics Unit, aptak@sgh.waw.pl

Prediction of Banks Distress – Regional Differences and Macroeconomic Conditions

Abstract: In this study we focus on distress events of European banks over the period of 1990–2015, using unbalanced panel of 3,691 banks. We identify 132 distress events, which include actual bankruptcies as well as bailout cases. We apply CAMEL-like bank-level variables and control macroeconomic variables (GDP, inflation, unemployment rate). The analysis is based on traditional logistic regression and k-means clustering. We find, that the probability of distress is connected with macroeconomic conditions via regional grouping (clustering). Bank-level variables that were stable predictors of distress from 1 to 4 years prior to event are equity to total assets ratio (leverage) and loans to funding (liquidity). From macroeconomic factors, the GDP growth is a reasonable variable, however with differentiated impact: for 1 year distance high distress probability is connected with low GDP growth, but for 2, 3 and 4 year distance: high distress probability is conversely connected with high GDP growth. This shows the changing role of macroeconomic environment and indicates the potential impact of favorable macroeconomic conditions on building-up systemic problems in the banking sector.

Keywords: banks distress, CAMEL, logistic prediction, macroeconomic variables

JEL: G20, G21

1. Introduction

The prediction of bank failures is found to be a challenge due to the limited number of these failures. The research on bank failures was initiated by Sinkey Jr (1975) for American banks and expanded mostly for the US. In Europe the number of studies is much more limited, however the wave of bank bailouts that swept across Europe during the global financial crisis (GFC), helped study bank problems in Europe (e.g., Poghosyan, Cihak, 2011; Altman, Cizel, Rijken, 2014; Betz et al., 2014; Iwanicz-Drozdowska, Laitinen, Suvas, 2018).

Due to the fact that the number of actual bank failures is low, cases of bailouts and forced mergers have been included in bank failure studies and called distress events (e.g. Arena, 2008; Betz et al., 2014; Altman, Cizel, Rijken, 2014). We follow this approach, which is now well-established in literature. However, still the number of distress events in Europe is much lower than in the US¹, which makes bank bankruptcy prediction a difficult task.

This study expands an already existing research by Iwanicz-Drozdowska, Laitinen and Suvas (2018), using the same database and accounting for macroeconomic factors and heterogeneity of countries. The goal of this paper is to identify the link between macroeconomic conditions and bank distress while accounting for heterogeneity of countries due to the level of the risk of bank distress. Addressing the heterogeneity of countries and advanced methods adds to the research in the field of banks distress prediction.

The remaining part of the paper is organized as follows. Section 2 provides a review of the literature and presents research hypothesis. Section 3 presents the data and methodology. Section 4 discusses the empirical results, and Section 5 addresses the policy implications and presents conclusions.

2. Literature review and research hypothesis

The analyses of bank financial difficulties have been conducted on macro- and microeconomic levels. On macroeconomic level, the research was focused on the detection of early warning signs of a banking crisis and was based to a large extent on macroeconomic and industry-level data. This stream of research has developed especially after the outbreak of the South Asian crisis in mid 1990s and again after 2008 (Drehmann, Juselius, 2014). On microeconomic level, the research was targeted at detecting banking failures, mostly for the US banking sector (e.g., Sinkey Jr, 1975; Peek, Rosengren, 1996; Wheelock, Wilson, 2000; Kolari et al.,

1 For 2000–2011, Hamusch and Shaffer (2016) registered 441 bank failures in the US, which is many more than distress events occurring among the European banks in the same period.

2002; Cole, White, 2012; Shaffer, 2012; Cox, Wang, 2014; López Iturriaga, Sanz, 2015). There is a limited number of studies of bank failures from other countries due to the fact that the number of failures is rather low. In this stream of research the bank-level data has been used to a large extent. Bank specific traits included in the analysis were based on the CAMEL² approach. CAMEL has been one of the most popular approaches to assess the banks' financial position, both in research and supervisory practice (e.g., Lopez, 1999), because it covers the most important aspects of bank risks and performance.

Against this background, we expand microeconomic analysis of bank problems using bank and macroeconomic characteristics. Therefore, our review of literature is selective and focuses on the studies which accounted for the impact of macroeconomic environment on bank distress. To the best of our knowledge, there are two studies fulfilling these criteria: Kapinos and Mitnik (2016) for the US banks and Betz et al. (2014) for European banks.

Kapinos and Mitnik (2016) conducted stress tests for medium and large banks showing how resilient individual banks and banking sector are to macroeconomic shocks. They applied principal component analysis (PCA) to macroeconomic factors and least absolute shrinkage and selection operator (LASSO) combined with PCA for bank-level characteristics. This study accounted for heterogeneity of banks. They identified macroeconomic drivers of bank-level variables and these bank-level factors that explain bank heterogeneity in reaction to macroeconomic shocks. Among macroeconomic variables appeared: VIX level, BBB spread, DJIA growth, growth of housing prices (one of the leading factor for the GFC), mortgage rates, GDP growth, unemployment rate and inflation rate. Bank-level traits included growth rates of assets and loans, leverage (equity to assets), tier 1 to assets, deposits to assets, trading assets to assets, non-performing loans (NPL), consumer to loans (and to assets), and real estate loans to loans (and to assets). From the CAMEL perspective two components were omitted: earnings and management (proxied in research by cost to income ratio, but it actually represents earnings). They modelled two variables, i.e. the pre-provision net revenue and net charge-offs on loans and leases, both representing earnings. The goal of Kapinos and Mitnik (2016) was to run stress tests and therefore we cannot follow their methodology in this study, but we apply similar list of macroeconomic and bank-level variables adjusted to the availability of data and specifics of European accounting.

Betz et al. (2014) based on quarterly data of European banks from 2000 to mid-2013 used CAMEL-like bank-level variables and macroeconomic and banking sector variables. They focus on banks with total assets of at least 1 billion EUR (546 banks in the sample and 28,832 bank-quarter observations). Macroeconomic var-

2 This methodology requires knowledge about the bank's capital adequacy (C), asset quality (A), management (M), earnings (E) and liquidity (L).

ables were similar to Kapinos and Mitnik (2016), but expanded by government debt to GDP, private sector credit flow to GDP and international investment position to GDP. The list of banking sector variables included e.g., total assets to GDP, debt to equity, loans to deposits, mortgages to loans, non-core liabilities. These variables duplicated the bank-level variables on sector level and we decided not to follow this approach. The inclusion of macroeconomic and banking sector variables helped improve performance of distress prediction models.

Peltonen, Piloiu and Sarlin (2015) introduced estimated network linkages into an early-warning model to predict bank distress among European banks. Authors used multivariate extreme value theory to estimate equity-based tail-dependence networks. This model links proxy for the markets' view of bank interconnectedness in case of elevated financial stress. This paper found that early warning models including estimated tail dependencies outperform bank-specific benchmark models without networks. This paper gave direct support for measures of interconnectedness in early-warning models.

Maghyereh and Awartani (2014) applied a simple hazard model for an early warning system of bank distress in the Gulf Cooperation Council countries (GCC). They identified a set of leading indicators of bank distress to predict the probability of bank failure in these countries. They covered a wide set of bank level variables and other variables like influence of bank management, competition, diversification, ownership and regulation. They found that good management lowers the probability of distress. The bank specific and other CAMEL type variables as well as the systematic shocks in the financial and macroeconomic environment were all found to be in line with the findings of related empirical studies. Finally, they found that a simple hazard model has performed fairly well in predicting bank distress.

Ravisankar and Ravi (2010) used unusual neural network architectures for bankruptcy prediction in banks: Group Method of Data Handling (GMDH), Counter Propagation Neural Network (CPNN) and fuzzy Adaptive Resonance Theory Map (fuzzy ARTMAP). Effectiveness of those techniques was tested by using four different datasets pertaining to Spanish banks, Turkish banks, UK banks and US banks. They selected top five (Spanish dataset) or top seven (in the case of Turkish and UK datasets) variables as input to GMDH, CPNN and fuzzy ARTMAP for classification purpose. The performance of these hybrid techniques of variables selection was compared with that of GMDH, CPNN and fuzzy ARTMAP in their stand-alone mode without variables selection. Cross validation was performed throughout the study. Results indicate that the GMDH outperformed all the techniques with or without variables selection. The results were much better than those reported in previous studies on the same datasets in terms of average accuracy, average sensitivity and average specificity.

Hájek, Olej and Myšková (2015) in their study proposed a model based on random subspace method to predict investment/non-investment rating grades of U.S.

banks. They showed that support vector machines SVM can be effectively used as base learners in the meta-learning model. Both financial and non-financial (sentiment) information were important categories of determinants in financial distress prediction.

SirElkhatim and Salim (2015) conducted a comprehensive review of the existing literature of prediction techniques that have been used to assist in the prediction of bank distress. They categorized the review results into groups depending on the prediction techniques method. They treated the literature from the period 1990–2010 as history of prediction techniques, and after this period until 2013 as recent prediction techniques. They then presented the strengths and weaknesses of both categories. There was no specific type fit for all bank distress issues although they found that intelligent hybrid techniques were the most reliable in term of accuracy and reputation.

Against the literature review done by SirElkhatim and Salim (2015) and selected papers described above, we test two research hypothesis:

H1: There is a clear link between macroeconomic conditions and distress risk in banking sector.

H2: There are homogenous clusters of countries (regions) with high, medium and low risk of distress in banking sector.

H1 is motivated by positive results of Betz et al. (2014) and the role of macroeconomic factors for banks financial capacity confirmed by Kapinos and Mitnik (2016). H2 is motivated by the fact that studies on European bank distress, due to the limited number of distress cases, did not account for heterogeneity of countries.

3. Data and methodology

We use data of 3,691 banks with financial statements (FS) for years: 1990–2015, extracted from BankScope database. A balanced sample of 132 distressed and 132 regular FSs has been selected for distress prediction, while full sample of 47,925 annual FSs has been used for clustering. Betz et. el (2014) also use the European banks data but on the quarterly FS level.

The modelling technique is the logistic regression with stepwise method (0.05 significance level at entry and at stay in the model) for distress prediction using CAMEL variables (like e.g., in Betz et al., 2014). Next, the cluster analyses with k-means method are used, based on the distress probability in each country, estimated on prediction model for balanced sample (mean distress probability 0.5). Countries are classified into three clusters due to the level of risk of distress. The clusters are characterized by low, medium and high risk of distress. Macroeconomic variables are used only on clusters level while Betz et al. (2014) use macro-economic variables for prediction in logit models.

Table 1 presents the list of bank-level CAMEL variables selected from a long list applied in other studies.

Table 1. CAMEL variables description

Variable Name	Description
<i>Growth_TA</i>	one-year total assets (TA) growth rate; A, M in CAMEL
<i>Growth_EQ</i>	one-year equity (EQ) growth rate; C in CAMEL
<i>Growth_G_Loans</i>	one-year gross loans (GL) growth rate; A in CAMEL
<i>ROA</i>	return on total assets ratio; E in CAMEL
<i>EQ_to_TA</i>	equity to total assets ratio; C in CAMEL
<i>Deposits_to_G_Loans</i>	total customer deposits to (gross) loans ratio; L in CAMEL
<i>L_Imp_to_G_Loans</i>	loan impairment charge (LI) = loan loss provisions (LLP) to gross loans ratio; A in CAMEL
<i>NIM</i>	net interest margin to interest-earning assets; E in CAMEL
<i>CI</i>	operating cost (expense) to operating income ratio; E, M in CAMEL
<i>Loans_to_TA</i>	net loans to total assets ratio; A, M in CAMEL
<i>Loans_to_Funding</i>	net loans to customer and short-term funding ratio; L in CAMEL
<i>Liquid_A_to_Funding</i>	liquid assets to deposits and short-term funding ratio; L in CAMEL

Source: own elaboration

Macroeconomic variables included in the clustering description are as follows: GDP growth, inflation rate and unemployment rate. These variables have been widely used in previous studies. As mentioned before, the macro-variables are not used for predictions.

4. Empirical results

Results for prediction models are presented in Table 2, while the accuracy of models is presented in Table 3.

Overall, accuracy of the prediction models was on the level of 72–76% (see Table 3). The highest percentage of correctly classified distressed banks occurred within 4 years prior to the distress.

Table 2. Prediction models results for 1, 2, 3 and 4 years prior to distress, based on balanced sample

Variable	DF	Estimate	Standard error	Wald ch-sqr statistic	<i>p</i> -value
1 year to distress					
<i>Intercept</i>	1	1.1288	0.6093	3.4321	0.0639
<i>ROA</i>	1	-49.8606	15.5835	10.2373	0.0014
<i>EQ_to_TA</i>	1	-20.7261	5.2050	15.8562	< .0001
<i>Loans_to_TA</i>	1	-4.3171	1.4075	9.4074	0.0022
<i>Loans_to_Funding</i>	1	3.5609	0.8880	16.0800	< .0001

Variable	DF	Estimate	Standard error	Wald ch-sqr statistic	p-value
2 years to distress					
<i>Intercept</i>	1	3.0040	0.9475	10.0509	0.0015
<i>EQ_to_TA</i>	1	-13.6464	3.7707	13.0978	0.0003
<i>CI</i>	1	-2.8896	0.9748	8.7863	0.0030
<i>Loans_to_TA</i>	1	-4.1489	1.2728	10.6252	0.0011
<i>Loans_to_Funding</i>	1	2.8466	0.7878	13.0576	0.0003
3 years to distress					
<i>Intercept</i>	1	2.4665	0.9181	7.2171	0.0072
<i>Growth_G_Loans</i>	1	3.6091	1.0728	11.3186	0.0008
<i>NIM</i>	1	-67.6595	14.8663	20.7133	<.0001
<i>CI</i>	1	-3.2462	1.0101	10.3277	0.0013
<i>Loans_to_Funding</i>	1	0.9571	0.4589	4.3501	0.0370
4 years to distress					
<i>Intercept</i>	1	0.7099	0.3751	3.5814	0.0584
<i>Growth_G_Loans</i>	1	7.8993	1.4262	30.6758	<.0001
<i>EQ_to_TA</i>	1	-16.5222	4.4606	13.7196	0.0002
<i>Deposits_to_G_Loans</i>	1	-0.6463	0.1870	11.9401	0.0005

Source: own elaboration using SAS 9.4

Analysed banks are situated in 17 different countries (see Table 4). In Denmark and Sweden probability of distress is low in 4, 3, 2 and 1 years prior to distress. In Ireland probability of distress is high in 4, 3, 2 and 1 years prior to distress. In UK the level of probability of distress is at medium level in all years prior to distress. There is a group of countries where probability of distress increases as the distance to distress decreases: Germany, Italy, Netherlands; and the group of countries where probability of distress decreases as the distance to distress decreases: Latvia, Cyprus, Greece, Iceland. In the rest of countries the trend is not clear.

Table 3. Prediction accuracy of the models

	1 year	2 years	3 years	4 years
AUC	0.848	0.797	0.816	0.815
Overall accuracy	76.1%	72.4%	76.5%	74.7%
% correct classified good	81.4%	68.8%	76.0%	69.9%
% correct classified bad	70.8%	76.2%	77.0%	78.6%

Source: own elaboration

Results for the 1 year to distress are presented in Table 5 and Figures 1 and 2. Low risk of distress 1 year prior to distress cluster is characterised by high value of *EQ_to_TA* ratio, high value of *ROA*, and high value of *Loans_to_TA* and *Loans_to_Funding* ratios. High risk of distress 1 year prior to distress cluster is characterised by low value of *EQ_to_TA* ratio, low *ROA*, but high value of *Loans_to_TA* and *Loans_to_Funding* ratios. Considering macroeconomic conditions, lower risk of distress is connected with lower unemployment rate, lower inflation and higher GDP growth.

Table 4. Clustering results

Country	Frequency – FSs	1 year to distress	2 years to distress	3 years to distress	4 years to distress
Austria	3153	3	2	3	2
Belgium	701	3	3	2	2
Cyprus	224	3	3	3	3
Denmark	1090	2	2	3	2
France	3809	3	2	3	2
Germany	23338	1	2	3	2
Greece	183	3	2	3	3
Iceland	101	3	3	2	3
Ireland	266	1	1	1	3
Italy	8382	1	3	3	2
Latvia	253	2	2	2	3
Netherlands	564	1	3	2	1
Portugal	632	3	2	3	2
Slovenia	218	3	3	3	1
Spain	1826	3	3	2	2
Sweden	1328	2	2	3	2
UK	1857	3	3	2	1

■ – Low probability of distress
■ – Medium probability of distress
■ – High probability of distress

Source: own elaboration.

Table 5. Clustering results CAMEL and macroeconomic variables – 1 year to distress

Variable	Cluster 1 high risk		Cluster 2 low risk		Cluster 3 medium risk	
	Mean	Median	Mean	Median	Mean	Median
1 year distress probability	0.411	0.410	0.230	0.234	0.332	0.324
CAMEL variables						
<i>Growth_TA_Mean</i>	0.109	0.112	0.183	0.104	0.154	0.104
<i>Growth_EQ_Mean</i>	0.450	0.335	0.292	0.271	0.486	0.374
<i>Growth_G_Loans_Mean</i>	0.150	0.149	0.211	0.112	0.187	0.135
<i>ROA_Mean</i>	0.006	0.005	0.007	0.007	0.006	0.005
<i>EQ_to_TA_Mean</i>	0.099	0.104	0.128	0.130	0.115	0.116
<i>Deposits_to_G_Loans_Mean</i>	11.284	1.698	2.041	1.277	2.273	1.886
<i>L_Imp_to_G_Loans_Mean</i>	0.009	0.009	0.018	0.012	0.003	0.013
<i>NIM_Mean</i>	0.023	0.024	0.033	0.031	0.028	0.027
<i>CI_Mean</i>	0.588	0.648	0.656	0.652	0.672	0.663
<i>Loans_to_TA_Mean</i>	0.539	0.568	0.582	0.592	0.554	0.563
<i>Loans_to_Funding_Mean</i>	0.832	0.845	0.836	0.887	0.778	0.766
<i>Liquid_A_to_Funding_Mean</i>	0.401	0.393	0.363	0.340	0.365	0.370

Variable	Cluster 1 high risk		Cluster 2 low risk		Cluster 3 medium risk	
	Mean	Median	Mean	Median	Mean	Median
1 year distress probability	0.411	0.410	0.230	0.234	0.332	0.324
Macroeconomic variables						
<i>GDP_growth_Mean</i>	0.018	0.015	0.025	0.022	0.018	0.017
<i>INF_Mean</i>	0.020	0.021	0.027	0.019	0.027	0.022
<i>UNEMPL_Mean</i>	0.073	0.079	0.081	0.070	0.082	0.074

Source: own elaboration.

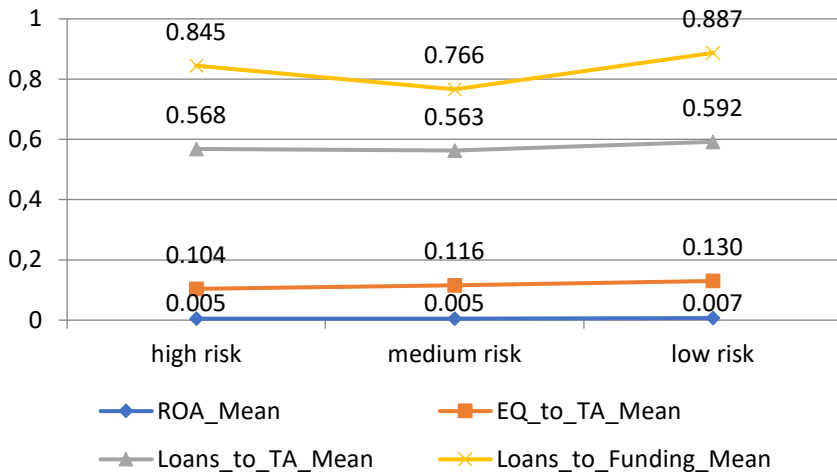


Figure 1. Mean values of financial ratios for clusters – 1 year to distress

Source: own elaboration

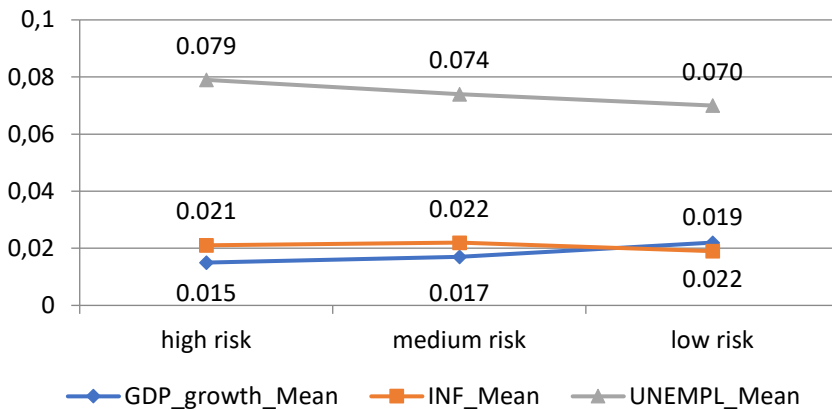


Figure 2. Mean values of macroeconomic variables for clusters – 1 year to distress

Source: own elaboration

Results for the 2 years to distress were presented in Table 6 and Figures 3 and 4. Low risk of distress 2 years prior to distress cluster is characterised by high value of *EQ_to_TA* ratio, high *CI*, and high value of *Loans_to_TA* and *Loans_to_Funding* ratios. High risk of distress 2 years prior to distress cluster is characterised by low value of *EQ_to_TA* ratio, low *CI*, and low value of *Loans_to_TA* and *Loans_to_Funding* ratios. Considering macroeconomic conditions, lower risk of distress is connected with lower inflation and lower GDP growth. Unemployment is not correlated with risk of distress 2 years prior to distress.

Table 6. Clustering results CAMEL and macroeconomic variables – 2 years to distress

Variable	Cluster 1 high risk		Cluster 2 low risk		Cluster 3 medium risk	
	Mean	Median	Mean	Median	Mean	Median
2 years distress probability	0.678	0.678	0.356	0.37	0.449	0.433
CAMEL variables						
<i>Growth_TA_Mean</i>	0.162	0.162	0.131	0.089	0.163	0.119
<i>Growth_EQ_Mean</i>	1.044	1.044	0.382	0.251	0.429	0.374
<i>Growth_G_Loans_Mean</i>	0.251	0.251	0.16	0.112	0.197	0.14
<i>ROA_Mean</i>	0.004	0.004	0.005	0.004	0.008	0.007
<i>EQ_to_TA_Mean</i>	0.093	0.093	0.114	0.123	0.115	0.117
<i>Deposits_to_G_Loans_Mean</i>	1.929	1.929	1.787	1.546	7.221	2.156
<i>L_Imp_to_G_Loans_Mean</i>	0.009	0.009	0.015	0.012	-0.002	0.01
<i>NIM_Mean</i>	0.011	0.011	0.029	0.028	0.029	0.027
<i>CI_Mean</i>	0.354	0.354	0.69	0.694	0.645	0.647
<i>Loans_to_TA_Mean</i>	0.43	0.43	0.573	0.569	0.554	0.573
<i>Loans_to_Funding_Mean</i>	0.624	0.624	0.782	0.767	0.842	0.807
<i>Liquid_A_to_Funding_Mean</i>	0.641	0.641	0.343	0.347	0.369	0.378
Macroeconomic variables						
<i>GDP_growth_Mean</i>	0.037	0.037	0.016	0.014	0.02	0.018
<i>INF_Mean</i>	0.022	0.022	0.023	0.019	0.028	0.023
<i>UNEMPL_Mean</i>	0.077	0.077	0.085	0.088	0.075	0.065

Source: own elaboration

Results for the 3 years to distress were presented in Table 7 and Figures 5 and 6. Low risk of distress 3 years prior to distress cluster is characterised by high value of *NIM*, high *CI*, and high value of *Loans_to_Funding* ratio but low value of *Growth_G_Loans* ratio. High risk of distress 3 years prior to distress cluster is characterised by low value of *NIM*, low *CI*, and low value of *Loans_to_Funding* ratio but high value of *Growth_G_Loans* ratio. Considering macroeconomic conditions, lower risk of distress is connected with lower inflation and lower GDP growth. Unemployment is not correlated with risk of distress 3 years prior to distress.

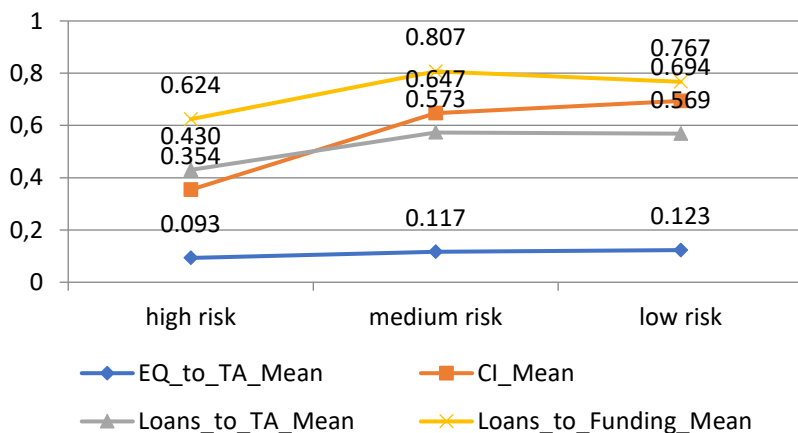


Figure 3. Mean values of financial ratios for clusters – 2 years to distress

Source: own elaboration

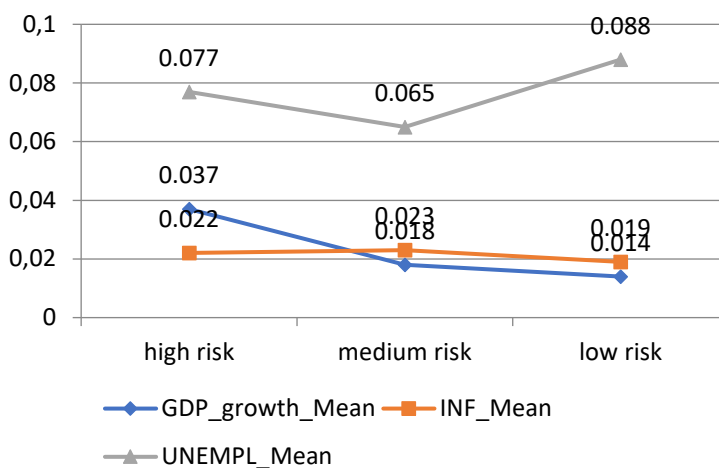


Figure 4. Mean values of macroeconomic variables for clusters – 2 years to distress

Source: own elaboration

Table 7. Clustering results CAMEL and macroeconomic variables – 3 years to distress

Variable	Cluster 1 high risk		Cluster 2 medium risk		Cluster 3 low risk	
	Mean	Median	Mean	Median	Mean	Median
3 years distress probability	0.775	0.775	0.534	0.529	0.399	0.403
CAMEL variables						
<i>Growth_TA_Mean</i>	0.162	0.162	0.182	0.119	0.127	0.1
<i>Growth_EQ_Mean</i>	1.044	1.044	0.393	0.316	0.413	0.348

Variable	Cluster 1 high risk		Cluster 2 medium risk		Cluster 3 low risk	
	Mean	Median	Mean	Median	Mean	Median
3 years distress probability	0.775	0.775	0.534	0.529	0.399	0.403
<i>Growth_G_Loans_Mean</i>	0.251	0.251	0.218	0.156	0.155	0.118
<i>ROA_Mean</i>	0.004	0.004	0.008	0.009	0.005	0.005
<i>EQ_to_TA_Mean</i>	0.093	0.093	0.121	0.122	0.111	0.114
<i>Deposits_to_G_Loans_Mean</i>	1.929	1.929	9.58	3.015	1.458	1.372
<i>L_Imp_to_G_Loans_Mean</i>	0.009	0.009	-0.003	0.008	0.013	0.012
<i>NIM_Mean</i>	0.011	0.011	0.025	0.024	0.032	0.028
<i>CI_Mean</i>	0.354	0.354	0.651	0.643	0.678	0.668
<i>Loans_to_TA_Mean</i>	0.43	0.43	0.519	0.561	0.59	0.588
<i>Loans_to_Funding_Mean</i>	0.624	0.624	0.787	0.748	0.827	0.791
<i>Liquid_A_to_Funding_Mean</i>	0.641	0.641	0.401	0.4	0.33	0.347
Macroeconomic variables						
<i>GDP_growth_Mean</i>	0.037	0.037	0.027	0.019	0.013	0.014
<i>INF_Mean</i>	0.022	0.022	0.031	0.024	0.022	0.019
<i>UNEMPL_Mean</i>	0.077	0.077	0.082	0.071	0.078	0.075

Source: own elaboration

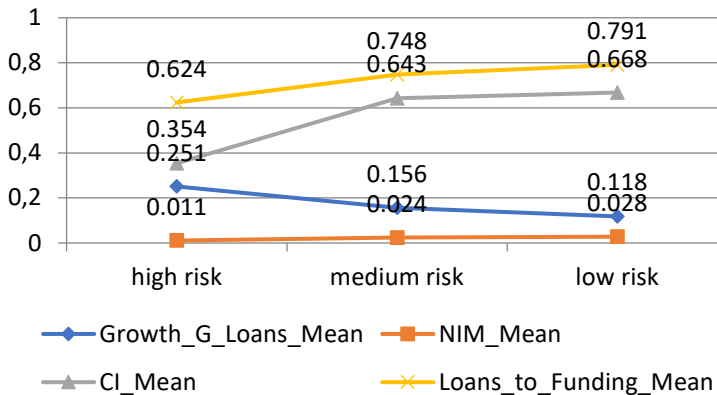


Figure 5. Mean values of financial ratios for clusters – 3 years to distress

Source: own elaboration

Results for the 4 years to distress are presented in Table 8 and Figures 7 and 8. Low risk of distress 4 years prior to distress cluster is characterised by high value of *EQ_to_TA* ratio, low value of *Deposits_to_G_Loans* and *Growth_G_Loans* ratios. High risk of distress 4 years prior to distress cluster is characterised by high value of *EQ_to_TA* ratio, low value of *Deposits_to_G_Loans* ratio and high value of *Growth_G_Loans* ratio. Considering macroeconomic conditions, lower risk of distress is connected with lower inflation and lower GDP growth. Unemployment is not correlated with risk of distress 4 years prior to distress.

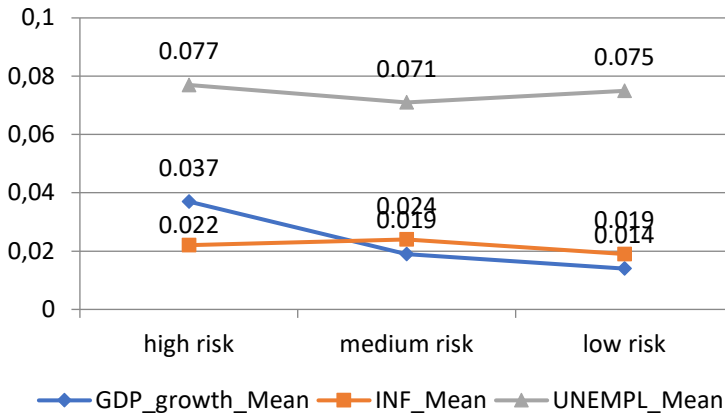


Figure 6. Mean values of macroeconomic variables for clusters – 3 years to distress

Source: own elaboration

Table 8. Clustering results CAMEL and macroeconomic variables – 4 years to distress

Variable	Cluster 1 medium risk		Cluster 2 low risk		Cluster 3 high risk	
	Mean	Median	Mean	Median	Mean	Median
3 years distress probability	0.383	0.385	0.294	0.3	0.468	0.471
CAMEL variables						
<i>Growth_TA_Mean</i>	0.126	0.122	0.082	0.083	0.280	0.321
<i>Growth_EQ_Mean</i>	0.487	0.438	0.436	0.284	0.430	0.324
<i>Growth_G_Loans_Mean</i>	0.152	0.145	0.104	0.112	0.344	0.346
<i>ROA_Mean</i>	0.006	0.004	0.007	0.006	0.006	0.004
<i>EQ_to_TA_Mean</i>	0.106	0.11	0.115	0.121	0.116	0.122
<i>Deposits_to_G_Loans_Mean</i>	14.607	2.228	2.267	1.624	1.955	1.929
<i>L_Imp_to_G_Loans_Mean</i>	0.009	0.005	-0.002	0.012	0.022	0.022
<i>NIM_Mean</i>	0.023	0.021	0.027	0.027	0.032	0.031
<i>CI_Mean</i>	0.646	0.653	0.665	0.679	0.622	0.64
<i>Loans_to_TA_Mean</i>	0.587	0.575	0.571	0.583	0.509	0.508
<i>Loans_to_Funding_Mean</i>	0.856	0.835	0.81	0.778	0.753	0.706
<i>Liquid_A_to_Funding_Mean</i>	0.339	0.385	0.322	0.347	0.485	0.520
Macroeconomic variables						
<i>GDP_growth_Mean</i>	0.019	0.02	0.013	0.016	0.031	0.037
<i>INF_Mean</i>	0.027	0.022	0.019	0.019	0.035	0.032
<i>UNEMPL_Mean</i>	0.058	0.062	0.087	0.080	0.080	0.077

Source: own elaboration

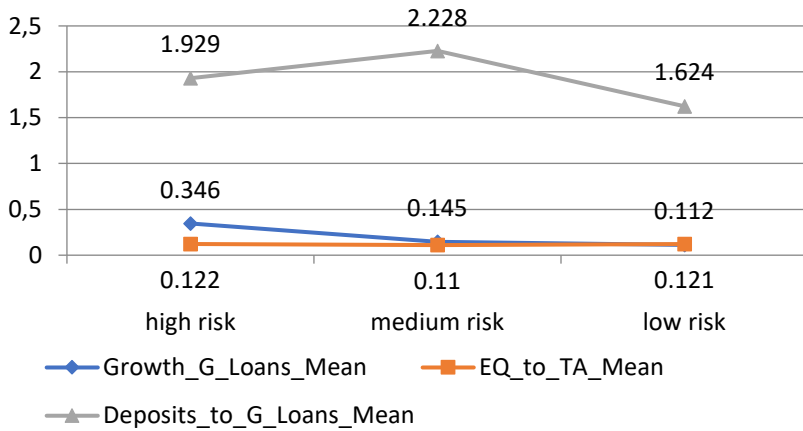


Figure 7. Mean values of financial ratios for clusters – 4 years to distress

Source: own elaboration

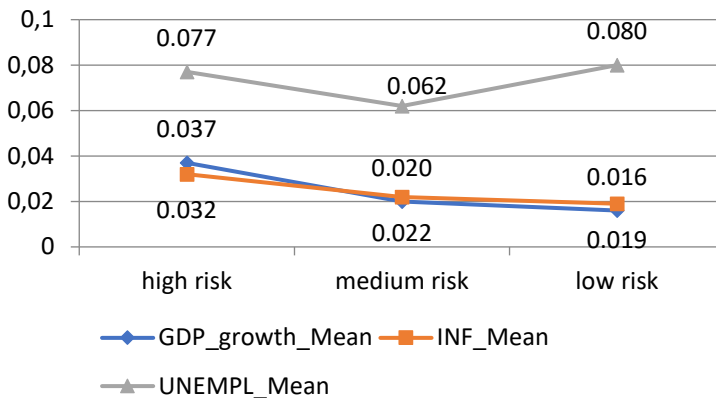


Figure 8. Mean values of macroeconomic variables for clusters – 4 years to distress

Source: own elaboration

5. Conclusions

To sum up, the following CAMEL-like ratios and macroeconomic variables were found to be significant in distress prediction and important in clustering of countries:

1. For 1-year distance to distress four variables were statistically significant: *ROA*, *EQ_to_TA*, *Loans_to_TA* and *Loans_to_Funding*. Among macroeconomic variables high distress probability is connected with low GDP growth. Lower distress was connected with lower unemployment. Only for 1 year distance to distress the unemployment was connected with distress.

2. For 2-year distance to distress four variables were statistically significant: *EQ_to_TA*, *CI*, *Loans_to_TA* and *Loans_to_Funding*. Among macroeconomic variables high distress probability is connected with high GDP growth, conversely to the 1-year distance to distress.
3. For 3-year distance to distress different variables were statistically significant: *Growth_G_Loans*, *NIM*, *CI* and *Loans_to_Funding*. Among macroeconomic variables high distress probability is connected with high GDP growth, conversely to the 1-year distance to distress.
4. For 4-year distance to distress only three variables were statistically significant: *Growth_G_Loans*, *EQ_to_TA* and *Deposits_to_G_Loans*. Among macroeconomic variables high distress probability is connected with high GDP growth, conversely to the 1-year distance to distress.
5. Betz et. el. (2014) obtained similar results, however the prediction period was not fully comparable (quarterly data, 8 quarters prior to distress, recursive model).

Based on the results of prediction models and clustering our research hypotheses were tested. We did not find support for H1 on clear correlation between macroeconomic conditions and distress risk in banking sector. The correlation between macroeconomic conditions and distress risk in banking sector is not clear or obvious. The correlation depends on the amount of time to distress.

We did not find support for H2 either. Clusters of low, medium and high distress risk regions are not homogenous. Homogeneity of clusters depends on the amount of time to distress.

References

- Altman E.I., Cizel J., Rijken H.A. (2014), *Anatomy of bank distress: the information content of accounting fundamentals within and across countries*, <http://dx.doi.org/10.2139/ssrn.2504926>
- Arena M. (2008), *Bank failures and bank fundamentals: A comparative analysis of Latin America and East Asia during the nineties using bank-level data*, "Journal of Banking and Finance", vol. 32(2), pp. 299–310.
- Betz F., Oprica S., Peltonen T., Sarlin P. (2014), *Predicting distress in European banks*, "Journal of Banking and Finance", no. 45, pp. 225–241.
- Cole R.A., White L.J. (2012), *Déjà Vu All Over Again: The Causes of U.S. Commercial Bank Failures This Time Around*, "Journal of Financial Services Research", vol. 42(1), pp. 5–29.
- Cox R.A.K., Wang G.W. (2014), *Predicting the US bank failure : A discriminant analysis*, "Economic Analysis and Policy", vol. 44(2), pp. 202–211.
- Drehmann M., Juselius M. (2014), *Evaluating early warning indicators of banking crises: Satisfying policy requirements*, "International Journal of Forecasting", vol. 30, issue 3, pp. 759–780.
- Hájek P., Olej V., Myšková R. (2015), *Predicting Financial Distress of Banks Using Random Subspace Ensembles of Support Vector Machines*, [in:] R. Silhavy, R. Senkerik, Z. Oplatkova, Z. Prokopova, P. Silhavy (eds.), *Artificial Intelligence Perspectives and Applications. Advances in Intelligent Systems and Computing*, vol. 347, Springer, Cham.

- Hambusch G., Shaffer S. (2016), *Forecasting bank leverage: an alternative to regulatory early warning models*, "Journal of Regulatory Economics", vol. 50(1), pp. 38–69.
- Iwanicz-Drozdowska (ed.) (2016), *European Bank Restructuring During the Global Financial Crisis*, Palgrave Macmillan, London.
- Iwanicz-Drozdowska M., Laitinen E., Suvas A. (2018), *Paths of glory or paths of shame? An analysis of distress events in European banking*, "Bank i Kredyt", vol. 49(2), pp. 115–144.
- Kapinos P., Mitnik O.A. (2016), *A Top-down Approach to Stress-testing Banks*, "Journal of Financial Services Research", vol. 49(2), pp. 229–264.
- Kolari J., Glennon D., Shin H., Caputo M. (2002), *Predicting large US commercial bank failures*, "Journal of Economics and Business", vol. 54(4), pp. 361–387.
- Lopez J.A. (1999), *Using CAMELS ratings to monitor bank conditions*, Federal Reserve Bank of San Francisco Economic Letter, no. 19.
- López Iturriaga F.J., Sanz I.P. (2015), *Bankruptcy visualization and prediction using neural networks: a study of U.S. commercial banks*, "Expert Systems with Applications", no. 42(6), pp. 2857–2868.
- Maghyereh A.I., Awartani B (2014), *Bank distress prediction: Empirical evidence from the Gulf Cooperation Council countries*, "Research in International Business and Finance", no. 30, pp. 126–147.
- Peek J., Rosengren E. (1996), *The use of capital ratios to trigger intervention in problem banks: too little, too late*, "New England Economic Review", September/October, pp. 49–58.
- Peltonen T.A., Piliou A., Sarlin P. (2015), *Network linkages to predict bank distress*, European Central Bank, Working Paper Series, no. 1828.
- Poghosyan T., Čihak M. (2011), *Determinants of Bank Distress in Europe: Evidence from a New Data Set*, "Journal of Financial Services Research", vol. 40(3), pp. 163–184.
- Ravisankar P., Ravi V. (2010), *Financial distress prediction in banks using Group Method of Data Handling neural network, counter propagation neural network and fuzzy ARTMAP*, "Knowledge-Based Systems", vol. 23(8), pp. 823–831.
- Shaffer S. (2012), *Bank failure risk: Different now?*, "Economics Letters", vol. 116(3), pp. 613–616.
- Sinkey J.F. Jr (1975), *A multivariate statistical analysis of the characteristics of problem banks*, "Journal of Finance", vol. XXX(1), pp. 21–36.
- SirElkhatim M.A., Salim N. (2015), *Prediction of Banks Financial Distress*, "SUST Journal of Engineering and Computer Sciences", vol. 16(1), pp. 40–55.
- Wheelock D.C., Wilson P.W. (2000), *Why do banks disappear? The determinants of U.S. bank failures and acquisitions*, "The Review of Economics and Statistics", no. 82(1), pp. 127–138.


Predykcja bankructwa banków – różnice regionalne i uwarunkowania makroekonomiczne

Streszczenie: Artykuł poświęcony został bankructwom banków europejskich z lat 1990–2015. Wykorzystane zostały w nim niezbilansowane dane panelowe dla 3691 banków. Zidentyfikowano 132 bankructwa – zarówno faktyczne, jak i wynikające z konieczności subwencji. Wykorzystano zmienne na poziomie banków typu CAMEL i kontrolne zmienne makroekonomiczne (PKB, inflacja, stopa bezrobocia). Analiza oparta została na tradycyjnym modelu regresji logistycznej do predykcji bankructwa i metodzie *k*-średnich do grupowania. Otrzymane wyniki wskazują, że prawdopodobieństwo bankructwa jest zależne od warunków makroekonomicznych poprzez wyniki klasteryzacji. Zmienne na poziomie banków, które są stabilnym predyktorem bankructwa od roku do czterech lat przed zdarzeniem, to: kapitał do aktywów ogółem (dźwignia) oraz kredyty do funduszy (płynność). Z czynników makroekonomicznych istotne znaczenie ma PKB, ale ze zróżnicowanym wpływem: dla roku

przed bankructwem wysokie prawdopodobieństwo bankructwa jest związane z niską dynamiką PKB, ale dla 2, 3 i 4 lat przed bankructwem wysokie ryzyko bankructwa jest związane z wysoką dynamiką PKB, czyli jest to zależność odwrotna. Pokazuje to zmienną rolę otoczenia makroekonomicznego i wskazuje na potencjalny wpływ sprzyjających warunków makroekonomicznych na powstawanie problemu systemowego w sektorze bankowym.

Słowa kluczowe: bankructwo, CAMEL, predykcja logistyczna, zmienne makroekonomiczne

JEL: G20, G21

	<p>© by the author, licensee Łódź University – Łódź University Press, Łódź, Poland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license CC-BY (http://creativecommons.org/licenses/by/3.0/)</p> <p>Received: 2019-01-12; verified: 2019-07-02. Accepted: 2020-01-03</p>
	<p>This journal adheres to the COPE's Core Practices https://publicationethics.org/core-practices</p>