

A NEW FINANCIAL STRESS INDEX FOR UKRAINE

VLADYSLAV FILATOV^{ab1}

^aNational Bank of Ukraine

^bNational University of Kyiv-Mohyla Academy

E-mail: vladyslav.filatov@bank.gov.ua

Abstract

This study improves on the methodology for calculating the financial stress index (FSI) for Ukraine by introducing a time-varying correlation into the aggregation of five subindices (representing the banking sector, households, the corporate sector, government securities, and the foreign exchange (FX) market). The index consists of 20 indicators selected from an initial list of 47 potential candidates. To check the performance of the indicators, subindices, and the index, we use the area under the receiver operating characteristic curve (AUROC) and logit tests. Each subindex is assigned a weight that reflects the impact of each market on the financial system. This new FSI peaks during periods of crisis that are in line with the consensus of financial experts and performs better than the previous FSI, which makes it more attractive for policy decisions. In particular, the new FSI can be used as a monitoring tool for the macroprudential policy of the National Bank of Ukraine.

JEL Codes

E44, G01, G18

Keywords

financial stability, financial stress index, indicator performance

1. INTRODUCTION

The objectives of modern central banks, and of the National Bank of Ukraine (NBU) in particular, are to ensure price and financial stability. The connections between these two objectives become more pronounced during periods of stress; the crisis in 2008-2009 is an example of this mutual relation. That is why the identification and assessment of risks to financial stability are some of the key functions of a central bank. Financial institutions develop financial conditions indices (FCIs) or financial stress indices (FSIs) to identify such systemic risks. The first FCI developed in the 1990s consisted of a small number of indicators. As financial markets became more complex, FSIs began to appear. The first inclusive FSI was developed by the Central Bank of Canada in 2003. After the financial crisis of 2007-2008, institutions started to develop their own versions of the FSI more actively. For instance, Bank of America developed the Global Financial Stress Index and domestic Financial Stress Indices. The Federal Reserve Banks in the U.S. constructed several local indices (the Kansas City Financial Stress Index, the St. Louis Fed Financial Stress Index, and the Cleveland Financial Stress Index). Moreover, the initial methodologies have been constantly updated in Sweden, Canada, and other countries. Hence, the FSI is a basic monitoring tool for financial institutions in 2020.

In 2017, the NBU developed its own FSI aimed at quantitatively measuring the degree of turbulence in the financial sector. This index comprised four subindices (a banking sector subindex, a corporate sector subindex, a government securities subindex, and a FX market subindex).

The weights of each subindex were set to be constant according to the volume of each market compared to GDP. However, the fixed weights for the aggregation of the subindices have some methodological weaknesses. This design implies that a substantial change in one indicator could cause a material surge in the FSI. Consequently, the FSI could produce signals that are misleading to stakeholders and could even increase uncertainty in the market. In essence, fluctuations in one or several indicators do not necessarily indicate stress in the financial sector as a whole and may send a false signal of increasing turbulence. The high volatility of the FSI due to spikes in the values of individual indicators distorts its explanatory power and makes the FSI less relevant and applicable for policy decision making. For instance, the current FSI includes the indicator Index of Ukrainian Stocks on the Warsaw Stock Exchange. There are approximately six companies traded on the Warsaw Stock Exchange, and the majority of them are agricultural companies. This means that a sectoral crisis in agriculture can significantly increase the FSI, even if there are no shocks to other markets.

Conversely, a relatively high index even in peaceful periods can lead to an underestimation of the stress level. First, the current FSI disregards spillover effects. During a real crisis, one sector's stress may spread to the whole economy as a domino effect. However, the constant-weights approach does not incorporate this co-movement between markets.

The current list of the indicators used in the FSI is also questionable. This list is selected by expert judgement

¹The author would like to thank Professor Jean-Guillaume Sahuc for his valuable guidance and comments. The author is also grateful to the BCC program, the Graduate Institute (Geneva), and the National Bank of Ukraine for providing the data and resources used in this study, as well as the Swiss State Secretariat for Economic Affairs (SECO) for funding this research. The views expressed are those of the author and do not necessarily reflect those of the National Bank of Ukraine.

without tests of the explanatory power of the indicators. Three years on, we see some of the indicators fail to perform as desired. Moreover, there is often a strong negative correlation between the indicators. The contribution of such indicators to the estimates of the overall level of financial stress remains to be determined.

Ex-post analysis gives us the opportunity to select the best stress indicators, exclude less significant indicators and add new ones. Furthermore, there have been several structural changes in the Ukrainian economy during recent years. First, an inflation targeting policy was implemented six years ago. Hence, the key interest rate now plays a major role in central bank's policy. Second, a new monetary regime provided for a flexible FX rate. Before 2014, the Ukrainian economy had a fixed FX rate and primarily experienced periods of rapid currency depreciation during crises. Now, we can observe periods of currency appreciation that might also be a source of risk.

In this paper, we propose a new FSI for Ukraine. We improve the selection of the indicators based on quantitative metrics rather than expert judgement. Moreover, we revise the normalization process and group the indicators into subindices. To reduce the frequency of false signals, we use time-varying correlations instead of fixed weights for the subindices. These updates significantly increase the explanatory power of the index, which makes it more useful for policy making.

Meanwhile, we keep the major blocks of the first version of the FSI for Ukraine. In particular, we continue to use the subindices approach, and we keep the daily frequency of the index. Daily frequency is not common to the FSI of other countries. It is always a trade-off between a time lag of results, data availability and policy purposes. In particular, the most influential indicator for the banking system is credit risk; however, we cannot estimate changes in it on a daily or even on a weekly basis.

Nevertheless, in this discussion, we decided to use daily frequency for the following reasons:

- Speed of the policy response. In times of crisis, the time of the policy response could be costly. Therefore, the daily FSI could demonstrate the most critical days, which need some immediate reaction from policymakers.
- Due to the previous point, policymakers should estimate the reaction of the market to some intervention. In case of a long time lag, the FSI could not give a clear picture of that reaction.
- The FSI is only one instrument of macroprudential monitoring. The NBU also has a quarterly heatmap, a quarterly Financial Conditions Index that consists of such indicators as credit risk. Therefore, the FSI should stay as a supplementary index instead of overlapping with others.
- Ukraine has a quite long list of related indicators with daily frequency. For instance, we tested 47 of them and chose 20 of them. There are no problems with data availability seen on emerging markets.

The paper is structured as follows. Section 2 describes the historical development of FSIs and the most relevant examples of the FSIs. Section 3 outlines the selection of the indicators. Section 4 describes the alternative methodologies for the aggregation of the indicators. The results of the paper

are discussed in Section 5. We highlight robustness testing in Section 6. Section 7 consists of policy recommendations and overall conclusions.

2. LITERATURE REVIEW

The era of FSI development is divided into two phases: before and after the publication of the Composite Indicator of Systemic Stress (CISS). During the pre-CISS era, the first composite indices, such as the Canadian FSI, were introduced, each of which has a completely different methodology. It was as though every team of authors built their own house and no one built a second floor for others. The CISS is an index developed by the European Central Bank. The methodology behind this index created a backdrop for the evolution of other domestic FSIs. First, the methodology sets daily frequency for data as the standard, with minimal delays to publication. Second, the authors significantly improved the methodology for raw indicator transformation. Moreover, the main contribution by the CISS creators was the use of time-varying correlations between subindices. The paper Portfolio-Theoretic Framework for the Construction of Composite Financial Stress Indices by Holló et al. (2012) describes this approach in detail.

The Swedish FSI methodology also adopts a time-varying correlation. The first version of the Swedish FSI (2011) uses a simple average of subindices as its aggregation process. However, Johansson and Bonthron (2013) improve on this methodology and make Swedish FSI 2.0 more advanced. They use a method parallel to the modern portfolio theory and apply the exponentially weighted moving average (EWMA) to build a correlation matrix. Their analysis shows that the new FSI is a better tool to measure financial stress. The authors emphasize that the correlation between the subindices reinforces the magnitude of the index during crises and more clearly highlights stress periods.

Chatterjee et al. (2017) use the basic ideas behind the CISS and improve the algorithm for indicator testing in the construction of the United Kingdom FSI. In particular, the authors use the AUROC and partial AUROC methodologies to test the explanatory power of the indicators. The dependent variable is a crisis dummy and independent variables are normalized indicators. If an indicator has a high AUROC value, it is a good crisis predictor. They test the EWMA and Generalized Autoregressive Conditional Heteroscedasticity (GARCH) approaches to construct the optimal dynamic correlation matrix to aggregate the subindices.

Duprey (2020) uses the Canadian FSI to estimate the relation between financial stress and GDP. The author suggests that a combination of economic decline and financial stress has the greatest negative impact on GDP. Duprey is also the co-author of the paper Dating Systemic Financial Stress Episodes in the EU Countries (2017) that uses a methodology that is parallel to that of the CISS. Duprey et al. (2017) describe the general algorithm for FSI construction. On the one hand, their method is a substitute for the CISS because it describes different approaches to fix the same issue. The authors perform more detailed robustness testing of the index and deeply analyze the normalization of the indicators. On the other hand, their method is complementary to the CISS because they use the basic principle of the CISS, and in particular, their approach to subindex aggregation is parallel to the modern portfolio theory.

A recent work by Drakopoulos et al. (2020) from the International Monetary Fund within the Global Financial Stability Report introduced the Local Stress Index for Emerging Market Economies. This index is primarily based on government bonds and FX markets, that totally corresponds to our two subindices. The authors used also a parallel to the modern portfolio theory approach to aggregate the indicators. For time-varying correlations, they used the EWMA, which we test in Section 4.

Vdovychenko and Oros (2015) suggested the first draft of the Ukrainian FSI. They use four subindices – those of banking, foreign exchange, the stock market, and government debt. Each subindex has only one indicator. On the one hand, the authors play with different specifications for the indicators, which may increase their explanatory power. For instance, they use the first difference in log-transformed variables for the banking sector and the GARCH model for the stock market. On the other hand, the small number of the indicators makes the index more volatile and less resilient to local shocks. Tyschenko and Csajbok (2017) go further and develop a modern version of the FSI in line with the practice of central banks all over the world. They follow Vdovychenko and Oros (2015) and also take four subindices: a banking sector subindex, a corporate sector subindex, a government securities subindex, and a FX market subindex. The authors use a simple average to aggregate indicators within the subindices and an average with constant weights to aggregate the subindices. The authors test different methods of normalization such as the MINMAX range, a cumulative distribution function, equal variance methods, and eventually choose the MINMAX range as their basis.

The figure of Ukrainian FSI 1.0 by Tyschenko and Csajbok is provided below. In recent years, Ukraine has experienced several crises, such as the financial crisis of 2008-2009, the crisis in 2014-2015 caused by the war in Donbas and the ongoing COVID-19 crisis in 2020. The FSI reacts to these periods with spikes. The highest level is 0.65; however, the crises in 2008-2009 and 2014-2015 are deep and comprehensive. Nevertheless, Figure 1 shows some volatility in the non-crisis periods, which is often driven by one factor. The average level of the index is approximately 0.22, even in macroeconomic stability periods. These are the main weaknesses of the FSI's performance we will highlight in this paper.

3. DATA PREPARATION AND SELECTION OF INDICATORS

This section features several stages. First, we identify the crisis periods. For this purpose, we use a dummy for GDP growth as a proxy for real economy developments and a survey of experts as a proxy for financial sentiments. After that, we test the ability of 47 potential indicators to mark off these crisis periods. Finally, we select the final list of the indicators that best define the crisis events and group them into subindices.

3.1. Identification of Crisis Periods

The performance of FSI 1.0 has not been measured, which is one of its main problems. The index shows both upturns and downturns, but there is no evidence of a correlation between true stress and the FSI. To solve this problem, we create several crisis dummies. One is based on GDP growth data, and the other two are derived from a survey of financial experts. This creates a good tool to measure the performance of the index and makes us more confident in our final estimation.

3.1.1. GDP Growth Crisis Dummy

GDP growth is a worldwide indicator of economic performance. However, GDP growth data is quarterly, which is too infrequent to create a dummy. Therefore, we use cubic spline interpolation to convert quarterly data to monthly data. A GDP decline year-over-year that lasts for more than four quarters indicates a crisis. Figure 2 depicts a monthly interpolation of GDP growth.

3.1.2. Expert survey dummies

Using the GDP growth dummy alone generates some controversial results. In general, financial stress correlates with GDP growth; however, there are also sometimes time lags between financial and economic crises. In fact, financial markets react to shocks and start to recover from them earlier than economic markets. Therefore, additional dummies were developed based on experts' judgements on the periods of financial stress.

Financial experts from Ukrainian investment institutions, banks, analytical centres, universities, and government institutions were surveyed about periods of crisis, particularly



Figure 1. FSI for Ukraine (version 1.0)

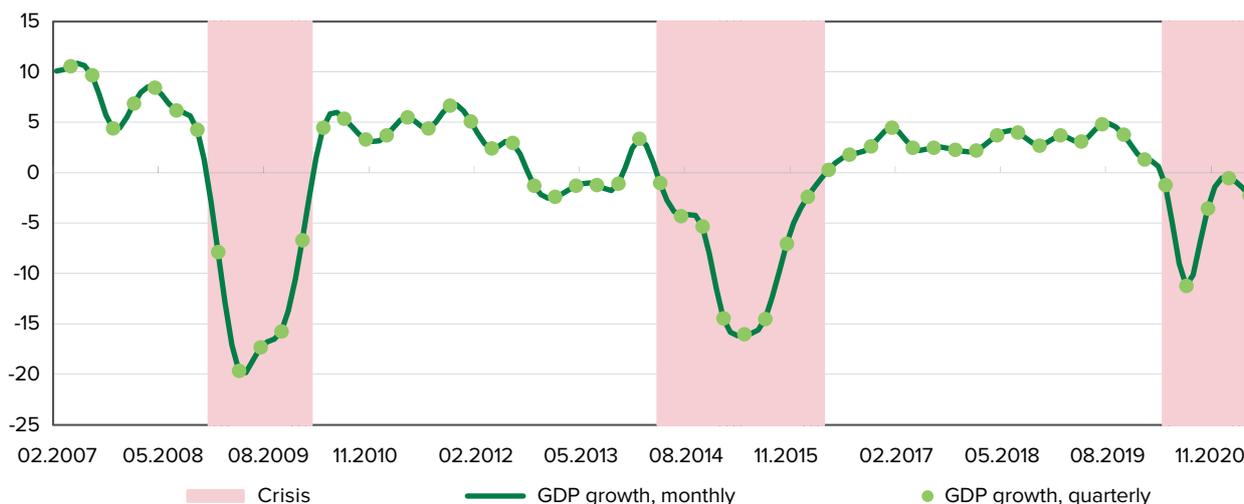


Figure 2. Cubic Spline Interpolation of GDP Growth in Ukraine (%)

regarding the month in which the crises began and ended. We also asked to indicate the worst months during the crisis. The value 1 was assigned to a dummy variable for any month that was marked as a crisis or crisis peak by more than 40 % of the experts.

As a result, we obtain a GDP growth dummy, a survey crisis dummy, and a survey crisis peaks dummy. More details about the dummies are given in Appendix C.

3.2. Selection and Grouping of Indicators

The selection of indicators is the basis for index development. If indicators are selected in the wrong way, the subsequent steps do not have any practical meaning. Therefore, this step is the most time-consuming and important. According to the current methodology and common practice for the majority of domestic FSIs, indicators are selected by a rule of thumb. Riksbank and ECB selection methodologies are also based on expert judgements. However, the UK’s FSI adopts an econometric approach to verify whether the selected indicators are statistically relevant or not. They proposed using a partial AUROC methodology to measure the explanatory power of each indicator.

Our selection process consists of several steps:

1) form a pool of all potential indicators
 First, we take all the indicators from FSI 1.0. Second, we add all the relevant indicators from other countries’ FSIs. As a result, there are both classical financial stress indicators as well as the retail price of gasoline and the price of Brent crude oil in the pool of potential indicators. At this step, we reject indicators only in case of missing data or different frequency of data. Third, we add different terms for each indicator when possible. For instance, we include the price of Ukrainian eurobonds’ credit default swaps (CDS’s) for six months, one year, two years, five years, and seven years. As a result, the pool of the potential indicators consists of 47 items. Then, we consider the different specifications of the indicators. For example, we add both the value of an indicator and its simple 30-day moving average.

2) estimate logistic regressions (logit) using the GDP crisis dummy and the indicators
 We construct a single-factor logit with the GDP growth dummy variable as the dependent variable and the

potential indicators as the independent variables. Indicators were preliminarily transformed into monthly data by averaging the daily data.

We use binary logistic regression with one predictor:

$$\log \frac{p}{1-p} = \beta_0 + \beta_1 x_1, \tag{1}$$

where p is the probability of a crisis (crisis=1), $\beta_{0,1}$ are parameters, and x_1 is an indicator.

For each logistic regression, we recorded the p-values and group together the indicators that were highly significant (P-value $\leq 1\%$), significant ($1\% < \text{P-value} \leq 10\%$) and insignificant (P-value $> 10\%$).

3) estimate the AUROC of the indicators for each dummy
 There are three dummies on our list: the GDP growth dummy, the survey crisis dummy, and the survey crisis peaks dummy. We estimated the AUROC of each indicator for each dummy. A high AUROC value means that the indicator explains the crisis well and produces minimal false signals in usual times. Chatterjee et al. (2017) use a loss function based on AUROC metrics in the UK’s FSI.

$$L(\theta) = \theta T_1 + (1-\theta)T_2 \tag{2}$$

where T_1 is the Type I error rate and is given by $C/(A+C)$. Similarly, T_2 is the share of Type II errors $B/(B + D)$. θ is a parameter from 0 to 1, that weights the loss from each type of error. Values for A, B, C, and D:

	Crisis	No crisis
Above threshold	A	B
Below threshold	C	D

where A is a true positive (TP), and D is a true negative (TN).

Figures 2 and 3 show the overlap in true positive signals and true negative signals. An AUROC value of 0.7 means that there is 70% chance that the model correctly distinguishes between crisis events and no crisis events.

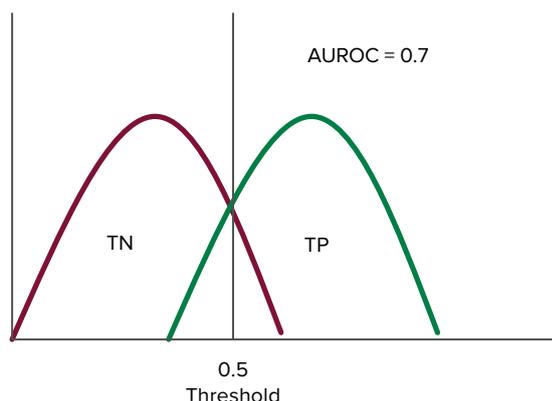


Figure 3. Overlap of True Negative and True Positive in AUROC Analysis

We obtain an AUROC estimate for all potential indicators with the full data available. During this step, we use data as of the start of 2020 (before the COVID-19 crisis).² Table 1 presents the results of the AUROC estimations.

Table 1. Minimum, Maximum, and Mean Values of the AUROC

	GDP growth	Survey crisis	Survey crisis peaks
MEAN	0.8232	0.786	0.836
MIN	0.3575	0.440	0.560
MAX	0.9922	0.967	0.970

The average values of the AUROC are close to each other. The overall average value of the AUROC is 0.8, and we use this as a reference when making further decisions.

4) comparing the indicators' standard deviation and mean for two months before the start of the crisis and the first two months during the crisis

We use data for the three crises: 2008-2009, 2014-2015, and 2020. For each indicator, we calculate its standard deviation and mean for two months before the start of each crisis and the first two months during the crisis.

We assume that during the crisis, the standard deviation increases significantly. Schwert G.W. (2011) points out that high volatility of markets is common for the onset of most financial crises. The mean value must increase

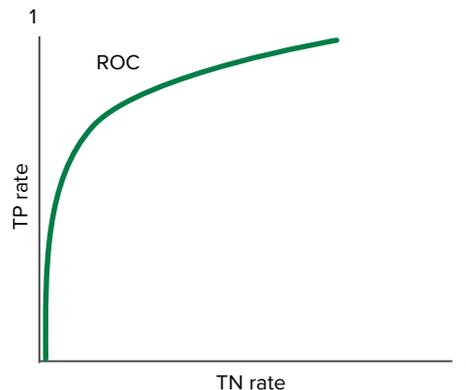
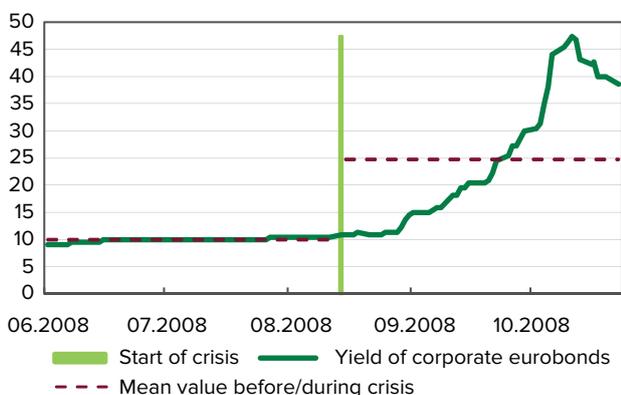


Figure 4. AUROC Curve Example

if the indicator is positively correlated with the crisis and vice versa. The indicator passes this test if the difference between the means and standard deviations in peaceful and crisis periods is higher than the standard deviation for the whole observed period. This should be true for all the three crises.

5) graphical analysis of the indicators

The graphical analysis was conducted for the whole observation period, as well as for each crisis. The indicators that were marked as “Good” have low volatility before the crisis and react immediately to the crisis.³ For instance, see the performance of the YTM of corporate eurobonds in Figure 5.

After all selection steps are completed, we finalize the list of the indicators. An indicator is selected for the final list if it passes all stages: it is significant in the logistic regression with the GDP growth; its AUROC is higher than 0.8 for all three dummies; the standard deviation and mean before the crisis and during the crisis differ significantly; and finally, it shows intuitive dynamics during the stress periods. We also choose only one indicator from among similar indicators (for instance, the CDS five-year spread and CDS two-year spread).

Based on the final list of the indicators, we decide to group them into five subindices: the four subindices from FSI 1.0 and a new one: a household behavior subindex. This new subindex shows the reaction of households to stress

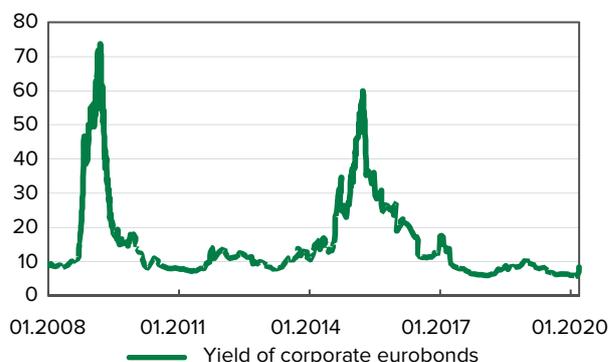


Figure 5. Yield of Corporate Eurobonds in 2008-2020

² We use data from 2020 and 2021 for an analysis of robustness testing in Section 6.

³ The first month of each crisis is chosen with the use of a survey crisis dummy. Within these months, we found the days when FSI 1.0 started to react. These days were marked and the reactions of the indicators were tested on these days.

events. In crises, households start to withdraw money from deposits, which increases the financial market's liquidity risk. Moreover, banks increase deposit rates to reduce this outflow, which creates additional interest rate risk.

All indicators grouped by subindices are given in Table 2.

6) policy check of the indicators.

Our selection process primarily was based on the explanatory power of the indicators. However, economic meaningfulness is crucially important here, especially for future policy interpretation. Due to crosschecking by senior experts the pre-final set of the indicators we use only short-term liquidity support by the National Bank in the indicator NBU's Support to Banks instead of all support.

Each subindex consists of a mix of short-term and long-term risks. For instance, the banking subindex consists of

liquidity stress indicators such as an LCR and the Price of Ukrainian banks' eurobonds that correspond to the risk premia of those banks (long-term risks). The government debt subindex consists of the yield of domestic bonds in UAH. The volatility of that indicator on the short-term horizon associates with liquidity issues faced by the government. Simultaneously, the price of government bonds' CDS's is related more to long-term credit risks.

However, there is also a household behavior subindex that corresponds only to short-term characteristics. Historically, in the time of stress, households start to withdraw short-term deposits, banks react and increase the interest rate for them (NBU, 2016). Therefore, this subindex is a good proxy for media sentiments and accordingly level of household uncertainty. This behavior is important because it could be a source of the banks' large liquidity and, on a greater scale, even bankruptcy.

Table 2. The Final List of Indicators

Name of Indicator	Description
BANKING subindex	
Ukrainian OverNight Index Average (UONIA)	Change in the overnight interbank interest rate. Calculated by the NBU.
Kiev Interbank Bid and Offer Rates (KIEIBOR), 1-month	Change in the interbank rate for a 1-month term. Calculated by the Association of Ukrainian Banks.
Price of Ukrainian banks' eurobonds	Price of eurobonds issued by Oschadbank, Ukreximbank, First Ukrainian International Bank, and PrivatBank (until the December 2016 bail-in)
Liquidity coverage ratio (LCR)	Average of banks' LCR, weighted by expected outflows in the LCR denominator
NBU's support to banks	Amount of the NBU's liquidity support transactions for the last 60 days. Only transactions with terms less than 90 days are included.
GOVERNMENT DEBT subindex	
Ukrainian CDS's, 5-year	Price of 5-year CDS's of Ukrainian sovereign bonds
Sovereign risk of Ukrainian eurobonds	Spread between the weighted average yield of Ukrainian eurobonds and the yield of 2-year U.S. Treasury bonds
Yield of domestic bonds in UAH	Simple average of yield to maturity (YTM) for Ukrainian domestic bonds in UAH
Bid-ask spread of Ukrainian eurobonds	Simple mean bid-ask spread of Ukrainian eurobonds on a given date
HOUSEHOLD BEHAVIOR subindex	
Ukrainian Index of Retail Deposit Rates (UIRD), 3-month	Change in retail 3-month deposit rates in UAH of the 15 largest banks. Calculated by Thomson Reuters.
Change of retail deposits in UAH	Percentage change in the stock of retail deposits in UAH over the last 30 days
Change of retail deposits in USD	Percentage change in the stock of retail deposits in USD over the last 30 days
CORPORATE subindex	
Yield of corporate eurobonds	Corporate bonds' YTM for Ukrainian enterprises. Calculated by Cbonds Agency.
Stock index	Deviation of the stock index from its maximum over the last year. The PFTS Index is used before 2012 and the Stock Index on the Warsaw Stock Exchange (WIG Ukraine) afterward.
Volatility of stock index	Standard deviation of the stock index over the last 30 days. The PFTS Index is used before 2012 and the Stock Index on the Warsaw Stock Exchange (WIG Ukraine) afterward.
FX MARKET subindex	
USD/UAH exchange rate	Deviation of the USD/UAH exchange rate from the maximum over the last year
Volatility of USD/UAH exchange rate	Volatility of the USD/UAH exchange rate over the last 30 days
UAH/USD non-deliverable forward	Difference between the non-deliverable forward (NDF) of UAH/USD for 3-month and spot UAH/USD exchange rates
Yield of non-deliverable forward, 3-month	Change in 3-month UAH/USD NDF's YTM
Currency intervention by the NBU	Net purchase/sale of foreign currency by the NBU on the interbank FX market.

We also check for an inadequate representative of some indicators due to illiquidity. There are no direct rules where the CDS market or the corporate eurobonds market is liquid or not. Therefore, we use judgmental expertise to estimate the required level of liquidity in each special case. For that purpose, we estimated historical comparison instead of comparison with peer countries. The liquidity of most indicators is on a growth path; therefore, it is the main reason to keep them in our final set.

4. METHODOLOGY FOR THE INDEX COMPOSITION

4.1. Normalization of Indicators and Their Aggregation to the Subindices

Each indicator has different units of measurement, which is why we normalize them. We test several approaches in keeping with the mentioned reference. One of them is the cumulative distribution function, an approach that, for Ukrainian data, gives many noisy and false signals. This is due to the high volatility of Ukrainian markets even in normal times. This method may be appropriate for developed economies; nonetheless, it is not applicable to such an emerging market as Ukraine. Another method is Z-score normalization, which Lang et al. (2019) use in the development of domestic systemic risk indicators. This method of normalization gives us stable, expected and plausible results. It is not sensitive to outliers and does not create much noise.

The formula for Z-score normalization is:

$$\frac{X_{it} - \mu_i}{\sigma_i}, \quad (3)$$

where μ_i is the mean value of the indicator and σ_i is the standard deviation of the indicator.

Another approach, MINMAX-range normalization, gives similar results. The weakness of this method is a need for retrospective recalculation when a new value appears that is historically the highest or the lowest observed.

The formula for MINMAX-range normalization is:

$$\frac{(X_{it} - \text{MIN}_i)}{(\text{MAX}_i - \text{MIN}_i)}. \quad (4)$$

Comparing these two methods, we have decided to use MINMAX because this method gives a complete and finite range [0:1], while Z-score values may be outside of this range. Second, we use this method in our FSI 1.0; therefore, we have decided to continue using it in the new FSI for Ukraine (FSI 2.0) to ensure the results are comparable.

To compose each subindex, we use a simple average of the normalized indicators. This is common practice for FSI methodologies, and it decreases the probability of one indicator dominating the index.

4.2. Weights for Subindices

We need to start with a fixed weight for the subindices regardless of the aggregation approach. There are several approaches to estimate the weights.

Tyschenko and Csajbok (2017) choose weights based on the importance of a sector according to its size relative to GDP. Lang et al. (2019) propose using the regression approach to estimate the weight of each subindex. The authors estimate regression coefficients and divide each coefficient by the sum of all coefficients. There are other methods, for instance, estimation by pairwise vector autoregression models (VARs) with GDP; however, this method is not suitable for short samples.

In this research, we replicate the estimations of sector size proposed by Tyschenko and Csajbok (2017). We also use Lang et al.'s (2019) approach. However, a logit model that includes all subindices gives us unstable and unintuitive results that could be a sign of multicollinearity. This is why we estimate a single-factor logit for each subindex. After that, we sum the coefficients of the five logit-regression models to find the weight for each subindex.

For weight robustness testing, we use the AUROC to examine each subindex. A subindex with a higher AUROC value should receive a higher weight. For instance, even if the ratio-to-GDP approach and the logit-regression approach assign low weights, we can increase the weight if the AUROC is high. Hence, robustness testing with the AUROC gives us more information for the final weight calibration.

We use the following metrics to compare sector size to GDP (ratio to GDP) for each market and subindex:

- 1) banking subindex—the total volume of loans to residents (non-financial corporations and households)
- 2) household behavior subindex — the volume of household deposits
- 3) corporate subindex — the size of stock market capitalization and the volume of the corporate bond market. This value was taken from Tyschenko and Csajbok (2017), which was estimated for FSI 1.0
- 4) government debt subindex — the volume of outstanding local-currency sovereign bonds and sovereign eurobonds
- 5) foreign currency market subindex — the share of financial assets and liabilities in foreign currency (loans and deposits dollarization).

Table 3 summarizes the results of the estimation based on several approaches.

The final weights are based on all values mentioned above. Choosing the weights for the banking and household behavior subindices is rather trivial, as various approaches show similar results.⁴ The weight for the FX market subindex varies in the range from 24.2% to 32%. Based on the AUROC results, we decided to choose 25%. Moreover, FX risks have decreased in recent years in Ukraine. The final reallocation of weight is between the government debt and corporate subindices. The AUROC for the corporate subindex is the highest, which is why we take the value from the high end of the 10%-22% range. Correspondingly, we choose the minimum value for the government debt subindex.

⁴ The banking subindex of FSI 1.0 consists of indicators included in the banking and household behavior subindices of the new FSI. The total weight of these subindices now is 35% compared to 30% in FSI 1.0.

Table 3. Estimation of Weights for Subindices

	Subindices				
	Banking	Household behavior	Government debt	Corporate	FX market
Logit-regression coefficient adjusted, % ⁵	19.5	13.5	20.8	22.0	24.2
Ratio to GDP, %	20	12	26	10	32
FSI 1.0, %	30	—	25	10	35
AUROC for survey crisis dummy	0.83	0.77	0.92	0.91	0.90
AUROC for survey crisis peaks dummy	0.87	0.84	0.87	0.94	0.92
Suggested weights, %	20	15	20	20	25

4.3. Aggregation of Subindices

The next step is to aggregate the subindices. A review of other index methodologies shows us different approaches that can be generalized into several groups. The first group of researchers uses a simple or weighted average to aggregate the subindices. Easy interpretability and understandability are the main advantages of this method. However, significant disadvantages are the sensitivity of the index to changes in one subindex and underestimation of the synergistic interactions between variables. The second group of authors experiments with different models, such as factor-augmented VARs and principal component analysis. These approaches demonstrate the best performance in particular countries and are able to capture some country-specific features. However, they are frequently not useful for other countries. The last group of authors uses an approach that is based on the modern portfolio theory. We decide to estimate the FSI, using a weighted-average method and the portfolio theory method.

4.3.1. The Weighted-Average Approach to Aggregation

The weighted-average aggregation approach is currently used in FSI 1.0.

$$FSI_t = \sum_1^i s_{i,t} \times w_i, \tag{5}$$

where $s_{i,t}$ is the value of subindex i in period t and w_i is the weight of subindex i .

In this case, the weights of the subindices are constant. However, in reality, the impact and size of each market may change over time. For example, the government debt-to-GDP ratio in 2008 and in 2016 is completely different. This means that after structural changes in the economy, we should re-estimate these weights to obtain precise estimates.

4.3.2. Portfolio Theory Approach to Aggregation

The main innovative feature in the design of the CISS is the use of the modern portfolio theory for subindex aggregation. After the introduction of the CISS, many other institutions have considered the modern portfolio theory for their domestic indices. Today, the Swedish FSI, the UK’s FSI, the European CISS, the European FSI by Duprey et al. (2017), and the Canadian FSI use this methodology:

$$FSI_t = (s_t \times w) C_t (s_t \times w)^T, \tag{6}$$

where s_t is the vector of the values of the subindices in period t , w is the vector of weights for the subindices and C_t is the dynamic correlation matrix for the subindices in period t , given by:

$$C_t = \begin{bmatrix} 1 & & & & & \\ \rho_{12,t} & 1 & & & & \\ \rho_{13,t} & \rho_{23,t} & 1 & & & \\ \rho_{14,t} & \rho_{24,t} & \rho_{34,t} & 1 & & \\ \rho_{15,t} & \rho_{25,t} & \rho_{35,t} & \rho_{45,t} & 1 & \end{bmatrix}, \tag{7}$$

where $\rho_{j,t}$ is the correlation between subindex j and subindex i in year t .

We can obtain this correlation matrix in different ways. As we mentioned in literature review, the EWMA and the multivariable dynamic conditional correlation GARCH (DCC-GARCH) are the most commonly used approaches to aggregate financial stress sub-indices.

EWMA

The exponentially weighted moving average is a moving average model. It allows for larger reactions to recent changes. The β -s parameter corresponds to the memory of the process. The higher this parameter is, the more resistant to recent data the correlation matrix is.

$$\rho_{j,t} = \sigma_{j,t} / \sigma_{i,t} \sigma_{j,t} \tag{8}$$

$$\sigma_{j,t} = \beta \sigma_{j,t-1} + (1-\beta) z_{i,t} z_{j,t} \tag{9}$$

$$\sigma_{i,t}^2 = \beta \sigma_{i,t-1}^2 + (1-\beta) z_{i,t}^2 \tag{10}$$

In line with Holló et al. (2012), we test different values for β -s. The range of β -s is from 0.89 to 0.98. The authors of the UK FSI and the Swedish FSI also use a β -s value from this range. Figure 6 shows the results of aggregation for $\beta=0.89$, $\beta=0.93$, and $\beta=0.97$.

Table 4 reports the AUROC results applied to these alternative indices. We observe that $\beta=0.97$ is the value that gives the highest AUROC value. Note that the index calculated based on this parameter best explains the crisis in 2014-2015. Other parameters show that stress in 2008-2009 is twice larger than in 2014-2015; however, real data show that in 2014-2015, at least the same level of stress was observed as in 2008-2009. Setting the value of β to 0.97 allows us to reproduce this empirical observation.

⁵ The outcomes of logit-regressions are available on request.

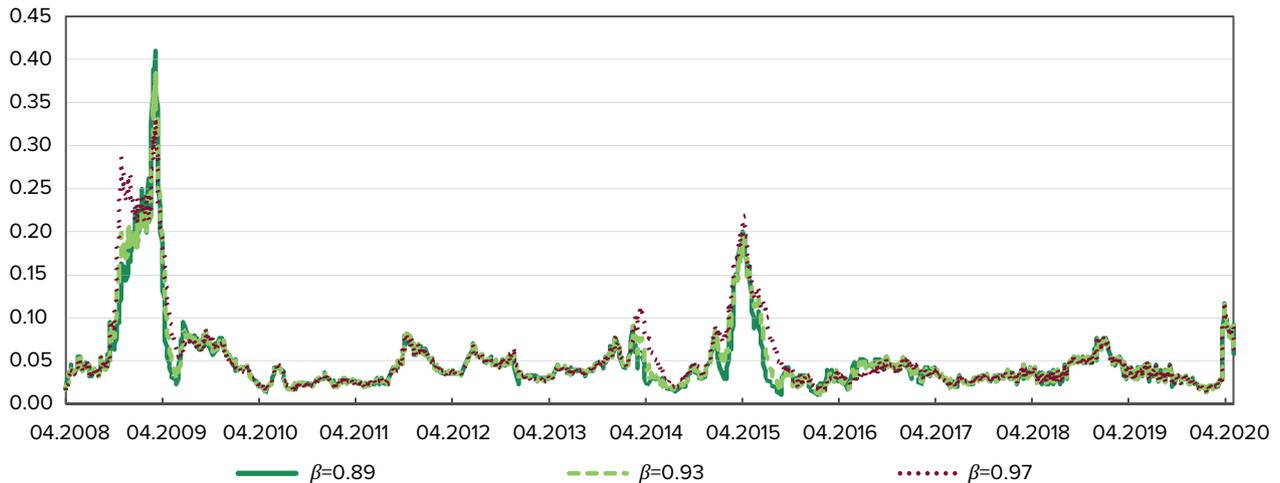


Figure 6. EWMA Aggregation with Different Values of β

Table 4. AUROC Testing of EWMA Results with Different Parameters

	Survey crisis	Survey crisis peaks
$\beta=0.89$	0.771	0.808
$\beta=0.91$	0.794	0.833
$\beta=0.93$	0.816	0.863
$\beta=0.95$	0.840	0.898
$\beta=0.97$	0.874	0.946

DCC-GARCH

The DCC-GARCH was introduced by Engle and Sheppard in 2001. Following Chatterjee et al. (2017), we use GARCH (1,1). The model includes two parameters (α, β), and we estimate them using the full sample. Details of the methodology are described in Appendix B. The results of the model estimation are presented in Appendix D.

With the use of this approach, we obtain a dynamic correlation matrix. In Figure 7, we can see periods of an almost perfect correlation between the subindices where the correlation reinforces the direct effect of the subindices. The periods of such perfect correlation are observed during crises. In normal times, correlations between the subindices usually decrease.

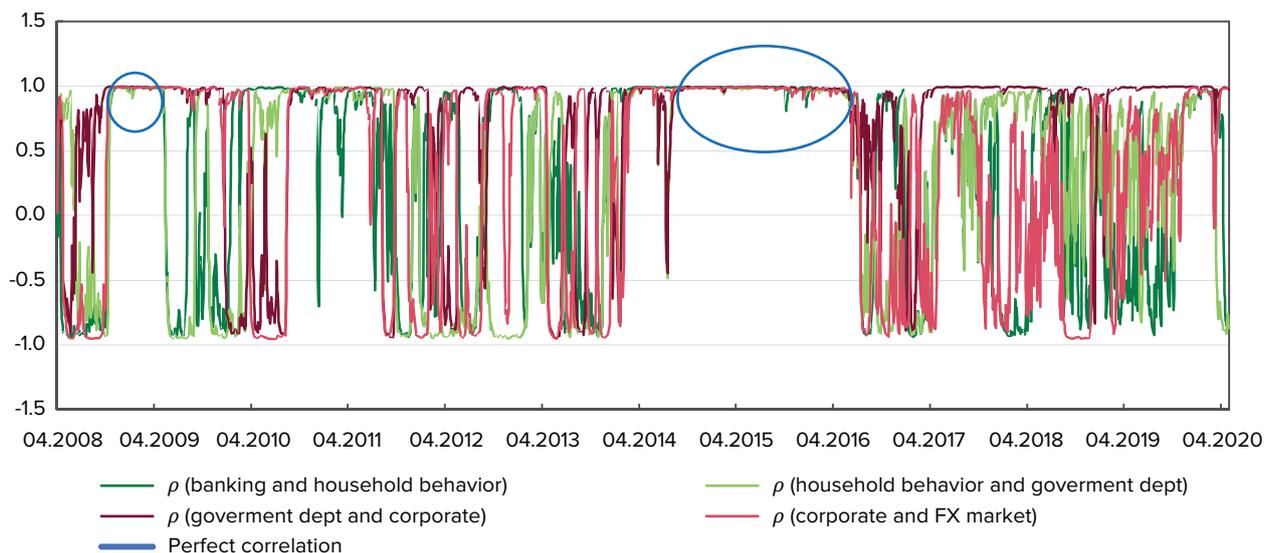


Figure 7. Correlations Between Subindices

5. RESULTS

Figure 8 reports the values of the three indices calculated with the use of different approaches to the subindex aggregation. In the figure, the weighted-average approach shows the highest level of stress during the crises; however, we should consider the specifics of each method. The weighted-average approach assumes a perfect correlation between all subindices in all periods. Moreover, it is calculated based on simple averaging, while two other methods require multiplication. This is why the EWMA and GARCH approaches by default have lower values than the weighted-average approach, including during crisis periods.

As the direct graphical comparison is inaccurate in this situation, we use other approaches to investigate the pros and cons of each index.

We consider several aspects:

- 1) values during the crisis of 2008-2009 (“+” if the index produces high values)
- 2) values during the crisis of 2014-2015 (“+” if the index produces high values)
- 3) values during the crisis of 2020 (“+” if the index produces medium values, as currently the impact of the crisis on the financial sector in Ukraine is moderate)

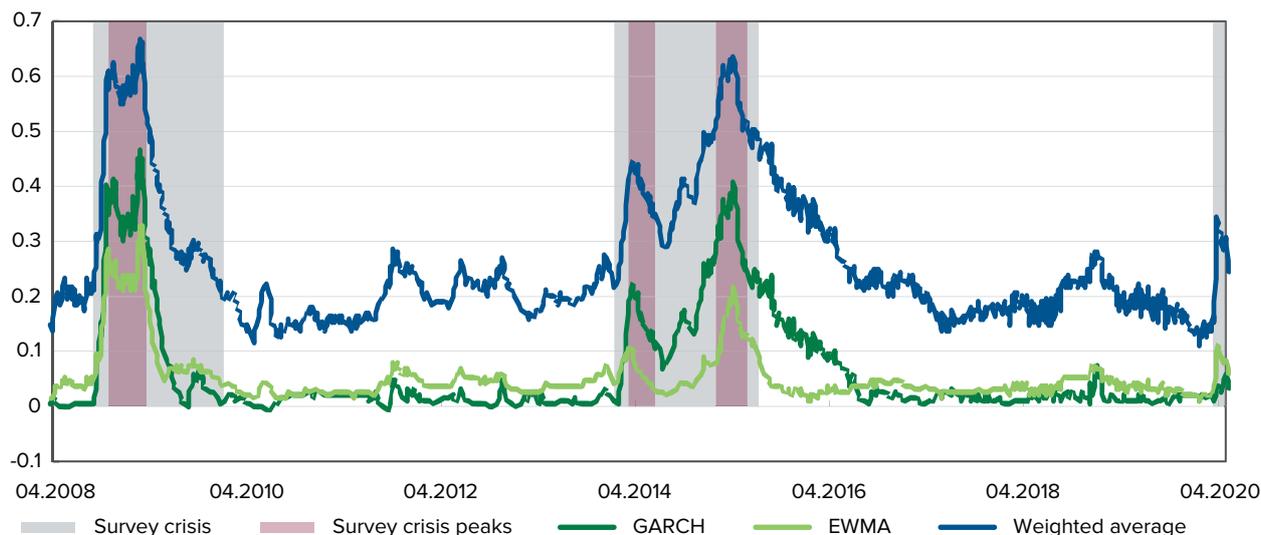


Figure 8. Comparison of the Indices

- 4) volatility in normal times (“+” if the index has low volatility in normal times)
- 5) peaks in crisis (“+” if the index identifies the peaks)
- 6) AUROC robustness testing.

Table 5 presents the comparison of the indices.

Based on these metrics, we decided to use the GARCH approach for aggregation. The main reason is that it generates fewer false signals in normal times and higher values during crises. The purpose of the index is to show the magnitude of the stress in the whole financial sector, not for one specific sector. To make the selected index more visually attractive, we normalize it with the use of the MINMAX methodology. This is done to (i) eliminate the negative values in the FSI; (ii) make the index more attractive: a range from 0 to 1 is much easier to interpret than a range from -0.1 to 0.5; and (iii) simplify the transition from FSI 1.0 to FSI 2.0.

Table 5. Comparison of Aggregation Methodologies

	Crisis 2008-2009	Crisis 2014-2015	Crisis 2020	Volatilities in usual times	Peaks in crises	AUROC survey crisis	AUROC survey crisis peaks
Weighted average	+	+	+	–	+ –	0.939	0.978
EWMA	+	–	+	+ –	+ –	0.874	0.946
GARCH	+	+	+ –	+	+	0.886	0.978

Figure 9 shows the final version of FSI 2.0, noting the main economic and political events. We can observe that after the Lehman Brothers bankruptcy, the index spikes drastically. We can also see a useful insight for policymakers: the peak of the crisis in 2014–2015 is just before the talks on Ukrainian external debt reprofiling were launched. However, FSI 2.0 has not reacted significantly to the COVID-19 quarantine, while the FSIs of other European countries have experienced a significant leap. The reason could be that the effect of the COVID-19 crisis is currently much lower for the Ukrainian financial sector than that of the previous crises.

The decomposition of the index shows us a variety of insights. It depicts the impact of the estimated time-varying correlation on the index value compared to the stable-correlation assumption. Figure 10 shows that correlation matters the most in crisis times. However, in times of macroeconomic stability, the correlation effect is low or even negative, which is intuitively expected.



Figure 9. The Final Version of FSI 2.0

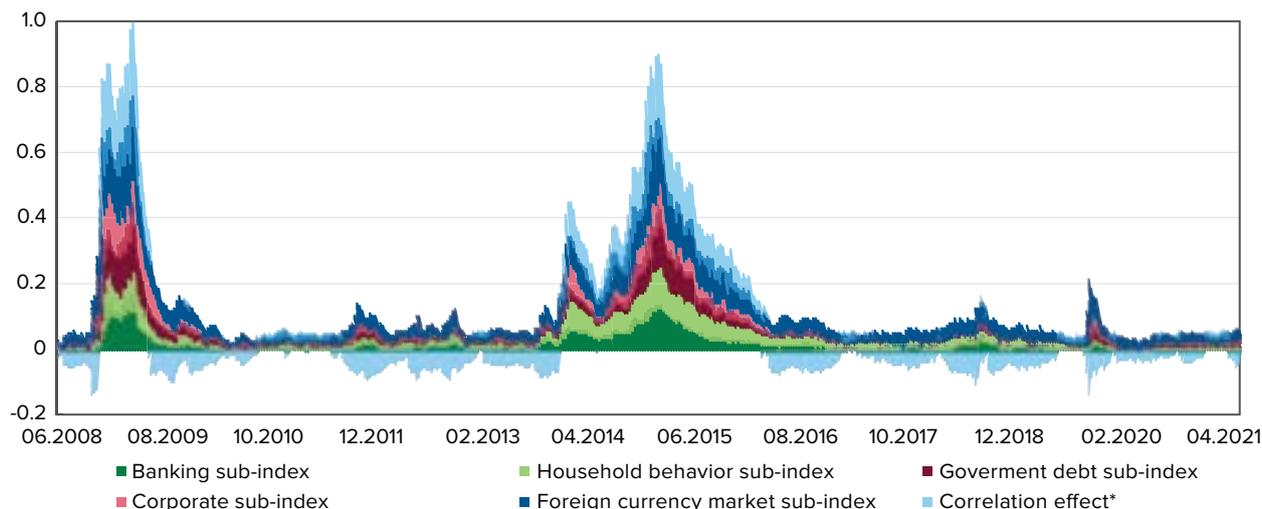


Figure 10. Decomposition of FSI 2.0

*A correlation effect is a net effect of the time-varying correlation (excluding the average correlation for the entire observation period)

6. ROBUSTNESS CHECKS OF EFFECTIVENESS

Our research primarily was done due to June 2020, when we observed only three months of the COVID-19 crisis. However, we processed additional one-year data since that period for robustness checks. H2 2020 – H1 2021 is the continuity of economic crises for most countries; therefore, the comparison is consistent in that view.

We compare the Financial Stress Indices of Canada, Sweden, the United Kingdom, the United States of America, the integrated FSI by the Office of Financial Research (OFR) with the Ukrainian FSI. This exercise is not trivial due to several reasons. At first, there are different methodologies for each index. Some indices focus more on market indicators; other indices show the financial sector solvency. There is also an issue of levels differentials. We use the MINMAX approach to normalize indices from 0 to 1 levels. Secondly, there are different frequencies. Most indices are monthly or weekly, while the Ukrainian one or that of the OFR are daily. We use the simple average for converting daily or weekly data to the monthly level. Thirdly, there is no open data for FSIs of developing countries /emerging markets; therefore, we compare Ukrainian financial stress with the stress of developed countries. The exception is the OFR's FSI, which consists of emerging economies' element.

Expectably, we observed a high spike in March 2020 due to the coronavirus pandemic. Before that, Sweden, the UK, and Ukraine had more volatility than Canada, the U.S., and the world economy overall (the OFR index). However, the dynamics after this spike are of more interest. We see a high correlation between all indices during the COVID-19 pandemic, and Ukraine is in line here.

We compare the original Ukrainian FSI with the OFR index daily. This comparison has less bias due to the absence of data conversion. We see that growth at the start of March 2020 is the same in amplitude. After some peak, there is a gradual decrease in stress levels. At the end of 2020, we observe local minimum for both indices.

Figure 11 and Figure 12 show us that 1) the modern Ukrainian FSI has consistent methodology with most of other FSIs; 2) Ukraine's financial system in 2020 had high interdependence with the global system. Moreover, COVID-19 did not have a significant and long-term effect on the financial sector in Ukraine, as it was in 2008 or 2014 due to other shocks.

The implemented robustness tests justifies the efficiency of the new FSI for Ukraine. We can use this methodology for monitoring stress development.

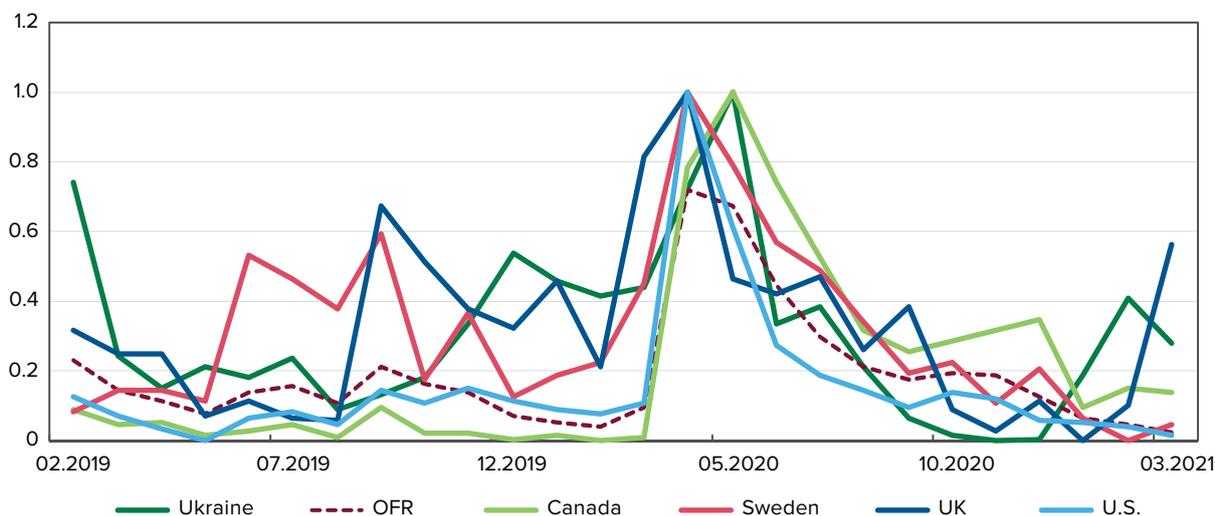


Figure 11. Ukrainian FSI vs Other FSIs

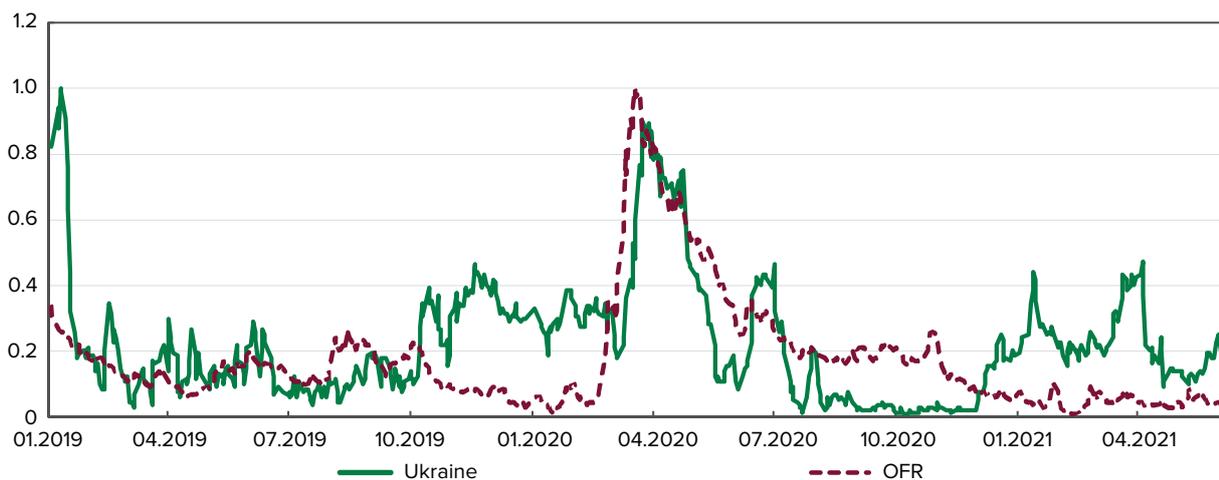


Figure 12. Ukrainian FSI vs OFR FSI

7. CONCLUSIONS AND POLICY RECOMMENDATIONS

In this paper, we build a new FSI for Ukraine, called FSI 2.0, with the objective to improve the performance of this tool, which can be employed as an important element of central bank macroeconomic monitoring system. The new FSI consists of 20 indicators grouped into five subindices (banking, household behavior, government debt, corporate, and FX market) and gives the opportunity to interpret their respective effects. The aggregation of the subindices is based on a dynamic conditional correlation (DCC) multivariate generalized autoregressive conditional heteroscedasticity (MGARCH) model. This methodology implies that the FSI shows significant growth only if several subindices demonstrate growth. In other words, it is not sensitive to one-factor movements.

This new FSI allows policymakers to more accurately assess the level of stress in real-time. In particular, it can be useful for determining anti-crisis policies of the central bank when timely reactions are very important. Currently, the NBU uses the FSI to monitor the ongoing situation due to the COVID-19 restrictive measures and to measure the level of systemic risk in the financial sector, particularly for decision making on FX control measures. From a macroprudential point of view, the FSI may trigger the release of the countercyclical capital buffer.

REFERENCES

- Chatterjee, S., Chiu, C-W., Duprey, T., Hoke, S. (2017). A Financial Stress Index for the United Kingdom. Staff Working Paper, 697. Bank of England. Retrieved from <https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2017/a-financial-stress-index-for-the-united-kingdom.pdf>
- Drakopoulos, D., Petrov, D. (2020). Local stress index for emerging market economies. Global Financial Stability Report, October 2020: Online Annex 2.1. Technical Note, 1-4. International Monetary Fund. Retrieved from <https://www.imf.org/en/Publications/GFSR/Issues/2020/10/13/global-financial-stability-report-october-2020#Chapter2>
- Duprey, T. (2020). Canadian financial stress and macroeconomic conditions. Staff Discussion Paper 2020-4. Ottawa: Bank of Canada. <https://doi.org/10.34989/sdp-2020-4>
- Duprey, T., Klaus, B., Peltonen, T. (2017). Dating systemic financial stress episodes in the EU countries. Journal of Financial Stability, 32, 30-56. <https://doi.org/10.1016/j.jfs.2017.07.004>
- Duprey, T., Ueberfeldt, A. (2018). How to manage macroeconomic and financial stability risks: A new framework. Staff Analytical Note, 2018-11. Ottawa: Bank of Canada. <https://doi.org/10.34989/san-2018-11>
- Haefcke, S., Skarholt, A. (2011). A Swedish financial stress index. Thesis in Finance. Stockholm School of Economics. Retrieved from <http://arc.hhs.se/download.aspx?mediumid=1418>
- Hakkio, C., Keeton, W. (2009). Financial stress: What is it, how can it be measured, and why does it matter? Economic Review, Second Quarter, 5-50. Federal Reserve Bank of Kansas City. Retrieved from https://www.kansascityfed.org/documents/432/PDF-09q2hakkio_keeton.pdf
- Holló, D., Kremer M., Lo Duca, M. (2012). CISS: A composite indicator of systemic stress in the financial system. ECB Working Paper Series, 1426. Retrieved from <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1426.pdf>
- Illing, M., Liu, Y. (2003). An index of financial stress for Canada. Working Paper, 2003-14. Ottawa: Bank of Canada. <https://doi.org/10.34989/swp-2003-14>
- Johansson, T., Bonthron, F. (2013). Further development of the index for financial stress for Sweden. Sveriges Riksbank Economic Review, 1, 1-20. Retrieved from http://archive.riksbank.se/Documents/Rapporter/POV/2013/2013_1/rap_pov_artikel_3_130321_eng.pdf
- Kliesen, K., Smith, D. (2010). Measuring financial market stress. Economic Synopses, 2, 1-2. Federal Reserve Bank of St. Louis. Retrieved from <https://files.stlouisfed.org/files/htdocs/publications/es/10/ES1002.pdf>
- Lang, J. H., Izzo, C., Fahr, S., Ruzicka, J. (2019). Anticipating the bust: a new cyclical systemic risk indicator to assess the likelihood and severity of financial crises. ECB Occasional Paper, 219. Retrieved from <https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op219~7483083881.en.pdf>
- NBU (2016). Financial Stability Report, June 2016. Kyiv: National Bank of Ukraine. Retrieved from https://bank.gov.ua/admin_uploads/article/FSR_062016_eng.pdf
- Oet, M., Eiben, R., Bianco, T., Gramlich, D., Ong, S. (2011). Financial Stress Index: Identification of systemic risk conditions, Working Paper, 11-30. Cleveland: Federal Reserve Bank of Cleveland. Retrieved from <https://www.clevelandfed.org/newsroom-and-events/publications/working-papers/2011-working-papers/wp-1130r3-the-financial-stress-index-identification-of-systemic-risk-conditions.aspx>
- Orskaug, E. (2009). Multivariate DCC-GARCH model – with various error distributions (Master's thesis). Trondheim: Norwegian University of Science and Technology. <http://hdl.handle.net/11250/259296>
- Schwert, G. W. (2011). Stock volatility during the recent financial crisis. European Financial Management, 17(5), 789-805. <https://doi.org/10.1111/j.1468-036X.2011.00620.x>
- Tyshchenko, L., Csajbok, A. (2017). A Financial Stress Index for Ukraine. Visnyk of the National Bank of Ukraine, 240, 5-13. <https://doi.org/10.26531/vnbu2017.240.005>
- Vdovychenko, A., Oros, G. (2015). Financial Stress Index: estimation and application in empirical researches in Ukraine. Journal of European Economy, 14(2).

APPENDIX A

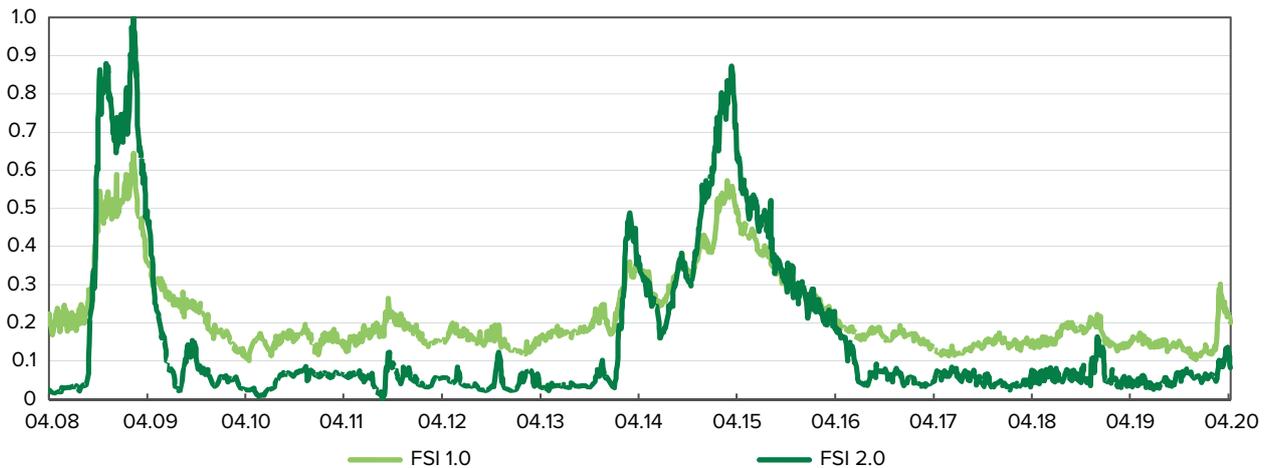


Figure A.1. FSI 1.0 vs FSI 2.0

FSI 1.0 and FSI 2.0 both increase during crisis periods. Moreover, their reactions generally coincide.

However, there are significant differences between the two indices. We can observe that FSI 1.0 stays at a level of approximately 0.2 in normal times, while the FSI 2.0 normal level is only 0.05. There is also a significant difference in magnitudes during crises. The value of FSI 1.0 in crises

is three times higher than in normal times, while FSI 2.0 demonstrates up to a tenfold jump during crises.

This means that by using FSI 2.0, policymakers will receive fewer false signals of crisis in normal times, and it should see an undoubtedly higher level of stress during a crisis.

APPENDIX B

DCC-MGARCH estimates the parameters of dynamic conditional correlation (DCC) multivariate generalized autoregressive conditional heteroscedasticity (MGARCH) models in which the conditional variances are modeled with the use of univariate generalized autoregressive conditional heteroscedasticity (GARCH) models and the conditional covariances are modeled as nonlinear functions of the conditional variances.⁶ The conditional quasi-correlation parameters that weight the nonlinear combinations of the conditional variances follow the GARCH-like process specified in Engle (2002).

The dynamic conditional correlation GARCH model is defined as (Orskaug, 2009):

$$r_t = \mu_t + a_t \quad (1)$$

$$a_t = H_t^{1/2} z_t \quad (2)$$

$$H_t = D_t R_t D_t \quad (3)$$

Notation:

r_t : $n \times 1$ vector of log returns for n assets at time t

a_t : $n \times 1$ vector of mean-corrected returns for n assets at time t ; i.e., $E[a_t] = 0$ and $Cov[a_t] = H_t$

μ_t : $n \times 1$ vector of the expected value of the conditional r_t

H_t : $n \times n$ matrix of conditional variances of a_t at time t

$H_t^{1/2}$: Any $n \times n$ matrix at time t such that H_t is the conditional variance matrix of a_t . $H_t^{1/2}$ may be obtained by a Cholesky factorization of H_t

D_t : $n \times n$, diagonal matrix of conditional standard deviations of a_t at time t

R_t : $n \times n$ conditional correlation matrix of a_t at time t

z_t : $n \times 1$ vector of independent and identically distributed errors such that $E[z_t] = 0$ and $E[z_t z_t^T] = I$.

⁶ <https://www.stata.com/manuals13/tsmgarchdcc.pdf>

APPENDIX C

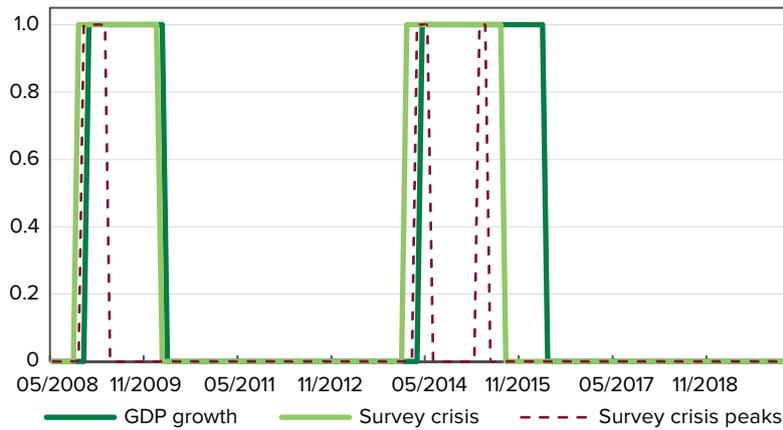


Figure C.1. Dummies for Crises

Figure C.1 shows the values of three dummies. The GDP growth dummy is a proxy for real economic development. We estimate GDP growth on a monthly basis. Values significantly less than 0 are considered a crisis. The survey crisis and survey crisis peaks dummies are proxies for financial sentiments. We calculated these values using a survey of eight Ukrainian financial experts. The GDP growth and survey crisis dummies are similar during the 2008-2009 crisis. However, the survey crisis dummy for 2014-2015 indicates an earlier beginning and end to the crisis. We can

conclude that the GDP growth dummy is somewhat lagged relative to the survey crisis dummy. This is reasonable because financial markets react to shocks rapidly, while the real economy reacts with inertia.

The financial stress index shows stress in the financial system, which is why we used the survey crisis and survey crisis peaks dummies to guide final decisions on the construction of the index.

APPENDIX D

Table D.1. Estimates of the DCC-MGARCH Model

Sample: 4/1/2008 - 4/30/2020, but with gaps

Distribution: Gaussian

Log likelihood = 28,769.66

		b (S.e.)
Bank	cons	0.1964*** (0.001)
ARCH_Bank	arch(-1)	1.0527*** (0.023)
	garch(-1)	0.0079*** (0.002)
	cons	0.0002*** (0.000)
House	cons	0.4152*** (0.003)
ARCH_House	arch(-1)	1.0263*** (0.246)
	garch(-1)	0.0224*** (0.006)
	cons	0.0006*** (0.000)
Gov	cons	0.1256*** (0.001)
ARCH_Gov	arch(-1)	1.0923*** (0.024)
	garch(-1)	0.0022* (0.001)
	cons	0.0002*** (0.000)
Corp	cons	0.1258*** (0.002)
ARCH_Corp	arch(-1)	1.0650*** (0.023)
	garch(-1)	-0.0007 (0.001)
	cons	0.0002*** (0.000)
FX	cons	0.2479*** (0.001)
ARCH_FX	arch(-1)	1.0030*** (0.233)
	garch(-1)	0.0074 (0.005)
	cons	0.0007*** (0.000)

Table D.1 (continued). Estimates of the DCC-MGARCH Model

	corr(Bank,House)	-0.0352 (0.046)
	corr(Bank,Gov)	0.2313*** (0.044)
	corr(Bank,Corp)	0.3036*** (0.041)
	corr(Bank,FX)	0.1375*** (0.043)
	corr(House,Gov)	-0.2619*** (0.052)
	corr(House,Corp)	-0.1233*** (0.047)
	corr(House,FX)	0.0637 (0.047)
	corr(Gov,Corp)	0.4522*** (0.033)
	corr(Gov,FX)	0.1106** (0.050)
	corr(Corp,FX)	0.1053** (0.046)
Adjustment	lambda1	0.3018*** (0.007)
	lambda2	0.6908*** (0.007)
N		2,988

Note: Standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.