National Bank
of Ukraine

VISNYK

OF THE NATIONAL BANK OF UKRAINE

VISNYK

OF THE NATIONAL BANK OF UKRAINE

No. 251
2021

PREFACE BY THE EDITOR-IN-CHIEF

Dear readers,

This edition of the *Visnyk of the National Bank of Ukraine* highlights subjects related to financial stability and central bank forecasting. Our authors examine the determinants of corporate loan pricing, improve existing approaches to more accurate measurement of financial stress, and hone macroeconomic forecasting tools.

The first article in the issue, *Determinants of Corporate Loan Interest Rate: Case of Ukraine*, by Solomiya Shpak, estimates the effects of loan, borrower, and bank characteristics on corporate loan pricing in Ukraine, using rich loan-borrower-bank monthly panel data for the period from 2013 to 2020. Examining an extensive set of fixed effects and determinants that might affect loan pricing, the author concludes that both loan and borrower characteristics are significant when estimating the effects of banks' health on the loan interest rate. Specifically, the study finds that larger borrowers with lower indebtedness who operate in concentrated industries get lower interest rates. In addition, interest rates are reduced if loans have shorter maturity and are secured by larger collateral value.

The second article, *A BVAR Model for Forecasting Ukrainian Inflation and GDP*, by Nadiia Shapovalenko, deals with the forecasting performance of a Bayesian Vector Autoregression (BVAR) model with a steady-state prior and compares the accuracy of the forecasts against the QPM and official NBU forecasts during the 2016–2020 period. Inflation forecasts produced by the BVAR model are found to be more accurate than those of the QPM model for two quarters ahead and are competitive for a longer time horizon. They also outperform the NBU's official inflation forecasts over the monetary policy horizon. Furthermore, the BVAR forecasts appeared to have shown better outcomes compared to the GDP growth forecasts based on the QPM for the whole forecast horizon.

In the third article, *A New Financial Stress Index for Ukraine*, Vladyslav Filatov considers methodology updates for calculating the financial stress index (FSI) for Ukraine. The author improves the FSI for Ukraine developed recently by Lesia Tyshchenko and Attila Csajbok. This new index could be more attractive for policymakers because it peaks during periods of crisis that are in line with the consensus of financial experts and outperforms the previously used index. Therefore, the suggested index is recommended by the author as a monitoring tool for the NBU's macroprudential policy.

The papers included in the issue offer important practical solutions that could be useful for economic agents and policymakers. We invite researchers and scholars to continue the discussion that started here, or to initiate own research in economics, finance, and banking. You are welcome to submit your original fundamental and applied studies for publication in the *Visnyk of the National Bank of Ukraine*. We are looking forward to cooperation.

Best regards,
Mihnea Constantinescu

DETERMINANTS OF CORPORATE LOAN INTEREST RATE: CASE OF UKRAINE

SOLOMIYA SHPAK[ab1]

[a]National Bank of Ukraine
[b]Kyiv School of Economics
 E-mail: sshpak@kse.org.ua

Abstract	This paper estimates the effect of loan, borrower, and bank characteristics on corporate loan pricing in Ukraine using rich loan-borrower-bank monthly panel data from 2013 and 2020 combined with data from borrowers' financial statements. Examining an extensive set of fixed effects, we find that larger loans, loans with a shorter maturity period and larger collateral value have lower interest rates even after controlling for borrower characteristics. We also find that larger borrowers, borrowers with more tangible assets, lower indebtedness, and a higher interest coverage ratio who operate in concentrated industries secure lower interest rates. Our findings suggest that it is crucial to take into consideration both loan and borrower characteristics when estimating the effects of banks' health on the loan interest rate.
JEL Codes	G21, E51, L11, P34
Keywords	credit supply, credit demand, cost of debt, Ukraine

1. INTRODUCTION

Banks are essential sources of external financing in many countries. Evidence shows that bank loan markets account for a larger share of external financing than equity or bond markets have in most economies (e.g. Drucker and Puri, 2007; Bae and Goyal, 2009). Therefore, it is of great academic and financial stability policy interest to investigate the factors that affect bank loan pricing. Discerning firm and bank determinants of loan rates is crucial for informing the public about policy decisions, given the importance of lending for financial stability and economic growth.

Bank loan financing in Ukraine has been slow. In recent years, demand from the corporate sector has been limited, and many potential borrowers were not ready to ensure the completeness and quality of information disclosure (Financial Stability Report, 2020). At the same time, the Ukrainian government has been calling for the resumption of large-scale loan financing for businesses. There have been extensive discussions about the high cost of borrowing for the corporate sector. Indeed, the Ukrainian economy and businesses need resources, and Ukrainian banks' liquidity is enough to meet the demand (Financial Stability Report, 2020). But among other factors, the ownership structure and financial reporting of potential borrowers are often nontransparent, preventing banks from financing Ukrainian businesses.

This study examines the factors that affect the cost of loans for Ukrainian businesses. We aim to establish the relative importance of firm-specific risk factors and bank-level characteristics and the effect of the length of a bank-firm lending relationship. Using matched data on firms, loans, and banks between 2013 and 2020, we can control for unobserved firm and bank characteristics to discern the unbiased estimates of the variables of interest.

The goal of this study is to answer the following research questions:

1. What are the firm-level determinants of loan prices?
2. Do bank-level characteristics affect the price of lending?
3. Do weak firms borrow from weak banks?

2. RELATED LITERATURE AND HYPOTHESES

We start with the observation that the loan interest rate charged by a bank to a borrower should reflect the borrower's risk characteristics and the bank's cost of funds. At the beginning of a lending relationship, the bank assesses the borrower's risks characteristics affecting the interest rate for the borrower. There is extensive literature showing that higher interest rates are charged to riskier firms such as smaller firms, those with low levels of tangible assets, less profitable businesses, and those with assets that have high information costs (see, for example, Strahan, 1999).

In this study, we explore 20 firm-level characteristics that might potentially affect the cost of borrowing for businesses. We can group these borrower characteristics observed by the bank in four categories: borrower size, borrower profitability, borrower indebtedness, and other characteristics. We

hypothesize that *larger firms* tend to pay lower interest rates as they encounter less information asymmetry on the credit markets than small firms do. Also, larger firms usually have longer track records and are followed by more financial analysts. We also expect more *profitable* firms to pay lower interest rates as they might have lower probabilities of default compared to smaller firms. One of the indicators of the borrower's *indebtedness* is the borrower's observable default risk. Other relevant borrower characteristics include such variables as *tangible assets* where we hypothesize that firms with more tangible assets are likely to secure lower interest rates as more tangible assets may offer higher recovery values in default states.

Another important empirical question is whether the borrower's *industry* has any influence on the cost of borrowing. Our hypothesis comes from Valta (2012), who showed that banks charge significantly higher loan spreads to U.S. publicly traded firms in industries with the high product-level competition. We are interested in verifying whether the same correlation exists in Ukraine: firms in more competitive environments might face a higher interest rate because more competition might mean a higher likelihood of defaults on interest payments. Also, competition affects a firm's liquidation value. We use the Herfindahl-Hirschman Index (HHI) for every NACE 2-digit industry to measure its competitiveness.

However, banks can only partly monitor firms' characteristics, which causes information asymmetry. This may include adverse selection and a moral hazard problem. Banks might use different ways to handle information problems, and one of them is through *repeat lending*. There is extensive literature on the importance of firm-bank relationship for credit access for small borrowers (Petersen and Rajan, 1994; Berger and Udell, 1995; Chakraborty et al., 2010) as well as large publicly traded borrowers (Ivashina and Kovner, 2011; Karolyi, 2018). As a result of multiple interactions with a borrower, a bank learns private information about the borrower. As the literature suggests, benefits of repeat lending might be realized through "the ability to share sensitive information (Bhattacharya and Chiesa, 1995), more flexible contracts compared to public debt (Berlin and Mester, 1992), the ability to monitor collateral (Rajan and Winton, 1995), and the ability to smooth out loan pricing over multiple loans (Berlin and Mester, 1999)."

When studying determinants of loan interest rates, it is also essential to consider *loan-specific characteristics* that, according to the literature, affect interest rates (Graham, Li, and Qiu, 2008). They include loan size, loan maturity, the currency of a loan, and the characteristics of collateral. There might be economies of scale in bank lending, so that loan size is likely to be associated with lower interest rates. Loan maturity is expected to be associated with a lower interest rate as banks face greater uncertainty and higher credit risk in loans with long maturities. Collateral might be used in loan contracts for two reasons: adverse selection and moral hazard. In the adverse selection models, collateral may signal a better-quality borrower, suggesting that better borrowers post collateral to obtain lower interest rates on the loan (Bester, 1985; Besanko and Thakor, 1987). On the other hand, the moral hazard model suggests banks might require riskier borrowers to post collateral to compensate for possible risk of nonrepayment and increase incentives for monitoring (Berger and Udell, 1990; Jimenez et al., 2006; Francis et al., 2012).

We are further interested to know whether *bank-level characteristics* affect loan costs, keeping borrower characteristics constant. Specifically, we are interested to see whether firms that borrow from banks with weak capital face a higher cost of debt. Our key regressors are indicators of banks' strength. The main measure of bank strength is the regulatory capital ratio. Other bank characteristics may be relevant for lending. For this reason, we follow the literature when defining a set of bank controls for lending regressions such as bank size, several measures of bank profitability and effectiveness as well as bank liquidity (Khwaja and Mian, 2008; Iyer et al., 2014).

There is also evidence from the literature about the *sorting of firms according to banks*. For example, previous research shows that foreign banks tend to lend to transparent, large, and less risky borrowers and offer them lower lending rates. Dell'Ariccia and Marquez (2004) also suggest that foreign banks face considerable information disadvantages and target more transparent clients (relying on transaction-based lending). In contrast, domestic banks tend to lend to firms based on soft information (relationship lending) (Althammer and Haselmann, 2011). More recent studies such as Michelangeli et al. (2020) find evidence of borrower-lender assortative matching where safer banks have more credit relations with less risky firms.

In this study, we first examine whether borrower-lender assortative matching exists in the Ukrainian economy; if present, we document and quantify it. We further employ methodology (described in Section 3) to examine possible firm-bank matching when studying the relative importance of firm- and bank-level characteristics.

3. DATA AND METHODOLOGY

This paper uses data from Form No. 613 combined with the information on firm performance from the balance sheet, report on financial results, and National Bank of Ukraine supervisory statistics on bank performance. Our goal is to create monthly panel data of firm-loan-bank relationships between 2013 and 2020.

Bank-borrower-loan-level data come from Form No. 613 *Report on Risk Concentration under Bank Exposures to Counterparties and Insiders*, which is submitted by banks to the National Bank of Ukraine monthly. In this form, the banks list active operations for all the borrowers for which the total amount of all claims of a bank and financial liabilities is UAH 2 million, or more. The data contain information on the loan amount, maturity, currency of the loan, and loan terms, among others. The unit of observation is loan contract l of firm i at bank j month-year t. We restrict the sample to new loans only as we are primarily interested in the cost of new loans for businesses. We classify a loan as a new loan on a specific date if it is the first month when the loan appears in Form No. 613.

Bank-level data come from the National Bank of Ukraine supervisory statistics on bank performance available quarterly from 2013 to 2020. We match each month-year from Form No. 613 to the corresponding quarter-year in the bank-level data. We restrict banks only to those that were solvent as of January 2020 and drop PrivatBank as the inclusion of insolvent banks might distort the results given the poor quality of their reporting. Out of 185 ever registered banks in Ukraine, a mere 71 banks make it to our sample. Firm-level data come from the balance sheet and financial results report

annually from 2013 to 2019. For each firm, we take the values of economic variables as of the beginning of each year.

In total, our final sample consists of 141,525 new loan contracts corresponding to 13,612 distinct firms taking loans from 71 banks between 2013 and 2020.

Methodology and Variables

To test our hypotheses about the determinants of corporate loan pricing, we estimate the following model:

$$Interest_{ijlt} = \alpha_i + \beta_j + \kappa_t + \gamma X_{ilt} + \lambda R_{it-1} + \delta B_{jt} + \epsilon_{ijlt}, \qquad (1)$$

where i, j and l index borrowers, banks, and loans, respectively, while t indicates month-year (e.g. January 2020). *Interest* is a natural logarithm of the interest rate charged by the bank j on the loan l for the borrower i. X represents nonprice loan characteristics such as currency of the loan, maturity, and loan size; R represents observed borrower risk characteristics, and B denotes bank-level characteristics that might affect the cost of loans for businesses. Borrower characteristics are measured as of the previous year to mitigate the possible impact of reverse causality. Both groups of variables are described below in detail. We also include month-year controls κ_t to consider aggregate shocks that affect all banks in month-year t. These include changes in the key policy rate, changes in a macroeconomic situation such as inflation, economic downturns as well as seasonal changes in interest rates, among others. We also control for NACE 2-digit industry-specific controls β_j capturing industry variation in the cost of loans. These unique data on loan-borrower-bank relationships on a monthly basis allow us to include these fixed effects and compare borrowers operating in the same 2-digit industry and receiving loan in the same month and year.

The estimation of (1) using the Ordinary Least Squares (OLS) model is based on the assumption that there is no correlation between observed loan, bank, and borrower characteristics and other factors that affect loan prices represented by the error term ϵ. The problem with this assumption is that unobserved borrower heterogeneity may introduce a nonzero correlation between the error term and the right-hand-side variables and lead to at least two estimation problems. First, if there is the correlation between the error term and loan characteristics X when, for example, a bank grants better loan terms to better firms, the estimates of loan characteristics $\hat{\gamma}$ will be biased upward. Second, sorting borrowers among banks according to private information might bias the estimate of bank effects $\hat{\delta}$ upward. Examples of such assortative borrower-bank matching include cases when firms with high unobserved risk tend to borrow from weak banks (Dell'Ariccia and Marquez, 2004), and foreign banks tend to lend to more transparent firms (Michelangeli et al., 2020). The richness of our data allows us to address potential unobserved borrower heterogeneity by the inclusion of borrower-level fixed effects α_i in some specifications.

Our loan-level characteristics X include loan size measured in the UAH equivalent, collateral value measured in the UAH equivalent, maturity measured in months, the loan currency (a dummy for the USD, euro, and other currency with loans in hryvnias as a base category), and the loan interest rate. All loan level characteristics are expressed as natural logarithms. Following Hasan et al. (2012), Francis et al. (2012), and others, we explore several firm characteristics that may affect the price of corporate loans in our analysis. These firm characteristics belong to four broad groups: borrower size, profitability, indebtedness, and other variables. Although we start with 20 borrower-level variables, our final regressions include only five of them as many of these variables are highly correlated and capture similar aspects of borrower risk. For example, initially, we look at five measures of borrower size such as total assets, revenue, gross profit, EBIT, and EBITDA, and only revenue makes it to the final regression. Similarly, we look at five measures of firm indebtedness, including net debt-to-asset, net debt-to-revenue, net debt-to-EBIT, net debt-to-EBITDA, and interest coverage ratios, but only net debt-to-EBIT and interest coverage ratios are included in the final estimation. Table A1 illustrates the correlations between 20 borrower characteristics initially explored.

In this paper, we measure the effect of firm size by *Log (Revenue)*, the natural logarithm of the firm's revenue in the previous year *t-1*. Previous literature shows that larger firms tend to secure lower interest rates as they suffer less from information asymmetries on the credit markets. Therefore, we expect to find a negative relationship between firm size and the interest rate. We proxy firm indebtedness with the *Interest Coverage Ratio* (ICR) measured as the ratio of EBIT to the net financial cost of the firm. Larger ICR values correspond to lower default risk, so we expect that borrowers with a higher ICR indicator will have lower interest rates. We also consider the borrower's *Tangibility* defined as the ratio of net property, plant, and equipment to total assets. We expect that firms with a higher share of tangible assets will enjoy a lower cost of borrowing, as tangible assets serve as a primary source of collateral thus are associated with the lower cost of financing. We also monitor the profitability of borrowers defined as the EBIT-to-revenue ratio. Finally, we include the *Current Ratio* (the ratio of current assets to current liabilities), which measures the borrower's ability to pay short-term liabilities with its short-term assets such as cash, inventory, and receivables. We expect that more profitable firms and firms with larger current ratios will have lower interest rates on corporate loans.

Following the literature, our regressions include bank characteristics that might affect the cost of credit. The first two measures – return on assets (ROA) and return on equity (ROE) represent bank profitability. We expect that borrowers who take loans at the more profitable banks will enjoy lower interest rates. We also consider regulatory capital ratio (H2) and liquidity ratios (H4, H5, H6).[2] Since there is a high correlation between some of the bank-level measures, we include only some of them in the final regressions. The share of nonperforming loans (NPLs) is another determinant of the corporate interest rate. It is important to note that the definition of NPLs changed during the sample period. Specifically, between 1 January 2013 and February 1, 2017, NPLs are measured by nonperforming exposures – exposures with payments past due 90+ days; individual exposures past due 30+ days with low counterparty financial class.[3] Starting 1 February 2017, NPLs are defaulted loans where default is determined by the fact of payments on assets past due 90+ days, or the inability of the borrower to repay the debt without repossession of collateral.[4]

Before estimating determinants of corporate loan pricing, we document patterns for loan, borrower, and bank characteristics by the borrowers' size group. Using the definition of firm size provided in Article 55 of the Commercial Code of Ukraine, we split all borrowers into four groups based on their revenue measured in the euro equivalent. Table 1 shows that most of the new loans between 2013 and 2019 were taken by firms with revenue between EUR 2 million and EUR 50 million. The largest firms (whose revenue exceeds EUR 50 million) enjoyed the most loan contracts per firm: around 15 contracts compared to only two contracts per firm for the smallest size group. On average, larger firms had lower interest rates and took loans that were larger in size and shorter in maturity. As we move from the smallest to the larger size group, firms tend to have a higher level of indebtedness (measured by the ratio of net debt to EBIT) and tangibility.

4. REGRESSION RESULTS

We start with using loan-level data to investigate loan-level determinants of corporate loan prices. The first three columns of Table 3 represent the results of the whole sample and illustrate that all loan characteristics have expected signs and are significant at the 1% level. These results show that larger loans face lower interest rates while loans with longer maturity tend to have a higher interest rate. This finding is consistent with Graham, Li, and Qiu (2008) and might indicate that banks face greater uncertainty and higher credit risk in loans with long maturities. Collateral value also has positive association with the interest rate implying that larger collateral might be used by Ukrainian banks to mitigate a moral hazard problem as suggested by the previous literature (Berger and Udell, 1990; Jimenez et al., 2006; Francis et al., 2012). As expected, currency matters for the pricing of loans, where loans in euros have the lowest interest rate, followed by loans

Table 1. Average Loan and Firm Characteristics by Borrower Size

Firm size		Loan			Borrower		
	No. of loan contracts	Interest rate	Loan size	Maturity	Net debt/ EBIT	Tangibility	Prior relationship
<=2 mln	41,075	18.65	189.20	19.54	2.55	0.11	0
<=10 mln	48,310	18.15	183.30	13.15	2.64	0.17	1
<=50 mln	47,525	17.58	311.20	8.52	3.43	0.18	2
>50 mln	23,524	15.32	1,737.10	5.85	4.12	0.21	2

Note: Firm size, loan size, and net assets measured in the EUR equivalent for a given year.

Table 2 illustrates that, on average, larger firms tend to borrow from larger banks and banks with lower ROA and ROE. In addition, there is some evidence of the relationship between firm size and the capital adequacy ratio (H2) and liquidity ratios (H4 and H6): larger firms tend to borrow from banks with lower values of all three ratios; however, this does not hold for the largest size group.

denominated in U.S. dollars compared to loans in domestic currency. Depending on the specification, we probe into month-year and industry-related fixed effects.

Although these fixed effects change loan characteristics, making them smaller in magnitude, they have consistent signs and significance across the specifications. The estimate of the number of prior relations with a bank in our

Table 2. Average Bank Characteristics by Borrower Size

Firm size	Bank net assets	Bank ROA	Bank ROE	H2	H4	H6
<=2 mln	1,706	2.13	12.91	19.23	52.88	93.43
<=10 mln	1,845	1.56	6.98	18.93	50.14	91.38
<=50 mln	1,876	1.14	3.34	18.81	51.32	89.72
>50 mln	2,150	0.93	2.02	19.84	53.58	89.83

Note: Firm size, loan size, and net assets measured in the euro equivalent for a given year.

These numbers provide preliminary evidence of borrower-lender assortative matching that we will consider when estimating our regression models. However, this sorting is based on observable characteristics only and does not capture sorting that might arise from the matching of borrowers by their private information not observable to the bank. Also, these tables do not take into account the effects of macroeconomic conditions and other shocks that might partly drive this sorting. We will take into consideration both unobservable borrower characteristics and time effects when testing our hypothesis in the formal regression setting.

preferred specification presented in Column 3 suggests that on average, a history of lending relationships does not affect the interest rate. If we split the sample into two subsamples based on firm size, we find that number of prior relations is negatively associated with the interest rate charged to small firms while the effect for large firms is not significant. This finding is in line with the previous literature on the importance of the firm-bank relationship for credit access for small borrowers (Petersen and Rajan, 1994; Berger and Udell, 1995; Chakraborty et al., 2010). The results also suggest that collateral value has larger positive association with the interest rate among small firms compared to large firms suggesting that collateral might be used as insurance against a moral hazard problem among small firms to the larger extent than among large firms.

Table 3. Loan-Level Determinants of Loan Pricing

	(1)	(2)	(3)	(4)	(5)
	All borrowers			Large and medium borrowers	Small borrowers
Loan size	-0.049***	-0.027***	-0.025***	-0.018***	-0.024***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Maturity	0.033***	0.027***	0.029***	0.032***	0.019***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Collateral value	0.042***	0.017***	0.016***	0.012***	0.022***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
USD	-0.836***	-0.797***	-0.806***	-0.768***	-0.853***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.007)
Euro	-1.010***	-0.967***	-0.983***	-0.948***	-1.019***
	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)
Other currency	-0.238***	-0.161***	-0.178***	-0.117***	-0.295***
	(0.028)	(0.026)	(0.027)	(0.030)	(0.057)
Prior relations	0.028***	0.003***	-0.001	0.001	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Month-year FE	No	Yes	Yes	Yes	Yes
Industry FE	No	No	Yes	Yes	Yes
N	141,525	141,525	141,523	71,873	69,650
R-squared	0.302	0.389	0.407	0.474	0.357

Note: Loan size, maturity, and collateral value are measured in natural logarithms. USD, Euro, and Other currency are dummies for currency of the loan, the Prior relations parameter measures the number of loan contracts in previous five years at a particular bank, Month-year Fixed Effects (FE) is a dummy for month and year of a loan contract. Standard errors in parentheses: $p<0.1$; $p<0.05$; $p<0.01$.

Although the regressions presented in Table 3 take into account a rich set of fixed effects, they do not consider the characteristics of the borrowers that take these loans. Excluding the borrowers' characteristics might be problematic if there is a selection of loans by the borrowers, as illustrated in Table 1, where large borrowers tend to take larger loans and those with shorter maturity. If we do not take this factor into consideration, our estimates of the loan characteristics will be biased. Indeed, as Table 4 shows, the inclusion of observable borrower characteristics lowers the magnitude of loan size, maturity, and currency coefficients across all specifications in the full sample. The results in our preferred specification (Column 3) suggest that doubling loan size is associated with a reduction in the interest rate by 1.5%.

Meanwhile, a twofold increase in maturity is associated on average with an interest rate increase of 2.3%. This implies that for an average loan in the sample with 12-month maturity and an average interest rate of 14.8%, our results suggest that an increase in maturity from 12 to 24 months will lead to an increase in the interest rate to 15.1% depending on industry-related and month-year effects.

Holding all other variables constant, the results suggest that compared to firms taking loans in domestic currency, firms receiving loans in U.S. dollars and in euros secure 79.4% and 97.8% lower interest rates, respectively. At the mean, this implies that compared to the cost of a loan in the

Ukrainian hryvnia set at 14.8%, a corresponding U.S.-dollar loan of the same size taken by a borrower with the same observable characteristics in the same month-year will have an interest rate of 3.05%.

All borrower characteristics in Table 4 have expected signs consistent with the findings of recent studies such as Hale and Santos (2009) and Hasan et al. (2012). We find that the coefficient of borrower revenue is -0.025 and is significant at the 1% level, indicating that a 10% increase in the revenue translates into a 0.25% decrease in the interest rate. The current ratio and tangibility are also negatively associated with the interest rate: a one-unit increase in tangibility (from 0 to 1) is associated with a 4% reduction in the interest rate, while a one-unit increase in the current ratio is associated with a mere 0.1% decrease in the cost of corporate loans. Firm indebtedness as measured by the net debt-to-EBITA ratio is positively associated with the interest rate: an increase in firm indebtedness by two standard deviations (10 units) increases the interest rate by 0.3%. Finally, the HHI coefficient suggests that loans taken by firms which operate in less competitive industries bear a lower interest rate.[5] This result is consistent with Valta (2012) and indicates that Ukrainian firms in more competitive industries face a higher interest rate because more competition might mean higher default risk for interest payments.

The estimation results of large-/medium- and small-sized firms' samples suggest that both revenue and tangibility

---

[5] We focus on Column 2 to interpret HHI, as the effect indicated in Column 3 is absorbed in industry-year interactions partly capturing industry concentration in a given year.

Table 4. Loan and Borrower Determinants of Loan Pricing

	(1)	(2)	(3)	(4)	(5)	(6)
	All borrowers			Large and medium borrowers		Small borrowers
Loan size	-0.027***	-0.017***	-0.015***	-0.007***	-0.006***	-0.021***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Maturity	0.020***	0.019***	0.023***	0.030***	0.034***	0.017***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Collateral value	0.036***	0.017***	0.016***	0.010***	0.002	0.022***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
USD	-0.814***	-0.784***	-0.794***	-0.751***	-0.743***	-0.861***
	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)	(0.008)
Euro	-0.997***	-0.957***	-0.978***	-0.947***	-0.926***	-1.018***
	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.009)
Other currency	-0.185***	-0.126***	-0.148***	-0.087***	-0.084***	-0.294***
	(0.028)	(0.026)	(0.027)	(0.030)	(0.029)	(0.059)
Prior relations	0.035***	0.010***	0.005***	0.006***	0.004***	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Revenue	-0.039***	-0.024***	-0.025***	-0.034***	-0.033***	-0.014***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Current ratio	-0.002***	-0.002***	-0.001***	0.003***	0.003***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
Tangibility	-0.100***	-0.064***	-0.044***	-0.071***	-0.091***	-0.026***
	(0.007)	(0.006)	(0.007)	(0.011)	(0.011)	(0.010)
Net debt/EBIT	0.002***	0.001***	0.000*	-0.001***		0.002***
	(0.000)	(0.000)	(0.000)	(0.000)		(0.000)
ICR					-0.002***	
					(0.000)	
HHI	0.263***	-0.150***	0.002	-0.029	-0.080***	0.023
	(0.015)	(0.015)	(0.019)	(0.023)	(0.023)	(0.034)
Month-year FE	No	Yes	Yes	Yes	Yes	Yes
Industry FE	No	No	Yes	Yes	Yes	Yes
N	138,722	138,722	138,720	70,504	64,053	68,216
R-squared	0.313	0.390	0.408	0.476	0.510	0.356

Note: Loan size, maturity, and collateral value are measured in natural logarithms. USD, Euro, and Other currency are dummies for currency of the loan, the Prior relations parameter measures the number of loan contracts in previous five years at a particular bank, Month-year FE is a dummy for month and year of a loan contract. All firm-level controls are estimated in the year prior to the loan's initiation. Revenue is measured as a natural logarithm of revenue, all other variables are measured in absolute terms. Standard errors in parentheses: *p<0.1; **p<0.05; ***p<0.01.

matter more among large- and medium-sized firms. At the same time, indebtedness is a more crucial determinant among small firms if we compare the results with those of the full sample. Interestingly, firm indebtedness as measured by a net debt-to-EBIT ratio has a negative sign for large firms: one unit increase in the net debt-to-EBIT ratio is associated with a 0.1% decrease in the interest rate. This means that large firms that have higher debts enjoy a lower interest rate compared to similar firms with a lower net debt-to-EBIT[6] ratio. If we measure indebtedness with the interest coverage ratio, we find that one unit increase in the ICR is associated with a 0.2% decrease in the average interest rate. This means that

firms with larger EBITDA-to-financial-cost ratios enjoy, on average, a lower interest rate.[7] Also, industry concentration does not have any effect on the interest rate charged to small firms.

Our final empirical question is about the extent to which bank characteristics affect the prices of a loan if we take into account the loan and borrower characteristics.[8] To illustrate the importance of examining the borrower and loan composition, we start our estimation with bank characteristics only. The results of this estimation are reported in Column 1 of Table 5 and suggest that with due regard for month-year

---

[6] Results on indebtedness hold if we alternatively use debt-to-EBITDA ratios available for large- and medium-firms only.
[7] Because results on two indebtedness measures for medium/large firms are not consistent, further research is needed to understand the mechanism of how they affect the corporate loan interest rate in Ukraine.
[8] Note that these regressions are based on the Q1 2013-Q3 2019 sample as NBU ceased calculation and publication of H4 in September 2019 (this liquidity ratio was replaced by the ICR) and we thus cannot estimate regressions for this period consistently.

Table 5. Bank Determinants of Loan Pricing

	(1)	(2)	(3)	(4)
Bank ROA	-0.005***	-0.006***	-0.005***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
Liquidity ratio	-0.101***	-0.016***	-0.004	0.007
	(0.005)	(0.004)	(0.004)	(0.004)
Bank assets	-0.028***	-0.039***	-0.039***	-0.024***
	(0.002)	(0.001)	(0.001)	(0.002)
% of NPLs	0.379***	0.213***	0.210***	0.102***
	(0.013)	(0.011)	(0.011)	(0.014)
Loan size		-0.029***	-0.018***	0.003**
		(0.001)	(0.001)	(0.001)
Maturity		0.030***	0.026***	0.028***
		(0.001)	(0.001)	(0.001)
Collateral value		0.020***	0.019***	0.003**
		(0.001)	(0.001)	(0.001)
USD		-0.778***	-0.777***	-0.741***
		(0.004)	(0.004)	(0.005)
Euro		-0.956***	-0.976***	-0.891***
		(0.006)	(0.006)	(0.007)
Other currency		-0.142***	-0.129***	-0.175***
		(0.023)	(0.024)	(0.023)
Prior relations			0.007***	0.025***
			(0.001)	(0.001)
Revenue			-0.021***	-0.003
			(0.001)	(0.003)
Net debt/EBIT			-0.000	0.001**
			(0.000)	(0.000)
Current ratio			-0.000	-0.000
			(0.000)	(0.000)
Tangibility			-0.007	0.115***
			(0.006)	(0.012)
HHI			0.034**	-0.093***
			(0.017)	(0.015)
Month-year FE	Yes	Yes	Yes	Yes
Industry FE	No	No	Yes	Yes
Firm FE	No	No	No	Yes
N	144,969	124,251	121,479	121,481
R-squared	0.140	0.438	0.461	0.368

Note: Loan size, maturity, and collateral value are measured in natural logarithms. USD, Euro, and Other currency are dummies for the currency of the loan, the Prior relations parameter measures the number of loan contracts in previous five years at a particular bank, Month-year FE is a dummy for month and year of a loan contract. All firm-level controls are estimated in the year prior to the loan's initiation. Revenue is measured as a natural logarithm of revenue, all other variables are measured in absolute terms. A bank's ROA is measured in absolute terms, a liquidity ratio (H4) and bank assets are measured in logs. Percent of NPLs is measured as a share of NPLs in total loans issued to legal entities and individuals. Standard errors in parentheses: *p<0.1; **p<0.05; ***p<0.01.

fixed effects, loans initiated at banks with lower ROA were associated with a lower interest rate. Also, stronger banks, as measured by the liquidity ratio, offered lower interest rates than those with lower values of these requirements. Loans initiated at large banks as measured by net bank assets and banks with lower NPL ratios were associated with a lower interest rate.

In Columns 2 and 3, we introduce controls for loan and borrower compositions of the banks, respectively. All coefficients stay significant and have comparable magnitude except for the liquidity ratio: its magnitude decreases from -0.016 to -0.005 and becomes insignificant. A slight change in other coefficients confirms the direction of the selection presented in Tables 1 and 2. Finally, if we expand

unobservable borrower characteristics by adding borrower-level fixed effects, we see that the magnitude of all bank controls decreases in absolute terms. It is important to note that a bank's effect is determined by the change of the bank's characteristics within the same firm. These results suggest that once we include loan characteristics, observable borrower characteristics and unobservable borrower attributes that are fixed over time, the effect of bank controls decreases and even disappears (for the liquidity ratio). The interpretation of the bank controls suggests that an increase in bank ROA by one unit is associated with a decrease in the interest rate by 0.3%. Meanwhile, doubling net bank assets leads to a reduction in the interest rate by 2.4%. A change of the NPL ratio from 0 to 1 (change from the minimum to maximum value in an extreme case) increases the corporate interest rate by 10%.[9]

5. CONCLUSIONS

This paper estimates the effect of loan, borrower, and bank characteristics on the corporate loan prices in Ukraine using rich loan-borrower-bank monthly panel data from 2013 and 2020 combined with the data from borrowers' financial statements. Examining an extensive set of fixed effects, we find that larger loans, loans with a shorter maturity period and larger collateral value have lower interest rates even after controlling for borrower characteristics. We also find that larger borrowers, borrowers with more tangible assets, lower indebtedness, and a higher interest coverage ratio who operate in concentrated industries secure lower interest rates. Our results suggest some preliminary evidence of the role of repeat lending for small borrowers. We also estimate the role of banks' health in the cost of corporate loans. Our findings suggest that it is crucial to control loan and borrower characteristics when estimating the effects of banks' health on the loan interest rate. We find that larger, more profitable banks and those with a smaller share of NPLs tend to offer lower interest rates even when we look into the loan and borrower composition of a particular bank.

---

[9] These results are also robust to the inclusion of a bank's cost-to-income-ratio (CIR) as an additional bank-level determinant of a corporate loan interest rate. The sign of CIR is as expected, but this coefficient is not significant across the specifications.

REFERENCES

Althammer, W., Haselmann, R. (2011). Explaining foreign bank entrance in emerging markets. Journal of Comparative Economics, 39(4), 486-498. https://doi.org/10.1016/j.jce.2011.03.002

Berger, A. N., Udell, G. F. (1990). Collateral, loan quality and bank risk. Journal of Monetary Economics, 25(1), 21-42. https://doi.org/10.1016/0304-3932(90)90042-3

Berlin, M., Mester, L. J. (1992). Debt covenants and renegotiation. Journal of Financial Intermediation, 2(2). 95-133. https://doi.org/10.1016/1042-9573(92)90005-X

Berlin, M., Mester, L. J. (1999). Deposits and relationship lending. The Review of Financial Studies, 12(3), 579-607. https://doi.org/10.1093/revfin/12.3.0579

Besanko, D., Thakor, A. V. (1987). Collateral and rationing: sorting equilibria in monopolistic and competitive credit markets. International economic review, 28(3), 671-689. https://doi.org/10.2307/2526573

Bester, H. (1985). Screening vs. rationing in credit markets with imperfect information. The American economic review, 75(4), 850-855. https://doi.org/10.1006/jfin.1995.1014

Bhattacharya, S., Chiesa, G. (1995). Proprietary information, financial intermediation, and research incentives. Journal of Financial Intermediation, 4(4), 328-357.

Botsch, M., Vanasco, V. (2019). Learning by lending. Journal of Financial Intermediation, 37, 1-14. https://doi.org/10.1016/j.jfi.2018.03.002

Dell'Ariccia, G., Marquez, R. (2004). Information and bank credit allocation. Journal of Financial Economics, 72(1), 185-214. https://doi.org/10.1016/S0304-405X(03)00210-1

Francis, J., LaFond, R., Olsson, P. M., Schipper, K. (2004). Costs of equity and earnings attributes. The Accounting Review, 79(4). 967-1010. https://www.jstor.org/stable/4093083

Graham, J. R., Li, S., Qiu, J. (2008) Corporate misreporting and bank loan contracting. Journal of Financial Economics, 89(1), 44-61. https://doi.org/10.1016/j.jfineco.2007.08.005

Hubbard, R. G., Kuttner, K. N., Palia, D. N. (2002). Are there bank effects in borrowers' costs of funds? Evidence from a matched sample of borrowers and banks. The Journal of Business, 75(4), 559-581. https://doi.org/10.1086/341635

Iyer, R., Peydró, J.-L., da-Rocha-Lopes, S., Schoar, A. (2014). Interbank liquidity crunch and the firm credit crunch: Evidence from the 2007–2009 crisis. The Review of Financial Studies, (27, 1), 347-372. https://doi.org/10.1093/rfs/hht056

Jimenez, G., Salas, V., Saurina, J. (2006). Determinants of collateral. Journal of financial economics, 81(2). 255-281. https://doi.org/10.1016/j.jfineco.2005.06.003

Khwaja, A. I., Mian, A. (2008). Tracing the impact of bank liquidity shocks: Evidence from an emerging market. American Economic Review, 98(4), 1413-1442. https://doi.org/10.1257/aer.98.4.1413

La Porta, R., Lopez-de-Silanes, F., Zamarripa, G. (2003). Related lending. The Quarterly Journal of Economics, 118(1), 231-268. https://doi.org/10.1162/00335530360535199

Michelangeli, V., Peydro, J.-L., Sette, E. (2020). Credit demand vs. supply channels: Experimental-and administrative-based evidence. Economic Working Paper, 1731. Barcelona: Universitat Pompeu Fabra. Retrieved from https://econ-papers.upf.edu/papers/1731.pdf

Rajan, R., Winton, A. (1995). Covenants and collateral as incentives to monitor. The Journal of Finance, (50, 4), 1113-1146. https://doi.org/10.1111/j.1540-6261.1995.tb04052.x

Strahan, P. E. (1999). Borrower risk and the price and nonprice terms of bank loans. FRB of New York Staff Report, 90. https://doi.org/10.2139/ssrn.192769

Valta, P. (2012). Competition and the cost of debt. Journal of Financial Economics, 105(3), 661-682. https://doi.org/10.1016/j.jfineco.2012.04.004

APPENDIX A

Table A1. Correlation Eable: Borrower Characteristics

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1 Assets	1.00																		
2 Revenue	0.55	1.00																	
3 EBIT	0.20	0.35	1.00																
4 EBITDA	0.49	0.45	0.90	1.00															
5 Gross profit	0.41	0.78	0.61	0.63	1.00														
6 EBIT/Assets	0.00	0.00	0.00	0.00	0.01	1.00													
7 EBIT/Margin	-0.02	-0.01	0.13	0.11	0.05	0.02	1.00												
8 EBITDA/Assets	0.00	0.00	0.00	0.00	0.00	1.00	0.01	1.00											
9 EBITDA/Margin	0.02	-0.03	0.11	0.12	0.03	0.02	0.95	0.01	1.00										
10 Profit/Margin	-0.02	-0.03	0.09	0.08	0.11	0.00	0.37	0.00	0.38	1.00									
11 Tangibility	0.07	-0.03	0.03	0.09	0.02	0.00	0.13	0.00	0.27	0.10	1.00								
12 Net debt/Assets	-0.02	-0.05	-0.03	-0.04	-0.07	-0.01	-0.10	-0.01	-0.11	-0.05	0.01	1.00							
13 Net debt/Margin	0.09	-0.05	-0.03	0.01	-0.05	0.00	-0.14	0.00	-0.02	0.14	0.18	0.49	1.00						
14 Net debt/EBITDA	0.01	0.03	-0.02	-0.01	0.00	0.00	-0.06	0.00	-0.06	-0.03	-0.02	0.11	0.08	1.00					
15 Net debt/EBIT	0.05	0.01	-0.06	-0.04	-0.04	-0.01	-0.23	0.00	-0.20	-0.11	0.04	0.40	0.34	0.12	1.00				
16 Interest coverage	0.02	-0.01	0.03	0.05	0.04	0.00	0.27	0.00	0.30	0.23	0.18	-0.30	-0.13	-0.06	-0.21	1.00			
17 Quick ratio	0.02	-0.12	-0.01	0.01	-0.10	0.00	0.27	-0.01	0.33	0.17	0.37	-0.42	-0.10	-0.07	-0.24	0.34	1.00		
18 Equity ratio	-0.03	-0.10	-0.03	-0.04	-0.10	0.00	0.19	0.00	0.21	0.14	0.07	-0.15	-0.04	-0.03	-0.12	0.23	0.52	1.00	
19 Payables turnover	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	-0.01	0.00	0.03	0.00	0.01	0.00	0.00	0.00	1.00
20 Inventories turnover	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.07	0.00	-0.01	0.02	0.00	0.00	0.02	0.01	0.01	0.65

A BVAR MODEL FOR FORECASTING UKRAINIAN INFLATION AND GDP

NADIIA SHAPOVALENKO[a]

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

Abstract   In this paper, I examine the forecasting performance of a Bayesian Vector Autoregression (BVAR) model with a steady-state prior and compare the accuracy of the forecasts against the QPM and official NBU forecasts during the Q1 2016–Q1 2020 period. My findings suggest that inflation forecasts produced by the BVAR model are more accurate than those of the QPM model for two quarters ahead and are competitive for a longer time horizon. The BVAR forecasts for GDP growth also outperform those of the QPM but for the whole forecast horizon. Moreover, it is revealed that the BVAR model demonstrates a better performance compared to the NBU's official inflation forecasts over the monetary policy horizon, whereas the opposite is true for GDP growth forecasts. Future research may deal with estimation issues brought about by COVID-19.

JEL Codes   C30, C53, E37.

Keywords   BVAR, forecast evaluation, inflation forecasting.

1. INTRODUCTION

In 2016, the National Bank of Ukraine moved de facto to an inflation-targeting regime. One of the necessary preconditions for the successful implementation of an inflation-targeting regime is the development of models capable of producing accurate and well-grounded forecasts. In this framework, forecasting inflation becomes an essential task.

Regular medium-term macroeconomic forecasts and monetary policy recommendations at the National Bank of Ukraine (NBU) are mostly based on a Quarterly Projection Model (QPM), which is the main element of the Forecasting and Policy Analysis System (FPAS). The QPM is a semi-structural, forward-looking New Keynesian model of a small open economy. Owing to the fact that the main role of the QPM is to produce story-telling and to incorporate some expert judgments, the issue of the forecasts' accuracy may fade into the background. For that reason, it is worth having an additional empirical model producing more accurate forecasts.

The aim of this research is to develop a Bayesian Vector Autoregression (BVAR) model for forecasting inflation and GDP in Ukraine, to examine the forecasting performance of the model, and to compare the accuracy of these forecasts against those of the QPM model and official NBU forecasts.

The forecasting evaluation exercise uses quarterly data for the period of 2016Q1–2020Q1. During this period, the QPM was the main forecasting model, while official NBU forecasts were systematically documented. This allows the forecasts based on BVAR models to be compared with both the QPM and official NBU forecasts.

A Bayesian approach to estimation was chosen given that the Ukrainian data is short and dimensionality problems may arise with the large number of parameters present in the model. The imposition of priors not only solves the dimensionality problem, but supplements the information contained in the data with the personal judgments contained in the prior. Hopefully, the use of different sources of information will sharpen macroeconomic analysis.

I employ a BVAR model with an informative steady-state prior as in Villani (2009) because this type of priors is widely used for inflation-forecasting in countries that have adopted an inflation-targeting regime, as it explicitly uses information about the inflation target and other equilibrium values.

To the best of my knowledge, I am the first to use a BVAR model with a steady-state prior for forecasting Ukrainian inflation.

The paper is organized as follows. Section 2 contains a literature review. The theoretical framework and some issues regarding the forecast conditioning procedure can be found in Section 3. Section 4 presents an overview of inflation dynamics in Ukraine during the past 15 years. Section 5 describes the data and presents some correlation analysis. Section 6 presents empirical specifications of the models and the priors. Section 7 describes the results and the forecasting performance. Finally, Section 8 offers some concluding remarks. Additional information and results can be found in Appendices A-C.

2. LITERATURE REVIEW

The recent forecasting literature points out that among empirical models, BVARs have superior abilities when it comes to forecast output and inflation. In this section, an overview of recent empirical papers using BVAR models for forecasting purposes is provided. The attention is focused on the papers that are using BVARs with steady-state priors.

Villani (2009) was the first who imposed priors directly on the steady state of the model. He argued that this form of priors can be very important, especially for long-term horizon forecasts. Indeed, prior beliefs regarding the steady state are often available in relatively strong form and seem to improve the forecasting ability of the models.

Iversen et al. (2016) compared forecasts made with a DSGE model with a BVAR model against judgmental forecasts published by the Riksbank and found that BVAR model inflation forecasts and the repo rate have outperformed DSGE model forecasts and Riksbank's published forecasts. They also evaluated the usefulness of conditioning information for model-based forecasts (the forecasts were conditioned on the international forecast and the short-term forecast) and found that the difference between conditional and unconditional forecasts is rather small for BVAR forecasts. However, for the DSGE-based forecasts, conditioning information was helpful.

Brázdik and Franta (2017) also came to the conclusion that over the monetary policy horizon, the BVAR approach provides a more precise inflation forecast than the official ones published by the Czech National Bank. In their study, they considered BVAR forecasts, conditioning on the foreign outlook and – for the period of the exchange rate floor – also on the officially announced exchange rate and interest rate commitments.

Beechey and Österholm (2010) emphasized that for inflation-targeting countries such as Australia, Canada, New Zealand and Sweden. the out-of-sample forecasts of the mean-adjusted autoregressive model are superior to those of the traditional specification, often by significant amounts.

Clark (2011) showed that a BVAR model with a steady-state prior and stochastic volatility improves the real-time accuracy of density forecasts and modestly improves

the accuracy of point forecasts. As he is dealing with the forecasting of U.S. indicators, his model is specified for a closed economy. The endogenous variables are GDP growth, the unemployment rate inflation, the federal funds rate, and the nominal exchange rate. One of the specifications also includes as an endogenous variable the long-term inflation expectation from the Blue Chip Consensus, which is used to measure trend inflation.

The model for the Swedish economy used in Villani (2009) and Iversen et al. (2016) has also foreign indicators and the endogenous variables of the model are foreign GDP growth, foreign inflation, foreign interest rate, domestic GDP growth, domestic inflation, domestic interest rate, and the real exchange rate. The model considered in Iversen et al. (2016) also has nominal wages, hours worked, and the trade-weighted nominal exchange rate instead of the real exchange rate.

The model of Brazdik and Franta (2017) for the Czech economy is similar to the Villani (2009), however it also has a nominal exchange rate instead of a real exchange rate.

To select the specification of a BVAR model for the Ukrainian economy, it is worth starting from the specifications used in the above-mentioned papers.

3. THE THEORETICAL FRAMEWORK

3.1. A BVAR Model with Steady-State Priors

VAR models are a common tool in empirical macroeconomics, used both in forecasting and for analyzing the impact of shocks to the economy. However, a generous parameterization of the model – together with a small data sample – can result in a poor forecasting performance. Moreover, since the levels at which the forecasts converge are a function of the model's estimated parameters, the forecasting performance at longer horizons may be even worse.

A BVAR with "informative priors" on steady state may be a solution to the problem because it relates the information contained in the data with the judgments about the long-run values of the model's variables. That's in contrast to the majority of BVAR models, which focus on the dynamic behavior of the BVAR model with "informative priors" on steady state deals with the deterministic component of the model.

The methodology was first described by Villani (2009). The author proposes to use a VAR model in a mean adjusted form:

$$A(L)(y_t - Fx_t) = \varepsilon_t, \qquad (1)$$

where $t = 1..T$, y_t is a $n \times 1$ vector of endogenous variables, x_t is a $m \times 1$ vector of exogenous variables, ε_t is i.i.d. $N(0,\Sigma)$, $A(L) = I - A_1 L - A_2 L^2 - ...A_p L^p$ is a p lag polynomial, $A_1...A_p$ are $n \times n$ matrices, and F is $n \times m$ is a matrix of coefficients for the m exogenous variables. In this framework Fx_t usually takes the form of a constant, a piecewise constant or a linear time trend.

Taking expectations on both sides of equation and rearranging the equation one has:

$$E(y_t) = Fx_t, \qquad (2)$$

That is, the long-run value of the variables of the VAR is determined by the model's exogenous component and Fx_t represents an unconditional mean of y_t. When the exogenous component includes only constant terms, Fx_t reduces to a vector of constants so that $E(y_t) = \mu$. Thus, the steady-state values for the data are μ.

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + Fx_t - A_1 Fx_{t-1} - A_p Fx_{t-p} + \varepsilon_t \quad (3)$$

After rewriting (3) into transposed form, stacking observations and gathering the regressors into matrices we get:

$$
\begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_T' \end{pmatrix} = \begin{pmatrix} y_0' & y_{-1}' & \cdots & y_{1-p}' \\ y_1' & y_0' & \cdots & y_{2-p}' \\ \vdots & \vdots & \ddots & \vdots \\ y_{T-1}' & y_{T-2}' & \cdots & y_{T-p}' \end{pmatrix} \begin{pmatrix} A_1' \\ A_2' \\ \vdots \\ A_p' \end{pmatrix} +
$$
$$
+ \begin{pmatrix} x_1' & -x_0' & \cdots & -x_{1-p}' \\ x_2' & -x_1' & \cdots & -x_{2-p}' \\ \vdots & \vdots & \ddots & \vdots \\ x_T' & -x_{T-1}' & \cdots & -x_{T-p}' \end{pmatrix} \begin{pmatrix} F' \\ F'A_1' \\ \vdots \\ F'A_p' \end{pmatrix} + \begin{pmatrix} \varepsilon_1' \\ \varepsilon_2' \\ \vdots \\ \varepsilon_T' \end{pmatrix} \quad (4)
$$

Or in compact notation:

$$Y = XB + Z\Delta + E, \quad (5)$$

where

$$
Y = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_T' \end{pmatrix}, X = \begin{pmatrix} y_0' & y_{-1}' & \cdots & y_{1-p}' \\ y_1' & y_0' & \cdots & y_{2-p}' \\ \vdots & \vdots & \ddots & \vdots \\ y_{T-1}' & y_{T-2}' & \cdots & y_{T-p}' \end{pmatrix},
$$

$$
B = \begin{pmatrix} A_1' \\ A_2' \\ \vdots \\ A_p' \end{pmatrix}, Z = \begin{pmatrix} x_1' & -x_0' & \cdots & -x_{1-p}' \\ x_2' & -x_1' & \cdots & -x_{2-p}' \\ \vdots & \vdots & \ddots & \vdots \\ x_T' & -x_{T-1}' & \cdots & -x_{T-p}' \end{pmatrix},
$$

$$
\Delta = \begin{pmatrix} F' \\ F'A_1' \\ \vdots \\ F'A_p' \end{pmatrix}, E = \begin{pmatrix} \varepsilon_1' \\ \varepsilon_2' \\ \vdots \\ \varepsilon_T' \end{pmatrix}
$$

Vectorizing (4) and compactly rewriting it we obtain:

$$y = \bar{X}\beta + \bar{Z}\delta + \epsilon, \quad (6)$$

where
$y = vec(Y), \bar{X} = I_n \otimes X, \beta = vec(B), \bar{Z} = I_n \otimes Z,$
$\delta = vec(\Delta), \epsilon = vec(E).$

Let:

$$vec(\Delta') = vec(F\, A_1 F \ldots A_p F) =$$
$$= \begin{pmatrix} I_{nm} \\ I_m \otimes A_1 \\ \vdots \\ I_m \otimes A_p \end{pmatrix} vec(F) = U\psi \quad (7)$$

where $\psi = vec(F)$,

$$U = \begin{pmatrix} I_{nm} \\ I_m \otimes A_1 \\ \vdots \\ I_m \otimes A_p \end{pmatrix} \quad (8 \text{ and } 9)$$

Note that there are now three blocks to estimate – β, which corresponds to the coefficients on the endogenous variables y_t; ψ, which corresponds to coefficients on the exogenous variables x_t and Σ - the residual variance-covariance matrix.

A diffuse prior for the error covariance matrix is assumed, while the prior on the other two sets of coefficients is normal.

$$p(\Sigma) \propto |\Sigma|^{-(n+1)/2}, \quad (10)$$

$$\beta \sim N(\beta_0, \Omega_0), \quad (11)$$

$$\psi \sim N(\psi_0, \Lambda_0), \quad (12)$$

Dieppe et al. (2016) argue that one can't set a flat prior for ψ as in the Minnesota scheme, because the very purpose of this type of prior is to add information about means into the estimation process. It is recommended to specify a subjective 95% probability interval for the prior values. Using the properties of the normal distribution, the prior mean of the distribution is determined as the mode of the specified subjective 95% probability interval, while the variance is obtained by the fact that the bounds of a subjective 95% probability interval are located at 1.96 standard deviations from the mean.

Villani (2009) shows the complete derivation of the posterior distribution. The steps of a Gibbs sampling algorithm for the BVAR with a steady-state prior can be found in Appendix A.

3.2. Hyperparameter Values

Many researchers obtain the optimal hyperparameters by maximizing the marginal likelihood over a grid of possible values. They include Del Negro and Schorfheide (2004), Schorfheide and Song (2015) and Carriero, Clark, and Marcellino (2015). The grid-search approach is also represented in the Bayesian Estimation, Analysis and Regression (BEAR) MATLAB toolbox developed by the European Central Bank (Dieppe et al (2016)).

However, being suitable for low-dimensional models, it may be computationally infeasible for higher dimensions. Therefore, more inference-based approaches to setting the hyperparameter have arisen. Giannone et al. (2012) use a hierarchical modeling framework.

Gustafsson, Villani and Stockhammar (2020) propose a new Bayesian optimization method. They focus on the common situation of maximizing a marginal likelihood evaluated by MCMC, where the precision is determined by the number of MCMC iterations. The authors argue that "the ability to choose the precision makes it possible for the algorithm to take occasional cheap and noisy evaluations to explore the marginal likelihood surface, thereby finding the optimum faster".

Chan et al (2019) apply Automatic Differentiation (AD) to calculate the gradient of the marginal likelihood with respect to the hyperparameters, which is then used as an input in an optimization routine. Authors claim that by computing the gradient efficiently using AD, the proposed method is substantially faster than the conventional grid-search approach.

To put it in a nutshell, the variety of methods is aimed at increasing the speed of finding the optimum. Whereas, they consider the same criterion, which is the maximization of marginal likelihood, which – according to Giannone et al. (2012) – "corresponds to maximizing the one-step-ahead, out-of-sample forecasting ability" of the model. In order to care more about the forecasting accuracy for the monetary policy horizon, it is probably worth trying to choose hyperparameters by maximizing the forecasting performance for 4-6 quarters over a presample. However, taking into account the small sample and low dimensionality of the models used in the research, a simple grid search was applied.

The procedure of the estimation of hyperparameters is the following. First, a range for each hyperparameter is specified, together with a step size defining the size of the increment within the range. Then the marginal likelihood is estimated for each model with every possible combination of hyperparameter values. The optimal combination, which is the one that maximizes the marginal likelihood, is then kept.

3.3. Forecasting

Iterated BVAR forecasts for up to six quarters[1] are simulated in the form of a posterior predictive distribution. The root mean squared error (RMSE) is used to evaluate the accuracy of the BVAR point predictions and to compare it against the QPM model and the NBU official forecasts. Also, a simple AR model is constructed to serve as a benchmark (the lag length of the AR model is selected minimizing the RMSE within the forecasting exercise period).

Together with unconditional forecasts, I compute forecasts conditioning on foreign indicators. I do so for a number of reasons. First, almost all mid-term forecasts at the NBU are based on some assumptions concerning either external or internal factors (e.g., conditioning on the interest rate, as it serves as a main instrument of monetary policy; or on external variables, as more precise forecasts of external indicators are available). Hence, conditioning allows forecasts to be more realistic. Moreover, it makes the interpretation of forecasts and story building around it easier. Second, conditioning on the same variables used in the QPM makes the comparison of the models more meaningful. Finally, I expect the conditional inflation forecast to be more precise and I am going to examine this hypothesis.

There are several options in the literature on how to incorporate external information into the forecasts of BVAR. The hard conditioning option was developed by Waggoner and Zha (1999), who derived a Gibbs sampling algorithm to construct the posterior predictive distribution of the conditional forecast.

A more efficient solution was suggested by Jarocinski (2010). In this framework, shocks are divided into constructive and non-constructive. Constructive shocks are the shocks on which a condition is imposed. However, conditioning may not be unique, meaning the same condition may be imposed on different shocks. Therefore, the researcher should carefully select the shocks generating the constraint, in order to produce sensible economic results.

In contrast to hard conditioning, in which the future values of variables are fixed at single points, soft conditioning is more flexible and deals with conditions that only restrict the future values within a certain range. Soft conditioning was also introduced by Waggoner and Zha (1999). However, an alternative methodology (entropic tilting) initially proposed by Robertson et al. (2005) and further developed by Krüger et al. (2017) allows for incorporating external information into model-based forecasts.

Comparing the soft conditioning by Waggoner and Zha (1999) with entropic tilting, Dieppe et al. (2016) argue that one of the main advantages of entropic tilting is its high flexibility. This is because the method of Waggoner and Zha (1999) only allows to set the center of the predictive distribution, whereas the entropic tilting method allows any moment associated with the distribution to be determined, along with quantile values.

The main idea of the entropy tilting method is to change the initial predictive distribution of the unconditional forecast to a new one that satisfies specified moment conditions, and to minimize the distortions in the other properties of the new distribution. In other words, to get a new distribution, one minimizes the relative entropy between the two distributions, subject to the restriction that the new distribution satisfies the specified moment conditions. So, by construction, conditional forecasts obtained through entropic tilting are as close to the initial distribution of unconditional forecast as possible. Further details on the technical implementation of entropic tilting can be found in Dieppe et al. (2016).

In this paper I use entropic tilting, presuming that it will produce more accurate forecasts.

4. INFLATION IN UKRAINE, AN OVERVIEW

During the Great Recession, Ukraine was hit by a sharp terms-of-trade shock: steel prices plunged (in 2008 steel represented about 40% of exports and 15% of GDP), while energy import prices remained high due to the phasing out of Russia's natural gas subsidies. The materialization of trade shock terms had a considerable impact on the real sector. In addition, major strains were building up in the banking system following a system-wide run on deposits. A loss of confidence domestically led to capital flight out of the hryvnia into foreign exchange cash. Altogether, this led to a massive devaluation of the currency, plummeting real GDP and a shrinking of the current account deficit in 2009.

In 2010-2011, the economy started recovering, inflation declined to single digits, the exchange rate was stabilized, and growth rebounded. In 2012-2013, inflation approached zero due to weak economic activity (annual GDP growth was 0.2% in 2012 and 0.0% in 2013). Keeping the exchange rate stable led to the accumulation of huge imbalances in the economy. In 2014, these imbalances – along with the military conflict in the east of the country – led to a severe economic crisis, with real GDP falling by 10% in 2015, with sharp depreciation of the hryvnia and inflation reaching its peak of almost 60% year-over-year in the spring of 2015.

It is worth noting that the nature of the two high inflation episodes (2008 and 2015) is different: the second inflationary spike was caused by the pass through of the hryvnia devaluation, whereas in 2008, rising inflation was a sign that the economy had been overheating.

---

[1] Typically, the monetary policy horizon is considered to be four- to eight-quarters ahead. Due to the short data sample, the forecasts are simulated up to six quarters, paying more attention to the results from fourth to sixth quarters.
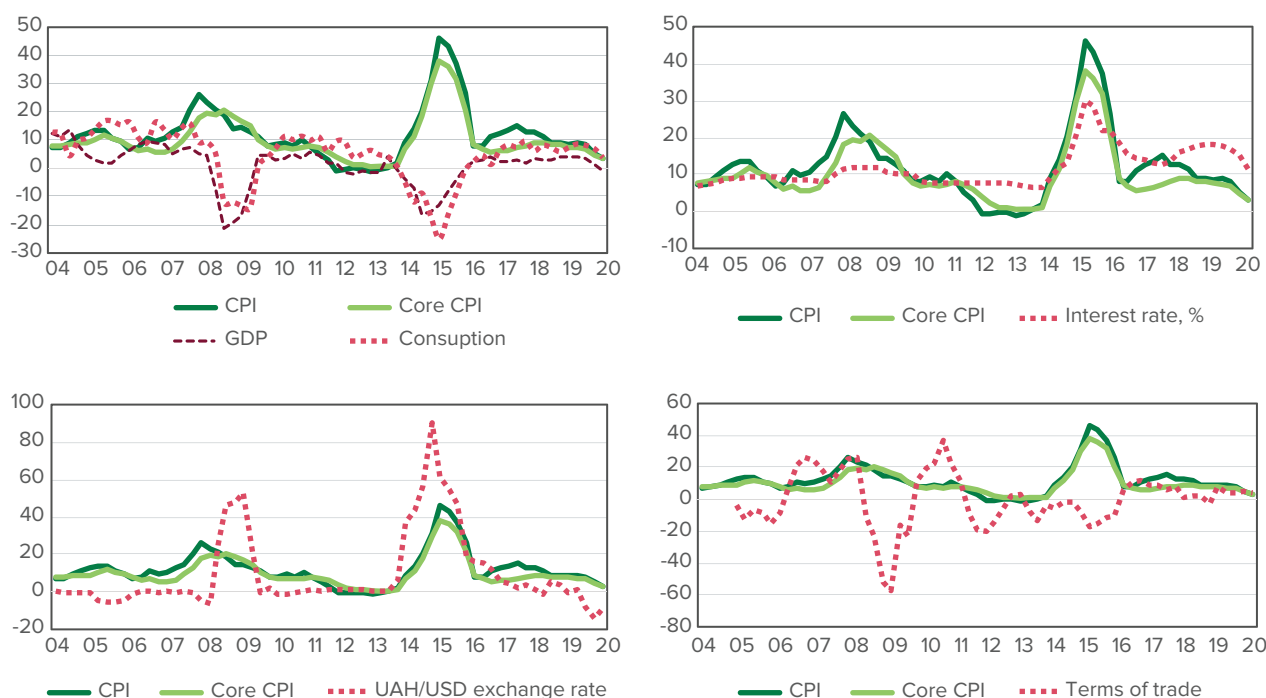
Figure 1. Main Economic Indicators, yoy, in logs

In August 2015, the NBU declared a transition to the inflation-targeting regime in order to break the upward inflationary trend and stabilize the economy. De facto, it moved to the inflation-targeting regime in 2016. The NBU announced it set its mid-term inflation target (year-over-year CPI growth) at 5% and to be achieved gradually in the following stages:

- 12% +/- 3 ppts as of the end of 2016;
- 8% ± 2 ppts as of the end of 2017;
- 6% ± 2 ppts as of the end of 2018;
- 5% ± 1 ppt as of the end of 2019 and further on.

It is well-known that the inflation targeting regime uses the policy rate as a main instrument. To bring inflation down to the target, the NBU should increase the interest rate to moderate demand and to cool inflationary pressure. Therefore, the gradual strategy of bringing inflation to its target was chosen deliberately in order to minimize the costs of disinflation for economic growth.

In general, the process of disinflation, which started in 2016, went well. In 2019, consumer price inflation gradually declined to a six-year low of 4.1%. Therefore, the NBU finally achieved its target of 5% ± 1 ppt. The average GDP growth was 2.8% in 2016-2019.

2020 brought a new challenge: the COVID-19 pandemic was a shock of unprecedented global severity affecting all areas of the economy. In this situation, swift and reasonable policy measures have been of great importance. In the near future, policy makers will need to find the right balance between supporting the economy using an accommodating policy and maintaining price stability.

To summarize, the recent economic developments in Ukraine show that along with domestic conditions, external ones are another important driver of inflation and should be used to forecast Ukrainian inflation.

5. DATA DESCRIPTION AND CORRELATIONS ANALYSIS

Although the methodology of meta-analysis is helpful I use quarterly foreign data, national accounts data, prices and exchange rates over the period of 2004Q1–2020Q1 (see Table B.1, Appendix B). Alternative measures of different variables are employed in order to find the one with the highest predictive power for inflation:

- In addition to weighted CPI, PPI is used for foreign price levels;
- CPI and PPI deflator-based REER[2] are used for the real exchange rate;
- vernight or 3-month LIBOR is used as the foreign interest rate;
- Various commodity prices are employed as an alternative for foreign price levels;
- I use two measures of terms of trade, constructed as the ratio between the index of export prices and the index of import prices for (1) goods, (2) most important groups of raw commodities;
- Monetary aggregate M2, nominal and real wage are used to reflect domestic factors.

All the data except interest rates are measured in natural logarithms. Growth variables in annualized quarter-over quarter terms are used. To choose the variable to be used in the forecasting exercise, I employed a simple correlation analysis. The figures – as well as the correlation coefficients between Ukrainian CPI, GDP and other variables – are presented in Figure C1, Appendix C.

The CPI is significantly correlated with both the inflation differential of trading partners and the NEER. However, there is no significant correlation of CPI with the weighted GDP of trading partners and various commodity prices, as these indicators may be more important for domestic production rather than consumption. CPI has a very weak negative correlation with the lagged policy rate, whereas the contemporaneous correlation has a positive sign. This can be explained by the fact that the interest rate hadn't been used as an instrument prior to 2016 (before the implementation of inflation-targeting). So the monetary transmission mechanism didn't work as it was supposed to.

In addition to unconditional correlations, correlations conditioned on the policy rate were analyzed. However, no serious differences with unconditional correlations were found (see Figure C.2, Appendix C). Domestic GDP is significantly correlated with foreign GDP, which means that for such a small open economy as Ukraine's, external demand is an important factor of GDP growth. The positive correlation of GDP with terms of trade and commodity prices reflects the fact that these indicators drive the Ukrainian business cycle, with Ukraine being a commodity net exporter.

For the same reason as with CPI, the correlation of GDP with the policy rate is weak. The correlation between monetary aggregates and wages suggests not to include them into the model.

To conclude, taking into account the results of the correlation analysis, together with the stylized facts from the Section 4 and the models described in the literature review, the following indicators were chosen for the BVAR model for the Ukrainian economy: the weighted[3] GDP of trading partners, the weighted inflation differential of trading partners, domestic GDP, domestic CPI, the domestic policy rate, the nominal effective exchange rate (NEER), and the terms of trade, constructed as the ratio of the most important groups of raw commodities and the non-energy commodity price index.

6. EMPIRICAL MODEL SPECIFICATIONS AND THE PRIORS

The selection of the variables for the model is based not only on correlations between the variables and is conducted in several steps. First, I analyzed the variables that were selected in other research papers (Villani (2009) and Iversen et al. (2016) for the Swedish economy, Brazdik and Franta (2017) for the Czech economy). Second, I described the recent economic developments in Section 4 with the purpose to better understand the relationships between macroeconomic variables and the main transmission channels in Ukraine. Finally, the correlation analysis was helpful in distinguishing between alternative measures of some economic variables.

The benchmark BVAR specification (MB) for Ukraine is the following:

$$y_t = (\Delta gdp_t^f, \pi_t^f, \Delta gdp_t, \pi_t, i_t, neer_t)' \qquad (13)$$

y_t includes foreign GDP growth (Δgdp^f), foreign inflation (π^f), domestic GDP growth (Δgdp), domestic inflation (π), domestic

interest rate (i), and the NEER ($neer$).

In order to find the best possible set of variables, two additional specifications are considered (MA_p and MA_tot):

$$y_t = (\Delta gdp_t^f, wbnonen_t, \Delta gdp_t, \pi_t, i_t, neer_t)' \qquad (14)$$

y_t includes non-energy commodity price index ($wbnonen$) instead of foreign inflation (π^f).

$$y_t = (\Delta gdp_t^f, \pi_t^f, \Delta gdp_t, \pi_t, i_t, neer_t, tot_t)' \qquad (15)$$

in addition to the variables from the benchmark model, y_t includes terms of trade (tot).

In order to take into account that Ukraine is a small open economy, foreign variables and terms of trade are treated as block exogenous. Namely, the block submatrices in A(L), corresponding to the effects of domestic variables on foreign ones, are set to zero.

Standard lag length criteria were used to select the lag length (see Table B.2, Appendix B). Different criteria suggest the use of lags from 1 up to 5. In general, specifications with, larger lag length were preferred. As a robustness check, more parsimonious specifications with two lags were also estimated and the results didn't significantly differ.

The hyperparameters for the models are set according to the results of grid search procedure (information regarding the grid search is in Table B.3, Appendix B). The values of hyperparameters, used as well as the information regarding number of lags and the number of iterations, is presented in Table 1.

Table 1. Hyperparameters and Lags

	MB	MA_P	MA_TOT
Autoregressive coefficient	0.5	0.4	0.4
Overall tightness (λ_1)	0.2	0.2	0.2
Cross-variable weighting (λ_2)	1	0.9	1
Lag decay (λ_3)	1	1	1
block exogeneity shrinkage λ_5:	0.001	0.001	0.001
total number of iterations:	10,000	10,000	10,000
bum-in iterations:	5,000	5,000	5,000
Lag length	3	4	5

The priors on the steady states are normally distributed. In order to account for changes in the monetary policy regime (the move to inflation targeting in 2016), two different sets of priors are employed. The first regime covers the period from 2004Q1 up to 2015Q4 and the second regime starts at 2016Q1.

To specify the moments of the prior distribution values, Dieppe et al (2016) recommend first to set a subjective 95% probability interval, and then calculate the mean and variance for each variable. Brazdik and Franta (2017), on the contrary, suggest calculating a 95% probability interval, based on the mean and variance.

---

[3] Aggregated foreign indicators for 40 countries were chosen.

I follow Brazdik and Franta (2017). The means of the priors are taken from the trends estimated in the QPM model. Variances are set using the information from other studies, keeping in mind that tighter interval would imply smaller prior variance and hence greater confidence that the steady-state value corresponds to the specified prior mean. On the other hand, a wider interval would imply larger prior variance and more weight given to the data.

The CPI prior for the 2nd regime is set in a different manner. Since in the inflation-targeting regime both target and the bounds for inflation are known, the bounds are used directly to set the values for a 95% interval. The means and variances from Villani (2009) and Brazdik and Franta (2017) are in Tables B.4 and B.5, Appendix B. Steady state priors for the Ukrainian model are presented in Table 2. In general, priors for Ukrainian model are looser than those for the Swedish model and tighter than those for the Czech model.

7. ESTIMATION RESULTS AND FORECASTING PERFORMANCE

7.1. Estimation Results

The priors and posterior estimates of the steady state for BVAR models are presented in Table 3. The estimates of the steady states are based on the reduced-form VAR, hence structural shock identification does not play any role. Impulse responses based on recursive identification can be found in Figures C.3-C.5, Appendix C.

Specifically, priors are reported for the 2nd regime, while posterior estimates are presented for 2020Q1 (which corresponds to the end of the sample, so that the full data set was used for the estimation).

There are some differences between the prior and posterior medians of the steady state for 2020Q1, as well as differences in the values of the posterior medians of three BVAR models that are worth discussing.

All three models have lower posterior medians for foreign GDP and CPI than the prior median value. However, for the MB model, this difference is more pronounced. The reason for the difference may be the fact that in contrast to MB model, both the MA_P and MA_TOT models contain additional information on the dynamics of commodity prices, which may influence the steady state values of foreign variables.

Lower steady state values of external demand in the MB model, in turn, affect domestic GDP growth. Hence, the posterior medians for the domestic variables of the MB model suggest lower steady-state value of GDP growth, inflation, the policy rate and a more pronounced NEER depreciation trend.

7.2. Forecasting Performance

In this subsection, the forecasting performance of the BVAR models is examined. The RMSE is employed as the measure of forecasting performance. In addition to RMSE, the equal forecasting accuracy of the models is statistically evaluated using the Diebold-Mariano test. The comparison is divided into four stages. During the first stage, the unconditional forecasts of MB, MA_P and MA_TOT are compared to the forecasts of the AR1 model.

Table 2. Steady State prior Distributions

	regime 1: 2004q1 2015q4				regime 2: 2016q1 2020q1			
	Mean	Var	95%	Interval	Mean	Var	95%	Interval
GDPW	4.0	0.5	3.0	5.0	3.0	0.5	2.0	4.0
CPIW	6.0	0.5	5.0	7.0	3.5	0.5	2.5	4.5
GDPUA	1.0	1.0	-1.0	3.0	2.0	0.5	1.0	3.0
CPIUA	11.0	2.0	7.1	14.9	8.0	1.0	6.0	10.0
IUA	12.5	0.7	11.1	13.9	11.0	0.6	9.8	12.2
NEER	-6.8	2.0	-10.7	-2.9	-2.5	1.0	-4.5	-0.5
TOT	-2.0	1.0	-4.0	0.0	3.0	0.5	2.0	4.0
PNONEN	6.0	1.0	4.0	8.0	1.0	0.5	0.0	2.0

Table 3. Priors and Posterior Estimates for 2020q1

	Prior, 2016q1 2020q1			Posterior, MB			Posterior, MA_P			Posterior, MA_TOT		
	Median	95%	Interval	Mean	Var	Interval	Median	95%	Interval	Median	95%	Interval
GDPW	2.0	3.0	4.0	2.7	1.9	3.7	1.8	2.7	3.6	1.8	2.7	3.6
CPIW	2.5	3.5	4.5	3.4	2.6	4.2				2.5	3.3	4.0
GDPUA	1.0	2.0	3.0	2.0	1.1	3.0	1.0	2.0	3.0	1.1	2.1	3.0
CPIUA	6.0	8.0	10.0	7.8	5.9	9.7	5.6	7.5	9.5	5.7	7.7	9.6
IUA	9.8	11.0	12.2	11.8	10.7	13.0	10.7	11.9	13.0	10.7	11.9	13.0
NEER	-4.5	-2.5	-0.5	-2 3	-4 2	-0.3	-4.4	-2.4	-0.4	-4.4	-2.3	-0.4
TOT	2.0	3.0	4.0							2.0	3.0	4.0
PNONEN	0.0	1.0	2.0				0.1	1.1	2.0			

Then, during the second stage, the forecasts of the BVAR model with the most accurate unconditional forecasts are compared with the forecasts of the same model, conditioned on external indicators – namely, foreign GDP growth and foreign inflation – in order to examine whether conditioning improves forecasting accuracy.

At the following stage, the conditional forecasts of the best model from the second stage are compared with the conditional forecasts of the QPM model.

Finally, at the fourth stage, conditional forecasts of the best BVAR model are transformed from annualized quarter-over-quarter indicators into year-over-year indicators and compared with official NBU forecasts. The data transformation is necessary because NBU forecasts are only available on a year-over-year basis.

Because the forecasting performance for inflation at the monetary policy horizon is of most interest, I focus attention on the forecast horizons from the fourth to sixth quarters.

Table 4. RMSEs for Unconditional Forecasts Relative to the AR1 Model

| period | CPI | | | GDP | | |
	MB	MA_P	MA_TOT	MB	MA_P	MA_TOT
1	1.05	1.02	1.05	0.89	0.99	0.89
2	1.02	1.02	0.93	0.51**	0.44**	0.48**
3	1.11	1.15	1.11	0.51**	0.50**	0.52**
4	1.04	1.02	0.97	0.67**	0.59**	0.60**
5	0.73*	0.75*	0.64**	0.74*	0.67**	0.73**
6	0.72*	0.69**	0.67**	0.83	0.75**	0.82*

Note: Asterisks indicate that according to the Diebold-Mariano test, the difference in forecasting performance relative to the AR1 model is statistically significant at 5% or 10% level (** and * respectively).

The forecast accuracy of the unconditional BVAR forecasts with different variable specifications is reported in Table 4. Plots of the forecasts can be found in Figure C.6, Appendix C. RMSE values are shown relative to those of an AR1 model in order to facilitate the comparison. Therefore, for the given model, a value below unity means it's better than the AR1 model's precision.

Regarding CPI, the BVAR model – which includes terms of trade (MA_TOT) – seems to have the best forecasting accuracy and outperforms AR1 at the horizon of interest. It is worth noting that in the short run, AR1 forecasts are more accurate. However, the difference is not statistically significant.

Regarding GDP, the BVAR models outperform the AR1 model from 2nd to 6th quarters and the differences are statistically significant. The added prior information may be the reason for the superior performance of the BVAR models at the longer horizons.

Since the BVAR model, which includes terms of trade (MA_TOT), has lower RMSE for inflation. It will be used in further comparison.

The value of incorporating external conditioning information can be judged by comparing the RMSE for the conditional and the unconditional BVAR forecasts (see Table 5).

Regarding CPI, on average, conditional forecasts are more accurate than unconditional ones, but the difference is rather small. For GDP, unconditional forecasts perform better than the conditional ones, still the difference is not significant.

Hence, I may conclude that adding external information probably does not play an important role in improving the forecasting accuracy of inflation and GDP.

Table 6 shows the results of the third stage (see also Figure C.7, Appendix C). For CPI inflation, the BVAR forecasts are superior for 4th and 6th quarters, while the RMSE of the QPM forecasts for 5th quarter is lower than that of the BVAR. Also, for the 1st quarter, both the BVAR and the QPM are inferior to the AR1.

The results are better for GDP. Both the BVAR and the QPM forecasts beat AR1 forecasts starting from the 2nd quarter, although not all differences are statistically significant. For the whole horizon, GDP forecasts of the BVAR model are more accurate than those of the QPM model.

Therefore, in general, for both inflation and GDP growth, the BVAR model forecasts are competitive at minimum, if not better than QPM forecasts.

The forecasting performance of the BVAR and official NBU forecasts for year-over year indicators is compared in Table 7. Plots of the forecasts can be found in Figure C.8, Appendix C.

Table 5. RMSEs for the Unconditional and Conditional Forecasts of MA_TOT Relative to the AR1 Model

| q | CPI | | GDP | |
	MA_TOT (cond)	MA_TOT	MA_TOT (cond)	MA_TOT
1	1.07	1.05	0.89	0.89
2	0.83*	0.93	0.79*	0.48**
3	1.02	1.11	0.72**	0.52**
4	0.83*	0.97	0.83	0.60**
5	0.78**	0.64**	0.91	0.73**
6	0.66**	0.67**	0.82*	0.82*

Note: Asterisks indicate that according to the Diebold-Mariano test, the difference in the forecasting performance – relative to the AR1 model – is statistically significant at 5% or 10% level (** and * respectively).

Table 6. RMSEs of Conditional Forecasts Relative to the AR1 Model

| q | CPI | | GDP | |
	MA_TOT (cond)	QPM	MA_TOT (cond)	QPM
1	1.07	1.17	0.89	1.04
2	0.83*	1.18	0.79*	0.84*
3	1.02	0.95	0.72**	0.97
4	0.83*	0.89	0.83	0.89
5	0.78**	0.74	0.91	0.95
6	0.66**	0.69**	0.82*	0.88

Note: Asterisks indicate that according to the Diebold-Mariano test, the difference in forecasting performance relative to the AR1 model is statistically significant at 5% or 10% level (** and * respectively).

Table 7. RMSEs of the Forecasts for the Indicators on a Year-over-Year Basis

q	CPI		GDP	
	MA_TOT (cond)	NBU	MA_TOT (cond)	NBU
1	1.07	0.99	0.89	1.02
2	0.95	0.85	0.74**	0.87
3	0.90	0.90	0.59**	0.68**
4	0.82	0.88	0.52**	0.60**
5	0.64*	0.92	0.62**	0.36**
6	0.55**	0.85	0.66**	0.52**

Note: Asterisks indicate that according to the Diebold-Mariano test, the difference in forecasting performance relative to the AR1 model is statistically significant at 5% or 10% level (** and * respectively).

Regarding CPI, the BVAR forecasts outperform the official NBU forecasts at the horizon of interest. However, in the short run, the official NBU forecasts are the most accurate: they beat both the BVAR and AR1 forecasts. This finding could be a consequence of the fact that the NBU is considering a broader information set during the forecasting process and different types of models specifically designed for short-run forecasting. Moreover, CPI is forecasted at the disaggregated level and for some groups of prices (e.g. administrative prices), expert judgments are included.

Regarding GDP, NBU forecasts have a better performance than BVAR forecasts at the horizon of interest, probably because GDP components are treated separately and expert knowledge is included (for example, the BVAR model doesn't explicitly have variables reflecting fiscal policy stance).

To conclude, the BVAR forecasts of inflation outperform the official NBU forecasts at the horizon of interest, whereas the opposite is true for the forecasts of GDP growth.

An interesting perspective can be added if we look at the forecast bias. The forecast bias is measured as the average forecast error at a certain horizon. In turn, the forecast error is calculated as the difference between the actual value and forecasted one. A non-zero bias indicates a possible persistent difference between the forecasts and the observed values.

Tables 8 and 9 present the values of CPI and GDP forecast bias for quarter-over-quarter indicators. Regarding CPI, the hypothesis of unbiasedness is rejected only for some forecast horizons for the MA_P and AR1 models. Also, except for the AR1 forecasts, a positive forecast bias is observed almost within the whole forecast horizon, meaning that the models on average underpredict inflation. As during the period of the forecasting exercise during which the disinflation occurred, lower forecasting values may mean that models are assuming faster convergence to the steady state than happened in real life. Moreover, the forecasts at the horizon of interest are less biased. In the 5th and 6th quarters, the forecast bias of the models is decreasing. For the MA_TOT model, it even becomes slightly negative in the 6th quarter.

Regarding GDP, the conditional forecasts of the BVAR have the smallest bias in absolute terms. However, contrary to other models, the errors of BVAR conditional forecasts have negative sign, meaning overprediction of GDP. Such a difference in the biases between conditional and unconditional forecasts may indicate the importance of conditioning for GDP forecasts.

Taking into consideration that official NBU forecasts are available only on a year-over-year basis, it is not possible to include them into the above comparison. However, if we look at Figure C.8, Appendix C, official NBU forecasts on a year-over-year basis seem to be biased towards under-forecast. There may be several reasons that contribute to this. First, under the inflation-targeting regime, the NBU may have tried to anchor inflation expectations by approaching forecasts to the target, which was lower than the actual inflation. Second, for both conditional BVAR forecasts and QPM forecasts, the observed values of external sector indicators are used. Whereas during the real forecasting process, these values are unknown and the values that are assumed may differ from actual ones.

Finally, I would like to address an issue that has received much attention lately. As the estimation period ends in 2020Q1, the forecasting accuracy of the BVAR during COVID-19 cannot be analyzed. However, the issue of dealing with COVID-19 outliers remains of key interest at the NBU because the developed BVAR model is going to be used for forecasting inflation and GDP in the years to come. There are several papers offering some solutions to the problem, which are applicable to the model I consider.

Table 8. Forecast Bias (CPI)

q	CPI					
	MB	MA_P	MA_TOT	MA_TOT (cond)	QPM	AR1
1	0.39	1.54	0.16	0.04	0.30	-0.53
2	1.78	2.57	1.52	1.41	0.38	-0.80
3	2.81	3.27**	2.44	2.19	1.44	-0.94
4	2.84	2.93**	2.13	1.91	1.65	-1.69
5	1.17	1.06	0.41	0.89	1.12	-3.57**
6	0.41	0.06	-0.53	-0.64	0.75	-4.49**

Note: Based on the results of a simple unbiasedness test[4], asterisks indicate that the null hypothesis of unbiasedness is rejected at the 5% level.

Table 9. Forecast Bias (GDP)

q	GDP					
	MB	MA_P	MA_TOT	MA_TOT (cond)	QPM	AR1
1	0.35	0.35	0.42	-0.23	0.99	1.99**
2	0.09	0.32	0.33	-0.64	0.79	2.59**
3	0.25	0.66	0.71	-0.29	0.50	2.97**
4	0.16	0.66	0.69	-0.13	0.31	2.84**
5	-0.14	0.24	0.29	-0.60	-0.52	2.37**
6	-0.34	0.11	0.03	-0.22	-0.75	2.08**

Note: Based on the results of a simple unbiasedness test, asterisks indicate that the null hypothesis of unbiasedness is rejected at the 5% level.

[4] Unbiasedness test for the forecast error e_{t+h} for forecast horizon h implemented with a t-test in the following regression: $e_{t+h} = y_{t+h} - y^f_{t+ht} = \tau_h + \varepsilon_t$, where the null hypothesis is $\tau_h = 0$

Foroni et al. (2020) consider simple methods to improve growth nowcasts and forecasts. Specifically, they combine forecasts across various specifications for the same model or across different models, extend the model specification by adding MA terms, and adjust the forecasts to put them back on track by a specific form of intercept correction etc. They find that among all these methods, adjusting the original forecasts by an amount similar to the forecast errors made during the financial crisis, as well as the following recovery, seems to produce the best results for the U.S., notwithstanding the different source and characteristics of the financial and COVID crises.

Lenza and Primiceri (2020) show how to handle a problem with COVID-19 outliers when estimating VAR models. Their solution consists of explicitly modeling the large change in shock volatility during the pandemic.

Type-1 and Type-2 publication biases, assessing both the extent of selecting only statistically significant estimates for publishing and the extent of selecting the estimates, which are consistent with economic theory. The results of the implemented tests have demonstrated that there is no statistically significant evidence of both types of publication biases in the estimates. The results of meta-regression have shown that the interconnection between interest rates and exchange rates is highly sensitive to a range of macroeconomic factors, especially when we are talking about the level of monetary freedom. Also, the effect was stronger for studies undertaken on post-1990 data.

Although due to data limitations, the inference about the possible effect of interest rate on exchange rate is made based on cross-country evidence rather than on the analysis of Ukrainian data, there is still a high probability that the same kind of relationship might be observed in Ukraine. Although there is no point in discussing the direct estimation of such a monetary policy instrument as the key policy rate on the national currency, the National Bank of Ukraine should take into account such an indirect inference while making its decisions regarding the key policy rate.

8. CONCLUSIONS

In this paper, I examined the forecasting performance of a Bayesian Vector Autoregression model with a steady-state prior for Ukrainian economy and compared the accuracy of the forecasts against the forecasts of the QPM model and official NBU forecasts. The RMSE is employed as the measure of forecasting performance. As the forecasting performance for inflation at the monetary policy horizon is of most interest, I focused on the horizon from the fourth to sixth quarters.

The BVAR model was estimated using both data for the Ukrainian economy and foreign indicators. In addition to the benchmark specification, models that include data on commodity prices and terms of trade were included in the alternative specifications to take into account the peculiarities of the Ukrainian economy. The model containing the terms of trade indicator happened to have the most accurate unconditional forecasts of inflation and GDP growth and outperformed the AR1 model at the horizon of interest. For this reason, it was further used to produce conditional forecasts.

The conditional forecasts of the BVAR model were compared to the forecasts of the QPM model. In general, for both inflation and GDP growth, the BVAR model forecasts are competitive with the QPM forecasts.

As the NBU forecasts are only available on a year-over-year basis, the conditional BVAR forecasts were transformed from annualized quarter-over-quarter indicators into year-over-year indicators to compare the forecast accuracy. The BVAR forecasts of inflation outperform the official NBU forecasts at the horizon of interest, whereas the opposite is true for the forecasts of GDP growth. In the short run, NBU forecasts dominate probably because the NBU is considering a broader information set during the forecasting process and different types of models specifically designed for short-run forecasting.

REFERENCES

Beechey, M., Österholm, P. (2010). Forecasting inflation in an inflation-targeting regime: A role for informative steady-state priors. International Journal of Forecasting, 26(2), 248-264. https://doi.org/10.1016/j.ijforecast.2009.10.006

Brazdik, F., Franta, M. (2017). A BVAR model for forecasting of Czech inflation. Working Papers 2017/7. Praha: Czech National Bank. Retrieved from https://www.cnb.cz/export/sites/cnb/en/economic-research/.galleries/research_publications/cnb_wp/cnbwp_2017_07.pdf

Carriero, A., Kapetanios, G., Marcellino, M. (2009). Forecasting exchange rates with a large Bayesian VAR. International Journal of Forecasting, 25(2), 400–417. https://doi.org/10.1016/j.ijforecast.2009.01.007

Chan, J., Jacobi, L., Zhu, D. (2019). Efficient selection of hyperparameters in large Bayesian VARs using automatic differentiation. CAMA Working Papers, 46/2019. The Australian National University. Retrieved from https://crawford.anu.edu.au/sites/default/files/publication/cama_crawford_anu_edu_au/2019-06/46_2019_chan_jacobi_zhu.pdf

Clark, T. E. (2011). Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility. Journal of Business and Economics Statistics, 29(3), 327–341. https://doi.org/10.1198/jbes.2010.09248

Del Negro, M., Schorfheide, F. (2004). Priors from general equilibrium models for VARs. International Economic Review, 45, 643–673. https://doi.org/10.1111/j.1468-2354.2004.00139.x

Dieppe, A., Legrand, R., van Roye, B. (2016). The BEAR toolbox. Working Paper Series, 1934. European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1934.en.pdf

Foroni, C., Marcellino, M., Stevanović, D. (2020). Forecasting the Covid-19 recession and recovery: lessons from the financial crisis. Working Paper Series, 2468. European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecb.wp2468~068eec9e3e.en.pdf

Giannone, D., Lenza, M., Primiceri, G., (2012). Prior selection for vector autoregressions. Working Paper Series, 1494. European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1494.pdf

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_en.pdf

Gustafsson, O., Villani, M., Stockhammar, P. (2020). Bayesian optimization of hyperparameters when the marginal likelihood is estimated by MCMC. Retrieved from https://arxiv.org/pdf/2004.10092.pdf

Iversen, J., Laseen, S., Lundvall, H., Söderström. U. (2016). Real-time forecasting for monetary policy analysis: The case of Sveriges Riksbank. CEPR Discussion Papers, 11203. Retrieved from http://archive.riksbank.se/Documents/Rapporter/Working_papers/2016/rap_wp318_160323.pdf

Jarocinski, M. (2010). Conditional forecasts and uncertainty about forecast revisions in vector autoregressions. Economics Letters, 108(3), 257–259. https://doi.org/10.1016/j.econlet.2010.05.022

Krüger, F., Clark, T. E., Ravazzolo, F. (2017). Using entropic tilting to combine BVAR forecasts with external nowcasts. Journal of Business & Economic Statistics, 35(3), 470–485. https://doi.org/10.1080/07350015.2015.1087856

Lenza, M., Primiceri, G. (2020). How to estimate a VAR after March 2020. Working Paper Series, 2461. European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecb.wp2461~fe732949ee.en.pdf

Robertson, J. C., Tallman, E. W., Whiteman, C. H. (2005). Forecasting using relative entropy. Journal of Money, Credit and Banking, 37(3), 383–401. https://doi.org/10.1353/mcb.2005.0034

Schorfheide, F., Song, D. (2015). Real-time forecasting with a mixed-frequency VAR. Journal of Business and Economic Statistics, 33(3), 366–380. https://doi.org/10.1080/07350015.2014.954707

Villani, M. (2009). Steady-state priors for vector autoregressions. Journal of Applied Econometrics, 24(4), 630-650. https://doi.org/10.1002/jae.1065

Waggoner, D. F., Zha, T., (1999). Conditional forecasts in dynamic multivariate models. The Review of Economics and Statistics, 81(4), 639–651. https://doi.org/10.1162/003465399558508

APPENDIX A

Gibbs sampling algorithm for BVAR with steady-state prior

1. Define the number of iterations It of the algorithm, and the burn-in sample u.

2. Define initial values β_0, B_0, Σ_0 for the algorithm. Obtain the initial value for U from β_0,

3. At iteration n, draw $\psi_{(n)}$, conditional on $\beta_{(n-1)}$, and $\Sigma_{(n-1)}$. Draw $\psi_{(n)}$, from a multivariate normal:

$$\pi(\psi|\beta_{(n-1)}\Sigma_{(n-1)}, y) \sim N(\bar{\psi}, \bar{\Lambda})$$

$$\bar{\Lambda} = \left[\Lambda_0^{-1} + U'\left(Z'Z \otimes \Sigma_{(n-1)}^{-1}\right)U\right]^{-1}, \bar{\psi} = \bar{\Lambda}\left[\Lambda_0^{-1}\psi_0 + U' vec(\Sigma_{(n-1)}^{-1}(Y - XB'_{(n-1)})Z)\right]$$

Reshape $\psi_{(n)}$, to obtain $F_{(n)}$.

4. Use $F_{(n)}$ to obtain $\hat{Y}$, $\hat{X}$ and $\hat{y}$.

5. Draw the value $\Sigma_{(n)}$, conditional on $B_{(n-1)}$ and $\psi_{(n)}$. Draw $\Sigma_{(n)}$, from an inverse Wishart distribution with scale matrix $\tilde{S}$ and degrees of freedom T:

$$\pi(\Sigma_{(n)}|B_{(n-1)}\psi_{(n)} \ , y) \sim IW(\tilde{S}, T)$$

$$\tilde{S} = \left(\hat{Y} - \hat{X}B_{(n-1)}\right)'\left(\hat{Y} - \hat{X}B_{(n-1)}\right)$$

6. Finally, draw $\beta_{(n)}$ conditional on $\Sigma_{(n)}$ and $\psi_{(n)}$, and reshape into $B_{(n)}$. Draw $\beta_{(n)}$, from a multivariate normal distribution with $\bar{\beta}$ mean and covariance matrix $\bar{\Omega}$:

$$\pi(\beta_{(n)} |\Sigma_{(n)}\psi_{(n)}, y) \sim N(\bar{\beta}, \bar{\Omega})$$

$$\bar{\Omega} = \left[\Omega_0^{-1} + \Sigma_{(n)}^{-1} \otimes \hat{X}'\hat{X}\right]^{-1}, \bar{\beta} = \left[\Omega_0^{-1}\beta_0 + \left(\Sigma_{(n)}^{-1} \otimes \hat{X}'\right)\hat{y}\right]$$

Update U from $B_{(n)}$.

7. Repeat until It iterations are realized, then discard the first Bu iterations.

Note that $\hat{y}_t$ is a demeaned data vector $\hat{y}_t = y_t - Fx_t$ and $A(L)\hat{y}_t = \varepsilon_t$ is a VAR in standard form conditional on F.

$\hat{X}$ and $\hat{Y}$ are defined as Y and X in accordance with (5) using $\hat{y}_t$ rather than $\hat{y}_t$.

APPENDIX B. TABLES

Table B.1. Data Used for the Research

Series	Name	Definition	Source
Foreign output	gdpw1 gdpw2 gdpw3	Trade-weighted index of real GDP of major trading partners (3,5 and 40* countries)	NBU staff calculations, based on the data from national statistics committees (NSC)
Foreign output	cpineerw1 cpineerw2 cpineerw3	Trade-weighted index of CPI of major trading partners (3*, 5 and 40 countries)	NBU staff calculations, based on the data from NSC
	ppineerw1 ppineerw2	Trade-weighted index of PPI of major trading partners (3 and 5 countries)	NBU staff calculations, based on the data from NSC
Commodity prices on foreign markets	wbnonen wben fao psteel pgrains	Non-energy commodities price index Energy commodities price index FAO price index Export price of steel Export price of grains	World bank commodity prices, FAO database, export and import prices from SSSU
Foreign interest rate	iw1 iw2	1-month LIBOR Rate* overnight LIBOR Rate	Thomson Reuters Datastream
Domestic output	gdpua	Ukrainian GDP at constant prices*	State Statistics Service of Ukraine (SSSU)
Domestic price level	cpiua	CPI*	SSSU
Domestic interest rate	iua	NBU policy rate*	NBU
Nominal effective exchange rate (+ means an appreciation)	neer1 neer2 neer3	Real effective exchange rate deflated by CPI (3*, 5 and 40 countries)	NBU staff calculations, based on the data from national statistics committees, Bloomberg, NBU data
Real effective exchange rate (+means an appreciation)	reercpiw1 reercpiw2 reercpiw3	Real effective exchange rate deflated by CPI (3 and 5 countries) (3,5 and 40* countries)	NBU staff calculations, based on the data from national statistics committees, Bloomberg, NBU data
	reerppiw1 reerppiw2	Real effective exchange rate deflated by PPI (3 and 5 countries)	
Terms of trade	totw1 totw2	Commodity terms of trade, based on IMF methodology Ratio between the index of export prices for grains and metals and the index of import prices for oil and gas*	NBU staff calculations, based on the SSSU data
Wage	nwage rwage	Average nominal wage Average real wage*	SSSU
Monetary aggregate	m2	M2	NBU

Note: the time series entering QPM model marked with asterisk.

Table B.2. Lag Length Criteria

	LR	FPE	AIC	SC	HQ
MB	3	2	5	1	1
MA_P	2	4	5	1	1
MA_TOT	4	2	5	1	1

Note: numbers in the Table 1 indicate lag order selected by the criterion:

LR: sequential modified LR test statistic (each test at 5% level)

FPE: Final prediction error

AIC: Akaike information criterion

SC: Schwarz information criterion

HQ: Hannan-Quinn information criterion

Table B.3. Lag Length Criteria

	Min value	Max value	Step size
Autoregressive coefficient	0.2	1.0	0.1
Overall tightness (λ_1)	0.05	0.20	0.01
Cross—variable weighting (λ_2)	0.1	1.0	0.1
Lag decay (λ_3)	1.0	2.0	0.2

Table B.4. Steady State prior Distributions (Villani)

	Regime 1: 1980q1 1992q4				Regime 2: 1993q1 2005q4			
	Mean	Var	95%	Interval	Mean	Var	95%	Interval
GDPw	2.5	0.8	1.0	4.0	2.5	0.3	2.0	3.0
CPIw	4.0	0.5	3.0	5.0	2.0	0.3	1.5	2.5
irw	7.0	0.5	6.0	8.0	5.0	0.3	4.5	5.5
GDP	2.3	0.6	1.0	3.5	2.3	0.1	2.0	2.5
CPI	7.0	0.5	6.0	8.0	2.0	0.2	1.7	2.3
ir	8.5	0.8	7.0	10.0	4.3	0.1	4.0	4.5
REER	3.9	0.3	3.4	4.5	3.9	0.0	3.9	4.0

Table B.5. Steady State prior Distributions (Brazdik and Franta)

	Regime 1: 2008q3 2010q1				Regime 1: 2010q2 2013q3				Regime 1: 2013q4 2016q4			
	Mean	Var	95%	Interval	Mean	Var	95%	Interval	Mean	Var	95%	Interval
GDPw	9.4	3.1	3.4	15.4	8.9	2.0	4.9	12.9	7.2	1.0	5.2	9.2
CPIw	2.0	1.5	-1.0	5.0	2.0	1.0	0.0	4.0	2.0	0.5	1.0	3.0
Euribor 3m	4.0	1.5	1.0	7.0	4.0	1.0	2.0	6.0	3.5	0.5	2.5	4.5
GDP	5.0	1.5	2.0	8.0	4.0	1.0	2.0	6.0	3.0	0.5	2.0	4.0
CPI	3.0	1.0	1.0	5.0	2.0	0.5	1.0	3.0	2.0	0.3	1.5	2.5
Pribor 3m	3.0	4.1	-0.5	15.4	3.0	1.3	0.5	5.5	3.0	0.8	1.5	4.5
CZK/Euro	-2.4	3.1	-8.4	3.6	-2.4	2.0	-6.4	1.6	-1.5	1.0	-3.5	0.5

APPENDIX C. FIGURES
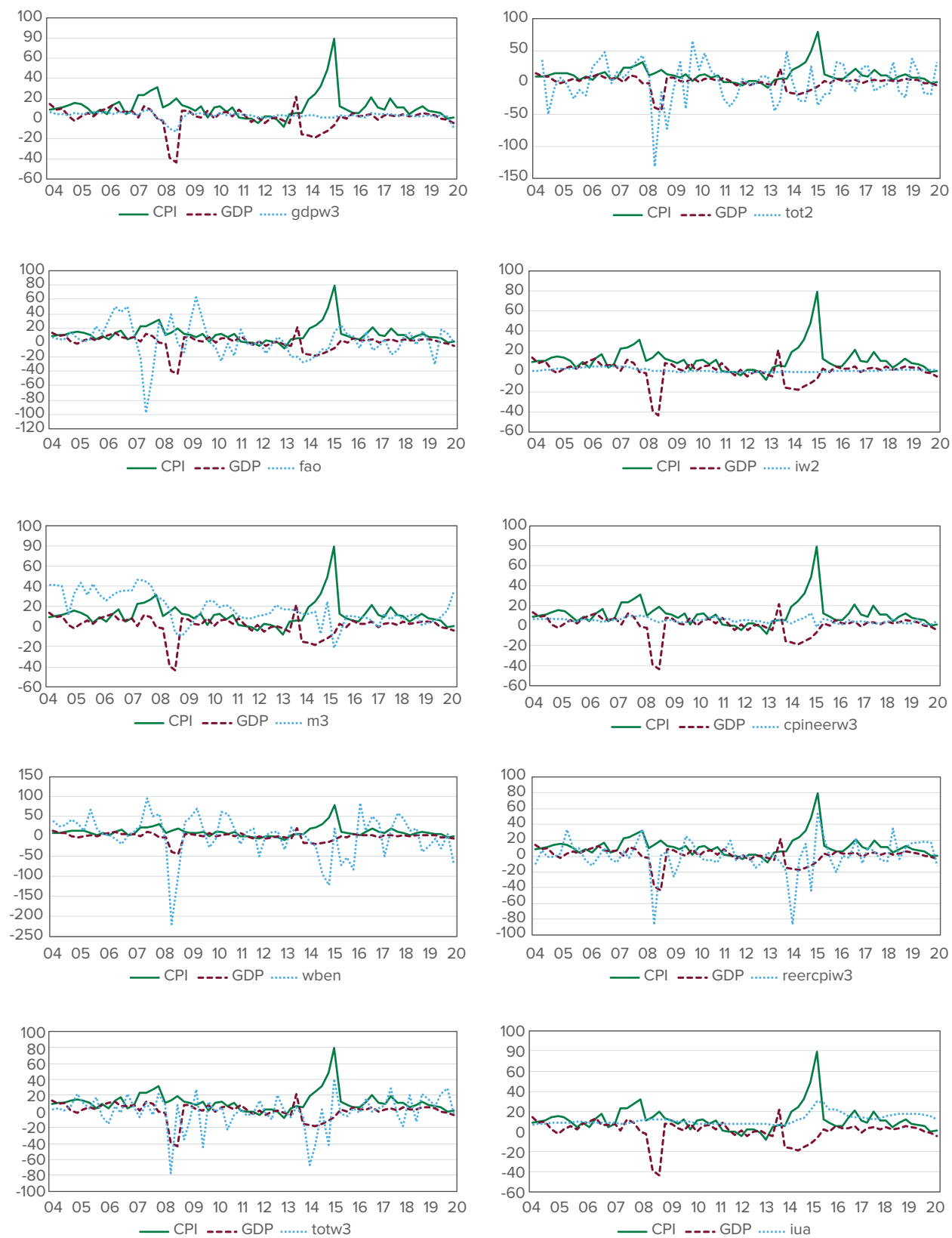
Figure C.1. CPI, GDP and Series Chosen for Models, yoy, in logs
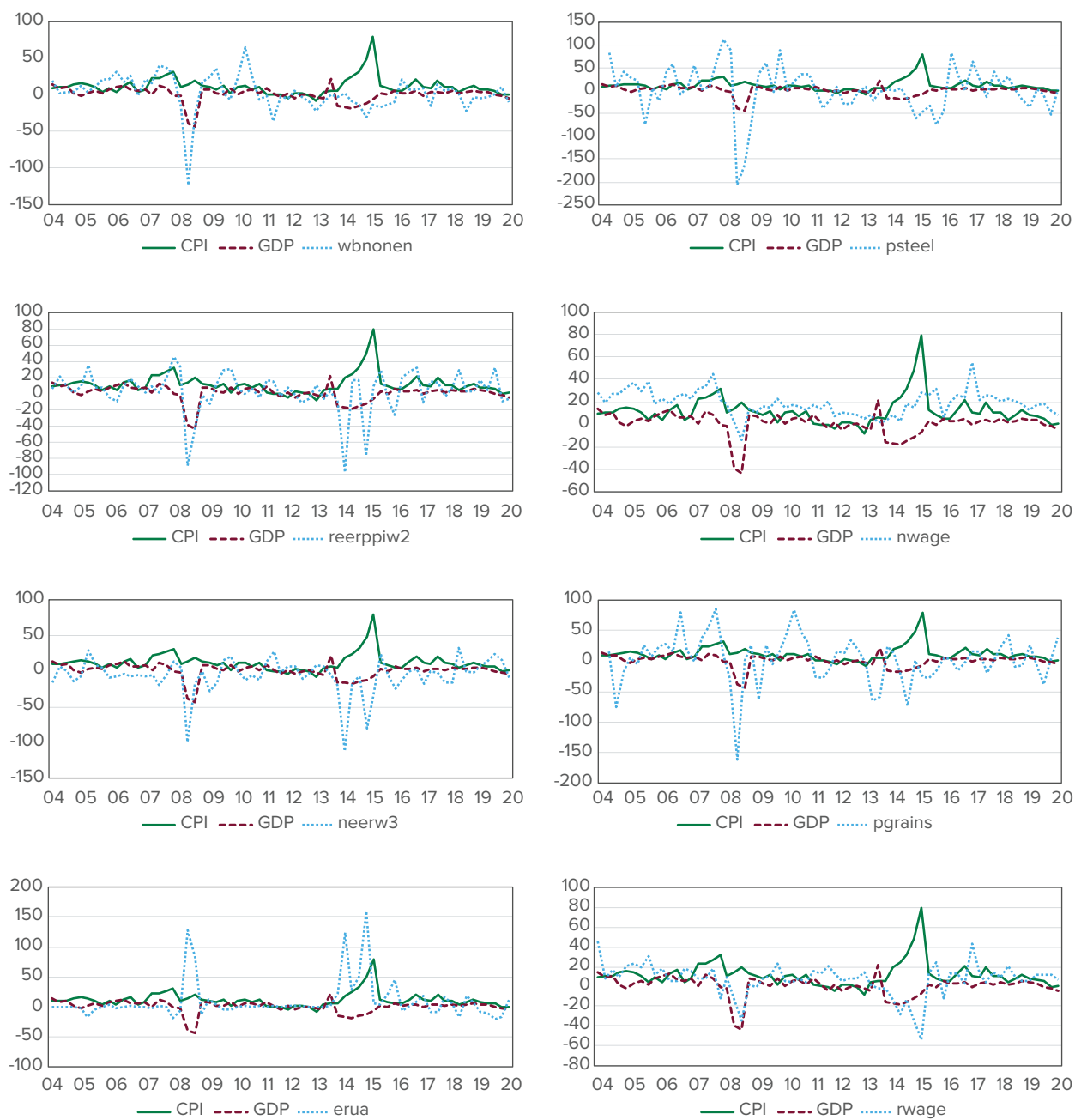
Figure C.1 (continued). CPI, GDP and Series Chosen for Models, yoy, in logs
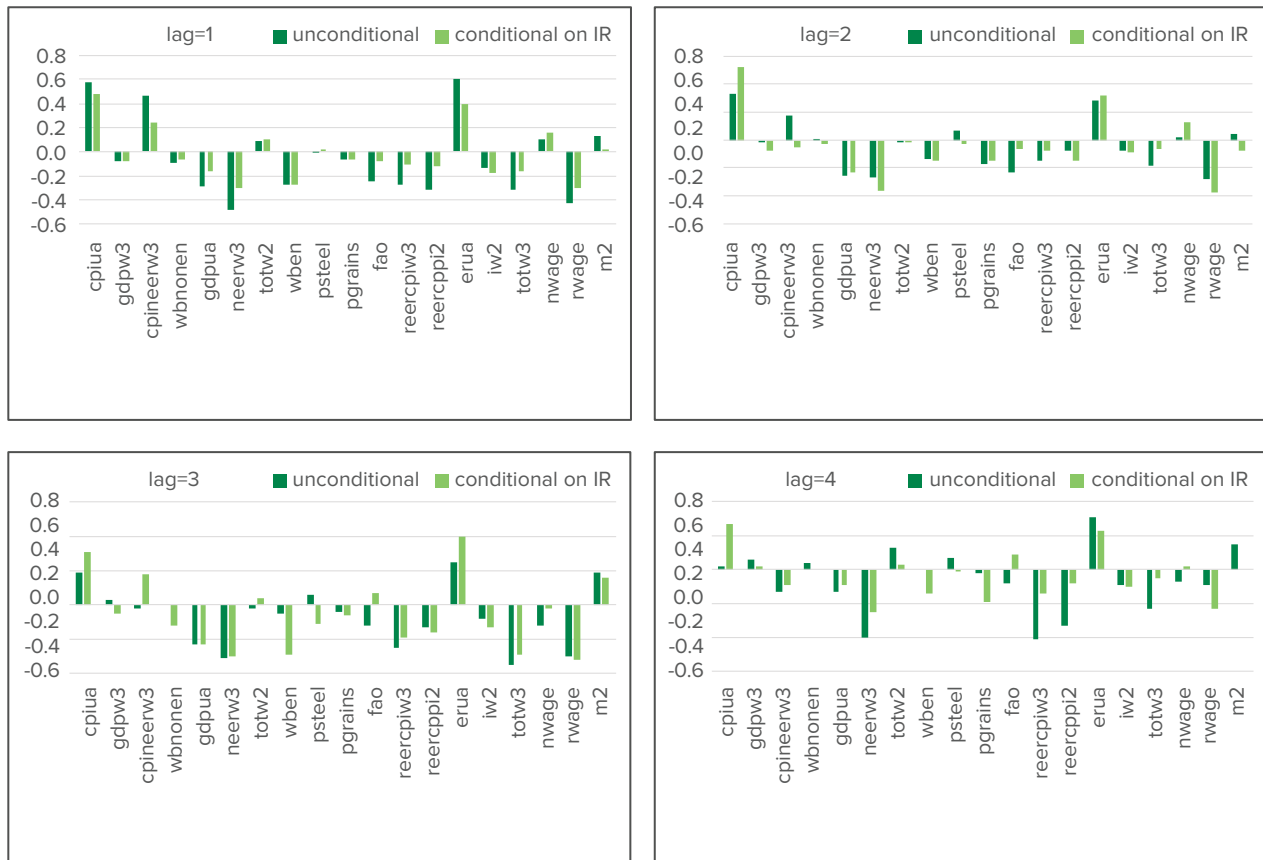
Figure C.2. Conditional and Unconditional Correlations
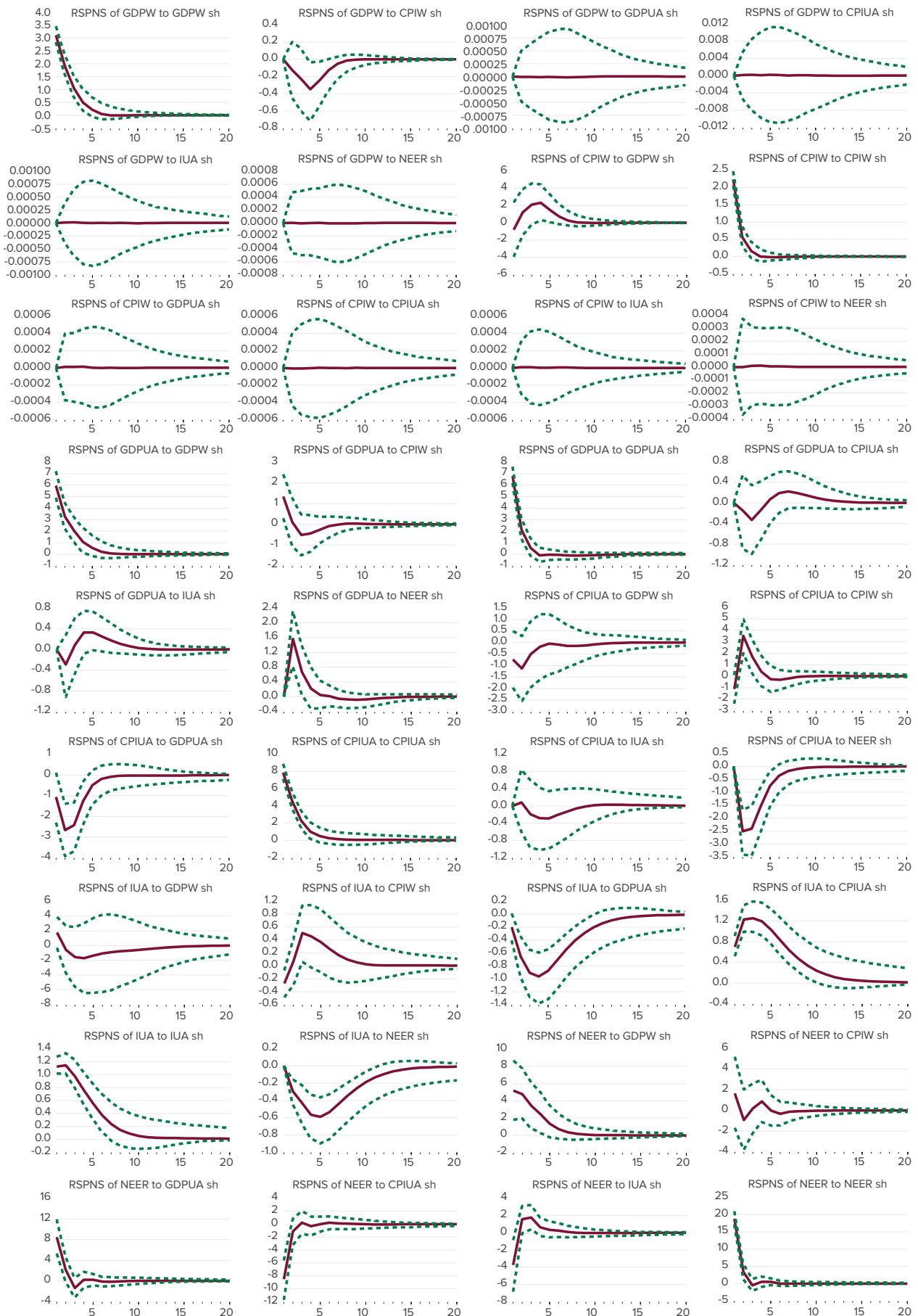
Figure C.3. Impulse Response Functions, MB Model
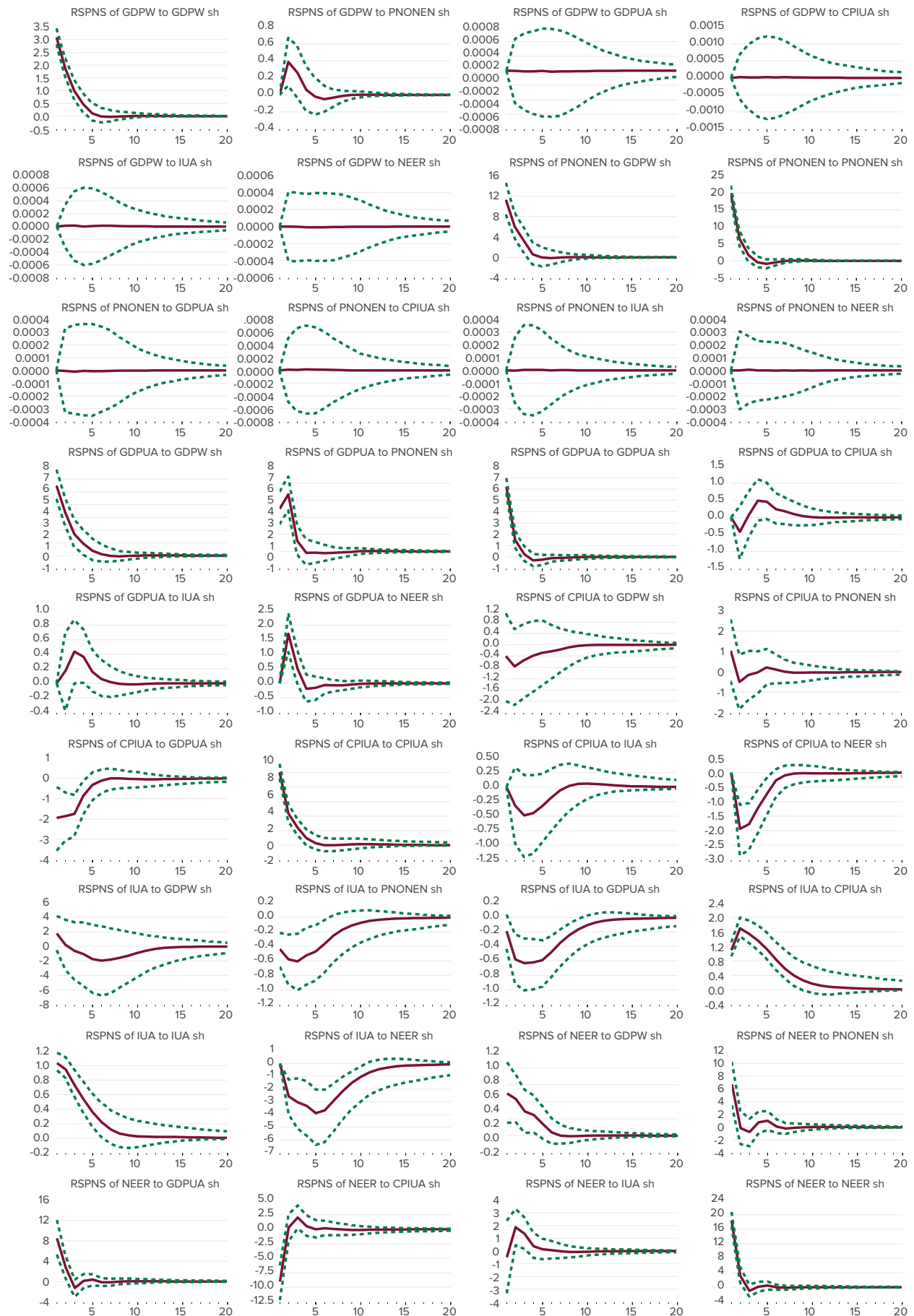
Figure C.4. Impulse Response Functions, MA_P Model
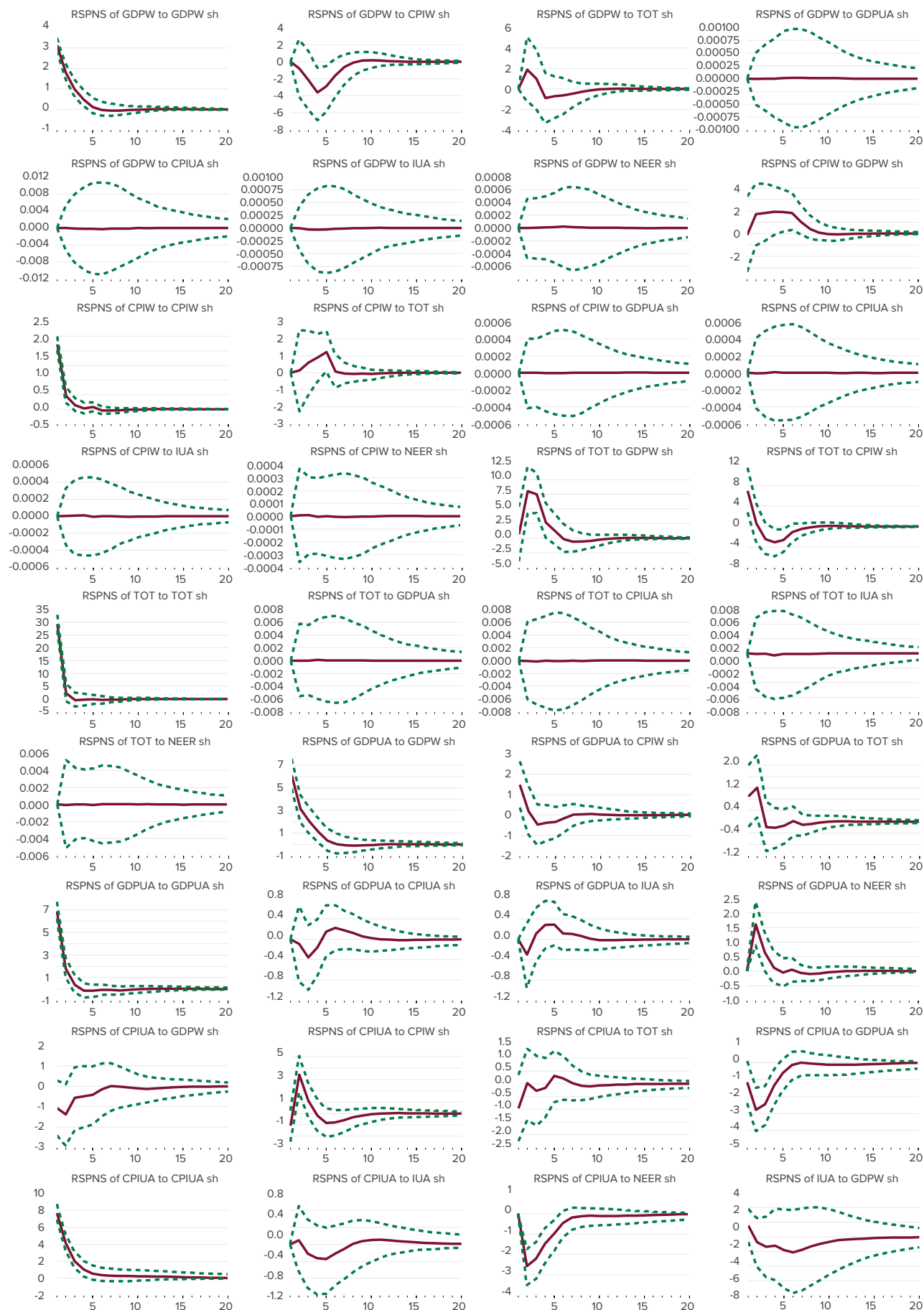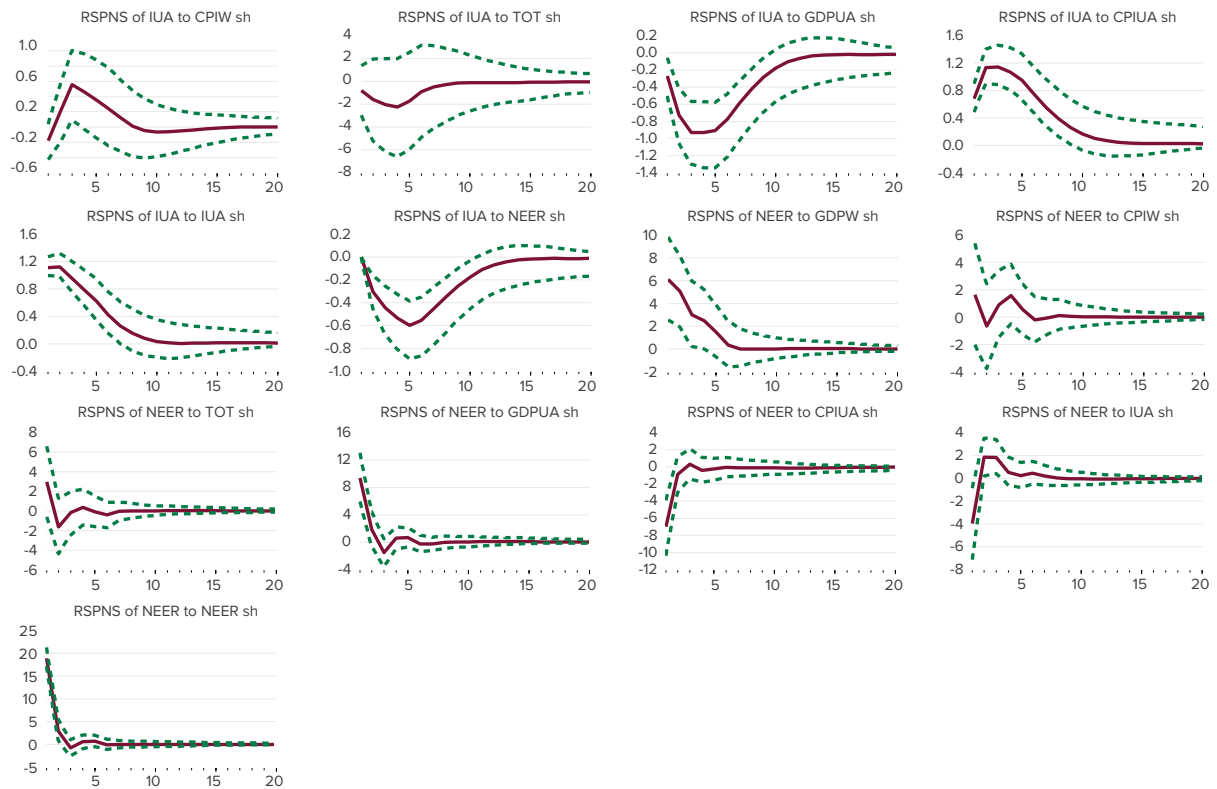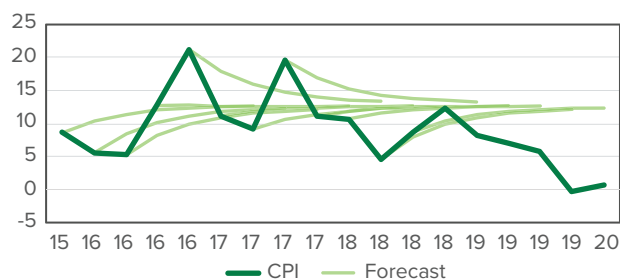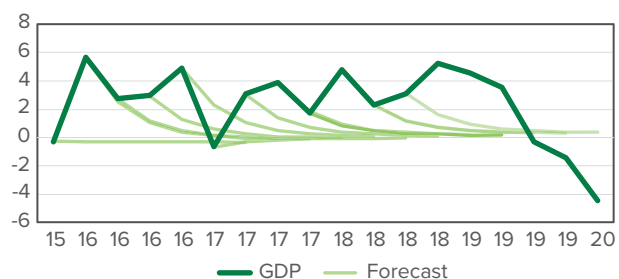
Figure C.5. Impulse Response Functions, MA_TOT Model

Figure C.5 (continued). Impulse Response Functions, MA_TOT Model

a) AR1

b) MA_TOT
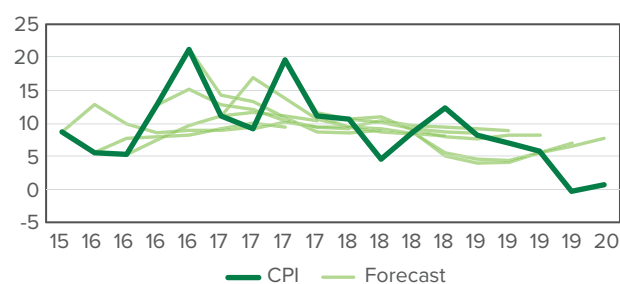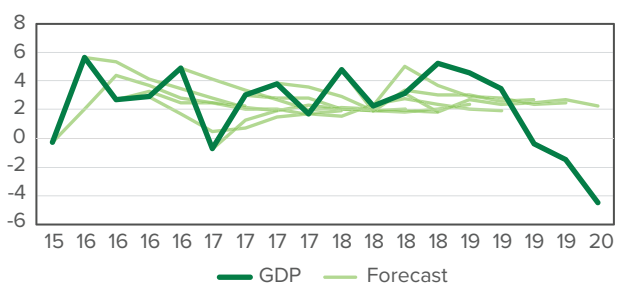
c) MA_P
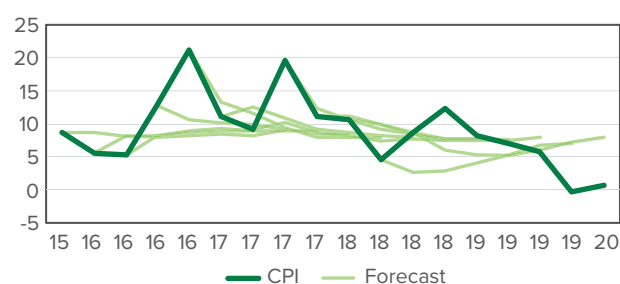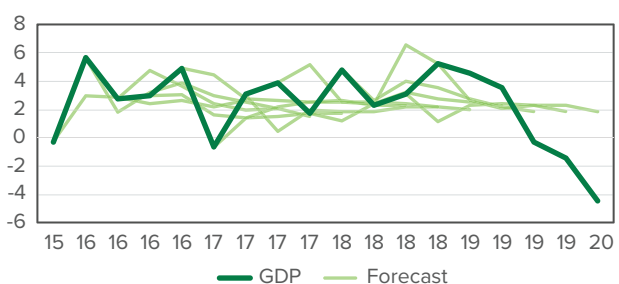
d) MB
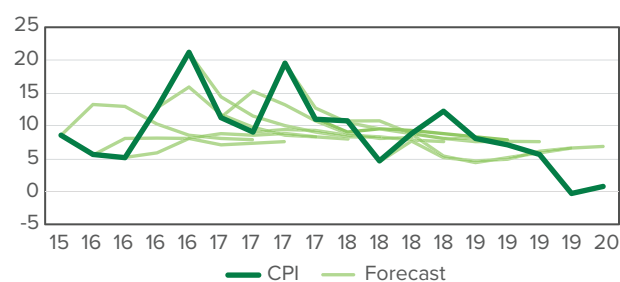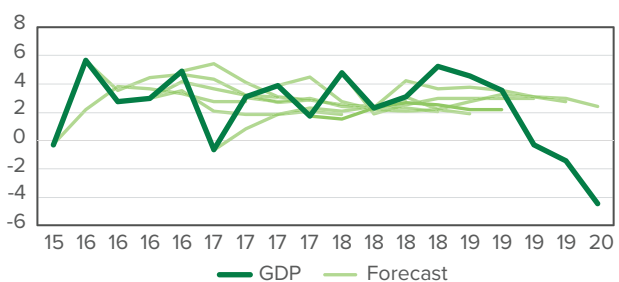
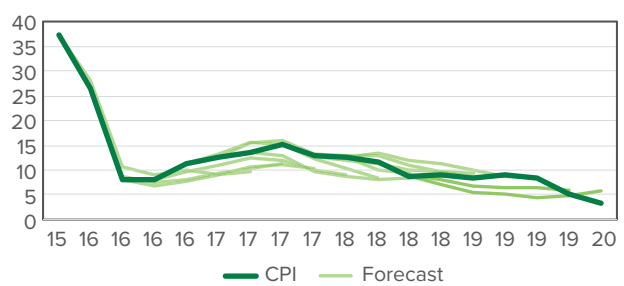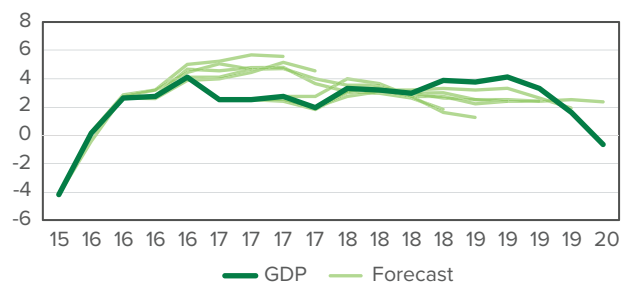**Figure C.6.** Unconditional BVAR and AR1 Forecasts

a) MA_TOT(cond)
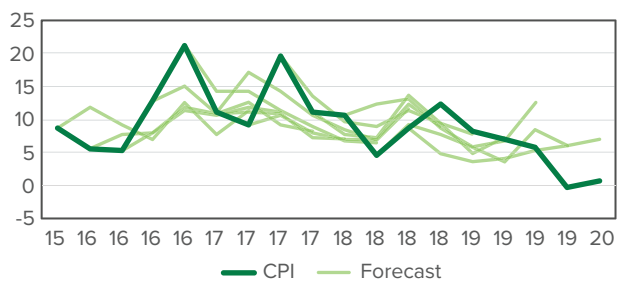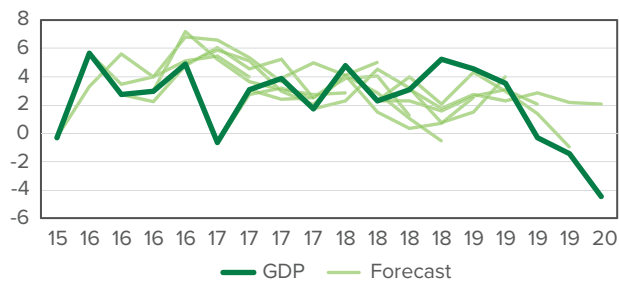
b) QPM
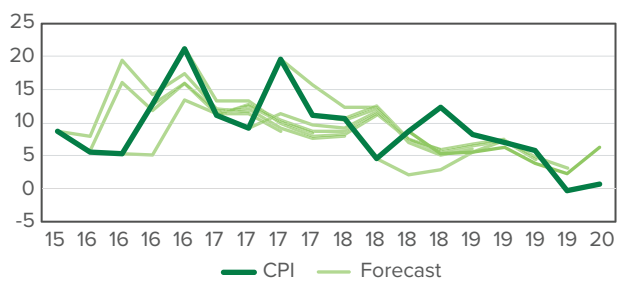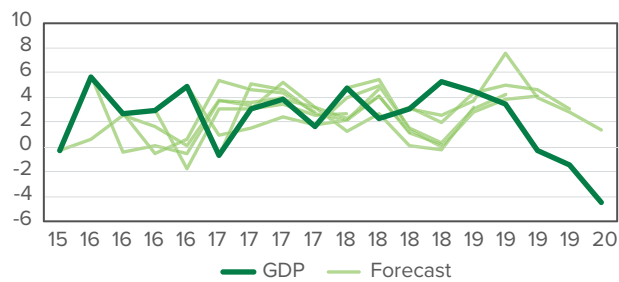
Figure C.7. Conditional BVAR and QPM Forecasts
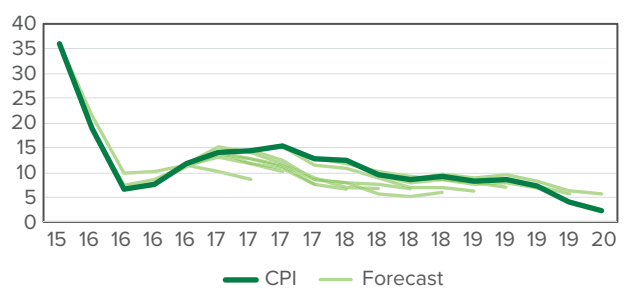
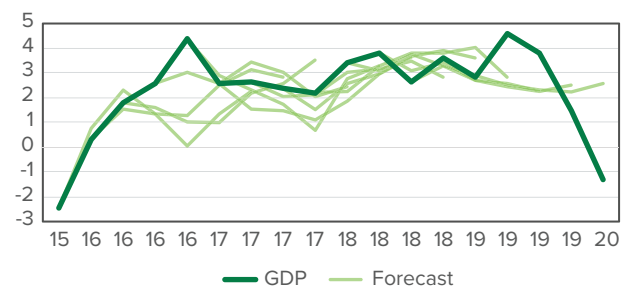a) MA_TOT(cond)

b) NBU

Figure C.8. Conditional BVAR and Official NBU Forecasts for the Indicators on Year-over-Year Basis

A NEW FINANCIAL STRESS INDEX FOR UKRAINE

VLADYSLAV FILATOV[ab1]

[a]National Bank of Ukraine
[b]National 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

---

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. Futhermore, 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
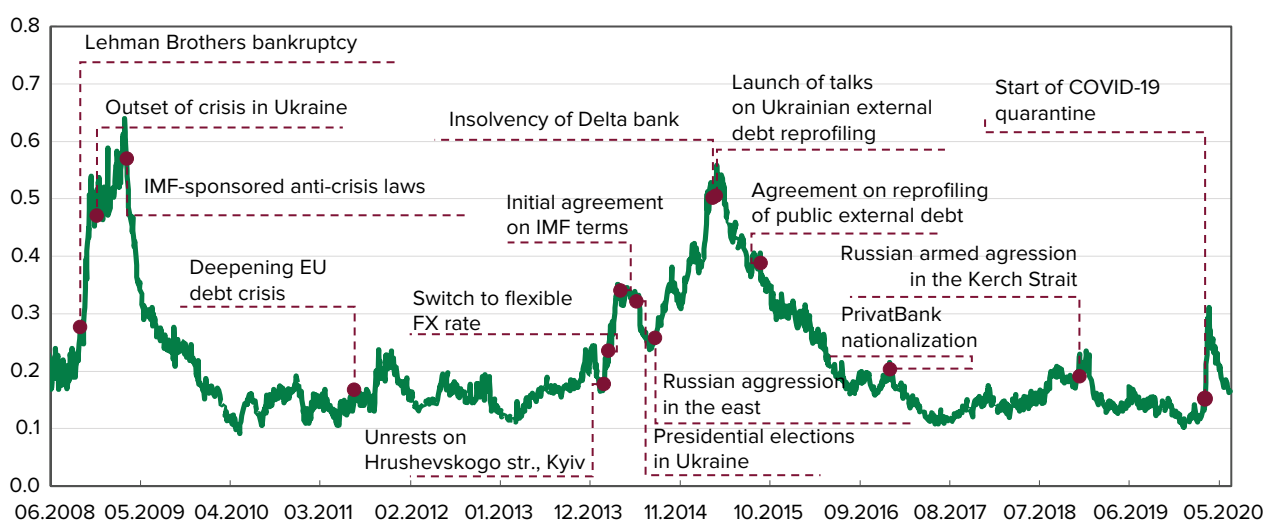

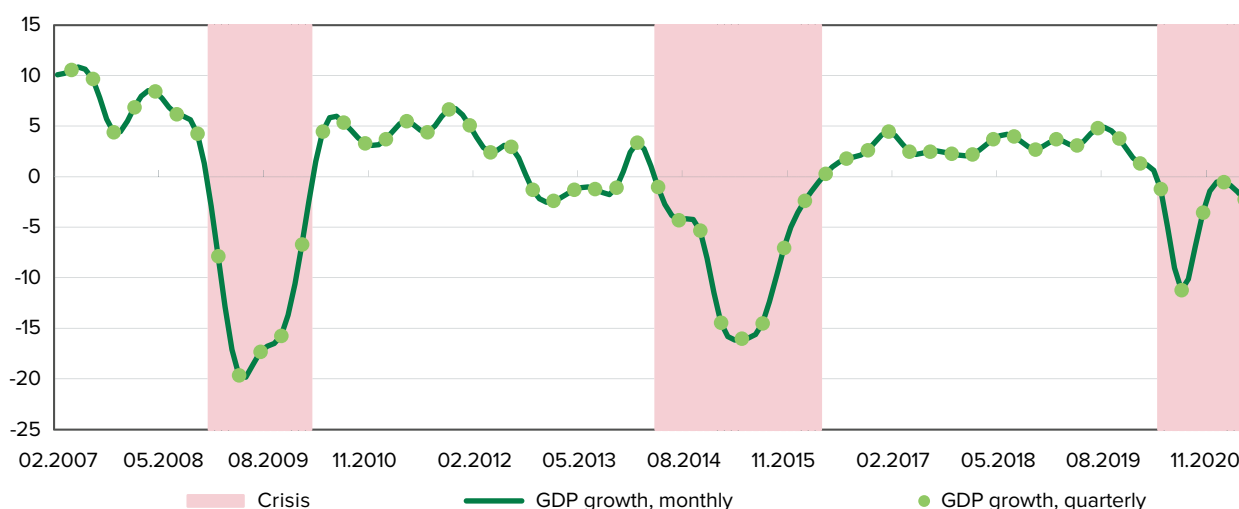**Figure 1.** FSI for Ukraine (version 1.0)

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:

$$log \frac{p}{1-p} = \beta_0 + \beta_1 x_1, \qquad (1)$$

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 ≤ 1%), significant (1% < P-value ≤ 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.

$$L(\theta) = \theta T_1 + (1-\theta)T_2 \qquad (2)$$

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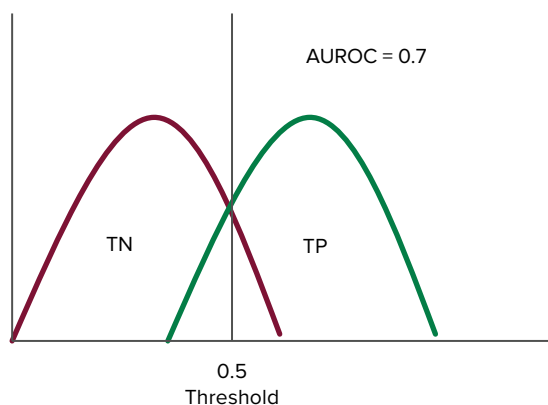
Figure 3. Overlap of True Negative and True Positive in AUROC Analysis

Figure 4. AUROC Curve Example

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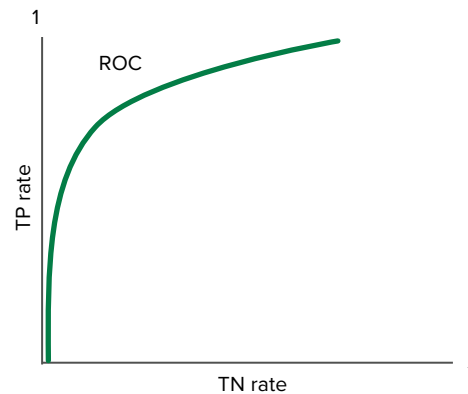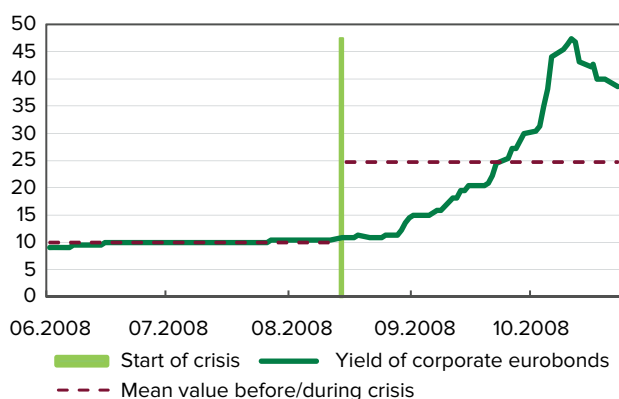
if the indicator is positively correlated with the crisis and vice versa. The indicator passes this test if the difference between the means and standard deviations in peaceful and crisis periods is higher than the standard deviation for the whole observed period. This should be true for all the three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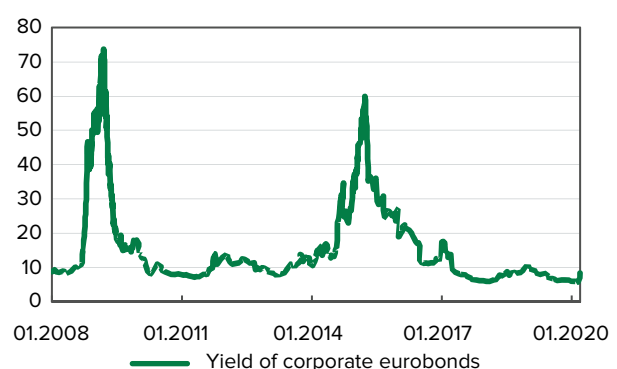

Figure 5. Yield of Corporate Eurobonds in 2008-2020

---

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:

$$\frac{X_{i,t} - \mu_i}{\sigma_i}, \tag{3}$$

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:

$$\frac{(X_{i,t} - MIN_i)}{(MAX_i - MIN_i)}. \tag{4}$$

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 $\rho_{ji,t}$ 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.

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

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

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

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 β=0.89, β=0.93, and β=0.97.

Table 4 reports the AUROC results applied to these alternative indices. We observe that β=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.

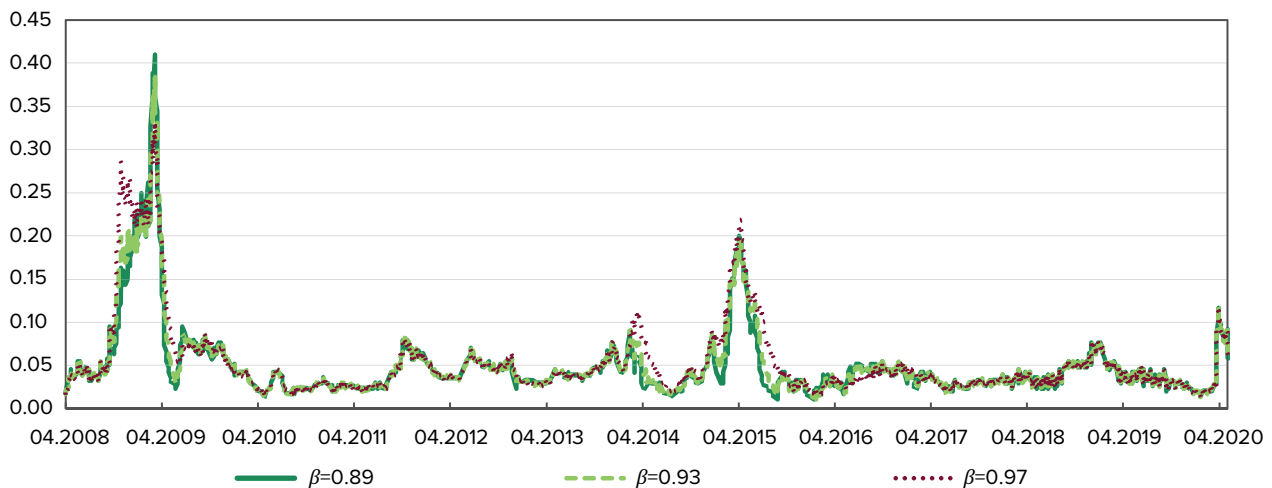[5] The outcomes of logit-regressions are available on request.

Figure 6. EWMA Aggregation with Different Values of β

Table 4. AUROC Testing of EWMA Results with Different Parameters

	Survey crisis	Survey crisis peaks
β=0.89	0.771	0.808
β=0.91	0.794	0.833
β=0.93	0.816	0.863
β=0.95	0.840	0.898
β=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.
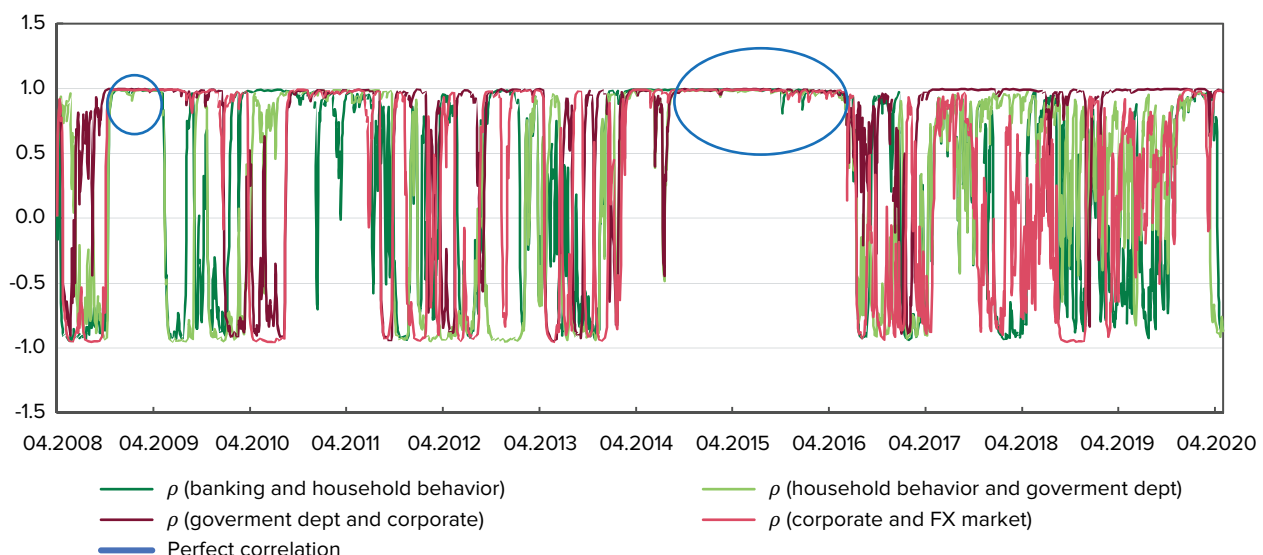
5. RESULTS

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)
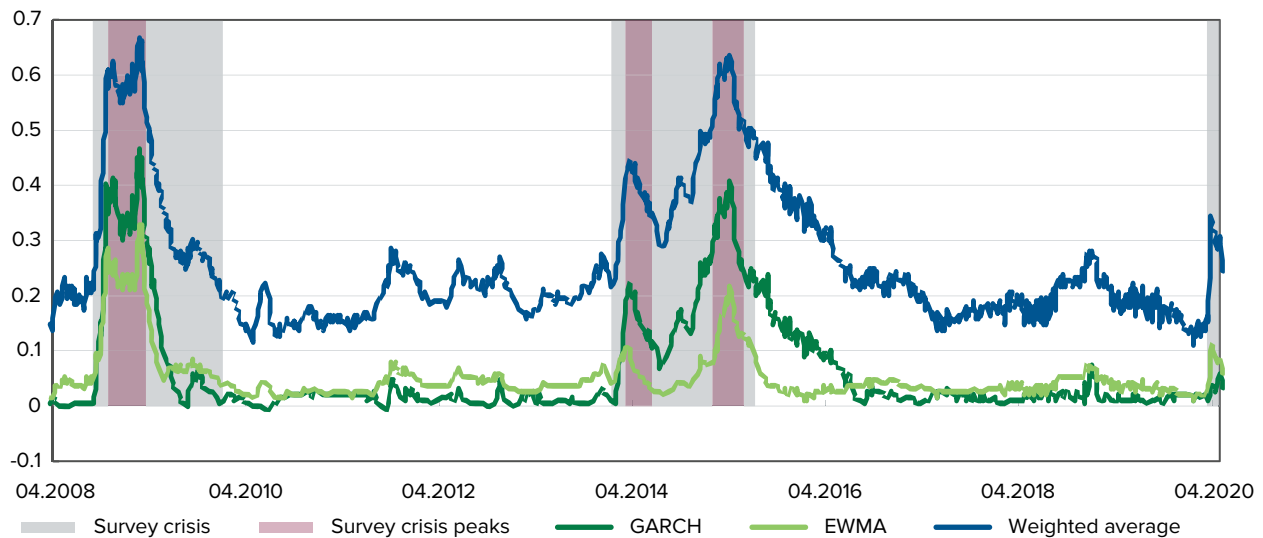

Figure 7. Correlations Between Subindices

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.

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.

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

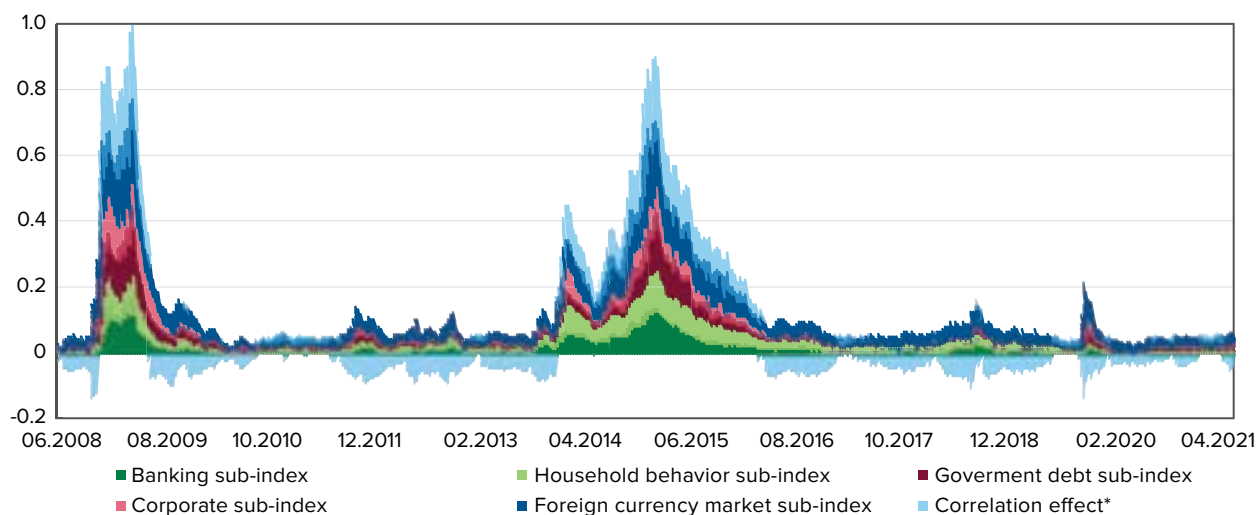**Figure 9.** The Final Version of FSI 2.0

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 COVD-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.
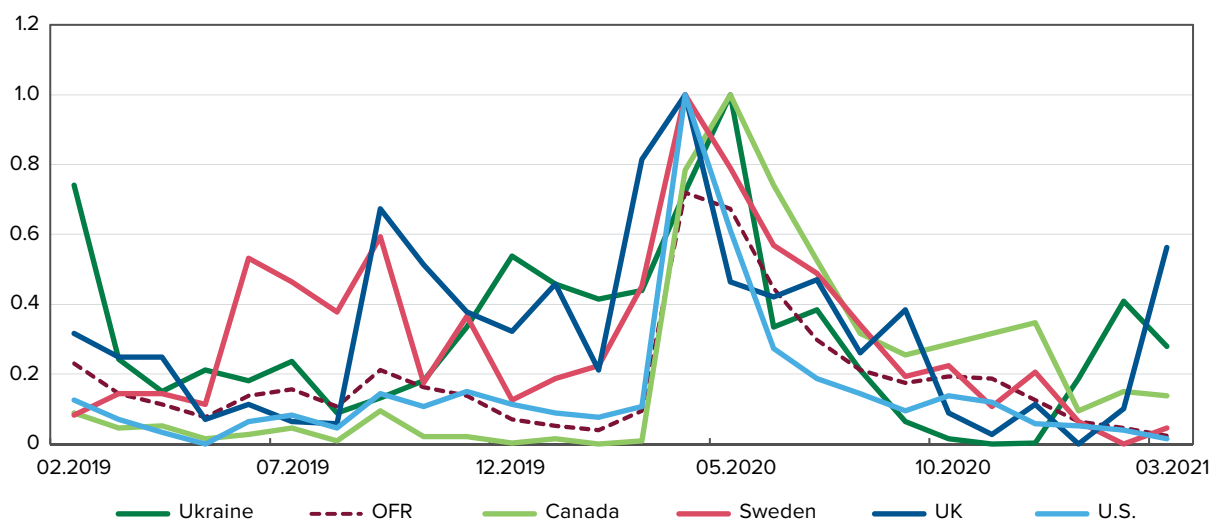

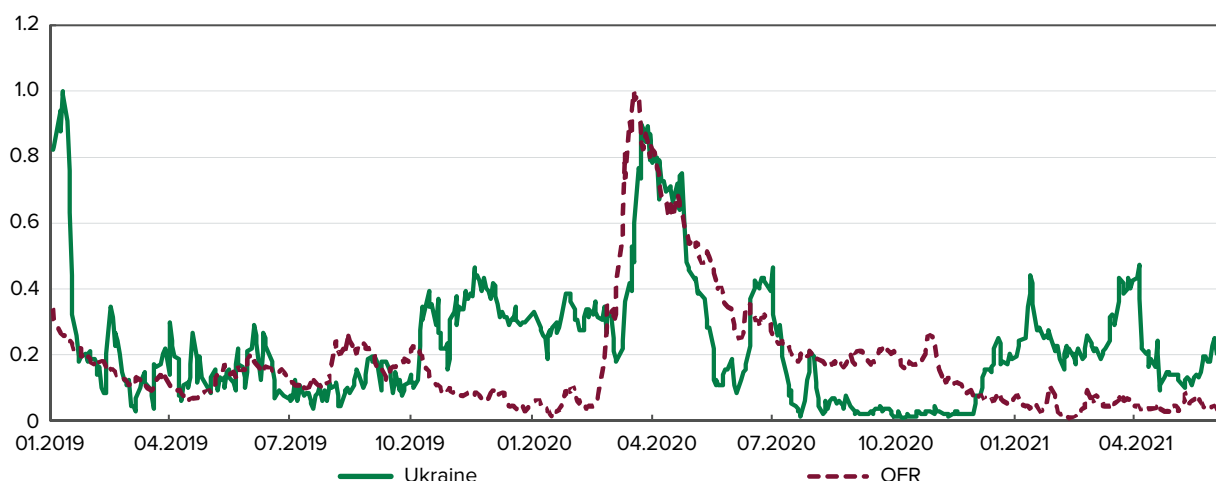**Figure 11.** Ukrainian FSI vs Other FSIs

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).
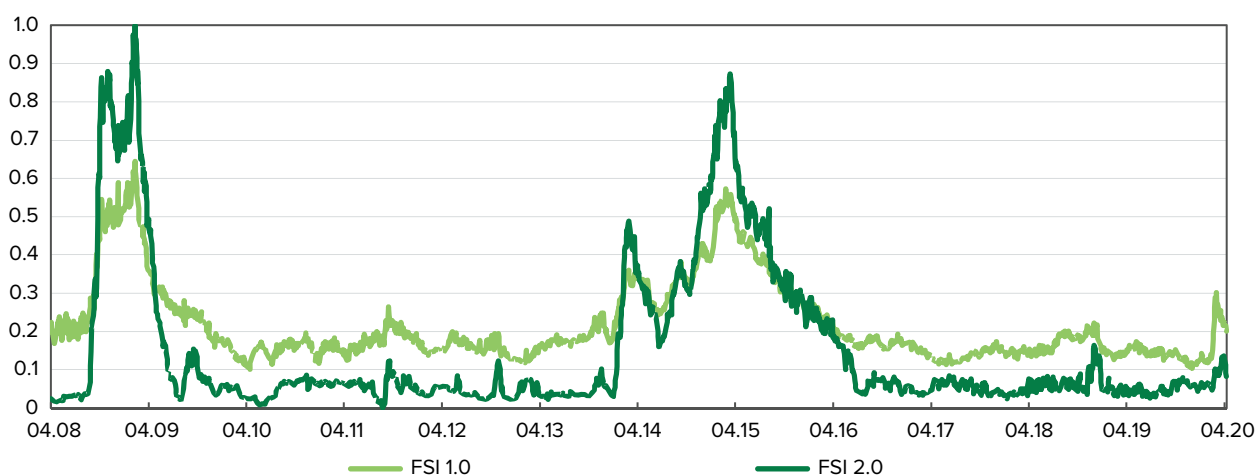
APPENDIX A

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):

$$r_t = \mu_t + a_t \tag{1}$$

$$a_t = H_t^{1/2} z_t \tag{2}$$

$$H_t = D_t R_t D_t \tag{3}$$

Notation:

r_t: n × 1 vector of log returns for n assets at time t

a_t: n × 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 × 1 vector of the expected value of the conditional r_t

H_t: n × n matrix of conditional variances of a_t at time t

$H_t^{1/2}$:   Any n × 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 × n, diagonal matrix of conditional standard deviations of a_t at time t

R_t: n × n conditional correlation matrix of a_t at time t

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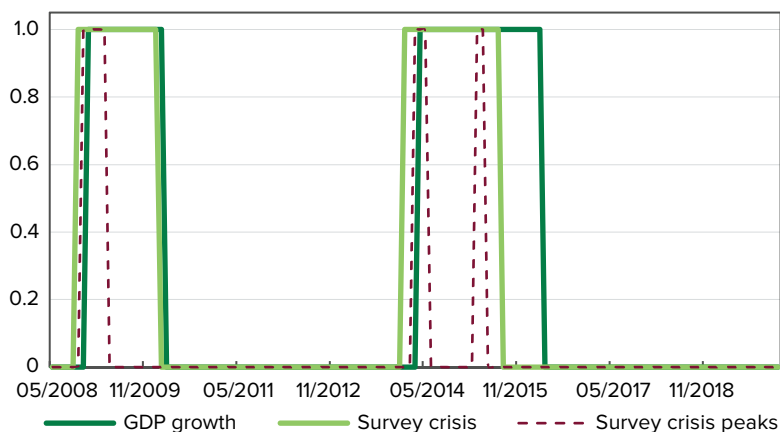
APPENDIX C

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)

APPENDIX D

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

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: $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.