National Bank of Ukraine

ISSN 2414-987X

VISNYK

OF THE NATIONAL BANK OF UKRAINE

No. 253
2022

VISNYK
OF THE NATIONAL BANK OF UKRAINE

No. 253
2022

https://journal.bank.gov.ua

Founder and publisher: National Bank of Ukraine

Editorial Board ensures compliance with ethical standards of publishing

For reprinting materials published herein the reference to the journal “Visnyk of the National Bank of Ukraine” is compulsory

The Editorial Board can publish materials being under discussion, not sharing the author’s opinion

Address: 9, Instytutska Street, Kyiv, 01601, Ukraine
E-mail: journal@bank.gov.ua

PREFACE BY THE EDITOR-IN-CHIEF

Dear Friends,

The full-scale aggression of russia against Ukraine has brought death and destruction and damaged the foundations of Ukraine's economy. Thanks to the brave Ukrainian Armed Forces, together with the resilience of the Ukrainian people and the support of our international partners, we have been able to resist russian aggression and continue the battle for our freedom, sovereignty, and democratic values. Due to earlier reforms, the Ukrainian financial sector has shown resilience in the face of war, while the National Bank of Ukraine, with its proactive policy and Power Banking initiative, has stood firm in defending the country's macrofinancial stability. The research contributors of our journal continue working hard to address challenging research questions that have become especially relevant in wartime.

This issue of the journal is devoted to questions that are relevant for central banks facing uncertain and turbulent conditions. The special focus of this edition is the role of information in the formation of inflation expectations, which, among other things, may be of particular importance for central bank communications policies. In addition, we also focus on early-warning monitoring techniques that allow central banks to recognize buildups of vulnerabilities and signals of upcoming crises, and to assess key risks to financial stability.

The issue starts with research by Tetiana Yukhymenko – *The Role of the Media in the Inflation Expectation Formation Process.* The author applies machine-learning techniques to a large array of news items and constructs news-based metrics to demonstrate the role of the media in forming inflation expectations in Ukraine. Inflation expectations are shown to be sensitive to news topics – a finding that is of great utility to central bank communications when anchoring inflation expectations and meeting inflation targeting goals in a post-war economy.

In the second paper, *A Heatmap for Monitoring Systemic Financial Stability Risks in Ukraine*, a team of authors, consisting of Adam Geršl, Pervin Dadashova, Yuliya Bazhenova, Vladyslav Filatov, Anatolii Hlazunov, and Roman Soltysiak, introduce a revised version of the NBU risk map for the Ukrainian financial sector. The suggested analytical tool assesses the financial system's resilience across key risks and captures a wide range of economic and financial vulnerabilities. The risk map is designed to identify and monitor the buildup and materialization of systemic risks. The authors test the validity of the developed methodology and demonstrate that the improved instrument clearly identifies signals of the buildup of vulnerabilities and crisis episodes in the past.

The findings presented in this issue of our journal will be of value to researchers and policy makers in wartime and during post-war recovery. Other research contributors are also welcome to submit their original fundamental and applied studies for publication in the Visnyk of the National Bank of Ukraine.

Do research, and support the Army of Ukraine in achieving our upcoming victory!

Glory to Ukraine!
Mihnea Constantinescu

THE ROLE OF THE MEDIA IN THE INFLATION EXPECTATION FORMATION PROCESS

TETIANA YUKHYMENKO[a]

[a]National Bank of Ukraine
E-mail: tetiana.yukhymenko@bank.gov.ua

Abstract This research highlights the role played by the media in the formation of inflation expectations among various respondents in Ukraine. Using a large news corpus and machine-learning techniques, I have constructed news-based metrics that produce quantitative indicators for texts, which show if the news topics are relevant to inflation expectations. I have found evidence that various news topics may have an impact on inflation expectations, and can explain part of their variance. Thus, my results could help in the analysis of inflation expectations – which is of value, given that anchoring inflation expectations remains a key challenge for central banks.

JEL Codes C55, C82, D84, E31, E58

Keywords inflation expectations, natural language processing, textual data, machine learning

1. MOTIVATION, THEORETICAL FRAMEWORK, LITERATURE OVERVIEW

Anchoring inflation expectations remains a key challenge for central banks, especially in developing economies. The process of forming inflation expectations it relevant to understanding macroeconomic dynamics and for designing optimal policies. A lot of research has been done in the area of inflation expectations, but there is still a great deal of uncertainty and inconsistency about the factors that determine them. The development of modern information technologies enables us to use new approaches to examine the processes that form inflation expectations. In particular, natural language processing and machine-learning tools can provide additional information that was previously inaccessible. They also makes it possible to supplement with new insights the results of existing studies that have become benchmarks in the industry. Many researchers are now turning to more modern data sources and analysis methods, but the field of research remains largely uncharted. In particular, there is still much uncertainty over how to transform unstructured data into economic indicators, how to take into account the tone of indicators, and how to assess their impact on inflation expectations. In addition, such studies have not yet been conducted on Ukrainian data. Thus, the prospect of being able to apply the latest technologies to already conventional approaches was the main motivation for researching the role of the media in shaping inflation expectations in Ukraine.

The rational expectations hypothesis has dominated the macroeconomic literature for many years. However, in a growing body of research, this hypothesis is being modified to account for information rigidities – expectations could be rational, but in a more realistic environment agents may be inattentive to relevant information due to the costs of acquiring and processing such information. The two leading models of information rigidity are the sticky information model of Mankiw et al. (2004) and the noisy information model developed by Woodford (2004) and Sims (2009). Mackowiak and Wiederholt (2009) also did work in this field. Coibion and Gorodnichenko (2012) proved that information rigidities have a large impact on macroeconomic variables, thus they should be integrated into modern macroeconomic policies in order to execute the optimal monetary policy. They also found that despite common wisdom, there is no significant difference in the degree of information consumption across agents – the speed of information processing by consumers is no lower than that by other agents. Among other things, this can be explained by the noisy information model. Similarly, Coibion and Gorodnichenko (2015a) find that the inflation expectations of professional forecasters from the U.S. Survey can be modeled with imperfect information models due to the existence of information frictions. Coibion and Gorodnichenko (2015b) also research economic agents' expectations in Ukraine, based on survey data on inflation and exchange rate expectations. The survey also shows that there is a strong positive correlation between the evolution of Ukrainian economic agents' expectations about inflation, and exchange rates. While some correlation might be expected from the pass-through of exchange rates into prices, a more likely rationale is that the exchange rate is being used as a straightforward proxy by households of

Available at https://doi.org/10.26531/vnbu2022.253.01

broader price movements within the economy, very much like households within the U.S. do with gasoline prices.

It can be assumed that survey respondents are also influenced by uncertainties regarding tax, tariff, spending, monetary and regulatory policy. These effects, however, are hard to detect because uncertainty is unobservable. However, people may obtain their views about the future path of the economy from the news media, directly or indirectly. So, news-based methods could be used to investigate the impact of the media environment on the formation of respondents' expectations.

For example, Carroll (2003) tested an epidemiological model of expectations in which information diffuses over time from professional forecasters to households. Pfajfar and Santoro (2013) complement this model with a measure of the actual perception of new information about prices. As a news metric, they used a question from a survey where participants have to indicate whether they have heard about positive or negative changes. Hearing news related to prices increases the probability of an adjustment in inflation expectations, while the quality of forecasts is not likely to improve. Similarly, Coibion et al (2019) researched how central bank communications impact expectations. Thus, they compare the answers of respondents after receiving eight different forms of information regarding inflation. They concluded that these messages to the public influence expectations by economically significant magnitudes. However, their effectiveness significantly decreases when channeled via news media. Mazumder (2021) proved that newspaper mentions of the Fed bring consumer and professional inflation forecasts closer, although this effect can vary depending on which newspaper it was published in and how the topic was covered by the author. Dräger and Lamla (2017) also found evidence of the impact of the media on the formation of inflation expectations. They analyzed the rotating panel dimension of the microdata in the University of Michigan Survey of Consumers, and found evidence that respondents are more likely to adjust their expectations if they have heard news about inflation.

However, most of these studies imply the use of supplementary questions in the survey, which can be costly. In addition, even if such questions are introduced, the results will not cover previous periods. Thus, measuring the impact of news and constructing relevant indexes requires novel sources of information and processing methods, as well as significant computational resources. Consequently, researchers are replacing these indexes with alternative indicators which could be related to news metrics. For example, Bauer (2015) used macroeconomic data surprises cumulated over the monthly or quarterly observation windows as an economic news metric. Thus, the data are macroeconomic indicators collected from traditional statistical sources, but their interpretation is somewhat different from the usual time series. Bauer found that several different survey measures of inflation expectations respond significantly to macroeconomic surprises. He also concluded that better anchoring of long-term inflation expectations can reduce the sensitivity of inflation expectations to macroeconomic news, and the variability of nominal rates as well. Garcia and Werner (2018) confirmed that early inflation releases had a significant impact on long-term inflation expectations, and that there was a weakening of the anchoring of inflation expectations in the EU in recent years. Nautz et al (2017) also found that euro area inflation expectation anchoring was undermined after the fall of 2011. They discovered that long-term inflation expectations respond significantly to macroeconomic news. As a news metric, a set of macroeconomic variables was used, including CPI, PPI, unemployment, GDP, trade balance, etc. D'Acunto et al. (2017) additionally found a relationship between the frequency and size of price changes.

Larsen et al. (2021) used a more sophisticated approach. They applied machine-learning algorithms to a large news corpus and examined the role of the media in the expectation formation process of households. It turned out that the news topics in the media are a good predictor of both inflation and inflation expectations. They also found that the degree of information rigidity among households varies across time, which can be explained by the relevant media coverage. Angelico et al. (2021) used a similar approach to build real-time measures of consumers' inflation expectations from tweets. They combined unsupervised machine-learning techniques with a dictionary-based approach to construct indices. Twitter-based indicators appear to be highly correlated with traditional measures of inflation expectations while having an advantage in their speed.

In this work, I focus on the analysis of news and its impact on the formation of inflation expectations. To this end, I explore approaches to transforming texts into quantitative indicators which can then be further used in traditional econometric analyses. These indicators should reflect news topics relevant to inflation expectations and accurately capture their intensity. These metrics should also be easily interpretable, as they aim to explain the impact of news on the formation of inflation expectations. All these tasks can be accomplished with text-mining techniques.

All measurement methods that are based on text-mining can be divided into two groups: 1) so-called naïve methods and 2) more complex methods, which are based on machine learning. Naïve methods are parsimonious, easy to use (as they do not require much computational power), and are recognized worldwide due to their simplicity. They are based mostly on term frequency and document frequency. For example, Baker et al. (2016) investigated the relationship between economic policy uncertainty and rates of investment, output, and employment growth. For these purposes, the authors developed an index of economic policy uncertainty (EPU) based on a monthly count of articles that contain specific terms. Their findings demonstrate that the EPU is a reasonable proxy for various types of important macroeconomic variables, and its results are consistent with theories that highlight the negative economic effects of uncertainty shocks.

However, these approaches have the potential to underestimate the actual level of uncertainty, as they require qualitative expertise and human resources. For example, most naïve methods involve dictionary construction. This problem can be resolved with more complex methods based on machine-learning techniques. Despite their relative complexity, machine-learning approaches have greater predictive power than the naïve methods, empirical results show. The fastest and easiest method is to use an unsupervised machine-learning technique.

One of the most popular unsupervised natural language processing tools is the Latent Dirichlet Allocation (LDA), introduced by Blei et al. (2003). This generative statistical model divides a collection of texts into subgroups, with each subgroup being characterized by keywords associated with a topic. This method estimates the likelihood of the probability of the occurrence of words for a different number of topics. The results indicate the most likely number of topics. LDA is an unsupervised machine-learning technique, which does not require a training dataset. However, the model's results are unpredictable and require careful analysis. But while the methodology has been applied heavily in the machine-learning literature and for textual analysis, surprisingly, in economics it has so far only seen a small number of successful applications, e.g., Larsen et al. (2021) or Azqueta-Gavaldon (2017). Tobback et al. (2016) chose a different route and tried to improve the first EPU index designed by Baker through applying supervised machine learning. Thus, they developed a classification model based on support vector machines (SVMs) and labeled articles into two classes – related to economic policy uncertainty or not. Further, they constructed an EPU SVM indicator based on this classification, and include it in different macroeconomic models. This helps improve the accuracy of these models' economic variable forecasts in the short term.

However, unsupervised models, like naïve methods, also have disadvantages – the absence of sentiment analysis. This could be mitigated by machine-learning and lexicon-based techniques that use a predefined vocabulary and assess the relative frequency of sentiment words in the text. For example, Taboada et al. (2011) presented Semantic Orientation CALculator (SO-CAL). This model uses dictionaries of words annotated with their semantic polarity and strength featuring intensification and negation. SO-CAL can be used on completely unseen data. VADER (Valence Aware Dictionary for sEntiment Reasoning) is another successful example of a lexicon-based sentiment analysis tool. To develop it, Hutto and Gilbert (2014) constructed a list of lexical features and combined them with general rules that embody grammatical and syntactical conventions for expressing sentiment intensity. VADER outperformed many other highly regarded sentiment analysis tools. However, the main downside of lexicon-based techniques is the lack of trained dictionaries in languages other than English.

The introduction of a new language representation model called BERT (Bidirectional Encoder Representations from Transformers), developed by Google researchers (Devlin et al., 2018) was a significant breakthrough in sentiment analysis. Like many other recent works in pre-training contextual representations, BERT makes use of an attention mechanism that learns contextual relations between words (or sub-words) in a text. But unlike many other models, it is designed to pre-train deep bidirectional representations from the unlabeled text (treating on both left and right context). As result, BERT can distinguish differences in the usage of even the same word, taking into account the context of the occurrence of a given word. A pre-trained BERT model can be fine-tuned for a wide range of tasks, including classification. Pre-trained versions of BERT are available in a wide range of languages (including Ukrainian and russian).

To narrow down the scope of this paper, I focus on the simpler naive methods and unsupervised machine learning, and leave the sentiment analyses for future research. Starting with the simplest naïve methods, I will continue with more complex machine-learning methods of text classifications such as LDA. Thereafter, I develop an econometric model to assess the impact of the constructed indices on the formation of inflation expectations.

The paper is organized as follows: The next section presents data characteristics divided into two parts – a text corpus of economic news and inflation expectations in Ukraine. Section 3 describes the construction and results of news-based indices and presents their statistical properties. Section 4 analyzes the empirical specifications of the models and describes the results. Finally, Section 5 offers some concluding remarks and future steps. Additional information and results can be found in the Appendices.

2. DATA CHARACTERISTICS

2.1. News Corpus

The general key criteria in the selection of news sources were the availability of a fairly long archive (at least for the last ten years) and the possibility to scrape data from the web, which significantly limited the available list. Also, the newspapers used should have mainly economic orientation and not be subject to the explicit influence of individual political forces. I was guided by a list of the most popular resources, from which I excluded those that did not meet the specified requirements. In particular, I selected Ukrainian Pravda, Liga, and UNIAN, which are included in the list of TOP-50 Ukrainian online media based on Gemius and TNS ratings (at numbers 6, 11, and 20 respectively). In addition, I included data from the site Finance.ua, which is narrowly specialized in economic topics, and was also among the top 200 most popular sites in Ukraine (according to alexa.com). According to Similarweb.com, these sites in total covered about 8.5% of the traffic of online news and media in Ukraine in August-October 2022. However, these news sites do not provide information on the volume of views or coverage of each news item. This information could be useful in determining the strength of the news impact in the form of a multiplier.

The official language in Ukraine is Ukrainian, but the russian language was very popular before before russia's full-scale invasion, so most national media published materials in two languages. At the same time, the natural language processing infrastructure for the russian language was slightly better developed at the time the research was conducted, containing richer libraries, and modules that could improve the results of the research. For this reason, it was decided to process the news articles in russian.

The web-scraped news corpus consists of more than 2 million articles published online covering a sample of 20 years from January 2000 to December 2020. Our dataset contains the full text of the article and the available metadata, which include, for instance, the date, the link, and the title. Some sources also have a subtitle or general topic. All work with textual data, starting from the web-scraping and preprocessing and ending with the index construction was carried out using the Python programming language.

The textual data is presented in a non-traditional format, which makes statistical inference challenging. Thus, it is essential to preprocess the corpus data into a machine-readable format. Preprocessing includes some steps to clean and reduce the raw dataset before analysis.

First, it is essential to lowercase all of the characters to avoid any case-sensitive processing. This should help to clean the dataset at least in two cases:

– words with the uppercase letters may not be detected as a stopword, as all the stopword lists are lowercase. Stopwords are words that have no significant contribution to the meaning of the text. For example, the most common stopwords are conjunctions and prepositions;

– for grammatical reasons, the same word can be treated differently due to its position in the sentence. For example, the first word of the sentence is always uppercased, ven if it is not a proper name. As a result, the same word could be given two different values.

Second, but fundamental, step in NLP methods is tokenization. Tokenization is a method of breaking a piece of raw text into smaller units called tokens and converting them into a list. Tokens can be words, characters, or subwords. A token is a unit that NLP tools can easily convert to a value suitable for further machine processing.

Third, I removed non-essential information like stopwords from the text to simplify data processing. The NLTK library in Python has rich corpora of stopwords in different languages, including russian. Additionally, I removed non-ASCII characters, links, and punctuations by using Regular Expressions (Regex).

The fourth step of text preprocessing is text normalization. The most common normalization techniques for Natural Language Processing are stemming and lemmatization. Stemming is a technique that chops off the ends of words. Due to this approach, words with the same meaning but having some variations according to the context or sentence are normalized. Lemmatization usually refers to a morphological analysis of words, and normally aims to return the base or dictionary form of a word (lemma). russian, as well as Ukrainian, is a morphologically rich language, characterized by free word order and various word forms. Almost all language parts are marked for many characteristics as number, gender, case, tense, aspect, or person, which agreed grammatically with each other (Rozovskaya and Roth, 2019). Therefore, despite the longer processing time and the need for greater computing capacity, lemmatization is preferable to stemming for russian text normalization. For this purpose, I used the morphological analyzer pymorphy2, which returns the dictionary form of a word (Korobov, 2015).

Finally, I received a cleaned and normalized textual dataset consisting of around 300 million words and nearly 800,000 unique tokens. On average, as can be seen in Figure 1, the article size did not change significantly during the observed period (100-140 words per article). However, the number of articles grew considerably from a few articles per month in the early 2000s to between 8,000 and 14,000 per month from 2008, while the distribution of article quantity between online sources also changed (for more details about the news corpus see Appendix C). The increase in the number of online articles in the mid-2000s is related to the rapid growth of internet penetration in Ukraine. Thus, according to the State statistics service of Ukraine, the number of active internet users in Ukraine exceeded 1 million people for the first time in 2007, having gradually increased from 200,000 in 2000. Since that time, the number of active users has skyrocketed to more than 23 million, and accordingly, internet penetration rose to 56% in 2019. This forced news media to move from traditional paper forms to online versions. As a result, the media not only fully transferred their articles online but also expanded the content of websites with additional materials that are not usually placed in newspapers.

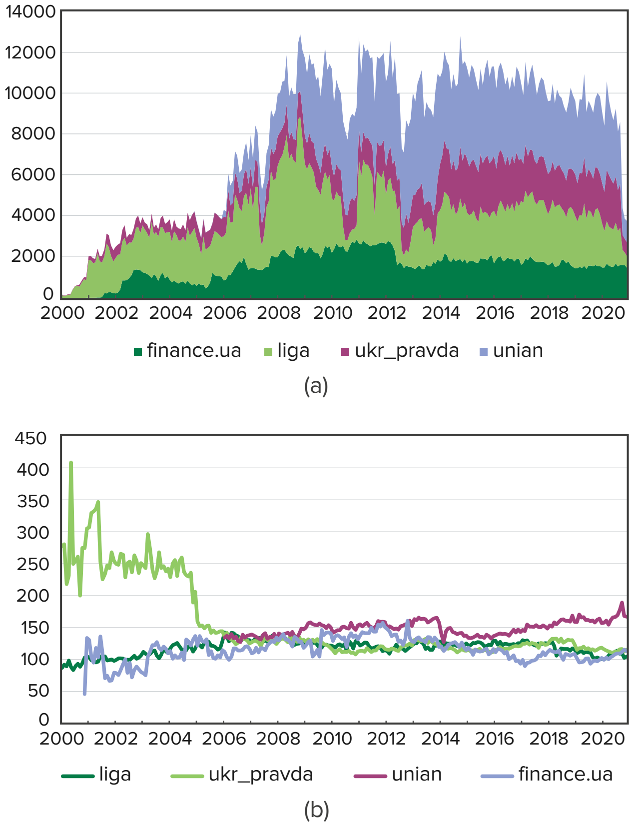


Figure 1. Number of News (a) and the Average Size of Articles (b) per Month

2.2. Inflation Expectations

The National Bank of Ukraine has been running surveys of inflation expectations for the next 12 months for several types of agents: households, banks, businesses, and professional forecasters. Before the adoption of the inflation targeting regime by the NBU in 2015, Coibion and Gorodnichenko (2015b) widely reviewed these surveys and discussed their limitations. In my paper, I will briefly describe the characteristics of the inflation expectations of all groups of respondents (Figure 2).

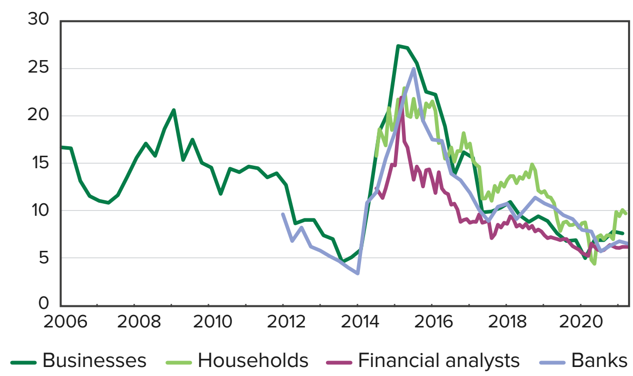


Figure 2. Inflation Expectations for the Next 12 Months, %

Source: NBU, GfK Ukraine, Info Sapiens.

Banks. The survey of banks covers at least 90% of the banking system's assets, excluding insolvent banks and banks in the process of liquidation. The NBU started to survey banks in 2012 and the data is available quarterly. Banks are surveyed during the first weeks of the quarter.

Businesses. This survey includes answers from about 700 non-financial sector enterprises. Enterprises are selected by the quota principle corresponding to the structure of Ukraine's economy, which ensures the representativeness of the sample. The business surveys have been conducted by the NBU quarterly since February 2006, however, they are conducted during the second month of the quarter.

Financial analysts. The NBU commenced monthly surveys of professional forecasters in July 2014 (during the second and third weeks of each month). Since November 2019, the frequency of this survey dropped to eight times a year to match the schedule of Monetary Policy Committee meetings. Responses from financial analysts are collected one week before the meeting date. The number of professional forecasters varies over time – from six to 12.

Households. Simultaneously with the surveys of financial analysts, a survey of households was launched in July 2014. Unlike the other surveys, the household survey is run monthly by the third-party company Info Sapiens (until 2019 it was run by GfK Ukraine). Every second and third week of the month approximately 1,000 consumers are surveyed about their inflation expectations, and on many other different social and economic issues. The sample is nationally representative and changes each month.

Banks, businesses, and households choose an interval of expected inflation for the next 12 months (more details in Appendix D). The resulting estimate is the weighted average of the midpoints of those intervals. The answer "hard to answer" is also available to households, and these answers are excluded from the calculation of average expectations. At the same time, financial analysts provide their discrete inflation forecasts (actual number, not an interval estimate), and their expectations are the simple average of these estimates. The latter can lead to periodic bias, as the number of experts in the survey is not constant.

Table 1 provides a brief statistical snapshot of inflation expectations in Ukraine. Historically, the expectations of professional forecasters have been lower than all other respondents (the mean is 2-4 pp lower than banks, businesses, and household expectations). However, they do not provide much more accurate forecasts, as forecast error fluctuates in different directions (figure 3), similar to other respondents. Thus, the RMSE of expectations of financial analysts is 12.0 p.p., which is higher than households' and businesses' expectations RMSE (11.4 p.p. and 11.3 p.p. respectively for the same period since July 2014). Banks' expectations show the worst forecasting power, with an RMSE of 13.1 p.p.

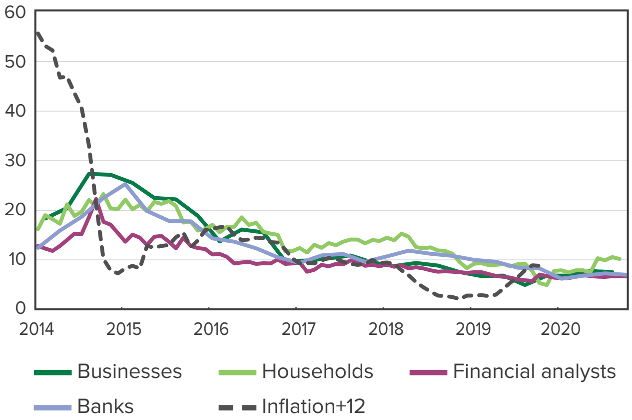


Figure 3. Inflation Expectations for the Next 12 Months and Actual Inflation (+12 months), %

Source: State Statistic Service of Ukraine, NBU, GfK Ukraine, Info Sapiens.

All the expectations have a positive skew, which means that the right tails are quite long. Meanwhile, household expectations are almost symmetrical, having only a small right-skewed tail. The distribution of household inflation expectations is flatter than normal, while all other expectations are more peaked.

3. CONSTRUCTING AGGREGATE NEWS INDEXES

The news content in the corpus is related mainly to economic, social, and political topics. Thus, the sample includes news that is associated not only with inflation developments or expectations (prices, supply of certain goods, tariffs, statistical information, forecasts, etc.). To focus

Table 1. Statistical Properties of Inflation Expectations

	Banks		Businesses		Households	Financial analysts
	Full sample	Since July 2014	Full sample	Since July 2014	Full sample (Since July 2014)	Full sample (Since July 2014)
Count (quarters or months)	38	28	61	27	81	75
Mean, %	10.660	12.120	13.070	13.430	13.780	9.880
std, p.p.	4.990	4.880	5.390	6.920	4.680	3.430
min, p.p.	3.500	5.800	4.700	5.100	4.510	5.340
25%, p.p.	6.890	9.150	9.000	7.800	9.790	7.200
median, p.p.	9.920	10.650	12.760	10.000	13.550	8.800
75%, p.p.	12.000	14.330	15.800	18.650	17.140	12.180
max, p.p.	24.900	24.900	27.300	27.300	22.890	21.900
Skewness	1.074	1.045	0.731	0.811	0.108	1.082
Kurtosis	0.653	0.556	0.132	-0.699	-1.002	0.852

only on the factors determining inflation expectations, I remove news items related to unrelated topics. I apply two different approaches to filter out the news. First, I use a dictionary-based approach to build a set of indexes based on the raw count of news. Second, I implemented a topic analysis using Latent Dirichlet Allocation (LDA) according to Blei et al. (2003).

Both approaches do not take into account the sentiments of news content. However, this may not be a huge problem, as usually the news is biased negatively – Hester & Gibson (2003) found that economic news was written in a negative tone more often than in a positive one. Additionally, they proved that negative news was a significant predictor of consumer expectations about future economic developments. Damstra & Boukes (2018) explain this negative bias of news well, giving a few main reasons:

– free media perform a crucial role in overseeing government, so negative events receive more attention, while positive ones do not meet such a need;

– in the process of judging the newsworthiness of real-world events negativity can be a key value, consequently, a “bad” news story is more likely to be selected by journalists;

– negative events have a stronger news impact than positive ones.

Moreover, Soroka et al. (2019) define the negative tone of news as an essential feature, while good news, in contrast, may be considered as the absence of news. Therefore, to construct simple indexes, it is possible to assume that there is a tendency for news to have a negative impact on perception and expectations.

Of course, a sentiment analysis could be useful to determine the impact of news on economic expectations. In particular, this approach could clean the data series of contradictory events that could have opposite effects. In addition, separating news into positive and negative would help in exploring the possible non-linearity of the impact of multidirectional sentiment. However, to apply this approach, it is important to create a high-quality training dataset, based on clear rules and with the involvement of several independent experts. In addition, sentiment analysis requires significant computing power and time, which may not be reasonable at the initial stages of a study.

3.1. Dictionary-Based Approach

The dictionary-based approach is the simplest approach to estimating the impact of news on various macroeconomic indicators. These indices are calculated as a share of articles related to the topic, commonly denoted as “document frequency”. The intuition behind these indices is that the more alarming the topic, the more articles would be written on the subject – for example in times of crisis.

Document frequency (df) is the fraction of the documents that contain a certain term, and is obtained by dividing the number of documents containing the term by the total number of documents:

$$df\left(t,D\right)=\frac{d}{N}, \tag{1}$$

where N is the total number of documents in the corpus D, and number of documents d where the term appears.

The dictionary-based approach to constructing news indices requires expertise in selecting the relevant keywords. In this case, to determine which prices concern Ukrainians the most, I turned to the consumer basket of the average household. Ukrainians spend the most of their income on food. In various periods, the share of spending on food and soft drinks was 40-60% for the period from 2000 to 2020, slightly decreasing in recent years. Accordingly, it is important to select news that contains mentions of basic foods: bread, meat, dairy, vegetables, fruits, etc.

Another essential component of household expenditure is utilities. Although the share of this type of spending is much lower than in many other countries, utility tariffs are important to Ukrainians and are often used as a political football by politicians. They therefore may have a visible impact on expectations. The most important utilities for Ukrainians are electricity and natural gas.

Fuel prices may also have a significant impact on the formation of households’ inflation expectations, even though not all people use private transport. For example, Kilian and Zhou (2020) found several episodes since 1990 in the United States when household inflation expectations growth could almost entirely be explained by hikes in fuel prices. On the one hand, this is due to the ubiquity of gas stations and fuel price boards, which allow their easy use for daily price monitoring. On the other hand, everyone is well aware that fuel is a component of the cost of most goods and services, explicitly or implicitly. In this case, I include news not only about fuel but also about oil, which is a defining cost component of fuel.

As stated in Coibion & Gorodnichenko (2015b) there is a strong positive correlation between inflation expectations and exchange rate developments – especially for households. In this case, logically, not only do exchange rate dynamics affect expectations, but also the coverage of this topic in the media.

In addition, I will analyze the index of news related to the word inflation itself, as such news often contains expert forecasts or analyses of the current situation. According to Zholud et al. (2019), inflation expectations in Ukraine are highly linked to current inflation trends and so have a future-oriented component. Therefore, it is advisable to check the impact of references to inflation in the media on the formation of expectations.

Figure 4 shows the resulting indices calculated using a dictionary-based approach. Since the earliest data on inflation expectations of respondents are from 2006, all the time series of news indices will start from 2005 (one year back to assess lag effects). However, given that the volume of news was much smaller in the early 2000s, this meant only about 10% of the articles had to be removed from the corpus, and there are still about 1.8 million articles left. News related to food has the smallest share among selected topics, while news about fuel appears to be the most important. Also, it can be seen that the document frequency of news related to utilities, in general, decreased over time – except in 2015 when there was a significant jump in the importance of this topic due to utility tariffs being raised to market levels. Interestingly, until 2014 the topic of exchange rate movements was mentioned more often in the news, which is probably due to the greater negative consequences of the sharp devaluation observed in Ukraine at that time with the abandonment of the fixed exchange rate regime. For more information about indices built with a dictionary-based approach, see Appendix E (Supplementary Materials).

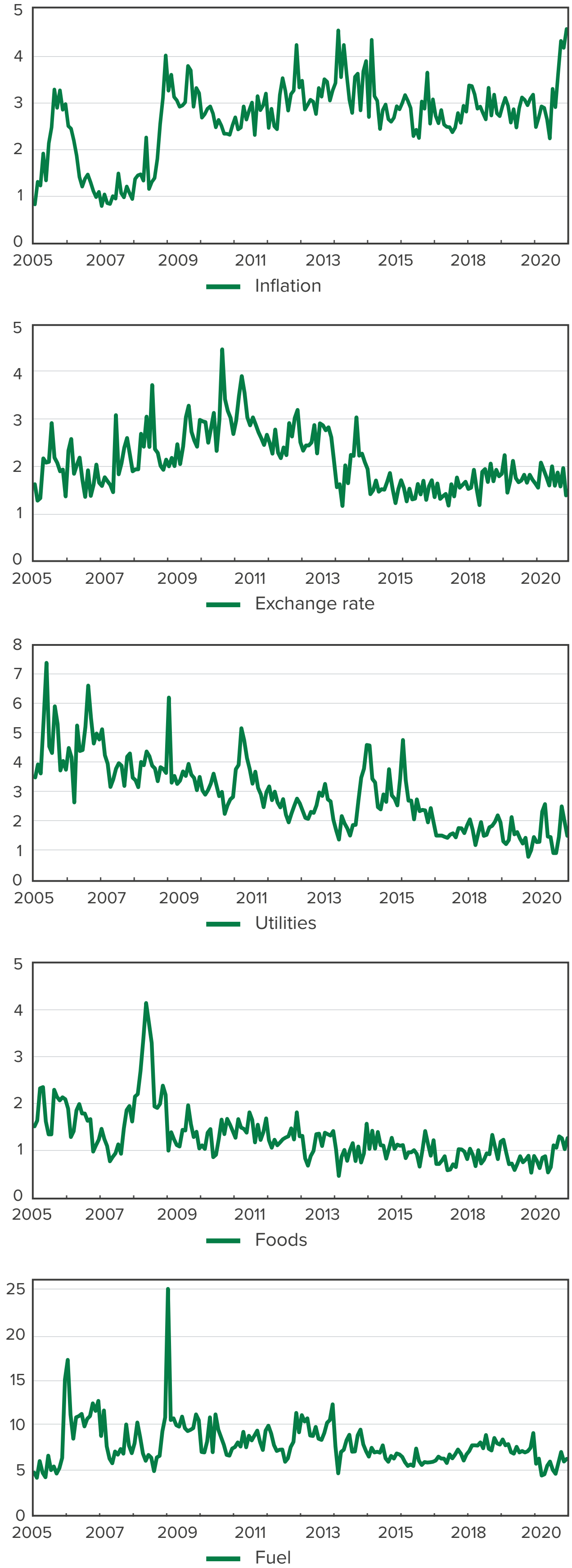


Figure 4. Document Frequency of Topics Relevant to Inflation Expectations

As the inflation expectations of the various respondents are collected at different periods and not evenly throughout the month, the impact of certain short-lived or even discrete news may be extremely important. Thus, some news may last only for several days, and due to the rapid loss of interest in the topic the effect on monthly document frequency can fade away. Therefore, the monthly indices may not reflect the real dynamics of the importance of individual events, and applying indices with a higher frequency may shed light on this issue. To assess this impact, I additionally computed similar indices in decade (10-day) terms for each month – a decade being a third of the month (results are in Appendix E (Supplementary Materials).

It is important to assume the independence of researched variables, which can be indicated by their stationarity. Stationarity is necessary when applying many statistical tools and procedures in a time series analysis. Indeed, if the data was generated by a stationary process, it will have the properties of a sample generated by such a process. According to the Dickey-Fuller test, the monthly time series for the share of utilities, exchange rate, and inflation topics in the news is non-stationary. Resultantly, the relevant indices in a decade-term-only time series for the share of utilities and inflation topics in the news are non-stationary. Additionally, the autocorrelation is high, and it seems that there is no clear seasonality. Therefore, to get rid of the high autocorrelation and to make the entire process stationary, in the same way, I take the first differences.

3.2. Unsupervised ML Approach

One of the important shortcomings of the dictionary-based approach is the availability of quality expertise and the selection of texts based on it. In particular, an article may contain keywords, but its topic may be a completely different issue. For example, the word “fuel” can be attributed to topics related to science and technology or car manufacturing. The solution here is to use unsupervised topic modeling algorithms. These statistical methods analyze the words of the collection of texts and divide them into subgroups, where each subgroup is associated with a set of keywords. Thus, the model finds combinations of words, rather than single ones. In our “fuel” example, articles with word combinations “fuel price” and “rocket fuel” would be distinguished. Most machine-learning models require the use of a part of a data set in which specially trained people classify information according to a predetermined procedure and therefore put labels on data. However, some methods do not need such labeled training samples. Latent Dirichlet Allocation (LDA), presented by Blei et al. in 2003, it today is a very common example of a topic modeling method that uses an unsupervised learning algorithm.

I used an extremely efficient implementation of LDA called LightLDA provided in the nimbusml Python module (Yuan et al., 2015). This state-of-the-art implementation incorporates several optimization techniques and can train a topic model on large document sets much faster. For example, our model produces 100 topics on a 2 million news item dataset in less than an hour, while using the full LDA at this scale takes days. Figure 6 (Appendix B) shows the distribution of topics received with LDA. The popularity of some topics changed over time, while others remained relevant throughout the observation period.

The number of topics in LDA is not fixed and can be set according to the task. I experimented with using a different number of topics. I observed that with a larger number of topics our main results do not change – some topics have very similar content and have to be combined in further analysis. At the same time, the interpretation of a larger number of topics becomes more complicated. With a lower number of topics, it is sometimes difficult to distinguish between different topics that have similar keywords. For example, topics related to exchange rates may include unnecessary information, as some articles contain similar words, but different content.

At this point, human intervention is necessary to analyze and label the topics of the received news clusters. Figure 7 (Appendix B) is a graph showing the relationship between the topics distributed by the LDA. Most of the news clusters are as expected attributed to politics, international relations, parliament, and government. At the same time, some topics are different and can be linked to economic topics that may affect inflation expectations. Most news topics do not belong to one, but to several clusters.

I identify a news cluster related to exchange rate movements, which includes six news topics defined with the help of LDA. As we already know, the situation in the foreign exchange market has some influence on the formation of inflation expectations in Ukraine. I also found a cluster related to commodities, including oil and gas. Additionally, topics describing the electricity market, budget, and government debt can be easily identified. Interestingly, LDA helped identify a topic related to the spread of coronavirus, for which the number of articles unsurprisingly increased from the end of 2019. In addition, LDA has a well-defined topic for the period of the war between russia and Ukraine from 2014 and subsequent years. However, LDA did not group articles related to food prices, utility tariffs, etc., in separate recognizable topics, which can be explained by their having similar structures, as well as their share of the news being relatively low. Increasing the number of topics does not fix this. In Figure 8 (Appendix B) I provide wordclouds for a few of the most relevant topics.

The popularity of certain topics closely corresponds to the historical development of events. In particular, the share of articles on the hryvnia exchange rate, seen in Figure 5, increased in 2008, when the hryvnia depreciated rapidly amid the global financial crisis. The next peak was observed in 2014-2015, when, due to the war between russia and Ukraine and the loss of control of a part of Ukraine's territory, the economy suffered a significant blow. At this time, the hryvnia also depreciated rapidly. But with the transition to a floating exchange rate and stabilization in the foreign exchange market, interest in this topic in the news began to wane.

News about gas and oil behaved in a similar way. Thus, in 2006-2008, the gas issue was extremely important for Ukraine against the background of difficult relations with russia. Problems with gas supply were repeated in 2014. In contrast, lower energy prices have contributed to less coverage of these topics in subsequent years.

I built the indexes the same way as in the dictionary-based approach, using equation 1 of document frequency. Thus, monthly indices were calculated to assess long-term impact, while decadal indices were calculated to assess short-term media shocks that may be important during the time of the inflation expectation survey, but then fade over the month. More details about the statistical characteristics of the indices constructed by LDA division can be found in Appendix F (Supplementary Materials). According to

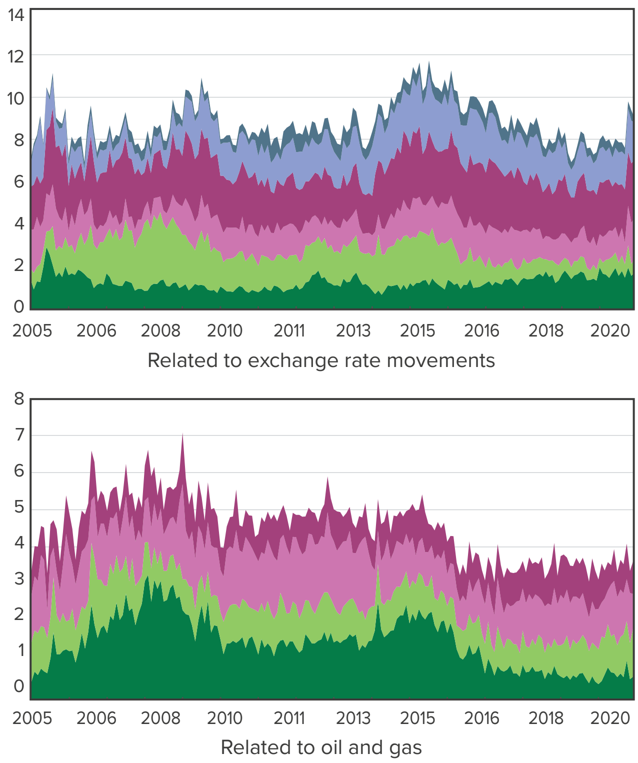


Figure 5. Share of Topics, Identified by LDA, document frequency, %

the Dickey-Fuller test, monthly time series for a share of energy news are non-stationary, while the share of news on exchange rate movements is stationary. Decade time series can be considered stationary with a 95% probability.

4. ESTIMATION RESULTS

As noted in previous sections, inflation expectations are largely shaped by past inflation (Zholud et al., 2019; Coibion and Gorodnichenko, 2015b). Therefore, for the analysis I used an extrapolative approach to the formation of inflation expectations (Lines and Westerhoff, 2010):

$$E\pi_t = \alpha + \beta\pi_{t-1} + \gamma\left(\pi_{t-1} - \pi_{t-2}\right) + \varepsilon, \tag{2}$$

where $E\pi_t$ is expected inflation in period t, π_{t-1} denotes inflation in the previous period, and $\pi_{t-1} - \pi_{t-2}$ stands for the change in inflation, α and γ – are coefficients of regression, while ε is the error. I use annual CPI change as a measure of inflation.

In this research, I assume that the formation of inflation expectations (equation 2) is influenced by the media environment rather than actual changes in inflation:

$$E\pi_t = \alpha + \beta\pi_{t-1} + \delta df_T^m + \varepsilon, \tag{3}$$

where df denotes document frequency of the news topic m in period T. T may be equal to t when testing the impact of news on the formation of inflation expectations in the same month a survey is conducted. However, some surveys are conducted at the beginning of the month, therefore, I test the impact of the frequency of news publications during the previous three months on the formation of inflation expectations. Accordingly, T can be equal to t – 1, t – 2, and t – 3. I test the monthly and decade frequency of T, as the inflation expectations survey is not conducted for a whole month, but for shorter periods. In addition, these periods also vary for different respondents. As quarterly surveys are not conducted throughout the quarter, I use matching months instead of aggregating news indices at the quarterly level. For example, bank surveys take place in the first month of the quarter, thus the same month for the news index was used as the base month.

I also test another variation of extrapolative inflation expectations, assuming that the respondents' expectations change in response to changes in current inflation. In this case, the formula of inflation expectations looks like this:

$$E\pi_t - E\pi_{t-1} = \alpha + \gamma\left(\pi_{t-1} - \pi_{t-2}\right) + \varepsilon. \quad (4)$$

I expanded formula 4 with changes in media environments in changes in current inflation:

$$E\pi_t - E\pi_{t-1} = \alpha + \gamma\left(\pi_{t-1} - \pi_{t-2}\right) + \eta\left(df_T^m - df_{T-1}^m\right) + \varepsilon. \quad (5)$$

In this case, all our variables are stationary and we can be sure that their properties do not change over time.

I also assume that the impact of the constructed news indices on inflation expectations is linear, so to estimate this effect I use the OLS regression.

I start with analyzing the impact of news on the formation of inflation expectations using the dictionary-based approach. Table 2 presents the coefficients and p-values (in brackets) of news indices built with a dictionary-based approach in OLS regressions of the inflation expectations of different groups of respondents. This table shows two different approaches: without transformations (equation 3), using all variables as they were computed, and the first differences of all variables, presenting an extrapolation of the change of inflation expectations (equation 5). R^2 for the first type of relationship is as expected much higher than for estimates of the first difference. However, the relatively low R2 for such studies is quite normal and typical for studies of human behavior (King, 1986).

As can be seen from Table 2 (Appendix A), all types of inflation expectations are dependent on current inflation trends as the coefficients are statistically significant. At the same time, only banks and businesses tie changes in inflation expectations to recent changes in inflation, while the relationship between changes in household and financial analyst expectations with recent inflation dynamics is insignificant. This is in line with the opinion that well-anchored long-term inflation expectations should not change in response to news about macroeconomic indicators, in particular inflation (Galati et al., 2011). However, it is too early to talk about anchored inflation expectations, given the difference between the central bank's inflation target and inflation expectations. Therefore, in this case, the result could be due to information rigidity.

The banks' inflation expectations are virtually independent of the current media environment on inflation. Most indicators are not statistically significant or contradict economic logic. For example, the banks' inflation expectations are negatively correlated with food news with a 90% probability. That is, the more this topic is talked about in the media, the faster it reduces the inflation expectations of banks. This might be explained by the tone and content of the news. However, without a more detailed study of the content of this news, this is impossible to determine.

At the same time, it is interesting that banks change their inflation expectations under the influence of changes in the information environment around utility tariffs and fuel, as well as news about inflation in previous periods. The rather significant lag of 2-3 months can be explained by the time needed for the preparation of macroeconomic forecasts, which are the basis for the answers in the survey.

Similar to banks, businesses' expectations may be significantly affected by news about past inflation trends, and by utilities. However, businesses are the most sensitive to the news about food. This can be explained by a high share of agriculture, food industry, retail, and wholesale trade (related to food) enterprises among the surveyed ones, which also corresponds to the structure of the economy of Ukraine. Food news is also an important factor in the estimation of changes in businesses' inflation expectations.

Households' expectations may be the most sensitive to the number of news items related to utility tariffs in the reported period and the previous quarter. This can primarily be explained by the high importance of utility tariffs for Ukrainian households. Thus, a significant share of tariffs are regulated by the government or local authorities, and changes in tariffs cause a substantial negative reaction from society. The share of utility tariffs in the CPI is relatively low, which is largely due to the non-monetary subsidies that were in place in previous periods. However, despite this, the average Ukrainian utility tariffs are some of the most important topics related to inflation, which is confirmed, among other things, by the results of our analysis. Households are also slightly sensitive to information about food in previous periods. Interestingly, these results are not confirmed by changes in households' inflation expectations. Thus, citizens change their estimates of future inflation under the influence of changes in the information field about the exchange rate in the previous three months, while the change in the importance of other topics has little effect on expectations dynamics.

The expectations of professional forecasters respond best to information on utility tariffs in the reporting and previous months, as well as on the exchange rate in the previous three months. In this case, financial analysts respond to both the amount of information, and its change. This may reflect approaches to forecasting for such analysts. As usual, changes in the exchange rate and the expected changes in utility tariffs have the greatest impact on the revision of forecasts.

I also test the hypothesis that shorter-term trends in the media environment may better explain the process of the formation of inflation expectations among different respondents. This is in line with the fact that most surveys are conducted in a shorter period than a month. To this end, I use decadal indices of frequency for mentions of these topics. Shocks in the news that last for several days can fade within a month, due to the rapid loss of interest in the topic, and therefore the monthly indices may not reflect the real dynamics of the importance of individual events. Thus, it is important to apply indices with a higher frequency. Going to the decade level, I get a mixed frequency in the OLS, so to switch to one frequency, I just used matching by month. Thus, I compare news indices separately for the first, second, and third decades of the reporting month with inflation expectations for the same reporting month. The procedure was repeated for individual news indices for the three decades of the previous month against the inflation expectations of the reporting month, as respondents also responded to the dynamics of the media environment in previous periods.

Table 3 (Appendix A) presents the results of OLS estimations of the impact of decade news indices on the formation of inflation expectations. As in the previous case, I add the latest available inflation data, which is published with delay, so I use actual inflation in period t-1.

Here we have a few interesting outcomes that deviate from our monthly estimations. For example, banks may be sensitive to the document frequency of news about the exchange rate in all decades of the previous month, while the monthly indices do not show this relationship. At the same time, food news may be more important in the last decade of the previous month, although monthly indices show significance for the current month. Businesses respond more to the news about utilities and fuel in the first decade of the reporting month. As with monthly indices, decade indices on utility tariffs may affect the formation of the inflation expectations of households. In this case, the first decades of the reporting and previous months are the most important. The expectations of financial analysts also proved to be most dependent on the frequency of news in the first decades of the reporting and previous months. However, in addition to utilities, they may follow the news about inflation (official figures are only published in the first decade of the month) and about fuel.

Another important opinion concerns the fact that the expectations of banks and enterprises are collected once a quarter. Therefore, the period for assessing the impact of news on inflation expectations was increased by applying a three-month moving average. This is especially important considering that the coefficients for the monthly indices are very volatile and can even change sign, depending on the applied lag. Table 4 (Appendix A) presents the results of OLS estimations of the impact of quarterly news indices (3-month moving average) on the formation of inflation expectations.

As can be seen, the hypothesis that banks and enterprises follow longer trends is mostly not confirmed. At the same time, the long-term change in the information space about inflation and the exchange rate is related to the change in inflation expectations of banks, and the change in the volume of food news affects the inflation expectations of enterprises. However, in both cases, this impact is limited to 1-2 quarters.

I repeated a similar procedure to reveal the impact of indices built by the LDA on the formation of inflation expectations. Table 5 (Appendix A) presents the coefficients and p-values (in brackets) of news indices built by the LDA in OLS regressions of inflation expectations of different groups of respondents. Similar to simple indices I tested two different approaches: without transformations, using all variables as they were computed, and the first difference of all variables, presenting an extrapolation of the change of inflation expectations.

According to the results of the regressions, I observe a weak correspondence between the news about energy and utility tariffs that were determined by LDA, and the formation of inflation expectations. Inflation expectations of households, as well as changes in business expectations, demonstrate significance in the reporting month at the 5% and 10% levels respectively, but the sign of the coefficients of these variables contradicts economic logic, which can be associated either with emotional coloring or a reflection of a coincidence of circumstances.

Instead, the situation is somewhat different with the exchange rate news set obtained by LDA. The frequency of such news in the reporting period was significant for the formation of expectations of businesses, households and financial analysts. For households and financial analysts, these indices were also important in recent months. Financial analysts and households were also sensitive to changes in the frequency of exchange rate news. However, households change their expectations in response to more recent developments, while financial analysts respond to a longer period.

Similar to simple indices, I identified short-term spot effects on the formation of inflation expectations by estimating decade indices. Table 6 (Appendix A) represent the results of this estimation.

Interestingly, for some groups of respondents, there is a clear relationship with the indices in the periods when the surveys are conducted. For example, bank surveys are usually conducted at the beginning of the month, and sometimes even cover the last week of the previous month. News about energy in the last decade of the previous month and in the first decade of the reporting month turned out to be significant. There is a similar situation with businesses and households. At the same time, the sign of the coefficients needs further study in terms of sentiment. Inflation expectations of businesses are formed under the influence of news about the exchange rate in the first decade of the reporting month, while all other respondents follow the news for previous periods.

I repeated the same procedure for determining the longer-term impact of news on the formation of inflation expectations using the current three-month average for banks and corporates. Table 7 (Appendix A) shows the main results of the estimations. However, the results indicate the absence of any long-term impact of news on the formation of inflation expectations. Only the expectations of enterprises have a significant connection with the change in the frequency of news about the exchange rate in the current quarter.

Thus, the formation of inflation expectations among different groups of respondents may depend on the media environment, namely both the volume of published articles and changes in this indicator. It is important to note that different groups of respondents may rely on different topics and different periods when estimating future inflation. It can also be seen that recent news, published during the month and even the decade preceding the survey, is mostly more important in shaping inflation expectations than older news. This may, among other things, be important for the central bank's communication policy.

5. CONCLUSIONS

In this paper, I describe the process of analyzing textual data to determine the role of news in the formation of inflation expectations among various types of respondents in Ukraine. I have scraped a news corpus from four Ukrainian online newspapers listed in the most popular online media in Ukraine, which mainly have an economic focus. Using natural language processing and machine-learning techniques, I have cleaned and transformed the textual data into news-based quantitative measures reflecting news topics relevant to inflation and inflation expectations.

I apply two different approaches to filter out the news: a dictionary-based approach and Latent Dirichlet Allocation

(LDA). Both approaches disregard the sentiments of news content, which I leave for future research. I compute all news indices as a "document frequency" following the intuition that the more alarming the topic, the more articles would be written on the subject.

I assume that the impact of the constructed news indices on inflation expectations is linear and estimate this effect with OLS regression. I have tested the impact on the level of inflation expectations, as well as on the change thereof. I have found evidence that different news topics may have a different impact on the inflation expectations of various groups. For example, I find there is a strong relationship between the inflation expectations of households and financial analysts with news about utilities, while businesses are sensitive to news about food. Additionally, financial analysts and households are also sensitive to levels and changes in the frequency of exchange rate news, as shown by LDA.

I also test the hypothesis that shorter-term trends in the media environment may better explain the formation of the inflation expectations of different respondents, as document frequency may vary during the month and the impact of the short-term news may fade away. I prove that for some groups of respondents there is a clear relationship with the indices within the periods when the surveys are conducted. I also show that recent news is mostly more important in shaping inflation expectations than older news.

As a result, the formation of the inflation expectations of different groups of respondents may depend on the media environment, namely the volume of published articles and changes in this indicator. Different groups of respondents rely on different topics and different periods when assessing future inflation. I also find that some events contradict economic logic, which could be a question for future research. In particular, an important issue is the impact of news indices in different periods (i.e. during periods of stability, accelerating inflation, or disinflation). Other future research questions may include an assessment of the tone of news, the relationship of the news indices to other macroeconomic indicators, as well as the predictive power of such indices. Another important issue may be the examination of nonlinearities. In particular, the impact of news may differ depending on the level of current inflation (e.g., economic agents may pay more attention to the news when inflation is high and vice versa) and the monetary policy regime.

These results complement previous studies on the formation of rational inflation expectations. In other words, the overall level of inflation expectations is generally determined by past inflation, and small fluctuations may well be explained by other factors, including the influence of the media environment. The results of this research can aid in understanding inflation expectations, which is important given that anchoring inflation expectations remains a key challenge for central banks. This may, among other things, be important for the central bank's communications policy, and help it to both articulate clear and effective messages and design optimal policy.

REFERENCES

Angelico, C., Marcucci, J., Miccoli, M., Quarta, F. (2021). Can we measure inflation expectations using Twitter? Bank of Italy Working Papers, 1318. Rome: Bank of Italy. Retrieved from https://www.bancaditalia.it/pubblicazioni/temi-discussione/2021/2021-1318/en_tema_1318.pdf

Azqueta-Gavaldón, A. (2017). Developing news-based Economic Policy Uncertainty index with unsupervised machine learning. Economics Letters, 158, 47–50. https://doi.org/10.1016/j.econlet.2017.06.032

Baker, S. R., Bloom, N., Davis, S. J. (2016). Measuring economic policy uncertainty. The Quarterly Journal of Economics, 131(4), 1593–1636. https://doi.org/10.1093/qje/qjw024

Bauer, M. D. (2015). Inflation expectations and the news. International Journal of Central Banking, 11(2), 1-40. Retrieved from https://www.ijcb.org/journal/ijcb15q2a1.pdf

Blei, D. M., Ng, A. Y., Jordan, M. I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3, 993–1022. Retrieved from https://dl.acm.org/doi/10.5555/944919.944937

Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. The Quarterly Journal of Economics, 118(1), 269–298. https://doi.org/10.1162/00335530360535207

Coibion, O., Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? Journal of Political Economy, 120(1), 116–159. https://doi.org/10.1086/665662

Coibion, O., Gorodnichenko, Y. (2015a). Information rigidity and the expectations formation process: A simple framework and new facts. American Economic Review, 105(8), 2644–2678. https://doi.org/10.1257/aer.20110306

Coibion, O., Gorodnichenko, Y. (2015b). Inflation expectations in Ukraine: A long path to anchoring? Visnyk of the National Bank of Ukraine, 233, 6–23. https://doi.org/10.26531/vnbu2015.233.006

Coibion, O., Gorodnichenko, Y., Weber, M. (2019). Monetary Policy Communications and their Effects on Household Inflation Expectations. NBER Working Paper Series, 25482. Cambridge: National Bureau of Economic Research. https://doi.org/10.3386/w25482

D'Acunto, F., Malmendier, U., Ospina, J., Weber, M. (2019). Exposure to daily price changes and inflation expectations. CESifo Working Paper, 7798. Munich: CESifo. Retrieved from https://www.cesifo.org/en/publications/2019/working-paper/exposure-daily-price-changes-and-inflation-expectations

Damstra, A., Boukes, M. (2018). The economy, the news, and the public: A longitudinal study of the impact of economic news on economic evaluations and expectations. Communication Research, 48(1), 26–50. https://doi.org/10.1177/0093650217750971

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 1, 4171–4186. http://doi.org/10.18653/v1/N19-1423

Dräger, L. Lamla, M. J. (2017). Imperfect information and consumer inflation expectations: Evidence from microdata. Oxford Bulletin of Economics and Statistics, 79(6), 933–968. https://doi.org/10.1111/obes.12189

Galati, G., Heemeijer, P., Moessner, R. (2011), How do inflation expectations form? New insights from a high-frequency survey. BIS Working Papers, 349. Basel: Bank for International Settlements.

Garcia, J. A. Werner, S. (2018). Inflation news and Euro area inflation expectations. IMF Working Papers, 167. Washington: International Monetary Fund. https://doi.org/10.5089/9781484363010.001

Hester, J. B., Gibson, R. (2003). The economy and second-level agenda setting: A Time-series analysis of economic news and public opinion about the economy. Journalism & Mass Communication Quarterly, 80(1), 73–90. https://doi.org/10.1177/107769900308000106

Hutto, C .J., Gilbert, E. (2014). VADER: A parsimonious rule-based M\model for sentiment analysis of social media text. Proceedings of the International AAAI Conference on Web and Social Media, 8(1), 216–225. https://doi.org/10.1609/icwsm.v8i1.14550

Kilian, L., Zhou, X. (2020). Oil prices, gasoline prices and inflation expectations: A new model and new facts. CEPR Discussion Paper, 15168. London: Centre for Economic Policy Research. Retrieved from https://cepr.org/publications/dp15168

King, G. (1986). How not to lie with statistics: avoiding common mistakes in quantitative political science. American Journal of Political Science, 30, 666. https://doi.org/10.2307/2111095

Korobov, M. (2015). Morphological analyzer and generator for russian and Ukrainian languages. In: Khachay, M., Konstantinova, N., Panchenko, A., Ignatov, D., Labunets, V. (eds) Analysis of Images, Social Networks and Texts. AIST 2015. Communications in Computer and Information Science, 542. Springer. https://doi.org/10.1007/978-3-319-26123-2_31

Larsen, V. H., Thorsrud, L. A., Zhulanova, J. (2021). News-driven inflation expectations and information rigidities. Journal of Monetary Economics, 117, 507–520. https://doi.org/10.1016/j.jmoneco.2020.03.004

Lines, M., Westerhoff, F. H. (2010). Inflation expectations and macroeconomic dynamics: the case of rational versus extrapolative expectations. Journal of Economic Dynamics and Control, 34(2), 246–257. https://doi.org/10.1016/j.jedc.2009.09.004

Maćkowiak, B., Wiederholt, M. (2009). Optimal sticky prices under rational inattention. The American Economic Review, 99 (3), 769-803. https://doi.org/10.1257/aer.99.3.769

Mankiw, N. G., Reis, R.F., Wolfers, J. (2003). Disagreement about inflation expectations. NBER Macroeconomics Annual, 18, 209–248. Cambridge: National Bureau of Economic Research. https://doi.org/10.1086/ma.18.3585256

Mazumder, S. (2021). The reaction of inflation forecasts to news about the Fed. Economic Modelling, 94, 256–264. https://doi.org/10.1016/j.econmod.2020.09.026

Nautz, D., Pagenhardt, L., Strohsal, T. (2017). The (de-) anchoring of inflation expectations: New evidence from the euro area. The North American Journal of Economics and Finance, 40, 103–115. https://doi.org/10.1016/j.najef.2017.02.002

Pfajfar, D., Santoro, E. (2013). News on inflation and the epidemiology of inflation expectations. Journal of Money, Credit and Banking, 45(6), 1045–1067. https://doi.org/10.1111/jmcb.12043

Rozovskaya, A., Roth, D. (2019). Grammar error correction in morphologically rich languages: The case of russian. Transactions of the Association for Computational Linguistics, 7, 1–17. https://doi.org/10.1162/tacl_a_00251

Sims, C. (2009). Inflation expectations, uncertainty and monetary policy. BIS Working Papers, 275. Basel: Bank for International Settlements. Retrieved from https://www.bis.org/publ/work275.htm

Soroka, S., Fournier, P., Nir, L. (2019). Cross-national evidence of a negativity bias in psychophysiological reactions to news. Proceedings of the National Academy of Sciences of the United States of America, 116(38), 18888–18892. https://doi.org/10.1073/pnas.1908369116

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M. (2011). Lexicon-based methods for sentiment analysis. Computational Linguistics, 37(2), 267–307. https://doi.org/10.1162/COLI_a_00049

Tobback, E., Naudts, H., Daelemans, W., Junqué de Fortuny, E., Martens, D. (2016). Belgian economic policy uncertainty index: Improvement through text mining. International Journal of Forecasting, 34(2), 355–365. https://doi.org/10.1016/j.ijforecast.2016.08.006

Woodford, M. (2004). Inflation targeting and optimal monetary policy. Federal Reserve Bank of St. Louis Review, 86(4), 15-41. St. Louis: The Federal Reserve Bank of St. Louis. Retrieved from https://files.stlouisfed.org/files/htdocs/publications/review/04/07/Woodford.pdf

Yuan, J., Gao, F., Ho, Q., Dai, W., Wei, J., Zheng, X., Xing, E.P., Liu, T., Ma, W. (2015). LightLDA: Big topic models on modest computer clusters. WWW'15: Proceedings of the 24th International Conference on World Wide Web, 1351–1361. https://doi.org/10.1145/2736277.2741115

Zholud, O., Lepushynskyi, V., Nikolaychuk, S. (2019). The Effectiveness of the Monetary Transmission Mechanism in Ukraine since the Transition to Inflation Targeting. Visnyk of the National Bank of Ukraine, 247, 19–37. https://doi.org/10.26531/vnbu2019.247.02

APPENDIX A. TABLES

Table 2. Relationship between Monthly News Indices and Inflation Expectations

Respondents	Variables	Without transformations					Variables	1st difference				
		Inflation	Exchange rate	Utilities	Food	Fuel		Inflation	Exchange rate	Utilities	Food	Fuel
Banks	π_{t-1}	0.2949*** (0.000)	0.2890*** (0.000)	0.2994*** (0.000)	0.2970*** (0.000)	0.2950*** (0.000)	$\pi_{t-1} - \pi_{t-2}$	0.1047** (0.022)	0.1545*** (0.005)	0.1642*** (0.003)	0.1658*** (0.003)	0.1737*** (0.001)
	df^m_t	-1.2029 (0.194)	0.9028 (0.510)	-0.7928 (0.356)	-4.0589* (0.099)	0.0736 (0.824)	$df^m_t - df^m_{t-1}$	0.6139 (0.379)	1.5717 (0.189)	0.096 (0.903)	1.6328 (0.460)	0.3832 (0.220)
	df^m_{t-1}	1.1547 (0.360)	-0.7359 (0.662)	2.2787* (0.099)	2.9647 (0.104)	-0.3926 (0.346)	$df^m_{t-1} - df^m_{t-2}$	1.5263 (0.123)	-2.2497 (0.153)	1.2997 (0.306)	0.8267 (0.629)	-0.1084 (0.772)
	df^m_{t-2}	1.3360 (0.153)	-0.7157 (0.591)	-1.7610 (0.143)	-0.4362 (0.840)	-0.7258 (0.113)	$df^m_{t-2} - df^m_{t-3}$	1.614** (0.031)	0.5906 (0.624)	-2.2488** (0.037)	-3.6356* (0.065)	-0.9720** (0.026)
	df^m_{t-3}	-1.0740 (0.264)	-0.6057 (0.675)	0.7886 (0.345)	-1.4298 (0.534)	0.7137* (0.100)	$df^m_{t-3} - df^m_{t-4}$	-1.037 (0.158)	-0.5096 (0.689)	0.9001 (0.246)	2.9091 (0.167)	0.8387** (0.033)
	C	6.4354* (0.093)	9.5205*** (0.000)	5.7224*** (0.000)	10.1442*** (0.001)	9.7139*** (0.000)	C	-8.3529*** (0.006)	0.9782 (0.546)	-0.4893 (0.664)	-2.2835 (0.388)	-1.1227 (0.573)
	R^2	0.8430	0.8250	0.8400	0.8370	0.8450	R^2	0.5700	0.3340	0.3720	0.3400	0.4130
Businesses	π_{t-1}	0.3452*** (0.000)	0.3698*** (0.000)	0.3173*** (0.000)	7.0696*** (0.000)	0.3552*** (0.000)	$\pi_{t-1} - \pi_{t-2}$	0.1434*** (0.003)	0.1344*** (0.004)	0.1467*** (0.002)	0.1326*** (0.003)	0.1400*** (0.003)
	df^m_t	-0.1827 (0.843)	1.2093 (0.245)	-0.2132 (0.768)	0.3255*** (0.000)	0.2268 (0.451)	$df^m_t - df^m_{t-1}$	1.0633 (0.122)	0.9183 (0.232)	-0.8341 (0.153)	0.9607 (0.277)	0.0557 (0.799)
	df^m_{t-1}	-2.0502* (0.073)	-0.5729 (0.672)	1.5272** (0.030)	3.3463*** (0.005)	0.2061 (0.293)	$df^m_{t-1} - df^m_{t-2}$	0.3751 (0.645)	0.0937 (0.926)	1.1851** (0.037)	-0.5685 (0.659)	0.0689 (0.631)
	df^m_{t-2}	3.1070** (0.026)	0.0296 (0.981)	-0.3270 (0.720)	-5.3502*** (0.002)	0.0973 (0.767)	$df^m_{t-2} - df^m_{t-3}$	-0.3006 (0.768)	-0.3382 (0.720)	-0.2001 (0.787)	2.7441* (0.054)	0.0326 (0.910)
	df^m_{t-3}	-0.6576 (0.541)	0.9159 (0.384)	0.3750 (0.584)	3.6920** (0.050)	-0.1574 (0.582)	$df^m_{t-3} - df^m_{t-4}$	-0.6195 (0.442)	-0.4283 (0.576)	-0.2405 (0.664)	-2.7642** (0.012)	-0.3174 (0.203)
	C	8.1511*** (0.000)	4.9233*** (0.010)	5.0653*** (0.000)	-0.1370 (0.924)	5.5163** (0.020)	C	-1.5599 (0.178)	-0.6747 (0.578)	-0.0132 (0.988)	-0.7796 (0.349)	1.1141 (0.489)
	R^2	0.6960	0.6850	0.7460	0.7360	0.6870	R^2	0.2270	0.1790	0.2190	0.2740	0.1880
Households	π_{t-1}	0.2314*** (0.000)	0.2412*** (0.000)	0.1257*** (0.000)	0.1901*** (0.000)	0.2328*** (0.000)	$\pi_{t-1} - \pi_{t-2}$	0.0403 (0.417)	0.0368 (0.440)	0.0538 (0.795)	0.0486 (0.315)	0.0363 (0.456)
	df^m_t	0.3198 (0.708)	0.7091 (0.609)	1.2605** (0.034)	2.2806 (0.146)	0.2741 (0.536)	$df^m_t - df^m_{t-1}$	0.6496 (0.106)	0.9103 (0.155)	-0.3622 (0.268)	0.9230 (0.239)	0.2290 (0.283)
	df^m_{t-1}	-0.1090 (0.905)	0.4006 (0.773)	0.5428 (0.469)	1.8611 (0.253)	0.1923 (0.691)	$df^m_{t-1} - df^m_{t-2}$	-0.5920 (0.169)	-0.5324 (0.406)	0.6644 (0.125)	-0.3849 (0.632)	0.0145 (0.951)
	df^m_{t-2}	0.9379 (0.320)	-0.4846 (0.725)	-0.1165 (0.876)	2.8517* (0.082)	0.3577 (0.463)	$df^m_{t-2} - df^m_{t-3}$	0.6681 (0.136)	-0.7833 (0.223)	-0.6464 (0.135)	1.3106 (0.106)	0.1944 (0.401)
	df^m_{t-3}	1.1311 (0.209)	1.0076 (0.461)	1.5535** (0.011)	1.7897 (0.280)	-0.0602 (0.890)	$df^m_{t-3} - df^m_{t-4}$	-0.0787 (0.861)	1.2966** (0.040)	0.3707 (0.261)	-1.3967* (0.081)	-0.3634* (0.080)

Table 2 (continued). Relationship between Monthly News Indices and Inflation Expectations

Respondents	Variables	Without transformations					Variables	1st difference				
		Inflation	Exchange rate	Utilities	Food	Fuel		Inflation	Exchange rate	Utilities	Food	Fuel
Households	C	3.3627 (0.305)	7.1208** (0.011)	4.9512*** (0.000)	2.2403 (0.289)	4.7291* (0.092)	C	-1.9764 (0.234)	-1.5633 (0.198)	-0.1267 (0.795)	-0.5184 (0.607)	-0.5716 (0.672)
	R^2	0.626	0.6100	0.7390	0.6690	0.6230	R^2	0.1070	0.1070	0.0590	0.0900	0.0680
Financial analysts	π_{t-1}	0.1624*** (0.000)	0.1704*** (0.000)	2.1542*** (0.000)	0.1344*** (0.000)	0.1667*** (0.000)	$\pi_{t-1} - \pi_{t-2}$	-0.0152 (0.696)	-0.0159 (0.668)	-0.0117 (0.752)	-0.0030 (0.940)	-0.0024 (0.953)
	df^m_t	0.1562 (0.794)	0.0426 (0.967)	0.0672*** (0.000)	2.1366* (0.086)	0.1165 (0.729)	$df^m_t - df^m_{t-1}$	0.5383* (0.100)	0.3330 (0.520)	0.3792 (0.148)	0.4323 (0.538)	0.0671 (0.708)
	df^m_{t-1}	0.9969 (0.133)	-0.1376 (0.892)	1.0333*** (0.004)	1.7494 (0.156)	-0.1388 (0.728)	$df^m_{t-1} - df^m_{t-2}$	0.6199* (0.086)	-0.8697* (0.092)	-0.4401 (0.215)	-0.3148 (0.646)	-0.0930 (0.667)
	df^m_{t-2}	0.7433 (0.257)	-0.297 (0.765)	0.0352 (0.939)	1.4050 (0.259)	0.2002 (0.628)	$df^m_{t-2} - df^m_{t-3}$	-0.8320** (0.021)	0.0794 (0.875)	0.8993*** (0.010)	-0.2915 (0.672)	0.1157 (0.595)
	df^m_{t-3}	1.2712* (0.052)	2.3240** (0.022)	0.7893* (0.077)	1.3447 (0.295)	0.1429 (0.677)	$df^m_{t-3} - df^m_{t-4}$	0.4078 (0.284)	1.5838*** (0.002)	-0.6829** (0.013)	0.6863 (0.333)	0.0082 (0.964)
	C	-2.1422 (0.355)	3.7398* (0.058)	1.2051*** (0.001)	1.1574 (0.491)	4.8133** (0.033)	C	-2.2581* (0.097)	-1.9795** (0.039)	-0.4214 (0.296)	-0.5979 (0.513)	-0.7690 (0.523)
	R^2	0.6750	0.6290	0.8380	0.6600	0.5980	R^2	0.1640	0.1750	0.1480	0.0230	0.0100

Notes: The table shows the results of OLS regressions where inflation expectations are the dependent variable. The time indicator T of document frequencies is set to t, t-1, t-2, and t-3. The first figures in the cells indicate coefficients and p-values are shown in parentheses *p<0.1; **p<0.05; ***p<0.01.

Table 3. Relationship between Decadal News Indices and Inflation Expectations

Respondents	Variables	Inflation		Exchange rate		Utilities		Food		Fuel	
		Curr.	Prev.	Curr.	Prev.	Curr.	Prev.	Curr.	Prev.	Curr.	Prev.
Banks	π_{t-1}	0.3025*** (0.000)	0.3007*** (0.000)	0.2962*** (0.000)	0.2924*** (0.000)	0.2924*** (0.000)	0.2767*** (0.000)	0.2982*** (0.000)	0.3023*** (0.000)	0.2959*** (0.000)	0.2892*** (0.000)
	df^m I	-0.2913 (0.719)	1.0018 (0.272)	-0.5795 (0.593)	-1.7788** (0.030)	0.3746 (0.652)	0.3437 (0.609)	-0.3568 (0.783)	-0.766 (0.452)	-0.5005 (0.287)	0.0193 (0.961)
	df^m II	-0.7864 (0.142)	0.8323 (0.346)	-0.3309 (0.742)	3.6288*** (0.007)	0.2311 (0.732)	0.8281 (0.291)	-1.3411 (0.334)	-0.2261 (0.849)	-0.1320 (0.648)	-0.3379 (0.225)
	df^m III	0.5547 (0.255)	-0.2232 (0.778)	0.4306 (0.680)	-2.2436** (0.012)	-0.0498 (0.932)	-0.2502 (0.738)	-0.7180 (0.637)	1.8529* (0.098)	0.1892 (0.568)	-0.0686 (0.799)
	C	8.6576*** (0.002)	2.2805 (0.494)	8.2388*** (0.000)	8.2280*** (0.000)	6.0084*** (0.000)	5.2585*** (0.000)	9.7445*** (0.000)	6.6276*** (0.000)	10.4832*** (0.000)	10.3404*** (0.000)
	R^2	0.8250	0.8250	0.8140	0.8670	0.8160	0.8280	0.8200	0.8300	0.8230	0.8290

Table 3 (continued). Relationship between Decadal News Indices and Inflation Expectations

Respondents	Variables	Inflation		Exchange rate		Utilities		Food		Fuel	
		Curr.	Prev.	Curr.	Prev.	Curr.	Prev.	Curr.	Prev.	Curr.	Prev.
Businesses	π_{t-1}	0.3429*** (0.000)	0.3468*** (0.000)	0.3645*** (0.000)	0.3590*** (0.000)	0.3285*** (0.000)	0.3165*** (0.000)	0.3300*** (0.000)	0.3378*** (0.000)	0.3517*** (0.000)	0.3527*** (0.000)
	df^m I	-0.0008 (0.999)	-0.5610 (0.409)	-0.0527 (0.940)	0.7648 (0.331)	0.9796* (0.052)	0.7777 (0.142)	1.3742 (0.104)	1.1391 (0.228)	0.7220** (0.019)	0.4026 (0.146)
	df^m II	-0.7326 (0.355)	0.0077 (0.991)	1.1940 (0.167)	-0.7624 (0.299)	-0.0830 (0.908)	0.5752 (0.206)	0.5473 (0.558)	-0.2134 (0.865)	-0.1173 (0.682)	0.0089 (0.955)
	df^m III	0.9513 (0.167)	0.1996 (0.723)	-0.0056 (0.995)	1.2836 (0.110)	0.2448 (0.725)	0.1232 (0.812)	-0.2146 (0.831)	-0.5889 (0.632)	-0.0325 (0.906)	0.0459 (0.853)
	C	8.0788*** (0.000)	9.5355*** (0.000)	5.9074*** (0.001)	5.8469*** (0.001)	5.7110*** (0.000)	4.7702*** (0.000)	6.5813*** (0.000)	8.1705*** (0.000)	3.9404* (0.051)	4.8362*** (0.006)
	R^2	0.6660	0.6590	0.6860	0.6890	0.7250	0.7480	0.6880	0.6640	0.6980	0.6900
Households	π_{t-1}	0.2332*** (0.000)	0.2338*** (0.000)	0.2273*** (0.000)	0.2341*** (0.000)	0.1569*** (0.000)	0.1519*** (0.000)	0.2257*** (0.000)	0.2177*** (0.000)	0.2307*** (0.000)	0.2337*** (0.000)
	df^m I	0.6075 (0.470)	0.3696 (0.641)	-2.1987** (0.031)	-0.6561 (0.507)	1.9786*** (0.003)	2.3171*** (0.001)	0.7161 (0.587)	1.3905 (0.280)	0.9265** (0.017)	0.6657 (0.098)
	df^m II	-0.5300 (0.317)	-0.4296 (0.421)	0.8797 (0.287)	0.6005 (0.512)	0.6423 (0.249)	0.3836 (0.485)	2.0484* (0.088)	1.5500 (0.192)	0.3043 (0.305)	0.4797* (0.083)
Households	df^m III	0.3656 (0.425)	0.3585 (0.484)	1.4200 (0.159)	0.7847 (0.463)	0.1211 (0.810)	0.2490 (0.621)	0.2029 (0.851)	0.4941 (0.682)	-0.4318 (0.120)	-0.4599 (0.136)
	C	8.7718*** (0.001)	9.1639*** (0.000)	10.3816*** (0.000)	8.9515*** (0.000)	5.5579*** (0.000)	5.2581*** (0.000)	7.0805*** (0.000)	6.7930*** (0.000)	4.4817 (0.074)	5.1578*** (0.043)
	R^2	0.6090	0.6050	0.6480	0.6140	0.7090	0.7180	0.6280	0.6270	0.6470	0.6380
Financial analysts	π_{t-1}	0.1604*** (0.000)	0.1623*** (0.000)	0.1652*** (0.000)	0.1662*** (0.000)	0.1028*** (0.000)	0.0922*** (0.000)	0.1508*** (0.000)	0.1492*** (0.000)	0.1655*** (0.000)	0.1668*** (0.000)
	df^m I	1.692*** (0.006)	1.4931** (0.011)	-1.1937 (0.134)	-0.8612 (0.263)	1.6138*** (0.001)	1.8996*** (0.000)	1.7747* (0.095)	1.4521 (0.131)	0.6686** (0.033)	0.5142* (0.100)
	df^m II	-0.1233 (0.775)	-0.0997 (0.805)	0.5855 (0.359)	0.7450 (0.299)	0.7149* (0.069)	0.5090 (0.175)	0.4292 (0.646)	0.3017 (0.733)	0.0583 (0.796)	0.1352 (0.544)
	df^m III	0.0259 (0.939)	0.4572 (0.218)	0.9114 (0.275)	0.4700 (0.568)	-0.0186 (0.958)	0.2341 (0.506)	0.7651 (0.354)	1.1162 (0.228)	-0.2818 (0.190)	-0.3239 (0.172)
	C	2.5993 (0.138)	1.7719 (0.325)	6.7942*** (0.000)	6.6295*** (0.000)	3.2151*** (0.000)	2.7488*** (0.000)	4.3551*** (0.000)	4.5878*** (0.000)	3.9331* (0.051)	4.7127** (0.023)
	R^2	0.6340	0.6420	0.6210	0.6110	0.7480	0.7690	0.6270	0.6230	0.6230	0.6130

Notes: The table shows the results of OLS regressions where inflation expectations are the dependent variable. The first figures in cells indicate coefficients and p-values are shown in parentheses *p<0.1; **p<0.05; ***p<0.01. Curr. stands for decades of reported month, Prev. – for the previous month. Indicators I, II, and III following document frequency indices denote the number of decades.

Table 4. Relationship between Quarterly News Indices and Inflation Expectations

Respondents	Variables	Without transformations					Variables	1st difference				
		Inflation	Exchange rate	Utilities	Food	Fuel		Inflation	Exchange rate	Utilities	Food	Fuel
Banks	π_{t-1}	0.2962*** (0.000)	0.2820*** (0.000)	0.2832*** (0.000)	0.3004*** (0.000)	0.2898*** (0.000)	$\pi_{t-1} - \pi_{t-2}$	0.0623 (0.154)	0.0903** (0.033)	0.1041** (0.015)	0.1014** (0.020)	0.0895** (0.033)
	df^m_t	-1.7001 (0.170)	1.1527 (0.467)	0.1101 (0.918)	1.3103 (0.571)	-0.0158 (0.973)	$df^m_t - df^m_{t-1}$	1.6473** (0.017)	1.5458* (0.094)	-0.4047 (0.516)	0.3991 (0.766)	0.2467 (0.368)
	df^m_{t-1}	0.7511 (0.669)	-1.7379 (0.440)	1.1679 (0.557)	-1.4879 (0.672)	-0.4129 (0.615)	$df^m_{t-1} - df^m_{t-2}$	-0.7856 (0.425)	-3.4065*** (0.010)	1.6354 (0.165)	-1.1473 (0.576)	-0.6507 (0.176)
	df^m_{t-2}	1.5584 (0.669)	-0.1744 (0.938)	-1.2997 (0.512)	-0.6431 (0.855)	-0.4109 (0.616)	$df^m_{t-2} - df^m_{t-3}$	0.0814 (0.933)	1.3132 (0.322)	-2.1069* (0.075)	-0.0197 (0.992)	0.1333 (0.782)
	df^m_{t-3}	-1.1049 (0.394)	-0.5921 (0.704)	0.8456 (0.421)	0.0465 (0.984)	0.3926 (0.399)	$df^m_{t-3} - df^m_{t-4}$	0.2624 (0.719)	0.6112 (0.508)	0.9212 (0.140)	0.7249 (0.582)	0.4089 (0.139)
	C	8.6963*** (0.000)	9.9577*** (0.000)	5.5146*** (0.000)	7.9478*** (0.000)	10.5852*** (0.000)	C	-3.6686*** (0.006)	-0.1601 (0.781)	-0.1348 (0.764)	0.0122 (0.987)	-1.0700 (0.174)
	R^2	0.8120	0.8220	0.8170	0.8100	0.8260	R^2	0.159	0.1140	0.0800	0.0590	0.1340
Businesses	π_{t-1}	0.3567*** (0.000)	0.3775*** (0.000)	0.3229*** (0.000)	0.3442*** (0.000)	0.3681*** (0.000)	$\pi_{t-1} - \pi_{t-2}$	0.1549*** (0.000)	0.1513*** (0.000)	0.1549*** (0.000)	0.1451*** (0.001)	0.1468*** (0.000)
	df^m_t	-1.3567 (0.305)	1.2189 (0.364)	0.6571 (0.428)	1.4312 (0.387)	0.3357 (0.245)	$df^m_t - df^m_{t-1}$	0.7995 (0.166)	0.8543 (0.162)	-0.3296 (0.416)	1.9479*** (0.009)	-0.0014 (0.991)
	df^m_{t-1}	0.2246 (0.914)	-0.788 (0.691)	0.7613 (0.596)	-1.2567 (0.665)	0.1356 (0.757)	$df^m_{t-1} - df^m_{t-2}$	-0.7318 (0.421)	-1.3392 (0.137)	1.3800* (0.051)	-3.1614** (0.015)	0.1989 (0.319)
	df^m_{t-2}	1.4987 (0.470)	-0.2494 (0.899)	-0.3459 (0.807)	-0.6994 (0.809)	-0.0457 (0.916)	$df^m_{t-2} - df^m_{t-3}$	0.8714 (0.335)	0.6937 (0.442)	-1.1423 (0.107)	1.8089 (0.159)	-0.3072 (0.120)
	df^m_{t-3}	0.0660 (0.961)	1.5895 (0.233)	0.4552 (0.576)	2.1718 (0.186)	0.1126 (0.676)	$df^m_{t-3} - df^m_{t-4}$	-0.8554 (0.144)	-0.0955 (0.875)	0.0518 (0.898)	-0.4610 (0.524)	0.0751 (0.541)
	C	7.4593*** (0.000)	4.5674*** (0.000)	4.6576*** (0.000)	6.6463*** (0.000)	4.0647*** (0.001)	C	-0.2782 (0.492)	-0.2835 (0.498)	0.0782 (0.798)	-0.2130 (0.475)	0.2446 (0.637)
	R^2	0.6760	0.6950	0.7460	0.6870	0.6960	R^2	0.1100	0.0880	0.1110	0.1110	0.0940

Notes: The table shows the results of OLS regressions where inflation expectations are the dependent variable. Time indicator T of document frequencies is set to t, t-1, t-2, and t-3 and corresponds to quarters. The first figures in cells indicate coefficients and p-values are shown in parentheses *p<0.1; **p<0.05; ***p<0.01.

Table 5. Relationship between Monthly News Indices Built by LDA and Inflation Expectations

Topic	Variables	Without transformations				Variables	1st difference			
		Banks	Businesses	Households	Financial analysts		Banks	Businesses	Households	Financial analysts
Energy	π_{t-1}	0.2707*** (0.000)	0.3437*** (0.000)	0.2051*** (0.000)	0.1626*** (0.000)	$\pi_{t-1}-\pi_{t-2}$	0.1731*** (0.003)	0.1598*** (0.001)	0.0532 (0.291)	-0.0016 (0.970)
	df^m_t	0.4646 (0.766)	-1.0941 (0.413)	-2.7212** (0.037)	-0.3453 (0.754)	$df^m_t - df^m_{t-1}$	0.7627 (0.622)	-1.8396* (0.059)	-1.1436 (0.111)	0.3425 (0.564)
	df^m_{t-1}	-1.8021 (0.296)	2.4938* (0.072)	-1.5848 (0.196)	-0.3938 (0.708)	$df^m_{t-1} - df^m_{t-2}$	-1.0257 (0.538)	1.2913 (0.197)	1.0800 (0.116)	0.2248 (0.693)
	df^m_{t-2}	-0.9249 (0.405)	0.5825 (0.657)	-1.7524 (0.151)	-0.4304 (0.677)	$df^m_{t-2} - df^m_{t-3}$	-0.2834 (0.809)	0.2077 (0.833)	0.0709 (0.916)	0.4140 (0.455)
	df^m_{t-3}	-0.5554 (0.758)	-2.0173 (0.112)	-1.8030 (0.148)	-1.5595 (0.165)	$df^m_{t-3} - df^m_{t-4}$	-0.1006 (0.956)	-0.8172 (0.374)	0.2491 (0.249)	-1.0749* (0.070)
	C	15.9023*** (0.000)	8.9195** (0.044)	32.4739*** (0.000)	14.6931*** (0.000)	C	1.9027 (0.579)	3.4848 (0.209)	-0.7805 (0.729)	0.1552 (0.936)
	R^2	0.8460	0.6850	0.7180	0.6210	R^2	0.276	0.2150	0.0720	0.0540
Exchange rate	π_{t-1}	0.2207*** (0.000)	0.2349*** (0.000)	0.0553** (0.028)	0.0443** (0.021)	$\pi_{t-1}-\pi_{t-2}$	0.1404** (0.011)	0.1296*** (0.005)	0.0376 (0.422)	-0.0148 (0.703)
	df^m_t	0.2808 (0.622)	1.9069*** (0.008)	1.2814*** (0.004)	0.7794** (0.022)	$df^m_t - df^m_{t-1}$	0.25 (0.667)	0.6776 (0.232)	0.6895** (0.020)	-0.0101 (0.968)
	df^m_{t-1}	1.0477 (0.254)	-0.6672 (0.334)	0.0629 (0.904)	0.9172** (0.032)	$df^m_{t-1} - df^m_{t-2}$	0.9799 (0.265)	-0.1744 (0.770)	-0.7571** (0.030)	0.6726** (0.031)
	df^m_{t-2}	0.7299 (0.397)	0.4674 (0.571)	0.7218 (0.165)	-0.2140 (0.608)	$df^m_{t-2} - df^m_{t-3}$	0.5198 (0.529)	0.4751 (0.509)	0.5903* (0.088)	-0.7020** (0.025)
	df^m_{t-3}	-0.4810 (0.520)	0.7308 (0.322)	1.0341** (0.022)	0.6904** (0.047)	$df^m_{t-3} - df^m_{t-4}$	-1.4949** (0.033)	-0.9657 (0.109)	-0.4956* (0.085)	0.0254 (0.917)
	C	-5.9118 (0.203)	-11.2109** (0.015)	-15.1982*** (0.000)	-10.7207*** (0.000)	C	-2.2814 (0.441)	-0.2290 (0.938)	-0.3152 (0.800)	0.0642 (0.953)
	R^2	0.8640	0.7560	0.8210	0.8090	R^2	0.397	0.2170	0.1470	0.1020

Notes: The table shows the results of OLS regressions where inflation expectations are the dependent variable. Time indicator T of document frequencies is set to t, t-1, t-2, and t-3. The first figures in cells indicate coefficients and p-values are shown in parentheses *p<0.1; **p<0.05; ***p<0.01.

Table 6. Relationship between Decade News Indices Built by LDA and Inflation Expectations

Respondents	Variables	Banks		Businesses		Households		Financial analysts	
		Curr.	Prev.	Curr.	Prev.	Curr.	Prev.	Curr.	Prev.
Energy	π_{t-1}	0.2788*** (0.000)	0.2715*** (0.000)	0.3554*** (0.000)	0.3647*** (0.000)	0.20600*** (0.000)	0.2151*** (0.000)	0.1588*** (0.000)	0.1605*** (0.000)
	df^m I	-2.2081** (0.038)	-1.2279 (0.176)	0.8904 (0.394)	-0.1089 (0.914)	-2.8889*** (0.004)	-2.5453** (0.013)	-1.3360 (0.103)	-1.3255* (0.098)
	df^m II	0.1589 (0.875)	0.6438 (0.413)	-1.9179** (0.043)	0.3250 (0.702)	-1.1732 (0.163)	-0.9812 (0.243)	-0.5387 (0.478)	0.1262 (0.858)
	df^m III	-0.1123 (0.891)	-1.8735** (0.017)	1.0673 (0.102)	1.2336 (0.136)	-1.1909* (0.078)	-0.9189 (0.197)	0.0937 (0.882)	-0.4895 (0.384)
	C	13.8197*** (0.001)	14.9455*** (0.000)	8.3880** (0.025)	4.0943 (0.210)	25.1004*** (0.000)	22.6818*** (0.000)	12.0712*** (0.000)	11.8456*** (0.000)
	R^2	0.8380	0.8650	0.6870	0.6730	0.6890	0.6650	0.6130	0.6150
Exchange rate	π_{t-1}	0.2496*** (0.000)	0.2161*** (0.000)	0.2851*** (0.000)	0.3141*** (0.000)	0.1237*** (0.000)	0.1121*** (0.000)	0.0774*** (0.000)	0.0704*** (0.000)
	df^m I	0.8894* (0.072)	0.8441** (0.049)	1.0070** (0.014)	-0.0078 (0.987)	1.8598*** (0.000)	1.9123*** (0.000)	1.5874*** (0.000)	1.2134*** (0.000)
	df^m II	0.8664** (0.038)	0.4398 (0.371)	-0.0442 (0.917)	0.3258 (0.463)	0.5687* (0.058)	0.7552** (0.012)	0.5167** (0.016)	0.5048** (0.022)
	df^m III	-0.6396 (0.110)	0.3476 (0.424)	0.7195 (0.132)	0.4304 (0.328)	-0.2831 (0.373)	-0.4095 (0.184)	-0.3247 (0.132)	0.0570 (0.824)
	C	-2.2563 (0.566)	-6.3928 (0.126)	-5.2484 (0.155)	2.4127 (0.520)	-7.5191*** (0.006)	-8.4340*** (0.003)	-7.5665*** (0.000)	-7.4620*** (0.000)
	R^2	0.8490	0.8640	0.7340	0.6820	0.7520	0.7560	0.7960	0.7720

Notes: The table shows the results of OLS regressions where inflation expectations are a dependent variable. The first figures in cells indicate coefficients and p-values are shown in parentheses *p<0.1; **p<0.05; ***p<0.01. Curr. stands for decades of reported month, Prev. – for previous month. Indicators I, II and III following document frequency indices denote number of decades.

Table 7. Relationship between Quarterly News Indices Built by LDA and Inflation Expectations

Topic	Variables	Without transformations		Variables	1st difference	
		Banks	Businesses		Banks	Businesses
Energy	π_{t-1}	0.2675*** (0.000)	0.3566*** (0.000)	$\pi_{t-1} - \pi_{t-2}$	0.1846*** (0.001)	0.1466*** (0.002)
	df^m_t	-1.3148 (0.744)	1.1802 (0.661)	$df^m_t - df^m_{t-1}$	0.3972 (0.914)	-2.3983 (0.206)
	df^m_{t-1}	-2.4590 (0.609)	1.6579 (0.695)	$df^m_{t-1} - df^m_{t-2}$	-5.1964 (0.261)	4.6212 (0.137)
	df^m_{t-2}	1.2399 (0.760)	-2.7387 (0.503)	$df^m_{t-2} - df^m_{t-3}$	2.3077 (0.550)	-3.1690 (0.275)
	df^m_{t-3}	-0.5603 (0.775)	0.4886 (0.857)	$df^m_{t-3} - df^m_{t-4}$	2.1470 (0.248)	-0.0908 (0.962)
	C	16.6815*** (0.000)	6.7721 (0.148)	C	0.9729 (0.758)	3.1107 (0.271)
	R^2	0.8440	0.6620	R^2	0.3410	0.1980
Exchange rate	π_{t-1}	0.2194*** (0.000)	0.2357*** (0.000)	$\pi_{t-1} - \pi_{t-2}$	-0.7182** (0.021)	0.1224*** (0.008)
	df^m_t	1.8980 (0.214)	2.8326* (0.059)	$df^m_t - df^m_{t-1}$	1.8418 (0.207)	2.6938 (0.022)
	df^m_{t-1}	-0.9758 (0.752)	-2.1270 (0.399)	$df^m_{t-1} - df^m_{t-2}$	-0.0609 (0.983)	-3.3550* (0.098)
	df^m_{t-2}	1.3292 (0.665)	0.1740 (0.941)	$df^m_{t-2} - df^m_{t-3}$	0.1379 (0.961)	1.8010 (0.341)
	df^m_{t-3}	-0.6878 (0.639)	1.4948 (0.375)	$df^m_{t-3} - df^m_{t-4}$	-1.8472 (0.156)	-1.1601 (0.388)
	C	-5.8439 (0.247)	-10.7359** (0.025)	C	0.1227 (0.807)	0.0370 (0.990)
	R^2	0.8590	0.7430	R^2	0.4320	0.2410

Notes: The table shows results of OLS regressions where inflation expectations are the dependent variable. Time indicator T of document frequencies is set to t, t-1, t-2 and t-3 and corresponds to quarters. The first figures in cells indicate coefficients and p-values are shown in parentheses *p<0.1; **p<0.05; ***p<0.01.

APPENDIX B. FIGURES

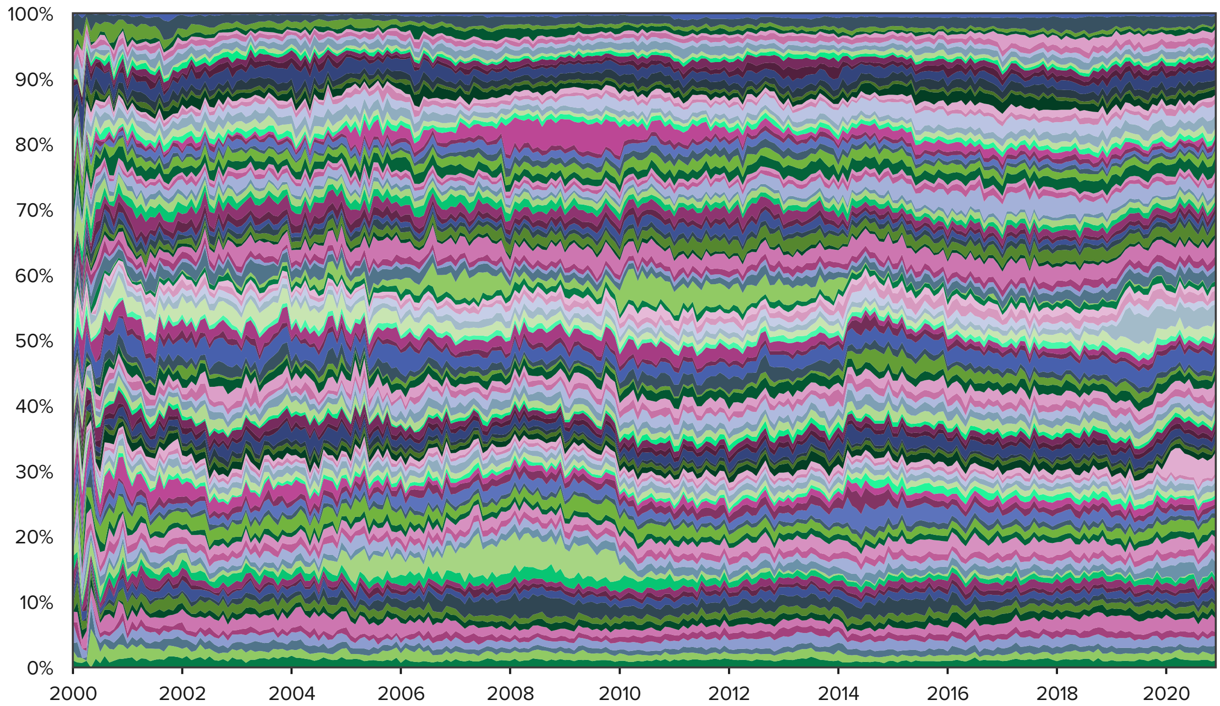


Figure 6. Distribution of Topics Received with LDA

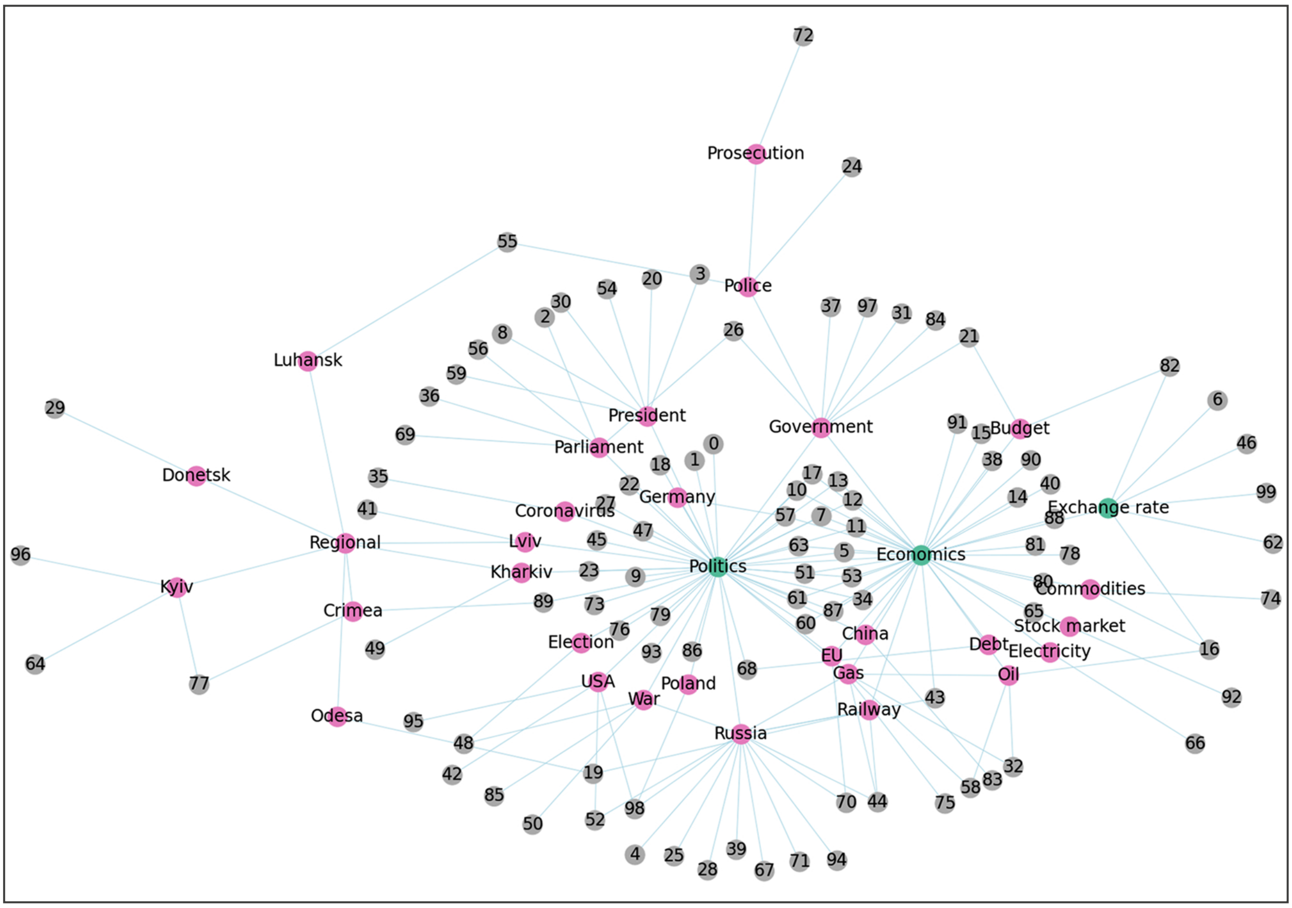


Figure 7. Graph of Topics in News Corpus. Grey Circles Refer to Topics Defined by LDA (pink circles - manually labeled clusters, and green circles – general topics)

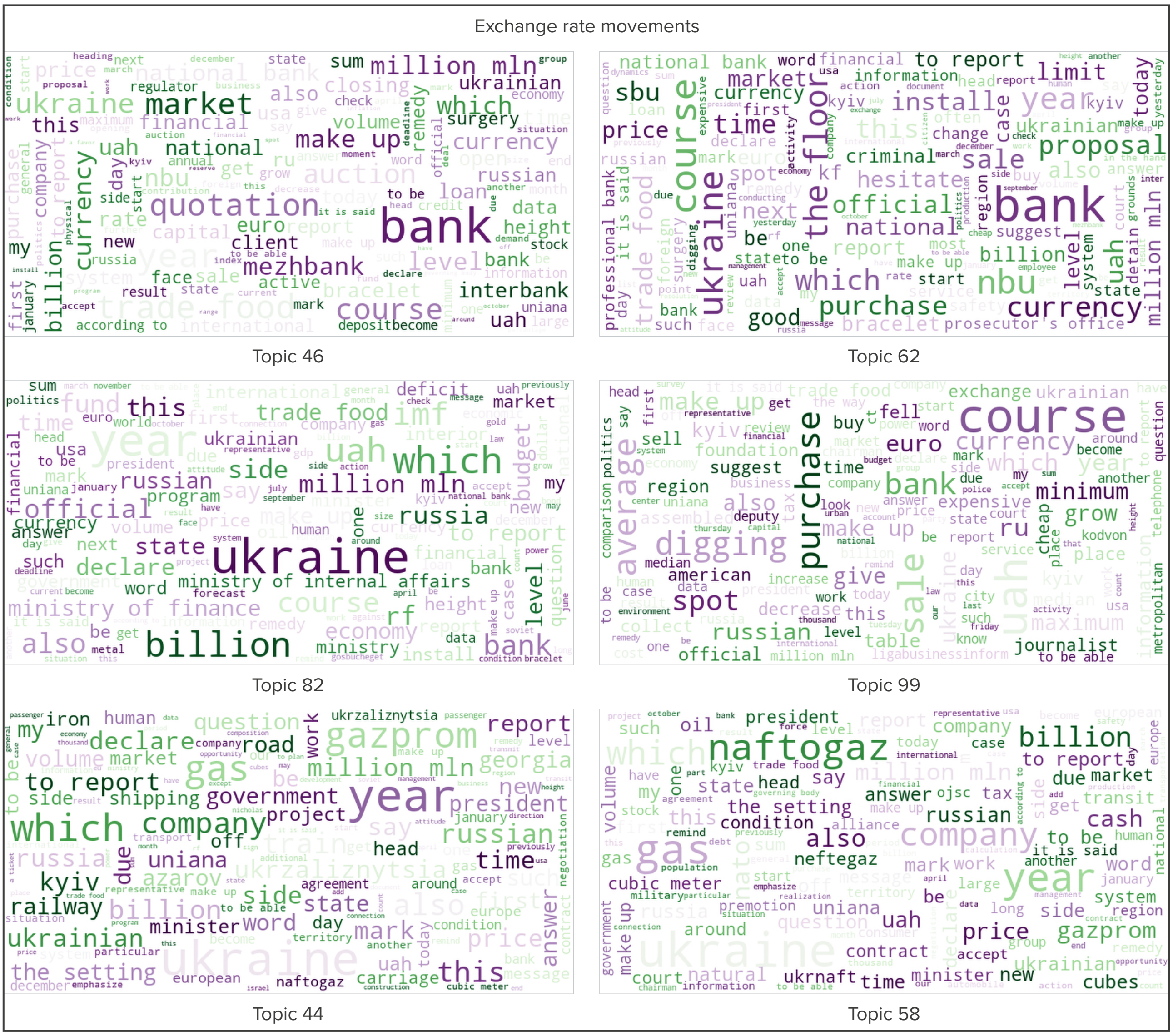


Figure 8. Wordclouds for Selected Topics

APPENDIX C. NEWS CORPUS

Originally, the news corpus consisted of 2,030,000 unique articles. However, after cleaning and filtering out items with various types of errors (parsing errors when web-page tags are wrongly placed, empty pages, corrupted symbols etc.), the number of articles decreased by 50,000 items. As this was only 2.5% of the total number of articles, I consider such a reduction quite acceptable, and that it will not affect the overall result.

Table 8. Article Size in News Corpus (after cleaning)

	finance.ua	liga	ukr_pravda	unian	Total
count	389,951	620,655	339,275	634,832	1,985,143
mean	120.6	121.5	131.0	151.5	132.5
std	96.8	79.6	94.2	123.9	102.1
min	0	2	4	3	0
25%	63	76	83	88	78
50%	99	108	115	127	113
75%	151	149	157	182	162
max	3,832	11,540	5,842	3,986	11,540
skewness	3.305	24.465	9.911	7.766	10.072
kurtosis	32.299	2,456.736	228.865	109.719	394.412

The articles in the corpus differ not only in content but also in writing style, size as measured by word count, and other features (Table 8). Expectedly, different sources of information have some dissimilarities in how the news is written, which is, for example, revealed in the article size. Unian has the largest articles on average, while finance.ua has the smallest articles. At the same time, the sizes of the articles from all sources are very close.

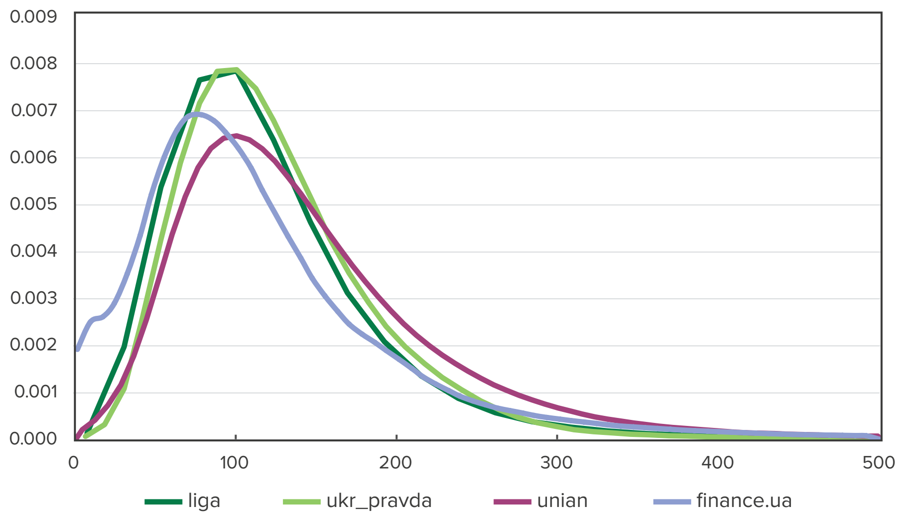


Figure 9. Distribution of Article Size by Sources

The distribution of article sizes for all news sources (figure 9) is highly asymmetrical. All values of the skewness are positive, and the tail of the distribution is longer towards the right-hand side of the curve. Articles from Liga are the most skewed. At the same time, distributions of article length are leptokurtic, which means they are tall and thin, and so near the mean. For example, the number of articles with a length of more than 500 words is less than 15,000, which is only 0.7% of the corpus. The number of articles with extremely small size is also negligible (around 2.5% of the total number).[1]

[1] The average sentence ranges from 15 to 20 words, I consider that the smallest article consists of a topic and one sentence = 30 words.

APPENDIX D. INFLATION EXPECTATIONS SURVEY DESIGN

In contrast to financial analysts, who are asked to answer open questions, banks, businesses, and households are asked to pick from a set of inflation intervals, for example:

“Inflation over the next twelve months will be:

a) less than zero (“prices will fall”),
b) between 0 and X percent,
c) between X and 2X percent,
d) between 2X and 3X percent,
e) between 3X and 4X percent,
f) over 4X percent.

In this example, inflation expectations would be computed by the formula:

$$E\pi = w_a \cdot \left(-\frac{X}{2}\right) + w_b \cdot \frac{X}{2} + w_c \cdot \frac{X+2X}{2} + w_d \cdot \frac{2X+3X}{2} + w_e \cdot \frac{3X+4X}{2} + w_f \cdot \left(4X + \frac{X}{2}\right), \quad (6)$$

where w is the share of respondents who pick the respective interval. Size of X as well as number of intervals is not fixed and changes over time to match the normal distribution of answers. Thus, in 2015 inflation in Ukraine accelerated drastically, so the maximum bracket was expanded to 50% and respondents selected from 12 intervals. Following disinflation in 2020 maximum bracket was decreased to 10% and number of intervals was cut to eight.

Since January 2018, the surveys of households have also included a question about inflation perceptions. Once a year, consumers are asked to answer one open question about perceived inflation over the previous 12 months. Additionally, households are also asked to pick answers from an interval question once a quarter. The construction of this question is similar to the inflation expectations question.

A HEATMAP FOR MONITORING SYSTEMIC FINANCIAL STABILITY RISKS IN UKRAINE[1]

ADAM GERŠL[a], PERVIN DADASHOVA[bc], YULIYA BAZHENOVA[b], VLADYSLAV FILATOV[c], ANATOLII HLAZUNOV[bc], ROMAN SOLTYSIAK[b]

[a]Charles University
E-mail: adam.gersl@gmail.com
[b]National Bank of Ukraine
E-mail: pervin.dadashova@bank.gov.ua
yuliia.bazhenova@bank.gov.ua
anatolii.hlazunov@bank.gov.ua
roman.soltysiak@bank.gov.ua
[c]National University of Kyiv-Mohyla Academy
E-mail: vladyslav.filatov@ukma.edu.ua

Abstract This study presents an updated risk map of the Ukrainian financial sector – an analytical tool for identifying and monitoring the buildup and materialization of systemic risks. The risk map methodology that the National Bank of Ukraine used until 2021 has been revised to ensure that risk assessment is based on reliable quantitative indicators rather than expert judgements, as well as to extend the list of risks considered. The instrument allows the stability of the financial system to be assessed across key risks, such as macroeconomic risk, the credit risks of households and non-financial corporations, capital adequacy risk, profitability risk, liquidity risk, and foreign exchange risk. We introduce indicators that capture a wide range of economic and financial vulnerabilities and group them by risks. Each risk category contains from four to seven indicators that combine both actual data and expectations. Statistical checks show that the indicators clearly signal previous crisis episodes, as well as the buildup of vulnerabilities during the research period. We find that macroeconomic risk and foreign exchange risk have the best explanatory and predictive power, while the weaker performance of other risks could result from structural changes in the banking sector over the past decades that have affected the overall risk profile of the financial sector.

JEL Codes G01, G10, G18, G21, G28

Keywords risk map, systemic risks, macroprudential policy, financial stability

1. INTRODUCTION

One of the fundamental goals of most central banks is to promote financial stability, which is a prerequisite for sustainable economic growth. To achieve this goal, they implement policies to prevent the buildup and materialization of systemic risks in order to reduce the probability and severity of crises, and to strengthen the resilience of the financial sector.

In Ukraine, the task of maintaining financial stability is especially relevant – over the past 30 years the country has experienced five deep crises. While a number of risks accumulated at the macroeconomic level, the severity and depth of Ukraine's systemic crises were exacerbated by the financial sector. Therefore, an appropriate risk assessment should be based on the analysis of the development of both the macroeconomic environment and the financial system.

As a macroprudential authority in Ukraine, the National Bank of Ukraine (NBU) promotes financial stability, including the stability of the banking system, provided that this does not impede the achievement of price stability. Its powers include the identification and monitoring of the buildup of systemic

[1] The authors thank the Bilateral Assistance and Capacity Building for Central Banks (BCC) programme for supporting this research. The BCC programme is funded by Swiss State Secretariat for Economic Affairs (SECO) and implemented by The Graduate Institute, Geneva. The views expressed in this paper are solely those of the authors and do not necessarily reflect those of the National Bank of Ukraine.

 Available at https://doi.org/10.26531/vnbu2022.253.02

risks, and the selection and introduction of macroprudential regulatory measures if the situation requires it.

The choice of macroprudential policy instruments depends on the type of risk that arises or is expected to arise at a particular moment. The NBU has a wide range of tools for monitoring the risks affecting financial stability. In 2016, the NBU developed a *risk map* of the banking sector, which captures such risk categories as credit risk, capital adequacy risk, liquidity risk, profitability risk, foreign exchange risk, and legal risk. The assessment of the risk level for each category was to a large extent based on expert judgements of NBU staff, which could lead to biased conclusions. Recently, we revised the risk map methodology to ensure that risk assessment is based on reliable quantitative indicators rather than personal views, as well as to extend the types of risks captured.

In this study, we present an updated risk map for the Ukrainian financial sector as an analytical tool for identifying and monitoring the buildup and materialization of systemic risks, and as a communication tool to raise stakeholder awareness of financial stability risks. The risk map allows for an assessment of financial system resilience across seven key risks, namely macroeconomic risk, credit risk of households, credit risk of non-financial corporations, capital adequacy risk, profitability risk, liquidity risk, and foreign exchange risk. We identified indicators in each risk category that reflect a wide range of economic and financial vulnerabilities. The selection of indicators is based on their ability to signal an accumulation and materialization of risks, as well as the availability of historical data and their comparability with data from other countries. The indicators were aggregated by simple averaging within each risk category. Finally, the obtained risk assessments were tested for the ability to predict crises.

According to the results, the aggregate risk level can explain and predict crises well. Macroeconomic risk and foreign exchange risk estimates have the better explanatory power compared to other risks. The weaker performance of other risk categories could be a result of structural changes in the banking sector over the past decades, which have affected the overall risk profile of the financial sector, and the limited availability of data for certain periods.

The paper is organized as follows. Section 2 describes the related literature. The methodology framework is presented in Section 3. Section 4 specifies data and indicators. The results of the paper are discussed in Section 5. Section 6 provides conclusions.

2. RELATED LITERATURE

This study builds on an extensive literature that seeks to find empirical evidence for the ability of a macroprudential toolkit to predict the probability of the occurrence of financial crises, and to assess their severity.

The early literature, motivated by emerging market crises in the 1990s, found that international reserves, domestic credit growth, real exchange rate volatility (Kaminsky et al, 1998; Kaminsky and Reinhart, 1999; Demirgüç-Kunt and Detragiache, 1999), and domestic inflation (Demirgüç-Kunt and Detragiache, 1998; Kaminsky et al., 1998) are good predictors of banking and currency crises.

Excessive growth in credit and asset prices have been identified in numerous studies as leading indicators of financial crises (Borio and Lowe, 2002; Mendoza and Terrones, 2008; Schularick and Taylor, 2012; IMF, 2011; Mitra et al., 2011; Dell'Ariccia et al., 2012; Arena et al., 2015).

Dell'Ariccia et al. (2012) identified factors frequently associated with the onset of credit booms: financial sector reforms, surges in foreign capital inflows, often in the aftermath of capital account liberalization. They also pointed out that credit booms generally start during or after a period of buoyant economic growth.

Mendoza and Terrones (2008) found major differences in credit booms in the industrial and emerging economies: (a) credit booms, and the macro and micro fluctuations associated with them, are larger in emerging economies; (b) not all credit booms end in crisis, but many of the emerging markets crises were associated with credit booms; and (c) credit booms in emerging economies are often preceded by large capital inflows and not by domestic financial reforms or productivity gains, while credit booms in industrial countries tend to be preceded by financial reforms or gains in total factor productivity.

Drehmann et al. (2010) and Drehmann et al. (2011) proved the importance of the credit-to-GDP gap as a leading indicator for predicting the expansion phase of the credit cycle, as well an as anchor for the countercyclical capital buffer setup. In response to the critics of the credit-to-GDP gap's relevance for emerging markets and transition economies (World Bank, 2010; Geršl and Seidler, 2015; RBI, 2013), Drehmann and Tsatsaronis (2014) emphasized the need to rely on a wide range of indicators rather than solely on the mechanical use of the credit-to-GDP gap.

This study contributes to the existing literature on financial stability risk measures. These metrics are commonly based on a set of indicators, which are aggregated into composite measures and visualized via heatmaps, risk dashboards, spider, radar, coxcomb or sun-burst charts etc. They either provide an assessment of risk evolution over time or a snapshot of risk at a given point in time. Some such tools for monitoring financial stability risks across countries are summarized in Table 2 (Appendix A).

Risk maps usually comprise indicators that characterize credit growth and debt burden in the non-financial private sector, current lending standards, banking sector leverage, liquidity and profitability, real estate price growth, macroeconomic imbalances, and financial market trends. Non-bank financial segments are also often captured. These indicators are typically grouped into different categories, which can be defined by intermediate macroprudential policy objectives according to the European Systemic Risk Board (Mencía and Saurina, 2016; NBB, 2019; Central Bank of Ireland, 2020), sectors of the economy (Aikman et al., 2018; IMF, 2019), or risks (Arbatli and Johansen, 2017; Lepers and Sánchez Serrano, 2017; Latvijas Banka, 2018; Venditti et al., 2018; EBA, 2020).[2]

Different techniques can be applied to aggregate risk assessments of indicators into groups or general risk level. This is often done linearly, by taking a simple or weighted

[2] According to ESRB (2013) the intermediate objectives of macroprudential policy should be to: (a) mitigate and prevent excessive credit growth and leverage, (b) mitigate and prevent excessive maturity mismatch and market liquidity, (c) limit direct and indirect exposure concentration, (d) limit the systemic impact of misaligned incentives with a view to reducing moral hazard, and (e) strengthen the resilience of financial infrastructure.

average of the standardized (or not) indicators within categories (Venditti et al., 2018; IMF, 2019; NBB, 2019; EBA, 2020). Commonly, the weights of indicators depend on their ability to predict a future crisis – indicators with better predictive power have higher weights. Mencía and Saurina (2016) also set weights depending on the correlation between indicators so as to avoid multiple counting of sources of the same risk – the lower the correlation, the higher the weight of the indicator. Some of the risk maps do not contain aggregate measures, such as those of Latvijas Banka (2018) and the Central Bank of Ireland (2020).

The setting of thresholds that determine the assignment of risk levels is another important aspect of risk map analysis. Typically, thresholds are set according to the national or cross-country historical distributions of the indicators (Mencía and Saurina, 2016; Aikman et al., 2018; IMF, 2019; EBA, 2020). Other approaches use early warning models, levels prescribed by legislation, guidelines or regulations, and expert judgments (Latvijas Banka, 2018; Venditti et al., 2018; NBB, 2019; Central Bank of Ireland, 2020).

In this study, we used the above-mentioned experience of other central banks and regulators to select indicators that can signal an incipient crisis, set thresholds for risk levels, and aggregate risk assessments, adjusting and supplementing them with information specific to Ukraine.

3. METHODOLOGY

When refining the risk map, we proceeded from the fact that the methodology should be straightforward and clear, so as to be easily interpreted by all stakeholders, such as policymakers, experts, media, and financial market participants. In the following, we describe the applied framework in more details.

The new risk map reflects risk assessments for the next 12 months based on quarterly data, as most macroeconomic and non-financial sector statistics are not available on a more frequent basis. Some of the indicators in the risk map show current distress, while some are able to provide an early signal of risk accumulation up to a year ahead.

3.1. Risk Categories

The set of risks was determined on the basis of the experience of other central banks, the significance of these risks for the financial system, and the impact of their materialization during previous crises. Since the Ukrainian financial sector is bank-centric and only banks bear systemic risks, the map is focused on risks to the banking sector.

We included the following categories in the map: macroeconomic risk, credit risk of households, credit risk of non-financial corporations, bank capital adequacy risk, bank profitability risk, bank liquidity risk, and foreign exchange risk.

We separated the credit risk of households and non-financial corporations, as these segments have different levels of indebtedness, loan quality, and sensitivity to crises. We also added a macroeconomic risk as a source of imbalances at the aggregate level. Even if the banking sector is healthy and resilient, risks can spill over into the financial system from the macroeconomic environment.

Risk assessments are presented in the heatmap both by risk categories and by indicators included in them, since proper macroprudential policy response requires clear understanding of the sources of risks. The overall risk level in the financial system is also calculated.

3.2. Selection of Risk Indicators

Each risk in the heatmap is measured by a set of indicators selected according to the following principles:

- There should not be too many indicators, while signals within risk categories should be effectively diversified.
- Indicators should be available at least on a quarterly basis and based on reliable statistics for a long enough time horizon.
- Indicators that can signal the accumulation and materialization of risks in advance should be included to ensure the forward-looking properties of the risk map.
- Risk indicators should be easy to interpret. We did not consider indicators with non-linear behavior relative to the level of risk.
- Highly correlated indicators should not be included, with the exception of indicators which clearly reflect different aspects of risk over the long term, even if they are correlated over a short horizon.

To start with, we compiled a list of indicators commonly used in risk dashboards and heatmaps by central banks, regulators and international financial organizations. These are primarily indicators of credit risk, bank solvency, profitability, and liquidity, which were supplemented by indicators used by the NBU to analyze the financial sector, and data from banking and economic activity surveys. We also added some macroeconomic and foreign exchange risk indicators that are of particular importance for Ukraine. For instance, indicators characterizing the foreign exchange rate dynamics were included, as FX rate volatility has a substantial effect on economic activity, inflation, the finances of households, and the corporate and public sectors.

As the next step, we excluded indicators related to areas that do not carry systemic risks for the Ukrainian financial sector. For example, non-banks currently do not bear systemic risks due to the small size of the sector, low interconnection with each other and with the banks, and their limited role in financial intermediation (NBU, 2020).[3] Given the weak development of the financial markets and financial instruments (derivatives, corporate shares and bonds, etc.) in Ukraine, the corresponding indicators were also discarded. Neither did we look at real estate market indicators, as mortgages are now at a low level, and the influence of banks on this sector is almost negligible. Nevertheless, the NBU constantly monitors and analyzes them, and also includes them in other analytical tools (for example, the Financial Cycle Index). Some indicators were withdrawn because of poor data quality or inconsistency. The final heatmap is to be used as a communication tool in the Financial Stability Report, showing the level of risks since 2015. Thus, we excluded indicators that contain missing data after Q1 2015, as well as those for which the calculation methodology has been fundamentally changed since then. Exceptions were made for the amount of overdue loans and the liquidity coverage ratio (LCR) as respective regulatory requirements emerged later.

At the next stage, we performed a visual and basic statistical analysis of the behavior of the indicators before

[3] Part 4. Non-Banking Sector Conditions and Risks.

and during crises. The crisis periods in Ukraine were set according to Filatov (2021). Some indicators signal the accumulation of risks in advance of the crisis, others – start signaling immediately the crisis occurs. We can use both to account for early warning signals and actual adverse events. At the same time, we omitted indicators that did not show any reaction before or during the crisis.

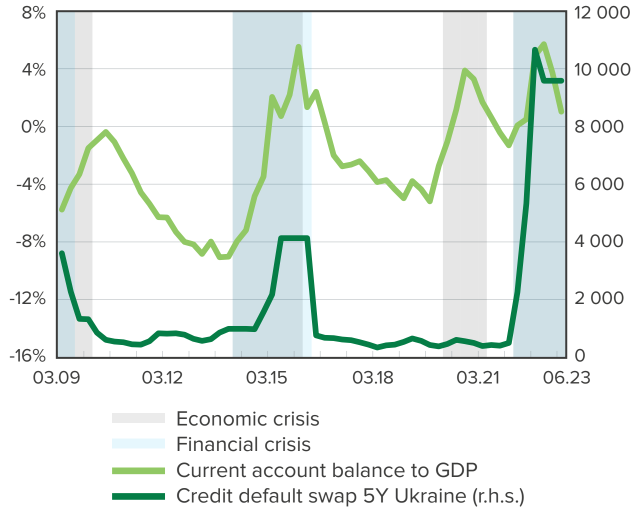


Figure 1. Dynamics while Current Account Balance to GDP Ratio and CDS 5Y Ukraine

For instance, both the current account balance to gross domestic product (GDP) ratio and the credit default swap (CDS) on 5-year Ukrainian sovereign debt reacted to previous crises (see Figure 1). The former indicator decreases robustly before the crises, then surges during the crisis and again declines after. Credit default swaps have spiked during all crises except the coronavirus crisis without prior reaction, thus we accept this indicator as coincident. Both can signal higher level of risks in the system either before or during the crisis.

Lastly, a correlation test was performed. Only one among highly correlated indicators within each of the risk categories was used. All of the others were omitted.

Ultimately, the final set of indicators encompasses 40 indicators grouped into seven risk categories. The number of indicators in each risk category varies from four to seven. A detailed description of the indicators is presented in Section 4 and Table 3 (Appendix A).

3.3. Color-Coding Scale and Threshold Selection

We retain the 1 to 10 color-coded scale used in the previous version of the risk map, where 1 is the lowest risk level, and 10 is the highest (see Figure 2). Thus, we had to set nine thresholds separating 10 intervals for each indicator to be able to assign a risk level for each observed value.

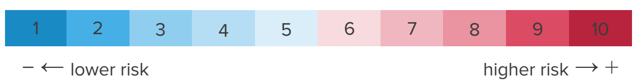


Figure 2. Color Bar Indicating Risk Score of Indicators

First, as a starting point, we created thresholds by dividing Ukraine's quarterly data from 2000, or since data became available, into deciles. Depending on the direction of the indicator – whether the higher values indicate higher or lower risk – we arranged the values either in descending or ascending order. Due to the short time series and several structural breaks in the data, we were unable to do this for every indicator. In such cases, we assumed that the values of the indicator should be more or less evenly distributed between its possible maximum and minimum. Thus, the historical data series were organized into 10 equally sized groups associated with the respective threshold and risk levels.

Setting the thresholds based on historical distribution or equally-sized intervals between potential minimum and maximum has advantages (high risk scores would reflect indicator values that are "historically high"), but it could also lead to a biased assessment if the time series is short and the observed values so far do not properly reflect the potential distribution of the indicator. In addition, for indicators where we created equally-sized intervals, the risk as captured by the indicator may change nonlinearly.

Second, we applied the decile-based method using data for other countries and analyzed their distribution.[4] This international dataset, which covers a large set of emerging markets, is available for a longer period and, at the same time, is more balanced, especially in terms of "good and bad times". We employed the same methodology to these data (percentile/decile distribution) and obtained another set of thresholds.

Finally, after analyzing the adequacy of the thresholds calculated for the Ukrainian data and for the relevant peer countries, we made final adjustments using expert judgments.

As an example of this three-step approach, we present here the calibration of the thresholds for the real GDP growth forecast (see Table 1). The NBU's real GDP growth forecast has been publicly available only since 2015, meaning it does not provide a data series long enough for there to be consistent thresholds. Hence, the historical distribution of data from peer countries is an important reference here.[5] We estimated thresholds based on both datasets separately. Some tail values of the peer countries data distribution were omitted as outliers. Then, we applied expert judgments to these estimates. For the higher risk intervals (8–10), we used an average between the Ukrainian and the peer countries' distribution threshold estimates. We adjusted the threshold for the 10th interval upward, so even a slight forecasted GDP decline is considered as high risk, as it usually is. For middle-risk intervals (4–7), we used the larger value of the two estimates. Usually, it leads to the selection of peer countries' values, as Ukrainian forecasts are highly concentrated closely to 2.5%, which is low, based on both comparative analyses and the expected potential GDP growth for Ukraine.[6] For lower risk intervals, we moved back to averaging. Additionally, for the lowest risk interval, we significantly decreased the value of the thresholds, considering that the probability of two-digit growth is relatively low in the observable future for Ukraine. It is also important that final threshold values were rounded to make the heatmap easily interpretable and more comprehensive.

[4] As a peer countries dataset, we used statistics from the emerging economies, Ukraine's trade partners, and economies with similar structures. It includes data from Albania, Armenia, Bolivia, Bulgaria, Chile, Columbia, the Czech Republic, Estonia, Georgia, Hungary, Latvia, Lithuania, Moldova, Poland, Romania, Slovakia, and Turkey.
[5] Database of IMF WEO forecasts across 1990-2020 for emerging markets
[6] According to Grui and Vdovychenko (2019), potential GDP growth in the steady state was calibrated at the level of 4%.

Table 1. Calibration of Thresholds for the Real GDP Growth Forecast

Risk score	Ukrainian distribution threshold estimate (decile value)	Peer countries' distribution threshold estimate (decile value)	Final thresholds (expertly corrected)		Expert correction explanation
			Lower (including)	Upper (excluding)	
10	–	–	–	-2.0%	Average of Ukrainian and peer countries' values*
9	-0.9%	-5.8%	-2.0%	0.0%	
8	1.9%	-1.5%	0.0%	1.0%	
7	2.2%	0.3%	1.0%	2.0%	Higher of Ukrainian or peer countries' values
6	2.5%	2.1%	2.0%	3.0%	
5	2.7%	3.6%	3.0%	4.0%	
4	2.9%	5.0%	4.0%	5.0%	
3	2.9%	6.5%	5.0%	6.0%	Average of lower threshold and peer countries' value**
2	3.1%	8.1%	6.0%	7.0%	
1	3.5%	15.1%	7.0%	–	

* – 10th interval threshold additionally adjusted 1 p.p. upwards; ** – 1st interval threshold additionally adjusted 4 p.p. downwards.

Having decided on the thresholds, we assigned risk levels that correspond to the indicator values in each period of time. Comparing the actual value of an indicator at each time with the thresholds led to a unique assignment of the corresponding risk score, ranging from 1 to 10.

3.4. Risk Level

Further, we determined a risk level for each risk group by simple averaging. Using a simple average for aggregation is a standard approach for a number of heatmaps (Venditti et al., 2018; IMF, 2019; NBB, 2019; EBA, 2020). This method is straightforward for interpretation and analysis, which is an essential feature for a communicational and policy tool. More sophisticated methodologies, such as a principal component analysis, cannot be applied here due to short data series, different lengths of the time series across the indicators, and difficulties in interpretation.

Finally, we calculated a simple average across all risk categories to arrive at a single aggregated risk score.

4. RISK INDICATORS

In this section, we describe in detail the indicators of each risk category.

The ***macroeconomic risk*** category encompasses macroeconomic variables to monitor risks stemming from the real economy, and the fiscal and external sectors. Key financial risks tend to raise during economic downswings, when it is more difficult for economic agents to service their debts, whereas investors demand higher returns on capital and look for instruments with low risk and high liquidity.

We considered the real GDP growth rate as a general measure of economic activity, low values of which indicate poor performance by the economy and a potential subsequent increase in risks to the financial sector. As an early warning indicator of a downturn, we looked at the NBU's real GDP growth forecast.

Fiscal sector vulnerabilities such as high public debt and budget deficits are of particular concern when assessing systemic risks. Excessive gross external and state debt carries liquidity and solvency risks, which can lead to the crowding out of private investments, an increase in the tax burden, and so on. Market participants' perception of the government's financial position is reflected in the required rate of return on government debt and the level of credit default swaps on sovereign bonds. Thus, higher required returns worsen the conditions for public and private borrowing. In addition, the transmission of fiscal risks to the financial sector is exacerbated by the banks' significant exposures to the government. To monitor these fiscal sector vulnerabilities, we included the ratios of the state and state guaranteed debt, gross external debt, and state budget balance to GDP, as well as the CDS rate.

To track external imbalances, we examined the ratio of the current account balance to GDP. An excessive current account deficit is a signal of an imbalance in foreign trade and greater dependence on financial inflows, which can cause economic vulnerabilities and even a currency crisis.

Credit risk is the risk of credit loss by a bank due to the inability or unwillingness of borrowers to repay their loans. The nature of lending to households and non-financial corporations is different, so we considered their credit risk separately.

The ***credit risk of households*** is higher when the debt burden becomes higher. Thus, the first indicator to be included is the ratio of gross retail bank loans to GDP. Simultaneously, even if the relative debt burden, as measured by the loans stock to GDP ratio, is low, high loan servicing costs can lead to a deterioration in payment discipline, especially during periods of economic downturn. This is particularly relevant for Ukraine, as expensive short-term consumer loans currently account for nearly 85% of total household debt. From this perspective, the debt service-to-income ratio (DSTI) at the aggregate level was incorporated. To capture a forward-looking view from the lender's side, we added an indicator of banks' expectations regarding the quality of the loan portfolio taken from the NBU's Bank Lending Survey. When filling out the questionnaire, banks take into account the available microdata on borrowers' current and projected indebtedness and solvency. As another indicator of debt-servicing problems, we included an index of economic expectations of households derived from a third-party survey, which covers both changes in personal financial standing and macroeconomic developments. Worsening expectations could have a negative impact on the payment discipline of borrowers even before their solvency is undermined.

The ***credit risk of corporates*** depends on the indebtedness of the borrowers and their financial condition. As the debt burden indicator, the ratio of net bank corporate loans to GDP was employed. We also looked at the ability of borrowers to service their debts, which was proxied by the ratio of total corporate debt to earnings before interests and taxes (EBIT) and the interest expenses coverage ratio. To characterize borrowers' financial performance, we included the return on equity of non-financial corporations: companies with low profitability or losses are considered to be more risky. On the other hand, we monitor the quality of the banks' loan portfolio, as represented by the frequency of defaults. Even a moderate increase in this indicator signals a higher credit risk. Similarly to households, we incorporated the banks' expectations of the credit risk level of non-financial corporations from the lending survey. We also added a business outlook index from another NBU survey, which is an aggregate indicator of the expected development of enterprises over the next 12 months. A deterioration in business expectations, among other things, may precede a future slowdown in economic activity, lower demand for corporate loans and an increase in credit risk.

To capture the ***capital adequacy risk*** of the banking sector, we consider indicators that assess the sufficiency of banks' capital to absorb risks. A higher level of capital ensures the banks are able to absorb unexpected losses resulting from economic shocks, meet their obligations, and remain solvent. We included here both core and total regulatory capital ratios, as they complement each other. To capture risks for capital that may arise from high level of non-performing loans (NPLs), we used the ratio of non-performing loans net of provisions to capital. Credit risk for these loans has already materialized, but they can still have a negative impact on capital.

In Ukraine, the capital adequacy requirements currently fully cover only credit, foreign exchange and partially operational risks. Therefore, we took additionally into account the ratio of capital to total net assets – leverage. This indicator covers other risks, in particular market risk, such as the risk arising from investing in government securities. The growing leverage may signal an increase in risk appetite and a possible lack of capital to cover other risks that are not fully reflected in capital adequacy ratios.

We assess the ***profitability risk*** using the banks' return on assets, return on equity, net interest margin, cost of risk, and cost to income ratio. All of them reflect the ability of the banks to generate net profit, which is an internal source of capital. Loss-making banks or those with deteriorating indicators typically face higher funding costs, limited ability to grow, and a larger probability of a capital shortfall. Return on equity (ROE) measures the return a bank earns on its equity. Return on assets (ROA) shows how efficiently a bank uses assets to make a profit. Both of these indicators were included, because ROA can signal risks in the case of possible ROE distortions caused by capital distributions, rather than higher profitability. Net interest margin shows the ability of banks to earn income from their core operations. Higher values of these ratios indicate a lower risk. The other two indicators in this group have opposite dynamics – higher values indicate a higher risk. These are the cost of risk (measured as annual provisions for expected losses per unit of bank loans) and cost-to-income ratio (total operating costs divided by total operating income). An increase in the cost of risk or cost-to-income ratio reveals threats to profitability that come from the worsening of loan quality or excessive operational expenses.

Liquidity risk indicators demonstrate the ability of banks to meet their liabilities to depositors and creditors in full and in a timely manner. It includes the liquidity coverage ratio (LCR), which is defined as the ratio of available high-quality liquid assets (HQLA) to net cash outflow expected over a 30-day horizon under adverse conditions. LCR is a relatively new ratio, which was introduced in Ukraine in 2018. To complement the LCR retrospectively, we have included another indicator – share of HQLA in total assets. Its dynamics are similar to that of LCR, but data for it is available for a longer period. We also look at the loan-to-deposit ratio as an indicator of liquidity risk. The logic behind this indicator is as follows: a low value of the ratio signals the availability of free funds, and, consequently, high liquidity. On the contrary, a high loan-to-deposit ratio reflects a greater need to raise funds from the wholesale markets, and thus higher funding and liquidity risks. To add forward-looking component, we include banks' expectations of changes in liquidity risk, derived from the NBU lending survey.

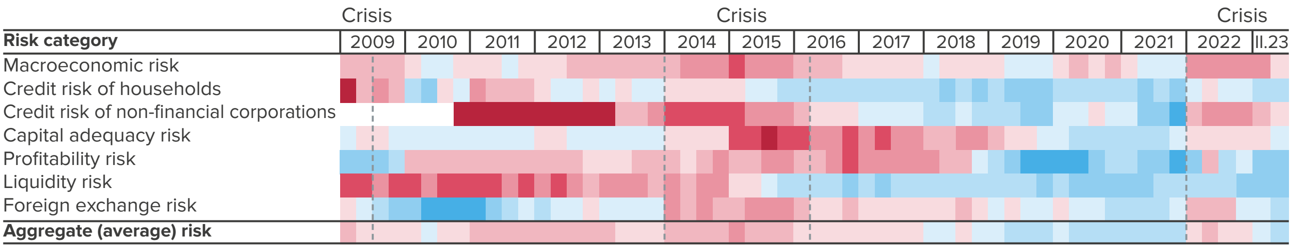
Foreign exchange risk shows to what extent adverse movements in exchange rates can affect financial stability. In fact, two aspects are captured here: the significance of the risk factors in the foreign exchange market and sensitivity of the financial system to those factors.

The first indicator in this category is exchange rate volatility. Higher volatility indicates higher risk. We have also included a leading indicator – the ratio of international reserves to imports. A higher level of this indicator shows a higher sufficiency of international reserves to mitigate possible adverse exchange rate fluctuations. Next, we have included the ratio of the banks' net open foreign currency position to regulatory capital. It reflects the exposure of banks to exchange rate fluctuations and their ability to cover foreign exchange risk by capital. Another indicator of the banks' vulnerability is their relative exposure to FX loans. Risk arises from a probable increase in the debt burden and the credit risk of borrowers who have loans in foreign currency but who do not have FX-linked income. We use the share of FX corporate loans in the total portfolio to capture this risk. FX-lending to households is not considered, as it has been prohibited since 2010. As a forward-looking indicator, we have added the banks' assessment of the foreign exchange risk level from the NBU lending survey. In addition, survey-based expectations of currency risks by corporates and households were added, as expectations may also determine their future behavior and influence risks.

5. RESULTS

In this section, we present the average risk level scores for all seven risk categories between Q1 2009 and Q4 2022.[7] The level of each risk category was calculated as a simple average across the indicators used in the risk category. This abbreviated format of our new heatmap is shown in Figure 3. We use colors to mark each risk level score. The color-coding scheme makes it easier to interpret the level of risk both

[7] In the Financial Stability Reports, the heatmap is shown since Q1 2015.



Notes: Financial crisis periods for Ukraine were derived following the methodology of Filatov (2021).

Figure 3. Heatmap for Risks Monitoring in Ukraine

for each indicator and for each risk category, as well as highlighting periods of higher and lower risk. The colors are the same as in the previous heatmap. The more detailed heatmap – presented with all risk indicators – is shown in Figure 12 (Appendix B).

The heatmap demonstrates a high level of risks in the crisis year of 2009. The following years, foreign exchange rate risk and capital adequacy risk eased, and macroeconomic conditions gradually improved. On the contrary, profitability risk increased. In 2012–2013, the situation worsened, signaling problems that materialized during the 2014–2016 crisis. At that time, most of the risks were at the highest level. A gradual improvement of all risk scores thereafter resulted in the lowest overall risk from 2019 to 2021, which was partially interrupted in 2020 due to the macroeconomic impact of the COVID-19 pandemic. Since the full-scale war began, the estimate of aggregate average risk has increased significantly. To sum up, we can conclude that risk scores calculated completely correspond to the actual situation during the illustrated period.

5.1. Testing the Explanatory Power of Risk Levels

To evaluate the explanatory power of our new heatmap, we employed the receiver operating curve (ROC). The ROC is a plot of the true positive rate against the false positive rate at various threshold settings. A summary measure of this curve – the area under the curve (AUC) measure – is a useful metric to assess predictive performance. An AUC of 0.5 indicates the predictive value of a coin toss. If the AUC is greater than 0.5, the respective factor (or combination of factors) has non-zero predictive power.

To test the ability of risk assessments to describe the current state, we estimated logit regression models for each risk category where explanatory variable is average risk score and dependent variable is the crisis event, which equals 1 if a crisis occurs, and 0 if one does not.[8] To test early warning capacity of heatmap, we built similar logit models for each risk category but dependent variables are crisis events one, two, three and four quarters ahead respectively. These regressions should indicate the ability of the heatmap to predict a crisis up to four quarters in advance. The higher the AUC value for each regression, the better the signaling and predictable power of the risk category scores.

To assess the predictive power of the heatmap more precisely, we employed additional accuracy metrics, which can be found in Table 4 (Appendix A).

In general, the results demonstrate that the heatmap can both show current and predict future crises (Figure 4). Aggregate, macroeconomic and foreign exchange risks explain and predict banking crises well. Profitability risk provides reliable advance signals of financial crises. The weaker performance of other risk categories could be a result of structural changes in the banking sector over the past decades, which have affected the overall risk profile of

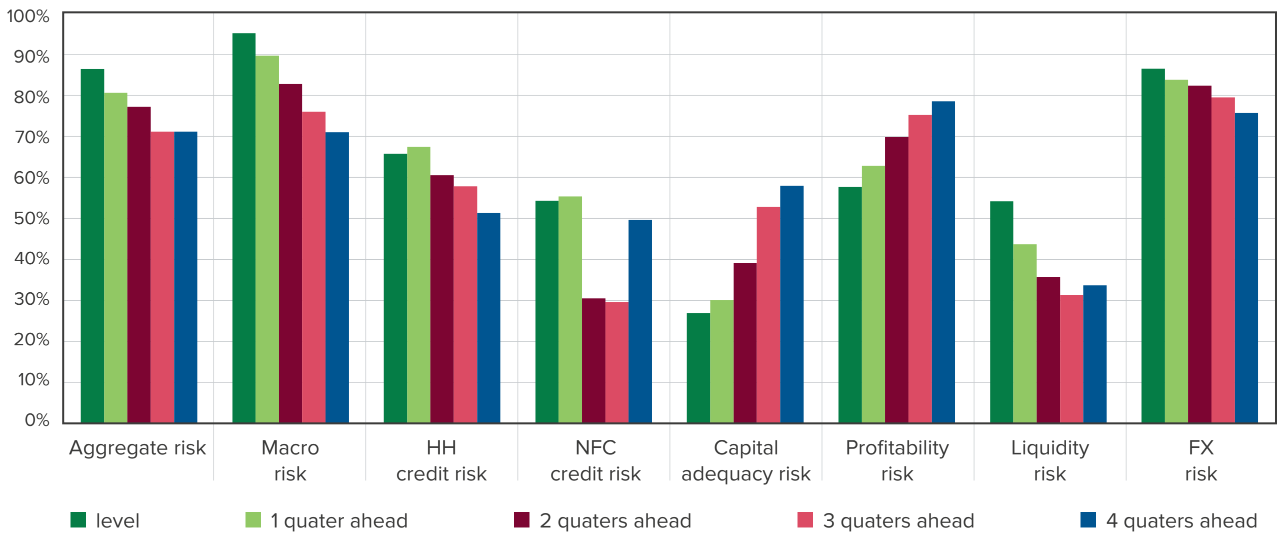


Figure 4. Cross-Validated AUC by Risks

[8] Financial crisis periods for Ukraine were derived following the methodology of Filatov (2021).

the financial sector, and limitations in data for certain periods. In particular, the liquidity risk and credit risk of non-financial corporations have the worst signaling power, mainly due to short time series available. Only two of the four liquidity risk indicators are available for the full time period, and none of the non-financial corporation credit risk indicators are available before 2012. Hence, we do not have enough crisis events in the sample to properly assess the predictive power of the risk estimates of these two risks. At the same time, we believe that these risks have been properly measured in recent periods: the heatmap reflects improved corporate credit risk up to 2022 and low liquidity risk.

Giving the proper signaling power of the heatmap, we discuss risk dynamics in more detail further in this section.

5.2. Dynamics of Average Risk Scores

Based on the dynamics of each risk category scores, we can also explain the key threats to the resilience of the financial system during the analyzed period.

The ***macroeconomic risk*** was building up prior to the 2014–2016 crisis (Figure 5). The imbalances in fiscal and monetary policy led to an increase in the budget and current account deficits in 2012–2013, which were reflected in a gradual increase in the level of risk. Along with a decline in real GDP growth and its forecast, this led to the highest level of macroeconomic risk during the 2014–2016 crisis. At the same time, the macroeconomic risk score was moderate prior to the COVID-19 pandemic and full-scale invasion in 2022. This is well explained by the unexpected and non-economic drivers of these crisis events. Risk scores were growing in response to adverse events of a non-economic nature.

The credit risk of households was among main triggers for financial stability distress in 2009 (Figure 6). At that time the highest level of this risk was observed, being associated with the excessive growth of FX mortgages and a further significant devaluation of the national currency, leading to the insolvency of borrowers. The share of non-performing FX mortgages surged. As a consequence, lending to households in foreign currency was prohibited.

Significant deleveraging followed, lending slowed due to the lower risk appetite of the banks and weak demand from households. Since then, the loan stock of households has remained low, as has lending penetration.

This explains the absence of strong signals from household credit risk prior to all subse-quent crises. During the crisis in 2020 and 2022, the total debt burden and the loan quality remained at appropriate levels, and this risk increased moderately.

The credit risk of non-financial corporations was driving financial system risks for some time before the 2014–2016 crisis (Figure 7). Indeed, that crisis for banks was caused by excessive lending to financially weak borrowers, a significant part of which were related parties. For example, Privatbank, the largest Ukrainian bank, provided more than 97% of corporate loans to companies related to shareholders. Besides that, there were a substantial number of captive banks that served business groups or were used to redistribute cash flows between them.

Crisis led to inability of some corporate borrowers to service their debts. The assets quality review revealed these hidden problems and forced banks to recognize the true quality of loans, leading to higher default rates. The regulatory reforms and measures introduced since 2016 have had a

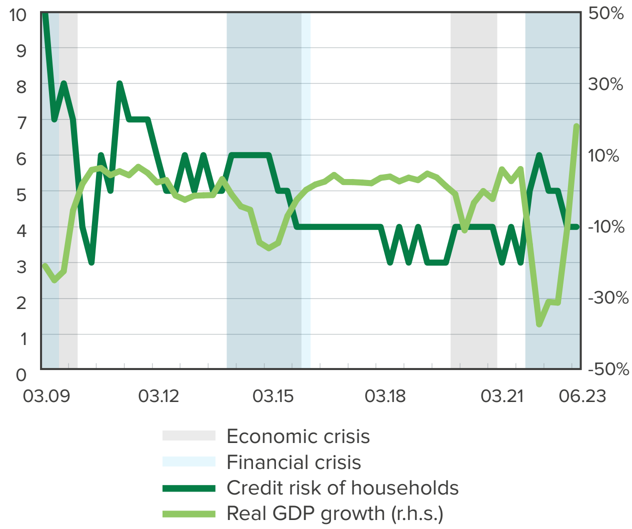


Figure 6. Credit Risk of Households

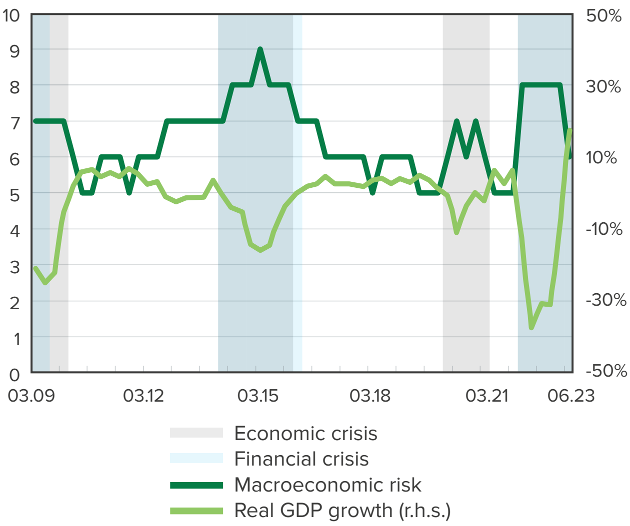


Figure 5. Macroeconomic Risk

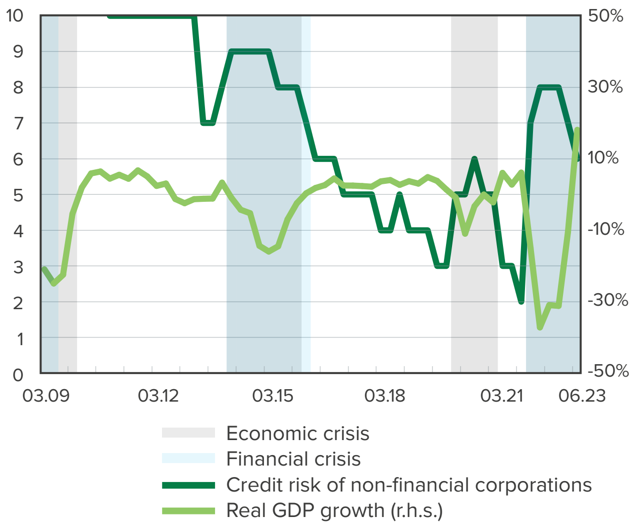


Figure 7. Credit Risk of Non-Financial Corporations

significant positive impact on the quality of the loan portfolio and the transparency of the banking sector. In particular, default rates have gradually decreased, and indicators of the financial state of borrowers have improved. This is fully reflected in the improvement in the corporate credit risk scores in recent years. The sudden surge in the level of credit risk in 2020 was primarily driven by adverse expectations of banks and enterprises, while the actual deterioration of the loan portfolio was moderate. Despite the high quality of the corporate loan portfolio prior to the invasion, the huge economic decline and damage to the real sector made credit risk one the key threats to the financial system in 2022.

Technically, ***capital adequacy risk*** revealed itself as a key risk to the financial system only in 2014 (Figure 8). This is explained by the fact that until then, banks rarely showed the true quality of loans and, accordingly, loan loss provisions. As a result, capital was inflated. Following an assets quality review, the banks were forced to reflect the real situation, and the risk increased sharply. Thus, the highest level of risk was observed in 2015, with gradual improvement seen since then. The banking sector passed through the COVID-19 crisis without significant capital losses. In 2022, capital ratios slightly deteriorated, leaving capital adequacy risk at a moderate level.

Low operating efficiency and a high share of poor-quality assets in the banks' portfolios were sources of high ***profitability risk*** in the financial system for many years (Figure 9). The crisis of 2014–2016 worsened the situation. After the crisis, operational costs surged, and increased default rates forced banks to recognize provisions, reducing profits significantly. After the regulatory reforms and the banking sector cleanup, the system was reborn from the ashes, like a phoenix. In particular, 2021 was the most profitable year in the last 30, despite the COVID-19 crisis. The system continued to generate high profits even in 2022. Hence, the risk scores remain in the "blue" low-risk zone.

Liquidity risk was high prior to 2015: most of the liquidity shortage occurred during the 2014–2016 crisis in small banks, which then left the market (Figure 10). After that crisis, the banks became much more prudent in funds allocation, keeping a high level of liquid assets. This was enhanced by the implementation of new liquidity requirements. Since then, liquidity risk has been low, even during the COVID-19 and war-related crises.

The ***foreign exchange risk*** was one of the triggers for the 2014–2016 crises (Figure 11). Maintaining a fixed exchange rate prior to the crisis required an enormous overdrawing

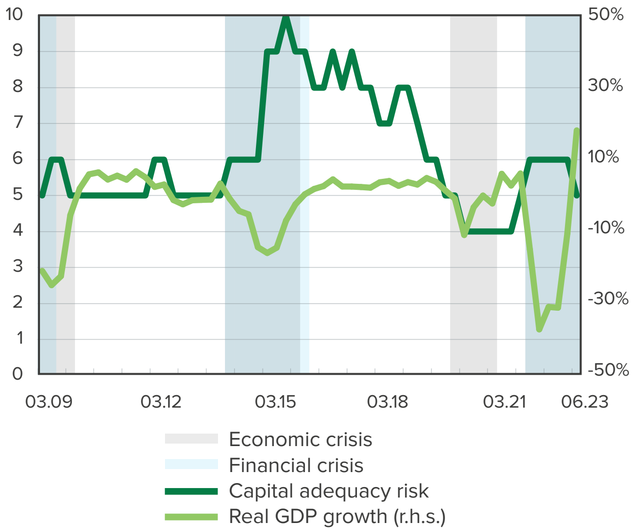


Figure 8. Capital Adequacy Risk

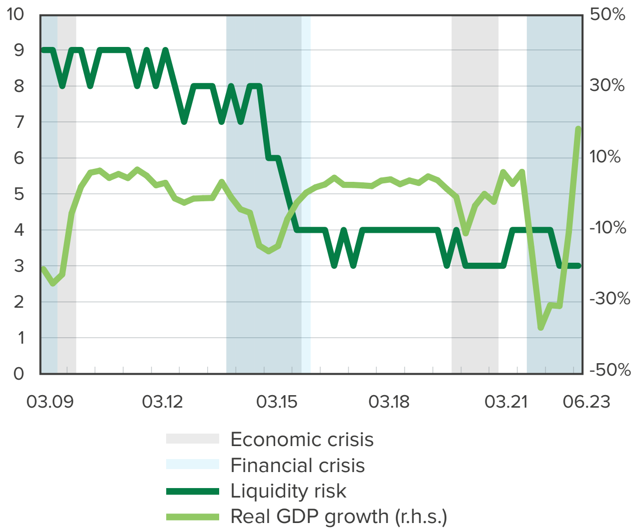


Figure 10. Liquidity Risk

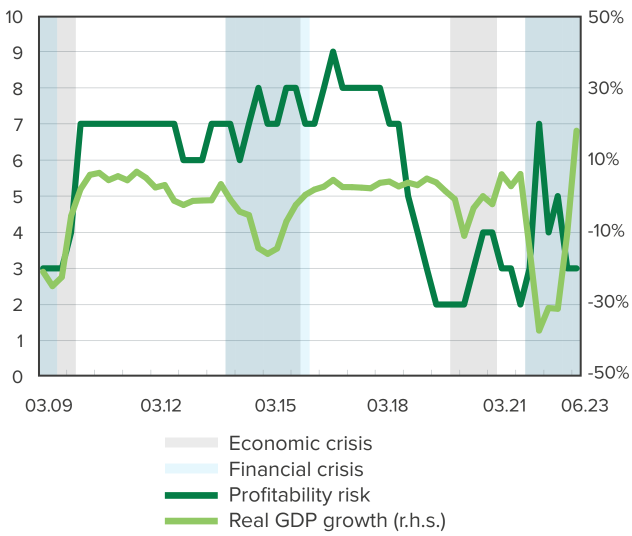


Figure 9. Profitability Risk

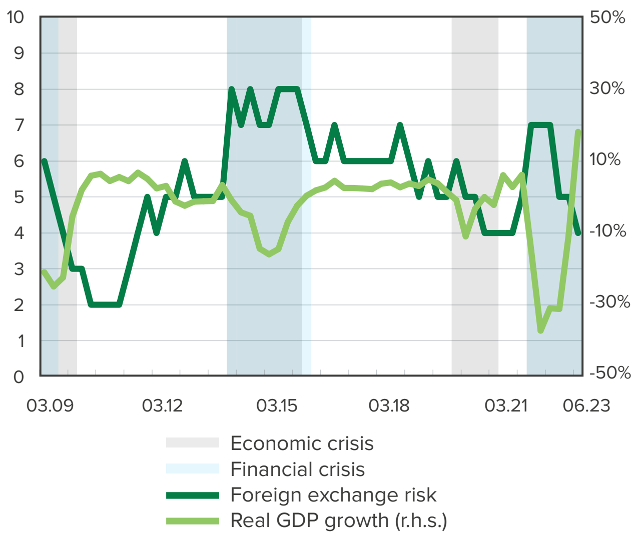


Figure 11. Foreign Exchange Risk

of international reserves. Their exhaustion pointed to an inevitable sharp devaluation, which created stress for the system. Since then, foreign exchange risk scores have improved on average. Currently the highest negative impact for the system can come from still high share of FX loans.

6. CONCLUSIONS

In this study, we present a refined risk map for monitoring systemic risks in Ukraine. The risk map is built on quantitative indicators rather than expert judgments. We identify 40 indicators capturing a wide range of economic and financial vulnerabilities and aggregate them into seven key risks: macroeconomic risk, credit risk of households, credit risk of non-financial corporations, capital adequacy risk, profitability risk, liquidity risk, and foreign exchange risk. The selection of indicators is based on international experience, data availability, and their ability to reflect risks to the financial system or the economy.

The values of the indicators used in the heatmap are assigned risk scores on a 1–10 scale with respective color-coding, with a set of threshold values being constructed for each indicator, using a combination of the historical data distribution in Ukraine, the historical data distribution in a pre-defined set of peer countries, and expert judgments. The color scheme makes it easier to visualize the risk assessment results for each indicator, highlighting periods of higher and lower risk. Finally, indicator risk scores in each risk category are averaged to obtain a score for each type of risk. The aggregate risk level is derived as an average score of all risks.

According to the results, which are also supported by a formal statistical analysis of the early warning properties, the new heatmap efficiently captures the vulnerabilities of the financial system and predicts financial crises up to a one-year horizon. Macroeconomic risk and foreign exchange risk have the best explanatory and predictive power. The weaker results from other risks are mainly due to structural changes in the banking sector and the short time series of data for the indicators.

The heatmap is a useful tool for macroprudential monitoring and will underpin regular risk surveillance and decision-making at the NBU. The forward-looking analysis could help predict crises; simultaneously, the backward-looking analysis could help better understand the causes of previous crises and market reactions to policy initiatives. We also regard the heatmap as a valuable communication tool to raise the awareness of stakeholders and the public about the nature of the risks that threaten financial stability in Ukraine. In addition, the risk map can be used together with indicators to calibrate macroprudential policy instruments.

REFERENCES

Aikman, D., Bridges, J., Burgess, S., Galletly, R., Levina, I., O'Neil,l C., Varadi, A. (2018). Measuring risks to UK financial stability. Staff Working Paper, 738. London: Bank of England. Retrieved from https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2018/measuring-risks-to-uk-financial-stability

Arbatli, E. C., Johansen, R. M. (2017). A heatmap for monitoring systemic risk in Norway. Staff Memo, 10. Oslo: Norges Bank. Retrieved from https://www.norges-bank.no/contentassets/c934013b17fc46259fa27c5da390236e/staffmemo_10_2017.pdf

Arena, M., Bouza, S., Dabla-Norris, E., Gerling, K., Njie, L. (2015). Credit booms and macroeconomic dynamics; stylized facts and lessons for low-income countries. IMF Working Paper, 2015/011. Washington: International Monetary Fund. Retrieved from https://www.imf.org/en/Publications/WP/Issues/2016/12/31/Credit-Booms-and-Macroeconomic-Dynamics-Stylized-Facts-and-Lessons-for-Low-Income-Countries-42615

Borio, C., Lowe, P. (2002). Asset prices, financial and monetary stability: exploring the nexus. BIS Working Papers, 114. Basel: Bank for International Settlements Communications. Retrieved from https://www.bis.org/publ/work114.pdf

Central Bank of Ireland (2020). Systemic Risk Pack, December 2020. Dublin: Central Bank of Ireland. Retrieved from https://www.centralbank.ie/docs/default-source/publications/systemic-risk-pack/systemic-risk-pack---december-2020.pdf

Dell'Ariccia, G., Igan, D., Laeven, L., Tong, H., Bakker, B., Vandenbussche, J. (2012). Policies for macrofinancial stability; How to deal with credit booms. Staff Discussion Notes, 12/06. Washington: International Monetary Fund. https://doi.org/10.5089/9781475504743.006

Demirgüç-Kunt, A., Detragiache, E. (1998). The determinants of banking crises in developing and developed countries. IMF Staff Papers, 45(1). Washington: International Monetary Fund. Retrieved from https://www.imf.org/external/pubs/ft/staffp/1998/03-98/pdf/demirguc.pdf

Demirgüç-Kunt, A., Detragiache, E. (1999). Monitoring banking sector fragility: A multivariate logit approach. IMF Working Paper, 99/147. Washington: International Monetary Fund. https://doi.org/10.5089/9781451856712.001

Drehmann, M., Borio, C., Gambacorta, G. J., Trucharte, C. (2010). Countercyclical capital buffers: exploring options. BIS Working Papers, 317. Basel: Bank for International Settlements Communications. Retrieved from https://www.bis.org/publ/work317.pdf

Drehmann, M., Borio, C., Tsatsaronis, K. (2011). Anchoring countercyclical capital buffers: the role of credit aggregates. BIS Working Paper, 355. Basel: Bank for International Settlements Communications. Retrieved from https://www.bis.org/publ/work355.pdf

Drehmann, M., Tsatsaronis, K. (2014). The credit-to-GDP gap and countercyclical capital buffers: questions and answers. BIS Quarterly Review, March 2014. Basel: Bank for International Settlements Communications. Retrieved from https://www.bis.org/publ/qtrpdf/r_qt1403g.pdf

EBA (2020). Risk Dashboard Data as of Q4 2020. Paris: European Banking Authority. Retrieved from https://www.eba.europa.eu/sites/default/documents/files/document_library/Risk%20Analysis%20and%20Data/Risk%20dashboard/

Q4%202020/972092/EBA%20Dashboard%20-%20Q4%20 2020%20-%20footnote%20%281%29.pdf

ESRB (2013). Recommendation of the European Systemic Risk Board of 4 April 2013 on intermediate objectives and instruments of macro-prudential policy (ESRB/2013/1). Official Journal of the European Union, C 170, 1-19. Retrieved from https://www.esrb.europa.eu/pub/pdf/recommendations/2013/ESRB_2013_1.en.pdf

Filatov, V. (2021). A new financial stress index for Ukraine. Visnyk of the National Bank of Ukraine, 251, 37-54. https://doi.org/10.26531/vnbu2021.251.03

Geršl, A., Seidler, J. (2015). Countercyclical capital buffers and credit-to-GDP caps: Simulation for Central, Eastern and Southeastern Europe. Eastern European Economics, 53(6), 439-465. https://doi.org/10.1080/00128775.2015.1102602

Grui, A., Vdovychenko, A. (2019). Quarterly projection model for Ukraine. NBU Working Papers, 3/2019. Kyiv: National Bank of Ukraine. Retrieved from https://bank.gov.ua/admin_uploads/article/WP_2019-03_Grui_Vdovychenko_ua.pdf

IMF (2011). World Economic Outlook, September 2011: Slowing Growth, Rising Risks. Washington: International Monetary Fund. https://doi.org/10.5089/9781616351199.081

IMF (2019). Global Financial Stability Report, April 2019. Washington: International Monetary Fund. https://doi.org/10.5089/9781498302104.082

Kaminsky, G. (1999). Currency and banking crises: the early warnings of distress. Working Paper, 99/178. Washington: International Monetary Fund. https://doi.org/10.5089/9781451858938.001

Kaminsky, G., Lizondo, S., Reinhart, C. (1998). Leading indicators of currency crisis. IMF Staff Papers, 45(1). Washington: International Monetary Fund. Retrieved from https://www.imf.org/external/pubs/ft/staffp/1998/03-98/pdf/kaminsky.pdf

Kaminsky, G., Reinhart, C. (1999). The twin crises: The causes of banking and balance-of-payments problems. American Economic Review, 89(3), 473–500. https://doi.org/10.1257/aer.89.3.473

Latvijas Banka (2018). Financial Stability Report, 2018. Rīga: Latvijas Banka. Retrieved from https://www.bank.lv/images/stories/pielikumi/publikacijas/FSR_2018_en.pdf

Lepers, E., Sánchez Serrano, A. (2017). Decomposing financial (in)stability in emerging economies. Working Paper Series, 39. Frankfurt am Main: European Systemic Risk Board. Retrieved from https://www.esrb.europa.eu/pub/pdf/wp/esrbwp39.en.pdf

Mencía, J., Saurina, J. (2016). Macroprudential policy: objectives, instruments and indicators. Documentos Ocasionales, 1601. Madrid: Banco de España. Retrieved from https://repositorio.bde.es/bitstream/123456789/6354/1/do1601e.pdf

Mendoza, E. G., Terrones, M. E. (2008). An anatomy of credit booms: evidence from macro aggregates and micro data. Working Papers, 14049. Cambridge: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w14049

Mitra, S., Beneš, J., Iorgova, S., Lund-Jensen K., Schmieder Ch., Severo T. (2011). Toward operationalizing macroprudential policies: When to act? Chapter 3 in Global Financial Stability Report, September 2011. Washington: International Monetary Fund. Retrieved from https://www.imf.org/-/media/Websites/IMF/imported-flagship-issues/external/pubs/ft/GFSR/2011/02/pdf/_ch3pdf.ashx

NBB (2019) Financial Stability Report, 2019. Bruxelles: National Bank of Belgium. Retrieved from https://www.nbb.be/doc/ts/publications/fsr/fsr_2019.pdf

NBU (2020). Financial Stability Report, December 2020. Kyiv: National Bank of Ukraine. Retrieved from https://bank.gov.ua/admin_uploads/article/FSR_2020-H2_eng.pdf

RBI (2013). Report of the Internal Working Group on Implementation of Counter-Cyclical Capital Buffer, Draft, December 2013. Kolkata: Reserve Bank of India.

Schularick, M., Taylor, A. M. (2012). Credit booms gone bust: monetary policy, leverage cycles and financial crises, 1870-2008. American Economic Review, 102(2), 1029-1061. https://doi.org10.1257/aer.102.2.1029

Venditti, F., Columba, F., Sorrentino, A. M. (2018). A risk dashboard for the Italian economy. Occasional Paper, 425. Rome: Bank of Italy. Retrieved from https://www.bancaditalia.it/pubblicazioni/qef/2018-0425/QEF_425_18.pdf

World Bank (2010). Comments on the Consultative Document Countercyclical Capital Buffer Proposal. Retrieved from https://www.bis.org/publ/bcbs172/worldbank.pdf

APPENDIX A. TABLES

Table 2. Tools for Monitoring Financial Stability Risks across Countries

Countries	Name	Risk categories	Aggregation method	Threshold setting and color coding	Source
22 advanced and 7 emerging countries	Matrix of financial vulnerability indices	Nonfinancial Corporates Households Banks Sovereigns Insurers Other financial institutions	Normalization by a pooled z-score, aggregation by an unweighted/ weighted arithmetic average of the z-scores	Percentiles of historical data	IMF (2019)
30 European countries	Risk indicators heatmap	Solvency Credit risk and assets quality Earnings and balance sheet structure	Weighted average	Percentiles of historical data	EBA (2020)
11 emerging countries	Heatmap of vulnerabilities	Valuation pressures and risk appetite Non-financial sector imbalances Financial sector vulnerabilities Global vulnerabilities	Aggregation of standardized series within each component to end up with an aggregated score for that component	By standardized risk score (from 0 to 1)	Lepers and Sánchez Serrano (2017)
Belgium	A risk dashboard for detecting and monitoring systemic risk	Indicators are grouped according to the ESRB's first four intermediate objectives[9]	A simple average of colors associated with all indicators in the sub-category	Mixed approach: early warning methodologies, international level, legislation or guidelines level, cross-country/ historical distribution, expert judgments	NBB (2019)
Ireland	Systemic risk heatmap	Indicators are grouped according to the ESRB's first four intermediate objectives	–	Historical or European average, guidelines level	Central Bank of Ireland (2020)
Italy	Risk dashboard	Interlinkages Credit markets Macroeconomic environment Funding conditions Financial markets Banking and insurance sectors	Standardized series are aggregated by simple and weighted average	Expert judgments or historical distributions	Venditti et al. (2018)
Latvia	Heatmap	External macrofinancial and domestic macroeconomic risks Credit risk of borrowers Liquidity and funding risks Solvency and profitability risks	–	Expert judgments, percentiles of historical observations	Latvijas Banka (2018)

[9] According to ESRB (2013) the intermediate objectives of macroprudential policy should be to: (a) mitigate and prevent excessive credit growth and leverage, (b) mitigate and prevent excessive maturity mismatch and market liquidity, (c) limit direct and indirect exposure concentration, (d) limit the systemic impact of misaligned incentives with a view to reducing moral hazard, and (e) strengthen the resilience of financial infrastructures.

Table 2 (continued). Tools for Monitoring Financial Stability Risks across Countries

Countries	Name	Risk categories	Aggregation method	Threshold setting and color coding	Source
Norway	Heatmap	Risk appetite and asset valuations Non-financial sector imbalances Financial sector vulnerabilities	Each indicator is normalized based on its empirical cumulative distribution function	Shading according to indicator changes from 0 to 1	Arbatli et al. (2017)
Spain	Aggregate heatmap	Potential risks: first four of the ESRB's intermediate objectives and macroeconomic imbalances. Materialized risks: real economy, as well as NPLs and dependence on central bank	Linear aggregation, weighted by adjustment factors: the capacity of indicators to anticipate future crises, the correlation between different indicators	Historical percentiles of the distribution	Mencía and Saurina (2016)
United Kingdom	Heatmap of the individual risk indicators	Private non-financial sector leverage (households, private non-financial corporations, external leverage) Asset valuations (financial and property) Terms of credit (residential and commercial property)	Aikman et al. (2017) approach: unweighted average of z-scores of individual series PCA-based weights "Intensity score" measure according to Kaminsky (1999)	Historical distributions	Aikman et al. (2018)

Table 3. Indicators Selected for Risk Assessment

Risk	Indicator	Description	Threshold	Start date
Macroeconomic risk	Real GDP growth	Is a measure of real economic performance, but is a lagging indicator of risk.	Historical and peer countries data	Q4 2002
	Real GDP growth forecast	Reflects the NBU's expectations regarding the growth or recession of the economy and is one of the main guidelines of economic policy; is an early warning indicator of crises at the macro level.	Countries data	Q1 2015
	Gross external debt to GDP	Indicates the level of debt burden of state residents to non-residents.	Historical and peer countries data	Q4 2003
	Current account balance to GDP ratio	Reflects the trade position of a country. An analysis of the indicator and its dynamics makes it possible to identify imbalances in the foreign economic relations of the state, which appear in a deficit or surplus of the current account of the balance of payments.	Expert judgments	Q4 2001
	External state and state guaranteed debt to GDP ratio	Is used to assess the level of the government's debt burden - a significant level of external debt denominated in foreign currency carries liquidity and solvency risks for the fiscal sector, can lead to the crowding out of private investment, an increase in the tax burden, etc.	Historical and peer countries data	Q4 2001
	State budget surplus/deficit to GDP ratio	Is used as a tool to measure the government's ability to meet its financial needs and ensure efficient public financial management.	Expert judgments	Q4 2005
	Credit default swap 5Y Ukraine	Reflects the probability of Ukraine's default on its obligations, should reflect in advance changes in the expectations of economic agents of the level of fiscal and financial stability of the country.	Historical and peer countries data	Q1 2007
Credit risk of households	Gross bank loans to GDP ratio	Allows the debt burden of households to be estimated.	Historical and peer countries data	Q1 2006
	Gross bank loans to disposable income ratio	Reflects the debt burden of households relative to their real income.	Expert judgments	Q1 2006
	Debt service ratio[10]	Measures the share of household disposable income spent on loan payments relative to total sector liabilities.	Historical and peer countries data	Q1 2012
	Loans at risk	The share of 30 days past due loans in gross performing loans to households.	Expert judgments	Q4 2016
	Index of economic expectations	Shows the expectations of households regarding changes in their financial situation and the development of the country's economy. Lower expectations lead to an increase in savings and a decrease in the purchasing power of consumers, which will ultimately slow down economic activity, and, accordingly, will lead to an increase in credit risk and a decrease in demand for loans in the future.	Expert judgments	Q1 2009
	Expected change in the loan portfolio quality over the next 12 months	Reflects the banks' expectations of changes in the credit risk of households (source – Bank Lending Survey, NBU).	Historical and peer countries data	Q1 2015
Credit risk of non-financial corporations	Net bank loans as a percentage of GDP	Gives an estimate of the debt burden of non-financial corporations at the macro level.	Historical and peer countries data	Q1 2012

[10] The inclusion of the indicators *Gross bank loans to GDP ratio* and *Debt service ratio* simultaneously in the category *Credit risk of households* is due to the following. The amount of debt can be small, so the debt to GDP ratio will not signal high credit risk. At the same time, the high cost of loans can lead to a deterioration in the debt service ratio.

Table 3 (continued). Indicators Selected for Risk Assessment

Risk	Indicator	Description	Threshold	Start date
	Gross corporate debt to EBITDA ratio	Reflects the ability of the corporate sector to meet its debt obligations from operating income; calculated at the level of individual companies and then averaged.	Historical and peer countries data	Q4 2013
	Return on equity (ROE)	Demonstrates how effectively non-financial corporations use capital to generate profits.	Historical and peer countries data	Q4 2013
	Interest coverage ratio	Shows the ability of non-financial corporations to cover interest costs from operating profit.	Historical and peer countries data	Q4 2013
	Default rate	Means the share of non-financial corporations with loans defaulted. This indicator reflects the quality of the corporate loan portfolio.	Expert judgments	Q4 2010
	Business outlook index for the next 12 months	The expectations of enterprises for their development over the next 12 months.	Expert judgments	Q2 2013
	Expected change in the loan portfolio quality over the next 12 months	Reflects banks' expectations regarding changes in the credit risk of non-financial corporations (source – Bank Lending Survey, NBU).	Historical and peer countries data	Q1 2015
Capital adequacy risk	Regulatory capital adequacy ratio	Reflects the banks' ability to pay their liabilities in a timely manner and in full.	Percentiles of historical data	Q4 2005
	Core (Tier 1) capital[11] adequacy ratio	Assesses the banks' ability to fully meet their obligations and remain solvent (going concern).	Percentiles of historical data	Q4 2005
	Net non-performing loans to capital ratio	Reflects the potential level of losses that may arise from the non-performing portfolio of banks, compared to their capital, and hence the banks' ability to absorb these risks and maintain solvency.	Expert judgments	Q1 2009
	Capital to total net assets ratio	Determines the financial leverage of banks, that is, the proportion of assets financed by borrowing. The indicator takes into account risks other than credit, in particular the risks that may arise from investing in government securities. A negative trend in the ratio may signal an increase in risk appetite and possible problems with capital adequacy, which are not fully reflected in the indicators of capital adequacy ratios.	Historical and peer countries data	Q1 2009
Profitability risk	Return on equity (ROE)	Shows how efficiently a bank uses capital to make a profit.	Historical and peer countries data	Q1 2010
	Return on assets (ROA)	Shows how effectively a bank manages its assets to make a profit. The indicator is related to the previous one, however, it should compensate for possible distortions in ROE by reducing capital, rather than increasing profits.	Historical and peer countries data	Q1 2010
	Net interest margin (NIM)	Gives an estimate of the profitability of the main operations carried out by banks.	Historical and peer countries data	Q1 2010
	Cost of risk (CoR)	Shows the level of losses from credit risk per unit of bank loans.	Historical and peer countries data	Q1 2010
	Cost-to-income ratio (CIR)	Is used to measure a bank's performance by comparing a bank's operating expenses with its operating income. Together with the NIM and CoR indicators, it provides a complete picture of the banks' ability to generate profits from core operations and possible risk factors for profitability.	Historical and peer countries data	Q1 2009

[11] Core capital in Ukraine is inherently analogue of Tier 1, but it does not include retained earnings.

Table 3 (continued). Indicators Selected for Risk Assessment

Risk	Indicator	Description	Threshold	Start date
Liquidity risk	Liquidity coverage ratio (LCR)	Is used to assess the state of banks' liquidity over a 30-day horizon. It sets the minimum required liquidity level to cover the net expected cash outflow within 30 calendar days, taking into account the stress scenario.	Expert judgments	Q4 2018
	High-quality liquid assets to total assets ratio	Reflects the volume of highly liquid assets available to banks in case of emergencies associated with a lack of liquidity. The indicator has similar dynamics to the LCR, but is available over a longer period, therefore, it is intended to complement the LCR retrospectively.	Expert judgments	Q1 2009
	Loan to deposit ratio	Indicates the activity of banks in lending, the level of direction of funds into lending operations. A low value indicates the availability of free funds, and therefore high liquidity, a high indicator indicates a greater need to raise funds and higher risks.	Historical and peer countries data	Q1 2009
	Expected change in the liquidity risk for banks over the next quarter	Reflects the dynamics of the liquidity risk during the next quarter according to the banks' assessment (source – Bank Lending Survey, NBU).	Expert judgments	Q4 2013
FX risk	US Dollar exchange rate volatility	Reflects the variability and frequency of changes in the official exchange rate of the Ukrainian national currency against the US dollar over time.	Percentiles of historical data	Q4 2000
	International reserves to import ratio	Shows the sufficiency of international reserves to reduce potential adverse exchange rate fluctuations and maintain the required level of international transactions.	Expert judgments	Q1 2006
	FX corporate loans to total corporate loans	Assesses the volume of credit claims on non-financial corporations that are vulnerable to currency fluctuations. For these loans, fluctuations in the exchange rate can lead to the materialization of both market risk and credit risk due to a negative impact on the solvency of borrowers.	Expert judgments	Q4 2005
	Net open FX position to regulatory capital ratio	Reflects the level of coverage by the capital of potential foreign exchange risks, taking into account the net open foreign exchange position of the bank.	Expert judgments	Q2 2014
	Corporate expectations of UAH/USD exchange rate for next 12 months	Deviation of expectations from the actual values of the exchange rate of the national currency against the US dollar.	Historical and peer countries data	Q2 2013
	Index of devaluation expectations of households	Reflects the expectations of households regarding the devaluation of the national currency against the US dollar.	Historical and peer countries data	Q1 2012
	Change in the currency risk for banks within the past quarter	Demonstrates the dynamics of the foreign exchange risk over the last three months according to the banks' assessment (source – Bank Lending Survey, NBU).	Historical and peer countries data	Q4 2013

Table 4. Predictive Power Performance of a Risk Measures

Metrics	Economic crisis dummy				
	level	1Q ahead	2Q ahead	3Q ahead	4Q ahead
	Aggregate (average) risk				
Accuracy	0.8933	0.8667	0.8267	0.8000	0.7467
Precision average	0.9394	0.9254	0.9058	0.8929	0.8699
Recall average	0.7647	0.7222	0.6579	0.6250	0.5476
F1 average	0.8139	0.7674	0.6880	0.6400	0.5122
Kappa	0.6350	0.5487	0.4080	0.3284	0.1316
AUC ROC	0.8631	0.8104	0.7650	0.7136	0.6680
AUC ROC cross-validated	0.8643	0.8061	0.7719	0.7116	0.7116
Observations	75	75	75	75	75
	Macroeconomic risk				
Accuracy	0.9067	0.8533	0.8400	0.7600	0.7467
Precision average	0.8876	0.8201	0.8310	0.6923	0.7101
Recall average	0.8357	0.7515	0.7190	0.6136	0.5767
F1 average	0.8577	0.7764	0.7500	0.6250	0.5709
Kappa	0.7161	0.5557	0.5087	0.2703	0.1963
AUC ROC	0.9615	0.9016	0.8412	0.7695	0.7152
AUC ROC cross-validated	0.9515	0.8966	0.8275	0.7597	0.7105
Observations	75	75	75	75	75
	Credit risk of households				
Accuracy	0.7069	0.7241	0.7069	0.6724	0.6552
Precision average	0.6250	0.6750	0.6646	0.6697	0.8246
Recall average	0.5861	0.6167	0.6066	0.5368	0.5238
F1 average	0.5897	0.6234	0.6092	0.4848	0.4391
Kappa	0.1958	0.2658	0.2427	0.0923	0.0600
AUC ROC	0.6793	0.6549	0.6269	0.5914	0.5515
AUC ROC cross-validated	0.6582	0.6748	0.6053	0.5781	0.5137
Observations	58	58	58	58	58
	Credit risk of non-financial corporations				
Accuracy	0.7179	0.6923	0.6667	0.6410	0.6154
Precision average	NaN	NaN	NaN	NaN	NaN
Recall average	0.5000	0.5000	0.5000	0.5000	0.5000
F1 average	NaN	NaN	NaN	NaN	NaN
Kappa	0.0000	0.0000	0.0000	0.0000	0.0000
AUC ROC	0.5828	0.5278	0.4660	0.5886	0.6097
AUC ROC cross-validated	0.5438	0.5540	0.3054	0.2969	0.4966
Observations	39	39	39	39	39

Table 4 (continued). Predictive Power Performance of a Risk Measures

Metrics	Economic crisis dummy				
	level	1Q ahead	2Q ahead	3Q ahead	4Q ahead
	Capital adequacy risk				
Accuracy	0.7119	0.6949	0.6780	0.6610	0.6441
Precision average	NaN	NaN	NaN	NaN	NaN
Recall average	0.5000	0.5000	0.5000	0.5000	0.5000
F1 average	NaN	NaN	NaN	NaN	NaN
Kappa	0.0000	0.0000	0.0000	0.0000	0.0000
AUC ROC	0.5056	0.5230	0.5342	0.5622	0.5934
AUC ROC cross-validated	0.2703	0.3019	0.3912	0.5286	0.5803
Observations	59	59	59	59	59
	Profitability risk				
Accuracy	0.6957	0.7391	0.7174	0.7174	0.7174
Precision average	0.6678	0.7250	0.6896	0.6896	0.6896
Recall average	0.6063	0.6688	0.6521	0.6521	0.6521
F1 average	0.6054	0.6783	0.6593	0.6593	0.6593
Kappa	0.2406	0.3699	0.3281	0.3281	0.3281
AUC ROC	0.5917	0.6271	0.7010	0.7542	0.7906
AUC ROC cross-validated	0.5771	0.6287	0.6985	0.7523	0.7855
Observations	46	46	46	46	46
	Liquidity risk				
Accuracy	0.6522	0.6522	0.6522	0.6522	0.6522
Precision average	NaN	NaN	NaN	NaN	NaN
Recall average	0.5000	0.5000	0.5000	0.5000	0.5000
F1 average	NaN	NaN	NaN	NaN	NaN
Kappa	0.0000	0.0000	0.0000	0.0000	0.0000
AUC ROC	0.6063	0.5396	0.4958	0.4688	0.5823
AUC ROC cross-validated	0.5421	0.4371	0.3582	0.3144	0.3376
Observations	46	46	46	46	46
	Foreign exchange risk				
Accuracy	0.9114	0.8667	0.8481	0.8101	0.7595
Precision average	0.9136	0.8532	0.8356	0.7754	0.6913
Recall average	0.8155	0.7602	0.7202	0.6746	0.6084
F1 average	0.8522	0.7917	0.7531	0.6990	0.6187
Kappa	0.7063	0.5875	0.5143	0.4102	0.2602
AUC ROC	0.8695	0.8392	0.8307	0.7975	0.7562
AUC ROC cross-validated	0.8648	0.8377	0.8235	0.7949	0.7566
Observations	79	79	79	79	79

The confusion matrix

		Actual: Yes	Actual: No
Predicted	Yes	True Positives (TP)	False Positives (FP)
	No	False Negatives (FN)	True Negatives (TN)
Total		P	N

The precision metric indicates how many predictions that we made were correct:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

The recall metric shows for the events that occurred, how many we predicted:

$$Recall = \frac{TP}{P} \tag{2}$$

The accuracy specifies how often the classifier is correct.

$$Accuracy = \frac{TP + TN}{P + N} \tag{3}$$

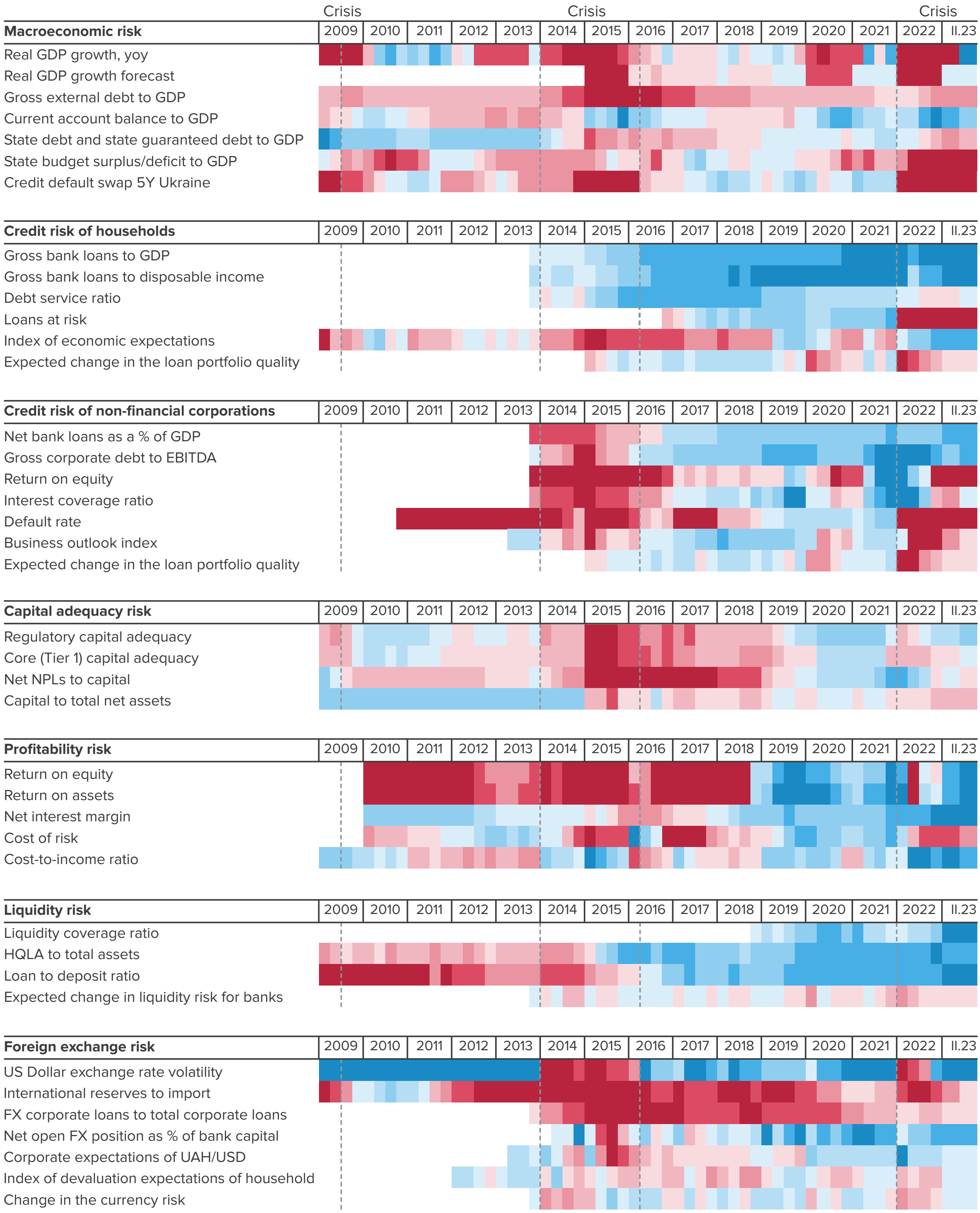
The F1 metric is defined as the harmonic mean (or a weighted average) of precision and recall.

$$F1 = \frac{2}{1 / Precision + 1 / Recall} \tag{4}$$

In addition, we calculated the kappa coefficient, which evaluates how well the classification performs compared to a map in which all values are just randomly assigned. The kappa coefficient can range from -1 to 1. A value of 0 indicates that the classification is as good as random values. A value below 0 indicates the classification is significantly worse than random. A value greater than 0 indicates that the classification is significantly better than random.

The receiver operating curve (ROC) is a plot of the true positive rate ($TP\,rate = TP / P$) against the false positive rate ($FP\,rate = FP / N$) at various threshold settings.

APPENDIX B. FIGURES



Notes: Financial crisis periods for Ukraine were derived following the methodology of Filatov (2021).

Figure 12. Heatmap Visualization by Indicators Risk Score