Hlazunov, Anatolii

Article

Determinants of corporate credit growth in Ukraine : the application of bank lending survey data

Visnyk Nacional'noho Banku Ukraïny

Reference: Hlazunov, Anatolii (2022). Determinants of corporate credit growth in Ukraine : the application of bank lending survey data. In: Visnyk Nacional'noho Banku Ukraïny (254), S. 4 - 14.
https://journal.bank.gov.ua/uploads/issues/254_eng.pdf.
doi:10.26531/vnbu2022.254.01.

National Bank of Ukraine

ISSN 2414-987X

VISNYK
OF THE NATIONAL BANK OF UKRAINE

No. 254
2022

VISNYK

No. 254
2022

OF THE NATIONAL BANK OF UKRAINE

https://journal.bank.gov.ua

Founder and publisher: National Bank of Ukraine

Editorial Board ensures compliance with ethical standards of publishing

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

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

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

PREFACE BY THE EDITOR-IN-CHIEF

Dear readers,

This issue of the Visnyk of the National Bank of Ukraine presents a few best studies submitted to the Annual Research Conference for Students and Young Researchers organized by the National Bank of Ukraine (NBU) and Kyiv School of Economics (KSE). The authors offer some noteworthy ideas that were highly rated by jurors of the conference and mature referees.

In the first paper included in the issue, *Determinants of Corporate Credit Growth in Ukraine: The Application of Bank Lending Survey Data*, Anatolii Hlazunov examines the determinants of corporate lending in Ukraine. Applying a panel-ordered logit model to disaggregated data from a Ukrainian bank lending survey, the author first transforms categorical data from the poll into a continuous index (CSI) that measures credit standard tightening. After this, the paper investigates how the index affects new corporate lending in both national and foreign currencies. The revealed negative impact of the CSI on corporate loans is concluded to be more pronounced for smaller banks.

The second paper authored by Diana Balioz, *Short-Term Forecasting of Global Energy and Metal Prices: VAR and VECM Approaches*, tests a set of multivariate models to forecast global prices of 1) crude oil, 2) natural gas, 3) iron ore, and 4) steel. The study verifies many fundamentals for metal and energy price predictions and confirms that Kilian's index of global real economic activity is a useful proxy for global demand to forecast such prices. No individual model is found to outperform others consistently throughout the forecast horizon. However, the short-term hands-on framework considered in the paper is argued to be a useful forecasting tool for central bank policymakers and researchers. The specific conclusions on the price projections obtained from the models could also be used further for the longer-term forecasting of commodity prices.

The third paper of the issue by Tetiana Stasiuk and Yuriy Kleban elaborates on the problems of *Cryptocurrency Price Forecasting*. This research uses the data of Binance (the most popular exchange in Ukraine) for the period from 7 June 2020 to 4 January 2023 to forecast the prices of Bitcoin, Ethereum, Ripple, and Dogecoin. Specifically, the authors compare the performance of machine learning techniques and traditional econometric methodologies to predict cryptocurrency prices. The recurrent neural network of long-term memory is shown to produce significantly better forecasting outcomes according to the RMSE, MAE, and MAPE criteria, compared to the forecasts made by a NAÏVE approach, results from the best-fitted ARIMA model, and the results of the FB Prophet.

We believe that the research insights, research results, and conclusions presented in the papers of the current issue of our journal will be useful for practical implementation. The NBU Visnyk calls academic researchers, experts with practical backgrounds in economics, banking, and finance, and policymakers to promote and discuss their research ideas at the NBU-NBP Annual Research Conference and the NBU-KSE Conference for Students and Young Researchers. We welcome research contributors to submit their original fundamental and applied studies for publication in the Visnyk of the National Bank of Ukraine. We look forward to cooperating with you!

Best regards,
Mihnea Constantinescu

Visnyk of the National Bank of Ukraine, 2022, No. 254, pp. 4-14

DETERMINANTS OF CORPORATE CREDIT GROWTH IN UKRAINE: THE APPLICATION OF BANK LENDING SURVEY DATA[1]

ANATOLII HLAZUNOV[ab]

[a]National Bank of Ukraine
[b]National University of Kyiv-Mohyla Academy
E-mail: anatolii.hlazunov@bank.gov.ua

Abstract This study investigates the determinants of corporate lending in Ukraine, with a focus on distinguishing between supply and demand factors. It uses a two-step process to build a credit standards index (CSI) based on disaggregated data from a Ukrainian bank lending survey (BLS). This paper describes the factors that are significant for corporate lending development in Ukraine. It contributes to the existing literature by developing a measure of corporate loan supply and analyzing its ability to explain corporate credit growth in Ukraine by using bank-level BLS data. First, a panel ordered logit model is used to transform categorical data into a continuous index that measures the likelihood of credit standard tightening. Second, the study examines how this index affects new corporate lending in both national and foreign currencies. It is found that the credit standard index is influenced by exchange rate movements (with depreciations leading to tighter standards), bank liquidity, and bank competition. It is also demonstrated that the CSI has a negative impact on corporate loans in national currency, with a more pronounced effect on smaller banks.

JEL Codes G22, E44, C33

Keywords supply and demand of corporate lending, bank lending survey data, bank lending standards

1. INTRODUCTION

Crisis events and their effects on the lending market are interesting due to the complex relationship mechanisms behind them. Pham et al. (2021) highlight the importance of exogenous shocks on bank lending. The researchers show that the military conflict with russia-backed separatists in Q1 2014 harmed Ukrainian banks. As a result, conflict-exposed banks generated higher levels of non-performing loans (NPLs) and issued fewer new loans to businesses following the onset of the crisis. These effects are observed more clearly in the local markets that are closer geographically to the conflict area. However, the 2014–2015 crisis was not the only cause of NPL accumulation, but rather a trigger. Vyshnevskyi and Sohn (2023) provide empirical evidence that the NPL problem in Ukraine was caused by related party lending and issues with the banks' capital adequacy. The Ukrainian lending market faced new crises in 2020 and 2022, the latter being triggered by russia's full-scale invasion. The war influenced both supply and demand for corporate loans (NBU, 2022). The decline in business activity decreased demand, and an unfavorable macroeconomic environment reduced the risk appetite of the banks, resulting in tighter lending conditions. This research offers insights into the factors that are significant for the development of corporate lending in Ukraine.

This study examines the determinants of corporate lending in Ukraine. It focuses on two main research questions: (i) What bank-level and macro factors influence proxy banks' decisions to change their lending standards for corporations? (ii) What are the effects of the factors in determining corporate lending in Ukraine, and specifically, what is the impact of corporate lending standards as a loan supply factor? To answer both questions, the author uses

[1] The author would like to thank Professor Ugo Panizza for his valuable guidance and comments. The author is also grateful to the BCC program, the Graduate Institute (Geneva), and the National Bank of Ukraine for providing the data and resources used in this study.

Available at https://doi.org/10.26531/vnbu2022.254.01.

a two-step process to distinguish between the supply and demand factors of corporate lending.

In the first step, a panel ordered logit model is used to transform categorical survey data into a continuous credit standards index (CSI).[2] In this set-up, a higher index value indicates an increased likelihood of tightening corporate lending standards. The results show that faster economic growth, higher liquidity, and competition among banks lead to looser credit standards for Ukrainian businesses, whereas hryvnia depreciation and elevated interest rates lead to stricter bank requirements for borrowers.

Second, this paper explores the influence of the CSI on new corporate lending, while controlling for economic activity, interbank interest rates, deposit growth, liquidity, and the share of non-performing loans (NPLs). This study demonstrates that in six months the negative effect of tighter lending standards starts to have a bearing on the lower level of new hryvnia corporate lending. Small banks experience more pronounced effects than large banks. Moreover, small banks significantly affect domestic and foreign currency loans. This paper also ascertains the effect of economic activity on total assets, depending on the share of government securities, government bonds, and deposit certificates. GDP growth is found to be positively correlated with both domestic and foreign currency corporate lending, whereas new NPLs are negatively correlated with new corporate lending.

The remainder of this paper is organized as follows. The following sections provide a short description of the corporate lending market in Ukraine. Section 3 surveys the related literature. Sections 4 and 5 describe the bank lending survey data and methodology for this research, respectively. Section 6 presents results, and section 7 provides conclusions.

2. CORPORATE LENDING IN UKRAINE

Corporate lending penetration in Ukraine has been low for many years (see Figure 1). This raises the question of whether the reasons for the slow lending lie with demand or with supply factors; in particular, whether corporations have suppressed the demand for loans, or banks have reduced their willingness to lend. There are several preconditions for the latter, primarily the numerous episodes of crises in Ukraine that decreased the banks' risk appetites and led to the tightening of credit risk assessment approaches.

Prior to 2014, corporate lending was reasonably active, but mainly driven by flawed practices and improper motives. Banks lent extensively to related parties or companies owned by politically influential people, who usually have a low operating income (Pham et al., 2021; Vyshnevskyi and Sohn, 2023). Sometimes, there was no intention to repay the loans. Eventually, when the crisis hit, these loans became NPLs. Moreover, Vyshnevskyi and Sohn (2023) indicate that when NPL shocks occur, then banks may even increase related parties lending.

The russian annexation of Crimea and the war that followed in the Donbas region caused an economic crisis in 2014. Businesses located in occupied territories were

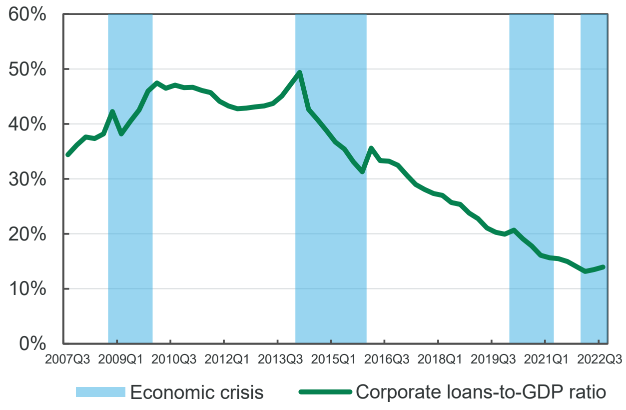


Figure 1. Corporate Loans to GDP Ratio

directly hit. External shocks triggered turmoil and systemic imbalances that had accumulated during the previous years exploded into a financial crisis, thus reinforcing the disruption. Consequently, the share of NPLs in the total portfolio increased significantly from 16.3% in 2014 to 52.2% in 2018.

The 2014–2015 crisis was a turning point. Since then, the National Bank of Ukraine (NBU) has considerably improved its supervision and regulations based on international standards, including prudential requirements for credit risk assessments. Thus, the banks were required to revise their credit standards, primarily for corporate loans, and tighten them significantly to improve loan quality. An unfavorable macroeconomic environment suppressed corporate demand for lending. Therefore, the loans to GDP ratio gradually decreased from its peak of 50% in 2014 to approximately 14% in 2022 (Figure 1).

3. THE RELATED LITERATURE

One challenge related to modelling the loan supply is that many of its drivers, such as internal bank loan policies, are non-observable. Qualitative data from bank lending surveys can help extract information about these unobservable variables (Lown and Morgan, 2006; Bassett et al., 2014).

Ricci et al. (2023) reveal that the level of bank lending standards explains heightened bank lending growth during extended periods of easing of lending standards, and the lower growth seen after they are tightened. This insight offers a potential indicator for macroprudential policy, complementing existing metrics such as the credit-to-GDP gap. Through counterfactual analyses focused on business lending in the Netherlands, the authors illustrate the relevance of their survey-based instrument at the macro-level. The study links strong credit growth and softer lending standards to early-warning signs of financial crises and subsequent economic downturns.

Apergis and Chatziantoniou (2021) employ the Autoregressive Distributed Lag (ARDL) method to investigate the role of lending standards in real business cycles. Their research shows that lending conditions are a major factor in defining business cycles, with robust findings over a range of time periods and countries. The study highlights the growing importance of lending standards in explaining real GDP changes prior to the global financial crisis.

[2] Lending and credit standards are used interchangeably.

Filardo and Siklos (2020) contribute to the literature by investigating how shifting bank lending criteria affect economic activity, with an emphasis on global spillovers. They utilize global VARs and construct a cross-country dataset with senior lending officer surveys from 17 economies, focusing on eurozone members. The results show that lending standards – rather than interest rates – play a crucial role in understanding credit patterns. Pre-crisis credit booms were greatly impacted by easier lending standards. In Europe, this effect was greater than in the United States because of Europe's more bank-dependent financial system. The study also emphasizes how looser lending requirements reinforce the stimulatory effects of quantitative easing on the local and international markets. In general, credit conditions and the efficiency of the monetary transmission mechanism are greatly influenced by lending requirements.

Rodano et al. (2018) investigate the impact of segmentation on lending conditions in the Italian banking sector during boom-and-bust periods. In the boom, substandard and performing firms display a 4% interest rate spread threshold. During a financial crisis, banks tighten lending standards, favoring performing firms with 39% more financing than comparable substandard firms. In later years, differences in lending were reduced, and the interest rate spread increased. The study's threshold analysis shows that segmentation explains a larger part of the observed credit differential during the bust. During the crisis, the interest rate spread is close to zero, indicating adjustments due to restricting substandard firms' credit access. The study also shows a progressively larger negative impact of a downgrade on credit allocations during crises and recovery.

This study contributes to the existing literature by using bank-level lending survey (BLS) data to develop a measure of corporate loan supply, and analyzing its ability to explain corporate credit growth in Ukraine. Some well-established literature analyzes credit growth factors using BLS data, but most researchers use aggregated information (Lown and Morgan, 2006). Usually, BLS data is confidential and not available for public use at a disaggregated level. Previous studies have used qualitative data from surveys to separate the supply and demand factors of lending, for instance, in the Euro area (de Bondt et al., 2010; Ciccarelli et al., 2015; Ciccarelli et al., 2013) and the United States (Bassett et al., 2014). However, only a handful of studies have employed bank-level BLS. Wośko (2015) used panel data from the Senior Loan Officers Opinion Survey to model corporate, mortgage, and consumer loan growth in Poland. Pintaric (2016) used bank-level data to develop a credit growth model for Croatia, and found that demand and credit standards have statistically significant effects on the growth of specific loan types.

Hempell and Kok Sørensen (2010) employed a cross-country panel based on a confidential dataset from the Eurosystem's bank lending survey and found that bank lending activity was generally influenced by the ability and willingness of banks to provide loans, especially during the financial crisis. There is also evidence that supply side constraints have a detrimental effect on loan growth – even after adjusting for demand-side effects. Altavilla et al. (2019) derived a measure of loan supply shocks from proprietary bank-level data on credit criteria from the euro area. Using a Bayesian vector autoregressive model, they found that tighter credit standards, internal bank regulations, and loan approval standards result in a prolonged decline in the amount of credit.

This study also contributes to the literature by exploring the imbalances in the Ukrainian banking system. Banks with liquidity surpluses tend to invest in government securities. The study finds that banks with a high share of government securities are susceptible to crowding-out effects, which result in reduced corporate lending and a potential hindrance to economic growth. The crowding-out effect of lending through government debt has also been discussed extensively in a series of recent studies. For instance, Pinardon-Touati (2022) argued that due to constraints on bank credit supply and segmentation across banks, an increase in local government lending can lead to a reduction in aggregate corporate credit and disproportionately affect firms' borrowing from the same bank, potentially leading to an inefficient allocation of resources and lower overall output. This phenomenon has been widely studied from different perspectives in China (Huang et al, 2020) and Mexico (Morais et al., 2021).

4. BANK LENDING SURVEY DATA DESCRIPTION

The NBU has been conducting a quarterly bank lending survey since 2011. The survey aims to help the central bank and other stakeholders better understand lending market conditions and trends from the banks' perspective. It provides general assessments and forecasts of changes in lending standards and conditions for the corporate sector and households, as well as fluctuations in lending demand.

The main question of interest for the research extracted from BLS is on lending standards: "How did the standards for approval of corporate loan applications change within the past quarter?" Figure 2 illustrates that according to

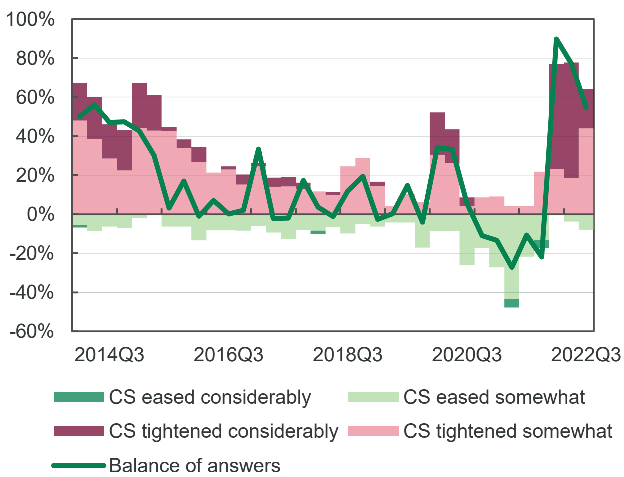


Figure 2. Distribution of BLS Answers for the Question: "How did the standards for approval of corporate loan applications change within the past quarter?"

Note: Background shows the share of answers in total (100%). The balance of answers[3] is weighted by the banks' net assets. A positive balance indicates a tightening of standards for the approval of loan applications.

[3] Balance of answers = 0.5*CS tightened considerably + 0.25*CS tightened somewhat + 0*CS remained unchanged – 0.25* CS eased somewhat – 0.5*CS eased considerably.

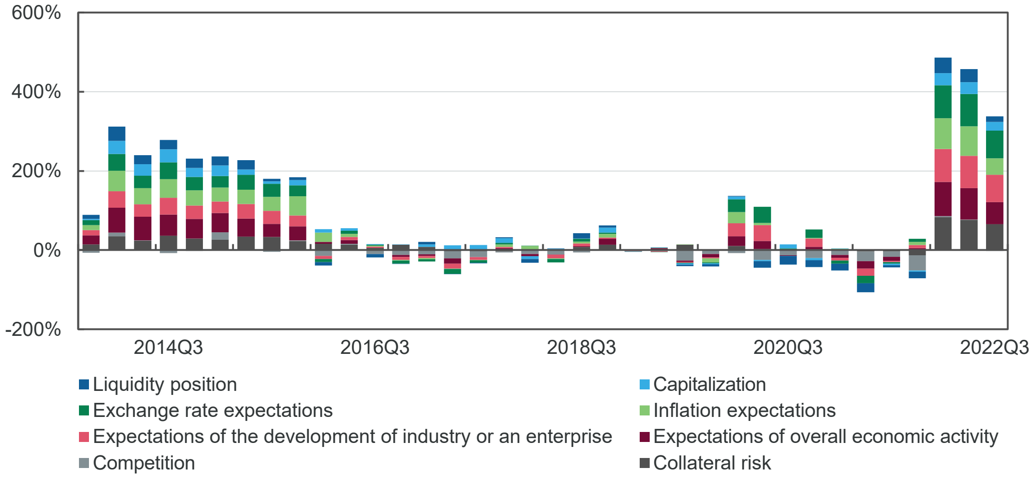


Figure 3. Factors Influencing Banks' Decisions to Change Credit Standards for Corporates According to BLS

the BLS responses, banks tightened their corporate credit standards (CS) in 2014–2015, 2020, and 2022 (all periods of economic crisis).

Economic, exchange rate and inflation expectations pushed banks to offer less favorable corporate lending conditions during crises (Figure 3). In normal times, better liquidity positions and competition encourage banks to loosen their standards. The study identifies proxies to quantitatively assess the factors that explain the decisions of banks to change their credit standards.

Only solvent banks provided BLS answers. Reliable quarterly data are available from Q4 2013 until Q3 2022. During 2015–2016, there was a decrease in the number of banks, and since 2020 the number of surveyed banks has dropped significantly. However, this reduction in respondents did not affect the representativeness of the data: the surveyed institutions have always represented more than 90% of net assets. The panel data are unbalanced, and include 56 banks and 1,249 observations.

5. METHODOLOGY

This study employs a two-step process similar to that described by Wośko (2015). First, categorical data from the BLS are transformed into a continuous CSI, which is a proxy for the supply of corporate loans. Second, we use the CSI to explain the evolution of new corporate lending.

In the first step, BLS answers regarding changes in corporate credit standards are used as a dependent variable. The answers come in five categories: "tightened considerably", "tightened somewhat", "remained unchanged", "eased somewhat" and "eased considerably". Fewer banks indicated that their lending standards eased or tightened considerably, thus, the five categories were combined into three: "eased", "unchanged", and "tightened". This allows for an increase in the number of observations in each remaining category and simplifies the estimation. As these answers are categorically ordered data, a panel ordered logit model, which explains the likelihood of a bank moving from one category to another, was employed.

The dependent variable takes values {1,0,-1} which represents the answers "tightened", "unchanged", and "eased" respectively.

The model for the first step is as follows:

$$z_{i,t} = \Sigma_q B^q X^q_{i,t}, \tag{1}$$

where $z_{i,t} = \log\left(\frac{P_{i,t}}{1 - P_{i,t}}\right)$ is a logit transformation of the probability that bank i during quarter t decides to tighten its corporate standards, $X^q_{i,t}$ is the qth control variable, and B^q is the respective coefficient. The following set of controls was used: regulatory capital adequacy ratio,[4] short-term liquidity ratio (ratio of assets to liabilities with the maturity of less than one year), real GDP growth, exchange rate change (positive values mean depreciation), interbank loan interest rates, and a dummy indicating whether BLS competition has led to tighter or looser credit standards. The fitted values from Model (1) are transformed into a CSI.

The fitted values from the ordered logit model are not limited and can take any real number. Higher fit values indicate an increased probability of tightening credit standards. The Model also estimates the cut-off points, allowing for the classification of the fitted values into categories. As there are three categories, the model produced two cutting points. For easier interpretation, the fitted values are rescaled to range from 0 to 100 using min-max normalization. These rescaled fitted values are used further as the CSI.

In the second step, the CSI is used as a measure of the supply side of corporate lending while controlling for macro variables and bank characteristics. An initial baseline model is then augmented with a series of interactions between the variables. All interactions are demeaned so that the main effects can be interpreted at the mean of the interacted variable.

The dependent variable in the second step represents corporate lending. In Ukraine, gross loans cannot be used because the share of NPLs is high owing to previous crises, and gross loan stock is significantly driven by NPL workouts. Net loans are a better proxy but depend on provisions that vary based on macro conditions. Hence, the volumes of new corporate loans were selected for all models. Separate models for national and foreign currency loans were estimated. To control for inflation and devaluation, the volumes of corporate loans provided during the quarter in national and foreign currencies were taken and then

[4] Descriptions of all the variables are provided in Table 5 (Appendix A).

adjusted to the cumulative change in inflation since 2007 and the exchange rate since 2014, when it became floating.[5]

The baseline model for the second step is the following panel fixed effects regression:

$$\log(\text{loans}_{i,t}) = \beta_0 + \beta_1 \text{CSI}_{i,t-2} + \text{ control variables} + \text{FE} + \epsilon_{i,t}, \quad (2)$$

where $\text{loans}_{t,i}$ are adjusted volumes of new corporate loans in bank *i* in period *t*. The control variables are the short-term liquidity ratio, real GDP growth, new deposit interest rates, new corporate loan interest rates, total deposit growth, the share of NPLs in the loan portfolio, and bank fixed effects (FE). The variable $\text{CSI}_{i,t-2}$ is the normalized values from the first-step model. An exploratory analysis suggests that the effect starts to be significant from the second lag.

Usually, smaller banks tend to be more flexible than larger banks, which allows them to have looser credit standards and to approve loan applications more quickly. Therefore, it was assumed that the effect of a change in credit standards could vary depending on bank size. *The first augmented* model includes the interaction of the CSI with bank size.

$$\log(\text{loans}_{i,t}) = \beta_0 + \beta_1 \text{CSI}_{i,t-2} + \beta_2 \text{size}_{i,t} + \beta_3 \text{CSI}_{i,t-2} \times \text{size}_{i,t} + \text{ control variables} + \text{FE} + \epsilon_{i,t}, \quad (3)$$

where $\text{size}_{i,t}$ is the share of net assets of bank *i* in total net assets during period *t*.

Following the crisis in 2014–2015, corporate lending penetration was low, resulting in increased bank liquidity. In Ukraine, banks invest excess liquidity in government bonds and deposit certificates because of their low credit risk and attractive interest yields. Additionally, frequent crisis episodes have increased the government's demand for supplementary financial resources, prompting the banks to build up government security portfolios. An adverse macro environment creates preconditions for the crowding-out effect; therefore, it is tested whether and how this effect influences corporate lending during normal and bad times. Consequently, in the *second augmented* model, the effect of real GDP growth interaction on the share of government securities is explored, controlling for periods of positive and negative real GDP growth:

$$\log(\text{loans}_{i,t}) = \beta_0 + \beta_1 \text{CSI}_{i,t-2} + \beta_4 \text{share\_gov}_{i,t} + \beta_5 d_t \times GR_t \times \text{share\_gov}_{i,t} + \text{ control variables} + \text{FE} + \epsilon_{i,t}, \quad (4)$$

where $\text{share\_gov}_{i,t}$ is the share of government bonds and deposit certificates in the total assets, GR_t is real GDP growth, and $d_{i,t}$ is a dummy variable controlling for the periods of positive and negative real GDP growth (1 if real GDP growth > 0, and 0 otherwise).

6. RESULTS

6.1. First Step

The results of the first step indicate that all the control variables, except for the capital adequacy ratio, are significant (Table 1). Faster economic growth and higher liquidity lead to the easing of credit standards, whereas elevated interbank loans interest rates and exchange rate depreciation stimulate tightening. According to the odds ratios, each additional percentage point in the interbank loan interest rate increases the probability of moving from easing credit standards to remaining unchanged, or from remaining unchanged to tightening, by 4.3%. An exchange rate depreciation of 1% increases the probability of such a move by 2.9%. In contrast, an increase of 1% in the short-term liquidity ratio increases the probability of banks loosening credit standards by 0.7%. If real GDP increases by 1%, the probability increases by 3.2%. Additionally, bank competition leads to looser credit standards. If the bank indicates in the BLS that bank competition eases credit standards, then there is a 93.3% probability that it will be in a category that loosens standards.[6]

To analyze the change in credit standards for the system the fitted values from the ordered logit model in the first step (Table 1, column 1) were aggregated. Aggregation (Figure 4) is conducted by averaging the fitted values and weighting them by each bank's net assets. The weighted average has a good ability to replicate aggregate BLS answers, but now it has clear drivers. The aggregate indicator signals that banks in Ukraine tightened their lending standards during episodes of economic crisis in 2014–2015 and in 2022. The model suggests that the banks generally did not ease their lending standards during most periods. Overall, the aggregated fitted values provide insight into the trends and patterns of lending standards in Ukraine over different periods, thus shedding light on the adjustments made by banks in response to economic conditions and external shocks.

Table 2 reveals that the accuracy of the model is 63.3%. The model has a poor ability to categorize banks that have eased or tightened their lending standards. It is assumed that this problem may be due to the uneven distribution of the BLS answers between categories. However, even if the model cannot clearly distinguish the change of the credit standards, it appears using survey responses as dummies imposes certain limitations. For instance, respondents signal only direction of credit standards change, but there is no scale. Therefore, the model is still useful since it can quantify the supply of corporate loans.

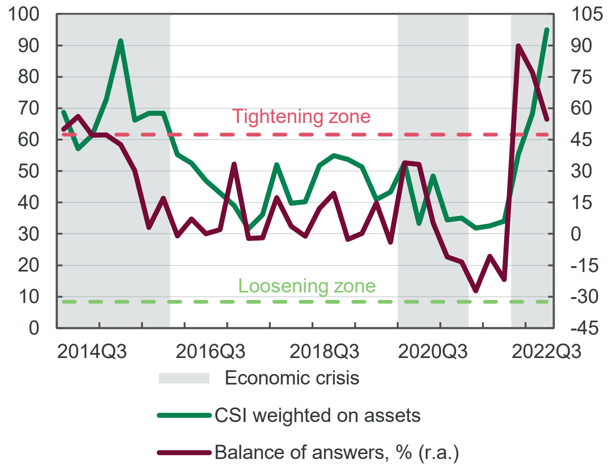


Figure 4. CSI Weighted by the Banks' Assets and Balance of Answers Regarding the Question about Corporate Lending Standards

5 The exchange rate was fixed before 2014 in Ukraine.

6 The following cutting points were also obtained: k1 = -0.4 and k2 = 3.6. Assume that p is fitted values. If p < -0.4, then the bank eased its corporate lending standards, if -0.4 < p < 3.6, then it left standards unchanged, and if p > 3.6, then the bank tightened its standards.

Using the fitted values from the model and estimated cut-off points, the decision to change credit standards is determined for each bank and every quarter. The results are compared with actual BLS answers (Figure 5). The estimated answers follow the main trends of the actual BLS answers.

Table 1. Results of the Ordered Logit Model in the First Stage

Variables	Ordered logit	Odds ratio	Pooled OLS
	(1)	(2)	(3)
Interbank loan interest rates	0.042* (0.017)	1.043* (0.017)	0.012** (0.004)
Capital adequacy ratio$_{t-1}$	-0.003 (0.003)	0.997 (0.003)	-0.000 (0.001)
Short term liquidity ratio$_{t-1}$	-0.007** (0.003)	0.993** (0.003)	-0.001* (0.001)
Real GDP growth$_{t-1}$	-0.032*** (0.010)	0.968*** (0.009)	-0.006** (0.002)
Exchange rate growth	0.029*** (0.006)	1.029*** (0.007)	0.007*** (0.002)
Dummy competition led to CS tightening	0.593 (0.349)	1.809 (0.631)	0.128 (0.082)
Dummy competition led to CS easing	-2.701*** (0.226)	0.067*** (0.015)	-0.607*** (0.045)
Constant			-0.006** (0.002)
Cutting point$_1$	-0.399 (0.802)	-0.399 (0.802)	
Cutting point$_2$	3.587*** (0.796)	3.587*** (0.796)	
Sigma	0.271* (0.119)	0.271* (0.119)	
Observations	1,174	1,174	1,174

Note: standard errors in parentheses; clustered on time.
*p < 0.05; ** p < 0.01; *** p < 0.001.

Table 2. Accuracy of the 1st Stage Model

	Eased	Unchanged	Tightened	Total
	(1)	(2)	(3)	(4)
BLS answers	127	754	368	1,249
BLS answers, % of total	10.1%	60.4%	29.5%	
Accuracy rate, % of right answers predicted by the model	11.8%	89.3%	27.7%	63.3%

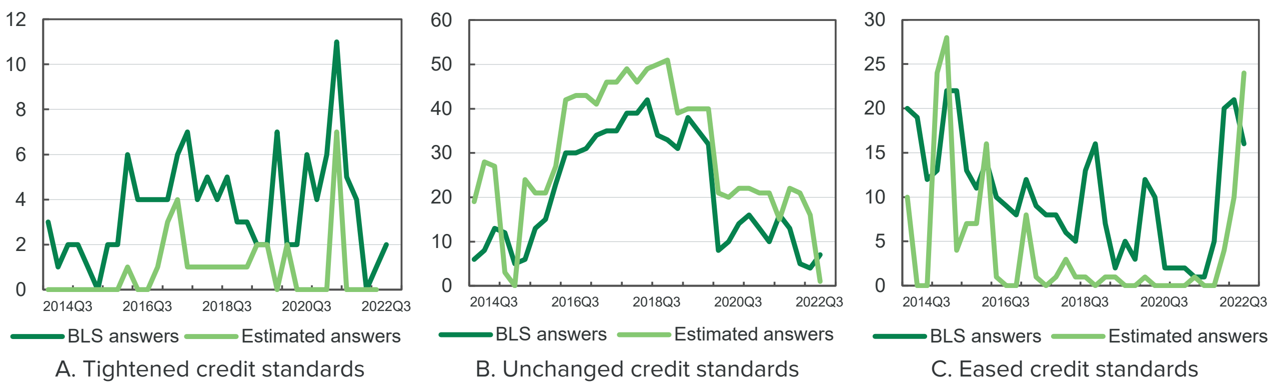


Figure 5. Number of Banks that Make Respective Decisions on Credit Standards According to the Estimated Model Results and Actual BLS Answers

6.2. Second Step

Table 3 presents the results of the analysis based on corporate lending in national currency (columns 1–4) and foreign currency (columns 5–8). Columns (1) and (5) present the baseline models without the CSI. Columns 2 and 6 use the answers from the BLS as dummy variables instead of the CSI. Two dummies are used – one that takes the value of 1 for banks that eased their credit standards, and the other that takes the value of 1 for banks that tightened their credit standards. Columns 3 and 7 present the CSI obtained in the first step. Finally, columns 4 and 8 contain both the CSI and the residuals obtained in step 1 (the residuals are orthogonal to the index). The residuals are computed from the OLS model in column 3 of Table 1.

The modeling results highlight several key relationships. Real GDP growth positively correlates with new corporate lending, thereby suggesting that higher GDP growth is associated with increased lending in both national and foreign currencies. For example, a 1% increase in real GDP is associated with a 1% increase in new corporate lending in the national currency and a 3% increase in foreign currency. Conversely, higher NPL levels negatively affect new corporate lending. For instance, a 1% increase in the share of NPLs is linked to an approximately 0.2% decrease in national currency lending and a 0.3% decrease in foreign currency lending.

The effect of the CSI is significant only for national currency loans. Specifically, an additional unit increase in the CSI decreases the volume of new corporate loans in the national currency by 0.7%, with the decrease starting to be material from the second quarter. Using dummy variables from the BLS, it is found that when banks indicate a decision to tighten their credit standards in the BLS, it leads to a 28.1% decrease in new corporate loans in foreign currency. However, when banks decide to ease their credit standards, new corporate loans in the national currency increase by 16.2%. Since the dummies are limited to two numbers and do not have magnitude, these effects have very wide confidence intervals of 95% and cannot be used in practice. To check for endogeneity, the residuals from the first-step OLS model (Table 1, column 3) were included in the CSI model. These residuals are orthogonal to the CSI. Thus, the insignificant coefficient of the residuals (Table 2, columns 4 and 8) indicates that only the credit standard component mediated by the variables included in the first step is significant for new corporate lending.

Table 4 presents the results of the augmented models for new corporate lending in national (columns 1 and 2) and foreign currencies (columns 3 and 4). In columns 1 and 3, the model includes an interaction term between the CSI and bank size. Columns 2 and 4 show the models with the interaction between real GDP growth and the share of government securities.

The results in Table 4 corroborate those in Table 3 and indicate that all of the interaction terms included in the benchmark models are significant. Additionally, given

Table 3. Results of Baseline Models

	National currency				Foreign currency			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔDeposit interest rates	-0.017 (0.014)	-0.016 (0.014)	-0.010 (0.013)	-0.010 (0.013)	-0.035 (0.023)	-0.032 (0.025)	-0.032 (0.023)	-0.032 (0.023)
ΔCorresponding currency loan interest rates	-0.021 (0.011)	-0.023 (0.013)	-0.021 (0.011)	-0.021 (0.011)	-0.107* (0.045)	-0.112* (0.048)	-0.066 (0.038)	-0.066 (0.038)
Log(NPL)	-0.210*** (0.032)	-0.220*** (0.032)	-0.240*** (0.030)	-0.240*** (0.030)	-0.290*** (0.050)	-0.290*** (0.051)	-0.320*** (0.058)	-0.320*** (0.059)
Short term liquidity ratio	0.002 (0.002)	0.001 (0.002)	0.001 (0.002)	0.001 (0.002)	0.004 (0.002)	0.003 (0.002)	0.003 (0.002)	0.003 (0.002)
Real GDP growth	0.020*** (0.003)	0.010*** (0.004)	0.010*** (0.004)	0.010*** (0.004)	0.030*** (0.007)	0.020** (0.007)	0.030*** (0.008)	0.030*** (0.009)
Deposits growth	-0.000 (0.001)	-0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)
CSI_{t-2}			-0.007*** (0.002)	-0.007*** (0.002)			-0.007 (0.004)	-0.007 (0.004)
BLS dummy indicating CS $tightening_{t-1}$		-0.082 (0.077)				-0.281* (0.114)		
BLS dummy indicating CS $easening_{t-1}$		0.162** (0.061)				0.048 (0.108)		
OLS residuals from 1st $step_{t-2}$				0.004 (0.054)				0.000 (0.111)
Constant	-1.120*** (0.083)	-2.310*** (0.536)	-0.790*** (0.112)	-0.790*** (0.111)	-2.900*** (0.140)	-3.970*** (0.497)	-2.540*** (0.252)	-2.540*** (0.250)
Bank fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	956	956	905	905	927	927	878	878

Note: Standard errors in parentheses; clustered on time.
*p < 0.05; ** p < 0.01; *** p < 0.001.

Table 4. Results of Benchmark Models

	National currency		Foreign currency	
	(1)	(2)	(3)	(4)
ΔDeposit interest rates	-0.005 (0.013)	-0.008 (0.012)	-0.028 (0.023)	-0.027 (0.022)
ΔCorresponding currency loan interest rates	-0.021 (0.011)	-0.021 (0.011)	-0.066 (0.038)	-0.068 (0.037)
Log(NPL)	-0.242*** (0.029)	-0.249*** (0.033)	-0.308*** (0.056)	-0.331*** (0.059)
Short term liquidity ratio	0.001 (0.002)	0.001 (0.002)	0.003 (0.002)	0.003 (0.002)
Real GDP growth	0.016*** (0.004)	0.014** (0.004)	0.032*** (0.008)	0.030*** (0.008)
Deposits growth	0.000 (0.001)	0.001 (0.001)	0.002 (0.002)	0.002 (0.002)
CSI_{t-2}	-0.008*** (0.002)	-0.007*** (0.002)	-0.008* (0.004)	-0.007 (0.004)
Size of the bank	-0.097*** (0.017)		-0.085* (0.033)	
CSI_{t-2} × size of the bank	0.001*** (0.000)		0.001* (0.000)	
Share of gov. securities		-0.005 (0.007)		-0.003 (0.008)
Real GDP growth<0 × share of gov. securities		-0.001*** (0.000)		-0.001** (0.000)
Real GDP growth>0 × share of gov. securities		-0.000 (0.001)		0.002 (0.002)
Constant	-0.675*** (0.106)	-0.789*** (0.114)	-2.421*** (0.242)	-2.557*** (0.260)
Individual fixed effects	Yes	Yes	Yes	Yes
Observations	905	905	878	878

Note: Standard errors in parentheses; clustered on time.
*p < 0.05; ** p < 0.01; *** p < 0.001.

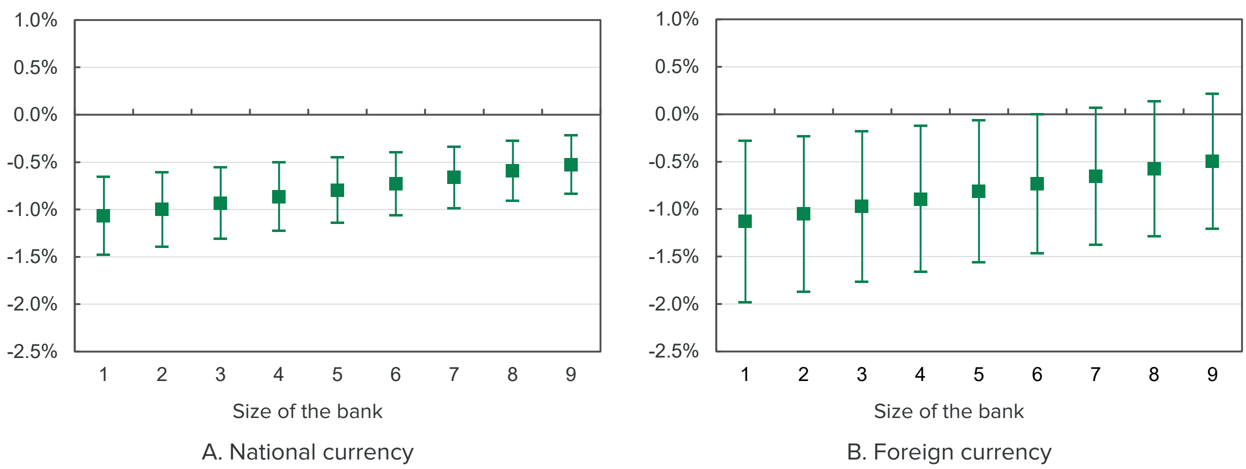
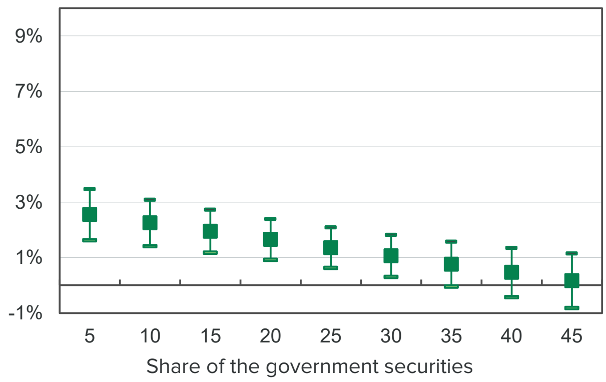
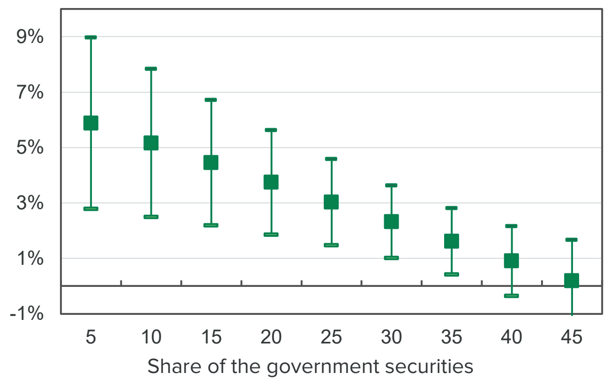


Figure 6. Marginal Effects of Tightening Credit Standards on New Corporate Lending, Depending on the Size of the Bank (Measured as a Bank's Share of Total Net Assets, %)

Note: whiskers indicate a 95% confidence interval.



A. National currency



B. Foreign currency

Figure 7. Marginal Effects of Real GDP Growth on New Corporate Lending Depending on the Share of Government Securities (Government Bonds and Deposit Certificates) in Total Assets

Note: whiskers indicate a 95% confidence interval.

the significant negative dependence between credit standard tightening and new corporate lending, I find that bank size matters, with the effect of a credit standards change being stronger for small banks. For small banks, the CSI has a negative impact on corporate loans in both national (Figure 6A) and foreign currencies (Figure 6B). An additional 1% increase in bank size enhances the CSI effect by 0.08% for foreign currency loans and by 0.07% for national currency loans.

From Table 4, we can conclude that the impact of real GDP growth interaction with the share of government securities is significant only during periods of negative real GDP growth. The positive correlation between GDP growth and new corporate lending is weaker for banks with a higher share of government securities in their assets (see Figure 7). During an economic decline, interest rates for risk-free assets increase. Therefore, having a high share of government securities in the portfolio provides banks with increased interest income and protects their ability to lend to corporations.

7. CONCLUSION

This study examined the determinants of Ukrainian banks' new corporate lending practices. The use of unbalanced panel data from the 4th quarter of 2013 to the 3rd quarter of 2022 shows that positive real GDP growth, bank competition, and higher liquidity lead to looser credit standards, whereas higher interest rates and exchange rate depreciation cause standards to tighten. Tightening credit standards decreases national currency corporate lending in half a year, and smaller banks experience a stronger effect in comparison with larger banks. A higher share of NPLs reduces loans in both national and foreign currencies. Real GDP growth positively correlates with new corporate loans in both national and foreign currencies. The effect of negative economic activity on loans in both national and foreign currencies is weaker for banks with a higher share of government securities.

Usually supply factors in corporate lending are latent and unobservable. The study helps to quantify the supply for business loans. Moreover, this paper explores the factors determining corporate lending development in Ukraine.

The study still has the potential to reveal more results through using other methodologies. Papers on credit growth determinants also implement time series models by applying aggregated data. However, the availability of data at the bank-level allows the use, for instance, of a local projection method, which produces comparable results.

REFERENCES

Altavilla, C., Darracq Pariès, M., Nicoletti, G. (2019). Loan supply, credit markets and the euro area financial crisis. Journal of Banking & Finance, 109, 105658. https://doi.org/10.1016/j.jbankfin.2019.105658

Apergis, N., Chatziantoniou, I. (2021). Credit supply conditions and business cycles: New evidence from bank lending survey data. Research in International Business and Finance, 55, 101332. https://doi.org/10.1016/j.ribaf.2020.101332

Bassett, W. F., Chosak, M. B., Driscoll, J. C., Zakrajšek, E. (2014). Changes in bank lending standards and the macroeconomy. Journal of Monetary Economics, 62, 23–40. https://doi.org/10.1016/j.jmoneco.2013.12.005

Ciccarelli, M., Maddaloni, A., Peydró, J.-L. (2013). Heterogeneous transmission mechanism: Monetary policy and financial fragility in the euro area. ECB Working Paper Series, 1527. Frankfurt am Main: European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1527.pdf

Ciccarelli, M., Maddaloni, A., Peydró, J.-L. (2015). Trusting the bankers: A new look at the credit channel of monetary policy. Review of Economic Dynamics,18(4), 979–1002. https://doi.org/10.1016/j.red.2014.11.002

de Bondt, G., Maddaloni, A., Peydró, J.-L., Scopel, S. (2010). The euro area bank lending survey matters: Empirical evidence for credit and output growth. ECB Working

Paper Series, 1130. Frankfurt am Main: European Central Bank. Retrieved from https://www.ecb.europa.eu//pub/pdf/scpwps/ecbwp1160.pdf

Filardo, A. J., Siklos, P. L. (2020). The cross-border credit channel and lending standards surveys. Journal of International Financial Markets, Institutions and Money, 67, 101206. https://doi.org/10.1016/j.intfin.2020.101206

Hempell, H. S., Kok Sørensen, C. (2010). The impact of supply constraints on bank lending in the euro area – crisis induced crunching? ECB Working Paper Series, 1262. Frankfurt am Main: European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1262.pdf

Huang, Y., Pagano, M., Panizza, U. (2020). Local crowding out in China. The Journal of Finance, 75(6), 2855–2898. https://doi.org/10.1111/jofi.12966

Lown, C. S., Morgan, D. P. (2006). The credit cycle and the business cycle: New findings using the loan officer opinion survey. Journal of Money, Credit and Banking, 38(6), 1575–1597. https://doi.org/10.1353/mcb.2006.0086

Morais, B., Perez-Estrada, J., Peydró, J.-L., Ruiz-Ortega, C. (2021). Expansionary austerity: Reallocating credit amid fiscal consolidation. CEPR Discussion Papers, 16511. Paris & London: Centre for Economic Policy Research. Retrieved from https://cepr.org/publications/dp16511

NBU (2022). Financial Stability Report, June 2022. Kyiv: National Bank of Ukraine. Retrieved from https://bank.gov.ua/admin_uploads/article/FSR_2022-H1_eng.pdf

Pham, T., Talavera, O., Tsapin, O. (2021). Shock contagion, asset quality and lending behaviour: The case of war in Eastern Ukraine. Kyklos, 74(2), 243–269. https://doi.org/10.1111/kykl.12261

Pinardon-Touati, N. (2022). The Crowding Out Effect of Local Government Debt: Micro- and Macro-Estimates. Retrieved from https://npinardontouati.github.io/files/JMP_PinardonTouati.pdf

Pintaric, M. (2016). What is the effect of credit standards and credit demand on loan growth? Evidence from the Croatian Bank Lending Survey. Comparative Economic Studies, 58(3), 335–358. https://doi.org/10.1057/s41294-016-0004-2

Ricci, L., Soggia, G., Trimarchi, L. (2023). The impact of bank lending standards on credit to firms'. Journal of Banking & Finance, 152, 106880. https://doi.org/10.1016/j.jbankfin.2023.106880

Rodano, G., Serrano-Velarde, N., Tarantino, E. (2018). Lending standards over the credit cycle'. The Review of Financial Studies, 31(8), 2943–2982. https://doi.org/10.1093/rfs/hhy023

Vyshnevskyi, I., Sohn, W. (2023). Nonperforming loans and related lending: Evidence from Ukraine. Emerging Markets Review, 57, 101069. https://doi.org/10.1016/j.ememar.2023.101069

Wośko, Z. (2015). Modelling credit growth in commercial banks with the use of data from Senior Loan Officers Opinion Survey. Warsaw: NBP Working Paper, 210. Narodowy Bank Polski. Retrieved from https://static.nbp.pl/publikacje/materialy-i-studia/210_en.pdf

APPENDIX A. TABLES

Table 5. Summary Statistics

Variable	Description	Data structure	Obs	Mean	Std.	Min	Max
Capital adequacy	Capital adequacy ratio, %	Bank-level	1,249	28.2	31.1	1.3	416.1
Liquidity	Short-term liquidity ratio, %	Bank-level	1,249	100.6	36.1	46.1	358.9
Inflation	CPI change, y-o-y, %	Macro	1,249	14.6	12.4	0.5	57.7
Exchange rate	Average exchange rate, UAH/USD	Macro	1,249	24.9	5.2	8.0	36.6
Economic activity	Real GDP growth, y-o-y, %	Macro	1,249	-1.9	10.5	-46.5	7.8
Interbank interest rates	Average quarterly interest rates on new interbank loans, %	Macro	1,249	13.7	4.4	5.4	23.3
Real corporate loans in foreign currency	Adjusted on exchange new corporate loans in foreign currency, bn. UAH	Bank-level	1,249	0.8	2.5	0.0	26.3
Real corporate loans in national currency	Adjusted on inflation new corporate loans in national currency, bn. UAH	Bank-level	1,249	2.2	4.8	0.0	38.6
Deposit interest rates	Quarterly averaged new deposits interest rates, %	Bank-level	1,249	9.9	3.6	0.0	22.0
National currency loan interest rates	Quarterly averaged new national currency loans interest rates, %	Bank-level	1,249	19.3	4.5	5.4	48.0
Foreign currency loan interest rates	Quarterly averaged new foreign currency loans interest rates, %	Bank-level	1,249	10.0	5.5	1.1	48.0
NPL level	Share of the non-performing loans in total portfolio, %	Bank-level	1,249	26.4	38.7	0.0	862.1
Deposits	Total deposits growth, y-o-y, %	Bank-level	1,249	22.1	45.9	-78.1	660.5
Share of government securities	Share of government bonds and deposit certificates in total assets, %	Bank-level	1,249	16.2	13.6	0.0	76.6
Size of bank	Share of the net assets in total, %	Bank-level	1,249	2.8	5.0	0.0	27.3
Dummy competition led to CS easing	1 if the bank indicated in the BLS that competition led to CS easing	Bank-level	1,249	0.1	0.2	0.0	1.0
Dummy competition led to CS tightening	1 if the bank indicated in the BLS that competition led to CS tightening	Bank-level	1,249	0.2	0.4	0.0	1.0

Note: The NBU ended the transition from a short-term liquidity ratio to more complex indicators (net stable funding ratio and liquidity coverage ratio (NSFR)) in 2022, and stopped calculating the short-term liquidity ratio. Therefore, the short-term liquidity ratio is approximated to the change in the NSFR during 2022.

Visnyk of the National Bank of Ukraine, 2024, No. 254, pp. 15–28

SHORT-TERM FORECASTING OF GLOBAL ENERGY AND METAL PRICES: VAR AND VECM APPROACHES

DIANA BALIOZ[a]

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

Abstract This study introduces a set of multivariate models with the aim of forecasting global prices of 1) crude oil, 2) natural gas, 3) iron ore, and 4) steel. Various versions of vector autoregression and error-correction models are applied to monthly data for the short-term prediction of nominal commodity prices six months ahead. The fundamentals for metal and energy price predictions include inter alia, stock changes, changes in commodity production volumes, export volumes by the largest players, changes in the manufacturing sector of the largest consumers, the state of global real economic activity, freight rates, and a recession indicator. Kilian's (2009) index of global real economic activity is found to be a useful proxy for global demand and a reliable input in forecasting both energy and metal prices. The findings suggest that models with smaller lag orders tend to outperform those with a higher number of lags. At the same time, selected individual models, while showing a standalone high performance, have varying forecast precision during different periods, and no individual model outperforms others consistently throughout the forecast horizon. Note that the price projections obtained from the models could be used further for the longer-term forecasting of commodity prices. Our short-term hands-on framework could be a useful forecasting tool for central bank policymakers and researchers.

JEL Codes C32, C53, Q02

Keywords forecasting, commodities, forecast evaluation, VAR models, VECMs

1. INTRODUCTION

Commodity prices play an increasingly important role in influencing global inflation and the macroeconomic environment. For many developing economies, primary commodities remain the main drivers of the balance of payments, while price fluctuations affect their macroeconomic performance. Energy transition, the COVID-19 pandemic, and russia's war against Ukraine have led to sharp price changes, highlighting the high volatility of the commodity markets and the vulnerability of commodity-dependent countries to price shocks (Baffes and Nagle (eds.), 2022). Therefore, a deeper understanding of commodity price movements and the factors behind them are crucial to policymakers, international institutions, and think tanks.

The approaches used to forecast the prices of energy and metals differ in many ways depending on the purpose of studies, the benchmarks chosen, the frequency of data, and forecasting techniques. There are papers that employ univariate techniques (Tularam and Saeed, 2016; Nademi and Nademi, 2018; Hosseinipoor et al., 2016), multivariate econometric models (Nick and Thoenes, 2014; Berrisch and Ziel, 2022), and machine learning approaches (Kriechbaumer et al., 2014; Li et al., 2020). Various research studies focus either on short- or long-term forecasting tools: they aim to predict the spot (nominal or real) prices or futures prices of commodities, and find the interrelation between commodity prices and their potential impact on one another (West and Wong, 2014).

This paper introduces the hands-on approach of multivariate models for the short-term forecasting of global prices for crude oil, natural gas, steel, and iron, and analyzes the forecasting performance of these techniques. More specifically, this study focuses on predicting the spot nominal monthly prices of commodities six months ahead, while the majority of papers develop models to predict either futures prices (Bowman and Husain, 2004; Reichsfeld and Roache, 2011; Ambya et al., 2020), spot real quarterly prices (Baumeister and Kilian, 2013, 2014; Wårell, 2018) or price indices (Chou et al., 2012). Our short-term hands-on framework could be a useful tool for central banks and analysts, while the price projections obtained from these models could be used further as assumptions for the longer-term forecasting of commodity prices.

The paper is organized as follows. Section 2 provides a review of relevant research literature on commodity price forecasting, and examines the modern econometric approaches used to predict oil, gas, iron ore and steel prices. Section 3 describes the general methodology of VAR/VECM models, and is divided into subsections to analyze in detail the specifications of the models and data for each of the four commodities discussed. In Section 4, we look at the results of our short-term models and assess their forecasting properties. Finally, Section 5 offers conclusions and recommendations on how this forecasting approach can be improved.

2. LITERATURE REVIEW

Forecasting commodity prices is generally considered a challenging task, and rightfully so, given their volatility, dependence on many economic and financial factors, trend changes over time, and the huge variety of methods and approaches used in forecasting. The literature on commodity price forecasting differs significantly, depending on the purposes of the studies, the techniques used and the features of each commodity market. There are papers that employ econometric approaches, namely univariate and multivariate forecasting models, and those that use machine learning and non-parametric techniques. Different studies also focus on short- or long-term forecasting tools; they aim to predict spot or futures prices of commodities, nominal or real (spot) prices; and seek to find the interrelation between different commodity prices and their potential impact on one another. In this section, we review the literature by the commodities of interest.

The literature on predicting global crude oil prices is probably the most extensive, compared to other commodity groups, due to the impact oil prices have on inflation and macroeconomic development. For example, Tularam and Saeed (2016) focus on univariate time-series models to predict oil price movements, and find that the ARIMA model is a better fit for daily WTI oil prices than the exponential smoothing and Holt-Winters models. Conversely, Nademi and Nademi (2018) find that the semiparametric Markov-switching AR-ARCH model outperforms other simple approaches, including ARIMA and GARCH, when it comes to forecasting OPEC, WTI and Brent oil prices. However, univariate forecasting models rely only on one input – in this case, the price of crude oil itself and its past patterns – and do not take into account other factors that might impact the price. Whereas multivariate models are more sophisticated and include the economic determinants of price movements.

It is worth noting that a number of research papers use oil *futures* prices to predict movements in *spot* prices. The intuition behind this is that oil is both a physical commodity and a financial asset. Thus, it is often argued that there is a theoretical link between its spot and futures prices, and the slope of oil futures prices may help predict the movements in spot prices. The empirical evidence, however, is mixed. Chernenko et al. (2004), for instance, argue that oil futures prices (just as natural gas futures prices) show little evidence for risk premiums and can be used to forecast spot prices. Some central banks tend to use futures curves for the short-term forecasting of oil spot prices as they are simple and easy to communicate. Reichsfeld and Roache (2011) prove empirically that futures-based forecasts outperform random walk models over a three-month horizon, but not over longer forecast periods. In contrast, Alquist et al. (2011) conclude that futures prices are not good predictors of nominal oil prices and do not outperform no-change forecasts. There is also an arbitrage relationship between oil futures and spot prices, which, inter alia, means that the slope of oil futures prices is rather flat relative to the changes in oil spot prices (Nixon and Smith, 2012). Moreover, due to oil being a physical and storable good with limited inventories, its futures price curve is downward sloping most of the time, except for some occasions of contango (i.e. an upward sloping curve) when there is ample supply and a high level of oil stocks. In general, futures-based models alone do not prove accurate in predicting spot oil prices, Therefore, other approaches or even combinations of different models should be used (ECB, 2015).

A growing number of recent research papers focus on vector autoregression (VAR) models to predict nominal and real *spot* prices on the commodity markets as these models take into account the economic determinants of price movements and market fundamentals. Such models consider each variable as a function of its own past values and past values of other variables in the model. They also provide estimates of the impact of supply and demand shocks on commodity prices, which makes such models a useful analytical tool. VAR and structural VAR models have smaller forecast errors and prove to be more accurate in forecasting oil price movements than other time-series techniques, especially in the short run (as discussed in Baumeister and Kilian (2013, 2014), Kilian and Murphy (2014) etc.). For example, Baumeister and Kilian (2014), in their seminal work, study real-time forecasting techniques, including forecast combinations, to predict the quarterly real price of oil over short-term horizons. The authors use market fundamental variables, such as a change in crude oil production, Kilian's (2009) index of global real economic activity, a change in oil inventories and so on to conclude that VAR models based on monthly data are the most accurate tools for predicting real oil prices on a quarterly basis. At the same time, one may argue that the accuracy and stability of individual forecasts are time varying, and different models might be suitable for different periods. Thus, the combination of individual forecasts should improve the accuracy of forecasts and help overcome the potential misspecifications of individual models. Baumeister and Kilian (2014) developed a number of forecasting models to test an equal-weighted combination of a monthly VAR model (as the best-performing one among the individual approaches) and the futures-based approach, which provides some MSPE improvement of the forecast of the U.S. real refiners' acquisition cost (RAC) and WTI price, while deteriorating directional accuracy. However, this combination method does not improve forecast accuracy for Brent oil real prices at all, thus the results are mixed. In their later study, Baumeister and Kilian (2013) demonstrate that the combination of four models (namely a VAR model, a model based on non-oil commodity prices, a method based on futures spreads, and a time-varying product spread approach) with inverse MSPE weights actually provides better forecast accuracy. The results hold for the U.S. refiners' acquisition cost for crude oil imports and WTI oil prices over January 1992 through September 2012, but there are no results for Brent oil prices, due to the lack of suitable data.

In contrast, Manescu and Van Robays (2014) focus on the current international benchmark price and prove empirically that for Brent oil prices for the period from Q1 1995 through Q4 2012, a four-model combination, which consists of futures, risk-adjusted futures approaches, Bayesian VAR, and a DGSE model, is the best forecasting technique. This equal-weighted model combination produces robust forecasts of oil prices over the studied period, reduces the forecast bias, and outperforms simple models in out-of-sample exercises. At the same time, this combination approach is found to improve the forecasts of the futures-based model and the random walk model (on average, up to 11 quarters ahead), but there is no evidence that it outperforms other forecasting approaches. When compared to benchmarks other than predictions by futures prices, it may produce worse results. Moreover, given the latest patterns on the global oil market, it is unclear if this particular combination of models can be robust over a more recent period than that discussed in a paper by Manescu and Van Robays (2014), i.e. after 2012.

The research papers mentioned above describe the methods of forecasting *real* oil prices or *real* RAC (refiner acquisition cost) based on global supply and demand variables, according to economic theory. These real price forecasts could then be used by analysts and policy makers. Nominal price forecasts, which are usually of interest, could be derived from them, using separate forecasts of the CPI. Thus, the models proposed in the aforementioned papers cannot directly predict the *nominal* price of oil and require other models or external forecasts for that. Meanwhile, Beckers and Beidas-Strom (2015) introduce CPI inflation into the VAR model to fill in this gap in the literature, and they find that this VAR model outperforms the random walk and futures-based forecasts. The authors also conclude that there is merit in combining forecasts of futures and VAR models, although only for horizons beyond 18 months.

Just as in the case of forecasting oil prices, the literature on predicting natural gas prices differs in terms of the purpose of forecasts, chosen benchmarks or markets, the use of additional price determinants, and forecasting methodology. For example, Hosseinipoor et al. (2016) apply the ARIMA/GARCH combined approach to predict Henry Hub (U.S. market) monthly spot prices in the long run. In contrast, Jin and Kim (2015) suggest using a combination of wavelet decomposition and the ARIMA model, for more precise forecasting of Henry Hub weekly spot prices.

A growing body of literature, however, studies the impact of additional factors on natural gas prices and suggests that multivariate models are more precise for forecasting. For instance, Nick and Thoenes (2014) develop a structural VAR model for the German (NCG) gas hub over the period of 2008-2011 and find that in the short-term gas prices depend on the temperature, storage and supply shortfalls, while in the long-term crude oil and coal prices have an impact on gas price developments. Moreover, the authors argue that while supply interruptions have an impact on NCG gas prices, their effects might be overestimated, since some of the supply shortfalls overlapped with extraordinary demand-side conditions. Thus, it is important to not only focus on the supply-side aspects of the gas market in order to improve its security, but also to address the flexibility side of market demand.

Hamie (2020) tests an extensive set of methodologies to model natural gas prices, including game theory, information theory, records theory, non-parametric approaches, and the multivariate regression analysis. As for multivariate models, the author employs the VECM (vector error-correction model) to account for the effects of fundamental variables on gas price formation. Hamie (2020) argues that natural gas prices in the German hub (NCG benchmark) are affected by the weather conditions measured by heating degree days, the storage utilization of gas, coal and crude oil prices, the euro-dollar exchange rate, as well as by the lags of their own prices. At the same time, it is argued that many other factors might determine natural gas prices, including oil storage inventories, extreme weather events, political factors, and financial market conditions. Similarly, Berrisch and Ziel (2022) use state-space models to forecast daily and monthly gas prices based on various factors, such as seasonality, air temperature, risk premiums, storage levels, the price of European Emission Allowances, and the prices of oil, coal and electricity. As can be seen even from the examples above, apart from the main supply and demand factors, it is common to use weather-related variables to model natural gas prices. Temperature has an impact on gas consumption as the primary usage of gas is for heating purposes. Moreover, gas is also used in hot weather to cool buildings. Heating degree days (HDD) and cooling degree days (CDD) are the measures that quantify respective energy needs depending on the temperature (Sharma et al., 2021).

Gao et al. (2021) develop a class of hybrid time-varying parameter models (i.e. combinations of TVPSV and Markov switching classes of models) for three gas markets, namely the U.S., and the European and Japanese markets. The authors find evidence that time-varying models are better for forecasting European and Japanese monthly gas prices than static models, while for the U.S. market a simple AR model outperforms other studied approaches.

In recent literature, more and more papers focus on the Dutch TTF price as a benchmark, because the TTF is currently Europe's main gas hub, and it is becoming widely internationalized. Hulshof et al. (2016), for example, prove that daily TTF gas prices predominantly depend on market fundamentals, such as weather and storage availability, while the linkage between crude oil and natural gas prices is not strong over the period of 2011–2014, and coal prices are insignificant for the day-ahead forecasting of gas prices. Obadi and Korcek (2020) examine month-ahead TTF contracts over 2016–2019 and find evidence that monthly gas prices are driven by demand and supply fundamentals, like the price of German power and the price of coal, changes in total demand for gas, storage capacity, and in LNG variables.

The literature on predicting iron ore and steel prices is interrelated, given the direct links between these two commodity markets. Iron ore is the primary raw material that is used in the production of steel and steel products. Almost all iron ore that is mined is used in steelmaking, thus the demand for iron ore is primarily defined by the demand for steel. Therefore, the factors impacting global iron ore and steel prices are related.

The Asian market, more specifically Chinese consumers, play a great role in shaping the global iron ore market. China is a dominant consumer of metals in general and iron ore in particular, as it is the world's largest producer of steel. The growing importance of the Chinese market is often seen as one of the main reasons behind the transition of the iron ore pricing mechanism from an annual negotiation system to spot market pricing in late 2008–2010 (Wårell, 2014; Wårell, 2018). The author also finds GDP growth in China to have the strongest impact on iron ore prices. Ma and Zhen (2020) analyze spot prices for iron ore in 2014-2018 and find evidence that China's steel production affects the volatility of iron ore prices, while the mean and volatility of prices are also influenced by changes in port stocks.

Mei and Chen (2018) study the factors influencing steel overcapacity on the Chinese market and find that they include the steel export rate, investment in fixed assets, the growth rate of real estate construction areas, concentration levels in the iron ore and steel industry, iron ore prices, and local government investment. The capacity utilization rate has an impact on market competition and prices. For example, moderately excess capacity can improve competition and boost technological innovation in the industry, whereas severe overcapacity might provoke vicious competition, weak prices and a deterioration in the business environment.

There is a growing bulk of literature suggesting that there are links between the prices of different commodity groups. For example, Campiche et al. (2007) find cointegrating relationships between crude oil, corn and soybean prices over 2006–2007. Nazlioglu and Soytas (2012) prove the presence of dynamic cointegration links between the global oil price and prices of twenty-four agricultural commodities over an extended period of 1980–2010. Meanwhile, West and Wong (2014) employ factor models to predict the monthly prices of energy, metals and agricultural commodities using a sample of 1996 to 2012. Ding and Zhang (2020) use cross-market information from long-run equilibrium models to predict commodity prices, such as oil, copper, cattle, corn, and gold.

A number of papers argue that energy and crude oil prices can determine the prices of other commodities, including metals. The intuition behind this argument is that oil constitutes an important operational expense and a power source for the shipping industry, and commodities, such as metals, are often transported by sea. Therefore, the literature provides empirical evidence of crude oil prices having long-term cointegration relations with other commodity prices. For instance, Chou et al. (2012) prove, using a VARMA model, that global steel prices measured by the CRU steel price index were cointegrated with, and affected by, crude oil prices over the period of 2000 to 2010. Similarly, Asmoro (2017) has evidence that hot rolled coil (HRC) and billet steel prices are impacted by crude oil prices over the period from May 1996–December 2016. Moreover, studies by Chou et al. (2012) and Asmoro (2017) suggest that there is a unidirectional causal relation between crude oil and steel, i.e. the steel price is impacted by the crude oil price, whereas changes in oil prices are not influenced by steel prices.

Therefore, the literature on modeling the prices of energy and metals is quite extensive. However there are still some shortcomings, which could be improved. One of them is the primary usage of quarterly frequency data to predict commodity prices, which makes forecasts less detailed and ignores some important price reactions to changes in fundamental factors. Another drawback is that many authors focus on predicting real prices, commodity indices or futures prices, while changes in nominal spot prices are often of higher interest to central banks, researchers and think tanks. Moreover, the models used in the literature do not always incorporate a sufficient number of explanatory variables, focusing rather on the impact of a limited number of factors on commodity price developments. This study adds to the literature by 1) focusing on monthly rather than quarterly data to predict commodity prices in the short-run, 2) predicting the nominal spot prices of the commodities of interest, 3) using up-to-date global benchmarks for commodity prices, and 4) accounting for the comprehensive set of supply and demand factors that determine price movements. Moreover, the models applied in this study do not completely repeat the specifications used previously in the literature, but represent a hands-on approach to predicting commodity prices, taking into account the perspectives of central banks.

3. DATA AND METHODOLOGY

3.1. General Methodology

We use monthly data for the periods from 2003, 2004 or 2008 (depending on the availability of the data required by model specifications) up to February 2023 to examine the determinants of global commodity prices and construct 6-month-ahead forecasts. For instance, the models that include global manufacturing PMI as a proxy for global demand have their estimation samples starting in 2008 due to the limited availability of this data. Overall, the chosen time span is sufficient to analyze the dynamics of commodity prices and allows for some adjustment of the models to previous episodes of price volatility, relatively similar to the current ones. Although macro-forecasting processes in central banks normally focus on quarterly data, monthly variables better capture price developments and market changes, while also allowing one to make a more precise prediction of commodity prices than with quarterly data. Moreover, the quarterly projections of prices can easily be derived from our monthly forecasts by averaging, and can be used further for macroeconomic forecasting (Figure 1).

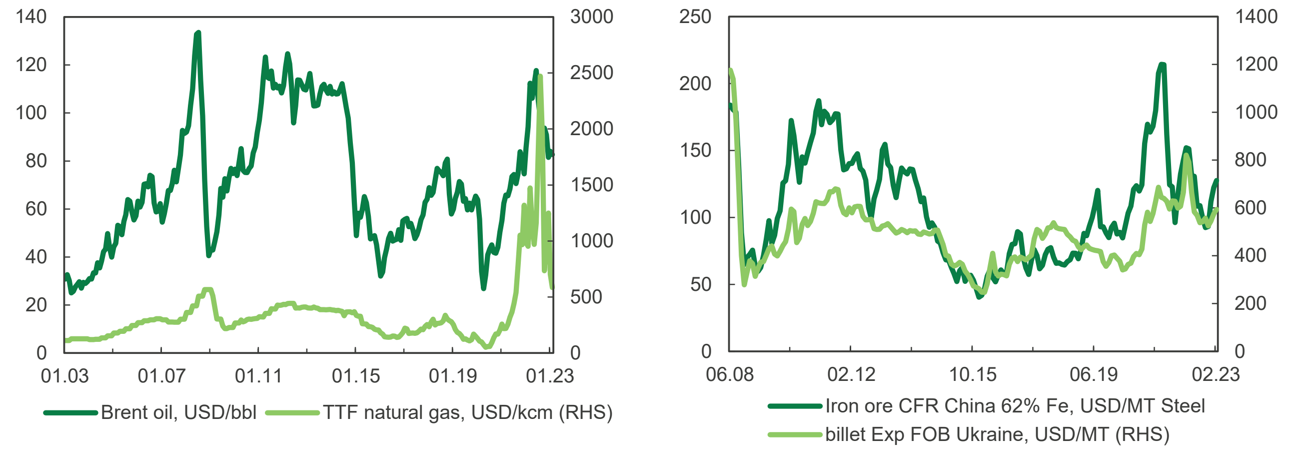


Figure 1. Nominal Global Prices of Energy Commodities and Metals

Source: World Bank, Thomson Reuters, Delphica.

In the models that are described below, the variables chosen have inter-links that are explained by economic theory. Moreover, they demonstrate Granger-causality relative to one another, which justifies the use of VAR models. All variables are tested for unit roots, and nonstationary series are transformed into stationary ones by simple differencing or log-differencing (see Table 5). In the case of cointegration relationships between the variables, error-correction models are used.

In order to account for the main demand and supply factors that affect price changes and for the impact of past observations on commodity prices, we employ standard vector autoregressive (VAR) and vector error-correction models (VECM) to regress the world prices of crude oil, natural gas, steel and iron ore, and to make projections. Although the choice of variables in the models and of forecasting approaches is based on the literature and economic theory, they do not completely repeat the models used previously in other studies. The methodology of this study represents a hands-on approach to predicting commodity prices, taking into account the perspectives of central banks, and shows the impact of various factor combinations on price changes. We compare the forecast accuracy of the models (measured by root mean square errors) to that of a random walk forecast and perform an out-of-sample forecasting exercise.

The VAR model

The general representation of a standard reduced-form VAR model with p lags can be written as follows:

$$y_t = c + B_1 y_{t-1} + B_2 y_{t-2} + \ldots + B_p y_{t-p} + u_t, \quad u_t \sim (0, \Omega), \quad (1)$$

where y_t and c are $K \times 1$ vectors of K monthly variables and constants, respectively, and $B_i, i = 1,\ldots,p$ are $K \times K$ matrices of coefficients. The residuals u_t are assumed to be i.i.d. N(0, Ω), where Ω is the variance-covariance matrix of innovations.[1]

Equation (1) indicates that any series in the model depends on the past values of all the K series through their lags. For example, if the number of variables K in the system equals 2, and the number of lags p = 2, the VAR(2) process can be rewritten as follows:

$$\begin{bmatrix} y_{t,1} \\ y_{t,2} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} B^1_{1,1} & B^1_{1,2} \\ B^1_{2,1} & B^1_{2,2} \end{bmatrix} \begin{bmatrix} y_{t-1,1} \\ y_{t-1,2} \end{bmatrix} + \begin{bmatrix} B^2_{1,1} & B^2_{1,2} \\ B^2_{2,1} & B^2_{2,2} \end{bmatrix} \begin{bmatrix} y_{t-2,1} \\ y_{t-2,2} \end{bmatrix} + \begin{bmatrix} u_{t,1} \\ u_{t,2} \end{bmatrix}, \quad (2)$$

where the subscripts indicate equation and variable numbers, and the superscripts refer to the lag number. Thus, the VAR(p) is an example of a seemingly unrelated regression (SUR) model with lagged variables and deterministic terms as common regressors (Zivot and Wang, 2003). In addition, some other deterministic terms (seasonal dummies, a linear trend, and a set of exogenous variables) can be included in the VAR system.

The VECM Model

When variables in such a system are cointegrated, the vector error correction model (VECM) should be used rather than a standard VAR model. Variables are said to be cointegrated if each of them are non-stationary with a unit root (I(1)), while there is some linear combination $a'y_t$ of these series which are I(0), i.e. stationary. Here a is a non-zero $K \times 1$ vector.

Let us consider again a VAR(p) process as in equation (1). In lag operator notation, this equation can be written as follows:

$$B(L) Y_t = c + u_t, \quad (3)$$

where $B(L) = I_K - B_1 L - \ldots - B_p L^p$. This VAR system is stable if the roots of the polynomial

$$\det\left(I_K - B_1 z - B_2 z^2 - \ldots - B_p z^p\right) = 0$$

lie outside the complex unit circle, or have a modulus greater than 1. If at least one series among $y_{K,t}$ is I(1), the VAR(p) is unstable, since $\Pi = -(I_K - B_1 - B_2 - \ldots - B_p)$ is singular, det(Π) = 0, and the roots lie on the unit circle.

In general, cointegration means that there is some long-term relationship between the individual elements of y_t, which is represented by the linear combination $a'y_t$. A VECM is a special type of VAR model, which introduces error-correction terms into the system. A VECM focuses on differences to account for short-run relationships between variables (as represented in a VAR), while its error-correction terms and cointegrating equations account for short-run adjustments and long-run cointegrating relationships. For example, the VECM(2) system for two variables $y_{t,1}$ and $y_{t,2}$ can be specified as:

$$\begin{aligned} \Delta y_{t,1} &= c + B^1_{1,1}\Delta y_{t-1,1} + B^2_{1,1}\Delta y_{t-2,1} + B^1_{1,2}\Delta y_{t-1,2} + \\ &+ B^2_{1,2}\Delta y_{t-2,2} - \lambda_1\left(y_{t-1,1} - \alpha_0 - \alpha_1 y_{t-1,2}\right) + u_{t,1} ; \\ \Delta y_{t,2} &= c + B^1_{2,1}\Delta y_{t-1,1} + B^2_{2,1}\Delta y_{t-2,1} + B^1_{2,2}\Delta y_{t-1,2} + \\ &+ B^2_{2,2}\Delta y_{t-2,2} - \lambda_2\left(y_{t-1,1} - \alpha_0 - \alpha_1 y_{t-1,2}\right) + u_{t,2} , \end{aligned} \quad (4)$$

where $y_{t,1} = a_0 + a_1 y_{t,2}$ is the long-run cointegrating relationship between the variables $y_{t,1}$ and $y_{t,2}$, and λ_1 and λ_2 are the error-correction terms. The error-correction terms measure the response of the variables $y_{t,1}$ and $y_{t,2}$ to deviations from long-run equilibrium. As in (2), the subscripts in the system indicate equation and variable numbers, and the superscripts refer to the lag number. If a VEC model has more than two variables, it means that there can be more than one cointegrating relationship. The number of these relationships can be determined using cointegration tests. Also one should note, that the VECM(2) in example (4) is derived from the VAR(3) model, since the VECM focuses on differences, and for a VAR(p) model the corresponding VECM would be with (p – 1) lags.[2]

3.2. Crude Oil Price Forecasting

To model global crude oil prices and to produce short-term forecasts, we use general VAR methodology introduced by Baumeister and Kilian (2013, 2014), with some adjustments. We use four standard VAR models with slightly different specifications to forecast monthly crude oil prices in the short run. The specifications of the models, which are based on the literature and economic theory, were adjusted in order to incorporate additional factors of interest that are relevant to the current period.

[1] The number of lags (p) is obtained on the basis of some theoretical models, by using a rule of thumb, or by statistical selection criteria, such as the Akaike information criterion (AIC), the Schwarz-Bayesian criterion (BIC) and Hannan-Quinn (HQ) criterion.

[2] For a more detailed description of VAR and VECM estimations, please refer to Hamilton (1994), Zivot and Wang (2003), Ouliaris et al. (2018).

They were also adjusted for analytical purposes to test the impact of various factor combinations on prices. For the first model, the vector of endogenous variables consists of 1) the real price of Brent oil (the nominal price deflated by the U.S. CPI) as an international benchmark, 2) the percentage change in crude oil production, 3) the percentage change in OECD petroleum inventories (as a proxy for changes in global inventories), and 4) the index of global real economic activity also known as the Kilian (2009) shipping index (as a proxy for global demand). One of the important factors affecting real commodity prices is the shift in demand for commodities which, in turn, is caused by unexpected fluctuations in real global economic activity. Kilian's index is a business-cycle indicator, which is derived from global bulk dry cargo shipping rates, and is expressed as a percentage deviation from the trend. Kilian's index proves to be a good monthly indicator of the state of, and changes in, the global economy. We also find that it is a more convenient indicator than the index of monthly GDP of OECD+Major 6 NMEs calculated by the OECD, as the latter is available with significant lags, and does not capture major global economic fluctuations. For the discussion of the advantages of Kilian's index, see Kilian (2009), and Kilian and Zhou (2018). Since our aim is to forecast *nominal* rather than *real* oil prices, we follow Beckers and Beidas-Strom (2015) and introduce a fifth variable, the U.S. CPI (index 1982 – 1984 = 100, seasonally adjusted, from the FRED database of the St. Louis Fed), into the vector of endogenous variables in our VAR models. That makes it possible to produce forecasts for both the real price of oil and the consumer price index, while also deriving a forecast of the nominal oil price. The vector of exogenous variables includes constants and eleven seasonal dummies, as in a paper by Beckers and Beidas-Strom (2015). The standard methodologies used in the aforementioned papers suggest including 12 lags as a rule of thumb for the models based on monthly data, or four lags for those based on quarterly data, respectively. Notwithstanding that, our standard VAR model has six lags, which is explained by the Akaike and Schwarz-Bayesian information criteria, and the model's estimation sample starts in August 2003. The real price variable and the CPI are log-differenced to make them stationary, Kilian's index is taken in the first difference, while two supply-side determinants expressed as percentage changes are already stationary.

The second model's specification is slightly different for analytical purposes and it shows the impact of a different combination of explanatory factors on price changes. Here we use the J. P. Morgan Global Manufacturing PMI as our global demand proxy, instead of Kilian's index to test if a model with a different demand-side variable would prove more accurate in terms of forecasting. We also change the representation of the oil production variable by expressing it as the log-difference of production levels rather than calculating the percentage change in production. Due to the limited availability of the PMI time series, the sample for this model is shorter and starts in April 2008. We also employ two more models with the same specifications and sample length as the second one, but with a different U.S. inflation variable, which is a non-seasonally adjusted index, 2010 = 100 (as a result, the real price of crude oil differs too). The model number three has six lags, like the first two models, which is based on the AIC and BIC. The fourth model has three lags in order to better capture the most recent movements in oil prices. Moreover, lag exclusion tests also show that a higher number of lags might be unnecessary for this model.

3.3. Natural Gas Price Forecasting

In order to model and forecast TTF gas prices, we apply VAR and VECM approaches. The choice of the explanatory variables is based on the research literature and the fundamentals for the European gas market (Nick and Thoenes, 2014; Hamie, 2020; Berrisch and Ziel, 2022). The first model is a VAR(3) that uses the price of Dutch TTF gas, the Brent oil price, Kilian's index of global real economic activity, the global manufacturing PMI, changes in natural gas reserves in the Netherlands, gas stock changes, and natural gas supply variables in first difference, as well as a vector of constants in exogenous variables. The price of oil is included into the model since it is a close substitute for natural gas, and the prices of these two energy resources normally tend to move in similar directions. The Kilian and PMI indices are used as proxies for global demand factors, whereas stock changes and gas supply and stocks represent the supply-side determinants of gas prices. The second model basically has the same specifications, except that natural gas stocks are an exogenous variable.

With a view to conducting an in-depth analysis as to whether or not gas prices have similar determinants as crude oil prices, we also apply two of the oil forecasting models to predict natural gas prices. Thus, model number three uses real rather than nominal prices of TTF gas, the change in oil production, the change in petroleum inventories, Kilian's index, and the U.S. CPI. It also incorporates the vectors of constants and seasonal dummies into the set of exogenous variables, as in the oil forecasting models, but has 12 lags, as suggested in the literature and is confirmed by the information criteria. The fourth model, VAR(6), incorporates the real price of gas, an oil production variable (the first difference of natural logarithms rather than a percentage change), the PMI, the change in petroleum inventories, and the CPI, which are all used as endogenous variables. Similarly, there are seasonal dummies and constants used in the vectors of exogenous variables. The number of lags for this model's specification is explained by the Akaike criterion and lag exclusion tests. These models are meant to test whether or not the factors influencing oil prices can be reliably used to model and predict natural gas prices, without including gas-specific market determinants.

Finally, the fifth model includes gas and oil nominal prices, Kilian's index and the PMI, gas supply, and gas stock changes, as in the case of our second model. However, it has a different set of exogenous variables, namely seasonal dummies (as in the third and fourth models) and a weather conditions proxy measured in degree days, or DDs. As described earlier in the literature review section, temperature conditions play a crucial role in shaping natural gas consumption and, consequently, prices. When air temperatures are abnormally low, there is greater demand for heating and natural gas, which leads to higher gas prices. Likewise, very high temperatures increase the need to cool buildings, and natural gas is also widely used for these purposes. Heating and cooling degree days (HDDs and CDDs, respectively) are weather-based technical indicators that measure the energy requirements of buildings in terms of heating and cooling. For example, if one compares energy needs in 1979 and

2022 in the EU, HDD values declined by 19% during this period indicating that heating needs in 2022 were roughly two-tenths lower on average than those in 1979. At the same time, CDD values in the EU were almost four times higher in 2022 compared to 1979, showing the increased need for air conditioning and higher energy consumption over decades (Eurostat, 2023).

Although we are interested in modeling gas prices for the Dutch TTF hub, which is a benchmark for the European market, degree days in the model are those related to weather conditions in Germany, rather than in the Netherlands. The intuition behind this is that Germany is the largest natural gas consumer in the EU. According to Eurostat's final energy consumption indicator, which measures the energy consumption of end-users (industry, transport, households, agriculture, and services), Germany accounted for almost 27% of the total consumption of natural gas in the European Union in 2021. For comparison, Italy takes second place, but its end-users consume only about 16% of the total natural gas quantities consumed in the EU. Moreover, according to annual data for 2021, Germany is the dominant leader in the final consumption of heat with a share of over 21% of the EU's total, and the top electricity consumer with a share that almost equals 20%. And these shares have been stable or even growing over the years. Therefore, we include degree days data for Germany in our model number five. Since DD is an exogenous variable, we need readymade forecast values of it for the whole forecast horizon, but getting these weather forecasts on a country level is too complicated. Thus, we focus on regional data and gather DDs for the most populous cities, such as Berlin and Munich, and important industrial towns, including Ludwigshafen am Rheine, Wittenberg, and Hamburg. According to the Federal Statistical Office of Germany, the latter three cities are leaders in natural gas consumption in German industry. We collect historical degree days from Eurostat, where these indicators are calculated as follows:

$$\textit{If } T_m \leq 15°\text{C} \quad \textit{Then } [\text{HDD} = \textstyle\sum_i (18°\text{C} - T_m^i)] \quad \textit{Else } [\text{HDD} = 0]$$

$$\textit{If } T_m \geq 24°\text{C} \quad \textit{Then } [\text{CDD} = \textstyle\sum_i (T_m^i - 21°\text{C})] \quad \textit{Else } [\text{CDD} = 0], \quad (5)$$

where T_m^i is the mean air temperature of day i. The base temperatures for HDDs and CDDs are set to 15°C and 24°C, respectively, in accordance with the general climatological approach. These calculations are made on a daily basis and then added up to monthly figures, which we use. We then calculate total monthly degree days (DDs) as the sum of HDDs and CDDs of all chosen cities as a proxy for German energy needs. As Eurostat updates degree days once per year for the full year that has passed, we gather daily air temperature data for the aforementioned cities from the website https://www.weather25.com/ and use the formulae to produce the missing actual values of monthly DDs. We obtain the projected HDD and CDD values that are needed for the model from https://cds.climate.copernicus.eu/. These values are modelled on the basis of historical averages over a 30-year period taking into account climatological inputs. This model is a VECM since there are cointegration relationships between the variables in this specification, and the inclusion of two lags is explained by tests and information criteria.

The estimation samples of all of the models that predict natural gas prices start in either April or May 2008 and run through February 2023, except for the third model, which has its estimation sample starting in February 2004, as data for that period was more readily available.

3.4. Iron Ore Price Forecasting

We apply standard VAR methodology to model iron ore prices. Again, we use four VAR models with slightly different specifications to better capture the impact that various combinations of factors might have on the price. After that, we construct a baseline forecast as an average of four approaches, which helps to combine the benefits of individual forecasting models and performs equally well during different periods. The models applied to use the monthly spot prices of iron ore fines, CFR China 62% Fe, from the World Bank database, as it is the most commonly used global benchmark. We deflate the nominal price by the U.S. CPI (index, 1982 – 1984 = 100) to obtain the real price. Similarly, the CPI is also included in the vector of endogenous variables, and the forecasts of the two variables are then used to obtain nominal price predictions.

The iron ore market is significantly influenced by the steel market since iron ore is primarily used in steelmaking. Moreover, China's large steel market makes it a big player, so it has a great impact on iron ore prices, especially from the demand side. China is the dominant consumer of metals in general and iron ore in particular, as it is the world's largest producer of steel. Moreover, the Chinese construction sector and infrastructure projects require substantial amounts of materials, such as steel. With that in mind, we use China's crude steel production (monthly growth rate expressed as a percentage) as a proxy for global demand for iron ore in all four models. Models one and two also use changes in the Baltic Dry Index, i.e. the Baltic Exchange's main sea freight index. We expect an increase in freight rates to push up commodity prices as well. The Baltic Dry Index is available on a daily basis, but we transform it into a monthly series and then calculate monthly percentage changes. From the supply side, we add a change in Brazil's exports of iron ore as this country, together with Australia, are traditionally major exporters of iron ore. As Australia's detailed export data was not made available when we were collecting data, we focused on Brazil's exports as the second largest iron ore exporter to China. The weather conditions in Brazil, other disruptions to its economy and the mining sector, as well as the policies of the mining giant Vale are known to influence Brazil's iron ore production and export volumes and, through these, world prices. Two out of four VAR models have these variables and a vector of constants to model iron ore prices, however, they have different numbers of lags – two and one, respectively. This is attributable to different lag suggestions by statistical information criteria, and to ambiguous test results for the optimal number of lags that should be used.

The third model incorporates China's Manufacturing Purchasing Managers' Index (PMI), which is published by the National Bureau of Statistics of China on a monthly basis. This variable is a proxy for the state of the Chinese economy in general and the health of its industrial sector in particular. China is the largest global importer and one of the biggest producers of iron ore, as a result of which its economic development is expected to impact world iron ore prices. The other endogenous variables are the same as in the first two models, except for the exclusion of Brazil's export variable, and the optimal testing-based number of lags for such a specification equals five. In contrast, the fourth model is more of a combination of the aforementioned specifications. This two-lag VAR model consists of the real price of iron ore, changes in steel production in China, and changes in the Baltic Dry Index, China's Manufacturing PMI, Brazil's iron ore exports, and in the U.S. CPI. Exogenous variables include constants (as in other models) and the recession dummy for the U.S., i.e. an NBER-based recession indicator, is available from the FRED database of the FRB of St. Louis. The recession dummy takes the value of 1 from January 2008 through June 2009 and from March 2020 through April 2020, which represents the recessionary periods in the U.S. Again, where necessary, the variables are transformed to log-difference or first difference forms or are winsorized to smooth out the outliers (like the change in Brazil's iron ore exports). Depending on data availability, and after some adjustments to the series, the estimation samples start either in April 2004, March 2005, or June 2005 and run through February 2023, as in all studied commodity groups.

3.5. Steel Price Forecasting

In order to model and predict global steel prices based on market fundamentals, we apply the VECM methodology. Error-correction is needed since there are cointegration relationships between the variables. As mentioned above, the global steel and iron ore markets are very interrelated, so the explanatory variables for steel prices are very similar to those used in iron ore forecasting.

As Ukraine used to be among the global top ten largest steel exporters (before russia's full-scale invasion of Ukraine and the port blockade), we are interested in forecasting prices for Ukrainian steel. As a benchmark, we use the monthly averages of daily steel billet prices, FOB Ukraine. For a period after 24 February 2022 (the start of the full-scale war), we use proxy prices for Ukrainian steel calculated based on either Turkey C&F steel billet prices (up until October 2022) or the Black Sea billet FOB UA prices. After calculating the monthly average steel price, we then take the log of it as we do for other price variables.

Explanatory variables in the first VECM include the price of iron ore fines (CFR China 62% Fe) and the price of coal (Australian thermal coal, FOB Newcastle, 6,000 kcal/kg futures price from the World Bank database) as inputs used to produce steel. Moreover, we use the Brent oil price (expressed in logs) as a determinant of global steel prices. There is a growing bulk of literature suggesting that there are cross-market price links between various commodity groups that can be used to predict prices (see, e.g., Campiche et al., 2007, Nazlioglu and Soytas, 2012, Ding and Zhang, 2020). More specifically, there is empirical evidence of crude oil prices having long-term cointegration relationships with other commodity prices and having an impact on their prices. Moreover, it is believed that there is a unidirectional relationship between crude oil and steel as suggested, for instance, by Chou et al. (2012) and Asmoro (2017).

Given the interrelations between the iron ore and steel markets, we also include changes in the Baltic Dry Index and China's Manufacturing PMI into this model, just as we did in the VARs for iron ore prediction. This VECM has one lag determined by the lag length selection tests and also incorporates a recession dummy in its exogenous variables. The second model is simplified for analytical purposes – it only includes the prices of steel and iron ore and the Baltic Dry Index. The optimal number of lags equals one, and there are no exogenous variables included. The third VECM is the price-only model as it includes the prices of steel, iron ore, coal and crude oil, and no other explanatory variables. This specification requires two lags, as chosen by the AIC and BIC statistical criteria and lag exclusion tests. Finally, the fourth model is a two-lag VECM incorporating steel, iron ore and coal prices, changes in the Baltic Dry Index and the manufacturing PMI for China, as well as a recession binary variable. After making necessary adjustments to the data and taking into account the time span of the data, the estimation samples start either in August 2008 or September 2008 and last until February 2023.

4. RESULTS

In this section, we provide the results of the forecasting performance of the models used to predict commodity prices six months ahead. We run the models to forecast oil and natural gas prices from the beginning of the models' estimation samples until the end of 2015 and then do out-of-sample six-month-ahead forecasting simulations starting from January 2016 through February 2023. For the models that forecast iron ore and steel prices, the out-of-sample exercise starts in January 2018 to better adjust the models to changes in the pricing regime of iron ore prices. Next, we calculate the root mean square errors (RMSEs) of all individual models based on forecast simulations and divide them by the RMSEs of the respective random walk (RW) models. Figures 2–5 (Supplementary Materials) depict the results of the out-of-sample forecast simulations six months ahead for the nominal prices of the four commodities of interest. Tables 1–4 provide a summary of the relative RMSEs of the models for one- to six-month-ahead horizons. Values below 1, highlighted in green, mean that the RMSEs of the given models are lower than the random walk RMSEs. This means that the forecasting power of a given

Table 1. The RMSEs of Individual Models Relative to RW RMSEs – Crude Oil

	# of lags	Forecast horizon, months ahead					
		1	2	3	4	5	6
VAR_1	6	**0.939**	0.776	**0.701**	0.737	0.823	0.803
VAR_2	6	1.067	0.826	0.827	0.816	0.896	0.875
VAR_3*	6	1.052	0.818	0.816	0.819	0.912	0.907
VAR_4*	3	1.005	**0.772**	0.747	**0.735**	**0.775**	**0.775**

* VAR models 3 and 4 use a different U.S. CPI index (non-seasonally adjusted index, 2010=100), which also causes variations in real prices. Therefore, the RMSEs of these models are compared to the RMSEs of a RW model, which also uses real prices calculated on the basis of a non-seasonally adjusted CPI index, whereas models 1 and 2 are compared to a RW model based on comparable real prices (where the CPI index, s.a., 1982-1984=100 is a deflator).

model is higher than that of an RW benchmark. Values above 1, highlighted in red, indicate that the given models fail to outperform the respective RW models over the given forecast horizons. The lowest relative RMSEs over each of the six horizons are presented in bold, indicating the best-performing models.

The models that predict oil prices demonstrate good forecasting performance over almost all forecasting horizons (Table 1). It is only in the one-month-ahead forecast that the random walk outperforms three out of four of the selected models. Note that models VAR_3 and VAR_4 use real oil prices calculated based on a different CPI index compared to models 1 and 2 (see the Data and Methodology section). Therefore, in order to make fair conclusions, we divided their RMSEs by the RMSEs of a different random walk model, which used comparable real prices.

The first model, which uses oil production, petroleum inventories, Kilian's index, and the CPI index as endogenous oil price determinants, improves benchmark RW forecasts over all forecast horizons, and is the best-performing model to predict Brent prices one and three months ahead. VAR model number 4, despite having the smallest number of lags – i.e. 3 lags, has the smallest number of forecast errors over two-, four-, five- and six-month-ahead horizons among all the other models. This may suggest that, in the case of the short-term forecasting of monthly oil prices, the information contained in only three lags of fundamentals might be enough to predict the future movement of oil prices. This finding is new and adds to the modern techniques of oil price forecasting, which normally rely on up to 12 lags of information. Models 2, 3 and 4 also have an estimation sample that is almost five years smaller than the sample of the first model due to the limited availability of the manufacturing PMI index that they use instead of Kilian's IGREA.

All of the models used to predict TTF gas prices in the short run outperform the RW benchmark model during all forecast periods, as shown in Table 2. Figure 3 (Supplemenary Materials) presents the results of the out-of-sample forecasting exercise.

The most consistent results are produced by the two-lag VECM model, which uses the weather conditions variable – degree days – as an exogenous one. This once again proves the importance of air temperature conditions for gas price forecasting. It is also worth noting that models 3 and 4, which use crude oil-related variables (such as the change in oil production and petroleum inventories) and general demand-side variables (such as Kilian's index and the PMI) and do not use gas-specific market fundamentals, also show good results and even outperform other models over some forecast horizons. These results add to the empirical evidence of the crude oil and natural gas markets being highly interrelated, making room for further investigation of cross-market energy price predictions.

Tables 3 and 4 present the relative RMSEs of the models that predict the prices of iron ore and steel, respectively, while Figure 4 and 5 (Supplemenary Materials) show the forecasting simulations of nominal prices. For these two commodities, we also test the models' performance compared to a random walk process. However, we include simple AR(1) models for comparison, since they are often used as a benchmark to test the forecasts of metal prices, or are extended to ARIMA models as standalone forecasting techniques (Pincheira and Hardy, 2019).

All of the selected models that forecast global iron ore prices in general produce better forecasts than both RW and AR(1) models. VAR_1 and VAR_2, which have the same specifications but a different lag number, are the best-performing approaches. This can be explained by these models having the most comprehensive set of supply and demand variables, including China's steel production, the Baltic Dry Index of freight rates and Brazil's exports of iron ore. As expected, the model with two lags produces better results over longer horizons. Model 4, which incorporates the recession dummy, also proves reliable. At the same time, model 3, which excludes Brazil's ore exports, performs worse than other models, despite having the highest lag number. However, the chart of out-of-sample simulations for model 3 indicates that it might have a better predictive power for more unstable periods, similar to those that we are currently observing on the markets, while other models are relatively better for stable periods (Figure 4 in Supplemenary Materials).

VECM number 1, which is used to predict steel prices, outperforms other approaches over half of the forecast periods. The model includes the prices of iron ore, coal and oil, the freight rates index, China's manufacturing sector proxy, as well as the recession dummy. It is also noteworthy that this VAR model has the smallest possible number of lags and shows better results than the two-lag models. AR(1) only slightly outperforms model 1 over one- and two-month horizons, but lags behind over longer forecast periods. Interestingly, the two-lag price-only model (VECM_p_3) which only uses the prices of iron ore, coal, and crude oil to predict steel prices, has the smallest number of RMSEs six months ahead (Table 4). This finding can be further developed and tested in future research into the longer-term forecasting of steel prices.

The results of individual models indicate that our choice of forecasting techniques and explanatory variables is reliable and the models can be used to predict commodity prices, at least in the short run. Moreover, given the generally high performance of these models and their varying forecast

Table 2. The RMSEs of Individual Models Relative to RW RMSEs – Natural Gas

	# of lags	Forecast horizon, months ahead					
		1	2	3	4	5	6
VAR_1	3	0.840	0.689	0.840	0.917	0.855	0.916
VAR_2	3	0.845	**0.681**	0.820	0.927	0.836	0.897
VAR_3	12	**0.791**	0.684	0.863	0.861	0.855	0.770
VAR_4	6	0.823	0.691	0.850	0.834	0.826	**0.766**
VECM_DD_5	2	0.829	0.695	**0.802**	**0.826**	**0.826**	0.850

Table 3. The RMSEs of Individual Models Relative to RW RMSEs – Iron Ore

	# of lags	Forecast horizon, months ahead					
		1	2	3	4	5	6
AR(1)	1	0.7893	**0.7449**	0.6906	0.6296	0.6801	0.6932
VAR_1	2	0.8190	0.7719	0.6916	**0.6260**	**0.6747**	**0.6926**
VAR_2	1	**0.7856**	0.7451	**0.6895**	0.6273	0.6794	0.6931
VAR_3	5	0.9374	0.8373	0.7445	0.6957	0.7141	0.7178
VAR_4	2	0.8414	0.7772	0.7084	0.6378	0.6840	0.6956

Table 4. The RMSEs of Individual Models Relative to RW RMSEs – Steel

	# of lags	Forecast horizon, months ahead					
		1	2	3	4	5	6
AR(1)	1	**0.8324**	**0.6943**	0.6855	0.7357	0.7443	0.7126
VECM_1	1	0.8918	0.7014	**0.6802**	**0.7275**	**0.7390**	0.7073
VECM_2	1	0.9001	0.7521	0.7620	0.8047	0.7801	0.7122
VECM_p_3	2	0.9592	0.9286	0.8478	0.8052	0.7456	**0.7031**
VECM_4	2	1.0598	0.8798	0.8586	0.9515	0.8573	0.7536

precision over different periods, it makes sense to apply a combination approach and to merge the models' benefits to generate a combined baseline forecast for each of the commodities.

5. CONCLUSIONS

This study offers a relatively simple hands-on approach to forecasting the global prices of crude oil, natural gas, iron ore, and steel. In line with the modern literature, we apply VAR and VECM approaches based on demand and supply factors to forecast commodity prices over the short term period. This paper adds to the literature in a few ways.

First, unlike most other similar papers, the forecasting models in this paper focus on predicting monthly rather than quarterly prices (while being developed from the central bank's perspective). The rationale behind this is that monthly time series are more detailed and contain more information about price movements. Thus, generating monthly rather than quarterly forecasts increases forecast precision. Moreover, generated monthly price forecasts can then be used to construct more reliable quarterly projections than those derived from smoothed quarterly data. This, in turn, could improve the forecast performance of other central banks' macroeconomic quarterly projection models that use commodity price forecasts as inputs or assumptions.

Second, in our models we forecast real prices as well as inflation indices in order to construct forecasts of nominal commodity prices, which are of greater interest to us. Furthermore, this study focuses on spot prices, and does not include futures-based predictions, which are still popular among many central banks and forecasters, despite their being rather inaccurate under the current conditions. There is no need to include such models in the set of our forecasting techniques as reliance on futures prices does not necessarily provide robust outcomes for forecasting spot prices.

Third, our findings suggest that, among the individual models in each of the four commodity groups, the models with the most balanced and comprehensively chosen fundamental explanatory variables, which cover supply and demand fundamentals equally, prove the most reliable in terms of forecasting. These fundamentals, which are important for commodity price prediction, include, inter alia, stock changes, changes in commodity production volumes, export volumes by the largest players, changes in the manufacturing sector of the largest consumers, the state of global real economic activity, freight rates, and a recession indicator. Seasonal factors play an important role in shaping commodity prices as well. Moreover, Kilian's index of global real economic activity is found to be a useful proxy for global demand and a reliable input in forecasting both energy and metal prices. In the case of iron ore and steel prices, developments in the Chinese economy prove to be essential inputs.

Furthermore, we demonstrate that when predicting energy and metal prices in the short run, the models with smaller lag orders tend to outperform those with a higher number of lags, which is a new finding in the literature, to our knowledge. While literature usually suggest using up to 12 lags (as a rule of thumb in vector models to predict monthly commodity prices), we show that the most important information in terms of short-term price prediction can be found in the most recent historical data, and there is no need to overload the models. The conducted lag length selection tests and information criteria demonstrate that our models require no more than six lags. Moreover, the models with a smaller number of lags in general show higher forecast accuracy, as can be seen from the RMSE tables. This finding can be used and further developed by both researchers and forecasters with the purpose of finding the best-fitting forecasting techniques.

Finally, we conclude that selected individual models, while showing standalone high performance, have varying forecast accuracy over different periods. Our findings show that no individual model outperformed the others consistently throughout the forecast horizon. Thus, it might make sense to apply a combination approach to merge the models' benefits and generate a combined baseline forecast for each of the commodities.

The methodology used in this study is a hands-on approach to forecasting commodity prices in the short run. That notwithstanding, there is room for further research and improvement of the models, for example by using Bayesian techniques or applying non-equally weighted model combinations based on forecast errors. Moreover, it can be observed from the results that the forecasting accuracy of some of the models deteriorated during the most recent periods, in particular for natural gas (see, for example, the out-of-sample simulations). Therefore, under the current conditions of high uncertainty and abnormal movements in commodity markets, further development of crisis-time forecasting techniques might be needed.

REFERENCES

Alquist, R., Kilian, L., Vigfusson, R. J. (2011). Forecasting the price of oil. Working Paper, 2011–15. Ottawa: Bank of Canada. Retrieved from https://www.bankofcanada.ca/wp-content/uploads/2011/06/wp2011-15.pdf

Ambya, A., Gunarto, T., Hendrawaty, E., Kesumah, F. S. D., Wisnu, F. K. (2020). Future natural gas price forecasting model and its policy implication. International Journal of Energy Economics and Policy, 10(5), 64–70. https://doi.org/10.32479/ijeep.9676

Asmoro, T. H. (2017). A VAR model to investigate the volatility of line-pipe steel prices using oil price as a referred Currency. PM World Journal, 6(11), 1–12. Retrieved from https://pmworldlibrary.net/wp--content/uploads/2017/11/pmwj64-Nov2017-Asmoro-VAR-Model-Line-Pipe-Oil-Price-featured-paper.pdf

Baffes, J., Nagle, P. (eds.) (2022). Commodity Markets: Evolution Challenges and Policies. Washington, DC: World Bank. http://doi.org/10.1596/978-1-4648-1911-7

Baumeister, C., Kilian, L. (2013). Forecasting the real price of oil in a changing world: a forecast combination approach. Working Paper, 2013–28. Ottawa: Bank of Canada. Retrieved from https://www.bankofcanada.ca/wp-content/uploads/2013/08/wp2013-28.pdf

Baumeister, C., Kilian, L. (2014). What central bankers need to know about forecasting oil prices. International Economic Review, 55(3), 869-889. https://doi.org/10.1111/iere.12074

Beckers, B., Beidas-Strom, S. (2015). Forecasting the nominal Brent oil price with VARs – one model fits all? IMF Working Paper, 15/251. Washington: International Monetary Fund. Retrieved from https://doi.org/10.5089/9781513524276.001

Berrisch, J., Ziel, F. (2022). Distributional modeling and forecasting of natural gas prices. Journal of Forecasting, 41(6), 1065–1086. https://doi.org/10.1002/for.2853

Bowman, C., Husain, A. (2004). Forecasting commodity prices: futures versus Judgment. IMF Working Paper, 04/41. Washington: International Monetary Fund. Retrieved from https://doi.org/10.5089/9781451846133.001

Campiche, J. L., Bryant, H. L., Richardson, J. W., Outlaw, J. L. (2007). Examining the evolving correspondence between petroleum prices and agricultural commodity prices. American Agricultural Economics Association Annual Meeting, Portland, OR, July 29-August 1, 2007. https://doi.org/10.22004/ag.econ.9881

Chernenko, S. V., Schwarz, K. B., Wright, J. H. (2004). The information content of forward and futures prices. International Finance Discussion Paper, 808. Board of Governors of the Federal Reserve System. Retrieved from https://doi.org/10.17016/ifdp.2004.808

Chou, M.-T., Yang, Y.-L., Chang, S.-C. (2012). A study of the dynamic relationship between crude oil price and the steel price index. Review of Economics and Finance, 2, 30–42. Retrieved from http://www.bapress.ca/Journal-7/A%20Study%20of%20the%20Dynamic%20Relationship%20between%20Crude%20Oil%20Price%20and%20the%20Steel%20Price%20Index.pdf

Ding, S., Zhang, Y. (2020). Cross market predictions for commodity prices. Economic Modelling, 91, 455–462. https://doi.org/10.1016/j.econmod.2020.06.019

ECB (2015). Forecasting the price of oil. Economic Bulletin, 4(3), 87–98. Frankfurt: European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/other/art03_eb201504.en.pdf

Eurostat (2023). Heating and cooling degree days at EU level. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Heating_and_cooling_degree_days_-_statistics#Heating_and_cooling_degree_days_at_EU_level

Gao, S., Hou, C., Nguyen, B. H. (2021). Forecasting natural gas prices using highly flexible time-varying parameter models. Economic Modelling, 105, 105652. https://doi.org/10.1016/j.econmod.2021.105652

Hamie, H. (2020). Econometric modeling of natural gas prices : different modeling approaches and case studies related to gas markets (dissertation). Technische Universität Wien. https://doi.org/10.34726/hss.2020.41262

Hamilton, J. D. (1994). Time Series Analysis. Princeton, New Jersey: Princeton University Press.

Hosseinipoor, S., Hajirezaie, S., Hadjiyousefzadeh, N. (2016). Application of ARIMA and GARCH models in forecasting the natural gas prices. Retrieved from https://ou.edu/content/dam/cas/economics/Student%20Journal%20of%20Economics%20publications/Saied%20Hosseinipoor_AppJOE.pdf

Hulshof, D., van der Maat, J.-P., Mulder, M. (2016). Market fundamentals, competition and natural gas prices. Energy Policy, 94, 480–491. https://doi.org/10.1016/j.enpol.2015.12.016

Jin, J., Kim, J. (2015). Forecasting natural gas prices using wavelets, time series, and artificial neural networks. PLoS ONE, 10(11). https://doi.org/10.1371/journal.pone.0142064

Kilian L. (2009). Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. American Economic Review, 99, 1053–1069. https://doi.org/10.1257/aer.99.3.1053

Kilian, L., Murphy, D. P. (2014). The Role of inventories and speculative trading in the global market for crude oil. Journal of Applied Econometrics, 29(3), 454–478. https://doi.org/10.1002/jae.2322

Kilian, L., Zhou, X. (2018). Modeling fluctuations in the global demand for commodities. Journal of International Money and Finance, 88, 54–78. https://doi.org/10.1016/j.jimonfin.2018.07.001

Kriechbaumer, T., Angus, A., Parsons, D., Rivas Casado, M. (2014). An improved wavelet-ARIMA approach for forecasting metal prices. Resources Policy, 39, 32–41. https://doi.org/10.1016/j.resourpol.2013.10.005

Li, D., Moghaddam, M. R., Monjezi, M., Armaghani, D. J., Mehrdanesh, A. (2020). Development of a group method of data handling technique to forecast iron ore price. Applied Sciences, 10(7), 2364. https://doi.org/10.3390/app10072364

Ma, Y., Zhen, W. (2020). Market fundamentals and iron ore spot prices. Economic Record, 96(315), 470–489. https://doi.org/10.1111/1475-4932.12564

Manescu, C., Van Robays., I. (2014). Forecasting the Brent oil price. Addressing time-variation in forecast performance. Working paper series, 1735. Frankfurt: European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1735.pdf

Mei, F., Chen, P. (2018). VAR analysis of the factors influencing overcapacity in the iron and steel industry. Journal of Scientific and Industrial Research, 77, 566–569. https://nopr.niscpr.res.in/bitstream/123456789/45115/1/JSIR%2077%2810%29%20566-569.pdf

Nademi, A., Nademi, Y. (2018). Forecasting crude oil prices by a semiparametric Markov switching model: OPEC, WTI, and Brent cases. Energy Economics, 74, 757–766. https://doi.org/10.1016/j.eneco.2018.06.020

Nazlioglu, S., Soytas, U. (2012). Oil price, agricultural commodity prices, and the dollar: a panel cointegration and causality analysis. Energy Economics, 34, 1098–1104. https://doi.org/10.1016/j.eneco.2011.09.008

Nick, S., Thoenes, S. (2014). What drives natural gas prices? – A structural VAR approach. Energy Economics, 45, 517-527. https://doi.org/10.1016/j.eneco.2014.08.010

Nixon, D., Smith, T. (2012). What can the oil futures curve tell us about the outlook for oil prices? Quarterly Bulletin, Q1. London: Bank of England. Retrieved from https://www.bankofengland.co.uk/-/media/boe/files/quarterly-bulletin/2012/what-can-the-oil-futures-curve-tell-us-about-the-outlook-for-oil-prices.pdf

Obadi, S. M., Korcek, M. (2020). Driving fundamentals of natural gas price in Europe. International Journal of Energy Economics and Policy, 10(6), 318–324. https://doi.org/10.32479/ijeep.10192

Ouliaris, S., Pagan, A. R., Restrepo, J. (2018). Quantitative Macroeconomic Modeling with Structural Vector Autoregressions – An EViews Implementation. https://www.eviews.com/StructVAR/structvar.pdf

Pincheira, P., Hardy, N. (2019). Forecasting aluminum prices with commodity currencies. MPRA Paper, 97005. https://mpra.ub.uni-muenchen.de/97005/

Reichsfeld, D. A., Roache, S. K. (2011). Do commodity futures help forecast spot prices? IMF Working Paper, 11/254. Frankfurt: European Central Bank. Retrieved from https://doi.org/10.5089/9781463923891.001

Sharma, V., Cali, U., Sardana, B., Kuzlu, M., Banga, D., Pipattanasomporn, M. (2021). Data-driven short-term natural gas demand forecasting with machine learning techniques. Journal of Petroleum Science and Engineering, 206, 108979. https://doi.org/10.1016/j.petrol.2021.108979

Tularam, G. A., Saeed, T. (2016). Oil-price forecasting based on various univariate time-series models. American Journal of Operations Research, 6(3), 226–235. http://dx.doi.org/10.4236/ajor.2016.63023

Wårell, L. (2014). The effect of a change in pricing regime on iron ore prices. Resources Policy, 41, 16–22. https://doi.org/10.1016/j.resourpol.2014.02.002

Wårell, L. (2018). An analysis of iron ore prices during the latest commodity boom. Mineral Economics, 31, 203–216. https://doi.org/10.1007/s13563-018-0150-2

West, K. D., Wong, K.-F. (2014). A factor model for co-movements of commodity prices. Journal of International Money and Finance, 42, 289–309. http://doi.org/10.1016/j.jimonfin.2013.08.016

Zivot, E., Wang, J. (2003). Modeling Financial Time Series with S-PLUS. New York: Springer-Verlag.

APPENDIX A. TABLES

Table 5. Augmented Dickey-Fuller (ADF) Test Results

Variable	tau-stat (p-value)
Brent real price 1	-2.840 (0.185)
Brent real price 1, dlog	-10.048*** (0.000)
Brent real price 2	-2.811 (0.195)
Brent real price 2, dlog	-10.061*** (0.000)
Petroleum inventories, dlog	-3.374*** (0.001)
Change in petroleum inventories, % mom	-3.135*** (0.002)
Crude oil production, dlog	-13.359*** (0.000)
Change in crude oil production, % mom	-11.583*** (0.000)
Manufacturing PMI, Global	-4.490*** (0.002)
U.S. CPI_1 (s.a. index, 1982-1984=100)	0.723 (1.000)
U.S. CPI_1 (s.a. index, 1982-1984=100), dlog	-9.165*** (0.000)
U.S. CPI_2 (n.s.a. index, 2010=100)	-1.329 (1.000)
U.S. CPI_2 (n.s.a. index, 2010=100), dlog	-10.061*** (0.000)
Kilian index (IGREA)	-2.819*** (0.010)
TTF natural gas, nominal price	-4.528*** (0.000)
TTF natural gas, nominal price, dlog	-3.000*** (0.003)
TTF natural gas, real price	-6.047*** (0.000)
TTF natural gas, real price, dlog	-2.887*** (0.004)
Gas stock changes, Netherlands	-4.597*** (0.001)
Supply of natural gas, Netherlands	-2.357 (0.401)
Supply of natural gas, Netherlands, first difference	-8.593*** (0.000)
Iron ore real price	-3.032 (0.126)
Iron ore real price, dlog	-9.827*** (0.000)
Steel nominal price	-4.445*** (0.003)

Table 5 (continued). Augmented Dickey-Fuller (ADF) Test Results

Variable	tau-stat (p-value)
Coal, Australian, nominal price	-4.309*** (0.004)
Change in China's crude steel production, % mom	-5.229*** (0.000)
Manufacturing PMI, China	-7.096*** (0.000)
Baltic Dry Index	-7.065*** (0.000)
Change in Baltic Dry Index, % mom	-13.306*** (0.000)
Change in Brazil's iron ore exports, winsorized, % mom	-15.659*** (0.000)

*** denote significance at the 1% level.

Visnyk of the National Bank of Ukraine, 2022, No. 254, pp. 29–42

CRYPTO CURRENCY PRICE FORECAST: NEURAL NETWORK PERSPECTIVES

YURIY KLEBAN[a], TETIANA STASIUK[a]

[a]National University of Ostroh Academy
E-mail: yuriy.kleban@oa.edu.ua
E-mail: tetiana.v.stasiuk@oa.edu.ua

Abstract This study examines the problem of modeling and forecasting the price dynamics of crypto currencies. We use machine-learning techniques to forecast the price of crypto currencies. The FB Prophet time-series model and the LSTM recurrent neural network were used to conduct the study. Using the example of data from Binance (the most popular exchange in Ukraine) for the period from 06.07.2020 to 01.04.2023, prices of Bitcoin, Ethereum, Ripple, and Dogecoin were modeled and forecasted. The recurrent neural network of long-term memory showed significantly better results in forecasting according to the RMSE, MAE, and MAPE criteria, compared to the results from the Naïve model, the traditional ARIMA model, and the FB Prophet.

JEL Codes C45, C53, G11

Keywords crypto currency, forecasting, time series, neural network

1. INTRODUCTION

The rapid development of digital currencies over the past decade is one of the most controversial and unpredictable innovations in the global economy to date. Significant fluctuations in the exchange rates of crypto currencies, the possibility of market perturbations due to false information and a lack of transparency, doubts over the legality of their use related to the anonymity of owners, as well as incomplete and contradictory legislation mean there are significant risks related to investing in crypto assets.

As for Ukraine, the discussion around the crypto currency market intensified with the adoption of the Law «On Virtual Assets» (2022) and the registration in November 2023 of draft Law No. 10225-1 on amendments to the Tax Code of Ukraine and other legislative acts of Ukraine regarding the regulation of the turnover of virtual assets in Ukraine (Verkhovna Rada of Ukraine, 2023). The most complex and most discussed issues of this draft law are related to tax conditions for individuals and businesses operating in the field of virtual assets (Malinovska, 2023). Crypto assets are becoming increasingly popular among economic agents as a form of investment asset (Corbet et al., 2019). Significant price volatility for crypto assets makes it necessary to develop and use models for forecasting prices effectively. The current economic literature actively applies various traditional statistical approaches and machine-learning methods (link) to assess the ability to forecast the prices of various kinds of digital currencies over a range of horizons.

In this study, we contribute to this large body of the literature, and analyze the effectiveness of machine-learning methods such as FB Prophet and LSTM in forecasting the price of crypto assets. The selected crypto currencies are Bitcoin and Ethereum, which have the largest capitalization, Ripple, a popular low-cost currency that is actively used by businesses around the world and has an affordable set of high-quality tools for managing financial resources, and Dogecoin, the so-called Meme-coin that applies to crypto assets, the value of those is determined mainly by community interests and online trends. Crypto currencies of a range of capitalizations and behaviors were chosen, as investors adjust their portfolio preferences depending on the market situation. It is also important to establish whether different forecasting methods should be used for crypto currencies of differing characteristics. The data is daily and was obtained from the service binance.com for the period 06.07.2020 to 01.04.2023.

The Naïve and ARIMA models demonstrated fairly high forecasting accuracy for crypto currencies with low prices that are affected most by market volatility (Ripple, Dogecoin). Prophet is best used for Bitcoin and Ethereum, the crypto currencies that effectively set the trends of the crypto currency market.

According to the results of the study, the recurrent neural network LSTM demonstrated the best forecast accuracy for all crypto currencies. Thus, the paper demonstrates that LSTM, despite the complexity of its use, is a powerful tool for modeling volatile and complex phenomena such as crypto currency prices.

This rest of this paper has the following structure: The second section provides a brief overview of the main characteristics of the crypto currency market. The third section reviews the literature that studies the problems of forecasting the prices of crypto assets. The fourth section describes the study methodology. The fifth section is devoted to the data used. The sixth section presents the key results of the study and evaluates the quality of the constructed models and forecasts, their description and quality analysis. The final section, conclusions, summarizes the results of this study.

2. THE NATURE OF THE CRYPTO CURRENCY MARKET

As in the case of traditional financial assets, the operation of the crypto currency market is based on the principle of a balance between supply and demand: when demand (supply) for crypto currencies increases (decreases), prices usually increase, and vice versa. Supply and demand, in turn, are influenced by various price factors (price stability of the market, the exchange rate price for Bitcoin), and non-price factors (crypto currency issues, news, legal restrictions).

The crypto currency market, where crypto currencies are bought and sold, is decentralized. While similar to a traditional market, unlike traditional ones the crypto currency market is available for trading 24/7. There are several specific features of the way the market functions:

- Decentralization. The crypto currency market is decentralized, meaning it does not have a central body that controls and regulates its operation. First, this technology does not have a central issuing authority, such as a central bank in the case of traditional currencies. Thus, the lack of centralized control contributes to greater autonomy and distribution of ownership. Second, all contracts and transactions made on the crypto currency network are reflected in a blockchain, which is a distributed database. This means that each node in the network has a copy of this database, so no centralized authority can manipulate or control these transactions. The third important aspect is the development of decentralized exchanges, where crypto assets are traded without the mediation of centralized structures. This is effected through smart contracts that ensure the execution of transactions directly on the blockchain, without the need for the involvement of a third party. Thus, decentralization in the context of crypto currencies and the blockchain technology allows for greater autonomy, security and transparency in financial and trading operations, avoiding centralized control and risks related to it. Crypto asset trading often takes place on centralized exchanges, where participants exchange crypto currencies under established rules, using infrastructure that is usually owned by a centralized company or organization. Despite the advantages of centralized exchanges, such as high liquidity and fast transaction execution, there are risks related to centralization, including the possibility of system hacking, fraud, and access restrictions. The disadvantages of centralized exchanges have led to the emergence of decentralized alternatives, where transactions are made directly between users using smart contracts on the blockchain.

- High volatility. Crypto currencies are known for their high volatility. This means that their prices can change very quickly depending on various factors, such as news, industry events, etc.

- Algorithmic trading. Algorithmic trading is widely used in the crypto currency market. This means that many users use special applications that analyze the market and automatically buy or sell crypto currencies depending on various factors.

- Risks. The crypto currency market carries certain risks, including those related to price volatility, possible cyber-attacks and security hacks, as well as legislative changes.

The level of legalization of virtual assets in different countries differs, because the lack of knowledge of the problem, the high risks of these assets and other factors related to the internal development of states do not allow the full implementation and use of crypto currencies as financial instruments. Legislation on virtual assets in various countries regulates the activities of crypto currency market participants. That means market regulators and other public authorities may exercise certain controls over the operations and activities of market participants that perform operations with virtual assets, in order to reduce asset volatility and reduce risk.

Most countries around the world have recognized crypto currencies as virtual assets and legalized them at the legislative level (Amase, 2023). In Ukraine, the legislative framework for regulating the turnover of virtual assets is still at the stage of formation and discussion by society.

The formation of legislation to regulate the crypto currency market is a reaction by countries to the spread and impact of blockchain technologies in the modern world. The main reasons include combating illegal activities, protecting investors and consumers, ensuring financial stability, implementing an effective tax policy, and so on.

3. LITERATURE REVIEW

The main limitation when forecasting prices for crypto assets is their high volatility and the difficulty of determining the main factors influencing the exchange rates of crypto currencies. Because of their high risk, investments and other operations in crypto currency require reasonable risk management and balanced management strategies, since quite often the future financial stability of the investor or shareholder depends on them.

Usually, forecasting the price of crypto currencies is considered as a time series problem (Persson, 2022). Using time series, the forecast of future values is based on previous observations for consecutive time intervals.

An important concept in time series analysis is stationarity, when dynamic series have the same behavior and the same statistical properties over a time period (Whittle, 1953). However, it is worth noting that crypto currency prices are non-stationary (Couts et al., 1966).

Pronchakov and Bugaienko (2019) compared several types of moving averages (simple, weighted, and exponential) for forecasting the prices of digital currencies during their study. They concluded that all moving averages had approximately the same trend, but the exponential model was closest to actual values and adapted faster to price changes.

A separate place in the studies is occupied by complex methods of forecast extrapolation, among the most common varieties of which are moving average and exponential smoothing methods (Pronchakov and Bugaienko, 2019; Pilipchenko et al., 2021; Gagnidze and Iavich, 2020). They are commonly used for noise smoothing, identifying fracture points, and short-term forecasting. For example, the intersection of moving averages is an important technical indicator according to the authors: when the moving average for a short period intersects with the moving average for a long period, this is a signal to buy or sell an asset.

Derbentsev et al. (2019) used time-series models, Binary Auto Regressive Tree (BART) and Autoregressive Integrated Moving Average (ARIMA), to build short-term forecast models for the crypto currencies with the highest market capitalization. The time periods contained different types of dynamics (stable, decline, growth, trend change). The results showed that the errors in the BART model were half those compared to the ARIMA model. However, the authors note that all models showed worse results during periods of sharp changes in trends.

The artificial intelligence sub-sector is also often used to forecast the price of crypto currencies: i.e. machine learning. It is based on the use of statistical methods by which a computer acquires the ability to "learn" from a data set.

Popular methods include Long Short-Term Memory and models derived from it (Livieris et al,, 2020; Luo et al., 2022; Ammer and Aldhyani, 2022) and Gated Recurrent Unit (Al-Nefaie and Aldhyani, 2022; Aljadani, 2022)

Aljadani (2022) used a Bidirectional LSTM and a Gated Recurrent Unit in his research. For the Bitcoin price time series, the highest value of the average absolute percentage error was 0.26% for the BiLSTM model and 0.22% for the GRU, given that for the Mean Absolute Percentage Error or MAPE indicator, a value of less than 10% means that the forecast model is considered to have a high level of accuracy.

In the work of Al-Nefaie and Aldhyani (2022) the use of deep learning to forecast the value of Bitcoin was studied, with the aim of helping investors make informed decisions and aiding authorities in evaluating crypto currencies. The authors used a GRU (Gated Recurrent Unit, a type of recurrent neural network (RNN) that was designed to work with sequential data such as text or time series) and an MLP (Multilayer Perceptron, a type of artificial neural network that consists of multiple layers of neurons, where each neuron in the previous layer is connected to each neuron in the next layer, creating a deep multi-layer architecture) models to analyze Bitcoin price time series between January 2021 and June 2022. Based on the results of the study, it was found that the MLP model achieved high regression efficiency with R2 = 99.15% at the training stage and R2 = 98.90% at the testing stage. The authors believe that these models may have a significant impact on portfolio management and optimization in the face of the unpredictability of the crypto currency market. However, it is necessary to keep in mind the need to use such models cautiously in conditions of high volatility in the crypto currency market, and take into account possible limitations and uncertainties when considering their results.

Research by Garlapati et al. (2022) was conducted into forecasting the value of Bitcoin using Facebook's Prophet and ARIMA. The authors compare the effectiveness of these two methods on the same data set covering the period from May 2016 to March 2018. To improve the accuracy of forecasting, control variables selected on the basis of correlation studies between crypto currencies and real currencies are added to the model. According to the test results, Prophet is more effective than ARIMA, demonstrating an R2 value of 0.94, compared to 0.68 for ARIMA.

A study by Cheng et al. (2024) uses empirical financial time series analysis and machine learning to forecast the price of Bitcoin using LSTM, SARIMA, and Facebook's Prophet models. The results show that LTSM has a marked improvement over SARIMA and Prophet in terms of MSE and MAE. Furthermore, the result confirmed that Bitcoin values are seasonally volatile and random, and are often influenced by external variables such as news, crypto currency laws, investments, or social media rumors.

The approaches used by researchers to building crypto currency price forecasts have shown that they may be quite effective in forecasting for the short term. However, there is a need to combine and compare these methods to ensure the stability and reliability of forecasts.

4. DATA

Four crypto currencies were selected for building models and making forecasts: Bitcoin (BTC) and Ethereum (ETH) having the highest capitalization, Ripple (XRP) is cheaper, but quite developed and popular among businesses, and the Dogecoin (DOGE) meme-coin. This was done in order to help investors choose the right asset to invest in, as most initial investors prefer both already well-capitalized and new, cheap crypto assets.

Historical data for the selected assets against the Tether (USDT) stable coin (which, in fact, is a virtual dollar) was taken from the most popular exchange in Ukraine – Binance, using the Application Programming Interface (API) of the exchange and the Python programming language. The paper uses 1,000 observations of the prices of the following trading pairs: BTC-USDT, ETH-USDT, XRP-USDT, DOGE-USDT. The data was obtained on a daily basis for the period 06.07.2020 to 01.04.2023. The data set consists of five characteristics: open – opening price, close – closing price, high – maximum price, low – minimum price, volume – amount of money in circulation. Descriptive statistics of the collected trading pairs are shown in Table 1 (Appendix A), and Figure 1 (Appendix B) illustrates the closing price of each of the crypto currencies for the selected period.

Before starting to model a time series, it is important to understand some of its basic properties. Understanding the properties of a time series may help determine which methods and models will be effective for forecasting. For further modeling, the closing price is selected as the main dependent variable.

The first step is to normalize the data. One of the most common methods of data normalization is min-max normalization: for each attribute, the smallest value is replaced with 0, the largest value is replaced with 1, and all other values are in the range from 0 to 1. The values are calculated using the following formula:

$$X_0 = \frac{X_i - \min(X_i)}{\max(X_i) - \min(X_i)}, \quad (1)$$

where $min(X_i)$ and $max(X_i)$ are the minimum and maximum values, 0 and 1, respectively.

An additional step in data preprocessing was checking for anomalies and outliers. In order to identify values that have statistical differences, we use an uncontrolled learning technique, namely neural networks (NNs). We used LSTM RNN and autoencoders to build a model of unsupervised learning.

Figure 2 (Appendix B) shows the distribution of the Mean Absolute Error (MAE) in the training and test datasets. In the training set, values greater than 0.2 are seen as unusual. This value was thus set as a threshold for outliers, i.e. values exceeding it will be outliers.

Figure 3 (Appendix B), we can see that there are abnormal values in the Bitcoin training sample. A total of 10 anomalies were detected in the training sample, while there were zero in the test sample. Similar steps have been applied to other crypto currencies. For clarity, the outliers were shown on graphs (Figure 4 in Appendix B).

Thus, Ethereum has 22 outliers, Ripple – 25, and Dogecoin – 20. No outliers were recorded in any of the test samples. For more accurate forecast models, anomalous values were smoothed out using a simple moving average (SMA).

Then, we check the time series for stationarity using the Augmented Dickey-Fuller Test (ADF), since for certain modeling methods, the stationarity of the data is a necessary condition. Table 2 shows the test results.

From the Table above, we can see that the selected price series for crypto currencies are characterized by non-stationarity. The data is reduced to a stationary form using a difference operation with different integration coefficients. The results of the data stationarity test after integration are shown in Table 3.

5. METHODOLOGY

To study, model, and forecast time series (for prices), we start with a naive method and use its results as a baseline for subsequent models, as this is primitive and the goal is to improve the accuracy of forecasts for other models. We make a naive forecast based on the last value in the data set, that is, we assume that the price of crypto currency tomorrow will be the same as today. The obtained forecasts based on the naive method shall be further compared with the results of more complex models in order to evaluate their effectiveness.

At the next stage, ARIMA models were built (Box et al., 2015) for the stationary time series of the prices for each of the crypto currencies. ARIMA is used to analyze and forecast series that may have a trend and/or seasonality. The main stages of build-up include data standardization, ensuring the stationarity of the time series. Then the parameters of the ARIMA model are determined, such as the autoregression order (p), the integrability order (d), and the moving average order (q). This is done on the basis of an analysis of autocorrelation and partial autocorrelation graphs. After that, the best model is selected and parameters are selected using criteria such as AIC or BIC. After the selection of the parameters, an ARIMA model is built and evaluated on the training dataset. Then, diagnostics and the verification of residual errors are performed to determine their randomness and the possibility of improving the model. Finally, the constructed model is used to forecast future values of the time series. This process may require iterations and parameter adjustments to achieve an optimal model for certain time series.

After building and evaluating ARIMA models, a forecast based on stationary and non-stationary data was made for the Facebook Prophet model, which is a both simple and powerful tool from Meta (Taylor and Letham, 2017). It can be used to forecast a time series without the user having to be an expert in data analysis. It was created because working with neural networks often used in forecasting, and this is quite difficult without having proper knowledge about the architecture of these networks.

Prophet's main algorithm is a generalized additive model, which can be decomposed into three main components: trend, seasonality, and public holidays. As mentioned above, seasonality and trend are two important but difficult-to-quantify components of a time series analysis, but Prophet takes both into account perfectly.

Because the model can be decomposed, it is relatively easy to obtain model coefficients to understand the impact of seasonality, trends, holidays, and other variables. For example, by forecasting the price of a crypto currency, we can obtain a demand ratio to find out how much demand affects price changes.

Prophet also contains a built-in cross-validation function for measuring forecast error using historical data. This is done by selecting boundary points in the history, and for each of them, a model is built using data only up to this boundary point. Then we can compare the forecasted values with the actual ones.

Table 2. Checking Time Series for Stationarity

	Bitcoin	Ethereum	Ripple	Dogecoin
ADF	-1.57	-1.77	-2.31	-2.46
p-value	0.49	0.39	0.16	0.12
Stationary	–	–	–	–

Note: calculated by the authors based on the collected data.

Table 3. Reducing the Time Series to Stationary Ones

	Bitcoin	Ethereum	Ripple	Dogecoin
ADF	-31.05	-11.31	-6.69	-5.76
p-value	0.00	0.00	0.00	0.00
Stationary	+	+	+	+

Note: calculated by the authors based on the collected data.

It is worth noting that Prophet is great for stationary data, i.e. time series that have the same behavior and the same statistical properties over a time period (Sivaramakrishnan et al., 2021). Prophet may also be used for non-stationary time series, but it does not always allow accurate forecasts to be made (Vasselin and Bertrand, 2021). In our study, the accuracy of the Prophet model was quite low for both stationary and non-stationary data.

Deep learning methods are also used to study time series. They are used to create multi-layered neural networks that ensure the high accuracy of results.

Recurrent Neural Networks (RNNs) are a popular class of neural networks that allows cyclic connections between nodes to be created. That is, the output of a node can affect subsequent inputs to the same node (Figure 5 in Appendix B). This allows them to use their internal state to process sequences of input data. By the same principle, a person reading a book understands each word by relying on knowledge gained earlier.

Nodes at different levels of the neural network are compressed and form a single layer of the recurrent neural network. At any given time t, the input data is a combination of the input data x<t> and x<t-1>, and the output data is returned to the network to improve the results.

Standard RNNs have some problems when exploring long-term dependencies and remembering them for a long time. These problems do not occur when using a Long Short-Term Memory (LSTM) – this is their typical behavior.

All RNNs have the same shape, i.e. a chain of repeating neural network modules. Regular models will have a simple structure with a single tanh (hyperbolic tangent) level (Feng et al., 2019), the LSTM, in turn, will have four interacting levels with unusual coupling (Figure 6 in Appendix B). Technical independent variables (lag values of the opening price, maximum and minimum price, and transaction volumes) are also used to build the LSTM model, which improves the forecasting ability of the model. At the same time, macroeconomic variables were not used as variables due to their low frequency and, consequently, their irrelevance in forecasting daily data.

To avoid overtraining with the LSTM network, the training sample was pre-tested through cross-validation.

The training sample was divided into five parts (as shown in Figure 7 in Appendix B), where green indicates the elements used for training the model, and blue indicates the elements used for testing. The metric for evaluating the model was set to Mean Absolute Percentage Error (MAPE).

For example, for the Bitcoin price, the mean error for the five evaluation stages was 2.80%, with a standard deviation of 0.82%. All models gave approximately the same results, which indicates the high stability of the model.

It is worth noting that models based on non-stationary data (Facebook's Prophet, LSTM) forecast the closing price of a candlestick (from a Candlestick Chart visual representation of trading data) for the day, while models based on stationary data (Naïve, ARIMA) forecast price movement between two candlesticks. Because of this, there are problems when converting results to their previous form as closing prices. In other words, models based on stationary series showed more accurate results for the average daily price, and forecasted changes in the movement of candlesticks, but these models could not forecast the closing price, so non-stationary data was used to compare the forecast strength of Prophet and LSTM, which does not violate the requirements for these models.

There are quite a few metrics for estimating the accuracy of mathematical models (Plevris et al., 2022). It is on the basis of such metrics that the best models are selected and their further use or implementation in production processes. Most of them may be divided into two categories based on the types of forecasts in ML models: classification and regression. Since the problem we are considering is a regression, we use popular indicators to evaluate forecast models, namely: RMSE, MAE, and MAPE.

MAE and RMSE are used together to diagnose changes in errors in the forecast set. RMSE will always be greater than or equal to MAE: the greater the difference between them, the greater the variance of the errors in the sample. RMSE and MAE measure the error in units of measurement of forecasted variables, while MAPE displays the result as a percentage, determines the accuracy of the forecast model, and allows it to be determined how accurate the forecasted values were compared to the actual ones.

In general, the machine-learning methods chosen have certain advantages over other models. They have the ability to analyze large amounts of data and make forecasts based on them. The algorithms built into them can significantly improve the accuracy of forecasting of values, and identify complex dependencies and their degree of impact on the resulting variable – in our case the closing price.

6. STUDY RESULTS

During modeling, the collected time series were divided into training and test samples in a 90/10 ratio (900 observations for training models and 100 for test ones). The large amount of training data gives the selected models a wider horizon for studying time-series patterns. Furthermore, given that cross-validation is applied to the selected methods and cross-validation of models occurs during their training, the error is averaged after each iteration of the training, and we obtain a more reliable estimate of the error and the accuracy of the model, respectively. It is worth noting that only the data of the training sample is involved in training the model using cross-validation, while the test sample is not accepted in this process and is used at the stage of checking the model for quality.

A basic forecast for the future was made with the assumption that the price of crypto currencies tomorrow will not differ from yesterday (Figure 8 in Appendix B). Figure 9 (Appendix B) shows the results of forecasting the price of crypto assets based on ARIMA.

The results of Prophet forecasting are shown below. The blue curve shows the real price, and the red curve shows the simulated price. The light-blue field is the 95% confidence interval. Figure 10 shows that Prophet performed well when forecasting with only historical crypto currency data.

During modeling, the settings of the model parameters were changed for each of the different crypto currencies. For example, for Ethereum, the influence of the trend and the number of points of change were increased. As a result, MAPE was significantly reduced, and a rather accurate forecast was obtained – although confidence boundaries widened. For Dogecoin, in contrast, the influence of the trend was weakened. As a result, the forecast gives good results, taking into account the abnormal values.

Increasing the change points makes the model more adaptive to dynamics – especially if there are significant price changes in different periods. The parameter that determines the impact of a trend characterizes how much the model will pay attention to possible changes in the trend and seasonality. In other words, this parameter is responsible for regulating the flexibility in finding change points in trend and seasonality.

Let's look at the results of long-term memory network modeling in the test sample (Figure 11 in Appendix B). The information is displayed only for the test sample, because due to the low error in the training sample, the lines are superimposed, which makes it difficult to visually analyze the series.

The errors of the LSTM model during cross-validation on the training and test data are almost identical, which indicates the stability of the model on the sets that were used in its building, and on the new data for the forecast. Although it is quite easy to visually trace the convergence of actual and modeled values, the graph still shows some lag between the real values of the crypto currency price and the simulated ones.

Let's compare the built models in more detail using different metrics. Tables 4 to 7 below show the relative percentage deviation of the RMSE, MAE, and MAPE values of the ARIMA, Facebook's Prophet, and LSTM models from the Naïve model values. The least successful model compared to the Naïve model is highlighted in red, and the most successful model is highlighted in green.

For the Bitcoin crypto currency, the forecasts based on the Naïve model and ARIMA had relatively similar results and had satisfactory accuracy indicators. LSTM proved to be the best model for forecasting, showing the lowest errors and the highest accuracy among the methods studied.

For Ethereum, Naïve and ARIMA showed better results than for Bitcoin, but as in the previous example, LSTM has significantly lower errors and higher accuracy in its forecast values.

The Ripple crypto currency has slightly different results. The worst model was Prophet, and the best one was LSTM. An interesting observation is how well the Naïve and ARIMA models performed. Both models have high accuracy rates, and their forecast errors are 2–3 times lower than for Prophet.

Table 4. Choosing the Forecast Method for Bitcoin

	RMSE	MAE	MAPE
ARIMA(1,2,1)	-13.2	-13.4	-13.4
FB Prophet	-48.8	-54.6	-48.8
LSTM	-94.1	-95.0	-94.7

Note: calculated by the authors based on the results of forecasting, the percentage deviation of model metric values from the Naïve model is shown.

Table 5. Choosing the Forecast Method for Ethereum

	RMSE	MAE	MAPE
ARIMA(2,2,3)	-15.0	-14.9	-15.0
FB Prophet	-36.3	-44.3	-43.0
LSTM	-85.0	-84.8	-83.5

Note: calculated by the author based on the results of forecasting, the percentage deviation of model metric values from the Naïve model is shown.

Table 6. Choosing the Forecast Method for Ripple

	RMSE	MAE	MAPE
ARIMA(3,2,3)	-16.7	-25.0	-20.7
FB Prophet	166.7	250.0	236.0
LSTM	-81.7	-75.5	-77.0

Note: calculated by the author based on the results of forecasting, the percentage deviation of model metric values from the Naïve model is shown.

Table 7. Choosing the Forecast Method for Dogecoin

	RMSE	MAE	MAPE
ARIMA (2,2,3)	-11.1	-14.3	-2.8
FB Prophet	33.3	57.1	57.4
LSTM	-64.4	-58.6	-57.5

Note: calculated by the authors based on the results of forecasting, the percentage deviation of model metric values from the Naïve model is shown.

Dogecoin has similar results to the Ripple currency. As in the case of Dogecoin, Prophet did the worst job with forecasting, and LSTM did the best.

When building models, various parameters and data sets were used (without and with smoothed anomalies), it turned out that for all crypto currencies forecasts based on smoothed data had a higher error.

From Tables 4 to 7, we may see that LSTM demonstrated the best results among the studied models. In half of the cases, Prophet showed the worst forecast results, while MAPE for Naïve and ARIMA did not exceed 11%.

7. CONCLUSIONS

This study examines the problems of forecasting the price dynamics of crypto currencies. The spread and impact of crypto currency technologies in the modern world is causing heated discussions about the place and role of crypto currencies in the modern economy. Research into methods of forecasting crypto currency prices is of great importance for the scientific community, financial analysts, investors, and traders.

In the course of the study, data was collected, cleaned, normalized, and selected as the resulting basis for forecasting the closing price of crypto currencies. When studying the time series data, it turned out that it is non-stationary, which limits the range of possible approaches for modeling. In order to use ARIMA, the data was transformed into a stationary time series, but experiments have shown that this, firstly, complicates the process of calculating the resulting variable closing price, and secondly, does not improve the accuracy of models. The final mathematical models selected as the best ones are built using machine-learning techniques and using non-stationary time-series prices.

The recurrent neural network of long-term memory showed significantly better results in forecasting, according to the calculated errors, compared to the naive forecast, and for all ARIMA models, as well as the results of Facebook's Prophet. It is worth noting that in half of the cases, even Naïve and ARIMA showed more accurate results than Prophet.

Modeling and forecasting the price of crypto currencies is a quite promising and still under examined area of scientific study. To improve the process of forecasting crypto assets in the future, it is necessary to take into account fundamental factors (news, events in the field of technology, regulation, etc.), and to study relationships with other financial markets and economic trends.

Mathematical models for forecasting crypto currency prices have already become the foundation for developing trading algorithms and bots based on them, portfolio management tools, and for budget planning. With the development of modeling methods themselves and the growth of computing power, the accuracy of forecasting in the short term, even in very volatile markets, will increase.

REFERENCES

Aljadani, A. (2022). DLCP^{2}F: a DL-based cryptocurrency price prediction framework. Discover Artificial Intelligence, 2(20). https://doi.org/10.1007/s44163-022-00036-2

Al-Nefaie, A. H., Aldhyani, T. H. H. (2022). Bitcoin price forecasting and trading: data analytics approaches. Electronics, 11(24), 4088, 2–18. https://doi.org/10.3390/electronics11244088

Amase, W. (2023). Countries where cryptocurrency is legal vs illegal. CoinGecko, 11 December 2023. Retrieved from https://www.coingecko.com/research/publications/crypto-legal-countries

Ammer, M. A., Aldhyani, T. H. H. (2022). Deep learning algorithms to predict cryptocurrency fluctuation prices: increasing investment awareness. Electronics, 11(15), 2349. https://doi.org/10.3390/electronics11152349

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control, 5th ed. New Jersey: Wiley.

Cheng, J., Tiwari, S., Khaled, D., Mahendru, M., Shahzad, U. (2024) Forecasting Bitcoin prices using artificial intelligence: combination of ML, SARIMA, and Facebook Prophet models. Technological Forecasting and Social Change, 198, 122938. https://doi.org/10.1016/j.techfore.2023.122938

Corbet, S., Lucey, B., Urquhart, A., Yarovays, L. (2019). Cryptocurrencies as a financial asset: A systematic analysis. International Review of Financial Analysis, 62, 182–199. https://doi.org/10.1016/j.irfa.2018.09.003

Couts, D., Grether, D., Nerlove, M. (1966). Forecasting non-stationary economic time series. Management Science, 13(1), 1–21. http://www.jstor.org/stable/2627909

Derbentsev, V., Datsenko, N., Stepanenko, O., Bezkorovainyi, V. (2019). Forecasting cryptocurrency prices time series using machine learning approach. SHS Web of Conferences, 65, 02001. https://doi.org/10.1051/shsconf/20196502001

Feng, J., He, X., Teng, Q., Ren, C., Chen, H., Li, Y. (2019). Reconstruction of porous media from extremely limited information using conditional generative adversarial networks. Physical Review E, 100, 033308. https://doi.org/10.1103/PhysRevE.100.033308

Persson, E. (2022). Forecasting Efficiency in Cryptocurrency. Markets A Machine Learning Case Study. Retrieved from https://www.diva-portal.org/smash/get/diva2:1711510/FULLTEXT01.pdf

Gagnidze, A., Iavich, M. (2021). Time-series forecasting of Bitcoin prices. East European University Collection of Scientific Papers, 2(2020), 46–53. https://orcid.org/0000-0003-4728-3709

Garlapati, A., Krishna, D. R., Garlapati, K., m. Srikara Yaswanth, N., Rahul, U., Narayanan, G. (2021). Stock price prediction using Facebook Prophet and Arima models. 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2021. https://doi.org/10.1109/I2CT51068.2021.9418057

Lindemann, B., Maschler, B., Sahlab, N., Weyrich, M (2021). A survey on anomaly detection for technical systems using LSTM networks. Computers in Industry, 131, 103498. https://doi.org/10.1016/j.compind.2021.103498

Livieris, I. E., Pintelas, E., Stavroyiannis, S., Pintelas, P. (2020). Ensemble deep learning models for forecasting cryptocurrency time-series. Algorithms, 13, 121. https://doi.org/10.3390/a13050121

Luo, C., Pan, L., Chen, B., Xu, H. (2022) Bitcoin price forecasting: an integrated approach using hybrid LSTM-ELS models. Mathematical Problems in Engineering, 2022, 2126518. https://doi.org/10.1155/2022/2126518

Malinovska, A. (2023). Crypto currency tax: percentage, who should pay and features (In Ukrainian). Fakty, 1 December 2023. Retrieved from https://fakty.com.ua/ua/ukraine/ekonomika/20231201-podatok-na-kryptovalyutu-vidsotok-hto-maye-splachuvaty-ta-osoblyvosti/

Pilipchenko, A., Kuzminsky, V., Chumachenko, O. (2021). Using methods of technical analysis to forecast the cryptocurrency market. Včenì zapiski unìversitetu “KROK”, 4(64), 28–35. https://doi.org/10.31732/2663-2209-2021-64-28-35

Plevris, V., Solorzano, G., Bakas, N., Ben Seghier, M. (2022). Investigation of performance metrics in regression analysis and machine learning-based prediction models. ECCOMAS Congress 2022, 5, 9 June 2022. Oslo. https://doi.org/10.23967/eccomas.2022.155

Pronchakov, Y., Bugaienko, O. (2019). Methods of forecasting the prices of cryptocurrency on the financial markets. Technology Transfer: Innovative Solutions In Social Sciences and Humanities, 13–16. https://doi.org/10.21303/2613-5647.2019.00927

Sherstinsky, A. (2020). Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. Physica D: Nonlinear Phenomena, 404, 132306. https://doi.org/10.1016/j.physd.2019.132306

Sivaramakrishnan, S., Fernandez, T. R., Babukarthikm R. G., Premalatha, S. (2021). 4 forecasting time series data using ARIMA and Facebook Prophet models. *In* Big Data Management in Sensing: Applications in AI and IoT, 47–60. River Publishers. Retrieved from http://ieeexplore.ieee.org/document/9606779

Taylor, S. J., Letham, B. (2017). Forecasting at scale. PeerJ Preprints, 5, e3190v2. https://doi.org/10.7287/peerj.preprints.3190v2

Vasselin, H., Bertrand, J. (2021). Is Facebook Prophet suited for doing good predictions in a real-world project? Artefact. Retrieved from https://www.artefact.com/blog/is-facebook-prophet-suited-for-doing-good-predictions-in-a-real-world-project/

Verkhovna Rada of Ukraine (2023). The Draft Law of Ukraine on Amendments to the Tax Code of Ukraine and Other Legislative Acts of Ukraine Concerning the Regulation of Virtual Asset Turnover in Ukraine No. 10225-1 dated 17.11.2023 (In Ukrainian). Retrieved from https://itd.rada.gov.ua/billInfo/Bills/Card/43232

Verkhovna Rada of Ukraine (2024). Law of Ukraine “On Virtual Assets” No 2074-IX Edition of 01.01.2024 (did not enter into force) (In Ukrainian). Retrieved from https://zakon.rada.gov.ua/laws/show/2074-20

Whittle, P. (1953). Estimation and information in stationary time series. Arkiv för Matematik, 2, 423–434. https://doi.org/10.1007/BF02590998

APPENDIX A. TABLES

Table 1. Descriptive Statistics of Collected Trading Pairs

Dataset	Column	count	mean	std	min	25%	50%	75%	max
Bitcoin USD (BTC-USDT)	Open	1000	32,116.56	15,386.39	9,069.41	19,307.39	30,306.58	44,405.35	67,525.82
	High	1000	32,990.02	15,828.64	9,145.24	19,632.81	31,394.45	45,804.72	69,000.00
	Low	1000	31,141.11	14,848.86	8,893.03	18,908.93	29,288.29	43,017.27	66,222.40
	Close	1000	32,136.01	15,369.77	9,069.41	19,314.61	30,306.59	44,404.55	67,525.83
	Volume	1000	116,894.17	109,911.34	15,805.45	47,413.36	73,239.47	146,121.37	760,705.36
Ethereum USD (ETH-USDT)	Open	1000	1,956.26	1,136.44	227.54	1,217.71	1,713.87	2,813.22	4,807.98
	High	1000	2,022.10	1,170.42	229.85	1,259.99	1,777.80	2,946.53	4,868.00
	Low	1000	1,881.64	1,094.06	223.05	1,185.72	1,659.09	2,720.82	4,713.89
	Close	1000	1,957.86	1,135.16	227.56	1,218.31	1,715.80	2,813.22	4,807.98
	Volume	1000	795,863.90	526,527.70	117,762.10	448,910.60	656,849.10	970,714.00	4,309,836.00
Ripple USD (XRP-USDT)	Open	1000	0.59	0.33	0.18	0.35	0.47	0.80	1.83
	High	1000	0.62	0.35	0.18	0.36	0.49	0.83	1.97
	Low	1000	0.57	0.31	0.17	0.34	0.45	0.77	1.65
	Close	1000	0.59	0.33	0.18	0.35	0.47	0.80	1.84
	Volume	1000	55,8190,600.00	627,478,500.00	59,622,710.00	242,894,300.00	365,348,400.00	612,346,800,00	8,608,358,000.00
Dogecoin USD (DOGE-USDT)	Open	1000	0.12	0.11	0.002	0.06	0.08	0.17	0.69
	High	1000	0.13	0.13	0.002	0.06	0.09	0.18	0.74
	Low	1000	0.11	0.10	0.002	0.05	0.08	0.17	0.60
	Close	1000	0.12	0.11	0.002	0.06	0.08	0.17	0.69
	Volume	1000	2,690,987,000.00	6,872,100,000.00	88,706,470.00	653,910,800.00	1,082,091,000.00	2,006,851,000.00	109,073,700,000.00

Note: calculated by the author based on the data.

APPENDIX B. FIGURES

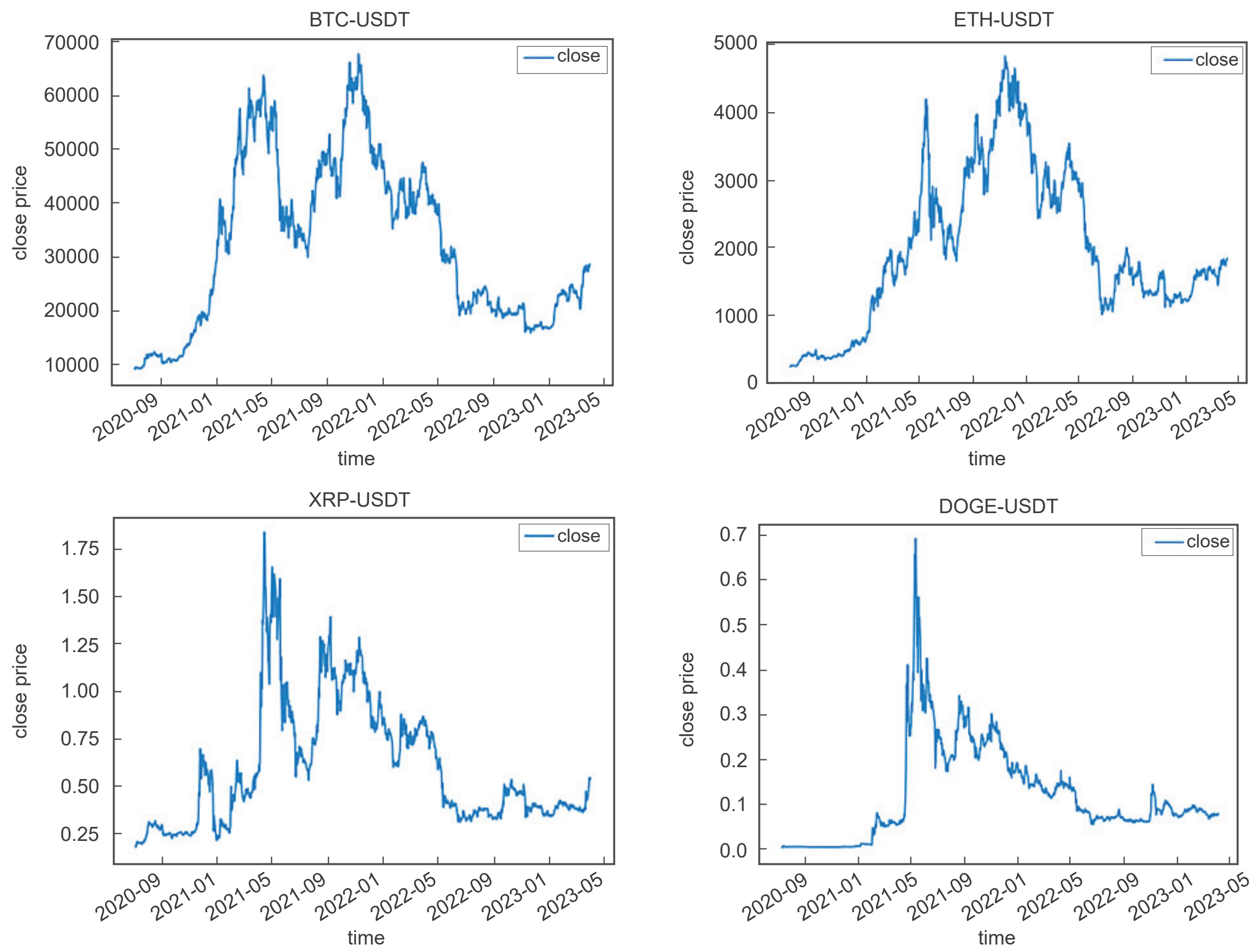


Figure 1. Dynamics of World Prices of Crypto Currencies for 06.07.2020 to 01.04.2023

Note: built by the author based on the collected data.

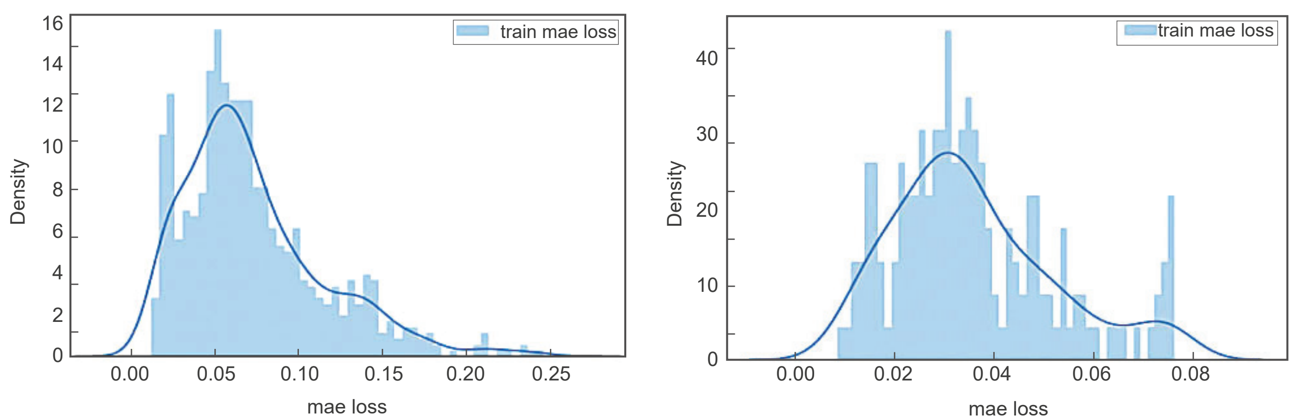


Figure 2. Distribution of the Mean Absolute Error in Bitcoin Samples

Note: built by the author based on the calculated errors.

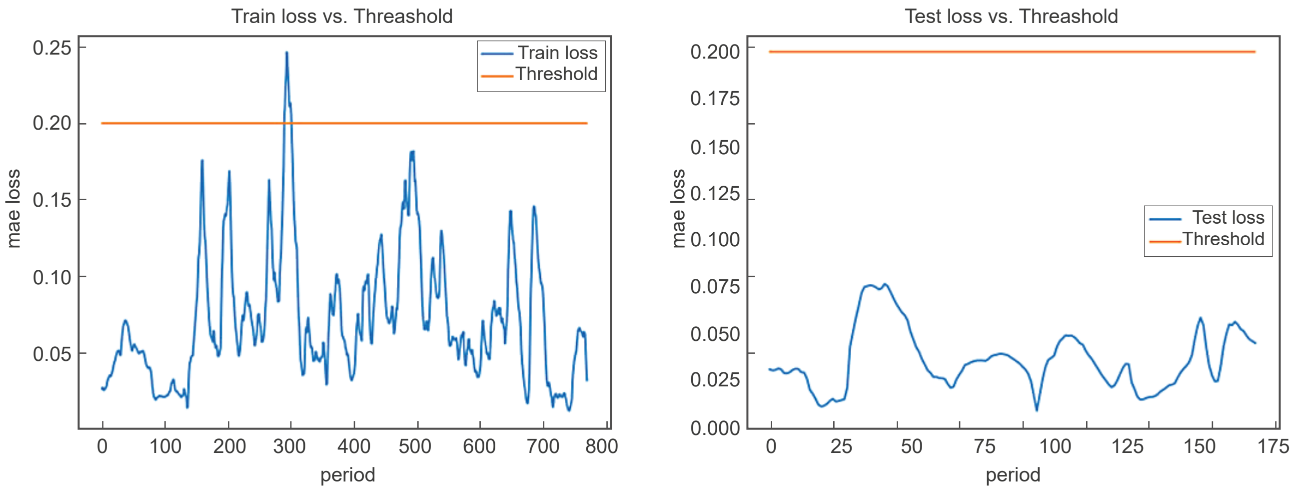


Figure 3. MAE and Bitcoin Sampling Anomaly Threshold
Note: built by the author basis on the calculated errors and the threshold value of the anomaly (Lindemann, 2021).

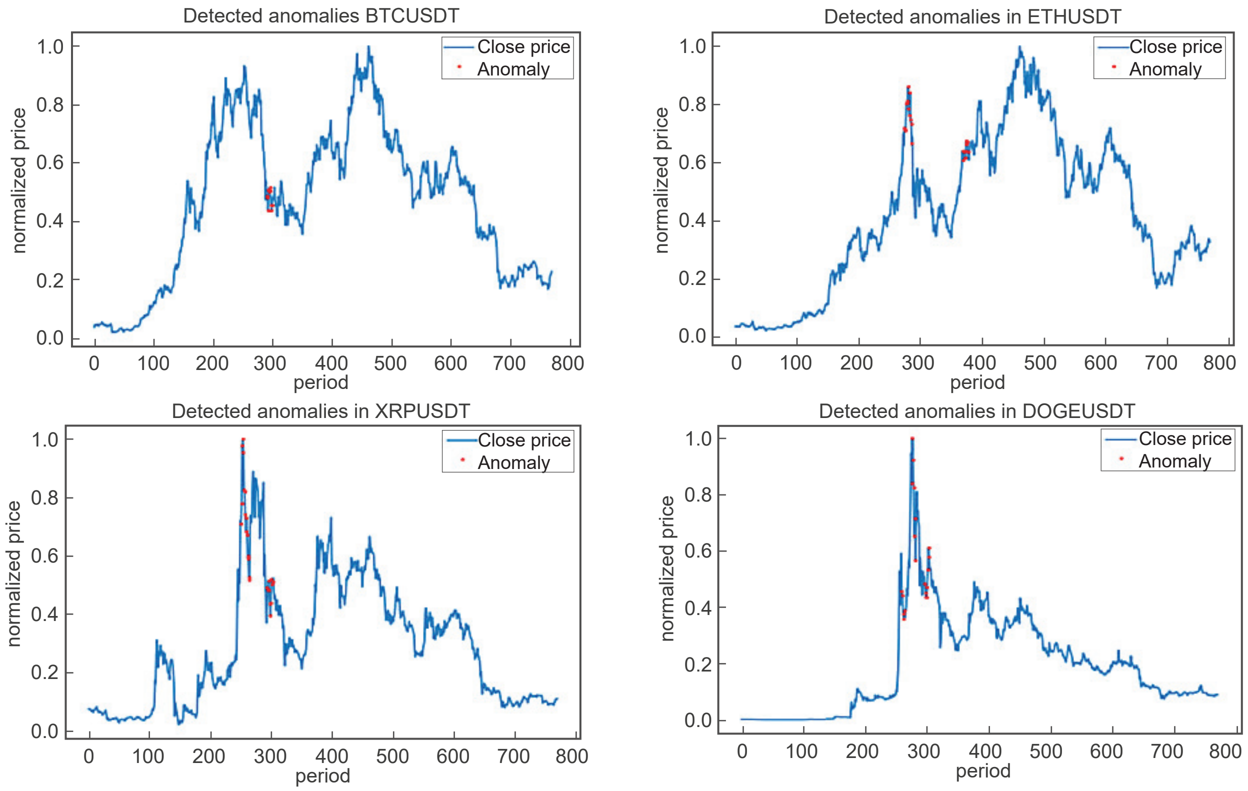


Figure 4. Abnormal Values in the Dynamics of Crypto Currency Prices
Note: built by the authors based on the detected anomalies.

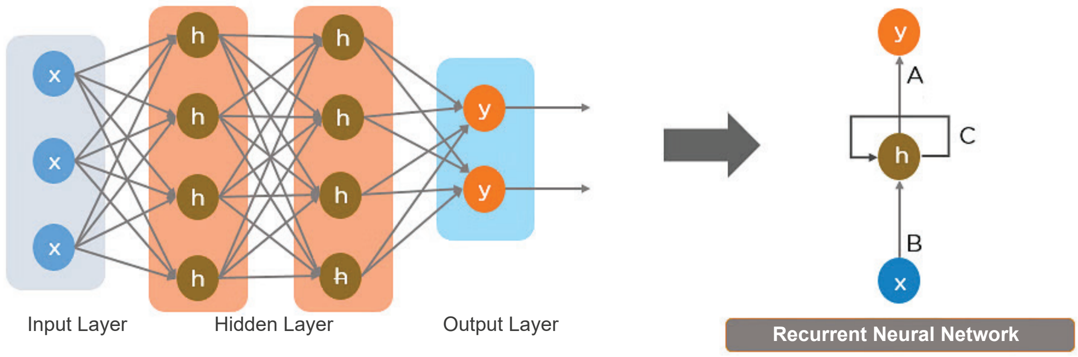


Figure 5. Structure of Simple RNN (Sherstinsky, 2020)

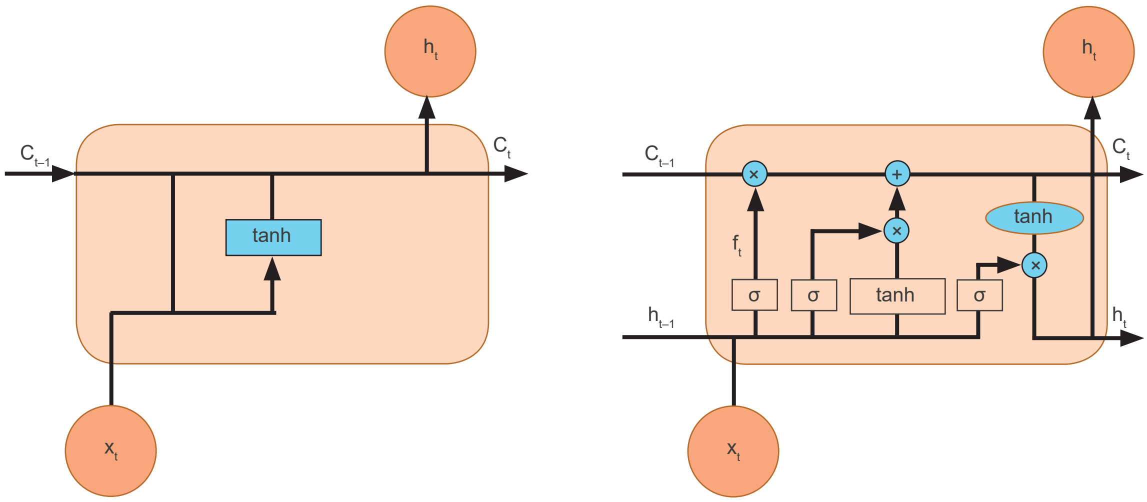


Figure 6. Module Structure for Regular RNN and LSTM Networks (Sherstinsky, 2020)

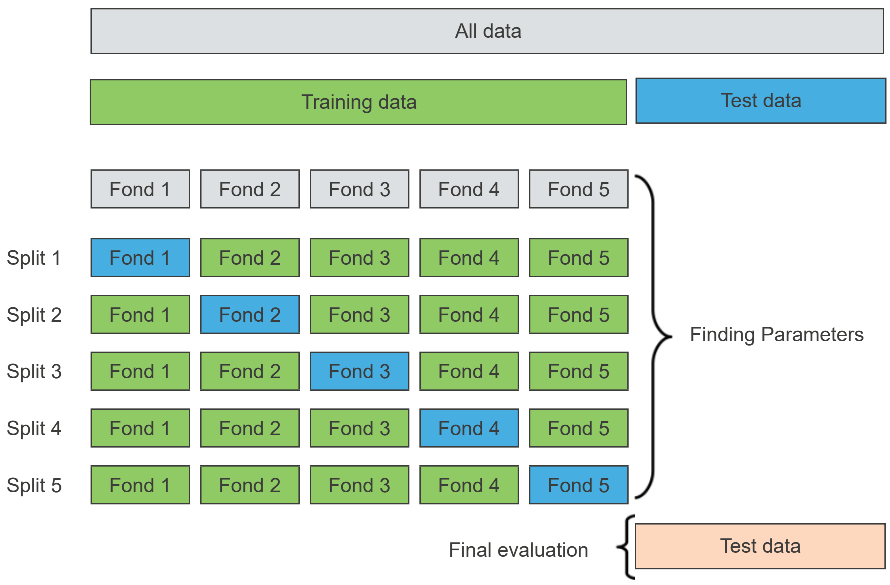


Figure 7. Cross-Validation Scheme *k-fold*[1]

[1] scikit-learn. Cross-validation: evaluating estimator performance https://scikit-learn.org/stable/modules/cross_validation.html#computing-cross-validated-metrics

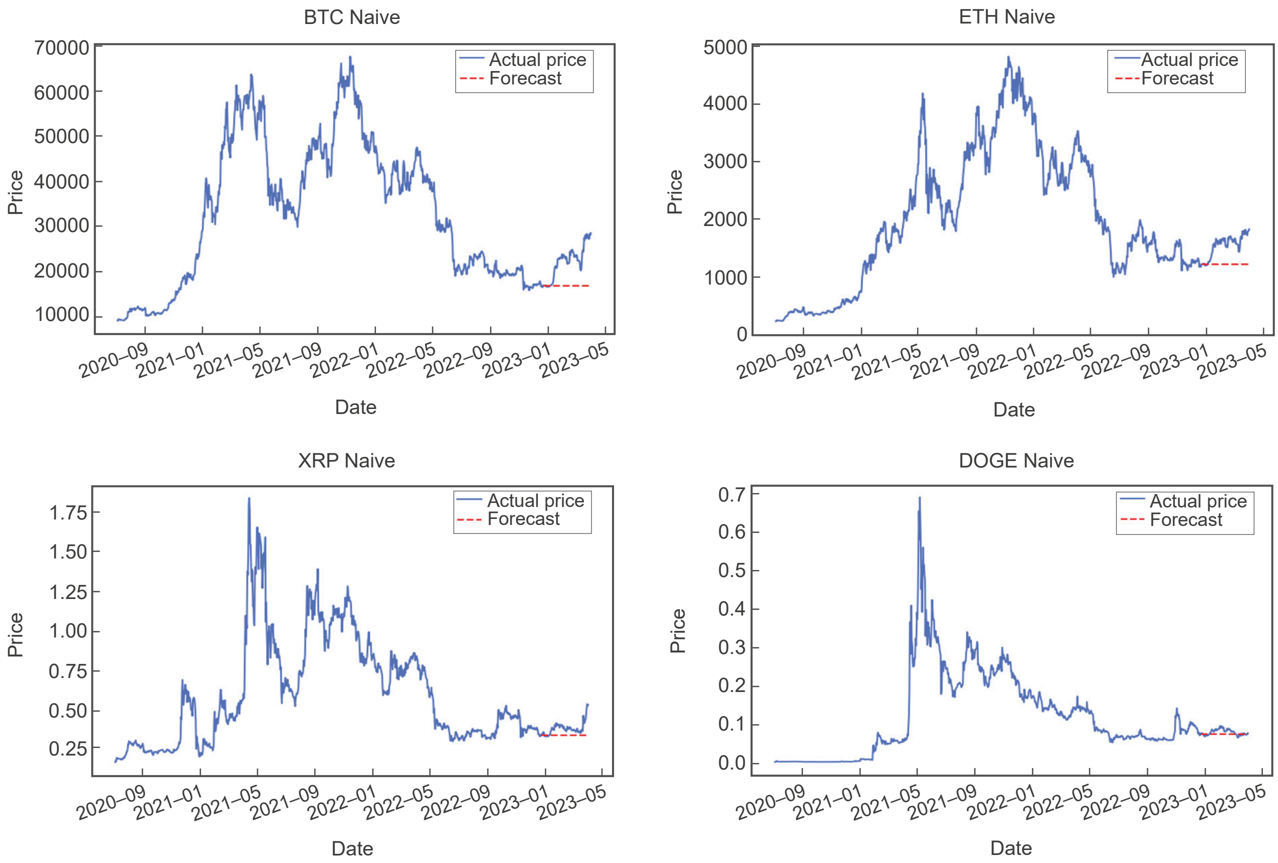


Figure 8. Naïve Forecast Models for Crypto Currencies for 06.07.2020 to 01.04.2023

Note: built by the authors based on the results of modeling.

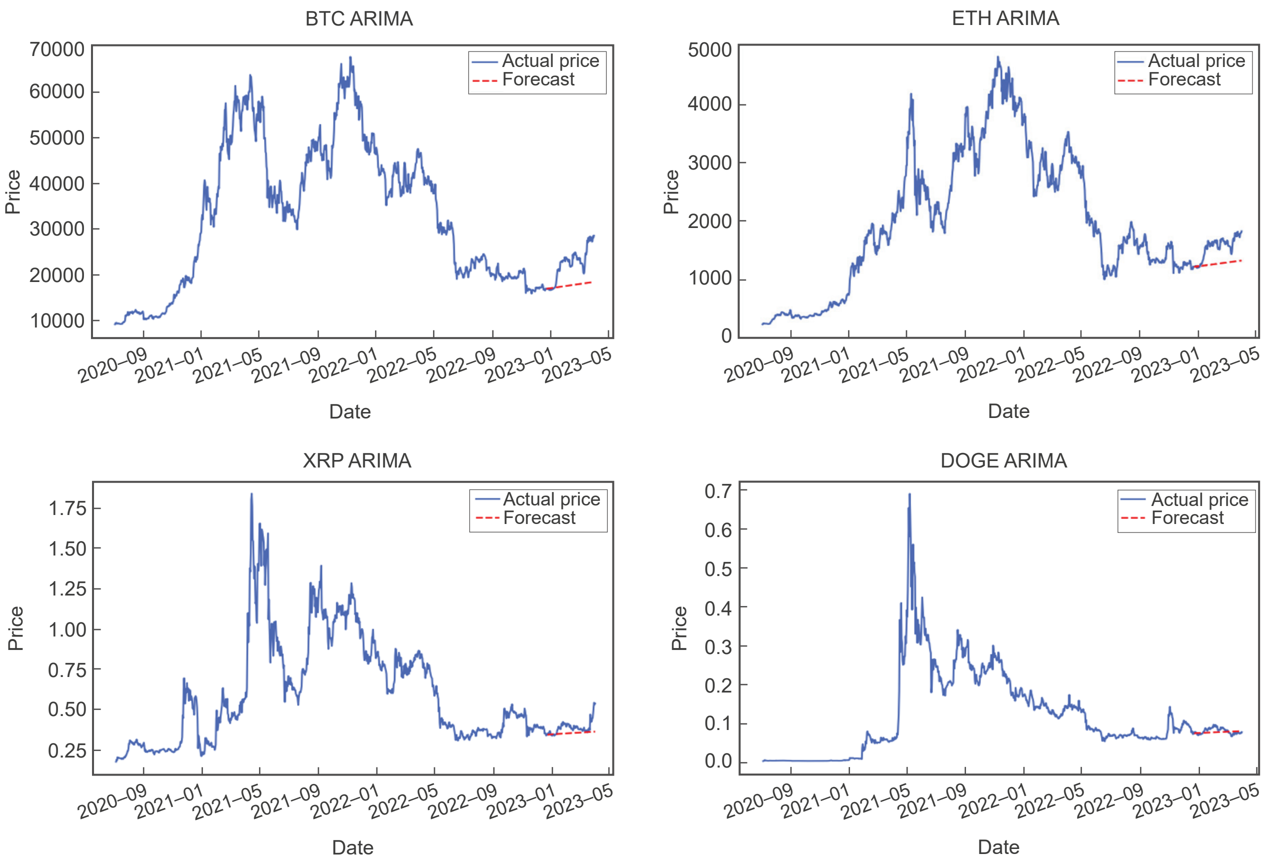


Figure 9. ARIMA Forecast Models for Crypto Currencies for 06.07.2020 to 01.04.2023

Note: built by the authors based on modeling results, ARIMA parameters are indicated in Tables 4 to 7.

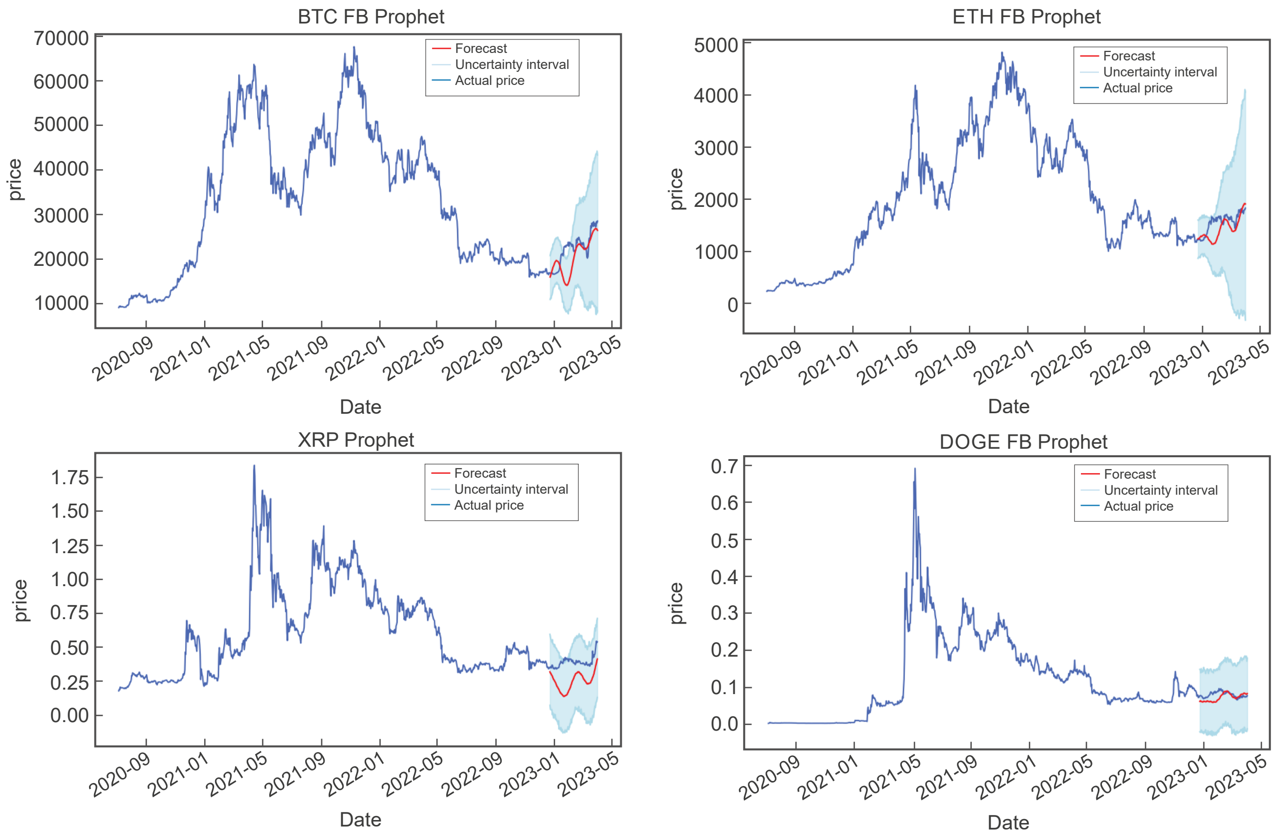


Figure 10. Prophet Forecast Models for Crypto Currencies for 06.07.2020 to 01.04.2023

Note: built by the author based on the results of modeling.

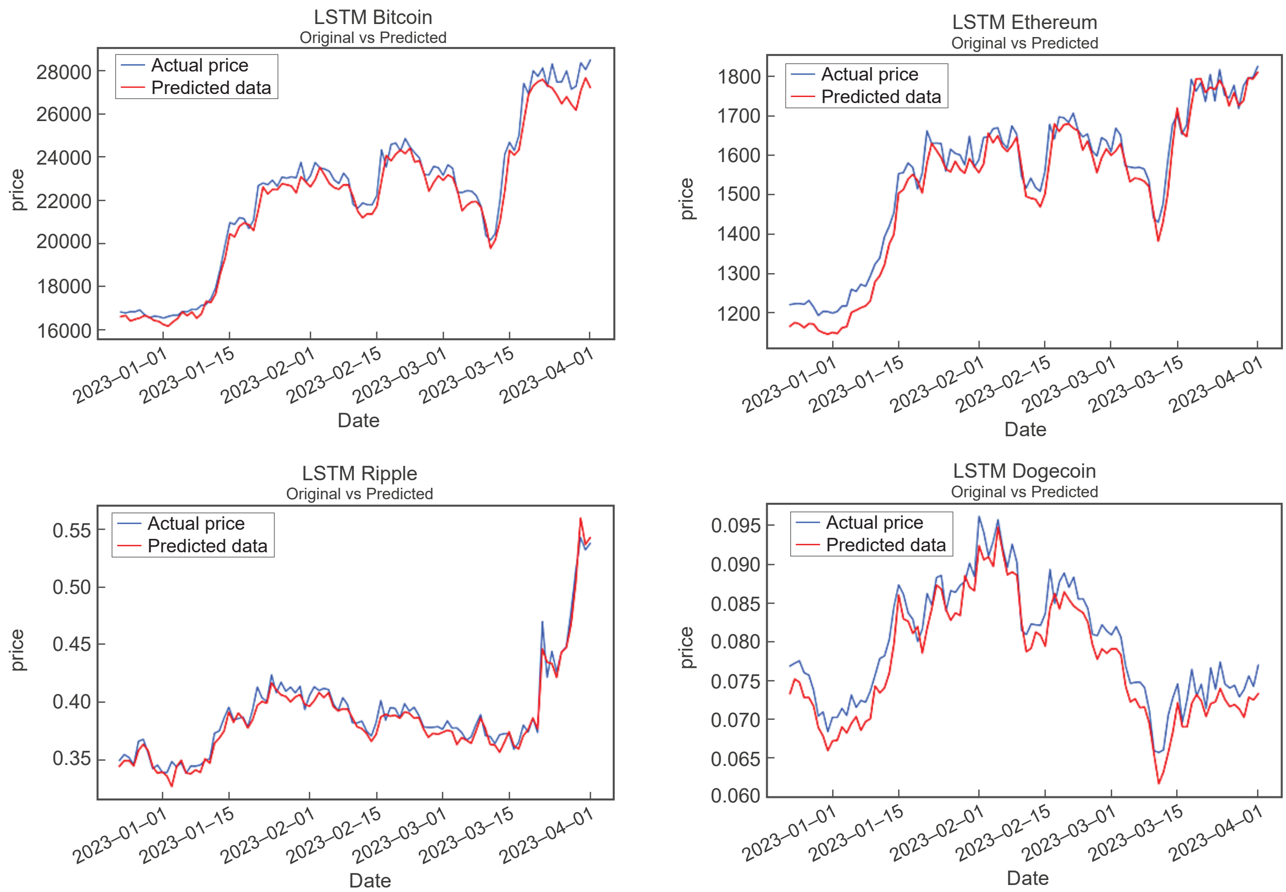


Figure 11. LSTM Forecast Models for Crypto Currencies for 23.12.2022 to 01.04.2023

Note: built by the author based on the results of modeling.