Tarnavskyi, Oleksandr; Kolomiiets, Viktor

Article

Identifying insurance companies' business models in Ukraine : cluster analysis and machine learning

Reference: Tarnavskyi, Oleksandr/Kolomiiets, Viktor (2021). Identifying insurance companies' business models in Ukraine : cluster analysis and machine learning. In: Visnyk Nacional′noho Banku Ukraïny (252), S. 37 - 55.
https://journal.bank.gov.ua/uploads/issues/252_eng_.pdf.
doi:10.26531/vnbu2021.252.02.

National Bank
of Ukraine

ISSN 2414-987X

VISNYK
OF THE NATIONAL BANK OF UKRAINE

No. 252
2021

VISNYK
OF THE NATIONAL BANK OF UKRAINE

No. 252
2021

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,

Although the full-scale war waged by russia against Ukraine is having significant adverse consequences for the financial and real sector, Ukrainian banks are resisting the challenges of wartime thanks to the previous reforms carried out in the banking sector, which allowed the banks to build up resilience margins. In contrast, some nonbank financial service providers – such as insurance companies – appear to have been unprepared for the challenges of wartime. Demand for their financial services has slumped significantly, and they have been forced to suspend or even roll back their activities. Moreover, in an environment of increased uncertainty, accurate forecasting and quick reactions by the National Bank of Ukraine to emerging operating risks are helping to tackle the negative effects of the war on the financial sector.

This issue of the Visnyk of the National Bank of Ukraine focuses on topics that have become particularly relevant recently, and that provide analytical support to the central bank's regulatory decision-making when responding to today's challenges.

In the first article of the current issue, *A Suite of Models for CPI Forecasting*, Nadia Shapovalenko addresses the topic of forecasting accuracy, testing the forecasting properties of the econometric models proposed for use by the National Bank of Ukraine for short-term CPI forecasting. Shapovalenko's findings suggest that these new models outperform benchmark AR models for almost all CPI components. The study also highlights the type of data restrictions experienced in wartime, and considers various avenues for improving the current suite of models for CPI forecasting.

In the second article*, Identifying Insurance Companies' Business Models in Ukraine: Cluster Analysis and Machine learning*, Oleksandr Tarnavskyi and Victor Kolomiiets study the business models of Ukrainian insurers and analyze their migration between clusters identified in the pre-war period. This approach allows the authors to determine the specific features of each model, and draw conclusions about insurers' financial stability. The results of this research are of particular relevance for insurance supervision and regulation.

We invite research contributors to investigate challenging research questions, such as examining the short-term and long-lasting effects of war, assessing current economic activity using new sources of data, regulating the banking and nonbanking sectors under conditions of continual shocks, designing post-war reconstruction and economic recovery, and so on. You are welcome to submit your 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, 2021, No. 252, pp. 4–36

A SUITE OF MODELS FOR CPI FORECASTING

NADIA SHAPOVALENKO[a]

[a]National Bank of Ukraine
Email: Nadiia.Shapovalenko@bank.gov.ua

Abstract This paper reviews the suite of models the National Bank of Ukraine uses for short-term forecasting of CPI components. I examine the forecasting accuracy of the following econometric models: univariate models, VAR, FAVAR, Bayesian VAR models, and Error Correction models. The findings suggest that for almost all components there are models that outperform benchmark AR models. However, the best performing individual model at each horizon for each component differs. Combined forecasts obtained by averaging the models' forecasts produce acceptable and robust results. Specifically, the combined forecasts are most accurate for core inflation, while they can beat the AR benchmark more frequently than other types of models when it comes to the raw food price index. This study also describes relevant data restrictions in wartime, and highlights avenues for improving the current suite of models for CPI forecasting.

JEL Codes C32, C51, C52

Keywords short-term forecasting, CPI, forecast evaluation

Acknowledgments: I appreciate the valuable comments of my colleges from the National Bank of Ukraine, especially Anton Grui and Andriy Tsapin. I would like to thank the anonymous reviewer for providing me with the ideas on how to restructure the paper and make some issues more comprehensive. I also extend my gratitude to Massimiliano Marcellino, professor of Econometrics in the Economics Department of Bocconi University, for his useful comments.

1. INTRODUCTION

In 2016 the National Bank of Ukraine (NBU) instituted de facto an inflation targeting (IT) regime. Under this framework, producing accurate and well-grounded forecasts of inflation is a difficult but essential task for the successful implementation of monetary policy. Usually, structural and semi-structural models are applied for a medium-term forecast, which covers a two- to four-year time horizon, whereas for the short-term forecast a variety of econometric models are used. The medium-term orientation gives central banks the flexibility to respond in an appropriate manner to the different economic shocks that may occur, however short-term projections are also of great importance for policy makers since they serve as a starting point for medium-term forecasts and policy analysis.

The development of the set of models for short-term CPI forecasting at the NBU started at the end of 2016 as part of a plan for a well-tailored and structured FPAS (Forecasting and Policy Analysis System). The first types of models developed were simple AR models and an ECM model aimed at forecasting one of the main CPI components – the raw food price index (RFPI). In the course of time, new types of models were developed to forecast the components of core inflation. Namely, in 2021 the set of models for the RFPI and Core CPI consisted of the following types: univariate models (AR, ARMA), vector autoregressive (VAR) models, factor augmented VAR (FAVAR) models, Bayesian VAR (BVAR) models, and error correction models (ECM). These models take into account the peculiarities of the Ukrainian economy and are based on the experience of peer central banks. It is important to point out that the final forecast combines the results of the model forecasts and expert judgments. Additionally, nowcasting based on web scraping is also used for the first month of the forecast. Detailed information regarding online price indexes which are used for nowcasting can be found in Faryna et al (2018).

The aim of this paper is to review the suite of econometric models used by the NBU for short-term CPI forecasting, examine the forecasting accuracy of these models, and elaborate recommendations on how to further improve the current models.

Various methods are usually applied for short-term inflation forecasting in central banks: starting from simple univariate models up to large dynamic factor models and Bayesian inference. Univariate models are a popular tool for producing bottom-up forecasts (Alvarez and Sanchez, 2017). Whereas multivariate models are able to incorporate a large amount of economic information into the short-term forecasting process (Akdogan et al., 2012). As a large amount of complex data is becoming available, increasing complexity in the data leads to increasing complexity in the models, with a growing number of parameters to estimate. One of the easiest ways to solve this issue would be to build a leading indicators model, either by regressing inflation on principal components derived from the indicators' data set, or to use each series individually and then combine forecasts. Dynamic factor models may be applied, as they not only benefit from exploiting information from large datasets but also account for the unbalanced data problem and have good forecasting properties. Another way to overcome dimensionality problems is to apply Bayesian techniques.

Available at https://doi.org/10.26531/vnbu2021.252.01

Several studies argue that in the presence of a large set of alternative forecasts, it is worth combining them rather than selecting one of them (Kapetanios et al. (2007) and Bjornland et al. (2008)). To verify the hypothesis that a combined forecast obtained by simply averaging all the alternative available forecasts tends to produce good and robust results in a variety of cases, I am going to compute combined forecasts and evaluate their accuracy.

The findings suggest that for almost all components there are models which outperform benchmark AR models. However, the best performing individual model differs at each horizon for each component. Combined forecasts obtained by averaging models' forecasts do produce good and robust results: for core inflation the combined forecasts are the most precise ones, while for the raw food price index they can beat AR benchmark more frequently than other types of models.

Due to data scarcity, especially for core components, a relatively short sample period for forecast evaluation is a considerable limitation. Moreover, the period of study covers the time of recovery from the financial crisis and military conflict, the switch to the IT regime in 2016, as well as COVID-19 pandemic. All these specific shocks may affect the behavior of macroeconomic variables and the relationships between them. That is why the research outcomes may be sensitive to the sample size, as well as the period studied.

To address the issue of IT-regime change, I estimate and analyze descriptive statistics for CPI components for "before IT" and "IT" subsamples. Two pre-IT samples were considered: one includes the whole period before 2016, while the alternative period excludes the beginning of 2015, the period when Ukraine experienced a huge nominal devaluation of the hryvnia. Note that there is not such a great difference in means if the devaluation period is omitted. This means that the difference was mainly explained by the effect of exchange rate pass-through to inflation. To solve the problem, I include the exchange rate as a control variable to multivariate models and include a dummy variable in univariate models.

For the COVID-19 crisis, I compare the percentages of types of models that have the best accuracy for various horizons and components for the whole sample of forecasting exercise and for the period of the COVID-19 pandemic. The results reveal that a different set of models are the most precise during COVID-19 pandemic times compared to the set for the whole forecasting sample. Namely, in crisis times models with a broad information set are more effective, and expert judgments may improve forecasts significantly.

This paper contributes to the existing literature by introducing a suite of models for short-term forecasting of inflation in Ukraine and analyzing their forecasting properties. The forecasts of inflation produced by this suite of models provide policy makers with a useful tool to assess current economic conditions and short-term developments.

The paper is organized as follows. In the next section techniques for short term inflation forecasting in CBs are examined. In section 3, the suite of models used for CPI forecasting at the NBU is described. The section contains both theoretical and empirical parts. In section 4, the forecasting properties of the models are reported and discussed. Finally, section 5 sets out conclusions and provides some recommendations on how to improve the forecasting performance of the models.

2. COMPARISON OF TECHNIQUES USED FOR CPI FORECASTING IN CENTRAL BANKS

Central banks usually apply a range of approaches and methods for short-term inflation forecasting. The following central banks are reviewed for their short-term forecasting methods: The Bank of Spain (BoS), the National Bank of Poland (NBP), the Central Bank of Bosnia and Herzegovina (CBBH), the Bank of England (BoE), the Central Bank of the Republic of Turkey (CBRT), the Bank of Norway (BoN), the Reserve Bank of New Zealand (RBNZ), the European Central Bank (ECB) and the Bank of France (BoF).

Information on the methods of the short-term inflation forecasting in these central banks, as well as the references, can be found in Table A.1, Appendix A.

The following conclusions can be drawn after reviewing the modeling techniques used by various central banks.

First, not all of the central banks focus on a model-based forecast of headline CPI. Some of them (BoS, BoN and CBRT) exclude food, energy or administrative prices (mostly prices on tobacco) from headline CPI because of the high volatility and poor predictability of these components. However, others argue that such an approach is not suitable for countries with a high share of these volatile groups (CBBH), as inflation excluding food and energy deviates significantly from the inflation faced by a typical household in the country. In most of the studies, the horizon of short-term forecasting varies from two to four quarters. The ECB has an even broader horizon of six quarters. RBNZ applies similar types of short-term forecasting models to those it uses for medium-term inflation forecasting as a cross-check for central forecasts, and thus has a forecasting horizon of eight quarters. The MAPI model of the BoF provides both monthly forecasts for 12 months, and quarterly forecasts for 12 quarters.

Second, all the reviewed banks use various types of models, starting from simple univariate models up to large dynamic factor models and Bayesian inference. Univariate models are a popular tool for producing bottom-up forecasts, but they are mostly applied when there is a high degree of disaggregation (for example, 120 components in BoS). Such a strategy enables more detailed information on each component to be incorporated into the forecast.

In contrast, the ECB and BoE use multivariate models to forecast a smaller amount of CPI components. Since there is a need to incorporate a large amount of economic information into the short-term forecasting process, in addition to standard VAR and single equation models, many central banks apply methods and approaches that can summarize the information contained in large datasets by reducing their dimensions (i.e. reducing the parameter space). The easiest way to proceed is to build leading indicator models (NBP, BoN) either by regressing inflation on principal components derived from the indicator data set, or to use each series individually and then combine forecasts. Dynamic factor models have also been increasingly popular at central banks (CBRT, NBP, BoN) as they not only benefit from exploiting information from large datasets, but also account for the unbalanced data problem (the so-called "ragged edge") and have good forecasting properties.

Another option for overcoming dimensionality problems is to apply Bayesian techniques. BVARs are used at the NBP, CBRT, BoN, ECB, RBNZ, and CBBH. The main strength of Bayesian estimation is precisely the fact that it is able to supplement the information contained in the data with expert information. When producing a forecast, BVAR models use a very large panel of data without exhibiting any signs of overfitting, and, as reported in the examined working papers, they produce good forecasting results. In Bayesian analysis a correct prior specification is a very important part of model creation. Various types of priors are used at central banks: a theoretical Minnesota-style (CBRT, ECB, CBBH) or an uninformative, conjugate Normal-inverse Wishart (BoN).

Some central banks also use modifications of the Phillips curve in the forecasting process, as it is considered to be a canonical economic model for forecasting inflation. Namely, both BoS and NBP add a backward- looking element to the equation. NBP uses unit labor costs (a proxy for marginal cost) instead of the output gap. While CBRT estimates the Phillips curve in a time-varying fashion.

Third, since it is important to estimate and report the uncertainty around the forecasts, the majority of central banks mostly use density forecasts instead of point forecasts. Moreover, as all banks have a suite of models, the question arises as to whether it is necessary to combine forecasts or to identify a baseline model and to use the others as supplementary ones. In the BoN paper, it is strictly recommended to combine some forecasts: “the next generation of macro modelers at Inflation Targeting central banks should adapt a methodology from the weather forecasting literature known as “ensemble modelling.” The NBP, CBRT, BoS and BoN report that the combined forecasting performance is better than that of any single model. However, the forecasts are combined in different ways. The BoS and CBRT use RMSE-based weights, whereas NBP uses log-predictive scores inside the models’ groups, and equal weights across the groups. At BoN, the weights attached to different models change within the quarter as new data is released. Some banks (BoN, CBRT, RBNZ, NBP) apply the strategy of building large sets of models of a similar type and then combining the forecasts from each type of model. The motivation for this is to avoid instabilities in the models caused by considerable uncertainty regarding the models’ specifications (e.g. choosing lag lengths, data-samples, variables to be included, etc.).

Fourth, depending on the type of model, the forecasts can be conditional or unconditional. The conditional forecast is based on the assumed future path of a set of inflation determinants (i.e. assumptions). Hence, conditioning allows forecasts to be more realistic. It makes the interpretation of forecasts and story building around them easier. However, the assumed values of these factors may vary from the actual ones and compound the forecasting error.

Fifth, many central banks are reporting that BVAR models have superior forecasting abilities in comparison to other models (CBBH, CBRT, ECB). For inflation in Spain, the best model is the multivariate one. Namely, a transfer function model that consists of single equation models describing the relationship between the main components of inflation and various explanatory variables. In the CBRT paper, the authors argue that models that use multivariate predictors outperform univariate models in terms of forecasting inflation, since “multivariate models exploit larger data sets, which are likely to contain more information about inflation, compared to univariate models.” In contrast, for inflation in Norway, the leading indicators model class shows the best performance most of the time, for all horizons. Thus, for BoN having a broad information set seems to add little extra value to performance. As for Phillips curve models, in general they tend to show poorer forecasting performance in comparison to other models, however they can provide some helpful insight as they seek to identify the effect on inflation of changes in demand.

To sum up, the NBU applies similar methods and techniques for short-term inflation forecasting as at peer central banks. As various banks use different measures of accuracy, and look at various forecast horizons and price indexes (CPI or different components of CPI), it is not possible to compare quantitatively the precision of the NBU forecasts to those at peer central banks. However, it is possible to compare whether the same techniques are claimed to be superior, and examine the issue of the accuracy of combined forecasts.

3. CPI FORECASTING IN UKRAINE

3.1. Stylized Facts of CPI

In the last two decades, inflation in Ukraine has been relatively high, the average year-over-year growth being around 10%. Since 2005, Ukraine has had two episodes with inflation exceeding 20%. In 2008, at the beginning of the World Financial Crisis, the Ukrainian economy was overheated. Despite the slowdown in GDP growth during the crisis, consumption growth together with a loose fiscal policy aimed at increasing social standards resulted in a substantial growth in minimum wages, which pushed prices upward.

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

In 2010-2011 the economy started recovering. Inflation fell to single digits and the exchange rate stabilized, while growth in consumption and nominal wages rebounded.

In 2012-2013 inflation approached zero due to weak economic activity (the annual GDP growth was 0.2-0.0%). Keeping the exchange rate stable led to an accumulation of huge imbalances in the economy. In 2014 these imbalances, along with the military conflict in the east of the country, led to a severe economic crisis with the real GDP falling by 10% in 2015, a sharp depreciation of the hryvnia, and inflation reaching a peak of almost 60% year-over-year in the spring of 2015. It is worth noting that the natures of the two episodes of high inflation (2008 and 2015) are different: the second inflationary spike was caused by the pass-through of the hryvnia devaluation, whereas in 2008 rising inflation was a sign that the economy had been overheating.

In August 2015 the NBU announced a transition to an IT regime in order to break the upward inflationary trend and stabilize the economy. De facto it moved to an inflation targeting regime in 2016, setting the following targets for inflation:

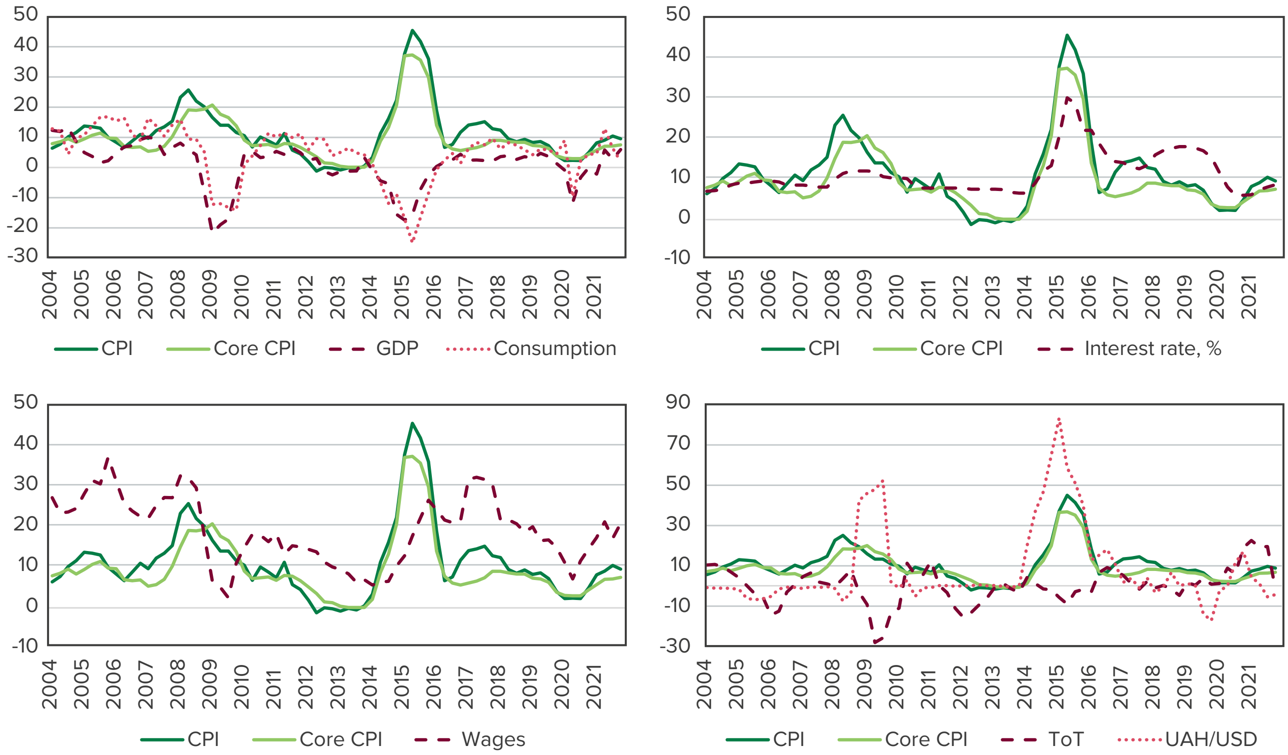


Figure 1. Main Economic Indicators, yoy in logs

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

The inflation targeting regime uses the short-term interest rate as its main instrument, and foreign exchange interventions as an additional one. To bring inflation down to the target, the NBU increases the interest rate to moderate demand and ease inflationary pressures. Thus, a gradual strategy of bringing inflation to its target was chosen deliberately in order to minimize the costs of disinflation for economic growth.

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

2020 brought a new challenge: The COVID-19 pandemic was a shock of unprecedented severity affecting all areas of the economy. At the beginning of the COVID-19 pandemic, households' consumer behavior changed. In the first half of 2020, during the stricter lockdown, some goods and services were not consumed, as selling them was prohibited or restricted. Thus, households cut spending on these items. The ability to work remotely affected demand for clothing and transportation services. Plummeting demand for many non-essential goods and services caused a decrease in prices. Prices for fuel also decreased significantly due to weak demand. However, prices for some raw food components increased substantially, due to both a lower-than-expected harvest and higher prices for food on the international markets.

Moreover, the structure of consumer spending was impacted by physical restrictions on the consumption of some goods and services, changes in demand on the back of the spread of remote working and studying, and high uncertainty over the course of the pandemic. The changes in consumer patterns during COVID-19 may not be fully reflected in official CPI estimates because according to "Consumer Price Index Manual: Theory and Practice" (2004), the stability of the price index weight structure has to be preserved. The NBU estimated a new price index with adjusted CPI weight structure [1] to analyze the impact of changes in consumption. According to NBU estimates of Covid inflation, by the end of 2020 it exceeded official inflation by 0.2–0.6 ppt. This corresponds in general to the results obtained by other countries. Moreover, considering the statistical properties of the CPI (see the means and deviations of the CPI components in Figure 2) such a deviation from official inflation probably doesn't affect the forecasting accuracy of the models significantly. In general, being lower than its target during 2020, inflation returned to its target in December 2020. However, in 2021 consumer inflation accelerated and exceeded its target largely due to rises in the prices of energy and some raw food items.

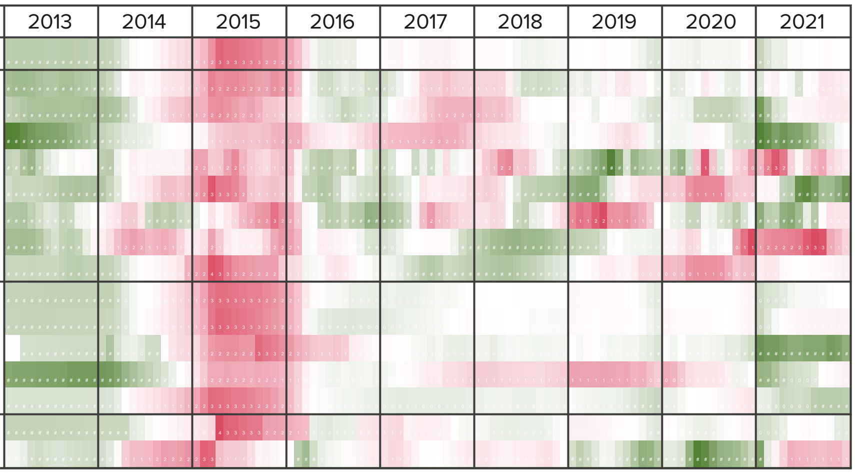
To sum up, the recent economic developments in Ukraine show that along with domestic conditions, external prices and the exchange rate are other important drivers of inflation and should be taken into account when forecasting Ukrainian inflation.

3.2. Factors Influencing The Dynamics of CPI Components

The NBU uses the year-over-year growth rate of CPI index as its target. CPI tracks changes in the market prices of a basket of consumer goods and services. It is comprised of 328 sub-indices. The weights of the items

[1] More information on estimates of COVID-19 inflation can be found Box 1. Covid Inflation in Ukraine from the NBU Inflation Report (January, 2021).

	Weight	Mean	St. dev.
CPI	100.0	13.4	14.4
RFPI	19.8	11.3	14.1
Meat	8.0	11.4	11.7
Milk	1.8	12.7	9.1
Eggs	1.0	16.7	34.3
Fruits	2.8	14.6	29.7
Vegs	3.0	6.6	25.3
Sugar	1.2	17.9	28.5
Cereals	1.9	19.0	30.3
Core CPI	58.3	10.6	11.8
Processed food	23.1	13.2	14.9
Clothes&Shoes	5.5	5.0	12.4
Services	13.0	10.3	4.9
Others	16.7	10.0	16.0
Adminstred	18.3	22.0	22.9
Fuel	3.5	13.7	21.6



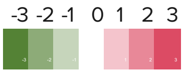


Figure 2. Heat Map of CPI Components

Note: Heat map is constructed for year-over-year, end of period percent change of CPI components, all indices are normalized. Weights are average for 2016-2021, means and standard deviations are calculated for 2013-2021.

from the basket are dynamic and can be adjusted to changes in the structure of consumption and in the type of items consumed.

The dynamic of sub-indices is not homogeneous. There are many indices, and their underlying characteristics vary widely in terms of both mean and standard deviation. One of the ways to simplify the analysis of these time series is to use a heat map that visually represents the relative inflation levels of various CPI components (as in McGillicuddy and Ricketts (2015) and Álvarez and Sánchez (2017)). Heat maps for some sub-indices of Ukrainian CPI are presented in Figure 2 (a more detailed heat map containing 92 items is presented in Figure B.1, Appendix B). It can be seen that for different CPI components, the periods of increase or decrease in prices as well as the causes of such dynamics are non-identical. For example, in mid-2020 only fuel prices decreased substantially as a consequence of a slump in global oil prices. In 2019, the increase in the prices of services was caused by a change in tariffs for transportation and communications, whereas an increase in prices on vegetables was spurred by unfavorable weather conditions. Such examples show that altogether with the analysis of common factors influencing inflation, it is worth splitting CPI into groups and looking at the factors which are specific for each group. will look into four major components of the CPI: core CPI, the raw food price index (RFPI), prices for fuel, and administrative prices.

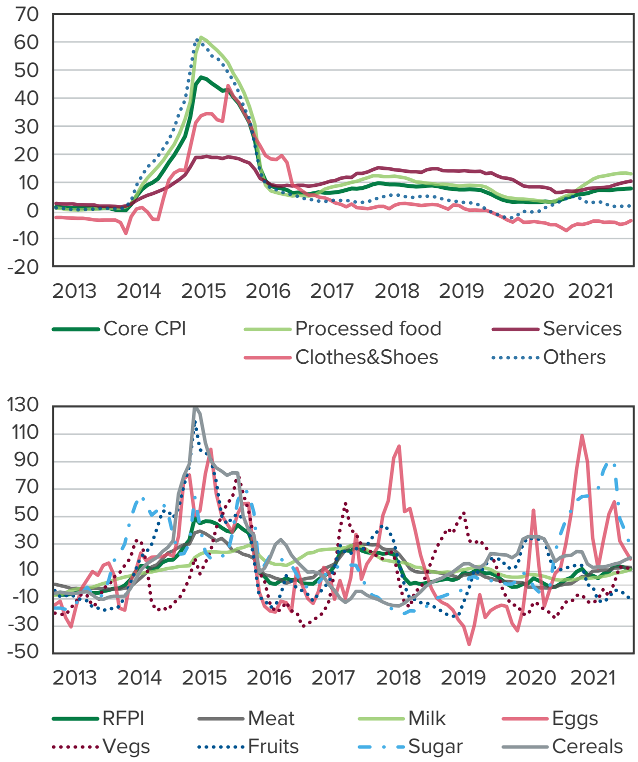


Figure 3. The Dynamics of Sub-Components of CPI, 12-mnths percent change

RFPI (Raw food price index)

The RFPI accounts for 19.8% of the CPI basket. The RFPI itself consists of the following components: “meat”, “milk”, “eggs”, “cereals”, “fruits”, “vegetables” and “sugar” (see the price dynamic of the components in Figure 3). The RFPI is considered to be the most volatile component of CPI for several reasons. First, raw food goods are demand inelastic, i.e., a consumer cannot eat twice as much food just because the price for that food has decreased substantially. Second, a quick adjustment to a supply shock in the short run is also difficult task, i.e., crop and livestock production are influenced by weather and diseases. If a crop is destroyed by severe weather conditions, it takes time to grow a new one.

The RFPI is driven more by supply than demand factors – production and harvest are important determinants of the RFPI dynamics. In order to decide which factors should be taken into account, i.e., which sub-components depend not only on harvest or production, but also on the exchange rate and prices on international markets, it is worth analyzing consumption balances – namely the share of exports in production and the share of imports in consumption. A full set of plots can be found in Figure B.2, Appendix B. In general, it is obvious from the consumption balances plots that for the “cereals”, “meat”, “milk” and “fruits” groups, external factors are important. As there is a trend for increasing exports of eggs and sugar, the exchange rate and prices on external markets could be considered for these groups of goods as well.

It is worth mentioning that seasonality in food items is more profound than in other items, and that this also depends on the share of domestic production in consumption (e.g., potato and vegetables are planted and consumed mostly domestically and have more intense seasonality than meat, which is traded internationally) and shelf life (the seasonality of the "processed food" group is less profound than that of the RFPI).

Core CPI

Core CPI accounts for 58% of the CPI basket and consists of four main components: "processed food", "clothes and shoes", "services" and "others".

As Ukraine has moved to an inflation targeting regime, it stands to reason that the policy rate should have an influence on the least volatile and most monetary policy relevant part of the CPI. However, taking into account the medium-term orientation of monetary policy and the fact that I am focused on short-term forecasting, it is also worth considering other indicators that are more applicable for the short-run.

As core inflation is considered to be more demand driven, the nominal wage indicator seems to be a good proxy for changes in demand, given that it is available on a monthly basis and assumptions regarding its dynamics during the forecasting period are also available.

Exchange rate dynamics seem to be another important factor: when the devaluation occurred in 2015 the "processed food" and "others" groups had the highest exchange rate pass-through (these groups have more intense color on the heat map in 2015). The main reason for such behavior is probably the high share of imported groups in these two components. In contrast, "services" had the smallest pass-through, reflecting the high share of non-tradable goods in this group.

Fuel Prices

Fuel prices account for 4% of the CPI basket. Fuel prices in domestic currency mostly depend on the nominal exchange rate and oil prices on international markets, as Ukraine is considered to be a net importer of energy goods. The prices for fuel are not forecasted within the framework of time series models, and need only assumptions for the nominal exchange rate, oil prices on international markets, and the excise tax.

Administered Prices

Administered prices account for 18% of the CPI basket. They mainly consist of prices for utilities, transportation services and alcohol and tobacco. As the dynamics of these prices mostly depend on information about the value of excise tax and information from local authorities regarding tariffs, it would be reasonable to use expert judgments instead of time series models when forecasting these prices.

To sum up, headline CPI is broken down into smaller components, each representing a different subset of goods and services. The suite of models is applied for two components

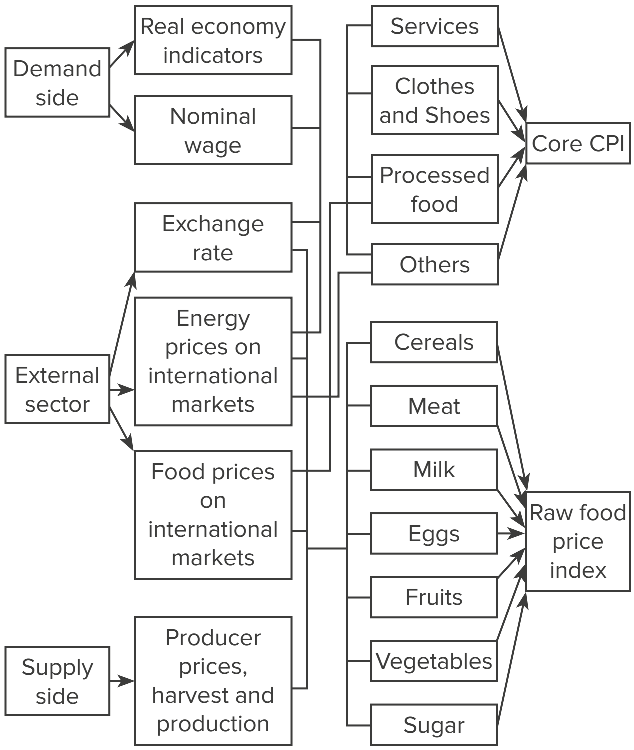


Figure 4. Factors that Drive the Dynamic of CPI Components

of CPI, namely, the RFPI and core CPI. Altogether they account for 78% of headline CPI. On the basis of the analysis conducted above, the indicators used in modelling are represented in Figure 4. More detailed information about the time series[2] used in the models is given in Tables A.1-A.2, Appendix A.

3.3. The Suite of Models used by the NBU

The NBU uses several types of models for the short-term forecasting of inflation in Ukraine: univariate models (AR, ARMA), vector autoregressive (VAR) models, factor augmented VAR (FAVAR) models, Bayesian VAR (BVAR) models, and error correction models (ECM). Each type is introduced and discussed below.

But before discussing the different types of models for the short-term forecasting of inflation, I would like to address the issue of sample stability. First, instability may arise due to a switch in the monetary policy regime. Namely, the implementation of the IT regime in 2016 could have changed the statistical properties of data, which can lead to huge forecasting errors if forecasts of price indices after 2016 are produced by models estimated using data from before 2016. To verify whether the statistical properties have changed, the means, standard deviations and AR coefficients of RFPI and Core inflation are analyzed (see Figure B.3). Two pre-IT samples were considered: one includes the whole period before 2016, and the other excludes the beginning of 2015, when Ukraine experienced a huge nominal devaluation of the hryvnia. We can see in the figure that if we do not consider the devaluation period, there is not such a great difference in means, indicating that the difference was mainly caused by the

[2] All data are measured in natural logarithms. As almost all levels of prices, production and harvest are I(1) processes according to the stationarity test, first differences of the variables are used. An identifiable seasonality test is used to decide whether a variable is to be seasonally adjusted by X12.

effect of exchange rate pass-through to inflation. To solve the problem, multivariate models contain the exchange rate as a regressor. Whereas for univariate models, a possible solution is the inclusion of dummy variables (as described in the ARMA models subsection below). It is clearly seen that during the IT period the values of the standard deviation for RFPI and Core inflation decreased, which is quite a common situation for countries implementing an IT regime. See, for example, how inflation deviation shrank after the implementation of the IT regime in New Zealand (Archer, 2000).

Second, sample instability may be caused by various factors that are specific to a certain group of goods. For example, changes in consumption or production patterns (an increase in the share of imports in consumption or exports in production) may influence the coefficients of a model. Similarly, dummies can be used to take these changes into account.

Autoregressive (AR) Models

Time series models, which in general extrapolate patterns in historical data, are considered to be the most appropriate for short-term forecasting (Galbraith and Tkacz (2006)). Univariate models, the simplest among them, are commonly used as a benchmark in the forecasting literature. Quite often, the forecasting properties of these models are found to be superior to large multiple-equation models such as vector autoregression and traditional structural macroeconomic models. Moreover, having few independent variables, they are believed to be convenient for short data samples.

Simple AR equations are estimated and used for forecasting. Lag length may be chosen according to various criteria (Akaike, Schwarz, Hannan-Quinn), however a first-order autoregressive model usually serves as a benchmark model.

AR equation can be written as:

$$dP_t^j = \alpha_0^j + \sum_{i=1}^{l^j} \alpha_i^j \cdot dP_{t-i}^j + \varepsilon_t^j \qquad (1)$$

where P_t^j is a price level of j-component [3] at time t, dP_t^j is a first difference at time t, l^j- is a l period lag, and ε_t^j is a randomly distributed error term.

AR equations are used both for forecasting the RFPI and core CPI components. The Schwarz criterion is used to find optimal lags. The results of the estimation are presented in details in Table A.4, Appendix A. The results show that the components of core inflation have more persistence than most components of the RFPI. This confirms the initial observation from the stylized facts section that the prices of most of raw food items are highly volatile. Moreover, some equations have quite a high S.E. (Standard Error) value. In other words, this type of model is not good at explaining the dynamics of certain prices. The results may be improved by using more sophisticated model structures, namely ARMA models.

Autoregressive Moving Average (ARMA) models

Another time series method for explaining variables in terms of their own past values is the ARMA (or more generally ARIMA [4]) model. In addition to autoregressive terms, this model has moving average terms. The notation ARMA (l^j, q^j)refers to a model with l^j autoregressive terms and q^j moving-average terms for each j-th price component:

$$dP_t^j = \alpha_0^j + \sum_{i=1}^{l^j} \alpha_i^j \cdot dP_{t-i}^j + \sum_{k=1}^{q^j} \beta_k^j \cdot \varepsilon_{t-k}^j + \varepsilon_t^j \qquad (2)$$

According to Box et al (2015), models containing processes of different types are considered to be more parsimonious. Namely, a model with small values l^j and q^j will do as well at explaining a process dP_t^j as a high order AR(l^{*j}) or MA(q^{*j}) process.

ARMA models are used to produce disaggregated forecasts of core inflation components (240 items). To account for excessive market movements and possible structural changes, an ARMAX type of model (ARMA with exogenous variables) was chosen. Namely, dummy variables were added into the specification:

$$dP_t^j = \alpha_0^j + \sum_{i=1}^{l^j} \alpha_i^j \cdot dP_{t-i}^j + \sum_{k=1}^{q^j} \beta_k^j \cdot \varepsilon_{t-k}^j + \gamma^j * D_t^j + \varepsilon_t^j \qquad (3)$$

where D_t^j is a dummy variable for j-th price component.

More detailed information about the model's structure and selection of dummy variables can be found in Krukovets and Verchenko (2019).

The main disadvantage of applying univariate models is that they do not use additional information that the available data may contain. In other words, such models don't refelect any structural relationships in the data, and lack economic meaningfulness. Thus, it is worth applying multivariate models to take into consideration additional information and increase the explanatory power of the model.

Vector Autoregressive (VAR) Models

VAR models are usually applied to describe relationships between different variables as well as between current and lagged observations. A standard VAR with l lags is expressed as:

$$Y_t = A_0 + \sum\nolimits_{i=1}^{l} A_i * Y_{t-i} + \varepsilon_t \qquad (4)$$

where $Y_t = [y_{1,t}, \dots . y_{n,t}]^T$ is a vector of variables, A_0 is a $n \times 1$ vector of constants, A_i is a $n \times n$ matrix of coefficients of Y_{t-i}, l- is number of lags and ε_t is a $n \times 1$ vector of residuals with multivariate normal distribution $\varepsilon_t \sim N(0, \Sigma)$, $E(\varepsilon_t \varepsilon_t') = \Sigma, E(\varepsilon_t \varepsilon_s') = 0$ if $t \neq s$.

Many empirical studies on the international transmission of shocks are based on VAR models that include only a few selected variables. However, Mumtaz and Surico (2009) argue that because of their small-scale, there may be a possibility of mis-specification of the models or incorrect interpretation of fundamental shocks. From a practical perspective, small scale VARs are also unable to provide inferences on a large number of variables that may be of

[3] The number of components is J, j=1..J, in our case J=11, namely 7 components of RFPI and 4 components of Core CPI.

[4] An ARIMA (autoregressive integrated moving average) model is a generalization of an ARMA model. ARIMA models are used when data show evidence of being non-stationarity. To eliminate non-stationarity, differencing is applied (as many times as an order of integration of the initial series). Since we model month-over-month changes in prices, which are supposed to be stationary, we do not need differencing. For ARMA models J=240.

interest. Hence, for the purposes of short-term forecasting, a wider information set can be used.

Though large VAR models disclose more information from data and are commonly used in forecasting, estimation of the parameters of such models requires long data samples, as the number of VAR parameters increases with the square of the number of variables. I.e., the number of observations must exceed the number of estimated parameters, which means being more than $k = n(n * l)$ for model (4).

One way to avoid the dimensionality problem, if the variable of interest is ${y^j}_{1,t}$ is to estimate $n^j - 1$ bivariate VARs of the form, as in Andersson and Löf (2007):

$$ {Y^j}_{b,t} = {A^j}_{b,0} + \sum_{i=1}^{l^j} {A^j}_{b,i} * {Y^j}_{b,t-i} + {\varepsilon^j}_{b,t}, b = 1 \dots n^j - 1 \quad (5) $$

where ${Y^j}_{b,t} = [{y^j}_{1,t}, {y^j}_{b+1,t}]^T$, ${y^j}_{1,t} = dP_t^j$, $y_{b+1,t}$ is the first difference of the b+1-th variable

At the end, each of models will produce forecasts for P_t^j. Thus, having $n^j - 1$ individual forecasts allows us to compute a variety of statistics and produce density forecasts for the variable of interest.

Bivariate VARs are estimated for seven components of the RFPI and four components of core CPI. The information set for each forecasted variable is given in Table A.5, Appendix A. Equal weights are used to construct a combined forecast for a set of bivariate VARs for each forecasted variable.

Another way to decrease dimensionality is to condense data in many variables into just a few variables, using factor analysis.

Factor Augmented VAR (FAVAR) Models

Bernanke et al. (2004) suggested adding an unobserved factor into a small-scale VAR model. Earlier, Stock and Watson (2002) forecasted inflation using factor estimation to account for more than several hundred variables. Further details regarding the optimal number of dynamic factors and tests for the factor restrictions can be found in Stock and Watson (2005). In our case, the model is similar to that in Lombardi et al (2012), who examined linkages across non-energy commodity price developments using a FAVAR model:

$$ \begin{bmatrix} Y_t \\ F_t \end{bmatrix} = \Phi(L) \begin{bmatrix} Y_{t-1} \\ F_{t-1} \end{bmatrix} + u_t \quad (6) $$

where $Y_t = [y_{1,t}, \dots . y_{m1,t}]^T$ is a vector containing the variable of interest and some fundamentals, $F_t = [f_{1,t}, \dots . f_{m2,t}]^T$ – factors extracted from information set $X_t = [x_{1,t}, \dots . x_{n,t}]^T$, $m1 + m2 \ll n, u_t$ is a $(m1 + m2) \times 1$ vector of residuals with multivariate normal distribution $u_t \sim N(0, Q)$.

There are several options for extracting factors for FAVARs. Since Bernanke et al. (2004) and Oskarsson and Lin (2018) found that applying more sophisticated methods rather than simple principal components analysis (PCA) did not yield significantly better results, I am going to use PCA as well.[5]

FAVARs estimated for seven components of the RFPI and four components of core inflation have similar bivariate structure to the models from the previous section. Namely, for j-th price component:

$$ {Y^j}_{favar,t} = {A^j}_{favar,0} + \sum_{i=1}^{l^j_{favar}} A_{favar,i} * Y_{favar,t-i} + {\varepsilon^j}_{favar,t} \quad (7) $$

where ${Y^j}_{favar,t} = [{y^j}_{1,t}, {pc^j}_{1,t}]^T$, ${pc^j}_{1,t}$ is the first principal component of the data set $[{y^j}_{2,t}, \dots {y^j}_{n^j,t}]$ which are the first differences of the variables and ${y^j}_{1,t} = dP_t^j$ is the forecasted variable. The information set for each component is similar to that used for bivariate VARs (see Table A.5, Appendix A). As mentioned, only the first principal component, which explains the most, was used. However, for some CPI components it is obvious that it is not sufficient to have only the first principal component.

So far, most of the described approaches did not consider models with exogenous variables. Hence, they do not require any assumptions on factors which allow a wider information set to be used.

However, it is also worth having models containing exogenous variables. Usually, forecasts are based on some assumptions about either external or internal factors (e.g., for the RFPI index it may be information regarding harvests or world prices dynamics, for core CPI – an increase in minimum wages announced by the government). Consequently, making forecasts based on assumptions allows the forecasts to be more realistic and consistent, as well as it making the interpretation of forecasts and building a story around them easier. Moreover, these models may provide us with a scenario analysis.

Bayesian VAR (BVAR) Model

Another alternative for dealing with the dimensionality problem by shrinking the parameters via the imposition of priors is a Bayesian VAR (BVAR) model. Given the fact that the sample size for the Ukrainian data is short, standard OLS estimates of parameters can be imprecise, thus making obtained impulse responses and forecasts unreliable. Banbura et al. (2008) show that with Bayesian shrinkage, it is possible to handle an unrestricted VAR with a large number of variables, where the data set can even be extended to incorporate disaggregated sectoral or geographical indicators.

The imposition of priors not only solves the dimensionality problem but supplements the information contained in the data with personal judgments contained in the prior. The recent literature on forecasting models points out that among a variety of empirical models, BVARs have superior abilities in forecasting.

One of the main challenges in this approach is the selection of prior distributions. I use the procedure developed in Litterman (1986) and impose Minnesota-style priors.

Let's consider a VAR with exogenous variables of the form of:

$$ Y_t = \sum_{i=1}^{l} A_i * Y_{t-i} + C * X_t + \varepsilon_t \quad (8) $$

where $Y_t = [y_{1,t}, \dots . y_{n,t}]'$ is a vector of variables, A_i is a $n \times n$ matrix of coefficients of Y_{t-i}, *l*- is the number of lags, C is a $n \times m$ matrix, $X_t = [x_{1,t}, \dots . x_{m,t}]$ is a $m \times 1$ vector of exogenous variables, and ε_t is a $n \times 1$ vector of residuals with multivariate normal distribution $\varepsilon_t \sim N(0, \Sigma)$, $E(\varepsilon_t {\varepsilon_t}') = \Sigma, E(\varepsilon_t {\varepsilon_s}') = 0$ if $t \neq s$

Reformulating the model for the whole data set [1…T] and vectorizing it we obtain:

[5] FAVAR models are also used to nowcast quarterly GDP figures. More detailed information can be found in Grui and Lysenko (2017).

$$y^{vec} = \bar{X}\beta + \varepsilon^{vec} \tag{9}$$

where $y^{vec} = vec(Y),\ \ Y = (Y_1, \dots, Y_T)'$

$$\bar{X} = I_n \otimes X, \qquad X = \begin{pmatrix} Y_0 & \dots & Y_{1-l} & X_1 \\ \dots & \dots & \dots & \dots \\ Y_{T-1} & \dots & Y_{T-l} & X_T \end{pmatrix}$$

$$\beta = vec(\mathrm{B}), \mathrm{B} = (A_1, \dots, A_l\ , C)'$$

$$\varepsilon^{vec} = vec(\mathrm{E})\ \mathrm{E} = (\varepsilon_1, \dots, \varepsilon_T)', \varepsilon^{vec} \sim N(0, \bar{\Sigma})$$

where $\bar{\Sigma} = I_T \otimes \Sigma$

multivariate normal assumption on ε_t gives:

$$(y^{vec}|\beta) \backsim N((X \otimes I_T)\beta, I_T \otimes \Sigma) \tag{10}$$

Bayesian estimation of VAR centers around the derivation of posterior distributions of β and Σ. It is assumed that β follows a multivariate normal distribution, with mean β_0 and covariance Ω_0

$$\beta \backsim N(\beta_0, \Omega_0) \tag{11}$$

Litterman (1986) proposed the following prior: As most observed macroeconomic variables seem to be characterized by a unit root, each endogenous variable included in the model presents a unit root in its own first lags, and coefficients equal to zero for further lags and cross-variable lag coefficients. In the absence of prior belief about exogenous variables, the most reasonable strategy is to assume that they are neutral with respect to the endogenous variables, and hence that their coefficients are equal to zero as well. In the case of variables known to be stationary, this unit root hypothesis may not be suitable, so that a value around 0.8 may be preferred to a value of 1.

Ω_0 is assumed to be a diagonal matrix. The diagonal elements, corresponding to endogenous i and j at lag l are specified by:

$$\delta^l_{0,i,j} = \begin{cases} \left(\dfrac{\lambda_1}{l^{\lambda_3}}\right)^2 \ for\ j = i \\ \left(\dfrac{\lambda_1\lambda_2\delta_i}{l^{\lambda_3}\delta_j}\right)^2 \ for\ j \neq i \end{cases} \tag{12}$$

where $\lambda_1\lambda_2$ and λ_3 are hyper-parameters and δ_i is the square root of the corresponding $(i,i)^{th}$ element of an initial estimate of Σ. The Minnesota prior also assumes that Σ is fixed, forming no prior on Σ.

More technical details can be found in Dieppe at all (2016).

The procedure for the selection of the models for the RFPI and Core CPI that have the best fit is organized in a following way:

1) Various exogenous variables are tried, the ones with minimum log likelihood are chosen;
2) Standard lag length criteria were used to select the lag length (see Table A.6, Appendix A);
3) A grid search similar to the procedure used by Giannone et al. (2012) is applied to find the values of the hyperparameters for the model (see Table A.7, Appendix A).

The best specifications are presented in Table A.8, Appendix A.

For the RFPI components, the best model contains the exchange rate and the FAO price index as exogenous variables.[6] The magnitude of the exchange rate and FAO price index shock varies and, in most cases, depends on the share of exports in domestic production.

The impulse responses for the BVAR model with four components showed that a shock to “processed food” prices is significant for “others” prices. As both groups have a high share of imported goods, it is probable that the price dynamics of both “others” and “processed food” are driven by a common factor – exchange rate movements. To check this hypothesis, “processed food” and “others” prices were combined in one group and a BVAR for three components was estimated. Overall, the response of price index of combined groups to the exchange rate turned out to be significant. In addition, a model with three components is more parsimonious than one with four components.

The best models for core inflation components contain two exogenous variables – nominal wages and the exchange rate. The impulse responses show that prices in the “services” group are highly sensitive to nominal wages, while “processed food” and “others” prices are mostly affected by exchange rate dynamics. This conclusion is in line with the fact that prices for “services” contain a significant share of nontradables, and are mostly driven by domestic factors. As already mentioned, “others” and “processed food” prices have high share of imported goods, and as a consequence have the strongest response to exchange rate shocks.

Similarly to the BVAR for RFPI components, the exogenous variables of the BVAR models for Core CPI components were tested for exogeneity using a Granger causality test. According to the test results, the direction of causality for the exchange rate was as expected: from the exchange rate to the price components. In contrast, the “services” component and “others” component doesn’t have causality with nominal wages in either direction. Thus, treating nominal wages as exogenous may lead to the fact that the model won’t be able to interpret or will misinterpret some relationships between the variables.

The latter issue deserves being explode in a separete study. There are two possible options: either endogenize nominal wages, or find another more relevant exogenous indicator. In case of endogenizing nominal wages, the model forecasts should be conditioned on the indicator for nominal wages in order to be coherent with the forecasts for nominal wages produced during the forecasting cycle.

Error Correction Models (ECM)

If one wants to take into consideration specific factors for each component, systems of equations can be used. For example, in the case of the RFPI components, a more detailed analysis of supply factors would be interesting: instead of combined data on harvests, it is worth looking at the relations between an RFPI component and its particular harvest (e.g., how the harvest of vegetables and potatoes influences prices for vegetables and potatoes).

Also, while analyzing the influence of the exchange rate and external prices, it is worth having the advantage of being able to incorporate both short-run dynamics and long-run equilibrium relations among variables. Thus, in addition to existing models, an ECM (error correction mechanism)

[6] To make sure that the FAO price index and the exchange rate can be treated as exogenous, a Granger causality test was conducted, indicating the correctdirection of causality for exogenous versus endogenous variables.

model is estimated. A similar approach is applied in De Charsonville et al. (2017) to forecast the main components of HICP for France.

In an ECM type model, equations in levels represent cointegrating relationships[7], which capture medium term dynamics, while the cointegration term derived from the equation accounts for the deviation of variables in the medium term. This approach thus provides us with a forecast of CPI components for both the short and medium term.

j-th equation in levels is the following:

$$P^j_t = \sum_{i=1}^{m^j} \theta^j_i \cdot X^j_{i,t} + \delta^j_t \qquad (13)$$

where P^j_t is a price level of the j-th component $X^j_1 \ldots X^j_{m^j}$ – is a set of exogenous regressors for j-th price level, both price level and exogenous regressors are of I(1), and δ^j_t is a normally distributed residual.

To derive the coi^j_t term, rewrite (13):

$$coi^j_t = (P^j_t - \sum_{i=1}^{m^j} \theta^j_i \cdot X_{i,t}) = \delta^j_t \qquad (14)$$

The equations in first differences contain the coi term (14):

$$dP^j_t = \sum_{i=1}^{l^j} \alpha^j_i \cdot dP^j_{t-i} + \beta^j \cdot coi^j_{t-1} + \\ + \sum_{k=1}^{n^j} \gamma^j_k \cdot dX^j_{k,t} + \varepsilon^j_t \qquad (15)$$

where dP^j_t is first difference of the j-th price level, $dX^j_1 \ldots dX^j_{n^j}$ is a set of first differences of exogenous regressors $X^j_1 \ldots X^j_{n^j}$ for the j-th equation, and $\varepsilon^j_{i,t}$ is a normally distributed residual. Note that the set of exogenous variables for the j-th equation ($X^j_1 \ldots X^j_{m^j}$) in levels doesn't necessarily coincide with the set of exogenous variables for the j-th equation ($dX^j_1 \ldots dX^j_{n^j}$) in differences.

The model for core inflation consists of:

- 4 equations for core CPI components (as in 15)
- 4 identities for coi terms (as in 14), where coefficients are estimated using equations in levels (as in 13)
- Identity for aggregated index

Non-zero residuals are used to adjust the forecasted value of the current month (according to nowcasting results) as well as to include expert judgments into the model. The model for the RFPI has the same structure, although it consists of seven components.

To account for specific structural breaks, individual dummies are used. Namely, this reflects a growing export share in production ("meat", "eggs", "processed food"), import share in consumption ("vegetables", "milk"), an asymmetric effect of currency appreciation ("cereals"), a surge in the minimum wage ("services"), pandemic events ("services", "clothes and shoes"). Also, I look at coefficients' recursive estimates to ensure the stability of the models' parameters.

[7] See Table A.2, A.3 Appendix A with the results of data stationarity tests for levels and differences, and also Table A.9 Appendix A with the results of an Engle-Granger cointegration test for equations in levels.

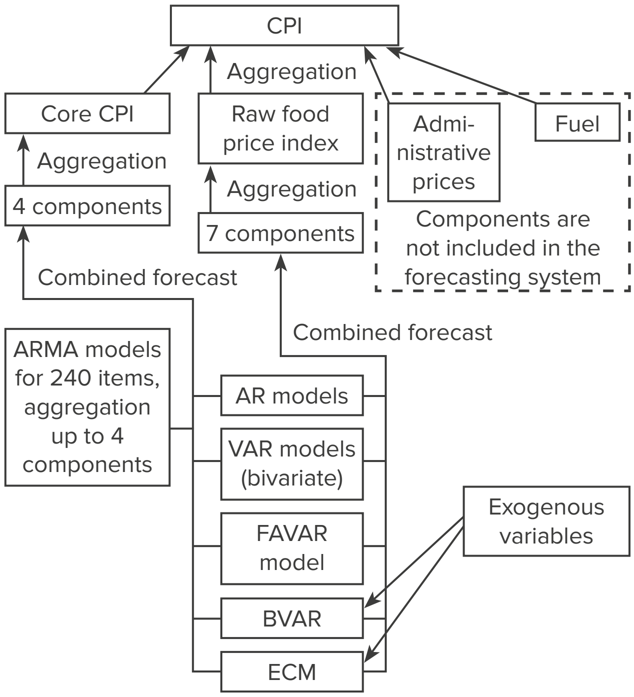


Figure 5. System for Short Term Forecasting of CPI

Additionally, I looked for various indicators that could be included into equations of Core CPI components in an effort to improve predictive accuracy and to better reflect relationships between economic variables. Namely, I estimated the specifications including (1) lags of interest rate and first difference of M2 to better capture monetary policy stance; (2) GDP gap and real marginal costs for Core CPI, taken from the QPM (Quarterly Projection Model) model, to replicate the elements of a Phillips curve; (3) data on surveys, such as the index of the propensity to consume and the index of consumer sentiment to account for changes in demand. However, in most specifications I either got the wrong sign or non-significant coefficients. Only real marginal costs for Core CPI components turned out to have forecasting power. This issue should be investigated more closely in further research: new specifications, similar to the ones used by De Charsonville et al. (2017) for French data and a "thick" Phillips curve approach (both specifications with and without inflation expectations), which is regularly employed in the Eurosystem's macroeconomic projection exercises to cross-check underlying inflation (Baumann U. et al, 2021), could be estimated.

The details on equations are represented in Table A.9, Appendix A. The magnitude of such factors as the exchange rate, nominal wages, and FAO prices is similar to that one produced in BVAR models. In addition, I may conclude that the exchange rate pass-through in the short run is smaller, and specific supply factors for each group of the RFPI are significant. The whole system of forecasting of CPI components is shown in Figure 5.

Forecasts of components are further aggregated to obtain forecasts of core and raw food indices. It is worth mentioning that State Statistic Office uses a complex system of dynamic weights, which is replicated by the NBU during the forecasting process. However, to simplify the calculations, yearly average weights are used in this paper.

4. FORECASTING PERFORMANCE

In this section I test the forecasting performance of the models. First, I describe what measures were used, and explain how the period of forecast evaluation exercise was chosen. Second, I look at forecasting performance of the models for the RFPI and its components, and then for Core CPI and its components. Finally, I analyze the forecast bias and address the issue of the quality of CPI forecasts during the COVID-19 pandemic.

4.1. Measures of Forecast Evaluation

The analysis is predominantly based on the RMSE (formula 16) indicator, as it is considered to be quite a widespread measure of a forecast's precision. The AR model serves as a benchmark. RMSE values are shown relative to those of an AR1 model in order to facilitate the comparison (formula 17). Thus, for the given model, a value of RMSE below unity means better than the AR1 model's precision.

Additionally, I compute a Theil index (formula 18), which also provides a measure of the distance of the true from the forecasted values. A Theil index always lies between 0 and 1, thus it makes the comparison of forecast evaluation for different components easier. For example, RMSE would usually be higher for the RFPI rather than core components because of the high volatility of raw food prices. Applying the Theil index, I can compare forecast accuracy of different indices using the same scale between 0 and 1: the closer the Theil index is to 1, the worse the forecasting accuracy.

I also analyze the forecasting bias, which is measured as the average forecast error at a certain horizon (formula 19). In turn, the forecast error is calculated as the difference between the actual value and the forecasted one. A non-zero bias indicates a possible persistent difference between the forecasts and the observed values. The formulas for the accuracy measures are the following:

$$RMSE_{j,m} = \sqrt{\sum_{t=T}^{T+h-1} (\hat{Y}_{t+j,m} - Y_{t+j})^2/h} \tag{16}$$

$$RMSE_{j,m}^{rel} = RMSE_{j,m}/RMSE_{j,AR} \tag{17}$$

$$Theil_{j,m} = \frac{\sqrt{\sum_{t=T}^{T+h-1}(\hat{Y}_{t+j,m} - Y_{t+j})^2/h}}{\sqrt{\sum_{t=T}^{T+h-1}\hat{Y}_{t+j,m}{}^2/h} + \sqrt{\sum_{t=T}^{T+h-1}Y_{t+j}{}^2/h}} \tag{18}$$

$$FBias_{j,m} = \sum_{t=T}^{T+h-1} (\hat{Y}_{t+j,m} - Y_{t+j})/h \tag{19}$$

where $RMSE_{j,m}$, $RMSE_{j,m}^{rel}$, $Theil_{j,m}$, $FBias_{j,m}$ are consequently RMSE, RMSE relative to AR, Theil index and forecast bias of the forecast of model m, forecast horizon j. The forecast sample of length h is the following $T, T+1 \dots . T+h-1$, $\hat{Y}_{t+j,m}$ is the forecast of model m started at time t for forecast horizon j, and Y_{t+j} is the actual value.

RMSE, the Theil index and the forecast bias are calculated for forecasts of different models as well as for combined forecasts. Equal weights are used to combine the forecasts of the following models: AR[8], VAR, FAVAR, ECM and BVAR for raw food components, and AR, VAR, FAVAR, ECM, 4BVAR and CARMA for core inflation components.

4.2. Evaluation of the RFPI Forecasts

The forecasting evaluation exercise uses monthly data for the period of 2016m9–2021m12 for the RFPI, and 2018m03-20212m12 for Core CPI as for these periods official forecasts of the components are available and can be compared with model forecasts. It should be noted that I am interested in forecasts made in particular months, namely months when the official inflation forecast of the NBU is released and published in the Inflation Report.[9] Assumptions are available for these particular months, which serve as exogenous variables for ECM and BVAR models. These assumptions are the same for both the other satellite models and the QPM model, which makes the short-term forecast of CPI components consistent with the predictions of other macroeconomic indicators produced by the NBU.

Table 1 shows the best performing models for each horizon and for each component of the RFPI. It can be seen

Table 1. Best Models for the RFPI and its Components (according to RMSE)

	Forecast Horizon					
	1 m	2 m	3 m	4 m	5 m	6 m
RFPI	IR	IR	CMB	CMB	AR	VAR
Cereals	IR	IR	VAR	AR	AR	VAR
Meat	BVAR	FAVAR	VAR	FAVAR	CMB	CMB
Milk	BVAR	BVAR	ECM	ECM	CMB	CMB
Eggs	IR	ECM	CMB	CMB	AR	BVAR
Vegs	BVAR	BVAR	ECM	BVAR	BVAR	CMB
Fruits	BVAR	AR	VAR	ECM	ECM	BVAR
Sugar	VAR	ECM	ECM	ECM	VAR	AR

[8] Further in the text, for simplicity, forecasts of the different models can be identified by the following abbreviations: AR- autoregressive model, VAR-combination of bivariate VARs, FAVAR- FAVAR model, ECM – ECM model, BVAR- BVAR model for the RFPI, 3BVAR and 4BVAR are BVAR models for core CPI with three and four components, CARMA- set of ARMA models, IR- official forecasts of the NBU, CMB-combined forecast of different models using equal weights.
[9] The Inflation Report reflects the opinion of the NBU as to the current and future economic state of Ukraine, with a focus on inflationary developments, which form the basis of monetary policy decision-making. The Inflation Report is published quarterly in accordance with the forecast periodicity.

that for most of the RFPI components, the best performing individual model differs, and the AR benchmark model is beaten in the majority of cases.

BVAR models show good forecasting performance for different components of the RFPI, especially at the beginning of the forecast horizon. That is consistent with the results of other studies that found that BVAR models with Litterman's prior outperform alternative models such as univariate time series models and VAR models (Akdogan et al. (2012), Bloor (2009) and Hasanovic (2020)).

Table 2. RMSE Relative to AR RMSE for the RFPI

	Forecast Horizon					
	1 m	2 m	3 m	4 m	5 m	6 m
FAVAR	1.006	1.031	1.008	0.992	**1.001**	1.000
VAR	1.001	1.020	0.996	0.993	1.004	**0.998**
ECM	1.030	1.000	1.021	0.992	1.129	1.007
BVAR	0.943	1.104	0.990	1.021	1.073	1.008
CMB	0.980	0.986	**0.971**	**0.988**	1.020	0.999
IR	**0.544**	**0.923**	1.087	1.052	1.097	1.111

Compared to other methods, the ECM model forecasts some of the RFPI components relatively well for horizons from two to five months. This may reflect a link with the long-run level of prices.

Official forecasts of the NBU published in the Inflation report (named IR), appear to be the best for "cereals", "eggs" and the RFPI for the horizon of the first month. The high forecasting accuracy of IR for the first month confirms the high precision and usefulness of nowcasting, and the importance of the incorporation of expert judgments for some components.

Combined forecasts are the best pick in around 21% of total cases: for the sixth and the fifth months of "meat" and "milk", the sixth month of "vegetables" and also for the third and the fourth months of "eggs". However, if we look at Table 2, presenting relative RMSE figures for the RFPI, it can be seen that even though combined forecasts are the best only for the third and the fourth months, they can beat the AR benchmark more frequently than other types of models. Accuracy might be improved even further if a more sophisticated system of weights is used. For example, Akdogan et al. (2012) uses inverse RMSE weights, whereas in Timmerman (2006) other generalizations are discussed.

The formula for the inverse RMSE weights is the following:

$$w_{jm} = \frac{RMSE_{j,m}^{-1}}{\sum_{m=1}^{M} RMSE_{j,m}^{-1}} \tag{20}$$

Where $m = 1..M$ is m-th type of model, $j = 1..h$ is the forecast horizon, and $RMSE_{j,m}$ is RMSE for m-th model for the j-th horizon.

The plots with RMSE and the Theil index for the RFPI components can be found in Figures B.4, B.6, Appendix B. As the scale of the Theil index is similar to each component, comparisons of accuracy between the groups can be made. According to the Theil index, the forecasts for "milk", "meat", "fruits" and "vegetables" prices seem to be more accurate than the forecasts of the other components. Forecasts of "sugar" from the second to the sixth month have the lowest precision. In general, the forecasts of the RFPI index are of decent accuracy. This may be due to the high accuracy of the forecasts of its main components. Another reason may be the fact that the error forecasts of different components are canceled out while aggregating the forecasts of components into RFPI forecasts. I further analyze forecast bias to check this hypothesis.

Plots of forecast bias can be found in Figure B.8, Appendix B. Indeed, for various components and models forecast bias is either negative or positive and has different patterns, which may support the hypothesis on the canceling out of errors while aggregating the RFPI forecasts. The forecast bias of the RFPI for most of the models is the smallest: it is slightly positive for the first three months, and then becomes slightly negative. Also, the forecast bias of more volatile components like "eggs", "vegetables" and "fruits" is larger. The bias patterns of various models for "meat", "milk" and the RFPI are different, this fact may lead to gains in forecast accuracy for combined forecasts of these prices.

Finally, I would like to discuss the models' accuracy during the COVID-19 pandemic, which covers the period from 2020m3 to 2021m12. As shown in Figure 6, the percentages of models that have the best accuracy for different horizons and components for the whole sample of forecasting exercise differ from those in COVID-19 pandemic times. The decrease in the percentage of combined forecasts by 17 ppt may be attributed to the above-mentioned fact that an equal weighting scheme is not optimal for combined forecasts. The increase in the percentage of multivariate models (FAVAR and VAR altogether) by 8 ppt shows the effectiveness of using models with a broad information set in times of crises and other extraordinary events. The better forecasting performance of the ECM model may have the same origin: the ECM model's equations contain a lot of factors individual to each group, namely, supply side factors (such as harvest and production) as well as a variety of international prices.

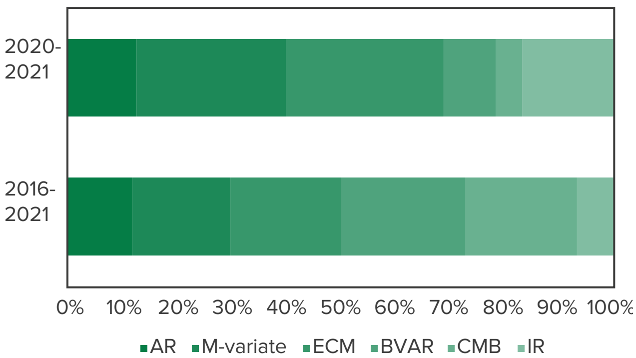


Figure 6. Percentages of Types of Models with the Best Forecasting Performance for Different Forecast Norizons and Components [10], RFPI

[10] The percentages for the whole sample of forecasting exercise (2016-2021) correspond to the frequency of the type of model represented in the Table 1. The similar was done for the part of forecasting exercise sample (2020-2021) to access the models' performance during COVID-19 pandemic. M-Variate includes FAVAR and VAR models.

Table 3. Best Models for Core CPI and Its Components (according to RMSE)

	Forecast Horizon					
	1 m	2 m	3 m	4 m	5 m	6 m
Core CPI	CMB	CMB	CMB	CMB	CMB	CMB
Processed food	IR	CMB	CMB	CMB	IR	3BVAR
Services	AR	CMB	CMB	CARMA	CMB	CMB
Cloth and shoes	CARMA	ECM	CARMA	CMB	CMB	CARMA
Others	IR	CMB	4BVAR	4BVAR	4BVAR	CMB

4.3. The Evaluation of Core CPI forecasts

Contrary to the RFPI components, for core CPI and its components the models that have the best precision are not so diverse: combined forecasts are the best for core CPI and partly for other components, CARMA is the best partly for “services” and for “clothes and shoes”, while BVAR is the best partly for “processed food” and for “others” (see Table 3).

For the “clothes and shoes” group CARMA turned out to be the best model for almost the whole forecast horizon. The good performance of the CARMA model means that the precision of a univariate model is higher than that of multivariate ones. The reason for this is likely due to the statistical properties of both of these groups (prices for “clothes and shoes” follow the ARMA process well enough and are less volatile than food prices) and the scarcity of the multivariate models' information dataset for the “clothes and shoes” group. A similar conclusion for a slightly richer information dataset was made by Aastveit et al. (2011) regarding inflation forecasting in Norway. Using new explanatory variables for these groups would probably improve the forecasts of multivariate models.

The combined forecasts of core CPI turned out to be the most precise, outperforming other models’ forecasts significantly (see Table 4), and confirming the conclusions of Kapetanios et al. (2007) and Bjornland et al. (2008) regarding the superiority of combined forecasts. Such large gains in precision were achieved because for some models the bias is positive, while for other models it is negative (see Figure B.9 in Appendix B). Thus, the combination of models’ forecasts led to a more precise and unbiased outcome.

For all of the core CPI components, except for the “clothes and shoes” component, the sign of the bias varies across the models. Forecasts for the “clothes and shoes” component show a consistent positive bias for all types of models. This may be partly explained by changes in the methodology. In 2014, the SSSU began including sales prices, thus decreasing the overall level of prices.

According to the Theil index (see Figure B.7 in Appendix B), forecasts of core CPI components are more precise than those of the RFPI components, “clothes and shoes” forecasts being the most accurate. However, for the “others” group the forecasts are the least accurate since it is very hard for this group to find appropriate indicators, explaining the price dynamic.

For Core CPI components, the shift in best types of models during COVID-19 pandemic times is more profound: there is a slump in the percentage of combined forecasts (from 57 to 7%%) and increase in the percentage of IR forecasts (from 10% to 30%%). The increase in precision of IR forecasts shows that in crisis times expert judgments may improve forecasts significantly. Similarly to the RFPI components, multivariate models such as FAVAR and VAR appeared to be highly precise in pandemic times: their percentage reached 27%.

The worsening in the performance of the models using exogenous variables (the percentage of BVAR and ECM

Table 4. RMSE Relative to AR RMSE for Core CPI

	Forecast Horizon					
	1 m	2 m	3 m	4 m	5 m	6 m
CARMA	0.706	1.100	1.077	0.902	1.120	1.184
FAVAR	0.877	1.055	1.011	0.964	0.965	0.976
VAR	0.878	1.006	1.017	1.008	1.017	1.000
ECM	1.042	1.815	1.768	1.414	1.622	1.473
4BVAR	0.914	1.333	1.118	1.011	1.227	1.134
3BVAR	0.982	1.438	1.498	1.261	1.412	1.260
CMB	**0.628**	**0.903**	**0.942**	**0.830**	**0.909**	**0.920**
IR	0.875	1.218	1.442	1.385	1.230	1.187

altogether decreased by 10ppt) may be due to the fact that the assumptions on the exogenous variables used in the models remarkably differed from the actual realizations of the data, thus worsening the forecasting ability of the models with exogenous variables.

To verify this, I calculated forecasts for ECM and BVAR models using actual realizations of the data instead of assumptions for exogenous variables. The RMSE of the models can be seen in Figure B.10, Appendix B. RMSEs of the models using actual data are lower than that ones' using assumptions, the difference between the RMSEs being wider for Core CPI. Thus, there is evidence that the difference between actual data and assumptions may further increase the forecasting error.

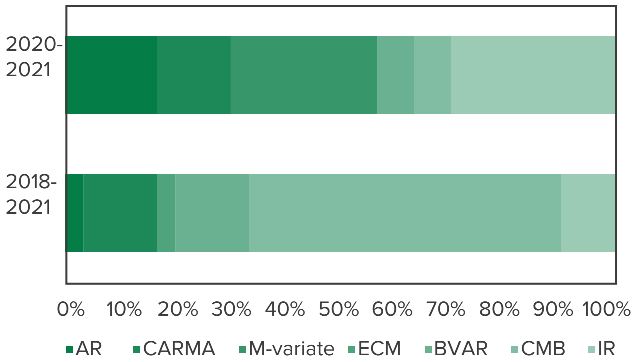


Figure 7. Percentages of Types of Models with the Best Forecasting Performance for Different Forecast Horizons and Components, Core CPI

4.3.1. Comparison of the QPM Forecasts and Combined Forecasts of the Suite of Models

The QPM[11] is the core model of the FPAS system at the NBU. At its core it has a set of theory-based relationships that capture the key part of the transmission mechanism. It provides the organizational framework for macroeconomic forecasting and story-telling, as well as having to be able to forecast the main economic indicators. Karam et al (2006) argue that in many central banks it is recognized that the model is much less accurate than experts at forecasting the first one or two quarters. That is why, to improve the forecasting qualities of the core model, the final forecast can be a hybrid of the QPM and short-term forecasting models, i.e the QPM forecast could include short-term tunes coming from satellite models.

Table 5. RMSE Relative to AR RMSE for q-o-q Core CPI

	Forecast Horizon	
	1 q	2 q
CMB	**0.79**	**0.84**
QPM	1.11	0.86

The best forecasts for core inflation i.e., combined forecasts, were transformed to quarterly frequency in order to compare the forecasting accuracy of the short-term forecasting system with that of QPM. CMB outperforms QPM in both the first and second quarters. However, in the first quarter the difference is more profound (the RMSE relative to AR can be found in Table 5).

Given the fact that the core inflation forecasts of the short-term forecasting models for the first and second quarters are more accurate than QPM's, we can incorporate the results of the suit of models into the QPM model in the form of short-term tunes. This will allow us to receive more precise short-term QPM forecasts.

5. CONCLUSIONS

This study reviews the suite of models used by the NBU for short term CPI forecasting and tests the forecasting properties of the employed models (univariate models (AR, ARMA), vector autoregressive (VAR) models, factor augmented VAR (FAVAR) models, Bayesian VAR (BVAR) models, and error correction models (ECM)), while the AR model serves as a benchmark. The forecasting evaluation exercises use monthly data for the period of 2016–2021. The findings of the paper suggest the following:

First, for almost all components of CPI there are models that outperform the benchmark AR models. BVAR models show good forecasting performance for different components of core CPI and the RFPI. This result is consistent with the conclusions of other studies arguing that BVAR models with Litterman's prior outperform alternative models, such as univariate time series models and VAR models. However, for the groups "services" and "clothes and shoes", the ARMA model forecasts turned out to be the most accurate. Similar results were observed in a Bank of Norway paper that showed that a rich data set added little extra value to multivariate models' performance.

Second, combined forecasts obtained by averaging models' forecasts produce acceptable and robust results, i.e. for core inflation the combined forecasts are the most precise ones, while for the raw food price index they beat the AR benchmark more frequently than other types of models. Thus, these findings confirm the conclusions of Kapetanios et al. (2007) and Bjornland et al. (2008) regarding the superiority of combined forecasts in comparison to individual model forecasts.

Third, the high forecasting accuracy of the official forecasts for the first month proves the precision and usefulness of nowcasting and the importance of incorporating expert judgments for some CPI components. In addition, the combined forecasts of core inflation for the first two quarters are more accurate than the forecasts produced by QPM. Hence, it looks promising to incorporate the results of a suite of models in the form of short-term tunes into the QPM model in order to receive more precise short-term QPM forecasts.

Fourth, the analysis of forecasting performance of the models during COVID-19 pandemic compared to the performance during the whole forecasting sample showed that models with a broad information set are more effective in times of crises or other extraordinary events. However, expert judgments also may improve forecasts significantly.

Even though this paper analyzes forecasts of inflation up to the end of 2021, it is worth briefly mentioning the influence of the russian invasion of Ukraine which began

[11] QPM is a semi-structural, forward-looking New-Keynesian model of a small open economy. It is a main element of the FPAS at the NBU. Detailed information regarding the QPM model can be found in Grui and Vdovychenko (2019).

on 24 February 2022 and highlight key challenges for forecasting CPI components in wartime. After the shock of the first weeks of war, economic activity began reviving in the relatively calm and liberated regions. In the first quarter of 2022, real GDP, decreased by 15.1% yoy. The slump in the second quarter was even deeper (-37.2%) due to large numbers of damaged and destroyed factories, enterprises and infrastructure. Additionally, there is a negative impact from the outflow of the labor force as well. In June 2022, consumer inflation accelerated to 21.5% yoy. The faster inflation was caused by both global trends (high energy prices) and internal factors (disrupted supply chains, higher production costs, and stronger household demand for some goods and services on the back of insufficient supply). Additionally, price pressure is uneven across the country's regions: the highest price hikes are in the temporarily occupied regions and in cities with active hostilities.

Measures taken by the government and the NBU partially offset the inflationary pressure caused by russia's full-scale invasion. The NBU was forced to fix the exchange rate and impose a number of administrative restrictions, including ones on FX transactions and capital movements, so as to maintain price and financial stability and to control inflation expectations. After that, on June 1, the NBU Board decided to raise the key policy rate to 25%. This is intended to spur investors' interest in hryvnia assets, while also easing pressures on international reserves and reining in inflation.

Thus, forecasting economic indicators in Ukraine in the near future will be very challenging, because of:

• Difficulties with data. The State Statistic Services of Ukraine announced that it will cease publishing the most of its official data during the war. The data on the CPI components, exchange rates, and international prices are still available but other information, such as that on economic output and labor statistics, has become heavily restricted or even unobservable and, therefore, researchers are forced to make numerous assumptions while modelling. It is not also clear how the statistics are being gathered in the occupied regions, and how the shift in consumption patterns had affected the data. In such circumstances limited data availability can cause some selection and estimation biases that might ultimately result in difficulty with forecasting. Soft data with a high frequency might be of benefit in this case.

• Structural changes. From the very beginning of the active phase of the war, drastic structural changes may have altered the statistical properties of data and the relationships between macroeconomic indicators. In these circumstances, expert judgements could be of great help (as during the COVID-19 pandemic). To learn of the experiences of other countries affected by wars, military conflicts, or natural disasters may also be beneficial for CPI forecasting in wartime.

Based on the results above, there are several ways to improve the existing suite of models and make its forecasts more accurate and better grounded:

• The BVAR model for Core CPI components could benefit from resolving the issue with the endogeneity of the existing indicators.

• The results from surveys, information obtained from non-conventional data sources like Google trends, etc. can enrich the dataset for FAVARs and bivariate VARs.

• New specifications, similar to ones used by De Charsonville et al. (2017) for French data and the "thick" Phillips curve approach (both specifications with and without inflation expectations), which is regularly employed in the Eurosystem's macroeconomic projection exercises (Baumann U. et al, 2021), could be estimated.[12]

• More sophisticated weighting schemes could be applied to combined forecasts in order to increase their precision.

REFERENCES

Aastveit, K. A., Gerdrup K. R., Jore, A. S. (2011). Short-term forecasting of GDP and inflation in real-time: Norges Bank's System for Averaging Models. Norges Bank Staff Memo, 09. Norges Bank. Retrieved from https://www.norges-bank.no/globalassets/upload/publikasjoner/staff-memo/2011/staff_memo_0911.pdf

Alvarez L. J, Sanchez I., (2017). A suite of inflation forecasting models. Occasional Papers, 1703. Bank of Spain. Retrieved from https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosOcasionales/17/Fich/do1703e.pdf

Akdoğan, K., Başer, S., Chadwick, M. G., Ertuğ, D., Hülagü, T., Kösem, S., Öğünç, F. (2012). Short-term inflation forecasting models for Turkey and a forecast combination analysis. Working Papers, 1209. Central Bank of the Republic of Turkey. Retrieved from https://tcmb.gov.tr/wps/wcm/connect/b9eb2da4-5028-4cf1-8baa-03f0465525e4/WP1209.pdf

Andersson M., Löf M. (2007). The Riksbank's new indicator procedures. Sveriges Riksbank Economic Review, 2007. Sveriges Riksbank.

Archer, D. (2000). Inflation targeting in New Zealand (a presentation to a seminar on inflation targeting, held at the International Monetary Fund, Washington, DC, March 20-21, 2000). Retrieved from https://www.imf.org/external/pubs/ft/seminar/2000/targets/archer.htm

Bańbura, M., Giannone, D., Reichlin, L. (2008). Large Bayesian VARs. Working Paper Series, 966. European Central Bank. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp966.pdf

European Central Bank (2021). Inflation expectations and their role in Eurosystem forecasting. Occasional Paper Series, 264. Retrieved from https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op264~c8a3ee35b5.en.pdf

Bernanke, B., Boivin, J., Eliasz, P. (2004). Measuring the effects of monetary policy: A Factor-Augmented Vector Autoregressive (FAVAR) approach. NBER Working Papers 10220. National Bureau of Economic Research. https://doi.org/10.3386/w10220

Bjornland, H. C., Gerdrup, K., Jore, A. S., Smith, C., Thorsrud, L. A. (2008). Improving and evaluating short

[12] We thank an anonymous referee for this suggestion.

term forecasts at the Norges Bank. Staff Memo, 04. Norges Bank. Retrieved from https://www.norges-bank.no/globalassets/upload/publikasjoner/staff-memo/2008/staff_memo_2008_04.pdf

Bloor C. (2009). The use of statistical forecasting models at the Reserve Bank of New Zealand. Reserve Bank of New Zealand Bulletin, 72, 21-26. Reserve Bank of New Zealand. Retrieved from https://www.rbnz.govt.nz/-/media/project/sites/rbnz/files/publications/bulletins/2009/2009jun72-2bloor.pdf

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., Ljung, G. M. (2015). Time series analysis: forecasting and control. Wiley Blackwell.

De Charsonville, F. Ferrière, C. Jardet, (2017). MAPI: Model for analysis and projection of inflation in France. Working Papers, 637. Banque de France. Retrieved from https://publications.banque-france.fr/sites/default/files/medias/documents/dt-637.pdf

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

Faryna, O., Talavera, O., Yukhymenko, T. (2018). What drives the difference between online and official price indexes? Visnyk of the National Bank of Ukraine, 243, 21-32. https://doi.org/10.26531/vnbu2018.243.021

Figueredo, F. M. R., Guillen, O. T. C. (2013). Forecasting Brazilian consumer inflation with FAVAR models using target variables (preliminary draft). Banco Central do Brasil. Retrieved from https://www.bcb.gov.br/secre/apres/FAVAR%20paper%20Figueiredo%20&%20Guillen%20prelim.pdf

Galbraith, J., Tkacz, G, (2007). How far can we forecast? Forecast content horizons for some important macroeconomic time series. Staff Working Paper, 2007-1. Bank of Canada. https://doi.org/10.34989/swp-2007-1

Grui, A., Lysenko, R. (2017). Nowcasting Ukraine's GDP using a factor-augmented VAR (FAVAR) model. Visnyk of the National Bank of Ukraine, 242, 5-13. https://doi.org/10.26531/vnbu2017.242.005

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

Hasanovic, E. (2020). Forecasting inflation in Bosnia and Herzegovina. IHEID Working Papers, HEIDWP07-2020. The Graduate Institute of International Studies. Retrieved from http://repec.graduateinstitute.ch/pdfs/Working_papers/HEIDWP07-2020.pdf

International Labour Office et al (2004). Consumer price index manual: theory and practice. Retrieved from https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/presentation/wcms_331153.pdf

Giannone, D., Lenza, M., Momferatou, D., Onorante, L. (2014). Short-term inflation projections: A Bayesian vector autoregressive approach. International Journal of Forecasting, 30 (3), 635-644. https://doi.org/10.1016/j.ijforecast.2013.01.012

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

Kapetanios, G., Labhard, V., Price, S. (2007). Forecast combination and the Bank of England's suite of statistical forecasting models. Bank of England Working Papers, 323. Bank of England. Retrieved from https://www.bankofengland.co.uk/working-paper/2007/forecast-combination-and-the-boe-suite-of-statistical-forecasting-models

Karam, P., Laxton, D., Berg, A. (2006). Practical model-based monetary policy analysis: a how-to guide. IMF Working Papers, WP/06/81. International Monetary Fund. Retrieved from https://www.imf.org/external/pubs/ft/wp/2006/wp0681.pdf

Krukovets, D., Verchenko, O. (2019). Short-run forecasting of core inflation in Ukraine: a combined ARMA approach. Visnyk of the National Bank of Ukraine, 248, 11-20. https://doi.org/10.26531/vnbu2019.248.02

Litterman, R. (1986). Forecasting with Bayesian vector autoregressions: five years of experience. Journal of Business & Economic Statistics, 4(1), 25-38. https://doi.org/10.2307/1391384

Mazur, G. (2022). Probabilistic inflation forecasting with model pooling (materials from the seminar held by the NBP in 22–23 February 2022).

McGillicuddy, J. T. Ricketts, L. R. (2015). Is inflation running hot or cold? Economic synopses, 2015(16). Federal Reserve Bank of St. Louis. https://doi.org/10.20955/es.2015.16

Mumtaz, H., Surico, P. (2009). The transmission of international shocks: a factor-augmented VAR approach. Journal of Money, Credit and Banking, Blackwell Publishing, 41(s1), 71-100. https://doi.org/10.1111/j.1538-4616.2008.00199.x

Oskarsson, M., Lin, C. (2018). A simplified approach in FAVAR estimation (Bachelor Thesis). Upsala University. Retrieved from https://www.diva-portal.org/smash/get/diva2:1215768/FULLTEXT01.pdf

Rummel, O. (2015). Economic modelling and forecasting (presentation). Bank of England.

Stock, J., Watson, M. (2002). Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association, 97(460), 1167-1179. https://doi.org/10.1198/016214502388618960

Stock, J., Watson, M. (2005). Implications of dynamic factor models for VAR Analysis. NBER Working Papers, 11467. National Bureau of Economic Research. http://www.nber.org/papers/w11467.pdf

Timmermann, A, (2006). Forecast Combinations. Handbook of Economic Forecasting, 1 (4), 135-196. Elsevier.

National Bank of Ukraine (2021). Inflation Report, January 2021. Retrieved from https://bank.gov.ua/admin_uploads/article/IR_2021-Q1_eng.pdf?v=4

APPENDIX A. TABLES

Table A.1. Methods and Techniques Used in CBs for Short-Term Forecasting

Authors	Bank	Object	Models	Models with best accuracy	Estimation period and reestimation	Forecast period	Density or point forecast	Combi-nation	Judge-ments
Aastveit et al. (2011)	BoN	CPI excl taxes and energy	AR models for main CPI components, bivariate VAR,VAR, BVAR, VECM, Factor models	bivariate VARs	1982Q4 or 1993Q1, depending on the approach	1-2Q	density	yes	yes
Akdogan et al. (2012)	CBRT	CPI excl unprocessed food and tobacco, no disaggregation	Univariate models, Nonlinear models, Phillips curve motivated time varying parameter model, VAR, BVAR, Dynamic factor models	BVAR	2003-2011	1-2Q	point	yes	yes
Alvarez, Sanchez (2017)	BoS	CPI, disagregated 120 items, CPI exluding food and energy	Univariate models, Transfer function models, Phillips curve motivated model	Transfer function models	sinse 2012	1-3Q	point	no info	yes
Bloor (2009)	RBNZ	CPI, GDP (both short and medium term forecasting)	1. VARs (both classical and bayesian VARs, VECM) 2. Leading indicators models (bivariate VARs ,bridge equations, AR) 3. Factor models	Combined forecast of leading indicators models	forecast exercise sample 2000-2008	1-8Q	not clear	not clear*	yes
De Charsonville et al. (2017)	BoF	HICP, 5 items and administered prices, 21 components for 3M	ARIMA, ECM	ECM	1996Q1/ 2007Q3**-2014Q4	1-12Q 1-12M	point	no	yes
Giannone et al. (2010)	ECB	HICP, 5 items, PPI	BVAR	BVAR	since 1991	1-18 M	density	no info	yes
Mazur (2022)	NBP	CPI, disagregated 42 components	S-ARIMA, Dynamic factor model, Liding indicator, BAR, BVAR	no info	no info	1-12M	density	yes	yes

Table A.1 (continued). Methods and Techniques Used in CBs for Short-Term Forecasting

Authors	Bank	Object	Models	Models with best accuracy	Estimation period and reestimation	Forecast period	Density or point forecast	Combi-nation	Judge-ments
Hasa-novic (2020)	CBBH	CPI	ARMA,VAR, BVAR	BVAR	2007-2017	1-12M	point	no info	no info
Rummel (2015)	BoE	CPI, disa-gregated 31 items	Naive sam-ple aver-age, ARMA, VARMA, FAVAR	no info	no info	1-6M	density	yes	yes

* As for 2009, RBNZ was considering the benefits of model averaging versus forecasts from individual forecasts, and the possible use of density forecast instead of point forecasts.
** The size of the sample varies for each component.

Table A.2. Time Series Used for Forecasting

Name	Description	Source	Beginning of Sample/ frequency	Source of data forecasts (for exogenous)	Seasonality test	Stationarity test		
						ADF Prob (level)	ADF Prob (1st diff)	ADF Prob (yoy diff)
	IMF prices of							
IMF_P_WHT	wheat	IMF	2004/m	IMF OUTLOOK	NP	0.27	0.00	0.00
IMF_P_BRL	barley				P	0.50	0.00	0.00
IMF_P_SOY	soybeans				NP	0.09	0.00	0.00
IMF_P_CHCK	chicken				NP	0.50	0.00	0.00
IMF_P_OIL	sunflower oil				NP	0.02	0.00	0.00
IMF_P_SGR	sugar				P	0.33	0.00	0.00
	FAO price index of							
FAO_P_F	food	FAO	2004/m	MPEAD assessments	NP	0.56	0.00	0.00
FAO_P_CRL	cereals				Probably NP	0.48	0.00	0.01
FAO_P_MT	meat				Probably NP	0.33	0.00	0.00
FAO_P_CHCK	chicken				NP	0.34	0.00	0.00
FAO_P_BF	beef				NP	0.76	0.00	0.00
FAO_P_PRK	pork				P	0.10	0.00	0.00
FAO_P_DAI	dairy products			REUTERS	NP	0.19	0.00	0.00
FAO_P_SGR	sugar				NP	0.29	0.00	0.00
	WB prices of							
WB_P_CRL	cereals	WB	2004/m	MPEAD assessments	P	0.46	0.00	0.03
WB_P_FUEL	energy				NP	0.16	0.00	0.00
WB_P_OIL	sunflower oil			WB OUTLOOK	NP	0.14	0.00	0.00
WB_P_FRT	fertilizers				NP	0.51	0.00	0.00
WB_P_BN	bananas				NP	0.78	0.00	0.00
WB_P_ORN	oranges				NP	0.00	0.00	0.00
WB_P_SGR	sugar				NP	0.29	0.00	0.00

Table A.2 (continued). Time Series Used for Forecasting

Name	Description	Source	Beginning of Sample/ frequency	Source of data forecasts (for exogenous)	Seasonality test	Stationarity test		
						ADF Prob (level)	ADF Prob (1st diff)	ADF Prob (yoy diff)
	Other indicators							
EC_P_EGG	Eggs prices in EU	EU commission	2004/m	MPEAD assessments	P	0.05	0.00	0.00
DISEL_P_UAH	Diesel prices in Ukraine	NBU/ web-scraping	2005/m		NP	0.65	0.00	0.00
ER_EU_USD	Euro/USD exchange rate	Reuters	2001/m		NP	0.12	0.00	0.00
ER_M	UAH/USD exchange rate	NBU	2001/m		NP	0.98	0.00	0.00
NWAGE	Nominal average wage	SSSU	2005/m		P	1.00	0.00	0.00
MINWAGE	Nominal minimum wage	SSSU	2005/m		NP	1.00	0.00	0.15
RMC_C	Real monetary costs for Core CPI	NBU (QPM)	2004/q					
PPI_EUD	Non-durable consumer goods (EU28 PPI)	OECD	2010/m		P	0.52	0.00	0.01
AGR	Average sale prices for agricultural products	SSSU	2005/m		Probably NP	0.79	0.00	0.00
	Production of							
PR_EGG	eggs		2001/m	MPEAD assessments	P	0.84	0.18	0.80
PR_MT	meat				P	1.00	0.00	0.00
PR_MILK	milk				P	0.37	0.00	0.00
	Harvest of							
CRL_H	cereals	SSSU	2001/y	MPEAD assessments				
FRT_H	fruits							
POTATO_H	potato							
SGR_H	sugar							
VGT_H	vegetables							
OIL_H	sunflower seeds							

Note: Results of the seasonality test for the Combined test, indicating whether there is the presence of identifiable seasonality. P stands for present, NP – not present. It is recommended that a series is adjusted in the cases of P and Probably NP, and not adjusted in the case of NP.
Stationarity test shows the p-values of ADF test for levels, 1st differences and yoy changes.
Quarter and year frequency data converted into monthly frequency using cubic spline.

Table A.3. Components of the CPI and the PPI

Name	Description	Source	Beginning of Sample/ frequency	Seasonality test	Stationarity test		
					ADF Prob (level)	ADF Prob (1st diff)	ADF Prob (yoy diff)
	CPI						
CPI_F	RFPI	SSSU	2004/m	P	0.86	0.00	0.07
CPI_MT	meat			P	0.91	0.00	0.00
CPI_MLK	milk			P	0.88	0.00	0.02
CPI_EGGs	eggs			P	0.73	0.00	0.00
CPI_FRT	fruits			P	0.74	0.00	0.01
CPI_VGT	vegetables			P	0.13	0.00	0.00
CPI_SGR	sugar			NP	0.88	0.00	0.00
CPI_CRL	cereals			NP	0.91	0.00	0.00
CPI_OIL	Sunflower oil, CPI			NP	0.86	0.00	0.00
CPI_FUEL	Fuel component of CPI		2004/m	NP	0.53	0.00	0.02
CPI_C	Core CPI		2012/m	NP	0.74	0.01	0.05
CPI_FC	processed food			NP	0.79	0.00	0.03
CPI_SRV	services			NP	0.98	0.01	0.03
CPI_CLSH	clothes and shoes			P	0.52	0.02	0.64
CPI_OTHR	others			NP	0.50	0.01	0.35
CPI_FOTHR	processed food and others			NP	0.63	0.01	0.08
	PPI						
PPI_F	processed food		2012/m	NP	0.79	0.00	0.00
PPI_MT	meat			NP	0.53	0.00	0.01
PPI_MLK	milk			NP	0.68	0.00	0.00
PPI_CRL	cereals			NP	0.88	0.00	0.01
PPP_SGR	sugar			Probably NP	0.84	0.00	0.00
PPI_CLSH	clothes and shoes			NP	0.96	0.00	0.03
PPI_COMP	computers			NP	0.57	0.00	0.58
PPI_AUTO	cars			NP	1.00	0.00	0.17

Note: Results of the seasonality test for the Combined test, indicating whether there is the presence of identifiable seasonality. P stands for present, NP – not present. It is recommended that a series is adjusted in the cases of P and Probably NP, and not adjusted in the case of NP.
Stationarity test shows the p-values of ADF test for levels, 1st differences and yoy changes.

Table A.4. AR Models for the RFPI and Core Inflation

Forecasted Variable	Lags	AR coefficient	Sample	S.E.
CPI_MT	1	0.65	2005m3-2021m12	1.09
CPI_MLK	2	0.69	2005m4-2021m12	0.86
CPI_EGGs	2	-0.11	2005m4-2021m12	8.76
CPI_FRT	1	0.36	2005m3-2021m12	3.93
CPI_VGT	1	0.31	2005m3-2021m12	6.39
CPI_SGR	1	0.38	2005m3-2021m12	5.21
CPI_CRL	4	0.50	2005m6-2021m12	3.10
CPI_FC	2	0.67	2014m4-2021m12	0.26
CPI_SRV	1	0.40	2014m3-2021m12	0.31
CPI_CLSH	1	0.29	2014m3-2021m12	0.61
CPI_OTHR	1	0.52	2014m3-2021m12	0.33

Table A.5. Data Sets for Bivariate VARs and FAVAR Models for Each Frecasted Variable

Forecasted variable	Data set
CPI_MT	fao_p_mt, fao_p_chck, fao_p_prk, fao_p_bf, fao_p_crl, imf_p_wht imf_p_brl, imf_p_soy, imf_p_chck, imf_p_oil, wb_p_crl, wb_p_fuel_l
CPI_MLK	fao_p_dai, fao_p_crl, imf_p_wht, imf_p_brl, imf_p_soy, imf_p_oil, wb_p_crl, wb_p_fuel"
CPI_EGGs	fao_p_crl, imf_p_chck, imf_p_wht, imf_p_brl, imf_p_soy, imf_p_oil, wb_p_chck, wb_p_crl, wb_p_fuel, ec_p_egg
CPI_FRT	fao_p_f imf_p_f imf_p_bn imf_p_orn wb_p_f wb_p_orn wb_p_bn wb_p_fuel wb_p_frt
CPI_VGT	fao_p_f, imf_p_f, wb_p_f, wb_p_fuel, wb_p_frt
CPI_SGR	fao_p_sgr, imf_p_sgr, wb_p_sgr, wb_p_fuel, wb_p_frt
CPI_CRL	fao_p_crl, imf_p_wht, imf_p_brl, wb_p_crl, wb_p_fuel, wb_p_frt
CPI_FC	agr, cpi_f, cpi_fuel, disel_p_uah, er_eu_usd, er_m, fao_p_f, imf_p_f, minwage, nwage, ppi_eund, ppi_f, rmc_c, wb_p_f
CPI_SRV	cpi_fc, er_eu_usd, er_m, minwage, nwage, rmc_c
CPI_CLSH	er_eu_usd, er_m, minwage, nwage, ppi_eud, rmc_c
CPI_OTHR	er_eu_usd, er_m, minwage, nwage, ppi_eud, rmc_c

Table A.6. Lag Length Criteria for BVAR Models

	LR	FPE	AIC	SC	HQ
BVAR_RFPI	4	2	2	1	1
BVAR_3CORE	2	2	2	1	1
BVAR_4CORE	2	2	2	1	1

Note: numbers in the Table A.6 indicate lag order selected by the criterion:
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

Table A.7. Grid Search

	Minimum value	Maximum value	Step size
Autoregressive coefficient	0.20	1.00	0.10
Overall tightness (λ1)	0.05	0.20	0.01
Cross-variable weighting (λ2)	0.10	1.00	0.10
Lag decay (λ3)	0.10	2.00	0.20
Exogenous variable tightness (λ4)	100	1000	100

Table A.8. BVAR Model Specifications

Endogenous variables	Exogenous variables	lags	Sample	Hyper parameters	total number of iterations:	burn-in iterations:
7 RFPI components (CPI_MT, CPI_MLK, CPI_EGGs, CPI_FRT, CPI_VGT, CPI_SGR, CPI_CRL)	ER_M(-1), FAO_P_F(-1)	2	2005 m1-2021m12	Mu1: 0.5, λ1: 0.05, λ2: 1, λ3: 1, λ4: 100	10000	5000
3 core CPI components (CPI_FOTHR, CPI_SRV,CPI_CLSH)	NWAGE, ER_M(-1)	2	2012m1-2021m12	Mu1: 0.5, λ1: 0.05, λ2: 1, λ3: 1, λ4: 100	10000	5000
4 core CPI components (CPI_FC,CPI_SRV,CPI_CLSH, CPI_OTHR)	NWAGE, ER_M(-1)	2	2012m1-2021m12	Mu1: 0.4, L1: 0.11, L2: 1, L3: 1, L4: 100	10000	5000

Table A.9. Equations for Components of the RFPI and Core CPI

eq name	SE	Coin-tegra-tion test	AR(1)	COI	factors	factors	factors	factors	factors
crl			D(CPI_CRL_L(-1))	CRL_COI2(-1)	D(ER_M_L(-1))	D(FAO_P_CRL_L)	D(WB_P_FRT_L(-1))	D(CRL_H_LQS(-5))	
	2.60	0.07	0.5	-0.06	0.40	0.10	0.10	-0.20	
mt			D(CPI_MT_L(-1))	MT_COI2(-1)	D(ER_M_L(-1))	D(FAO_P_MT_L(-1))	D(CPI_OIL_L)	C	
	0.90	0.06	0.60	-0.05	0.10	0.01	0.10	0.20	
mlk			D(CPI_MLK_L(-1))	MLK_COI2(-1)	D(ER_M_L(-1))	D(FAO_P_DAI_L(-2))	D(CPI_OIL_L)	C	
	0.80	0.08	0.60	-0.03	0.04	0.04	0.10	0.30	
egg			D(CPI_EGG_L(-1))	EGG_COI2(-1)	@MOVAV(D(ER_M_L(-1)),3)	D(EC_P_EGG_L(-0))	D(CPI_OIL_L(-2))	D(PR_EGG_L(-1))	D(PR_EGG_L(-2))
	7.80	0.00	0.20	-0.32	0.30	0.20	0.40	-0.30	-0.40
frt			D(CPI_FRT_L(-1))	FRT_COI2(-1)	D(ER_M_L(-1))	D(DISEL_P_UAH_L)	D(IMF_P_BN_L(-1))	D(FRT_H_LQS(-2))	
	3.50	0.13	0.30	-0.07	0.30	0.20	0.10	-0.30	
vgt			D(CPI_VGT_L(-1))	VGT_COI2(-1)	D(ER_M_L(-1))	D(VGT_H_LQS(-2))	D(POTATO_H_LQS)	D(FAO_P_F_L)	@SEAS(7)
	5.80	0.02	0.30	-0.14	0.20	-1.40	-1.00	0.30	4.40
sgr			D(CPI_SGR_L(-1))	SGR_COI2(-1)	D(ER_M_L(-1))	@MOVAV(D(FAO_P_SGR_L),3)	D(DISEL_P_UAH_L)	D(SGR_H_LQS(-7))	
	4.0	0.00	0.3	-0.10	0.3	0.2	0.3	-0.4	
fc			D(CPI_FC_L(-1))	FC_COI2(-1)	D(ER_M_L(-1))	D(ER_M_L(-0))	D(CPI_F_L)		
	0.50	0.02	0.60	-0.05	0.20	0.03	0.10		
srv			D(CPI_SRV_L(-1))	SRV_COI2(-1)	@MOVAV(D(ER_M_L(-0)),2)	@MOVAV(D(NWAGE_L),5)			
	0.30	0.02	0.60	-0.03	0.10	0.20			
clsh			D(CPI_CLSH_L(-1))	CLSH_COI2(-1)	@MOVAV(D(ER_M_L(-1)),6)	D(@MOVAV(COVDUM,4))	RMC_C(-3)		
	0.90	0.00	0.10	-0.25	0.20	-3.90	0.10		
othr			D(CPI_OTHR_L(-1))	OTHR_COI2(-1)	D(ER_M_L(-1))	D(ER_M_L(0))	D(DISEL_P_UAH_L)		
	0.40	0.00	0.60	-0.02	0.10	0.10	0.00		

Note: Co-integration test shows z-statistic of the Engle-Granger Co-integration test for the long-run equation. Value less than 0.05/0.10 rejects the null hypothesis of no co-integration at a significance level of 5/10%%.

APPENDIX B. FIGURES

	Mean	St. dev.
CPI	13.4	14.4
Fuel	13.7	21.6
Food and non-alcoholic beverages	12.4	14.0
Rice	15.8	33.4
Bread	17.4	15.9
Pasta	12.9	16.6
Beef and veal	12.3	12.7
Pork	10.4	15.1
Poultry	12.7	15.3
Other meats	11.8	9.0
Fish and seafood	11.7	19.6
Fresh whole milk	12.7	9.1
Yoghurt	12.7	7.9
Cheese and curd	12.0	7.2
Eggs	16.8	34.4
Butter	13.6	9.2
Margarine and other vegetable fats	14.1	14.2
Olive oil	12.6	21.8
Other edible oils	18.0	29.8
Fruit	14.6	29.7
Citrus fruits	10.5	31.5
Banana	11.1	30.2
Apples	18.5	43.9
Dried fruits	15.2	34.2
Vegetables	6.6	25.3
Cabbage	28.5	84.1
Cucumbers, tomatoes, pepper, zuccini	-5.1	28.7
Potatoes	9.9	40.4
Preserved or processed vegetables	12.2	13.9
Potatoes	19.0	40.7
Borsch vegetables	24.7	57.8
Sugar	17.9	28.5
Chocolate	14.0	27.2
Coffee, tea and cocoa	13.5	24.8
Mineral waters, soft drinks, fruit and vegetable juices	11.1	9.4
Alcoholic beverages	12.7	10.3
Tobacco	22.2	10.3
Clothing	4.5	11.7
Other articles of clothing and clothing accessories	8.4	10.1
Cleaning, repair and hire of clothing	13.1	6.5
Shoes and other footwear	5.1	13.9
Repair and hire of footwear	11.5	6.9
HOUSING, WATER, ELECTRICITY, GAS AND OTHER FUELS	26.6	40.7
Actual rentals for housing	7.6	4.3
Imputed rentals for housing	2.5	3.3
Maintenance and repair of the dwelling	11.1	11.4
Water supply and miscellaneous services relating to the dwelling	22.3	18.2
Electricity	22.4	26.4
Gas	57.3	125.7
Solid fuels	6.4	11.0
Heat energy	26.2	33.4
Furniture and furnishings, carpets and other floor coverings	9.3	11.8
Household textiles	9.2	14.7
Household appliances	8.8	15.4
Glassware, tableware and household utensils	9.7	15.1
Tools and equipment for house and garden	9.0	15.5
Goods and services for routine household maintenance	9.3	16.3
Medical products, appliances and equipment	12.9	16.2
Out-patient services	12.4	6.4
Hospital services	10.8	4.9
Purchase of vehicles	13.8	26.5
Operation of personal transport equipment	13.2	17.9
Fuels and lubricants for personal transport equipment	13.7	21.6
Passenger transport by railway	11.3	7.7
Passenger transport by road	15.3	11.0
Passenger transport by air	7.9	10.7
Postal services	21.4	27.6
Telephone and telefax equipment	-0.3	13.6
Telephone and telefax services	8.3	6.6
Audiovisual, photographic and information processing equipment	2.8	14.7
Other recreational items and equipment, gardens and pets	11.5	18.9
Recreational and sporting services	8.8	3.5
Cultural services	12.6	6.5
Newspapers, books and stationery	9.2	12.0
Package holidays	16.3	27.7
Pre-primary and primary education	18.9	14.7
Secondary education	12.3	5.4
Tertiary education	10.2	5.0
Education not definable by level	9.6	3.6
Restaurants, cafés and the like	10.4	6.5
Canteens	13.4	9.3
Accommodation services	8.0	5.2
Hairdressing salons and personal grooming establishments	10.7	4.0
Other appliances, articles and products for personal care	11.2	17.6
Personal effects n.e.c.	8.0	14.4
Insurance	10.5	13.3
Financial services n.e.c.	8.6	7.2
Other services n.e.c.	10.1	5.7

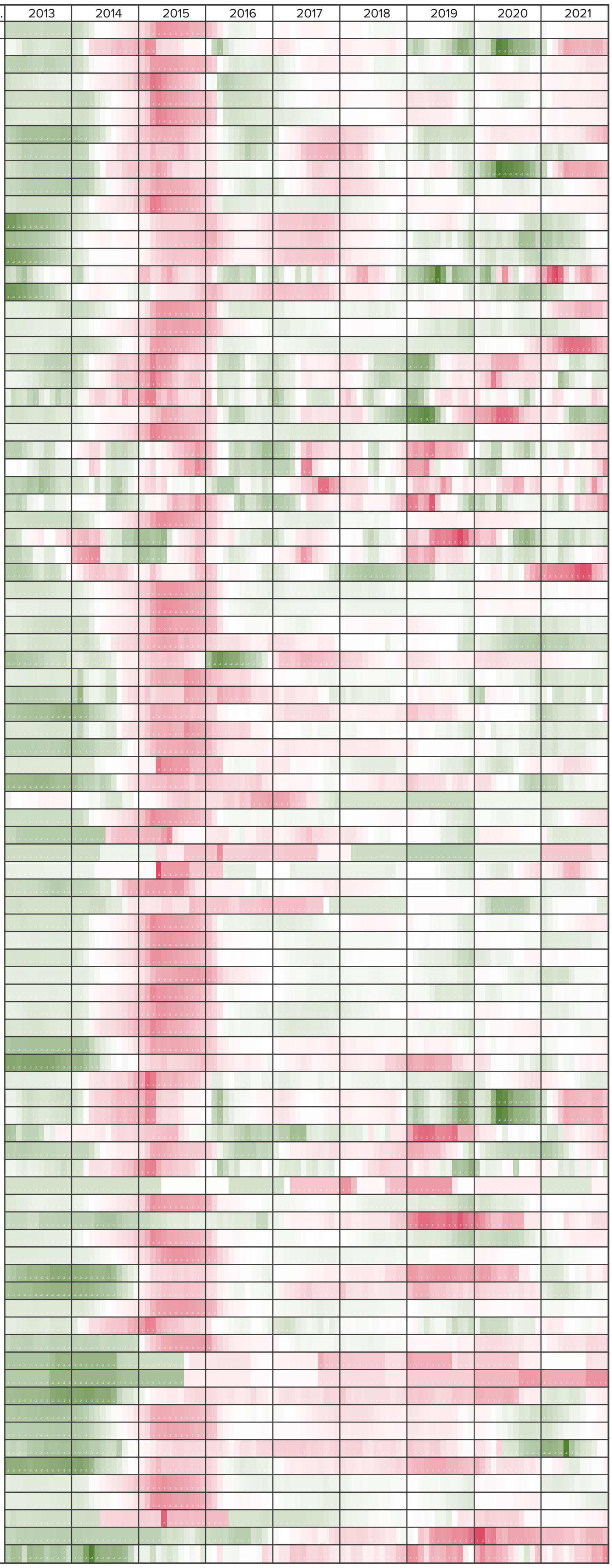


Figure B.1. Heat Map

Figure B.2. Consumption Balances

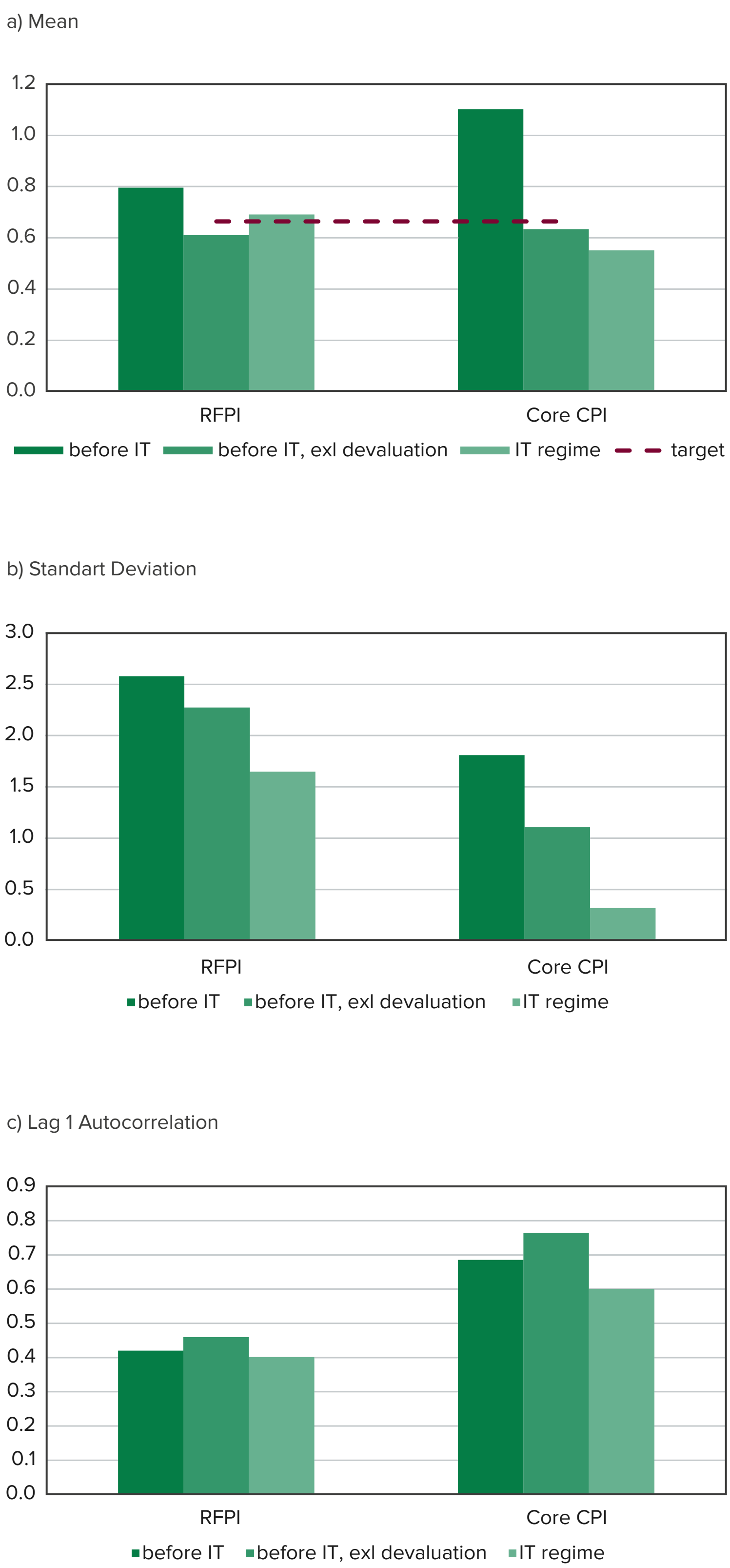


Figure B.3. Statistical Properties of the Data on RFPI and CPI for pre-IT- and IT Regime Data Samples

Note: Data samples for RFPI and Core CPI start in 2005m1 and 2012m1 respectively.

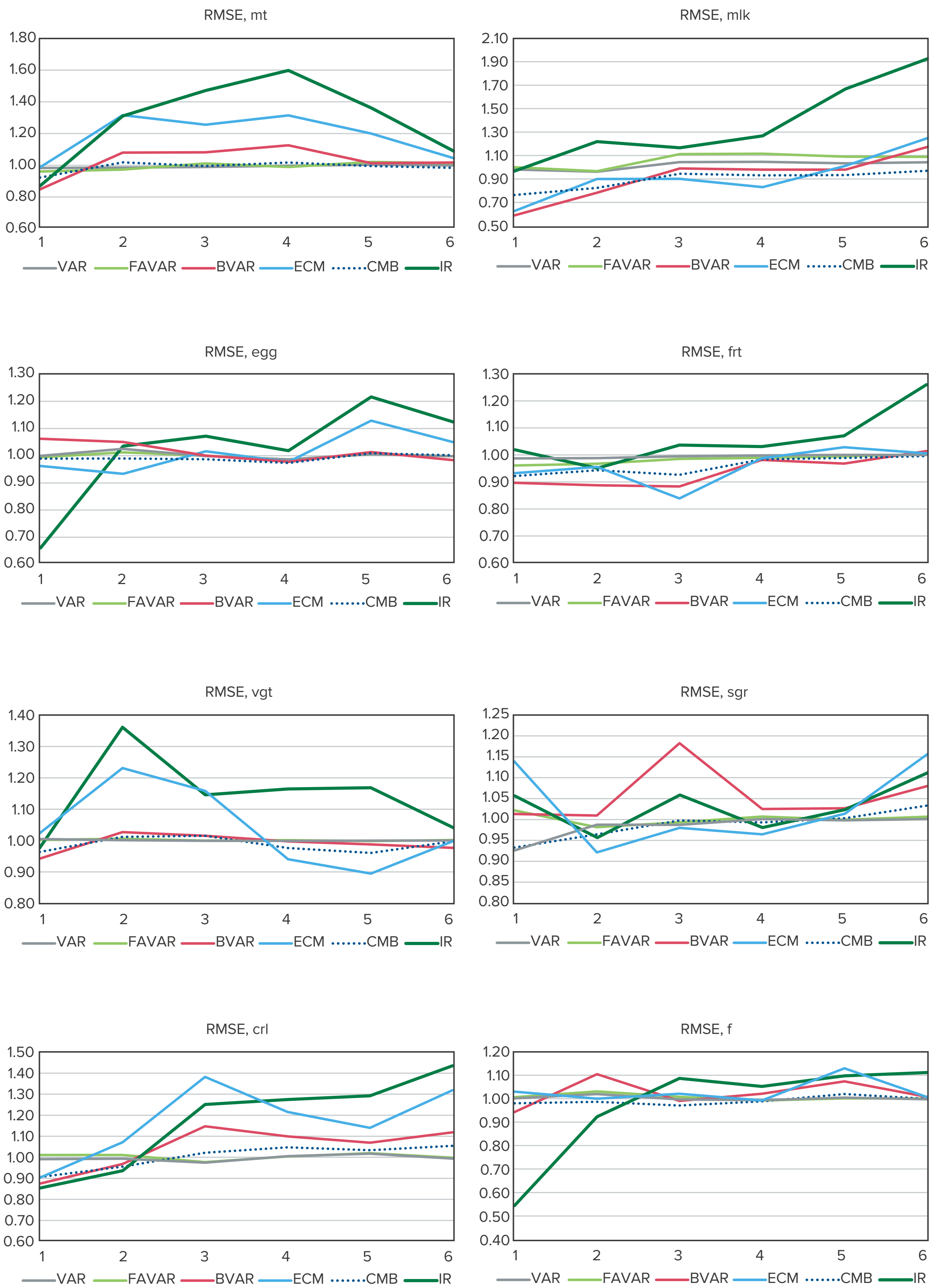


Figure B.4. RMSE, Relative to AR's Model RMSE, the RFPI

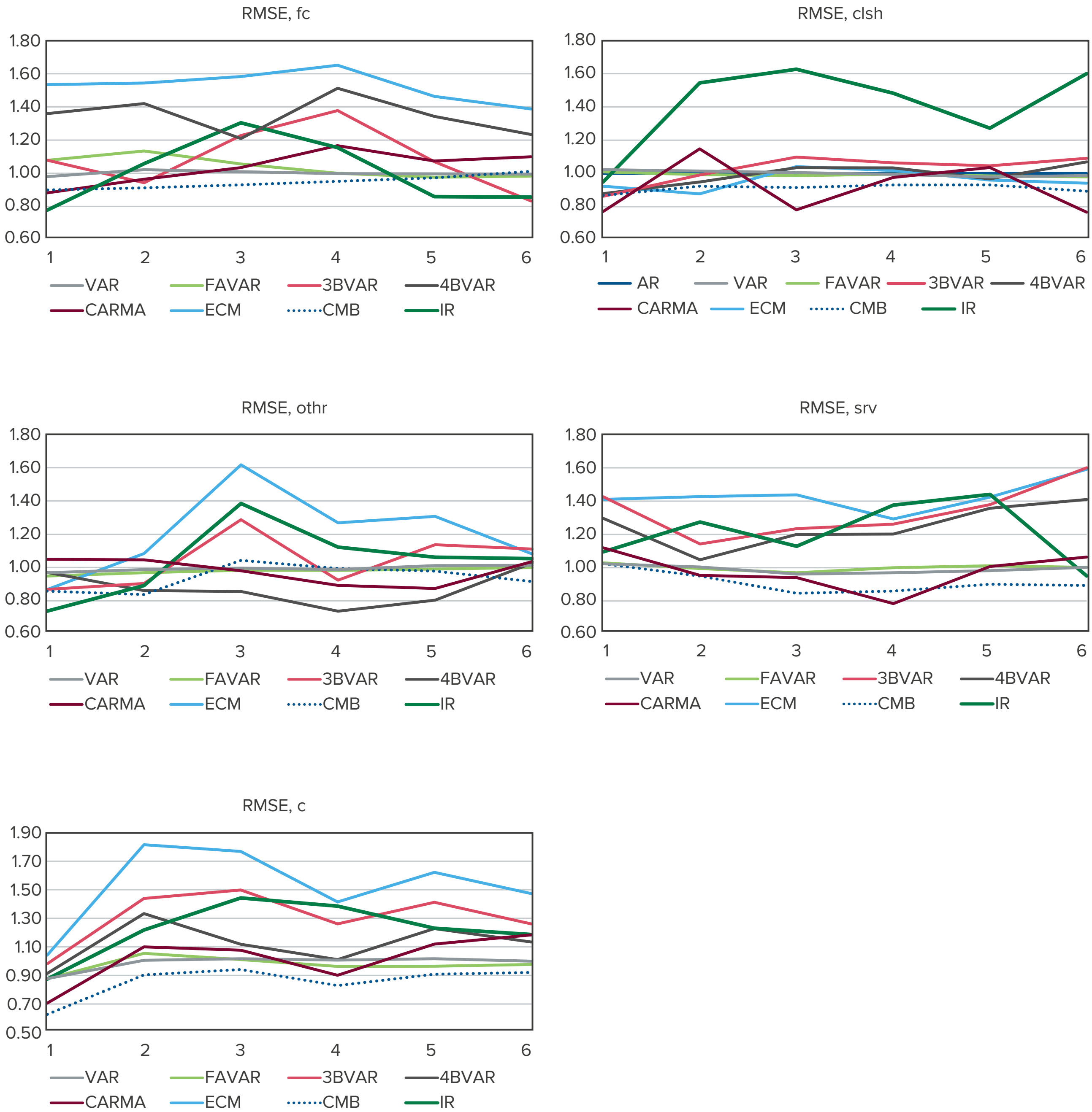


Figure B.5. RMSE, Relative to AR's Model RMSE, core CPI

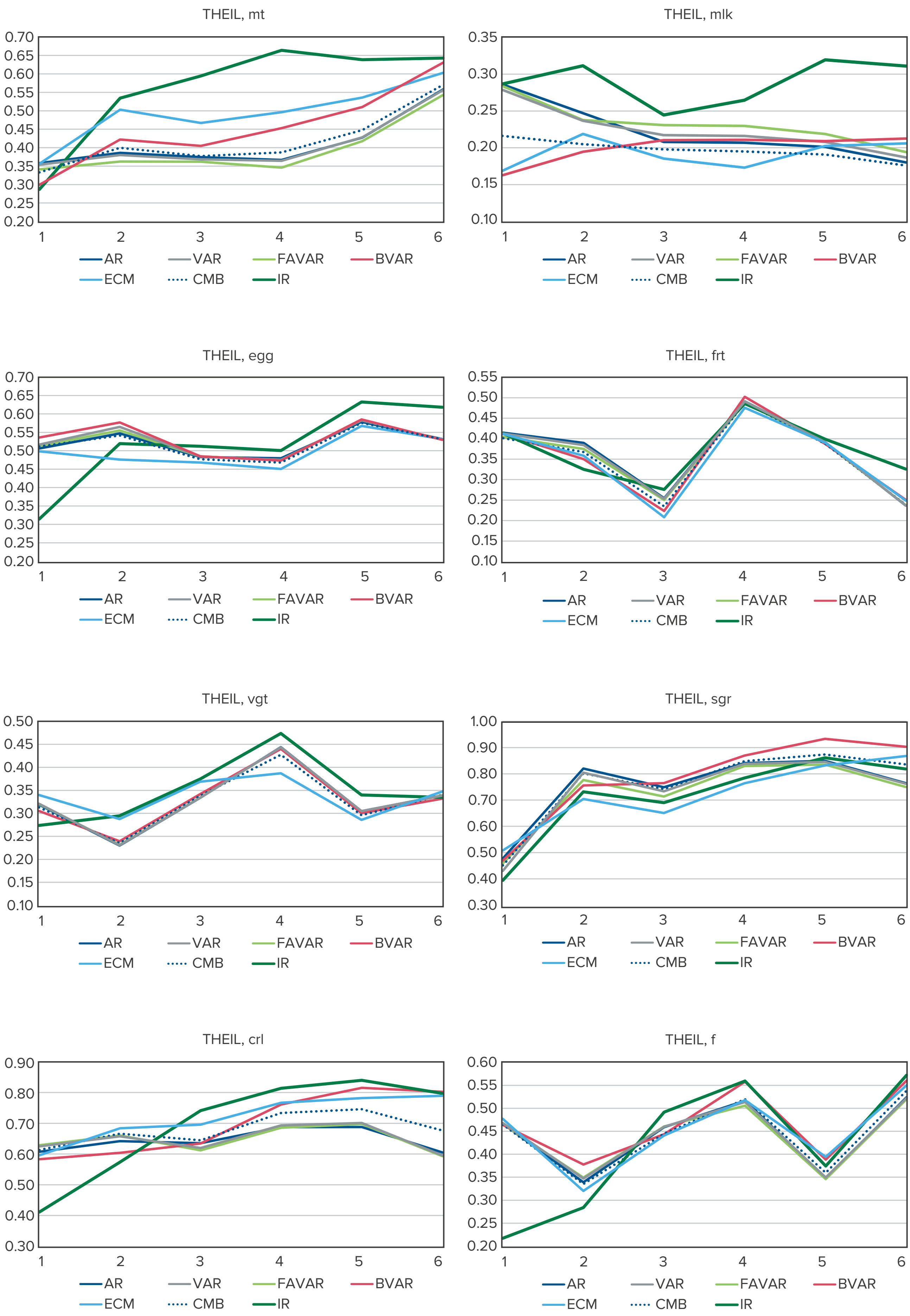


Figure B.6. Theil Index, the RFPI

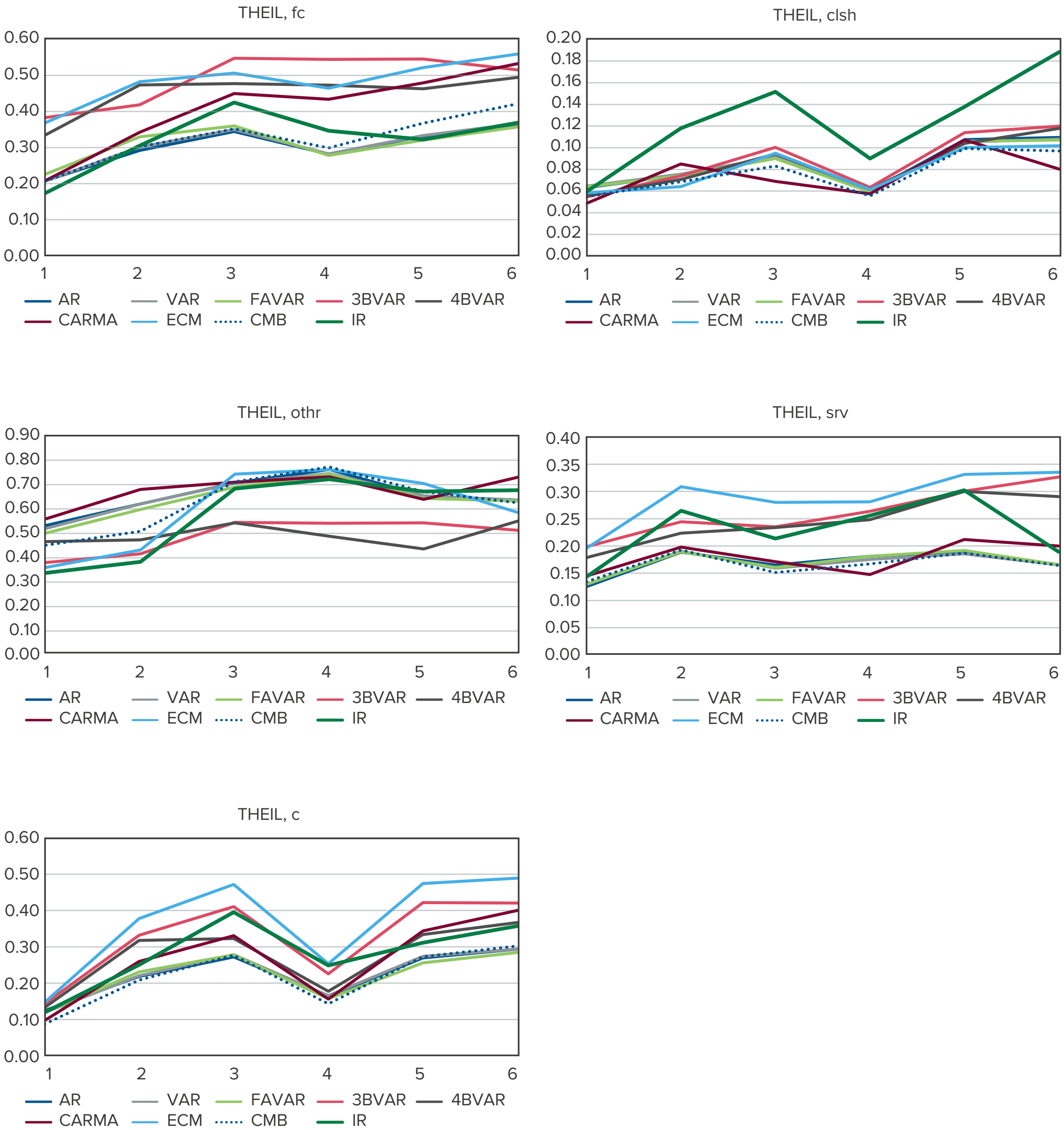


Figure B.7. Theil Index, core CPI

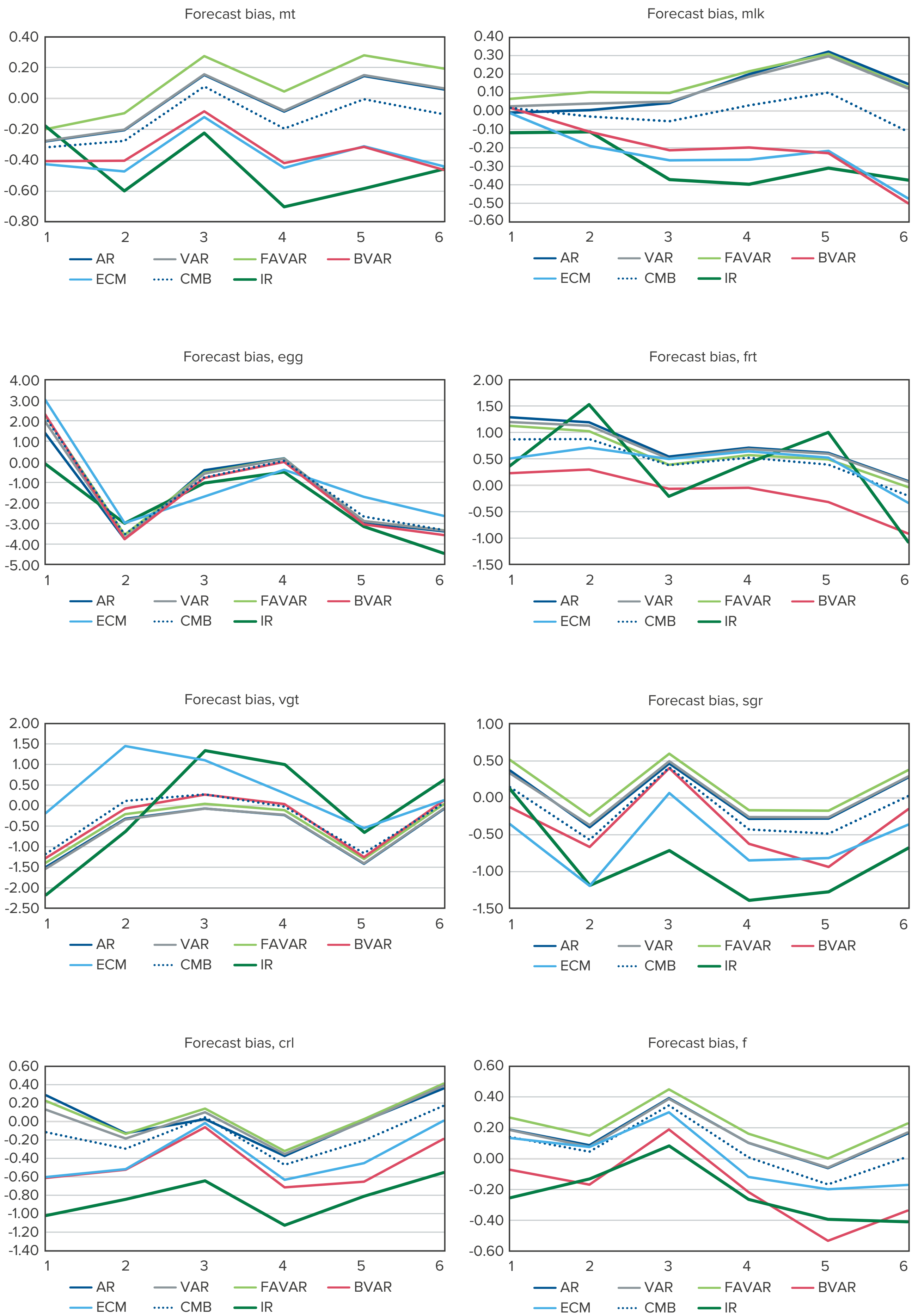


Figure B.8. Forecast Bias, the RFPI

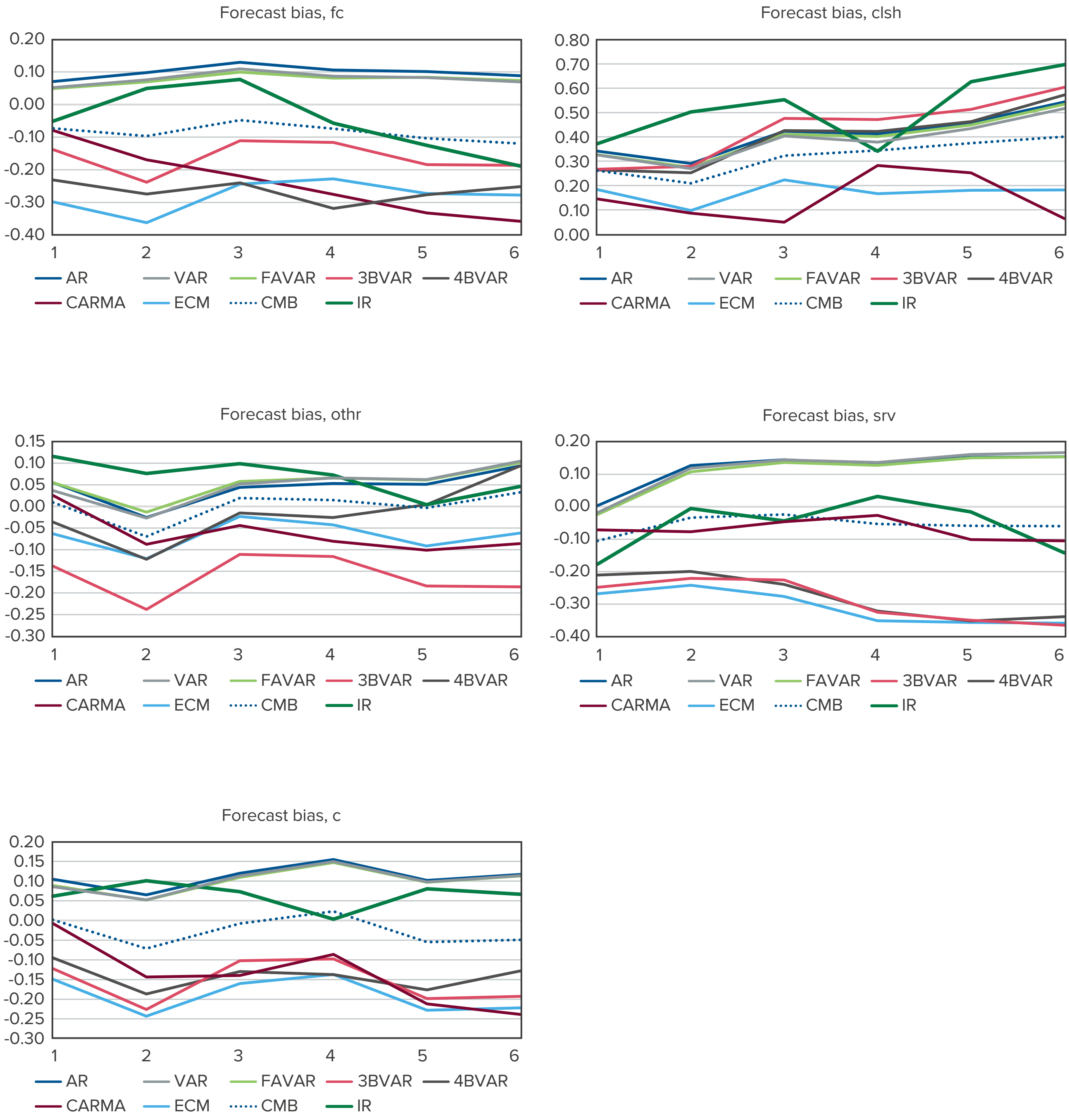


Figure B.9. Forecast Bias, core CPI

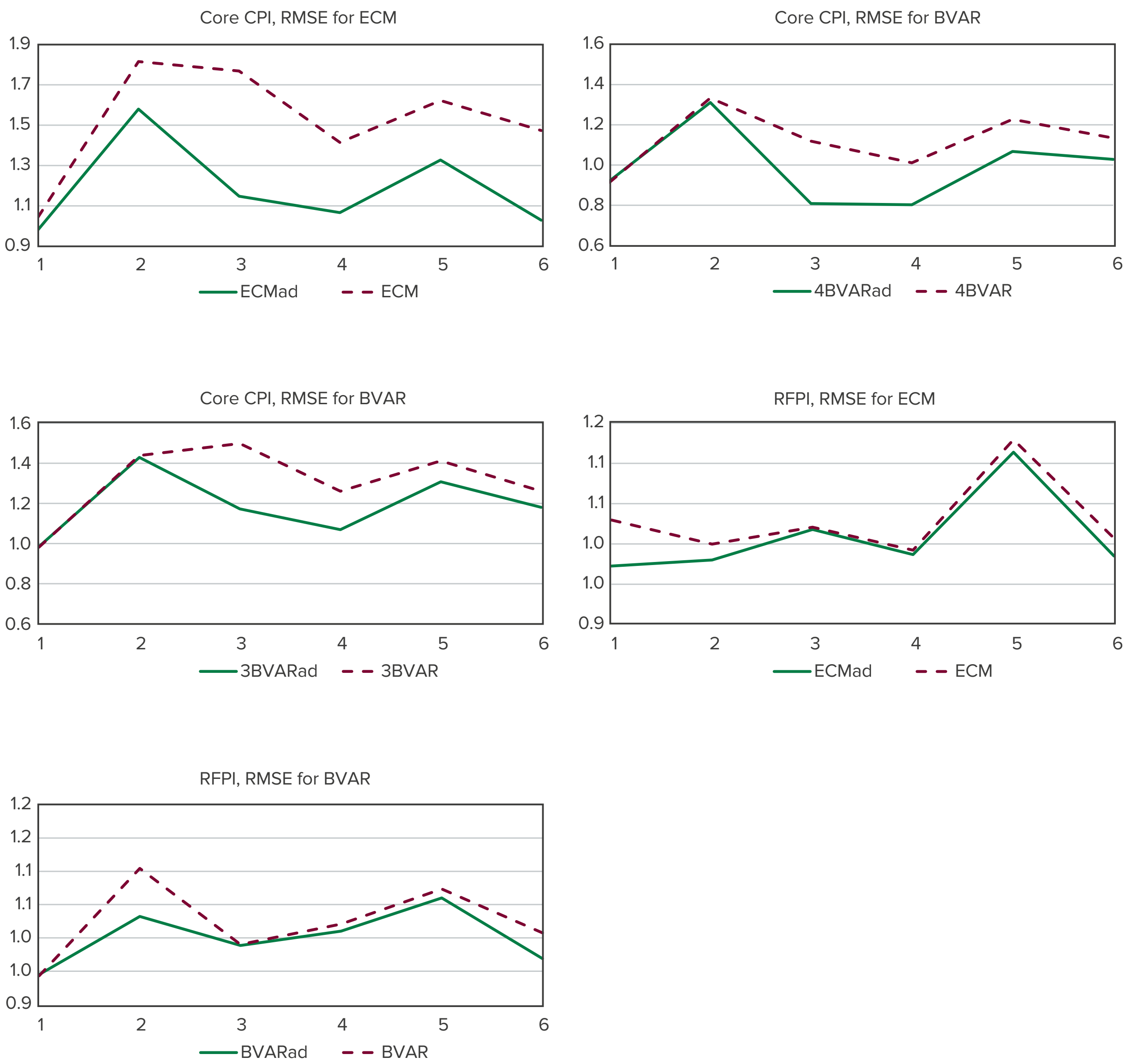


Figure B.10. RMSE, Relative to AR's Model RMSE (models with assumptions and actual data)

Note: ad stands for actual data for exogenous variables instead of assumptions.

Visnyk of the National Bank of Ukraine, 2021, No. 252, pp. 37–55

IDENTIFYING INSURANCE COMPANIES' BUSINESS MODELS IN UKRAINE: CLUSTER ANALYSIS AND MACHINE LEARNING

OLEKSANDR TARNAVSKYI[ab], VIKTOR KOLOMIIETS[a]

[a]National Bank of Ukraine
[b]National University of Kyiv-Mohyla Academy
E-mail: Oleksandr.Tarnavskyi@bank.gov.ua
Viktor.Kolomiiets@bank.gov.ua

Abstract This study examines the performance of the nonlife insurance companies that operated in Ukraine in 2019–2020. Specifically, we employ a set of clustering techniques, e.g. the classic k-means algorithm and Kohonen self-organizing maps, to investigate the characteristics of the Retail, Corporate, Universal (represented by two clusters), and Reinsurance business models. The clustering is validated with classic indicators and a migration ratio, which ensures the stability of the clusters over time. We analyze the migration of companies between the identified clusters (changes in business model) during the research period and find significant migration between the Reinsurance and Corporate models, and within the Universal model. Analysis of the data on the termination of the insurers' ongoing activity allows us to conclude that companies following the Universal business model appear to be the most financially stable, while their peers grouped into the Reinsurance cluster are likely to be the least stable. The findings of this research will be valuable for insurance supervision and have considerable policy implications.

JEL Codes G22, D22

Keywords cluster analysis, neural networks, business model, insurance

1. INTRODUCTION

Starting from 1 July 2020, the NBU began implementing the reform of the financial sector, extending the requirements of transparency, reliability, and efficiency to the nonbanking financial sector. The primary goal of the reform is to improve the quality of insurance services and protect policyholders' interests.

Effective supervision, control and implementation of reforms on the nonbanking financial market require an understanding of the market structure and how its participants conduct their business. For example, different business models may have quite different risk profiles. The identification of homogeneous groups of companies with similar risks allows for a more detailed analysis of the stability and solvency of insurance companies, and the effective prediction of crisis events. This research aims to contribute to the understanding of the Ukrainian insurance market's structure as well as its companies' operational and risk profiles. It identifies Ukrainian insurers' business models and their key features using quantitative indicators to assist supervision of the insurance market.

To achieve our goal, we attempted to answer these questions: Can homogenous and stable groups (business models) of insurance companies be identified through the analysis of regulatory data? What are the key characteristics of these business models? How did companies change their business models during the research period? Can certain business models be associated with increased risks?

For this paper, we conducted a cluster analysis of the Ukrainian insurance market to determine the business models used by insurers. We apply a number of clustering methods, including hierarchical, nonhierarchical, and machine-learning ones. We identify five clusters with the k-means method that correspond to four business models – Corporate, Retail, Universal (divided into two clusters), and Reinsurance. Before applying clustering algorithms, an artificial cluster named "Inactive" was formed (it comprised companies that were not very active or did not engage in insurance activities at all, but had a license and were present in the sample). A number of calculated coefficients, namely the migration ratio and the silhouette coefficient helped us assess the quality of our research.

We analyze the business models by using both the features by which the clustering took place and additional variables that are not used in the clustering algorithms. Thus, the companies with the Corporate business model mostly ensure legal entities, while those using the Retail model, on the contrary, work with individuals. Companies with the Universal business model tend to use sales offices widely as a sales channel, while those with the Reinsurance business model do not use them at all. The further text provides a more detailed description of the clusters.

Next, we show how insurers migrated between the clusters in the period from 2019 to 2020. We observed significant migration between the Reinsurance and Corporate business models and within the Universal business model. We also find significant migrations to the artificially created Inactive cluster, i.e. in cases when insurers terminated their insurance activity. Based on these migrations, it is possible to empirically draw conclusions about the riskiness of a particular business model. Thus, the largest share of the companies that left the market during the studied period belonged to the cluster using the Reinsurance business model; more than half of the companies in this cluster ceased operations in 2020. Significant migration to the Inactive group was also observed in relation to entities using the Corporate and Retail business models.

The paper is structured in the following way. The second section provides an overview of the relevant literature. The third section highlights the methodology, data, and software used in this analysis. The fourth section presents the key findings of the research and shows the riskiness of each of the identified business models. The fifth section briefly summarizes the results of the research and outlines promising directions for future research.

2. LITERATURE REVIEW

The development of research on banks' business models was facilitated by the Basel regulatory framework and the implementation of the Supervisory Review and Evaluation Process (SREP). Studying insurance companies' business models are also seen as a promising area of research.

Most of the existing work in the area is aimed at the analysis and segmentation of insurers' clients. Research by Wang and Keogh (2008), and Zaqueu (2019) is devoted to a clustering analysis for target group identification. Clustering techniques were used to identify customer profiles based on datasets derived from policy transactions and policyholder information. Wang and Keogh used self-organizing maps (SOMs) and the k-means algorithm. The k-means method was also used in a publication by Abolmakarem et al. (2016) that used segmentation to identify the most profitable customers for companies. Velykoivanenko and Beschastna (2018) use SOMs to rate insurance companies in terms of their financial performance into three groups. Then they combine clustering results with experts' ratings to arrive at integral indicator of company financial stability.

Most researchers use a combination of the two clustering methods. In particular, a study by Kramarić et al. (2018) groups European insurance companies into seven clusters. Unlike previous studies, her research groups companies, rather than their customers, into clusters. Using a combination of hierarchical clustering and k-means clustering, 119 insurers are divided into seven groups by country of origin and company type. Bach et al. (2020) use a Kohonen map in combination with a hierarchical cluster analysis to investigate fraud risks in the leasing industry. The neurons of the Kohonen self-organizing map are combined into five clusters using Ward's method, after which the risk characteristics of the clusters are analyzed.

A study by Ahmar et al. (2018) is an example of a cluster analysis outside of the financial sector. In their study, they use the k-means method to group the provinces of Indonesia. Such a grouping, according to the authors, should help to classify the regions of that country so that social problems can be tackled more effectively. Abbas et al. (2020) compared the k-means and k-medoids methods using data on women during pregnancy. The k-medoids method is inherently very similar to the k-means method, so they are often used in combination. The authors show that the k-medoids method is more accurate than k-means for specific data.

Rashkovan and Pokidin (2016) identified business models of banks in Ukraine using a Kohonen self-organizing map, and drew parallels between business models and indicators of various types of risk to which a bank may be exposed. In terms of methodology, our work is very similar to this study. However, unlike Rashkovan and Pokidin, who base their research findings on a Kohonen self-organizing map, we use this method in addition to the k-means method. Ferstl and Seres (2012) also used a cluster analysis to identify business models. Unlike previous researchers, they utilized the k-means algorithm based on the use of the Mahalanobis distance. Their work identifies five business models of banks, based on five indicators.

Most authors use the simplest clustering models, including the k-means method. In our work, we intend to develop a methodology that helps to determine the distribution of companies according to their business model. To implement the research, we used a wide set of clustering tools, but the conclusions were based on the k-means method. Kohonen self-organizing maps are a convenient visualization tool in our work. This research for the first time evaluates the quality of clustering through the use of the migration ratio – an indicator that characterizes the stability of clusters.

3. DATA AND METHODOLOGY

Description of the Data Used

To conduct a cluster analysis of insurance companies, we gathered data from the regulatory reports of 247 Ukrainian insurers for two years, from 2019 to 2020. During the research period, the number of active insurers decreased significantly. Thus, as of the end of 2020, the database consisted of entries for 185 insurers. The data were taken from a regulatory database.

To identify a business model, we aimed to select indicators that would answer the following questions about an insurance company:

Who are its target customers?

What types of insurance does it focus on, and how explicitly?

What sales channels does it use?

Table 1. Indicators Used for Clustering

No.	Indicator	Variable name	Formula
1	Return on assets	ROA	Net income / Total assets
2	Number of offices	Offices	Total number of used offices that are not the head office
3	Share of premiums from mandatory types of insurance in the total amount of collected premiums	% of mandatory	Amount of premiums from mandatory types of insurance / Total amount of premiums
4	Share of premiums from legal entities	Corporate	Amount of premiums from legal entities / Total amount of premiums
5	Share of inwards (assumed) reinsurance in premiums	Re-to-premiums	Amount of reinsurance premiums / Total amount of premiums

We selected indicators that would simultaneously help to find the answers to these questions, and which would allow the insurers to be optimally sorted into clusters according to certain quantitative metrics – the ratio of migration and the silhouette coefficient (as described further in the text). According to the values of the metrics obtained, an optimal set of indicators for clustering was selected. Next, a different set of indicators that allowed for a broader description of the clusters and the risks inherent in them was chosen separately. These indicators were not included in the model, as partitioning based on them led to worsened clustering quality. Rather, they were used for the broader characterization of the identified clusters. Table 1 and Table 2 describe these two groups of indicators.

After calculating the indicators, their values were standardized (to mean 0 and unit variance). This is necessary because of the clustering algorithms' sensitivity to variance in the data. We also detected outliers in the data. Observations that were more than three standard deviations away from the mean were rounded to the nearest value within the range of three standard deviations. The distribution of the observations before and after this procedure is given in Figure B.2.

Companies whose total premiums for the reporting period did not exceed UAH 5 million were grouped into an artificial cluster (group) named "Inactive." Such companies in 2020 accounted for less than 1% of total market share (in premiums).

Further are the descriptive statistics of the data for 2020.

We can see that most companies in 2020 were slightly profitable or unprofitable, in contrast to the higher levels of profitability observed in previous years (Table A.1). The reason for such a drop in profitability could be attributed to an increase in health insurance claims as an effect of the COVID-19 pandemic. Indeed, loss reserves for health insurance increased significantly in the periods that there were peaks in new COVID-19 cases.

Most of the market focuses its activities on low-priced contracts, which are most likely to be sold to individuals. The median "average check" is about UAH 3,000.

According to our data, most companies did not have sales offices. On the one hand, this may indicate the predominance in the market of business models that do not use offices as a sales channel. On the other hand, such a strong skew indicates a possible risk that some companies are misreporting. We are not able to verify this. We assume that any misreporting companies are evenly distributed across the clusters and do not significantly shift cluster centers. It is worth noting that a similar structure is observed for the data for 2019.

There is also a high concentration of one type of insurance on the market. About half of the companies had a share of premiums from one of the groups of insurance types that exceeded 60%. This indicates the presence of

Table 2. Indicators Used for Additional Description of Clusters

No.	Indicator	Variable name	Formula
1	Ratio of the share of reinsurers in insurance reserves to the total amount of insurance reserves	Re-to-provisions	Amount of reinsurance recoverables / Amount of insurance reserves
2	Loss ratio	Loss ratio	(Insurance claims paid + Expenses associated with claim settlement + change in loss reserves / (Premiums + change in unearned premium reserves)
3	Average size of premium collected	Mean premium	Amount of premiums collected / Number of insurance contracts
4	Ratio of wages to premiums collected	Wages/ Premiums	Amount of wage expenses / Amount of premiums collected for the reporting period
5	Maximum concentration on a group of types of insurance	Concentration	Maximum value of premiums among 7 categories* / Total amount of premiums

* List of categories: 1. Nuclear insurance; 2. Motor insurance (other); 3. Motor insurance; 4. Liability insurance; 5. Personal insurance (health, accident, pension insurance, etc.); 6. Property insurance; 7. Other.

specialization by companies in a certain insurance segment. Thus, the portfolios of many companies can be described as weakly diversified.

Research Methodology and Models

Clustering algorithms are a convenient tool for dividing observations into homogeneous groups based on given features. The literature review cites only some of the successful cases of using cluster analysis in the study of social and economic phenomena. An important advantage of such algorithms is to reduce the influence of a researcher's judgment about a phenomenon under study on the findings of the research. We use classic hierarchical and nonhierarchical clustering methods, along with machine learning methods, to study the business models of insurers. The following is a brief summary of the applied methods.

The k-means method is the most commonly used nonhierarchical method. It suggests iterative minimization of the distance between constituents of a cluster, while the number of clusters is set at the beginning – that is, the model does not determine the optimal number of clusters. The centroid coordinates, the number of which corresponds to the number of clusters, are set randomly at the initial stage. As a result, the division into clusters can be unstable and can depend on the initial centers.

We use the method of seeding the initial centers for k-means, called k-means++, proposed by Arthur and Vassilvitskii (2006) to avoid this problem. Denoting the input data sample χ, and the shortest distance between an element of the sample x_i and the closest center $D\left(x_i\right)$, the algorithm can be described stepwise:

1. Choose the initial center c_1 from χ at random

2. Choose the next center c_j from χ, selecting each element with probability

$$p\left(c_j = x_i\right) = \frac{D\left(x_i\right)^2}{\sum_{x \in \chi} D\left(x\right)^2}$$

3. Repeat Step 2 until the required number of the centers has been chosen

4. Proceed with the classic k-means algorithm.

It can be seen from Step 2 that the elements χ located farther from the initial center are selected with a higher probability. That is, the centers are located so that they are different from each other. Clusters based on the k-means++ procedure were evaluated 100 times to select the clustering with the minimum within-cluster sum of squares (WCSS).

In summary, the cluster centers are first chosen from the sample elements so as to be located farther from each other, then iteratively change their coordinates to describe the largest possible group (cluster) of the sample elements.

For the k-means method, it was decided to use five clusters, which is the optimal number of clusters with regard to the elbow method. The elbow method results are given in Figure B.2. A division into five clusters was used for all further methods.

The k-medoids method, first described by Kaufman and Rousseeuw (1990), is inherently very similar to the k-means method. The key characteristic of the k-medoids method is a partitioning technique of clustering to choose data points as centers. Such data points, which are exemplars for their cluster, are called medoids.

The Partitioning Around Medoids (PAM) algorithm is used to choose medoids iteratively in such a way as to reduce the average distance from data points to the centers of their clusters. Modern algorithms of the k-medoids method offer faster optimization, but PAM remains one of the most accurate algorithms for solving this problem. That is why we chose this method. We selected the initial medoids using the k-medoids++ algorithm, which is identical to k-means++ and ensures cluster stability.

Hierarchical methods do not require the number of clusters to be known before applying the algorithm. They build a tree-like structure called a dendrogram. First, each dataset forms a separate cluster. Further, datasets (clusters) based on the selected criterion are combined into new, larger clusters until they are all combined into one cluster, which includes all observations. Ward's method was chosen for our purpose. The number of clusters is determined by the researcher based on the dendrogram produced by applying the algorithm.

According to Ward's method, a separate cluster is combined with the cluster and their combination will lead to the smallest increase in the distance between data points within the cluster. This distance, which is similar to the WCSS metric of k-means, is displayed on the dendrogram along the vertical axis.

Table 3. Descriptive Statistics of the Models' Variables

	ROA	Offices	Re-to-premiums	% of mandatory premiums	Corporate
Mean	0.02	8.16	0.08	0.19	0.51
Std. Deviation	0.10	24.99	0.21	0.26	0.33
Min	-0.67	0.00	0.00	0.00	0.00
Q(25%)	0.00	0.00	0.00	0.00	0.22
Q(50%)	0.01	0.00	0.00	0.03	0.46
Q(75%)	0.05	0.00	0.03	0.35	0.86
Max	0.40	200.00	1.00	0.86	1.00

Table 4. Descriptive Statistics of Companies' Parameters

	Re-to-provisions	Loss ratio	Mean premium	Wages/ Premiums	Concen-tration
Mean	0.21	0.39	116.41	0.06	0.68
Std. Deviation	0.24	0.51	405.89	0.06	0.19
Min	0.00	-0.80	0.00	0.00	0.29
Q(25%)	0.02	0.11	0.89	0.02	0.54
Q(50%)	0.12	0.35	3.22	0.04	0.66
Q(75%)	0.36	0.52	29.79	0.07	0.83
Max	0.93	4.08	3,534.63	0.46	1.00

The following is a brief description of the Kohonen map, which is a machine learning method capable of clustering. It is described in Kohonen's work (1982). A self-organizing map is an artificial neural network consisting of two layers:

1. Sample data are present in the input layer. The dimensionality of this layer corresponds to the number of features used to cluster datasets into distinct groups.

2. The output layer, which is actually a map consisting of neurons arranged in two (in the case of this study) dimensions and has predetermined arbitrary dimensionality.

All of the neurons on the grid are connected to all of the inputs, and these connections have strengths, or weights, associated with them. That is, each neuron has a set of weights that can be interpreted as a description of the neuron in the features of the data in the input layer. The learning algorithm of the Kohonen map can be described step-by-step:

1. The weights of neurons are initialized to sufficiently small random values.

2. The feature vector x_i from χ is supplied to the input layer and the distance is calculated (this study uses the Euclidean distance) between the vectors x_i and w_j, where w_j is the vector of the weights of the neuron j in the output layer of the grid.

3. The neuron that is closest to x_i based on Step 2 is called the best matching unit (BMU).

4. Taking the radius $\sigma(t)$, the neighborhood parameter is computed for each neuron of the map based on the Gaussian function

$$N(t)_{BMU,j} = exp\left(-\frac{D(BMU,j)^2}{2\sigma(t)^2}\right),$$

where $D(BMU,j)$ is the topographic distance between the BMU and the neuron j.

5. The weights of the neurons on the map are updated according to the formula $\Delta w_j = \alpha(t)N_j(t)(x_i - w_j)$, where $\alpha(t)$ is the learning rate, which is a decreasing function of time.

6. Steps 2-5 are repeated for a given number of epochs (training cycles), as determined by the researcher. At the same time, it is customary to pay attention to the quantization error, which reflects the average distance between the input data and the BMU, and to the topographic error, which reflects the number of data samples for which the first BMU (BMU1) is not an adjacent neighbor of the second BMU (BMU2).

As a result of training, the neurons become "similar" to the input data. As training proceeds, the parameters $\alpha(t)$ and $\sigma(t)$ gradually decrease. Thus, the further the training progresses, the slower the neurons adapt their weights and the less "interaction" they demonstrate. The decreasing function used in this study to describe the dynamics of the parameter $\alpha(t)$ has the following formula:

$$\alpha(t) = \frac{\alpha(0)}{1 + (t/(MI/4))} \tag{1}$$

where $\alpha(0)$ is the initial value of $\alpha(t)$ set by the researcher;
MI is the maximum number of epochs (iterations) set by the researcher;
t is a sequence number of an epoch.

We set $\alpha(0)$ for this study at 0.5, while MI equals 10,000. Thus, the learning rate gradually decreases from 0.5 to 0.1. The dynamics of the parameter $\sigma(t)$ in the process of learning for the model are similar, with an initial value of 1. Honkela (1998) describes the self-organizing map algorithm in more detail.

Given the number of observations in a dataset, a 10x10 map (100 neurons in total) with a rectangular topology was chosen for this study. It is common to initialize neurons' weights based on the principal components observed in the data. However, given that neurons are activated (become BMUs) evenly on the map (Figure B.3) and the learning time is acceptable, we do not use this approach. The dynamics of the topographic error and the quantization error are presented in Figure B.4.

After training, the neurons were clustered by applying the k-means method to their weights in order to be able to compare the findings of the Kohonen map with those of other methods.

The described methods were implemented with Python tools using open-source machine learning libraries such as Scikit-learn (Pedregosa et al., 2011) and MiniSom (Vettigli, 2019).

Evaluation of Clustering Results

Each of the methods has its advantages and disadvantages: assessments of them are given in Table 5.

Table 5. Comparison of Clustering Methods

Features	k-means	k-medoids	Ward's method	Kohonen maps
Ease of interpretation of findings	+	+	+	-
Availability of graphic tools	-	-	+	+
Resistance to outliers	-	+-	+-	+-
Applicability of evaluated model to different datasets	+	+	-	+
Ease of use	+	+	+	-

As we see, none of the methods stands out as the best. Therefore, when applying cluster analysis, the method chosen is most often the one best suited to the available data and numerical criteria for clustering quality.

To evaluate the quality of the models, we used a classic indicator, the Calinski-Harabasz score (CH score), which evaluates the quality of clustering into groups based on the distances between observations. The stability of clusters over time is also important for our purposes. Business models reflect stable behavior (a strategy), and so in order to draw conclusions about business models and their risks, it is important that the clusters do not change significantly over time. To assess stability, we used the migration ratio.

The CH score was first described by Calinski and Harabasz (1974). Also known as the Variance Ratio Criterion, it is the ratio of the sum of between-cluster dispersion and of inter-cluster dispersion for all clusters, both weighted by their respective degrees of freedom. The indicator is calculated as follows:

$$CH\ score = \frac{\left(\frac{BCSS}{k-1}\right)}{\left(\frac{WCSS}{n-k}\right)} \tag{2}$$

where $BCSS$ is the variance between clusters;
$WCSS$ is the variance between datasets within clusters;
k is the number of clusters;
and n is the number of datasets.

There is no critical value for this indicator. However, a larger value indicates a more definite grouping into clusters. The value of the CH score is larger when the centers of the clusters are farther from each other, and the datasets in the clusters are close to their centers.

High-quality clustering forms groups (clusters) that do not change significantly over time. In our example, this is fundamentally important because a business model is a stable feature of a company that does not change significantly under normal operating conditions. To assess the stability of clusters, the migration ratio was calculated:

$$\text{Migration ratio} = \frac{n_m}{n_{2019 \cap 2020}} \tag{3}$$

where n_m is the number of companies that, based on the model, migrated between clusters from 2019 through 2020.
$n_{2019 \cap 2020}$ is the number of companies that were active in 2019 and 2020.

Migration between clusters occurs as a result of two factors – a change in a company's business model and clustering errors. Therefore, an overly large value of the migration ratio indicates inaccurate clustering, and an overly small value hints that a model is "overfit." There is no critical value of this indicator.

A pseudo migration ratio was calculated for Ward's method. Since the estimated model cannot be applied to data from another year, we calculated the pseudo migration ratio. To do this, the model was evaluated on the basis of the most recent data. Next, based on the centroid-based classification method described by Tibshirani et al. (2002), we identified clusters for data from the previous year and applied the formula (3).

Table 6 assesses the clustering quality for the applied models.

Table 6. Comparison of the Quality of Clustering Methods

Indicator	k-means	k-medoids	Ward's method	Kohonen maps
CH score	68.807	68.806	77.101	-
Migration ratio	15.8%	19.0%	20.6% (pseudo)	15.8% (between clusters), 76.9% (between neurons)

It can be seen that Ward's method gives the best clustering outcome according to the CH score criterion. However, the clusters are significantly less stable compared to all of the other methods. Given this criterion, which is of great importance from the point of view of the applicability of the model, we decided not to draw conclusions based on Ward's method. The findings of Ward's method are presented in Figure B.5 for reference. The clusters were named similarly to the main model for ease of comparison.

The k-means and k-medoids models have very close CH score values. Although these values are lower than those obtained by applying Ward's method, the difference is not very significant. The k-means model shows more stable clusters than the k-medoids model does. Also, in the presence of biased data, the k-medoids model may not fully characterize the clusters, as it bases its conclusions on a single observation. For example, this model characterizes four out of five clusters as business models that do not use offices in their activities at all. Such a conclusion is erroneous, as can be seen from the findings of the k-means model presented below. Therefore, due to this data distortion, we did not base our conclusions on the findings of the k-medoids model. The findings of the k-medoids method are presented in Table A.2 for reference.

Two types of migration ratio were calculated for the Kohonen map. The first is based on the five clusters into which the neurons are grouped. The second is based on all one hundred neurons of the model. As expected, the former is much smaller than the latter. It is interesting to observe that the migration between the clusters when applying the Kohonen map is almost identical to the case with the k-means method.

Given the findings of the quality assessment of the models, we decided to base our conclusions on the findings of the k-means model. Also, since the Kohonen map findings are similar to those of the k-means method, we used the map as a cluster visualization tool.

Next, we built a silhouette graph (Rousseeuw, 1987) for the k-means findings to evaluate the outcome in more detail. The silhouette coefficient is calculated for each observation as:

$$s_i = \frac{b(i) - a(i)}{max\{a(i), b(i)\}} \tag{4}$$

where $a(i)$ is the average distance from the sample i to the members of its cluster; and
$b(i)$ is the average distance from the sample i to the members of the nearest neighboring cluster.

The value of the ratio for the model is the average value of the silhouette coefficient of all observations. A silhouette coefficient value of 1 indicates that clusters are clearly distinguished; 0 means that clusters are indifferent; and -1 means that clusters have been assigned wrongly.

The graph of silhouettes (Figure 1) shows that there is only one observation for which members of the neighboring cluster are "closer" on average than members of its own cluster. This observation relates to cluster 1 (Universal "Large" model). It, and those with a silhouette value close to 0, may be located "on the edge" of the cluster. The overall value of the ratio (0.41) indicates a sufficiently high-quality clustering; in addition, the graph shows that the Reinsurance business model (Cluster 4) is the best defined.

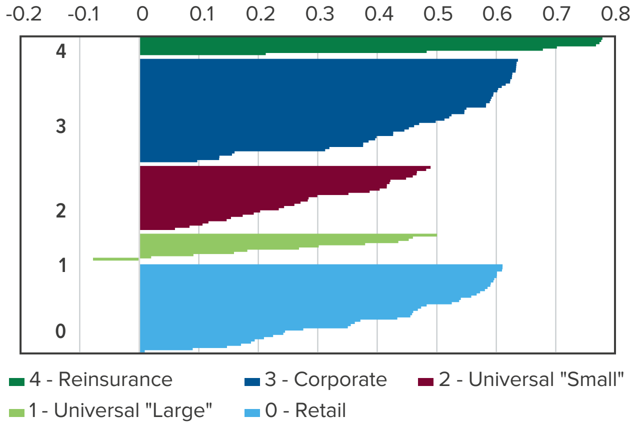


Figure 1. Cluster Silhouettes (the abscissa is a silhouette coefficient, the ordinate is a cluster number)

4. RESEARCH FINDINGS

Description of Business Models Based on Clusters

The model was evaluated on the basis of data for the year 2020 and applied to all years in the sample (2019-2020). The features (the coordinates of the centroids) of the identified clusters are shown in Figure 2. The coordinates in the figure are standardized.

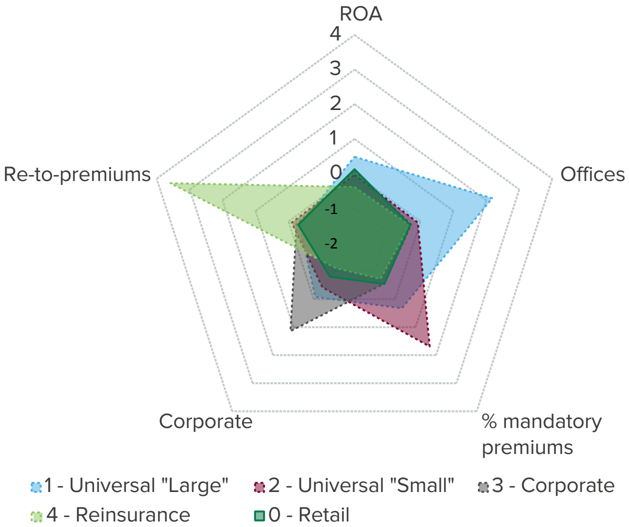


Figure 2. Features of Identified Clusters (standardized)

The identified clusters were numbered and named for convenience. We do not rule out that other homogeneous groups of companies may be identified from the data and may distort the features of the clusters we have identified. However, relying only on the data and the methods described above, we managed to identify those business models that are best separated in terms of quantitative criteria. Then we provide a brief overview of the identified clusters (business models) based on the features used for clustering and on descriptive indicators. Table A.3 summarizes the clustering findings in an unstandardized form.

Business model 0 – Retail, focuses on insuring individuals (who account for 74% of premiums) and has an average level of return on assets (3%). With a small number of offices and a significant ratio of wages to premiums collected, companies with this business model use their own agents as a channel to acquire customers. That is, the companies "hunt" for customers rather than customers themselves coming to their offices. Insurers using this business model offer mostly voluntary types of insurance, their average share of reinsurance in premiums is 1.5%. It is worth noting that this cluster has the highest concentration indicator (76%) on one of the insurance types, which can be a risk factor for the companies in this group. In 2020, the companies of this cluster posted the highest losses, and the share of reinsurance in their insurance reserves was moderate, which indicates the companies' vulnerability to underwriting risk. In 2020, this cluster included 40 companies, which accounted for 19% of the market by premiums.

Business model 1 – Universal "Large" insurers serve both legal entities and individuals and have a distribution between mandatory and voluntary insurance of 28% and 72%, respectively. Their return on assets is the highest among all selected clusters (6.5%). A characteristic feature of this cluster is the wide use of its own offices (they have about 62 offices on average). One sign that companies of this business model actively use both their own offices and agents as sales channels is the high share of wages in premiums collected (6%). Thus, these companies try to diversify their ways of acquiring customers. Companies in this cluster also have the second largest share of reinsured risk (25%), which means lower underwriting risk, as well as the fact that the companies run the risk of counterparty (reinsurer) default. Some 12 companies in 2020 used this business model and had the largest share of gross premiums, estimated at 35.5%. It is worth noting that given the high market share combined with the lowest average premium, these companies tend to sell low-priced policies on a large scale.

Business model 2 – Universal "Small", is characterized by a relatively even (compared to other models) distribution in premiums of mandatory and voluntary insurance and individuals and legal entities (64%/36% and 63%/37%, respectively). However, the share of premiums from mandatory types of insurance in this business model is the largest of all the clusters. Premiums from Motor Third Party Liability (MTPL) insurance account for 71% of the premiums from mandatory insurance types for this cluster. This cluster also has the second lowest rate of return on assets among all groups, and companies own an average of six offices and have the second highest share of reinsurance in premiums (5.6%). In addition to focusing on mandatory insurance, this model differs from Universal "Large" by a significant difference in the average premium, which could be evidence that these companies try to insure more expensive risks. In 2020, this cluster included 29 companies, which together accounted for about 16.5% of the market by premiums.

Business model 3 – Corporate, is characterized by an 89% share of legal entities in premiums, as well as a low rate of return on assets (2.7%) and a small number of offices, while its share of mandatory insurance is close to zero. For companies that do not use a reinsurance business model, their share of inwards (assumed) reinsurance in premiums is significant (27%). With a relatively high level of average premium (UAH 254,000), the companies of this cluster have a fairly low loss ratio compared to other business models (22.5%). A high share of a reinsurer in the insurance reserves is predictable, as such insurers often need to share a corporate client's large exposure. However, this creates the risk of counterparty (reinsurer) default for the companies in this cluster. This cluster encompasses the largest number of companies (47), which, based on the premiums collected in 2020, together account for 19% of the market.

Companies of business model 4 – Reinsurance, have an average share of reinsurance in premiums of about 81%. The return on assets of the companies of this cluster is negative on average, and the average premium is more than UAH 304,000. Companies in this cluster do not use offices as a sales channel at all and have the lowest share of wages in premiums. The share of voluntary insurance in premiums approaches 100%. Reinsurers themselves are weakly reinsured, which may indicate a potential vulnerability to underwriting risks that they do not share (diversify) among themselves. However, the low value of the loss ratio compared to other business models indicates that the underwriting risk may be insignificant. The cluster included eight companies according to 2020 data (10% of the market by premiums).

Histograms with the features of the grouped clusters are shown in Figure B.6.

Neurons on the Kohonen self-organizing map in the process of training become "similar" to the input data in terms of their weights, i.e. they reproduce clusters. The Kohonen self-organizing map provides a convenient tool for visualizing the similarities between different observations and the characteristics of those observations. With the help of the map, one can see the samples that lie on the border of the clusters and how far they are from other elements of the cluster.

The maps of the features in Figure 3 show a map's neuron weights correspond to data features (coordinates are standardized). They should be interpreted as follows: companies for which the neuron with coordinates (1;10) (upper left corner) after training is the BMU (the closest one) have an average market share of premiums from mandatory types of insurance, lots of offices, and almost no reinsurance in premiums.

Neighboring neurons are quite similar due to the mechanism of "cooperation" during learning, so they can be combined into groups. For this, the k-means algorithm with a number of clusters equal to five was used to interpret the results in a similar way as the results of the k-means division. We do not indicate the centroids of these clusters, since grouping by the k-means method was carried out only to mathematically estimate the boundaries of the clusters on the Kohonen map, and such centroids would essentially reproduce the centroids of the previous model.

Figure 4 shows the results of combining neurons into clusters. The dots in the figure indicate companies for which a particular neuron is the BMU after training. As one can see, considering the combination of neurons and feature maps, it is possible to identify business models that are similar to those identified by the k-means method.

However, the map allows us to see the distance (similarity) between the observations. The topographic distance can immediately be seen on the map – neighboring neurons are similar. The Euclidean distance between neurons after training can be seen in Figure B.7. Companies for which the BMU is located on the topographic boundary of a cluster are "weak" representatives of the cluster and may change their cluster over time. It is these companies to which we refer when validating the results of the k-means model.

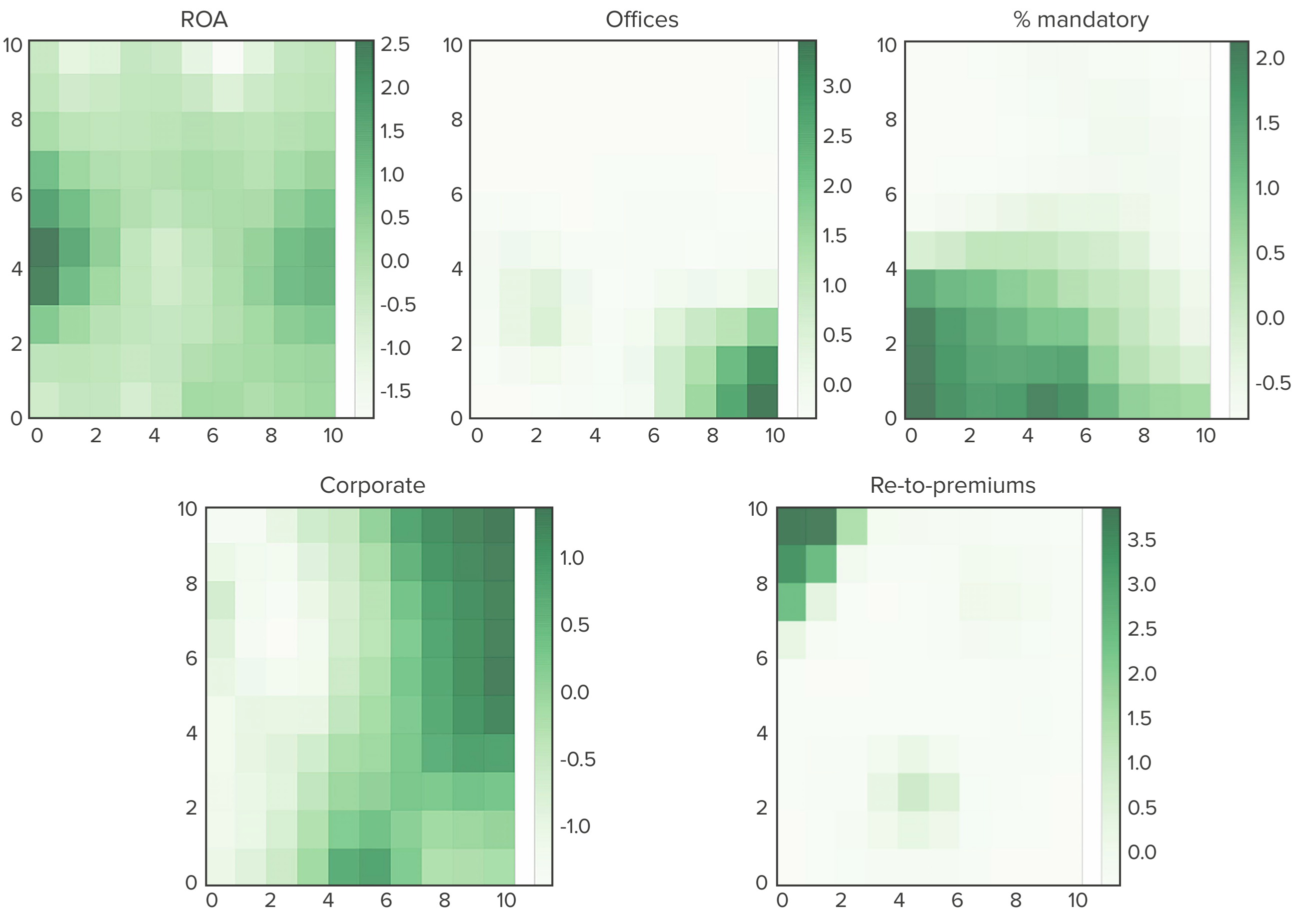


Figure 3. Kohonen Maps of Features

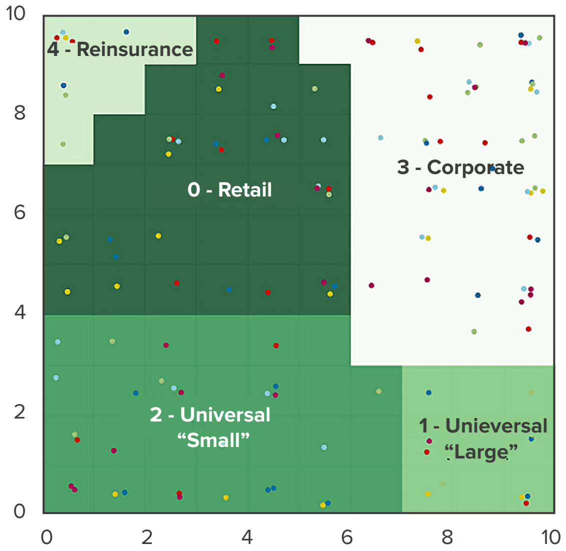


Figure 4. Grouping of Clusters on the Kohonen Map

The migration ratio for the Kohonen map between the five clusters is 16%, the one between 100 neurons is 77%. This indicates that migration within the clusters is greater than between the clusters. However, given this indicator, conclusions regarding the business models are given based on the results of k-means clustering. But it is worth noting that the economic essence of the models determined using the Kohonen map coincides with the results of the K-means method.

Analysis of Migration Between Clusters

Having identified clusters (business models), it is possible to study the dynamics of their constituents throughout the period under review. The migration coefficients from each cluster in the period from 2019 to 2020 were calculated.

As described above, migration between clusters can occur under the influence of two factors – model errors and changes in a company's business model. Knowing that clustering is not an exact method, we considered migration between clusters to be significant if more than 10% of cluster constituents migrated from it. The application of this threshold shows the migrations of at least two companies from a cluster, and most migrations that are greater than the value of the migration coefficient of the model (15%), to be significant.

Since the companies that earned less than UAH 5 million in premiums per year were not included in the k-means algorithm and were assigned to the artificially created Inactive cluster, migration from the selected business models to the Inactive cluster was also observed.

Figure 5 shows significant migrations of companies between the business models. It can be seen that there was a considerable migration in 2020 from the Universal "Large" model to the Universal "Small" model. This is to be expected, as the difference between the clusters and business models is not significant. There are less obvious reasons for the migration of companies from the Reinsurance model to the Corporate model. However, if one looks at the centers of the clusters, it can be seen that the Corporate model is closer to the Reinsurance model than the others, as the companies of the Reinsurance model have a small share of premiums from legal entities, and the companies of the Corporate model have a share of premiums from inwards reinsurance.

Given the migrations to the Inactive group, companies using the Corporate, Retail, and Reinsurance business models have a greater risk of exiting the market, and are therefore seen as less stable. More than half of the companies of the Reinsurance model exited the market in 2020, which may serve as a signal to the regulator that closer supervision is required.

Describing the business models, we noted that insurers in the Corporate model widely use the outward reinsurance of their risks, that is, they depend on the Reinsurers in their operations. Therefore, it is logical that there is a high level of simultaneous exiting from the market among companies of these two clusters.

One can assess the robustness of the conclusions based on migrations to the Inactive group. This group includes the companies whose annual premiums were less than UAH 5 million, so very small companies for which premiums of about UAH 5 million are normal could migrate due to a change in premiums from year to year. A mere 24% of companies that migrated to the Inactive cluster had

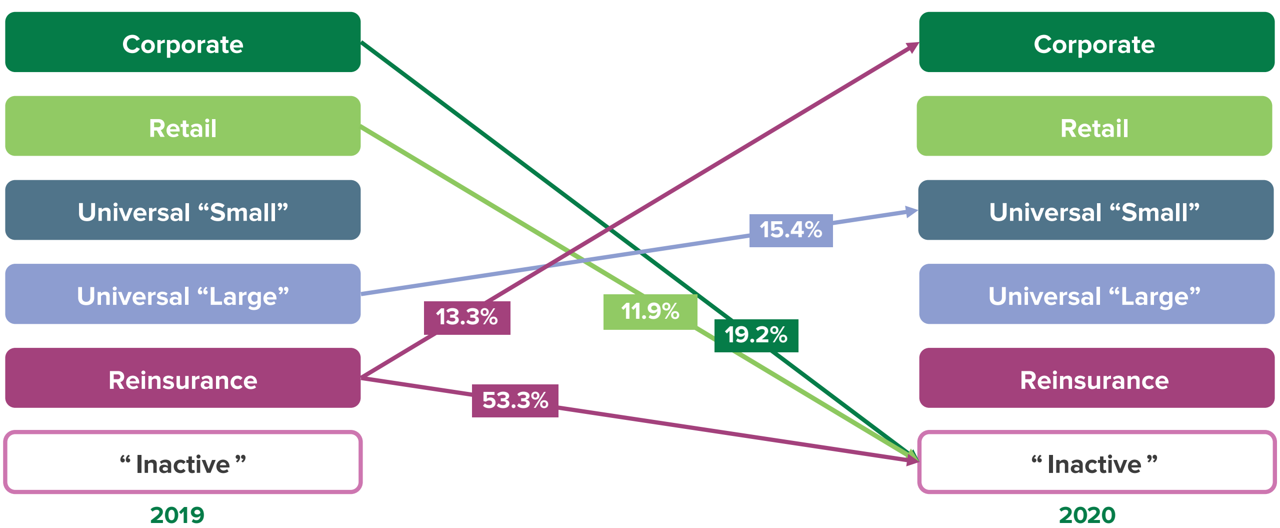


Figure 5. Migrations between the Business Models

less than UAH 10 million in premiums in 2019. The median value of premiums among the companies that migrated to the group is UAH 175 million, the average is UAH 397 million. Since a sharp reduction in the volume of premiums from such values indicates a crisis in a company's activities, the results can be considered robust. A mere 16% of these companies had a nonzero volume of premiums in 2020. That is, the majority completely discontinued insurance activity.

Migrations between the clusters and their causes can be assessed in more detail using the Kohonen map. Figure 6 shows the companies that were included in the respective clusters in either 2019 or 2020. It is worth noting that while Figure 5 shows one-way migrations, the Kohonen map shows two-way migrations (all companies that migrated from cluster to cluster).

This figure demonstrates that migration occurs mostly between neighboring clusters and neurons. The ratio of migration between clusters for the Kohonen map is close to the k-means model, at 15.8%.

For example, it can be seen that a significant migration between the Reinsurance cluster and the Corporate cluster, which was identified by both models, is due to the sharp curtailment by reinsurers of their activities and the start of the servicing of corporate clients. Since this migration does not occur between neighboring neurons and clusters, it can be argued that the companies were not on the edge of the clusters, but significantly changed their business model.

Unlike the k-means model, the Kohonen map does not show a significant migration within the Universal business model. Both clusters of this model demonstrate a slight migration of companies to and from other clusters, which we cannot deem significant. It is interesting that migration for small companies of this business model occurs mostly with neighboring neurons on the edge of the cluster, while migration for large ones happens only far from the edge of the cluster.

By depicting on the Kohonen map the companies that have discontinued providing insurance services, it is possible to highlight those of its zones that are characterized by high risk (Figure 7).

The findings of the Kohonen map are consistent with those of the k-means model; Universal can be considered the safest business model. The companies using the Reinsurance and Corporate business models are empirically the least stable.

The upper-right corner of the map is a particularly risky area of the Corporate model. There are companies whose share of legal entities in premiums is close to 100% and which offer voluntary types of insurance. It is worth noting that unprofitability is hardly the cause of these companies' high risk, as their ROA is close to the average for the market.

The lower part of the cluster on the map is risky for the Reinsurance model. These are companies that provide both reinsurance and direct voluntary insurance services. It can be concluded that more stable reinsurers are engaged solely in reinsurance activity.

It can be seen that the area of the map characterized by the highest return on assets shows absolutely no migration to the Inactive group. That is, profitable operation increases the stability of companies.

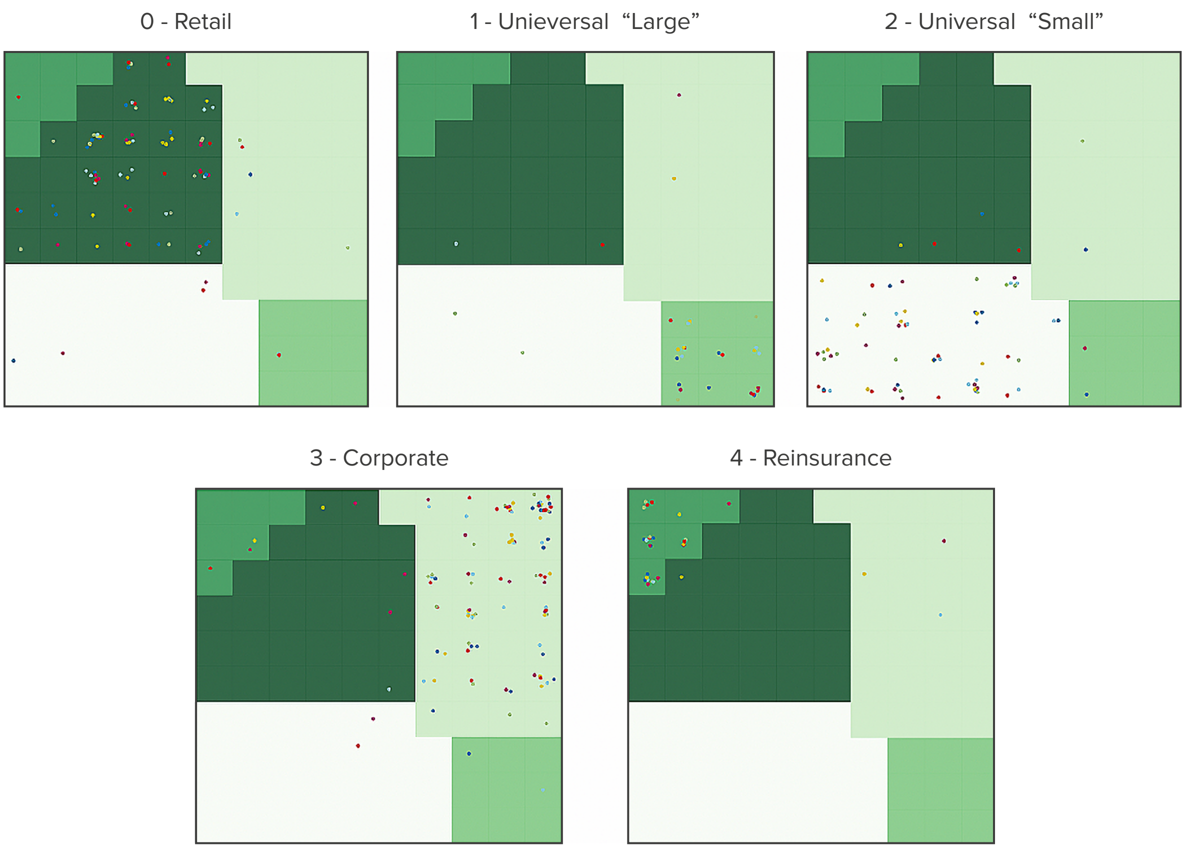


Figure 6. The Number of Neuron Activations on the Kohonen Map

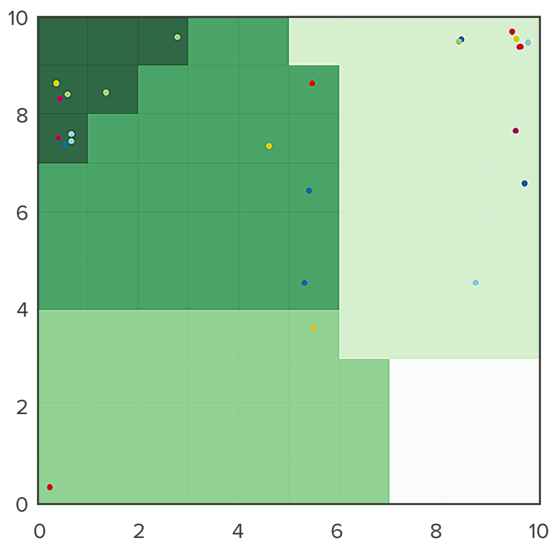


Figure 7. Companies That Migrated to the Inactive Group

5. CONCLUSIONS

In this paper, we study the nonlife insurance market in Ukraine. Particularly, we find persistent and economically reasonable ways of doing business that companies use (business models). The business model highlights not only the operational but also the risk characteristics of a company. Thus, knowledge of business models and the ability to identify which model a specific company uses is of great importance in the supervision process.

First, we decide which quantitative indicators can help to describe the business model of an insurance company. Then, we apply a set of clustering techniques to company-level indicators calculated from the regulatory database and group the companies into clusters. If we know that these groups are stable in time and are formed on the basis of indicators that describe their business model, then the description of a cluster is itself a description of a business model. The use of well-defined algorithms and performance metrics allows us to rule out personal judgment to a great extent. Finally, as we divided companies into clusters, we can study how companies changed clusters over the research period.

We apply a set of clustering algorithms to our data. Specifically, we perform clustering using the hierarchical Ward method, k-means, k-medoids, and Kohonen's SOM. We find that k-means provides the best combination of the quality of clusters' separation and their stability overtime. We also use SOMs as convenient visualization tools for clustering results, as SOM clusters carry the same economic meaning as k-means clusters.

We identify the following four different business models of insurers on the Ukrainian market based on quantitative data: Retail, Corporate, Universal (divided into two clusters, large and small), and Reinsurance. The sixth cluster is formed artificially – it includes insurance companies whose gross premiums for the year amounted to less than UAH 5 million and which were considered inactive for the purposes of this study. The research also describes the mentioned business models on the basis of the key quantitative indicators that characterize them.

Companies whose business model is retail insure individuals and tend to focus on certain types of insurance. This focus, and a low level of outward reinsurance, make them vulnerable to underwriting risk.

Large universal insurers are mostly well-known insurance companies that enjoy the trust of consumers, have many offices, and that have high profitability. They focus on selling a large number of low-priced policies.

Small universal insurers are inclined to provide mandatory types of insurance, in particular MTPL. Thus, the risks of this business model are closely related to the risks of civil liability insurance of vehicles. These companies tend to have low profitability.

Corporate insurers focus on legal entities and insure expensive risks. They make extensive use of outwards reinsurance to reduce underwriting risk. However, this makes them vulnerable to the risk of counterparty default.

We also conclude that reinsurers are the least profitable on the market, reinsuring mostly voluntary types of insurance. We reveal that reinsurers are themselves insufficiently reinsured, which makes them exposed to underwriting risk.

Then, the study shows insurance companies' migration between clusters. According to the model, companies using the Corporate and Reinsurance business models from 2019 to 2020 most often exited the market, which may indicate that such companies need more attention from the supervisor. At the same time, the Retail and Universal business models are the most stable, and therefore may be considered the least risky. Therefore, the proposed combination of methods can be considered effective for market supervision purposes.

This study provides a foundation for further research in two directions. First, we consider the clusters identified in this work to be quite broad, although they correspond to the key areas of the companies' activities. Therefore, identifying more narrowly oriented business models based on the clusters described in this study would be a logical continuation of the development of the topic. Second, in view of the described empirical dependence of an insurer's risk level on the type of business model it uses, it is extremely important to look into the risk factors that affect companies from different clusters. We see the availability of detailed and reliable data on companies on the insurance market in Ukraine as a key factor that would contribute to the further development of this topic.

REFERENCES

Abbas, S. A., Aslam, A., Rehman, A. U., Abbasi, W. A., Arif, S., Kazmi, S. Z. H. (2020). K-Means and K-Medoids: Cluster analysis on birth data collected in city Muzaffarabad, Kashmir. IEEE Access, 8, 151847-151855. https://doi.org/10.1109/ACCESS.2020.3014021

Abolmakarem, S., Abdi, F., Khalili-Damghani, K. (2016). Insurance customer segmentation using clustering approach. International Journal of Knowledge Engineering and Data Mining, 4(1), https://doi.org/10.1504/IJKEDM.2016.082072

Ahmar, A. S., Napitupulu, D., Rahim, R., Hidayat, R., Sonatha, Y., Azmi, M. (2018). Using K-Means clustering to cluster provinces in Indonesia. Journal of Physics: Conference Series, 1028, 012006. http://doi.org/10.1088/1742-6596/1028/1/012006

Arthur, D., Vassilvitskii, S. (2007). k-means++: the advantages of careful seeding. SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms (p./pp. 1027--1035), Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.

Bach, M. P., Vlahović, N., Pivar, J. (2020). Fraud prevention in the leasing industry using the Kohonen self-organising maps. Organizacija, 53(2), 128-145. https://doi.org/10.2478/orga-2020-0009

Caliński, T., Harabasz, J. (1972). A dendrite method for cluster analysis. Communications in Statistics, 3(1), 1-27. https://doi.org/10.1080/03610927408827101

Honkela, T. (1998). Description of Kohonen's self-organizing map. In Honkela, T. Self-Organizing Maps in Natural Language Processing. Helsinki: Helsinki University of Technology. Retrieved from http://www.mlab.uiah.fi/~timo/som/thesis-som.html

Kaufman, L., Rousseeuw, P. J. (1990). Partitioning around medoids (Program PAM). In Kaufman, L., Rousseeuw, P. J. Finding Groups in Data: An Introduction to Cluster AnalysisFinding Groups in Data, pp. 68-125. John Wiley & Sons. https://doi.org/10.1002/9780470316801.ch2

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. Biological Cybernetics, 43, 59–69. https://doi.org/10.1007/BF00337288

Kramarić, T. P., Bach, M. P., Dumičić, K., Žmuk, B., Žaja, M. M. (2017). Exploratory study of insurance companies in selected post-transition countries: Non-hierarchical cluster analysis. Central European Journal of Operations Research, 26(3), 783–807. https://doi.org/10.1007/s10100-017-0514-7

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(85), 2825–2830. Retrieved from http://jmlr.org/papers/v12/pedregosa11a.html

Rashkovan, V., Pokidin, D. (2016). Ukrainian banks' business models clustering: Application of Kohonen neural networks. Visnyk of the National Bank of Ukraine, 238, 13-38. https://doi.org/10.26531/vnbu2016.238.013

Rousseew, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Tibshirani, R., Hastie, T., Narasimhan, B., Chu, G. (2002). Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proceedings of the National Academy of Sciences of the United States of America, 99(10), 6567-6572. https://doi.org/10.1073/pnas.082099299

Velykoivanenko, H., Beschastna, G. (2018). Analysis of the stability and rating of Ukrainian insurance companies. Modelling and Information Systems in Economics, 95, 65-81. Retrieved from https://ir.kneu.edu.ua:443/handle/2010/30980

Vettigli, G. (2019). MiniSom: Minimalistic and Numpy-Based Implementation of the Self Organizing Map (release 2.1.5. 2019). Retrieved from https://github.com/JustGlowing/minisom t

Wang X., Keogh E. (2008) A clustering analysis for target group identification by Locality in motor insurance industry. Soft Computing Applications in Business. Studies in Fuzziness and Soft Computing, 230, 113–127. https://doi.org/10.1007/978-3-540-79005-1_7

Zaqueu, J. R. (2019). Customer Clustering in the Health Insurance Industry by Means of Unsupervised Machine Learning: An Internship Report. University of Lisbon, Information Management School. Retrieved from https://run.unl.pt/bitstream/10362/89468/1/TAA0043.pdf

APPENDIX A. TABLES

Table A.1. Descriptive Statistics of Variables in 2019

a) Descriptive Statistics of Variables in the Model

	ROA	Offices	Re-to-premiums	% of mandatory premiums	Corporate
Mean	0.10	6.26	0.17	0.49	0.13
Std. Deviation	0.25	18.44	0.25	0.34	0.27
Min	-0.22	0.00	0.00	0.00	0.00
Q(25%)	0.02	0.00	0.00	0.18	0.00
Q(50%)	0.05	0.00	0.03	0.44	0.01
Q(75%)	0.09	0.00	0.29	0.83	0.08
Max	2.19	115.00	0.90	1.00	1.00

b) Descriptive Statistics of Companies' Additional Characteristics

	Re-to-provisions	Loss ratio	Mean premium	Wages/Premiums	Concentration
Mean	0.23	0.49	243.71	0.02	0.70
Std. Deviation	0.25	0.50	1,199.14	0.02	0.18
Min	0.00	-0.63	0.00	0.00	0.33
Q(25%)	0.03	0.09	0.79	0.00	0.57
Q(50%)	0.14	0.43	2.69	0.01	0.69
Q(75%)	0.37	0.73	26.32	0.02	0.81
Max	1.72	3.33	13,768.22	0.07	1.00

Table A.2. Findings of the k-medoids Method

a) Coordinates of Cluster Centers

	ROA	Offices	% of mandatory premiums	Corporate	Re-to-premiums
0 – Retail	0.002	0.000	0.000	0.242	0.004
1 – Universal "Large"	0.067	44.000	0.242	0.433	0.017
2 – Universal "Small"	0.001	0.000	0.663	0.342	0.068
3 – Corporate	0.011	0.000	0.001	0.913	0.004
4 – Reinsurance	0.001	0.000	0.000	0.038	0.893

b) Number of Companies Included in the Clusters in the Period Under Review

	2019	2020
0 – Retail	44	40
1 – Universal "Large"	14	14
2 – Universal "Small"	28	28
3 – Corporate	52	46
4 – Reinsurance	14	8

Table A.3. Description of the Clusters

a) Coordinates of Cluster Centers

	ROA	Offices	% of mandatory premiums	Corporate	Re-to-premiums
0 – Retail	0.030	0.350	0.050	0.242	0.015
1 – Universal "Large"	0.065	62.330	0.277	0.482	0.008
2 – Universal "Small"	0.014	6.140	0.637	0.368	0.057
3 – Corporate	0.027	1.190	0.047	0.887	0.027
4 – Reinsurance	-0.020	0.000	0.001	0.141	0.810

b) Additional Descriptive Characteristics of the Clusters (2020)

	Re-to-provisions	Loss ratio	Mean premium	Wages/ Premiums	Concen-tration
0 – Retail	0.186	0.457	5.424	0.065	0.759
1 – Universal "Large"	0.250	0.381	1.669	0.060	0.466
2 – Universal "Small"	0.168	0.390	41.900	0.064	0.637
3 – Corporate	0.275	0.225	254.20	0.051	0.701
4 – Reinsurance	0.119	0.049	303.97	0.001	0.679

c) Number of Companies Included in the Clusters in the Period Under Review

	2019	2020
0 – Retail	42	40
1 – Universal "Large"	13	12
2 – Universal "Small"	29	29
3 – Corporate	52	47
4 – Reinsurance	15	8

APPENDIX B. FIGURES

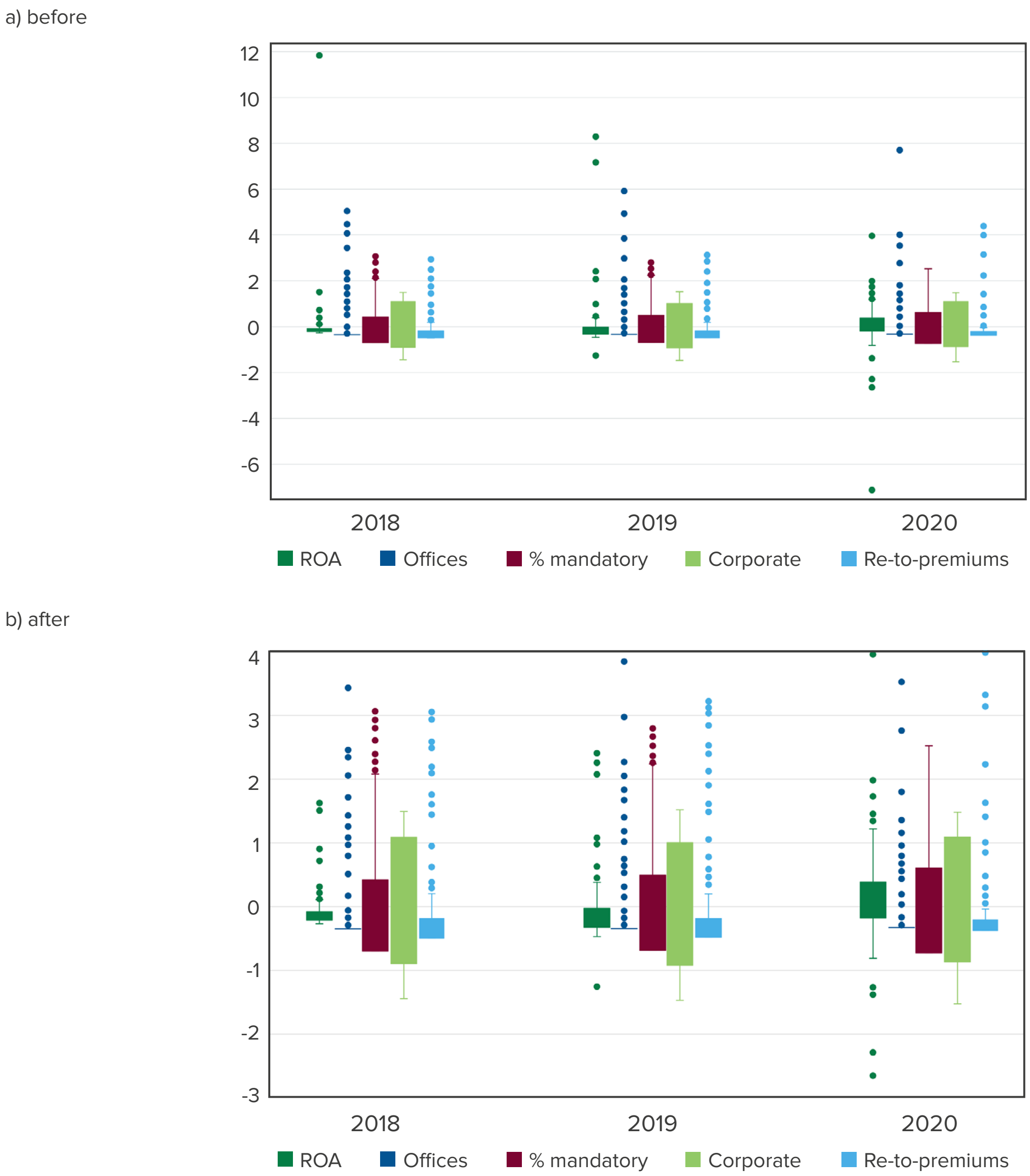


Figure B.1. Distribution of Values of Variables Before and After Adjustment of Outliers, value, years.

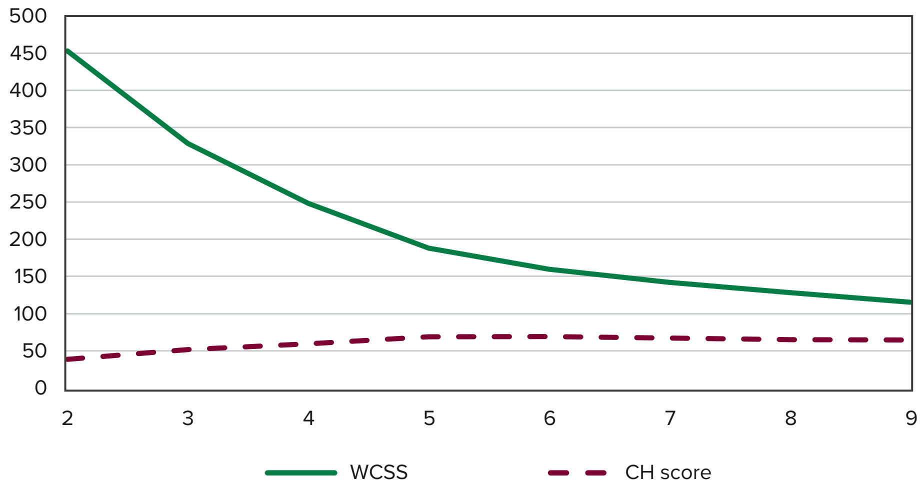


Figure B.2. Criteria for Choosing the Number of Clusters, Elbow method

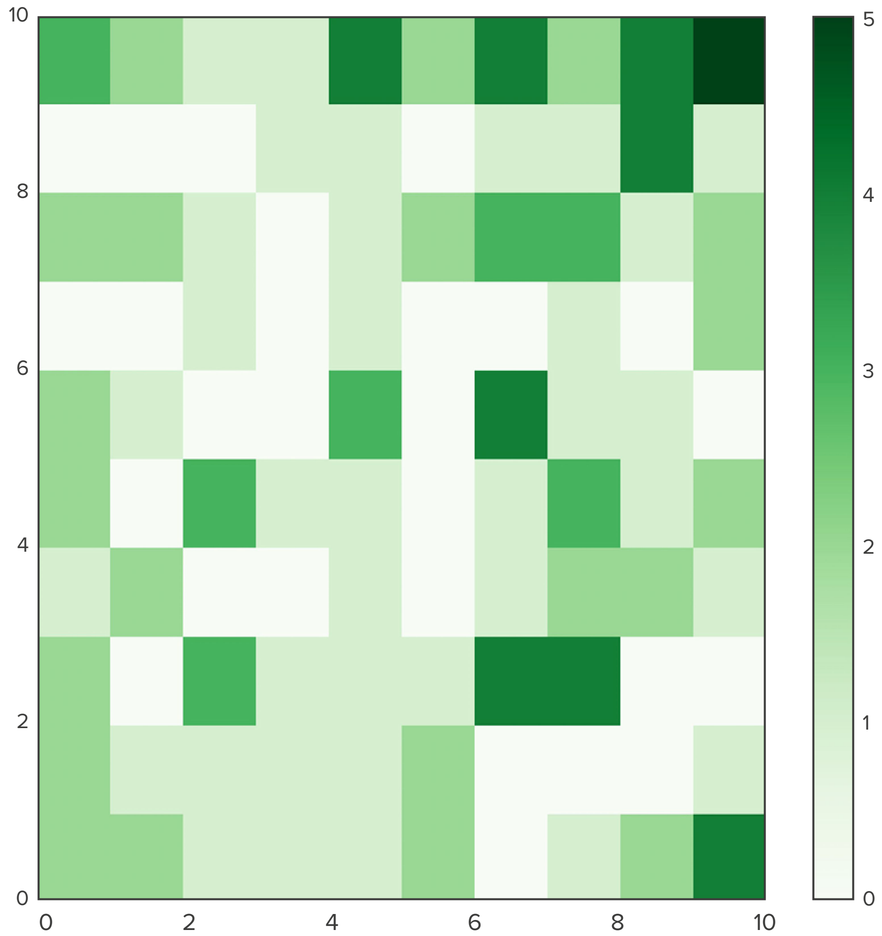


Figure B.3. The Number of Neuron Activations on the Kohonen Map

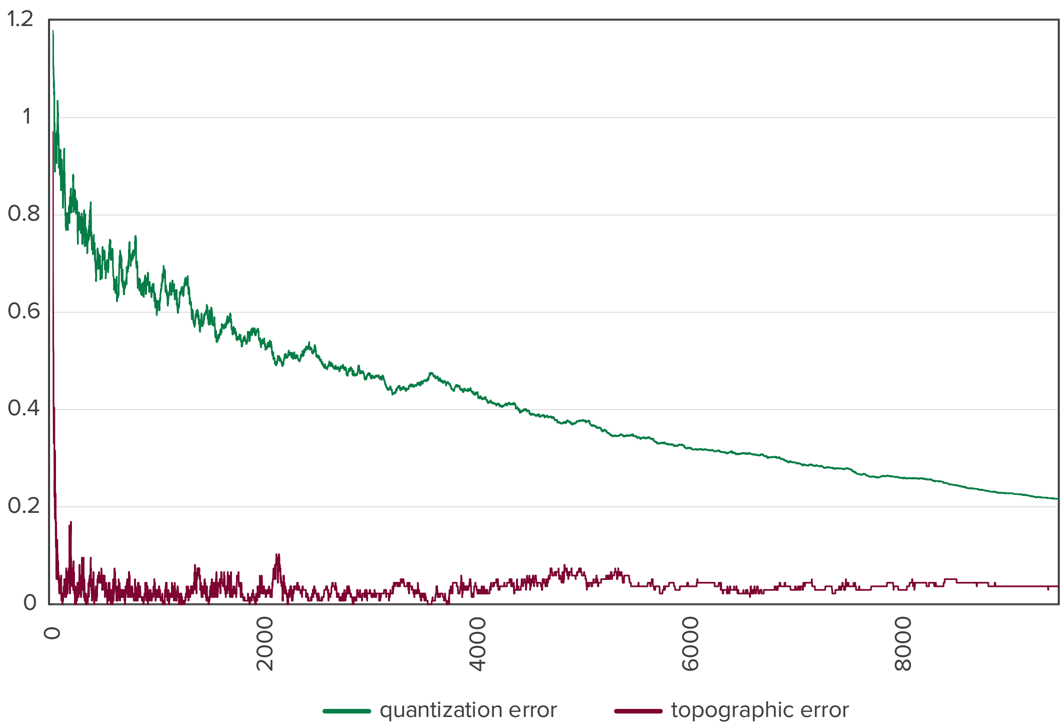
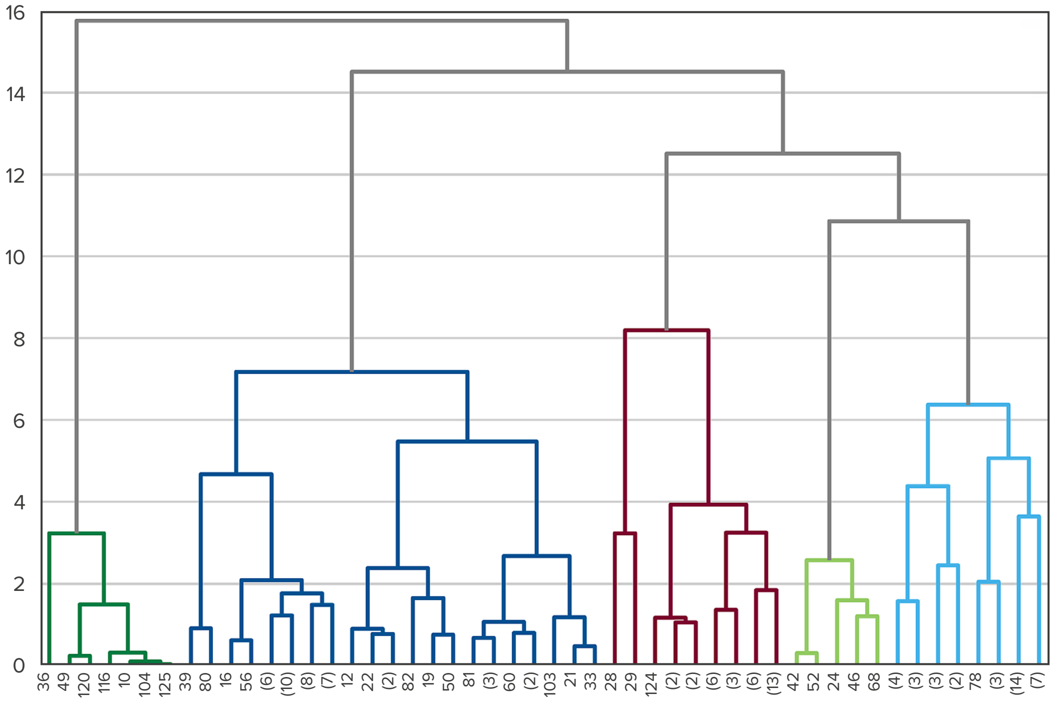


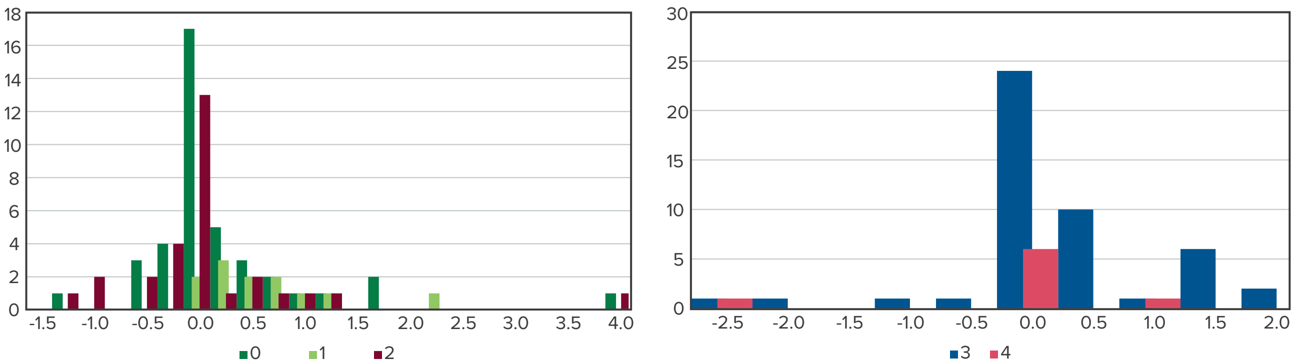
Figure B.4. The Dynamics of Kohonen Network Learning, error, iteration index



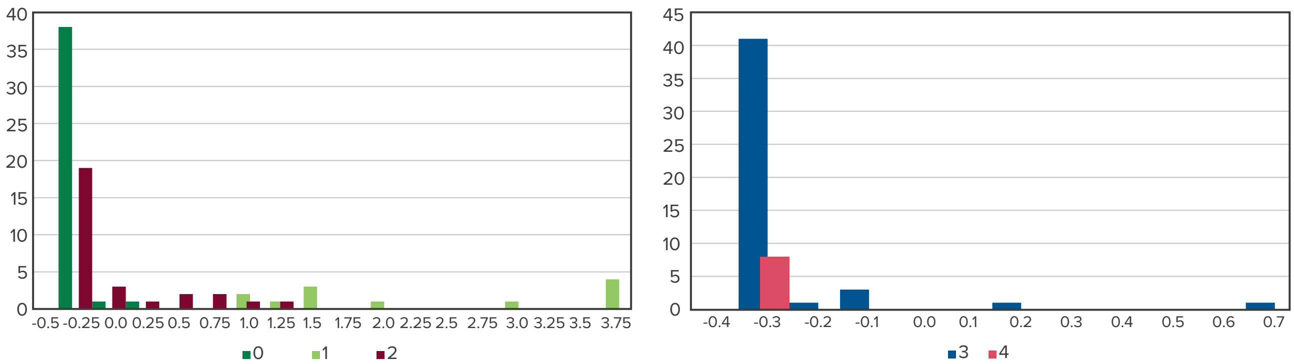
	ROA	Offices	% of mandatory premiums	Corporate	Re-to-premiums	Number of companies (2020)
0 – Retail	0.046	0.143	0.044	0.207	0.009	35
1 – Universal “Large”	0.057	92.800	0.278	0.487	0.001	5
2 – Universal “Small”	0.002	6.838	0.537	0.393	0.066	37
3 – Corporate	0.034	5.269	0.062	0.847	0.026	52
4 – Reinsurance	-0.032	0.000	0.001	0.096	0.849	7

Figure B.5. Findings of Ward’s Method

a) ROA



b) Offices



c) % of mandatory premiums

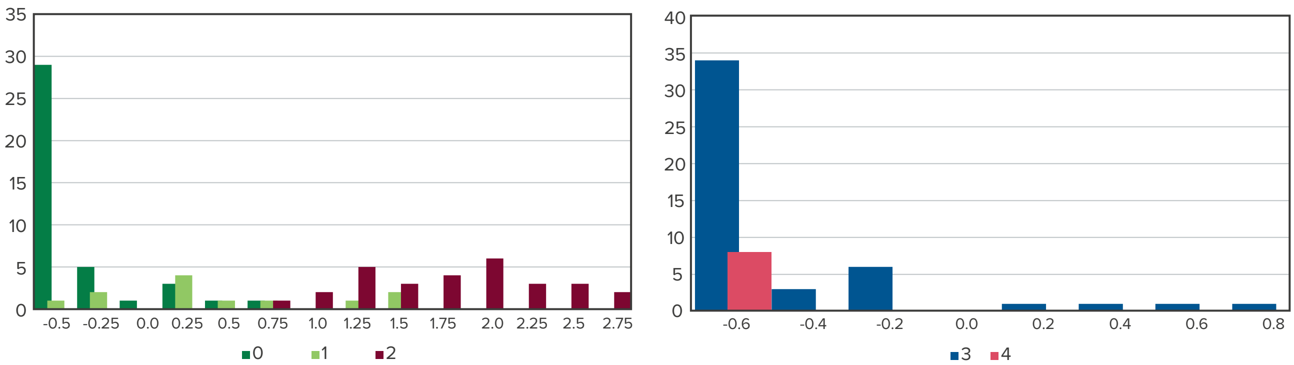
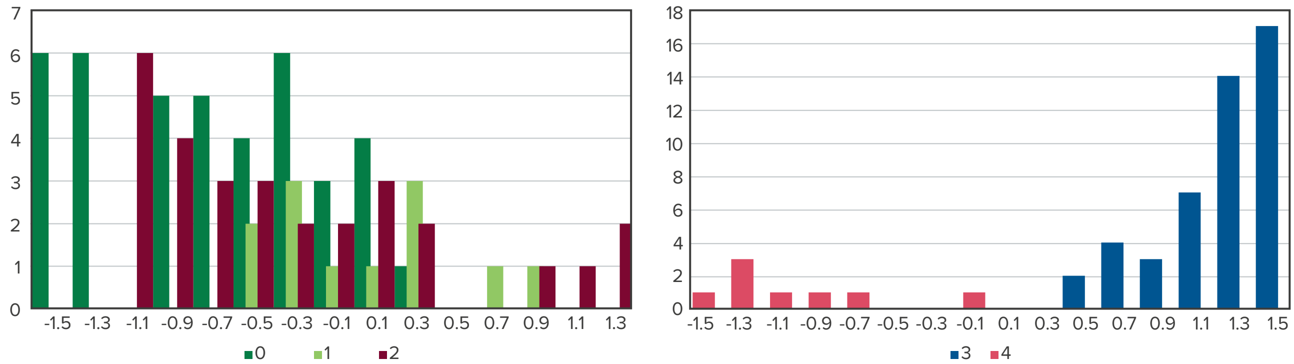


Figure B.6. Histograms of Features of the Identified Clusters (standardized)

d) Corporate



e) Re-to-premiums

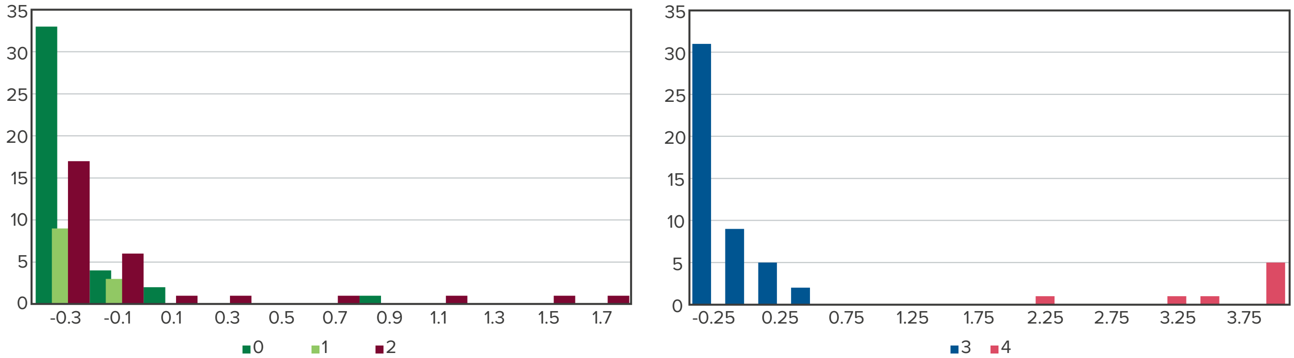


Figure B.6 (continued). Histograms of Features of the Identified Clusters (standardized)

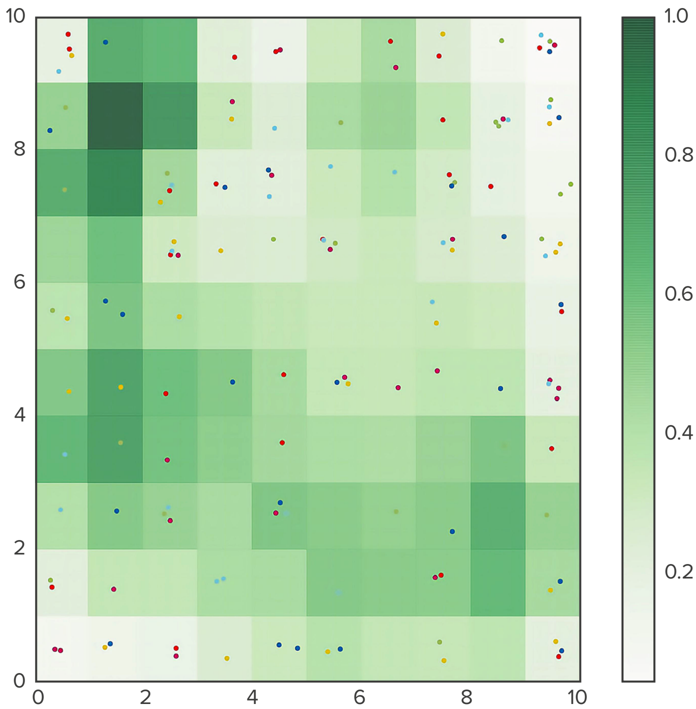


Figure B.7. Euclidean Distance Between Neurons (normalized)