In search of a measure of banking sector distress: Empirical study of CESEE banking sectors

Paola Bongini

Milan-Bicocca University

Piazza dell'Ateneo Nuovo, 1 - 20126, Milano, Italy

phone: +39 02 6448 3012

Małgorzata Iwanicz-Drozdowska (corresponding author)

Warsaw School of Economics

al. Niepodległości 162, 02-554 Warsaw, Poland

e-mail: miwani@sgh.waw.pl

phone: +48 22 564 93 20

Paweł Smaga

Warsaw School of Economics and National Bank of Poland, Warsaw, Poland

al. Niepodległości 162, 02-554 Warsaw, Poland

e-mail: psmaga@sgh.waw.pl

phone: +48 22 564 60 00

Bartosz Witkowski

Warsaw School of Economics, Warsaw, Poland

al. Niepodległości 162, 02-554 Warsaw, Poland

e-mail: bwitko@sgh.waw.pl

phone: +48 22 564 60 00

In search of a measure of banking sector distress: Empirical study of CESEE banking sectors

We tested the reliability of different versions of the Z-Score and CAMELS-based financial strength indices (aggregated from bank-level) data in detecting periods of banking crises on a sample of 20 Central, Eastern and Southeastern European (CESEE) countries between 1995–2014. We demonstrated that the predictive power of both types of accounting-based measures is weak. Our results cast some doubts on their use in academic research and in macroprudential monitoring framework for emerging countries. Thus, there is a need to strengthen the informational content of accounting data through more frequent and higher quality data disclosures, including exposures allowing for analysis of interconnectedness and network effects for systemic banking risk monitoring.

Keywords: financial strength; Z-Score; CAMELS, crisis

JEL codes: C43, F36, G21

1. Introduction

The increased number, frequency and costs of banking crises, both in developed and emerging countries, have reinvigorated interest in models capable of identifying the occurrence of the crisis in a timely fashion reflected in the change in bank financial ratios.

Bank supervisors usually monitor the potential for idiosyncratic failures in individual banks through the use of the so-called early warning systems (EWS), typically based on accounting data from bank reporting frameworks. Since the 1970s, the CAMELS methodology has been one of the most popular approaches for assessing the financial strength of individual banks. In its basic form, this methodology requires knowledge about a bank's capital adequacy (C), asset quality (A), management (M), liquidity (L), earnings (E) and sensitivity to market risks (S). In more recent academic studies, the CAMELS variables have been substituted with another accounting-based indicator, the Z-score, because of its main advantage of being less data-demanding while delivering similar results (Chiaramonte *et al*, 2015). This ratio requires a limited number of bank-specific variables: the ratio of equity to total assets (ETA), the return on assets (ROA) and its standard deviation (σ_{ROA}). Indeed, the Z-score has become a widely used proxy of bank soundness and has, since the global financial crisis (GFC), been frequently used to analyse the determinants of bank risk-taking in the pre-crisis period (Laeven and Levine, 2009; Foos *et al*, 2010; Altunbas *et al*, 2011; Bertay *et al*, 2013; ECB, 2016).

However, since the outbreak of the global financial crisis in 2007, the policy discussion has been concentrated on how to detect systemic problems so that broader, system-wide banking failures, not only individual ones, could be properly identified and addressed.

Systemic risk has received increased attention in policy discussions and both supervisors and academics have been studying it from different angles. First, an important question in the debate is how to measure systemic risk, with the academia proposing a variety of market-based measures (e.g. CISS by Holló *et al*, 2012 or CoVar by Adrian and Brunnermeier, 2016) and supervisors relying on accounting and confidential supervisory data (for a review see Bongini and Nieri, 2013). A second, related issue is how to deal with those financial institutions deemed to be too systemic/TBTF, in terms of enhanced regulation, supervision and additional prudential requirements (FSB, 2011). A third question concerns taking into consideration the interconnectedness that exists among financial institutions, thus linking the risk of individual institutions with systemic risk. In this regard, a promising literature has made sensible efforts in the study of the structure and dynamics of financial networks and how they react to shocks (Minoiu and Reyes, 2013; Castrèn and Rancan, 2014; Cimini, 2015).

Finally, there is an increased interest in understanding how well aggregate prudential ratios identify banking crises. Substantial efforts have been devoted on national and international level, with the IMF and ECB in the lead, to define and compile so-called financial soundness – or strength – indicators (FSI). These are aggregate measures of the health of a country's financial sector that comprise a key and integral part of the macroprudential toolkit (Sundararajan *et al*, 2002; Costa Navajas and Thegeya, 2013). The IMF's core FSIs cover capital adequacy, asset quality, earnings and profitability, and sensitivity to market risk. Indeed, the monitoring and analysis of systemic stability usually employs, on an aggregate country level, the same prudential ratios that are used to detect individual bank's distress (Čihák and Schaeck, 2010). However, such approach might be put into question, as the soundness of individual banks reflected in accounting-based measures does not capture structural dimension of systemic risk such as contagion and interconnectedness. We thus try to explore this issue by testing the reliability of using aggregated individual bank condition measures to signal systemic banking crises in CESEE countries.

We follow this last stream of studies and attempt to answer the main question as to whether aggregated accounting-based prudential ratios of individual banks are able to identify banking problems on a country level[1]. Our goal is to test, on country-level data, the performance of financial strength indices (FSIs) and various versions of the Z-Score in the detection of banking crises in CESEE countries. This extends the approach by Čihák and Schaeck (2010) and Costa Navajas and Thegeya (2013), who tested the effectiveness of different indicators in signalling banking crises on cross-country datasets, as well as the work of Männasoo and Mayes (2009), which focused on identifying common features of bank distress in the 19 Eastern European transition countries from 1995–2004. We contribute to the literature on early warning systems (Cole and Gunther, 1998; Goldstein *et al*, 2000; Davis and

Karim, 2008; Wong *et al*, 2010) for supervisory purposes by assessing the reliability of Z-Score and CAMELS-based ratios. This warrants improvement of bank risk assessment frameworks and the need to further develop more accurate bank strength measures, including using market-based data, being complementary to the accounting-based approach. This is especially important, as the use of CAMELS and CAMELS-like systems by banking supervisors for the assessment of individual banks is popular in CESEE countries (Green and Petrick, 2002), due to its simple numerical approach.

The structure of the paper is as follows. In the second section, we review the relevant literature, while the third section explains the data sources and methodology of Z-Scores and FSIs. In the fourth section, we present results for the accuracy of Z-Scores and FSIs in identifying past crisis episodes using area under curve (AUC). The final section provides some conclusions.

2. Literature review

The recent global financial crisis that mainly hit developed banking systems has made the use of the so-called early warning systems (EWS) more necessary than ever, not only as a supervisory tool for detecting individual problem banks but also for preventing system-wide collapse.

Typically, EWS can be based on macroeconomic variables or microeconomic variables and can take two different approaches: a univariate signalling approach or a multivariate approach.

The non-parametric, univariate approach looks at the behaviour of individual variables – macro or micro or a combination of the two – around crisis episodes and tries to extract warning signals based on specific thresholds (see, i.a. Kaminsky and Reinhart, 1999; Borio and Lowe, 2002; Borio and Drehmann, 2009; Drehmann and Juselius, 2013). The second, multivariate approach tries to obtain early warning signals by applying statistical techniques, including multiple discriminant analysis, logit or probit regression models, Bayesian model averaging, or classification and regression trees to a set of accounting and/or market information to identify the *ex post* determinants of the distress event (Cole and Gunther, 1998; Bongini *et al*, 2001; Demirgüç-Kunt and Detragiache, 2005; Babecký *et al*, 2014; Joy *et al*, 2015).

EWS based on macro variables are important tools for the timely detection of system-wide banking crises, as highlighted by Demirgüç-Kunt and Detragiache (2000) and Kaminsky and Reinhart (1999) on the East Asian financial crisis. However, these models have several disadvantages. First, they do not allow for determining the cause of bank distress: an exogenous shock or the accumulation of operational and managerial weaknesses within a single institution before the outbreak of the crisis. Second, they leave policymakers with insufficient information as to which specific institutions are the most fragile within the banking system and increase the risk of dealing with banking sector problems at the aggregate level by applying one-size-fits-all solutions to bank distress.

Therefore, EWS based on individual institutional data are the most popular, both in academic literature and as supervisory tools: The early detection of bank distress enables supervisory authorities to undertake prompt corrective actions, with respect to each individual institution, so as to minimise the cost of bank resolution and reduce the risk of domino effects. Individual institutional data can be grouped into two broad categories: market-based and accounting-based. EWS based on market data rely mostly on prices to estimate, alternatively, (i) equity market-based distance to default (Hagendorff and Kato, 2010; Hagendorff and Vallascas, 2011); (ii) bond spreads in the secondary market (Flannery, 1998; Sironi, 2000; Morgan and Stiroh, 2001; Gropp *et al*, 2002); and (iii) CDS spreads (Constantinos, 2010; Volz and Wedow, 2011; Chiaramonte and Casu, 2013). Although market-based indicators of bank distress have the main advantage of being forward-looking in nature since they incorporate market participants' expectations, they also share two major weaknesses that prevent their widespread use as supervisory tools. In fact, the quality of data, i.e., market prices, is conditional on the degree of liquidity, transparency and efficiency of financial markets where bank equities, bonds and CDS are traded. Moreover, the number of listed banks or listed bank bonds tends to be rather limited, especially in emerging countries, as in CESEE, reducing the attractiveness of EWS based on market data.

Therefore, supervisory authorities have, since the 1970s, relied on the CAMELS approach to assess bank risk and vulnerability levering on accounting values (Sinkey, 1979). Financial and accounting ratio proxies for capital adequacy, asset quality, managerial capability, earnings and liquidity and, more recently, sensitivity to market risks are considered relevant signals of incoming imbalances at the individual-bank level. Indeed, the empirical literature on individual-bank distress has widely confirmed the ability of CAMELS ratings to assess bank vulnerability and predict bank distress (for a review see Poghosyan and Cihak, 2011). However, in recent studies, CAMELS variables have been substituted with another book-based indicator, the Z-score, for its main quality of being less data-demanding while delivering similar results (Chiaramonte *et al*, 2015). Indeed, the Z-score has become a widely used proxy of bank soundness, and has, since the global financial crisis (GFC), frequently been used to analyse the determinants of bank risk-taking in the pre-crisis period (Laeven and Levine, 2009; Foos *et al*, 2010; Altunbas *et al*, 2011; Bertay *et al*, 2013; ECB, 2016). The Z-score is in fact a proxy of a bank's distance to default; yet, being an accounting-based measure, it could be computed for both unlisted and listed banks.

It should be noted that EWS based on individual institutional information typically requires the identification of the trigger event (i.e., the bank's distress) at the bank level; these methodologies look at each bank separately; when transposed to detect system-wide crises, these are ‘identified’ when multiple distresses at a single-bank level take place at the same time.

More recent, are those studies that use micro data to create system-wide indicators of financial soundness to measure the health of a country’s financial system. Exceptions are the works by Čihák and Schaeck (2010) and Costa Navajas and Thegeya (2013). The first study tested the effectiveness of different financial soundness indicators in signalling banking crises and highlighted significant correlations between some FSIs and the occurrence of banking crises. The latter study assessed the viability and efficiency of EWS to predict banking crises. The authors also suggested different statistical methodologies, according to whether a global or a country-specific crisis was under investigation. They also underlined the need to consider the policymakers’ objectives when designing predictive models and setting related thresholds, since there is a sharp trade-off between correctly identifying crises and false alarms.

3. Methodology and data

The sample consisted of 20 CESEE countries over the period 1995–2014. Given the partly incomplete data, the final sample consisted of 355 country-year observations. We used bank-level data from Bankscope for computation of the Z-Scores and FSIs. In the sample, we included all banks that reported data to Bankscope[2] in the analysed period, both listed and non-listed. As already mentioned, the share of listed banks varied in CESEE countries, from 0% to 60%, with on average only one-fifth of banks being listed in each post-communist country (see Table 1). There does not seem to be a pattern indicating that a larger number of total banks in a given country imply a higher share of listed banks.

Table 1. Number of banks and share of listed banks.

Country	Total number of banks	Share of listed banks
Albania	18	0%
Bosnia and Herzegovina	37	54%
Bulgaria	31	13%

Belarus	28	0%
Czech Republic	46	4%
Estonia	11	9%
Croatia	61	20%
Hungary	57	4%
Kosovo	5	0%
Lithuania	10	10%
Latvia	20	0%
Moldova	13	38%
Montenegro	14	57%
Macedonia	20	60%
Poland	68	22%
Romania	35	9%
Serbia	38	16%
Slovenia	26	0%
Slovakia	24	17%
Ukraine	158	24%

Note: Number of banks covers only active commercial banks as of end 2015 (excluding national central banks). The share of listed banks is given as a ratio of listed banks to total number of banks

Source: Author's calculations based on Bankscope data.

The Z-Score is usually computed as:

$$Z - Score = \frac{ROA+ETA}{\sigma_{ROA}} \quad (1)$$

Given the three basic ingredients, the final estimate depends on how each variable is measured. In the literature (see table 2), we find a plethora of approaches, mixing current and average values of the variables used for the numerator and rolling windows or sample period observations for the denominator.

Table 2. Summary of approaches used to calculate Z-Score.

Study	ROA	ETA	Standard deviation of ROA	Z-Score formulas used
Boyd and Graham (1986) Hannan and Hanweck (1988)	Current value	Current value	Over 3-years rolling time window	
Boyd *et al* (2006)	Current value	Average over 3-years rolling time window	Over 3-year rolling time window	Z-Score (4); 'instantaneous' standard deviation estimates in each year as INST $\sigma_{ROA\,t}$ = ROA_t - μ_{ROA} (over the whole period) and ETA_t and ROA_t
Beck and Laeven (2006)	Current value	Current value	Over the full sample period	Z-Score (2); ETA_t and ROA_t and σ_{ROA} for each bank for the whole period
Maecheler *et al* (2007)	3-year rolling average	Average over 3-years rolling time window	Over 3-year rolling time window	Z-Score (5); for all variables we use averages on a 5-year rolling time window

Hesse and Cihak (2007)	Average value over sample period	Current value	Over the full sample period	Z-Score (2); ETA_t, ROA_t and σ_{ROA} for each bank for the whole period
Yeyati and Micco (2007)	Average value over sample period	Current value	Over 3-year rolling time window	Z-Score (7); ETA_t and μ_{ROA} for the whole sample, as well as σ_{ROA} over a 5-year rolling time window
Laeven and Levine (2009)	Average value over sample period	Average value over sample period	Over the full sample period	
Bertay *et al* (2013)	Average value over sample period	Average value over sample period	Over the full sample period	
Lepetit and Stroebel (2013)	Average value over sample period	Current value	Over the full sample period	Z-Score (1), ETA_t and both μ_{ROA} and σ_{ROA} for each bank for the whole sample
Chiaramonte *et al* (2015)	Current value	Current value	Over 3-year rolling time window	
ECB (2016)	Current value	Current value	Over 5-year rolling time window	Z-Score (6); ETA_t and ROA_t while $moving\ \sigma_{ROA}$ is calculated over a 5-year rolling window

Note: For Z-Score (3) we included ROA_t and both $\mu_{\text{EQ\_to\_TA}}$ and σ_{ROA} for each bank for the whole sample.

Source: Author's compilation based on the referred literature.

To ensure the robustness of our results we calculate the Z-Score according to seven alternative formulas (see Boyd *et al* 2006; Beck and Laeven 2006; Maecheler *et al*, 2007; Yeyati and Micco 2007; Lepetit and Strobel 2013, 2015; ECB 2016):

$$Z - Score_1 = \frac{\text{ETA}_t + \mu_{ROA}}{\sigma_{ROA}}, \qquad (2)$$

$$Z - Score_2 = \frac{\text{ETA}_t + \text{ROA}_t}{\sigma_{ROA}}, \qquad (3)$$

$$Z - Score_3 = \frac{\mu_{\text{ETA}} + \text{ROA}_t}{\sigma_{ROA}}, \qquad (4)$$

$$Z - Score_4 = \frac{\text{ETA}_t + \text{ROA}_t}{\text{INST}\ \sigma_{ROA}}, \qquad (5)$$

$$Z - Score_5 = \frac{\text{moving}\ \mu_{\text{ETA}} + \text{moving}\ \mu_{ROA}}{\text{moving}\ \sigma_{ROA}}, \qquad (6)$$

$$Z - Score_6 = \frac{\text{ETA}_t + \text{ROA}_t}{moving\ \sigma_{ROA}}, \qquad (7)$$

$$Z - Score_7 = \frac{\text{ETA} + \mu_{ROA}}{moving \sigma_{ROA}}, \qquad (8)$$

We followed the approach for constructing a compound FSI by aggregating weighted and normalised variables similar to IMF financial soundness indicators, which use bank balance sheet data (see e.g. Geršl and Heřmánek 2006; Jahn and Kick 2012; Ginevičius and Podviezko 2013; Petrovska and Mihajlovska 2013). Most attempts have tried to provide an *ex-post* assessment of banking sector conditions and are only single-country studies used for analytical purposes by the central banks. Moreover, there have been few cross-country studies with FSIs (e.g. Cardarelli *et al*, 2011; Slingenberg and de Haan, 2011; Cevik *et al*, 2013; Vermeulen *et al*, 2015), yet none for a wide panel of CESEE countries.

The FSI is similar to the CAMELS approach (Lopez, 1999). We included in the FSIs five bank-specific characteristics (after empirical normalisation) that took into account the availability of complete data for capital adequacy, profitability, liquidity (two variables) and assets quality. We calculated two indices: (a) using equal weights – we call it FSI, given by formula (9); and (b) with weights assigned using principal components analysis (PCA)[3] – we call it FSI_PCA. PCA has been used by, e.g., Klomp and de Haan (2012) and Demirgüc-Kunt *et al* (2015) for comparative analysis. The PCA approach is used to determine a small number of unobserved factors that explain the maximum of variance in the data (Suhr, 2005). In our index, PCA was based on a country group of variables that were assumed to be linearly correlated, while the proportion of variance described by each extracted factor was assumed to be time-constant. According to Kaiser–Guttman's rule, we retained only those characteristics with eigenvalues higher than 1.

$$FSI = 0.2 \cdot \mathrm{ETA} + 0.2 \cdot \mathrm{ROA} + 0.2 \cdot \mathrm{LAF} - 0.2 \cdot \mathrm{LD} - 0.2 \cdot \mathrm{LITA} \quad (9)$$

where:

LAF – liquid assets to total funding ratio

LD – loans to customers to deposits from customers ratio

LITA – loans impairment charges to total assets ratio

Higher values of FSIs illustrate better situation of the banking system, thus their interpretation is similar to that of a Z-Score. Nevertheless, Z-Score reflects only the profitability and solvency aspects of the bank intermediation activity, while our FSI also accounts for liquidity issues. In the construction of the ratios we assumed that prolonged illiquidity might lead to insolvency.

In order to estimate country-level scores, we weighted each Z-Score (or FSIs) with total assets of the given bank to calculate the aggregated asset-weighted values on the country level for each year over the period 1995–2014.

The average asset-weighted values of Z-Scores, FSI and FSI_PCA for all countries are presented in Figure 1.

[FIGURE 1 HERE]

4. Empirical results

While higher values of Z-Scores and FSI are expected to be associated with a lower risk of banking crisis, the question is what values of Z-Score and FSI should be interpreted as 'likely to indicate a crisis'. Let C_{it} be 1 if there is a crisis in country *i* in year *t* and 0 otherwise.

Crisis events in CESEE are presented in Table 3; and, in the considered sample, there were 11% crisis observations. Banking crises in CESEE were most frequent in the early-to-mid-1990s, yet very few of them suffered the direct impact of the GFC. Crises in CESEE were linked to the process of economic transformation towards a market economy and the liberalisation of banking systems. Banking sector vulnerabilities in CESEE were caused mainly by the high legacy stock of NPLs, foreign exchange volatility and significant penetration of cross-border financial services (Dietrich *et al*, 2011). Yet, the use of banking crisis dates in CESEE as criteria to assess the effectiveness of the signalling power of bank condition measures must be assessed with caution. This is because the timing of crises in the CESEE was at the beginning of our sample period, when financial data reporting standards were underdeveloped in banks in CESEE at that time.

Table 3. Banking crisis events in CESEE countries.

Country	Crisis
Albania	1995-1996
Bosnia and Herzegovina	1995-1996
Bulgaria	1995-1998
Belarus	1995
Czech Republic	1996-2000
Estonia	1995-1996; 1998
Croatia	1996; 1998-1999
Hungary	1995; 2008
Lithuania	1995-1996; 1998-1999; 2013
Latvia	1995-1999; 2008-2010
Macedonia	1995
Poland	1995
Romania	1995-1996
Slovenia	2008; 2011-2013
Slovakia	1995; 1998-2002
Ukraine	1995-1999; 2008-2010

Note: There were no banking crises identified in Kosovo, Moldova, Montenegro and Serbia, because according to the referenced studies, instability in banking systems in those countries did not meet the crisis classification criteria which are based i.a. on cost of policy interventions in the banking sector and output gap.

Source: Based on Caprio and Klingebiel (2003), Laeven and Valencia (2012), Chaudron and de Haan (2014), and our own work.

Let $c_{M_{it}}$ be 1 if a crisis is predicted in country *i* in year *t* based on indicator *M* and 0 otherwise, where *M*∈{Z-Score1, …, Z-Score7, FSI}, *i*=1,…,20, *t*=1995,…,2014. Further, let m_{it} be the value of indicator *M* in country *i* in year *t*. We conclude that

$$c_{M_{it}} = \begin{cases} 1 \ if \ m_{it} \leq \tau_M \\ 0 \ if \ m_{it} > \tau_M, \end{cases} \tag{10}$$

where τ_M is an unknown threshold, specific for indicator *M* and constant over time and space.

We propose to find the optimal τ_M for each of the indicators *M* in the considered sample by maximising the total number of correct predictions in the sample, that is, by solving (with respect to τ_M)

$$max_{\tau_M} \sum_{i,t} I_{c_{M_{it}}=c_{it}}, \tag{11}$$

where $I_{condition}$ is the indicator function, which equals 1 if the logical condition is true and 0 otherwise. Notably, we need to restrict the set of possible τ_M to the values of m_{it}, *i*=1,…,20, *t*=1995,…,2014, which are present in the sample. This is because the value of the optimised function would be the same for all τ_M between two neighbouring values of m_{it} put in ascending order, thus making the problem unidentified. The thresholds (cut-off points) τ_M obtained by solving (11) are given in Table 4. We also provide the number of in-sample cases for which the existing crisis was correctly predicted $n_{C=1,c=1}$, the non-existing crisis was predicted $n_{C=0,c=1}$, the existing crisis was not predicted $n_{C=1,c=0}$ and the non-existing crisis was correctly predicted as a non-crisis $n_{C=0,c=0}$. The last two columns present the sensitivity and specificity of the model. The former is defined as a conditional probability $P(c_{M_{it}} = 1|C_{it} = 1, \tau_M)$, which is estimated in the sample as the $n_{C=1,c=1}/(n_{C=1,c=1}+n_{C=1,c=0})$ and should be interpreted as the fraction of correctly predicted existing crises with thresholds equal to τ_M. The latter is, in turn, given by $P(c_{M_{it}} = 0|C_{it} = 0, \tau_M)$, which is estimated as $n_{C=0,c=0}/(n_{C=0,c=1}+n_{C=0,c=0})$, and represents the fraction of correctly predicted non-crisis periods.

Table 4. Z-Scores and FSI cut-offs τ_M found by optimising the total number of correct predictions.

Indicator M	cut-off τ_M	$n_{C=1,c=1}$	$n_{C=0,c=1}$	$n_{C=1,c=0}$	$n_{C=0,c=0}$	sensitivity	specificity
ZS1	-5,55995	0	1	39	315	0	0,996835
	1,729434	3	4	36	312	0,076923	0,987342
ZS2	4,847415	11	11	28	305	0,282051	0,96519
	4,461099	9	9	30	307	0,230769	0,971519
ZS3	4,702194	9	4	30	312	0,230769	0,987342
	4,609509	8	3	31	313	0,205128	0,990506
	5,035521	10	5	29	311	0,25641	0,984177
ZS4	-1292,62	0	1	39	315	0	0,996835
ZS5	0,862776	4	3	35	313	0,102564	0,990506
ZS6	1,392625	6	5	33	311	0,153846	0,984177
ZS7	1,054973	5	4	34	312	0,128205	0,987342
FSI	-0,13301	0	1	39	315	0	0,996835
	-0,06682	1	2	38	314	0,025641	0,993671
FSI_PCA	-0,12725	0	1	39	315	0	0,996835
	-0,06096	1	2	38	314	0,025641	0,993671

Source: *own calculations*

Although the number of correct predictions $n_{c=1,c=1}+n_{c=0,c=0}$ is in each case between 315 and 319 out of the 355 considered cases (88–90%), such a success rate is due to predicting hardly any crises by the indicators, and thus high specificity, yet very low (in some cases even zero) sensitivity of the predictions. As a result, the globally optimal algorithm fails to fulfil its role—predicting the crises. We thus repeat the optimisation process, additionally imposing a constraint requiring sensitivity for a given indicator to be at least s:

$$max_{\tau_M} \sum_{i,t} I_{c_{M_{it}}=C_{it}} \text{ subject to } max_{\tau_M} \sum_{i,t} I_{c_{M_{it}}=C_{it}} \quad (12)$$

We arbitrarily, yet realistically, apply s=0.8. A high value is required if the indicators are not supposed to omit crises in a high fraction of cases, which should be their main feature. The resulting optimal τ_M as well as numbers of particular combinations of C_{it}, $c_{M_{it}}$, sensitivity and specificity are given in Table 5.

Table 5. Z-Score and FSI cut-offs τ_M found by optimising the total number of correct predictions subject to a minimum 80% sensitivity constraint.

Indicator M	cut-off τ_M	$n_{C=1,c=1}$	$n_{C=0,c=1}$	$n_{C=1,c=0}$	$n_{C=0,c=0}$	sensitivity	specificity
Z-Score 1	15,2372	32	234	7	82	0,820513	0,259494
Z-Score 2	14,98685	33	223	6	93	0,846154	0,294304
Z-Score 3	14,94567	32	218	7	98	0,820513	0,310127
Z-Score 4	32,12521	32	264	7	52	0,820513	0,164557
Z-Score 5	27,21621	32	150	7	166	0,820513	0,525316
Z-Score 6	24,76684	32	136	7	180	0,820513	0,569620
Z-Score 7	24,64899	32	144	7	172	0,820513	0,544304
FSI	0,288821	35	294	4	22	0,897436	0,069620
FSI_PCA	0,304056	33	289	6	27	0,846154	0,085443

Source: own calculations

While in this case sensitivity fulfils the imposed constraint, and the considered indicators detect most of the crises, the rates of false alarms are unacceptably high. As a result, the total fraction of correct forecasts does not exceed 25%.

As it can be seen, the share of successful judgements or forecasts (usually referred to as count R-squared) can be misleading if used to describe the predictive power of the model or indicator used. In the considered case, highly asymmetric distribution of the crisis variable, which suggests existence of a crisis in only 11% of the cases, results in the values of count R-squared of over 82% for each of the Z-Score and FSI indicators as long as no constraints of minimum sensitivity are imposed in the optimisation process. However, this is achieved at the cost of an almost 'always guess no crisis' policy, which, in turn, results in a very low model sensitivity, that is, the probability of identifying a truly existing crisis. We thus further explore the discriminative power of the considered indicators measured by the commonly used AUC (area under ROC curve). AUCs values for subsequent indicators are given in Table 6.

Table 6. AUCs for Z-Scores and FSI.

Variable	AUCs
Z-Score 1	0,653116
Z-Score 2	0,661068
Z-Score 3	0,683382
Z-Score 4	0,459023
Z-Score 5	0,746592
Z-Score 6	0,751542
Z-Score 7	0,748215
FSI	0,405875
FSI_PCA	0,396787

Source: own calculations

The discriminative power of compound measures (FSI, FSI_PCA), being more comprehensive than Z-Scores, as well as some Z-Scores (1–4), is poor. Only Z-Scores 5–7 performed relatively better than the other measures; however, their AUCs were only average. Even the inclusion of country-specific time persistent effects does not improve the results. This shows that distress detection should not be based only on accounting data reported by banks. On one hand, this may be attributed to the changes in the accounting standards (e.g., in the EU, implementation of IAS/IFRS from 2004) and too much discretion in their application. On the other hand, the two waves of banking crises in the CESEE region (the mid-90s and the GFC) were much different in their roots. Economic and political transformation caused the so-called ‘transformation crisis’ in the mid-90s, while later on, the main reason was excessive risk taking. Moreover, if the crisis had occurred at the beginning of the year, the intervention of the safety net players or the owners may have improved the financial position of banks by the end of the year. This suggests the need to use quarterly data rather than annual data for the computation of such indices.

5. Conclusions

Our study points out the weaknesses of aggregated bank-level accounting-based measures as predictors of system-wide bank distress. In contrast to Männasoo and Mayes (2009), we cannot confirm that ‘CAMELS’ factors play an important role in banking sector distress identification and early warning in CESEE. Consequently, an important lesson is that the use of Z-Scores for measuring the financial strength of the overall banking system (e.g., as in Demirgüç-Kunt *et al*, 2008; Chiaramonte *et al*, 2015) needs to be reconsidered. Our results support the conclusions of Cole and Gunther (1998) that the informational quality of CAMELS-based ratios for assessing bank strength is rather weak. Although we acknowledge that CAMELS framework initially was not built to handle systemic crises, results show that it would need to be significantly enhanced to be effectively used in macroprudential monitoring in CESEE countries. The results are also in line with Cihák and Schaeck (2010), who cautioned against using aggregate prudential indicators for identifying banking crises, as they may disguise problems in individual banks or groups of banks and are typically based on backward-looking data. Given that accounting-based measures are widely used in academic research and as supervisory tools, we therefore argue that their deficiencies might be partially reduced by increased reliability in accounting information and limiting any potential for smoothing reported figures.

However, these measures cannot be easily complemented by market-based measures. In CESEE countries, as in many other emerging countries, the low number of listed banks, partly due to the ownership structure of the banking sector, which is largely held in the hands of foreign capital, makes marked-based measures a questionable choice to serve as a proxy for the strength of the banking sector. Additionally, those ownership ties and cross-border financial exposures in CESEE countries represent the interconnectedness among financial institutions and network effects, which would have to be taken into account in constructing a ratio powerful enough to identify systemic – not only individual – bank distress, thus constituting a valuable future research venue.

FUNDING AND ACKNOWLEDGEMENTS

This work was supported by the Polish National Science Center (NCN) under Grant number UMO-2014/13/B/HS4/01619. The opinions expressed herein are those of the authors and do not reflect those of the associated institutions.

Notes

1 At the same time, we are unable to use complementarily market-based measures, given their limited reliability in CESEE countries. An inadequate number of listed banks, combined with relatively small and less liquid stock markets deliver an insufficient quality for market-based bank risk measures.

2 We are constrained to rely on only publicly available data, although we acknowledge that supervisors can rely on better quality and more detailed, yet confidential, data from bank reporting templates.

3 The authors would like to thank Karol Rogowicz for his helpful assistance.

REFERENCES

Adrian, T. and Brunnermeier, M.K. (2016) CoVaR. *American Economic Review* 106.7:1705-1741.

Altunbas, Y., Manganelli, S. and Marques-Ibanez, D. (2011) *Bank Risk During The Financial Crisis Do Business Models Matter?* ECB Working Paper 1394.

Babecký, J., Havránek, T., Matějů, J., Rusnák, M., Šmídková, K. and Vašíček, B. (2014) Banking, debt, and currency crises in developed countries: Stylized facts and early warning indicators. *Journal of Financial Stability* vol. 15(C): 1-17.

Beck, T. and Laeven, L. (2006) *Resolution of failed banks by deposit insurers: Cross-country evidence.* World Bank policy paper 3920.

Bertay, A.C., Demirgüç-Kunt, A. and Huizinga, H. (2013) Do we need big banks? Evidence on performance, strategy and market discipline. *Journal of Financial Intermediation* Vol. 22 (4): 532-558.

Bongini, P. and Nieri, L. (2014) Identifying and Regulating Systemically Important Financial Institutions. *Economic Notes* Vol. 43: 39-62.

Bongini, P., Claessens, S. and Ferri, G. (2001) The political economy of distress in East Asian financial institutions. *Journal of Financial Services Research* 19: 5–25.

Borio, C. and Drehmann, M. (2009) *Assessing the risk of banking crises- revisited*. BIS Quarterly Review, March.

Borio, C. and Lowe, P. (2002) *Assessing the risk of banking crises*. BIS Quarterly Review, December.

Boyd, J., De Nicoló, G. and Jalal, A. (2006) *Bank risk-taking and competition revisited: New theory and new evidence.* IMF Working Paper 06/297, IMF.

Boyd, J.H. and Graham, S.L. (1986) *Risk, regulation and banking holding company expansion into nonbanking*. Quarterly Review 10, Federal Reserve Bank of Minneapolis.

Caprio, G. and Klingebiel, D. (2003) *Episodes of Systemic and Borderline Financial Crises*. World Bank.

Cardarelli, R., Elekdag, S. and Lall, S. (2011) Financial stress and economic contractions. *Journal of Financial Stability* 7:78–97.

Castrèn, O. and Rancan, M. (2014) Macro networks: an application to euro area financial accounts. *Journal of Banking and Finance* 46: 43-58.

Cevik, E.I., Dibooglu, S. and Kutan, A.M. (2013) Measuring financial stress in transition economies. *Journal of Financial Stability* Volume 9 Issue 4 December: 597-611.

Chaudron, R. and de Haan, J. (2014) *Identifying and dating systemic banking crises using incidence and size of bank failures.* DNB Working Paper No. 406, DNB.

Chiaramonte, L. and Casu, B. (2013) The determinants of bank CDS: Evidence from the financial crisis. *European Journal of Finance* 19: 861–887.

Chiaramonte, L., Croci, E. and Poli, F. (2015) Should we trust the Z-score? Evidence from the European banking industry. *Global Finance Journal* 28: 111–131.

Čihák, M. and Schaeck, K. (2010) How well do aggregate prudential ratios identify banking system problems? *Journal of Financial Stability* Volume 6 Issue 3: 130-144.

Cimini, R. (2015) Eurozone network "connectedness" after fiscal year 2008. *Finance Research Letters* 14: 160-166.

Cole, R.A. and Gunther, J.W. (1998) Predicting Bank Failures: A Comparison of On- and Off-Site Monitoring Systems. *Journal of Financial Services Research* Vol. 13: 103.

Constantinos, S. (2010) *Rethinking market discipline in banking: lessons from the financial crisis.* Policy Research Working Papers 5227, World Bank.

Costa Navajas, M. and Thegeya, A. (2013) *Financial soundness indicators and banking crises.* IMF Working Paper WP/13/263, IMF.

Davis P.E. and Karim, D. (2008) Comparing early warning systems for banking crises. *Journal of Financial Stability* Volume 4 Issue 2: 89-120.

Demirgüç-Kunt A., Kane, E. and Laeven, L. (2015) Deposit insurance around the world: A comprehensive analysis and database. *Journal of Financial Stability* 20: 155–183.

Demirgüç-Kunt, A. and Detragiache, E. (2000) *Monitoring banking sector fragility: a multivariate logit approach with an application to the 1996-97 banking crisis*. World Bank Economic Review, Vol. 14, No. 2: 287-307.

Demirgüç-Kunt, A. and Detragiache, E. (2005) *Cross-Country Empirical Studies of Systemic Bank Distress. A Survey*. IMF Working Paper WP/05/96 , IMF.

Demirgüç-Kunt, A., Detragiache, E. and Tressel, T. (2008) Banking on the principles: Compliance with Basel Core Principles and bank soundness. *Journal of Financial Intermediation* Volume 17 Issue 4: 511-542.

Dietrich, D., Knedlik, T. and Lindner, A. (2011) Central and Eastern European countries in the global financial crisis: a typical twin crisis? *Post-Communist Economies* Vol. 23, Issue 4.

Drehmann, M. and Juselius, M. (2013) *Evaluating early warning indicators of banking crises: Satisfying policy requirements*. BIS Working Paper Series No. 421, BIS.

ECB (2016) *Recent trends in euro area banks' business models and implications for banking sector stability*. Financial Stability Review, May.

Flannery, M.J. (1998) Using market information in prudential banking supervision. A review of US evidence. *Journal of Money, Credit and Banking* 30: 273–305.

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

FSB (2011) Policy Measures to Address Systemically Important Financial Institutions.

Geršl, A. and Heřmánek, J. (2006) *Financial stability indicators: Advantages and disadvantages of their use in the assessment of financial system stability*. Financial Stability Report, CNB.

Ginevičius, R. and Podviezko, A. (2013) The evaluation of financial stability and soundness of Lithuanian banks. *Ekonomska istraživanja* 26(2): 191-208.

Goldstein, M., Kaminsky, G.L. and Reinhart, C.M. (2000) *Assessing Financial Vulnerability: An Early Warning System for Emerging Markets*. Washington: Peterson Institute.

Green D. and Petrick, K. (eds.) (2002) *Banking and Financial Stability in Central Europe. Integrating Transition Economies into the European Union*. Cheltenham, Northampton: Edward Elgar Publishing Ltd.

Gropp, R., Vesala, J. and Vulpes, G. (2002) *Equity and bond market signals as leading indicators of bank fragility*. ECB Working Paper Series 150.

Hagendorff, J. and Kato, P. (2010) Distance to default, subordinated debt, and distress indicators in the banking industry. *Accounting and Finance* 50: 853–870.

Hagendorff, J. and Vallascas, F. (2011) CEO pay incentives and risk-taking: Evidence from bank acquisitions. *Journal of Corporate Finance* 4: 1078–1095.

Hannan, T. H. and Hanweck, G. A. (1988) Bank insolvency risk and the market for large certificates of deposit. *Journal of Money, Credit and Banking* 20: 203–211.

Hesse, H. and Cihák, M. (2007) *Cooperative banks and financial stability*. IMF Working Paper WP/07/2, IMF.

Holló, D., Lo Duca, M. and Kremer, M. (2012) *CISS – a composite indicator of systemic stress in the financial system*. ECB Working Paper Series No. 1426.

Jahn, N. and Kick, T. (2012) *Early warning indicators for the German banking system: A macroprudential analysis*. Bundesbank Discussion Paper No. 27/2012.

Joy, M., Rusnák, M., Šmídková, K. and Vašíček, B. (2015) *Banking and currency crises: differential diagnostics for developed countries*, ECB Working Paper Series No. 1810.

Kaminsky, G. and Reinhart, C. (1999) The twin crises: The causes of banking and balance-of-payments problems. *American Economic Review* 89 (3): 473-500.

Klomp, J. and de Haan, J. (2012) Banking risk and regulation: Does one size fit all? *Journal of Banking and Finance* Vol. 36 (12): 3197-3212.

Laeven, L. and Levine, R. (2009) Bank governance, regulation and risk taking. *Journal of Financial Economics* 93: 259–275.

Laeven, L. and Valencia, F. (2012) *Systemic banking crises database: An update.* IMF Working Paper WP/12/163, IMF.

Lepetit, L. and Strobel, F. (2013) Bank insolvency risk and time-varying Z-score measures. *Journal of International Financial Markets, Institutions and Money* 25: 73–87.

Lepetit, L. and Strobel, F. (2015) Bank insolvency risk and Z-score measures: A refinement. *Finance Research Letters* 13:214–224.

Lopez, J.A. (1999) *Using CAMELS ratings to monitor bank conditions*. Federal Reserve Bank of San Francisco Economic Letter No. 19.

Maecheler, A., Srobona, M. and DeLisle, W. (2007) *Decomposing financial risks and vulnerabilities in Eastern Europe*. IMF Working Paper 248, IMF.

Männasoo, K. and Mayes, D. G. (2009) Explaining bank distress in Eastern European transition economies. *Journal of Banking and Finance* Volume 33 Issue 2: 244-253.

Minoiu, C. and Reyes, J.A. (2013) A network analysis of global banking: 1978-2010. *Journal of financial stability* 9: 168-184.

Morgan, D. and Stiroh, K. (2001) *Bond market discipline of bank: Is the market tough enough?* Federal Reserve Bank of New York Working Paper No. 95.

Petrovska, M. and Mihajlovska, E. M. (2013) Measures of Financial Stability in Macedonia. *Journal of Central Banking Theory and Practice* vol. 2, no. 3.

Poghosyan, T. and Čihák, M. (2011) Distress in European banks: An analysis based on a new dataset. *Journal of Financial Services Research* 40: 163–184.

Sinkey, J. (1979) *Problem and failed institutions in the commercial banking industry*. Contemporary studies in economic and financial analysis, 4, Amsterdam, NL: JAI Press.

Sironi, A. (2000) *Testing for market discipline in the European banking industry: Evidence from subordinated debt spreads*. Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series No. 40.

Slingenberg, J.W. and de Haan, J. (2011) *Forecasting financial stress*. De Nederlandsche Bank Working Paper No. 292.

Suhr, D. D. (2005) *Principal Component Analysis vs. Exploratory Factor Analysis*. Paper 203-30, SUGI 30 Proceedings, University of Northern Colorado.

Sundararajan, V., Enoch, C., San José, A., Hilbers, P., Krueger, R., Moretti, M. and Slack, G. (2002) *Financial soundness indicators: analytical aspects and country practices*. IMF Occasional Paper No. 212.

Vermeulen, R., Hoeberichts, M., Vašíček, B., Žigraiová, D., Šmídková, K. and Haan, J. (2015) Financial Stress Indices and Financial Crises. *Open Economies Review* 26: 383–406.

Volz, M. and Wedow, M. (2011) Market discipline and too-big-to-fail in the CDS market: Does banks' size reduce market discipline? *Journal of Empirical Finance* 18: 195–210.

Wong, J., Wong, T-Ch. and Leung, P. (2010) Predicting banking distress in the EMEAP economies. *Journal of Financial Stability* Volume 6 Issue 3:169-179.

Yeyati, E.L., and Micco, A. (2007) Concentration and foreign penetration in Latin American banking sectors: Impact of competition and risk. *Journal of Banking and Finance* 31:1633–1647.