Visnyk of the National Bank of Ukraine, 2020, No. 250, pp. 33–44

PREDICTING BANK DEFAULTS IN UKRAINE: A MACRO-MICRO PERSPECTIVE

ANATOLII HLAZUNOV[a], OLESIA VERCHENKO[b]

[a]National Bank of Ukraine
E-mail: anatolii.hlazunov@bank.gov.ua
[b]Kyiv School of Economics
E-mail: verchenko@kse.org.ua

Abstract This paper develops an early warning model (EWM) for a micro-macro analysis of individual and aggregated bank vulnerabilities in Ukraine. We applied a stepwise logit for predicting defaults at Ukrainian banks based on a panel bank and macro-level data from Q1 2009 to Q3 2019. Next, we aggregated individual bank default probabilities to provide policymakers with information about the general state of the financial system with a particular focus on generating a signal for countercyclical capital buffer (CCB) activation. Our key findings suggest that the probability of default exceeding 11% could signal about a vulnerable state in a bank and, in the aggregated model, in a financial system in general. The aggregated model successfully issues an out-of-sample signal of a systemic crisis four periods ahead of the start of the 2014-2015 turmoil.

JEL Codes E44, E58, G01

Keywords early warning models, bank default probability, countercyclical capital buffer

1. INTRODUCTION

In 2015-2017, Ukraine went through a period of a large-scale banking system “cleaning”, the outcome of which was a decrease in the number of commercial banks from 163 at the beginning of 2015 to 82 at the end of 2017. As a result, the Ukrainian banking system emerged much more resilient to external and internal shocks, the evidence of which is in its remarkable stability during the current turmoil of the COVID-19 crisis.

The scale of bank closures over such a short period was probably unprecedented in recent economic history. There were also remarkable differences in their causes. Some banks were unwilling or unable to disclose their ownership structure in line with the new banking system transparency requirements. The National Bank of Ukraine (NBU) revoked the licenses of 24 such banks. Sixty-eight banks left the market due to an inability to comply with capital requirements and/or other regulated liquidity and financial stability ratios. Some banks experienced financial distress (having to restructure their NBU loans) but managed to stay in operation.

From an academic standpoint, this data provide very rich grounds for empirical analyses of individual bank responses to the new environment and their aggregate effect on the banking system. This is in contrast to numerous existing studies for other countries, which almost uniformly point out that bank distress events are quite rare (e.g., Betz et al, 2014, for the European Union, Rosa and Gartner, 2018, for Brazil, etc.).

While bank closures related to noncompliance with transparency (financial monitoring) requirements deserve a separate in-depth analysis, the focus of this paper is on distress events related to capital and other financial requirements. Though the period of 2015-2017 provides a particularly interesting case study on banking system reform, to make our conclusions more general, we decided to look beyond this period and considered all available data on bank distress events in Ukraine starting from Q1 2009. This also gained us additional degrees of freedom to produce forecasts and test their accuracy.

More specifically, the goal of this study is to conduct a micro-macro analysis of individual and aggregated bank vulnerabilities in Ukraine. For this, we first estimate a bank-level early warning model containing both bank-specific and macroeconomic factors. Early warning models for bank distress events are an important tool for the banking supervision framework as defined by Pillar 2 of the Basel recommendations developed by the Basel Committee on Banking Supervision (2010) to strengthen the regulation, supervision and risk management of banks.

While such models can be used to derive probabilities of individual banks being in vulnerable states, they can also be used to provide policymakers with information on the general state of the financial system, and to signal for a call to action. This is the second step of our analysis: to aggregate bank-level results with the purpose of detecting a buildup of systemic imbalances and possibly intervening with additional macroprudential measures. In this regard, we attempt to apply an aggregation methodology to our baseline bank-level model as a framework for operationalizing the countercyclical capital buffer (CCB) in Ukraine.

The CCB is a part of Basel III – a new iteration of the Basel framework in response to the financial crisis of 2007-09. In 2015, the NBU started implementing Basel III recommendations in Ukraine: it performed extensive stress testing and then gradually introduced new requirements to capital, liquidity and other ratios (National Bank of Ukraine, 2015). It also announced its intent to launch several new capital buffers: a CCB, a capital buffer for systemically important banks, a systemic risk buffer and a concentration buffer.

The CCB is an additional capital requirement called to counteract the procyclical behavior of the banking system, which tends to exhibit excessive credit growth during the expansionary phase of business cycles and disproportionate credit contractions during recessions. A gradually increasing CCB requirement during business cycle expansion is expected to discourage banks from issuing too many new loans, and a declining CCB requirement during recessions is expected to stimulate banks to transform their now excessive reserves into new loans, which is exactly what an economy needs to dampen the cycle. Excessive credit growth is only one of many possible reasons for a buildup of systemic imbalances in the banking system, and the CCB buffer's role is to counteract all such buildups, as well as to provide an additional layer of protection against them. The BCBS recommends accumulating an additional cushion of risk-weighted assets in a range of 0% to 2.5% of capital adequacy ratio as a CCB. The key question for policymakers is to identify the right moment to activate ("turn on") and deactivate ("turn off") the CCB requirement so that its effect is indeed stabilizing. In addition, banks should get sufficient advance notice about CCB activation for it not to be overly disruptive to their operations. Early warning models can potentially serve this purpose as they could help to build a system of indicators to signal the accumulation of such systemic risk.

To summarize, the main research questions of this paper are:

1) What bank-specific and macro-related variables can predict bank distress events in Ukraine?

2) What is the explanatory power of such early warning models for individual bank performance and on the aggregate level (in sample)?

3) What is the forecasting power of such early warning models for individual bank performance and on the aggregate level (out of sample)?

To answer these questions, we used an unbalanced panel of quarterly bank-level data for the period Q1 2009-Q3 2019 to estimate a logit model for the probability of an individual bank being in a vulnerable state in the future, using a set of bank-specific and macroeconomic variables. As a robustness check, we did all estimations both with and without "distorting" banks (banks that exited the market due to non-compliance with reporting requirements, rather than capital adequacy issues).

For meaningful comparisons of model predictions with actual data, we then needed to set a value for the threshold parameter (θ), such that if the model-based probability of a bank going into a vulnerable state exceeds θ, the model is said to produce a (positive) signal, which can then be easily compared to the actual state of things (1 for crisis, 0 for no crisis). Following the literature, we set this parameter in such a way that the in-sample "relative usefulness" of the model is maximized. The relative usefulness measure can be used to evaluate both the in-sample and the out-of-sample performance of our model. A set of out-of-sample forecasts of future vulnerabilities was generated by the model on an expanding-window basis.

Our main results were the following.

1. The best model produces a signal (of 1 or 0), whose precision is 62.6% higher than a no signal case (always either 1 or 0), as measured by the relative usefulness indicator. This is comparable to the findings of other authors for other countries. Therefore, we used this model for aggregation, and an algorithm based on mean probabilities of individual bank defaults can indeed predict a banking crisis four periods ahead of the start of a crisis in real-time.

2. The relative usefulness of our preferred model in out-of-sample predictions is quite low, which is consistent with other studies. In particular, the model tends to produce distress signals during tranquil periods, which is consistent with the typical for this literature assumption that a central bank is much more tolerant to false positive bank default signals than to false negative ones.

3. Our analysis suggests that the CCB should be activated when the mean probability of bank default exceeds 11%. However, this figure needs extra validation since the data period –for which model-produced forecasts are available – does not contain any crises.

2. LITERATURE REVIEW

Early warning models became the main tool for the analysis of financial distress in the late 1970s. Martin (1977) was among the first to use a logistic modelling approach in this context. His model included such explanatory variables as financial ratios for asset quality, capital adequacy, and earnings. Barth et al. (1985) complemented this model with liquidity ratios, and Thomson (1992) added management quality as one more predictive factor. Not incidentally, these five variables are constituents of the CAMEL (Capital adequacy, Asset quality, Management, Earning and Liquidity) classification system introduced by banking regulators in the U.S. in 1979 as a tool for evaluating the strength of financial institutions.[1] Numerous other authors (e.g., Sinkey, 1975; Altman, 1977; Arena, 2008; Cole & White, 2012) have also used these variables in their studies, though some additional variables have also been suggested in the literature, e.g., market prices of financial instruments (Flannery, 1998; Bongini et al., 2002) and deposit rates (Kraft & Galac, 2007).

Most of these studies focused on U.S. bank closures, though other countries have received some attention as well. For example, Poghosyan and Čihák (2009), Cipollini and Fiordelisi (2012), and Betz et al. (2013) considered bank defaults in the European Union; Bongini et al. (2001) and Arena (2008) looked at Easter-Asian banks, and González & Hermosillo (1999) and Rosa and Gartner (2018) analyzed Latin America. Among the recent international studies are

[1] In 1996, the "sensitivity to market risk item" was added to the abbreviation currently known as CAMELS.

Altman et al. (2014), who used a sample of banks from 15 European countries and the U.S. during the period of 2007–2012.

In an earlier paper, Pazarbasioglu and Hardy (1998) built a multinominal logit model, which links the likelihood of bank distress events to country-specific and regional peculiarities. They analyzed the banking crises in 38 countries during the period of 1980-1997 and found that bank distress was associated with fall in real GDP growth, high inflation, declining capital-output ratio, and adverse trade shock, as well as decreases in FDI inflows and international reserves, high growth of domestic credit, increases in interest rates rise and an overvalued real exchange rate. Most importantly, the authors suggested that severe banking difficulties were mainly domestic in origin and effect.

Demirgüç-Kunt and Detragiache (1999) also built an EWM as a multinominal logit model using the data on 65 countries during the 1980-1995 period. They analyzed the probability of type I and type II model errors, the unconditional probability of a banking crisis, and the decision maker's preferences parameter for a preventive action in response to the anticipated crisis. They also performed the in-sample and out-sample analysis estimating the predictive power of the model. As a result, two monitoring tools were developed: a particular threshold of the probability indicator and a bank rating system. At the same time, the authors cautioned that aggregated variables convey information about general economic conditions, while the individual bank or specific segment data might point out weaknesses that could lead to contagion, but be invisible in the aggregated data.

Behn et al. (2013) focused on forecasting financial crises based on credit and other macro-financial variables in a sample of 23 EU countries during the period from Q2 1982 to Q3 2012. For validation purposes, they did an out-of-sample prediction of vulnerable states in Finland and Sweden in the early 1990s, and Italy and the U.K. in the mid-1990s preceding the financial crisis in those countries. They found that the loans-to-GDP gap is the best domestic indicator among other credit-related indicators. Moreover, the results showed that more global indicators, i.e. aggregated on a regional level, are outperforming domestic indicators, i.e. aggregated on the local level. However, they also pointed out to the caveat that the evaluation period included the global financial crises, but not the episodes of country-specific crises.

One of the most important papers for our study is by Lang et al. (2018). These authors provided a detailed framework on building an EWM as either an explanatory or a predictive tool. The model that they suggested was aimed at predicting potential future crises at the micro (using aggregation method) and macro level using a large dataset of EU banks. The model exhibited quite satisfactory out-of-sample and in-sample signaling ability with 11 risk drivers and lead time of 1-8 quarters. For evaluation purposes, they used the loss function approach and cross-validation to find a model specification with optimal for the policymaker, real-time, out-of-sample forecasting power. The authors also illustrated how the model's output could assist policymakers by providing EWM visualization.

Our paper contributes to the existing literature by developing an EWM based on the Ukrainian quarterly bank-level data – over the period that includes several domestic crises – and by identifying variables and instruments that will help policymakers understand whether vulnerabilities are accumulating or not.

3. DATA AND METHODOLOGY

3.1. Premodeling

In general, early warning models are used for identifying vulnerable conditions before distress events. As a result, we can view our problem as a two-class identification process, in particular, whether an object is in a vulnerable state or not. According Lang et al. (2018), EWM modeling includes three stages: pre-modeling (purpose, forecast horizon and event indicators), modeling (evaluation criterion, modeling technique, model selection, and evaluation exercise) and post-modeling (policy-relevant dimensions, visualization).

Following Lang et al. (2018), we consider three types of events as bank distress events: bank bankruptcy, default, and NBU refinance.[2] There are 86 such distress events in our sample, with most of them happening in small private banks.[3] Banks that exited the market during the sample period because of the war, occupation of territory, merger or self-liquidation were not included in this count.

As Figure 1 demonstrates, the peak in the number of bank distress events in Ukraine was in Q4 2014. Interestingly, this indicator is highly correlated with a financial stress index (FSI) developed by Filatov (2020) for Ukraine, using 20 indicators containing information about the level of financial stress in the system. As Figure 1 shows, the FSI jumped in Q1 2014, signaling the start of the 2014-2015 crisis. We can also observe a sharp increase in the number of defaulted banks at the same time. Our goal is to produce a model that can signal a potential crisis at least one year in advance to leave banks with a sufficiently long window to accumulate a buffer once a policymaker "turns on" the CCB requirement. The crisis of 2014-2015 will serve as the main testing grounds for our model performance both in sample and out of sample.

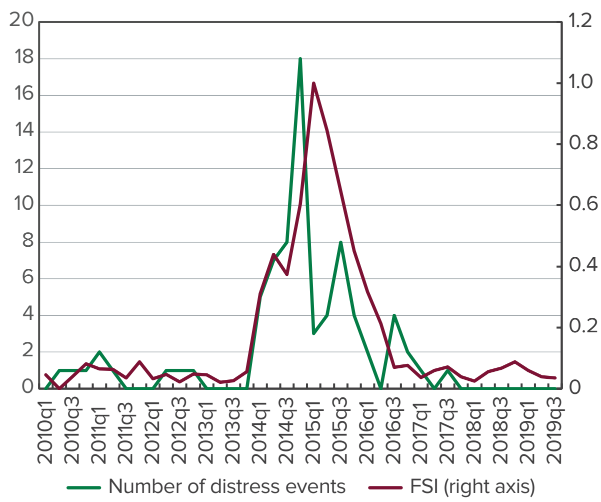

Figure 1. Total Number of Distress Events

[2] During the sample period, the NBU shut down 24 banks due to non-compliance with financial monitoring requirements or an unclear stakeholder structure. These banks are excluded from the main sample as "distorting" banks since these events are not directly related to financial distress events, but are kept in robustness check regressions.

[3] Eight banks had a NBU loan refinance and all of them stayed in the market; 78 banks defaulted or declared bankruptcy and 68 of them left the market.

The goal of an EWM is to generate a signal about a potential distress event in the future, which is typically defined not as one particular quarter N periods from now, but as any time between the next quarter and a quarter N periods from now. The more forward-looking the model is, the more time policymakers and banks have to introduce preventive measures. On the other hand, longer horizons mean lower degrees of freedom for estimations (which is particularly important given our relatively short sample) and, if financial vulnerabilities are building up quite quickly, become redundant after some point. As Lang et al. (2018) point out, there is no consensus about the time horizons in the literature, and this is an empirical question. In particular, we experimented with five different future time horizons (TH) ranging from five to nine quarters.

Following Bussiere and Fratzscher (2006) and numerous other authors, we processed the data in the following way: first, we assigned the value of “1” to each period during the chosen time horizon before the distress event (corresponding to *vulnerable* states), and the value of “0” to all other (tranquil) periods for each bank. Then, the data points containing distress events and four subsequent periods were excluded from the sample due to the noise in the data caused by a crisis. We added up all signals across banks for each period and for different time horizons and received aggregated signals for various horizons (Figure 2). The optimal horizon will be chosen based on model’s relative usefulness criterion explained in the next section.

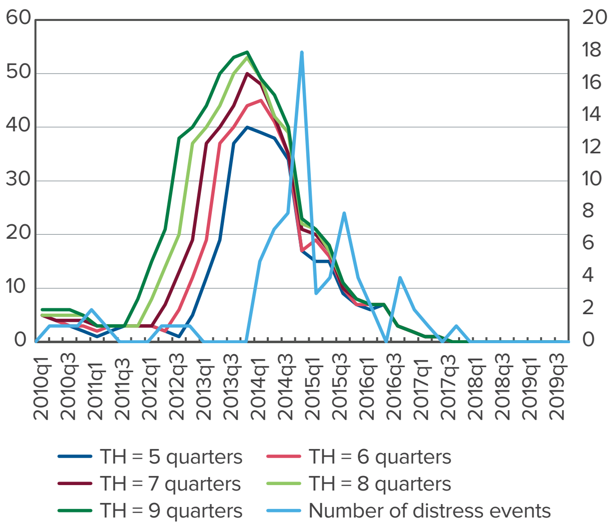

Figure 2. Aggregated Signaling Horizons

3.2. Modeling

The next step is to set up the modeling and evaluation approach. This involves stipulating the evaluation criterion, modeling technique, optimal model complexity, specification, and setting up evaluation procedure.

A bank distress event can be described as a binary variable $I_{i,t}$Î {0,1} which at time *t* signals the state of bank *i*. If $I_{i,t} = 0$, then it is a tranquil period, and if $I_{i,t} = 1$, then the bank is in a vulnerable state and could potentially have a distress event in five to nine periods depending on the selected time horizon. To estimate the probability that a bank will enter a vulnerable state, we suggest using the following logit model[4]:

$$p(I_{i,t} = 1 \mid X_{i,t} = x_{i,t}) = \frac{exp\,(\beta' x_{i,t})}{1 + \exp\,(\beta' x_{i,t})}, \qquad (1)$$

where, p(Yi,t = 1|Xi,t=xi,t) denotes the probability that in period *t* bank *i* is in a vulnerable state. As independent variables, the vector $x_{i,t}$ includes credit, macro-financial, and balance sheet variables (Table 3). Betz et al. (2014) state that the frequency of banking crises corresponds to a fat-tailed error distribution, which makes a logit model more appropriate than a probit model.

The real-economy variables and credit-related variables are used as independent variables by Drehmann et al (2011), Detken, et al. (2014) and Behn, et al. (2013) and other authors. Lang et al. (2018) supplemented this explanatory variable by bank balance sheet data. In particular, these authors start with more than 100 balance-sheet variables and apply a selection operator (a recursive LASSO regression) to reduce the number of variables in the final model.

Our model also includes a wide range of balance sheet variables (according to the CAMEL methodology). In particular, we have 15 variables + four lags for each bank balance sheet variable (36 variables in total). We added lags to control for delays in the financial reporting. We applied the same stepwise selection operator to decide on the set of variables to be included in the final model specification.

Once the model is estimated, we can use the fitted values of $p_{i,t}$ to construct a binary variable $P_{i,t}$ that mimics the behavior of $I_{i,t}$. In particular, when $p_{i,t}$ exceeds a certain threshold $\theta \in [0,1]$, then $P_{i,t} = 1$ or $P_{i,t} = 0$ otherwise. Table 1 describes the relationship between $P_{i,t}$, and $I_{i,t}$ as a contingency matrix, and classifies the outcomes in terms of their signaling quality.

As in many other econometric applications, here we face a trade-off between Type I error (false negative) and Type II error (false positive). Starting from simple univariate early warning models by Drehmann et al. (2011), the area under the receiver operator curve (AUROC) is used as a standard instrument to measure the performance of such models. The receiver operator curve plots A/(A+C), or the model’s sensitivity, against B/(B+D), or the false positive rate for a chosen threshold parameter θ.[5] The area under the curve is the measurement of the model’s performance, as this

Table 1. Contingency Matrix

		Actual class $I_{i,t}$	
		Crisis	No crisis
Predicted class $P_{i,t}$	Signal	A True positive	B False positive
	No signal	C False negative	D True negative

[4] We also suggested conditional (fixed-effect) logistic regression. However, more than half of the sample was dropped due to the specific sample selection as the fixed effects logit model used only within variation while ignoring between variation. Moreover, standard errors became significantly larger due to the smaller sample, resulting in inadequate thresholds since most of the fitted values were concentrated near 0 or 1.

[5] Sensitivity is as same as true positive rate and is equal to A/(A+C) from Table 1. Specificity is the same as true negative rate and is equal to D/(B+D) from the same table. In some literature, sensitivity is called a *signal ratio* and (1-specificity) is referred to as a *noise ratio*.

measurement is independent of its objective: explanatory or predictory (Sarlin 2013).

An AUROC based on the whole sample can sometimes be overoptimistic, and the way to correct for this is to use a K-fold cross-validation AUROC. It is a validation technique for assessing how the estimated model will generalize to an independent data set. Its logic is to split data into k folds, build a model on k-1 folds (training sample), test its classification performance by using AUROC on the last k_{th} fold (test sample) that is independent of the training sample and repeat these steps for each fold. Afterwards, we averaged the AUROCs corresponding to each fold and applied the bootstrap procedure to the cross-validated AUROC to obtain statistical inference and 95% bias corrected confidence intervals (CI).[6]

Betz et al. (2014) and Lang et al. (2018) applied recursive methods for optimizing their models and thresholds. Recursive methods help researchers to compute the floating threshold (a threshold that differs across periods). A recursive approach uses an in-sample period, i.e. data available at the beginning (before t=1,2...T), for training the model and computing the optimal threshold (the one at which the model exhibits maximum usefulness, or, equivalently, minimum loss for the policymaker). The next step is to make predictions during the out-of-sample period, i.e. the next quarter (t), with the in-sample threshold and collect the results. The final step is to recursively re-estimate the model at t=t+1 and repeat all the previous steps till t≤T. As a result, we have more precise estimates of the model's usefulness and thresholds.

The concept of model usefulness, which is a standard performance evaluation criterion in this strand of literature, is closely linked to the concept of the policymaker's (in our case, the central bank's) loss function. Following Sarlin (2013), we assume that it is of the following form:

$$L(\mu) = \mu P_1 T_1 + (1-\mu) P_2 T_2, \qquad (2)$$

where $P_1 = P(I_{i,t} = 1)$ and $P_2 = P(I_{i,t} = 0) = 1-P_1$ are the estimated frequencies of the classes (unconditional probabilities: $P_1 = (A+C)/(A+B+C+D)$ and $P_2 = (B+D)/(A+B+C+D)$, $T_1 = C/(A+C)$ is the false positive rate and $T_2 = B/(B+D)$ is the false negative rate and μ is a preferences parameter.[7] Both T_1 and T_2 are functions of the threshold parameter θ: a higher threshold reduces the false positive rate T_1 and at the same time increases the false negative rate T_2, and vice versa. The optimal value of the threshold parameter for each μ is the one that minimizes the policymaker's loss function.

Policymakers could get a loss μP_1 when the model never signals a crisis or $(1-\mu)P_2$ when the model always issues a signal. Therefore, the loss is equal to $\min[\mu P, (1-\mu)P_2]$ if a policymaker does not apply the early warning model (the CCB is always either on or off). We can then compute the absolute usefulness of the model, Ua, by subtracting the loss associated with using the model from the loss incurred from ignoring it:

$$U_a(\mu) = \min[\mu P_1, (1-\mu)P_2] - L(\mu), \qquad (3)$$

Along the same lines, the *relative* usefulness of the model, U_r, is the ratio of this "model-induced loss recovery" to the baseline "no-model" loss:

$$U_r(\mu) = \frac{U_a(\mu)}{\min[\mu P_1, (1-\mu)P_2]} = 1 - \frac{L(\mu)}{\min[\mu P_1, (1-\mu)P_2]}, \qquad (4)$$

Notice that if $L(\mu) = 0$, then $U_a(\mu) = \min[\mu P_1, (1-\mu)P_2]$ and $U_r(\mu) = 1$ meaning that the model is working perfectly. The relative usefulness criterion is our main model performance criterion, in particular for identifying the optimal threshold (θ) and the optimal time horizon (TH).

3.3. Postmodeling

Once the model is estimated, and all evaluation exercises are performed, it is important to decide how the model output could be analyzed and visualized. Taking into account that we have bank-level data, it is important to analyze the aggregate effect on the financial system. For this purpose, we experimented with two alternative aggregation approaches: we took either the mean or the median estimated default probabilities among all banks in each period. Then, we compared the aggregated results with FSI to identify which approach gives more precise results within the in-sample analysis.

Next, we conducted an out-of-sample analysis using the following algorithm. First, our preferred model is estimated based on shorter, ever-expanding sub-samples of the data. Then, the estimated model coefficients are used to forecast the probabilities of individual bank defaults over the future time horizon, along with the optimal threshold parameter. Finally, these projected individual bank default probabilities are aggregated using the preferred aggregation approach to produce a single system-wide signal on accumulated vulnerabilities.

3.4. Data Description

Our data set contains information on 209 banks in the period from Q1 2009 to Q3 2019 (5,632 observations in total). We identified 86 relevant distress events in our sample.[8] The data was collected from the websites of State Statistics Service of Ukraine (SSSU) and the NBU. It is unbalanced panel bank-level quarterly data (Table 3). Some banks had reporting gaps during the sample period. Such banks were eliminated from the sample. A detailed data description could be found in Table 3. We did not exclude outliers as we believe they could contain important information on financial system vulnerabilities.

4. ESTIMATION RESULTS

Below we present the estimation results of models with different time horizons based on the entire sample. As described in the methodology section, we focus on the relative usefulness statistics (U_r) as the main indicator of the model's explanatory power. Therefore, an early warning model with the optimal time horizon (TH) will have the highest relative usefulness (U_r). Table 2 reports the results of the estimated absolute and relative usefulness of our models for varying time horizons (TH) and different values of the preference paramerter μ. The highest relative usefulness is for the models with six- and seven-quarter horizons and the preference parameter μ=0.9. These results are quite

[6] Following Lang and Peltonen (2018), we use ten folds, which is also a standard in Stata.
[7] Following the literature, we experimented with several values of the preferences parameter μ between 0.6 and 0.9, implying that it is relatively more costly for the central bank to miss a crisis than to issue a false signal.
[8] According to the NBU website, the number of banks in Ukraine dropped from 175 in 2008 to 77 in 2019.

comparable to the findings of other authors, suggesting that this modeling approach is as valid for Ukrainian data as it is for other countries.

Though there is some variability in the results of the models with different time horizons, no rigorous tests are available to evaluate whether these differences are statistically significant, and the top model must be chosen on a mostly ad hoc basis. For the lack of more formal arguments, we decided to use the middle-horizon (seven quarter) model as the baseline since it has higher relative usefulness than the short-horizon (five quarter) model and either slightly underperforms or noticeably outperforms the long-horizon (nine quarter) one. Fortunately, the conclusions are robust to the choice of time horizon, including those concerning the optimal threshold parameter θ, as Table 2 clearly demonstrates.

expenses per liability increases this probability. As a result, the model shows the balance sheet's effective management is a key indicator of bank solvency.

Return on assets (ROA) is robust for the first and second lags, but the sign is positive, meaning that better return on assets increases the probability of the signal being issued. The most likely explanation for this is that higher returns on assets are also associated with their higher riskiness, which in turn makes banks more vulnerable.

The higher ratio of provisions to total assets reduces the probability of a signal being issued: more generous provisions against expected losses add to bank stability. The cumulative (over all lags) effect of the ratio of total equity to total assets is consistently negative for all model specifications, which is also in line with what the theory suggests.

Table 2. In-Sample Analysis of the Models Based on Five to Nine Quarter Time Horizons

	5 quarter	6 quarter	7 quarter	8 quarter	9 quarter
μ	Absolute usefulness				
0.6	0.011	0.016	0.021	0.024	0.027
0.7	0.018	0.025	0.030	0.034	0.039
0.8	0.028	0.036	0.043	0.049	0.055
0.9	0.041	0.051	0.056	0.052	0.048
μ	Relative usefulness				
0.6	0.243	0.296	0.328	0.330	0.335
0.7	0.331	0.390	0.405	0.404	0.422
0.8	0.455	0.494	0.510	0.520	0.518
0.9	0.605	0.626	0.626	0.590	0.555
μ	Threshold				
0.6	0.422	0.397	0.401	0.432	0.410
0.7	0.312	0.340	0.330	0.352	0.357
0.8	0.242	0.225	0.252	0.246	0.248
0.9	0.132	0.123	0.107	0.131	0.137
	AUROC				
cvMean AUC:	0.876	0.881	0.881	0.881	0.872
Bootstrap bias corrected 95%	0.850	0.861	0.852	0.857	0.855
Confidence intervals	0.895	0.901	0.893	0.895	0.890

By analyzing the results of the benchmark model, we can distinguish what bank-specific and macro-related variables can predict bank distress events in Ukraine (Table 4). We consider only those variables that are significant for the best model and compare them with other models for robustness check. If most of the models have significant results (with same signs), then we consider such results as robust. Therefore, we describe below only robust results.

Net interest income/total assets, net commission/total assets, and interest expenses/total liabilities indicate the efficiency of the assets and liabilities respectively as these ratios show net commission and interest income per asset and expenses per liability.[9] The rising income per asset reduces the probability of issuing a signal and rising

Among the macroeconomic variables, the consistently significant ones for all time horizons are the ratio of international reserves to GDP, the house price index and the ratio of state budget surplus to GDP. All these variables have expected signs and are very robust in terms of magnitudes across all specifications. Interestingly, the real GDP growth variable is insignificant in the preferred (mid-horizon) specification, but is negative and significant for the five-quarter horizon, and positive and significant for the nine-quarter horizon. This might suggest that this variable contains some interesting cyclical features, and further analysis is required to properly map them into our variables of interest.

The next step is to go from bank-level results to system-level outcomes. To aggregate the estimated probabilities

[9] There is a negative sign, as the original data expenses also have a negative sign, meaning that rising expenses will increase the probability of signaling.

of individual bank vulnerabilities into a single figure that reflects the state of the banking system as a whole, we experimented with two aggregation strategies: by taking either the mean or the median default probability across all banks for each period. From Figure 3, we can observe that before the 2014-2015 crisis, the mean default index gave an earlier and more pronounced signal about the accumulated vulnerabilities than the median indicator. However, after the peak of Q2 2014, aggregation by mean tends to drag on in terms of overly strong signals and becomes quite volatile. Still, since the mean approach produces an earlier signal than the median one, we chose it as our preferred approach for further analysis.

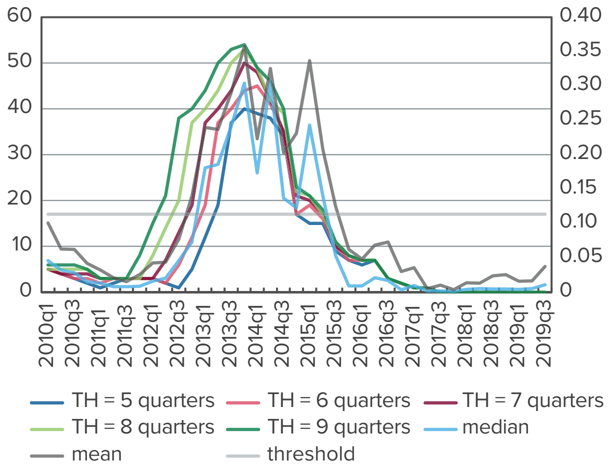

Figure 3. Aggregated Probabilities of Bank Distress Events (in-sample analysis)
Note: Time horizons are aggregated signaling horizons from Figure 2.

Figure 4 depicts the mean-based aggregated probabilities of bank defaults and the FSI index for clearer comparisons. Here we can see that the mean algorithm gave a signal in Q4 2012, and the crisis started to unravel in Q1 2014, according to the FSI. That means the (seven-quarter horizon) model is issuing an accurate positive signal on the whole banking system crunch five periods ahead of the critical time. This result is very encouraging since a policymaker has five quarters to implement CCB before the start of a crisis, which should potentially help with reducing the negative costs associated with it or even preventing it entirely.

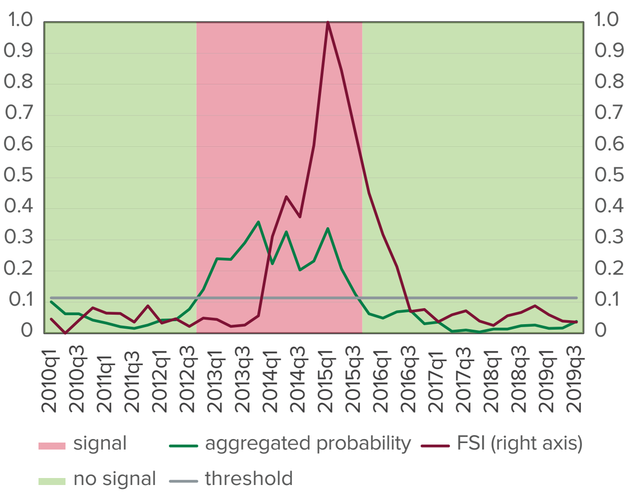

Figure 4. Comparison Combined Aggregated Results with FSI

Next, we do recursive model estimations based on shorter, ever-expanding sub-samples. For each of them, we estimate optimal thresholds (for the preference parameter μ=0.9, which corresponds to the most apprehensive and vigilant central bank among all considered alternatives), generate predictions for each individual bank and then calculate a mean-based measure of aggregate probability of crisis. Quite expectedly, the out-of-sample relative usefulness of the model falls and now is only 37%. Still, using the model is much better than not using any model at all.

Figure 5 summarizes other important out-of-sample results. Due to data restrictions, the first available prediction is for Q1 2013. The threshold parameter varies between 3% and 22%, and more or less stabilizes around 12% once the sub-sample length reaches 20 quarters. Most importantly, starting from Q1 2013, the model produces a positive out-of-sample signal that gives policymakers four quarters before the crisis starts in Q1 2014. At the same time, we observe that the out-of-sample aggregated signal is relatively unstable: it produces a false signal in Q3 2016. However, given that the thresholds are also estimated with errors, this signaling mistake could actually be quite within the threshold confidence bounds.

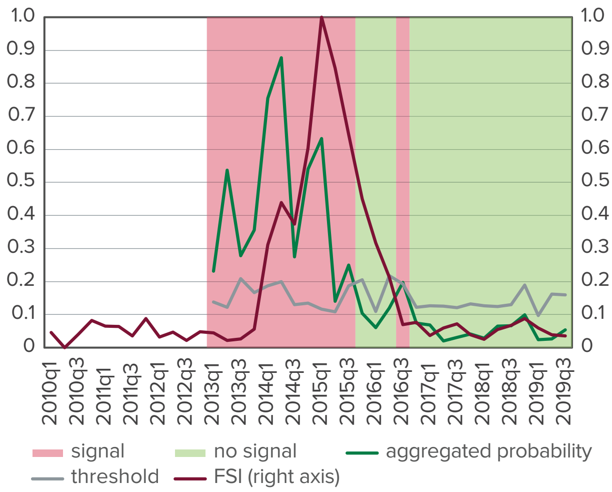

Figure 5. Recursive Estimation (out-of-sample analysis)

5. CONCLUSIONS

The main goal of setting a CCB is to protect the banking sector from the excessive aggregate credit growth associated with broad systemic risk. For this instrument to be effective, its timing (the ”on” and “off” switch) must be as precise as possible. Numerous economists analyzed a wide range of indicators and thresholds that signaled when to activate CCB and concluded that the credit-to-GDP gap is one of the most accurate indicators for many countries. However, this was not the case for East European countries, including Ukraine, as they have structural changes and a relatively short observation period.

An alternative approach is to use bank-level data to identify individual bank vulnerabilities and then aggregate them into a system-wide risk measure. We use an early-warning bank default model as a building block for this approach. The model contains both bank-specific and macro-level variables. To be consistent with the literature (e.g., Sarlin, 2013), we use a relative usefulness measure to evaluate its empirical performance both in sample and out of sample. We can conclude that the model based on a seven-quarter signaling period (benchmark model) and $\mu = 0.9$ is the best one according to its relative usefulness (62.5%, which indicates that the explanatory power of the model is quite high). This model gives a signal when the probability of default exceeds the (optimal) threshold of 11%.

To aggregate the individual bank data, we use the mean-based approach, in which the mean of the estimated individual-bank probabilities of default is the statistic to be compared to the threshold value. This approach produces a positive signal about the 2014-2015 crisis (and therefore the need to launch the CCB accumulation) five periods ahead of its start with in-sample estimations and 4 period ahead with out-of-sample estimations. We conclude that the model can potentially be used as a predictive instrument to help policymakers identify a build-up of system vulnerabilities and substantiate the need for policy intervention.

REFERENCES

Altman, E. I. (1977). Predicting performance in the savings and loan association industry. Journal of Monetary Economics, 3(4), 443-466. https://doi.org/10.1016/0304-3932(77)90015-0

Altman, E. I., Cizel, J., Rijken, H. A. (2014). Anatomy of bank distress: The information content of accounting fundamentals within and across countries. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=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 & Finance 32(2), 299-310. https://doi.org/10.1016/j.jbankfin.2007.03.011

Barth, J. R., Brumbaugh, R. D., Sauerhaft, D., Wang, G. H. (1985). Thrift institution failures: Causes and policy issues. Proceedings of the 21st Annual Conference on Bank Structure and Competition, 184-216. Chicago: Federal Reserve Bank of Chicago. Retrieved from http://webhome.auburn.edu/~barthjr/publications/Thrift%20Institution%20Failures%20Causes%20and%20Policy%20Issues.pdf

Basel Committee on Banking Supervision. (2010). Guidance for National Authorities Operating the Countercyclical Capital Buffer. Bank for International Settlements. Retrieved from https://www.bis.org/publ/bcbs187.pdf

Behn, M., Detken., Peltonen, T. A., Schudel, W. (2013). Setting countercyclical capital buffers based on early warning models. Would it work? ECB Working Paper, 1604. Frankfurt am Main: Europen Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1604.pdf

Betz, F., Oprica, S., Peltonen, T. A., Sarlin, P. (2014). Predicting distress in European banks. Journal of Banking & Finance, 45, 225-241. https://doi.org/10.1016/j.jbankfin.2013.11.041

Bongini, P., Laeven, L., Majnoni, G. (2002). How good is the market at assessing bank fragility? A horse race between different indicators. Journal of Banking & Finance, 26(5), 1011 – 1028. https://doi.org/10.1016/S0378-4266(01)00264-3

Bussiere, M., Fratzscher, M. (2006). Towards a new early warning system of financial crises. Journal of International Journal of International Money and Finance, 25(6), 953-973. https://doi.org/10.1016/j.jimonfin.2006.07.007

Cipollini, A., Fiordelisi, F. (2012). Economic value, competition and financial distress in the European banking system. Journal of Banking & Finance, 36(11), 3101-3109. https://doi.org/10.1016/j.jbankfin.2012.07.014

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, 42(1-2), 5-29. https://doi.org/10.1007/s10693-011-0116-9

Demirgüç-Kunt, A., Detragiache, E. (1999). Monitoring banking sector fragility: A multivariate logit approach. IMF Working Paper, 99/147. International Monetary Fund. Retrieved from https://www.imf.org/external/pubs/ft/wp/1999/wp99147.pdf

Detken, C., Weeken, O., Alessi, L., Bonfim, D., Boucinha, M. M., ... Welz, P. (2014). Operationalising the countercyclical capital buffer: indicator selection, threshold identification and calibration options. Occasional Paper Series, 5. European Systemic Risk Board. Retrieved from https://www.esrb.europa.eu//pub/pdf/occasional/20140630_occasional_paper_5.pdf

Drehmann, M., Borio, C., Kostas, T. (2011). Anchoring countercyclical capital buffers: The role of credit aggregates. International Journal of Central Banking, 7(4), 189-240. Retrieved from https://www.ijcb.org/journal/ijcb11q4a8.pdf

Filatov, V. (2020). A new financial stress index for Ukraine. Working Papers, HEIDWP15-2020. Switzerland: Graduate Institute of International and Development Studies. Retrieved from http://repec.graduateinstitute.ch/pdfs/Working_papers/HEIDWP15-2020.pdf

Flannery, M. J. (1998). Using market information in prudential bank supervision: A review of the U.S. empirical evidence. Journal of Money, Credit and Banking, 30(3), 273-305. https://doi.org/10.2307/2601102

Kraft, E., Galac, T. (2007). Deposit interest rates, asset risk and bank failure in Croatia. Journal of Financial Stability, 2(4), 312-336. https://doi.org/10.1016/j.jfs.2006.08.001

González-Hermosillo, B. (1999). Determinants of ex-ante banking system distress: A macro-micro empirical exploration of some recent episodes. IMF Working Paper, 99/33. Retrieved from https://www.imf.org/en/Publications/WP/Issues/2016/12/30/Determinants-of-Ex-Ante-Banking-System-Distress-A-Macro-Micro-Empirical-Exploration-of-Some-2908

Lang, J. H., Peltonen, P., Sarlin, T. A. (2018). A framework for early-warning modeling with an application to banks. ECB Working Paper, 2182. Frankfurt am Main: European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecb.wp2182.en.pdf

Martin, D. (1977). Early warning of bank failure: A logit regression approach. Journal of Banking & Finance, 1(3), 249-276. https://doi.org/10.1016/0378-4266(77)90022-X

Pazarbasioglu, C., Hardy, D. C. (1998). Leading indicators of banking crises: Was Asia different? IMF Working Paper, 98/91. International Monetary Fund. Retrieved from https://www.imf.org/external/pubs/ft/wp/wp9891.pdf

Poghosyan, T., Cihak, M. (2009). Distress in European banks: An analysis based on a new data set. IMF Working Paper, 09/9. International Monetary Fund. Retrieved from https://www.imf.org/en/Publications/WP/Issues/2016/12/31/Distress-in-European-Banks-An-Analysis-Basedon-a-New-Dataset-22547

Rosa, P. S., Gartner, I. R. (2018). Financial distress in Brazilian banks: an early warning model. Revista Contabilidade & Finanças, 29(77), 312-331. https://doi.org/10.1590/1808-057x201803910

Sarlin, P. (2013). On policymakers' loss functions and the evaluation of early warning systems. ECB Working Paper, 1509. Frankfurt am Main: European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1509.pdf

Sinkey, J. F. Jr. (1975). A multivariate statistical analysis of the characteristics of problem banks. Journal of Finance 30(1), 21-36. https://doi.org/10.1111/j.1540-6261.1975.tb03158.x

Thomson, J. B. (1992). Modeling the bank regulator's closure option: A two-step logit regression approach. Journal of Financial Services Research, 6, 5-23. https://doi.org/10.1007/BF01046114

APPENDIX A. TABLES

Table 3. Descriptive Statistics

	Mean	Standard deviation	Max	Min
Bank balance sheet variables				
Net Interest income per asset	0.034	0.046	2.051	-0.106
Net commission income per asset	0.012	0.023	0.415	-0.470
Net interest expenses per liability	-0.055	0.288	0.018	-21.159
Provisions/total assets	-0.101	0.393	0.649	-22.552
Total equity/total assets	0.235	0.454	1.000	-27.763
Common equity/total assets	0.261	0.296	6.837	0.000
ROE	0.177	9.708	537.000	-47.824
ROA	-0.021	0.465	0.549	-31.175
Credit-related variables				
State budget surplus/deficit to GDP	-0.034	0.015	0.003	-0.062
Money supply growth	0.024	0.037	0.080	-0.100
Central government debt service to GDP	0.023	0.011	0.046	0.004
Real-economy variables				
Real GDP growth	-0.017	0.083	0.067	-0.249
Current account growth as % of GDP	0.052	0.039	0.206	0.001
House price index	51.892	13.054	72.547	31.493
REER	0.892	0.104	1.014	0.660
Reserves as % of GDP	0.684	0.202	1.174	0.394
Observations	5,632			

Table 4. All Results of the Models

	TH=5 quarters	TH=6 quarters	TH=7 quarters	TH=8 quarters	TH=9 quarters
Net commission income per asset		-14.4000* (7.262)	-14.1300* (6.758)	-13.0000* (6.351)	-14.9700* (5.894)
Net commission income per asset(-1)	-12.8900 (7.339)	-14.2300 (8.015)	-15.4600* (7.132)	-17.2800* (7.050)	21.5700*** (5.725)
Net commission income per asset(-2)	-13.9300* (6.281)	-7.7940 (5.158)	-8.1320 (5.007)	-7.9300 (5.255)	
Net commission income per asset(-4)	-9.5140 (5.214)				
Net Interest income per asset	-21.3000*** (5.279)	-21.1500*** (5.188)	-25.7800*** (4.800)	-27.9900*** (5.431)	-30.1600*** (5.300)
Net Interest income per asset(-1)	-16.0200* (6.519)	-23.2400*** (6.160)	-18.4400*** (5.476)	-21.0400*** (5.202)	-23.5600*** (4.363)
Net Interest income per asset(-2)	-13.6900** (5.165)	-6.3380 (4.538)	-6.3120 (4.283)	-5.4620 (4.058)	
Net Interest income per asset(-4)				7.2360 (3.888)	8.9800* (3.756)
Corporate deposits growth					0.0122 (0.009)
Corporate deposits growth(-1)				0.0121 (0.008)	0.0117 (0.008)
Net interest expenses per liability	-2.0740 (1.147)	-2.1780 (1.141)	-2.9980* (1.223)	-3.4680** (1.270)	-4.0510** (1.431)
Net interest expenses per liability(-1)	-22.7600*** (2.933)	-23.9700*** (2.819)	-24.1100*** (2.723)	-24.0900*** (2.650)	-24.6900*** (2.278)
Net interest expenses per liability(-2)	-14.4200*** (3.058)	-13.6800*** (2.884)	-12.9200*** (2.731)	-11.8100*** (2.621)	-8.4370*** (1.872)
Net interest expenses per liability(-3)	-6.0300*** (1.363)	-6.3320*** (1.506)	-4.1770** (1.447)	-3.0580* (1.283)	-3.1630** (1.183)
Provisions/total assets					2.0700* (0.883)
Provisions/total assets(-1)	-2.8920* (1.322)	1.8390 (1.013)	-2.3170* (0.952)	-1.9150* (0.938)	
Provisions/total assets(-2)	-2.5340 (1.573)				
Provisions/total assets(-4)	-1.5880 (1.228)	-2.9470** (0.956)	-3.3200*** (0.874)	-2.5510** (0.879)	-2.2570** (0.803)
Total equity/total assets	2.7890 (1.945)	2.8770 (1.854)			
Total equity/total assets(-1)	-6.2800* (2.663)	-6.4910* (2.541)	-4.7740** (1.805)	-5.2690** (1.768)	-5.1430*** (1.392)
Total equity/total assets(-2)	-5.0380* (2.059)	-5.4140* (2.562)	-6.7210** (2.158)	-3.0700 (1.778)	
Total equity/total assets(-3)		-4.1040 (2.627)			-2.8480* (1.355)
Total equity/total assets(-4)		4.6130** (1.717)	3.1120* (1.249)		

Table 4 (continued). All Results of the Models

	TH=5 quarters	TH=6 quarters	TH=7 quarters	TH=8 quarters	TH=9 quarters
ROE		0.0424	0.0460	0.0503	0.0591*
		(0.032)	(0.027)	(0.027)	(0.023)
ROE(-2)	-0.0567				
	(0.043)				
ROE(-4)					-0.0910*
					(0.042)
ROA	5.8510	5.1200	5.9420*	6.5840*	5.5820*
	(3.626)	(3.202)	(2.948)	(2.900)	(2.346)
ROA(-1)	11.3100**	11.6500**	11.2700***	10.5300***	9.9440***
	(3.816)	(3.738)	(3.257)	(3.134)	(2.641)
ROA(-2)	4.6730**	4.1420**	4.3850**	3.4500*	2.5830*
	(1.527)	(1.516)	(1.439)	(1.491)	(1.300)
ROA(-3)	2.2630*	2.5050*			
	(0.957)	(1.032)			
ROA(-4)			1.9940	2.6010	3.6610*
			(1.215)	(1.328)	(1.586)
Real GDP growth	6.0800***	2.8440		-2.1480	-4.0690**
	(1.834)	(1.762)		(1.568)	(1.546)
Money supply growth M3	-7.1090**		-3.5060	-4.0410	
	(2.752)	(2.604)	(2.342)		
Reserves as % of GDP	-7.0010***	-7.1400***	-6.9070***	-6.6130***	-6.5150***
	(0.654)	(0.620)	(0.536)	(0.542)	(0.523)
House price index	0.0589**	0.0474**	0.0562***	0.0740***	0.0624**
	(0.019)	(0.017)	(0.016)	(0.015)	(0.021)
Central government debt service to GDP					-36.4700
					(19.036)
State budget surplus to GDP	-35.2300***	-35.6300***	-30.7500***	-15.4100*	
	(8.281)	(7.660)	(6.883)	(5.984)	
REER	-6.6740***	-3.7270*	-2.4140	-2.7100	-2.2590
Constant	3.8060***	1.9460*	0.7230	0.6260	2.3980
Observations	4,107	4,107	4,107	4,107	4,107
Pseudo R-squared	0.361	0.368	0.369	0.362	0.355
AIC	1,519.692	1,690.021	1,852.620	2,035.838	2,198.058
BIC	1,702.985	1,873.314	2,023.273	2,212.810	2,375.030

p-values $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$
Standard errors are in parentheses. TH stands for time horizon.