DIGITAL INSIGHTS:
TRANSFORMING GDP
NOWCASTING IN UKRAINE WITH
HIGH-FREQUENCY DATA

by

Nataliia Horoshko

A thesis submitted in partial fulfillment
of the requirements for the degree of

MA in Economic Analysis.

Kyiv School of Economics

2024

Thesis Supervisor: ______________________ Professor Mihnea Constantinescu

Approved by ____
Head of the KSE Defense Committee, Professor

____

____

____

Date ______________________________

Kyiv School of Economics

Abstract

DIGITAL INSIGHTS:
TRANSFORMING GDP
NOWCASTING IN UKRAINE WITH
HIGH-FREQUENCY DATA

by Nataliia Horoshko

Thesis Supervisor: Professor Mihnea Constantinescu

The ongoing war in Ukraine has disrupted the timely availability of official economic statistics, which are crucial for assessing the state of the economy. With frequent migration, military operations, and fluctuating public sentiment, Ukraine's economic conditions can vary significantly on a monthly basis. This necessitates the development of alternative approaches to measure economic activity, particularly for institutions relying on data for decision-making. One promising method is the integration of high-frequency digital data, such as internet search trends, which can provide real-time insights into population behavior and economic dynamics. However, combining data sources with different frequencies poses challenges. This work explores mixed-frequency model specifications for nowcasting the growth rates of 10 GDP components in Ukraine, using both official statistics and Google Trends data. By leveraging the strengths of diverse data sources, this study aims to enhance the accuracy and responsiveness of economic monitoring in crisis situations.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES - Continued

LIST OF TABLES

LIST OF TABLES - Continued

LIST OF ABBREVIATIONS

GRP. Gross Regional Product.

BEM. Bridge Equation Model.

MIDAS. Mixed-Data Sampling.

LASSO. Least Absolute Shrinkage and Selection Operator.

VAR. Vector Autoregression.

AR. Autoregression.

ARIMA. Autoregressive Integrated Moving Average.

ADL. Autoregressive Distributed Lag.

NBU. National Bank of Ukraine.

SSSU. State Statistics Service of Ukraine.

RMSE. Root Mean Square Errors.

MAE. Mean Absolute Error.

VAT. Value Added Tax.

Chapter 1

INTRODUCTION

The economic impact of the 2022 invasion of Ukraine by russia has been devastating and unprecedented. The war has disrupted the normal functioning of the economy, creating uncertainty and volatility in various sectors and regions. The availability and reliability of official economic statistics have also been affected, making it difficult to assess the current and future state of the economy. In such a situation, timely and accurate economic data are essential for policymakers, businesses, and citizens to make informed decisions and cope with the crisis.

However, traditional economic models that rely on stable and predictable data are not well suited for capturing rapid and complex economic changes. These models often use low-frequency data, such as quarterly GDP, that are released with a considerable lag and cannot reflect real-time economic conditions. Moreover, these models could not account for the heterogeneity and nonlinearity of the economic relationships during crisis periods.

To address these limitations, this thesis reexamines the nowcasting methodology, a well-established technique used for over a decade, and seeks to enhance it by integrating contemporary high-frequency data sources. Nowcasting is a key method in modern economic analysis, especially when economic conditions change quickly, and conventional data sources are insufficient or outdated. This approach uses high-frequency data, such as internet search trends, social media sentiment, and transactional data, to provide a more immediate and comprehensive picture of economic activity. In the context of Ukraine, where traditional data sources can be disrupted or delayed due to the ongoing war, nowcasting offers a dynamic and adaptive approach to tracking the economy.

The main objective of this thesis is to improve Ukraine's GDP nowcasting by incorporating high-frequency digital data, such as search engine queries, into a mixed-frequency model. The research focus of this paper is to evaluate the impact of the data frequency on the accuracy and responsiveness of GDP nowcasting for Ukraine. The hypothesis is that higher-frequency monthly data can better capture subtle signals and fluctuations in the economy despite being noisier and more volatile, while lower-frequency quarterly data can provide more stable and consistent estimates but may miss some critical information. To test this hypothesis, this thesis uses a bridge equation model (BEM), a flexible nowcasting technique, and compares the performance of different model specifications using monthly and quarterly alternative data.

This research has demonstrated the value of incorporating Google Trends data to enhance the accuracy and timeliness of GDP nowcasting. By integrating high-frequency digital data with conventional economic statistics within a mixed-frequency modeling framework, the research has captured real-time economic fluctuations. The inclusion of Google Trends in the nowcasting models has revealed potential in predicting short-term economic series, underscoring the role of digital data in crisis contexts where rapid economic assessments are crucial for informed decision-making.

This thesis is structured as follows: Chapter 2 provides a literature review on the theory and practice of nowcasting, focusing on the use of digital data and the challenges and opportunities of nowcasting in crises. Chapter 3 describes the methodology used in this study. Chapter 4 describes the data sources, including the selection and transformation of the variables. Chapter 5 presents the estimation of the BEM and discusses the main results and findings of the empirical analysis and nowcasting Ukraine's GDP. Chapter 6 concludes the thesis with a summary of the main contributions, policy implications, and limitations of the study and suggests some directions for future research.

C h a p t e r 2

LITERATURE REVIEW

Nowcasting is a term coined by Bańbura et al. (2013) to describe a method that predicts the current, the near-term, or the recent past state of the economy using high-frequency data indicators. These indicators are available in real-time or with a short delay and can capture the economic conditions more promptly than low-frequency data, such as quarterly GDP. Nowcasting aims to produce accurate estimates of the current economic activity and provide timely information to policymakers and economists for effective decision-making (Giannone et al. 2008). This method is particularly relevant when the economy undergoes rapid and complex changes that demand fast and flexible responses.

The emergence of nowcasting began decades ago alongside the exponential growth of big data. As economies have expanded, higher volumes of data have enabled more profound insights into the behavior of economic agents, enhancing GDP estimates by illuminating the economy's current state.

A common nowcasting approach surveys households, firms, banks, and institutions to gauge business activity and expectations as predictive economic trajectory and risks indicators. In a study analyzing the use of Euro area surveys for GDP nowcasting, Basselier et al. (2018) investigated applying such data. They determined that manufacturing and business climate indicators significantly improved accuracy. Surveys' strength lies in capturing qualitative, sector-specific attitudes about the economy. However, the study found consumer confidence oddly disconnected from the actual economic performance. It highlights the complexity of indicator selection, as predictive power varies by survey type. Overall, it reveals both the potential and limitations inherent in leveraging surveys for small open economies' nowcasting.

Ukraine's central bank has demonstrated the strategic adoption of surveys for GDP nowcasting, as discussed by Lysenko and Kolesnichenko (2016). Incorporating Business Outlook Survey results into econometric models generates short-term output predictions like GDP estimates. Accessing timely business sentiments and expectations through surveys facilitates responsive policymaking. However, biases may emerge, with subjective responses not always reflecting reality. Moreover, variability in volatile periods impacts the consistency and reliability of projections. Additionally, conducting comprehensive surveys can prove impossible amidst war.

Given the significant limitations of survey data for GDP nowcasting, including subjectivity, delayed reporting, and resource intensiveness, dynamic and high-frequency data sources have emerged as essential alternatives during turbulent periods of disrupted data collection, most critically amid conflicts and wars. Traditional economic surveys face acute hurdles under such conditions. Thus, continuous and widely available data streams provide a pathway to overcoming restricted traditional statistics. News-based data offers more real-time economic visibility, offering a broader scope of events. The use of news-based indices represents a significant advancement in the field of economic nowcasting, leading to more accurate and timely assessments of GDP.

Thorsrud (2016) made a compelling case for news-based GDP nowcasting over traditional models, showing a superior performance by as much as 15% during volatile turning points in business cycles. This suggests that it is uniquely capable of tracking rapid shifts. In 2018, Thorsrud expanded on this by demonstrating that news-based indices competed well against top forecast combinations and judgments for GDP prediction. Meanwhile, Ashwin et al. (2021) revealed that news could improve Euro area GDP nowcasts, catching early-quarter signals. However, their euro-centric focus hinders their application to developing economies.

While news-based nowcasting boasts speed and sensitivity, capturing economic subtleties, this approach faces acute challenges in turbulent regions like war-ridden Ukraine. Specifically, propaganda and oligarchic media control severely distort the wartime news landscape. State and non-state actors intentionally skew coverage to misportray conditions, advance interests, and even censor vital economic data. Consequently, the veracity and objectivity of reporting suffer, undercutting indices' reliability.

However, even with the advantages of news-based indices, there is a continual search for even more dynamic and predictive tools. In this quest, the emergence of Google Trends as a tool for enhancing GDP nowcasting presents a novel perspective. As digital search continues to gain popularity for gauging consumer activity, researchers have increasingly leveraged Google Trends data to forecast economic indicators from private consumption (Vosen and Schmidt 2011) to automobile sales (Carrière-Swallow and Labbé 2013). Building on this, D'Amuri and Marcucci (2017) demonstrate the value of Google Trends data for improving the accuracy of unemployment rate forecasts in the United States. Their findings highlight the unique capacity of search query activity to capture shifts in underlying economic conditions with immediacy and granularity that surpass traditional data sources. These studies establish search query volume as a timely proxy for tracking both consumer behavior and broader economic vitality.

However, can the search query data describe GDP components in detail to strengthen official data in nowcasting the GDP itself? Bantis et al. (2021) specifically examine the efficacy of Google Trends in nowcasting and predicting GDP growth across developed and developing economies. They determine more significant nowcasting improvements for the United States compared to Brazil, suggesting variability in effectiveness based on a country's economic attributes. This indicates the potential of search query data to enhance GDP predictions, contingent on the national data landscape.

Ferrara and Simoni (2019) provide a valuable contribution by employing bridge equation models to take advantage of the early availability of Google Trends data within a quarter. Their approach demonstrates the predictive capacity of digital search queries before conventional economic indicators are released. However, they note that the forecasting accuracy of Google Trends declines as more traditional statistical data becomes available over a quarter, which limits its ongoing utility for continuous economic monitoring.

Similarly, Götz and Knetsch (2019) utilize bridge equation models to improve the nowcasting of German GDP based on Google Trends data. Their research shows how this unconventional data source can capture economic fluctuations from unexpected events, offering a timelier alternative to traditional nowcasting techniques. Though capable of detecting these signals, their study suggests Google Trends data also contains noise, requiring careful modeling to derive meaningful economic insights.

The application of Google Trends in the realm of nowcasting exhibits a complex and context-dependent utility. While they offer a more accessible and immediate data source compared to conventional surveys, their effectiveness is not uniform across all situations. The highlighted literature suggests that Google Trends can augment other high-frequency economic indicators to enhance GDP predictions. However, their performance is not consistently superior as it varies by country and over time. In well-structured economies, they have shown marked improvements in nowcasting accuracy, but they do not consistently outperform traditional indicators. Within the context of Ukraine, given the ongoing war, Google Trends could provide valuable real-time insights when traditional data sources are compromised. They could be used alongside both 'hard' indicators, like production statistics, and 'soft' indicators, such as consumer sentiment, to offer a composite and current view of the economy.

The above-mentioned studies leverage variations of bridge equation models (BEMs) commonly used in economic nowcasting. Klein and Sojo (1989) pioneered the concept of Bridge Equation Models, deploying them as a regression system aimed at forecasting GDP growth. This approach models each component of the National Accounts separately, allowing for a more granular analysis of GDP dynamics. The main idea of this model is to "bridge" the gap between high-frequency data, such as monthly or weekly indicators, and low-frequency data, such as quarterly or annual GDP. This framework presents a compelling case for their efficacy over alternative modeling techniques such as MIDAS, VAR, and factor models. Schumacher (2014) discusses the benefits of BEMs for forecasting, noting their ability to incorporate high-frequency data such as Google Trends while mitigating overfitting and stability issues. This framework provides a more direct linkage between current indicators and lower-frequency outcomes like quarterly GDP, a useful capability for timely predictions.

Furthermore, an analysis by Smith (2014) compares model efficacy for nowcasting UK GDP during recession periods. The study determines that BEMs achieve end-quarter nowcasting accuracy similar to the often more intricate MIDAS and factor models. This comparable precision implies that simpler BEMs could be an efficient substitute, especially when prioritizing straightforward implementation and computational practicality.

With highlighting the strengths of BEM in nowcasting, it is also crucial to consider the challenges inherent in the data sources these models employ. Specifically, when incorporating Google Trends data, researchers face unique hurdles. Shi and France (2018) underscore the necessity of aggregating and combining Google Trends search volume data to maintain consistency, an endeavor that becomes particularly challenging in unstable environments, such as war-torn countries where economic data may be erratic or scarce. This suggests that while Google Trends can provide timely insights, the data must be handled carefully to ensure its reliability. Meanwhile, Constantinescu (2023) adds a cautionary note on the unexamined escalation of alternative data like Google Trends in nowcasting. The study warns of the risks of incorporating such data without thorough analysis, emphasizing the need to sift through the white noise and irrelevant signals that can distort the economic picture. This transition in the narrative underscores the importance of not only choosing the suitable nowcasting model but also ensuring the data it relies on is robust and accurately interpreted.

Enriched by the literature reviewed, the evolving landscape of GDP nowcasting compels us to adopt an interdisciplinary approach that harmonizes high-frequency digital indicators with traditional economic measures. This synthesis not only enhances the robustness of real-time economic assessments but also serves as a bulwark against the data irregularities imposed by geopolitical upheavals. Consequently, the future of economic forecasting is envisioned as a confluence of established econometric models and emergent data sources, meticulously curated and critically evaluated to inform the discerning analyses of contemporary economists and policymakers.

This thesis contributes by applying advanced nowcasting techniques to enhance real-time tracking of Ukraine's GDP amidst the disruptions caused by the ongoing war. It builds upon previous work using Google Trends data and bridge equation models for GDP nowcasting but extends the approach to a crisis-affected developing economy where conventional data sources are compromised. By integrating high-frequency digital indicators like Google search queries with other economic data in a mixed-frequency bridge model, the research evaluates whether the use of monthly Google Trends can improve nowcasting accuracy for Ukraine's key GDP components. The findings will offer insights into the relative value of different data frequencies for nowcasting in volatile environments, addressing a gap in the literature's focus on more stable, developed economies.

Chapter 3

METHODOLOGY

Building on the findings of the Literature Review, this research aims to improve the accuracy of short-term forecasts for Ukrainian GDP and its components by incorporating Google Trends data. This approach is similar to the one used by Götz and Knetsch (2019) for Germany. However, due to the limited availability of high-frequency survey data in Ukraine — which was primarily used for building the baseline model to compare efficiency with alternative data — we will focus only on comparing the performances of models with different Google Trends frequencies.

An important point to highlight is that the authors define two types of monthly indicators used in the model: "hard" and "soft." Hard indicators refer to objective, quantitative economic data such as industrial production or retail sales figures obtained from official statistical sources. In contrast, soft indicators such as surveys and sentiment indices provide an early, subjective glimpse into economic activity before official data is released. Notably, this soft data is always available from non-official online sources.

The first step involves estimating a model for the soft indicators. Following the approach of Götz and Knetsch (2019), an autoregressive process (AR) was employed to estimate and forecast future trends. However, to ensure model stability, the data was first transformed to achieve stationarity, as recommended by the authors. While forecasting the next 12 months using the AR model, the predictions resulted in a straight line representing the mean, particularly zero. Therefore, to capture both the level and dynamics of the series more effectively, an autoregressive integrated moving average (ARIMA) model was used instead.

$$\hat{G}_{mt} = \mu_G + \phi_G(L^{1/3})G_{m,t-1/3} + \theta_G(L^{1/3})\epsilon_{m,t-1/3} + \epsilon_t , \qquad (1)$$

where $t = 1, \ldots T$ denotes the time periods; $\hat{G}_{mt}$ is a vector of differenced estimated soft indicators; $G_{m,t-1/3}$ is a vector of the differenced soft indicators' lagged values of 1 month; $\phi_G = \sum_{i=0}^{p-1} \phi_{G,i+1} L^i$ is the lag operator with coefficients $\phi_{G,i+1}$ measuring the influence over p lags of Google Trends values on the current month; $\theta_G = \sum_{i=0}^{q-1} \theta_{G,i+1} L^i$ captures the moving average component, which represents the effect of the previous month's forecast error on the current month's forecast; $\epsilon_{m,t-1/3}$ is the vector of the lagged differences of the forecast errors; $L^{1/3}$ is a high-frequency lag operator, which is the modification of the quarter lag operator L for further fitting the high-frequency data into the lower frequency model structure.

The second step involves estimating models for the hard indicators, similar to the approach taken for the soft indicators. However, in addition to using the lagged values of the hard indicators themselves, the model also incorporates the estimated soft indicators from the previous step as predictors. In this case, an autoregressive distributed lag (ADL) is employed. It is crucial to ensure stationarity in the time series to avoid spurious regression and obtain reliable results when using the ADL model. Therefore, the trend component was removed from the estimated Google Trends categories series and the transformed hard indicator series if needed. The transformation applied to the hard indicators was the year-over-year difference of logarithms, which represents the indicators' percentage change over the year.

$$\hat{X}_{mt} = \mu_X + \rho_X(L^{1/3})X_{m,t-1/3} + \delta_x(L^{1/3})\hat{G}_{mt} + \epsilon_{tX}, \qquad (2)$$

where $\hat{X}_{mt}$ is a vector of monthly hard indicators; $X_{m,t-1/3}$ is a vector of the hard indicators' lagged values from 1 to 3 months prior; $\hat{G}_{mt}$ is a vector of estimated soft indicators from the previous step; $\rho_X(L) = \sum_{i=0}^{p-1} \rho_{X,i+1} L^i$ and $\delta_x(L^{1/3}) = \sum_{i=0}^{q} \delta_i L^i$ are lag operator polynomials capturing the influence of past hard indicator values through coefficients $\rho_{X,i}$ and the influence of soft indicators through coefficients δ_x both using lags at the monthly frequency.

The Google Trends categories are selected using the LASSO method. This method performs both variable selection and regularization to enhance the prediction accuracy and interpretability of the statistical model. LASSO solves some of the limitations of traditional regression methods by imposing a constraint on the sum of the absolute values of the model parameters. This constraint allows not only shrinking coefficients towards zero but also setting some coefficients exactly to zero, thus effectively selecting a simpler, more interpretable model from a potentially large set of predictors. The function which LASSO minimizes has the form:

$$\sum_{i=1}^{n} (X_i - \beta_0 - \sum_{j=1}^{p} \beta_j \hat{G}_{ij})^2 + \lambda \sum_{j=1}^{p} |\beta_j|, \qquad (3)$$

where X_i is a vector of observed hard indicators; $\hat{G}_{ij}$ is a vector of previously modeled Google Trends; β_0 is an intercept; β_j are the coefficients for the predictors G_{ij}; λ is the regularization parameter that controls the strength of the penalty applied to the coefficients; n is a number of observations and p is the number of predictors.

The final step involves estimating the benchmark equation that links the GDP components to the modeled hard indicators. This benchmark model takes the form of a dynamic linear equation relating the quarterly growth rates of GDP components to their own lags as well as the lags of the estimated hard indicators:

$$\hat{Y}_t = \mu_y + \rho_Y(L)Y_{t-1} + \beta(L)\hat{X}_{qt} + \epsilon_{ty}, \tag{4}$$

where $\hat{Y}_t$ is a vector of estimated growth rates of GDP components; Y_{t-1} is a vector of lagged values of these quarterly GDP components; $\hat{X}_{qt}$ is the vector of estimated hard indicators aggregated to a quarterly frequency; $\rho_Y(L) = \sum_{i=0}^{p-1} \rho_{Y,i+1} L^i$ is a lag polynomial capturing the influence of past GDP component values through coefficients $\rho_{Y,i+1}$; $\beta(L) = \sum_{i=0}^{q} \beta_i L^i$ is a lag polynomial measuring how current and past values of the hard indicators impact the GDP components through coefficients β_i.

To examine the accuracy of the final models' predictions, loss functions such as the mean absolute error (MAE) and root mean squared error (RMSE) will be employed. MAE represents the average of the absolute differences between the actual and predicted values. This metric treats all errors equally, regardless of their magnitude, making it less sensitive to outliers.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \hat{Y}_i|, \tag{5}$$

In turn, RMSE gives more weight to larger errors due to the squaring operation, making it more sensitive to outliers. It can be mathematically represented as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}. \tag{6}$$

C h a p t e r 4

DATA

The key variables of interest are the 10 main components that make up Ukraine's GDP, which is calculated using the production approach, along with related official high-frequency economic indicators and data from Google Trends. The quarterly GDP component values and monthly indicators such as industrial production, retail trade, construction activity, and foreign trade statistics are obtained from publicly available databases maintained by the National Bank of Ukraine (NBU) and the State Statistics Service of Ukraine (SSSU).

Additionally, monthly Google search query data is collected across 62 categories considered as potentially relevant to economic activity based on the hypothesis that they encompass major industries, consumer sectors, and public services. These categories include Construction & Maintenance, Nuclear Energy, Industrial Materials & Equipment, Property Development, Consumer Electronics, Military, Politics, Restaurants, Food Production, Business Operations, Shopping Portals, etc. All data series, including the Google Trends information, are aggregated at the national level for Ukraine.

However, some challenges have arisen since the origin of Google search data formation. This data can be analyzed at different levels of specificity, with keywords being the most granular and categories being the broadest. Categories aggregate multiple topics within a general sector, providing a macro-level overview of trends but lacking the details available from narrower keywords or topics, which cluster multiple related keywords. While more specific keywords and topics clarify precise changes in narrower sub-segments or individual products and services, the category-level data indicates broad shifts across an industry or consumer segment. Therefore, categories are best suited for generalized tracking of high-level economic and consumer trends across a sector. Nevertheless, in some cases, they can overgeneralize and show inconsistency and unavailability of useful categories due to Google's conditions and policy changes. For these cases, alternative high-frequency data should be examined and included, such as truck registration instead of the 'Trucks & SUVs' category, which contains inappropriate topics and search queries.

The study period spans January 2015 to December 2023, covering major economic shocks experienced by Ukraine over the past decade. This nine-year period allows assessment of model performance across normal periods as well as volatile crisis events.

The focus is on estimating the 10 biggest quarterly GDP components: Mining and Quarrying, Manufacturing, Electricity and Gas Supply, Construction, Energy and Water Supply, Transportation, Agriculture, forestry and fisheries, Public Administration and Defense, and Taxes on products (VAT). Estimating these components employs related monthly series, including the volume of industrial production sold in mining and manufacturing, electricity and water supply metrics, construction production volumes, retail turnover, price index of agricultural products, monthly government spending, and VAT. Tables 1 and 2 present the descriptive statistics for the GDP components and their related series, respectively

To enable a thorough analysis, we transform the GDP components and their related hard indicators into logarithmic differences. This approach not only normalizes the data but also allows us to calculate the growth rates of these variables. Furthermore, we adjust all values to 2015 prices to ensure consistency in our comparisons.

Table 1. Descriptive statistics of GDP components

Statistic	Period Start	Period End	N	Mean	St.Dev	Min	Max
GDP Components							
ΔLog(Mining and quarrying)	2015:Q1	2023:Q4	36	-0.03	0.1	-0.5	0.2
ΔLog(Manufacturing)	2015:Q1	2023:Q4	36	-0.03	0.1	-0.4	0.3
ΔLog(Electricity supply)	2015:Q1	2023:Q4	36	-0.1	0.2	-0.5	0.1
ΔLog(Water Supply)	2015:Q1	2023:Q4	36	-0.1	0.2	-0.7	0.3
ΔLog(Construction)	2015:Q1	2023:Q4	36	-0.1	0.2	-0.5	0.1
ΔLog(Retail Sales)	2015:Q1	2023:Q4	36	-0.1	0.1	-0.5	0.2
ΔLog(Transport)	2015:Q1	2023:Q4	36	-0.04	0.4	-1.4	0.5
ΔLog(Agriculture)	2015:Q1	2023:Q4	36	-0.03	0.1	-0.4	0.3
ΔLog(Public Administration)	2015:Q1	2023:Q4	36	0.05	0.1	-0.1	0.5
ΔLog(VAT)	2015:Q1	2023:Q4	36	-0.1	0.2	-0.6	0.2

Table 2. Descriptive statistics of economic series related to GDP components

Statistic	Period Start	Period End	N	Mean	St.Dev	Min	Max
Hard Indicators							
ΔLog(Mining Production)	2015:Q1	2023:Q4	36	-0.04	0.4	-0.9	0.8
ΔLog(Industrial Production)	2015:Q1	2023:Q4	36	-0.1	0.3	-0.8	0.5
ΔLog(Electricity Production)	2015:Q1	2023:Q4	36	0.03	0.3	-0.6	0.7
ΔLog(Water Supply)	2015:Q1	2023:Q4	36	-0.02	0.4	-1.4	0.5
ΔLog(Construction Production)	2015:Q1	2023:Q4	36	-0.1	0.2	-0.7	0.4
ΔLog(Retail Turnover)	2015:Q1	2023:Q4	36	-0.1	0.2	-0.7	0.2
ΔLog(Transport)	2015:Q1	2023:Q4	36	-0.05	0.2	-0.6	0.2
ΔLog(Price Indices of Agricultur: Products)	2015:Q1	2022:Q4	32	-0.2	0.2	-0.7	0.1
ΔLog(Government Spending)	2015:Q1	2023:Q4	36	0.1	0.2	-0.2	0.9
ΔLog(Taxes on Products)	2015:Q1	2023:Q4	36	0.01	0.2	-0.4	0.3

In contrast, Google Trends data doesn't require specific transformations. Google Trends uses query shares to measure the relative popularity of a query category over time. A query share of 1 means that the category reached its peak frequency, while a query share of 0 means that the category had no searches in that period. To ensure statistical consistency and model stability, the data was made stationary. Table 2 shows the quarterly-aggregated categories corresponding to the variables and their descriptive statistics.

Table 3. Descriptive statistics of Google Trends categories

Statistic	**Period Start**	**Period End**	**N**	**Mean**	**St.Dev**	**Min**	**Max**
Google Trend Categories							
Energy & Utilities	2015:Q1	2023:Q4	36	0.5	0.1	0.4	0.8
Water Supply & Treatment	2015:Q1	2023:Q4	36	0.5	0.1	0.3	0.8
Nuclear Energy	2015:Q1	2023:Q4	36	0.4	0.1	0.3	0.8
Food Grocery Retailers	2015:Q1	2023:Q4	36	0.5	0.1	0.3	0.9
Retail Trade	2015:Q1	2023:Q4	36	0.6	0.1	0.4	1.0
Shopping Portals	2015:Q1	2023:Q4	36	0.5	0.1	0.4	0.9
Construction & Maintenance	2015:Q2	2023:Q4	36	0.8	0.1	0.6	1.0
Gardedning	2015:Q3	2023:Q4	36	0.6	0.1	0.4	0.9
Manufacturing	2015:Q1	2023:Q4	36	0.8	0.1	0.6	0.9

Before delving into the estimation details, it is insightful to visually examine the relationship between the Google Trends query categories and the official economic series. The reason is that categories could exhibit similar patterns of peaks and troughs as the official economic indicators.

This visual inspection will help to indicate which categories could potentially serve as relevant predictors during the formal model estimation process. The graphs below depict the trends over time for the selected Google Trends categories alongside the corresponding official economic series from national statistics. A few examples are shown below.

As evident from Figure 1, only the 'Construction & Maintenance' and 'Property Development' categories exhibit trend patterns similar to the official Construction Production series. The 'Gardening' category, however, displays strong seasonality and deviates from the overall trajectory of the construction indicator, suggesting a weaker relationship.

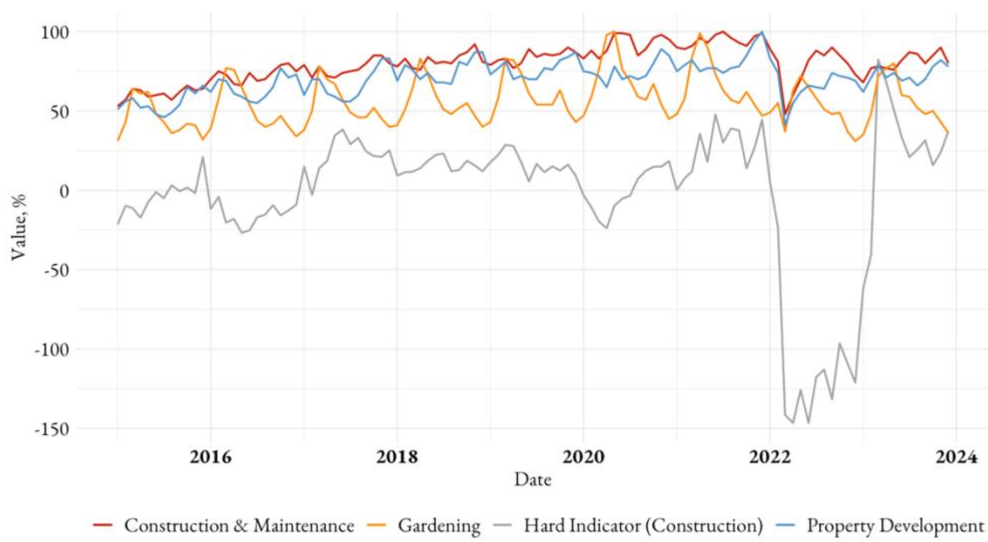

Figure 1. The dynamics of Construction and related Google Trends categories

Figure 2 illustrates a relatively weaker correlation between the selected query categories and the corresponding economic series for most of the time period. However, it is noteworthy that these categories exhibited a significant decline in February 2022, mirroring the overall drop observed in the official sales data.

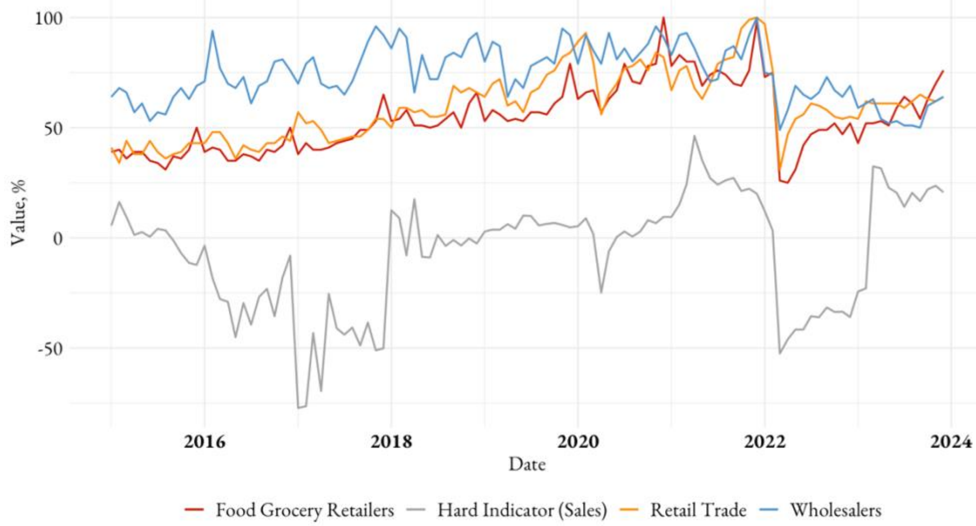

Figure 2. The dynamics of Sales and related Google Trends categories

This synchronous pattern suggests that the Google Trends categories could still possess predictive value in capturing shocks in the economic variable despite their inability to track the overall trend consistently.

Chapter 5

ESTIMATION RESULTS

This chapter presents the estimation results, concentrating on the models for the hard indicators and the bridge equation models linking these indicators to the GDP components. The estimation of soft indicators is not covered in detail, as these models follow the ARIMA process without providing additional insightful information. The core value lies in examining how the hard economic indicators and Google Trends data can enhance the nowcasting of the official GDP components.

We will first assess the accuracy of the in-sample GDP predictions, which are presented in Table 4. The estimations and predictions were made based on training and test samples. The training sample included all the data from 2015 to 2021, while the test sample covered the data from 2022 to 2023. The evaluation metrics reported below are calculated based on the test sample.

Table 4. Loss Function Results

GDP Component	**MAE**	**RMSE**
Mining and quarrying	0.119	0.152
Manufacturing	0.148	0.177
Electricity Supply	0.096	0.111
Water Supply	0.055	0.069
Construction	0.222	0.278
Retail Sales	0.135	0.169
Transport	0.156	0.192
Agriculture	0.184	0.218
Public Administration	0.052	0.074
VAT	0.118	0.138

The maximum RMSE of 0.278, or 27.8%, has the model of construction GDP component growth, while the minimum RMSE of 0.074, or 7.4%, has the model of public administration GDP component growth. This suggests that the models with high RMSE could not effectively capture the underlying dynamics and factors influencing the GDP component growth. Figure 3 and Figure 4 show the actual and predicted values of the construction and public administration growth rates, respectively. The in-sample predictions for all GDP components are provided in Appendix A.

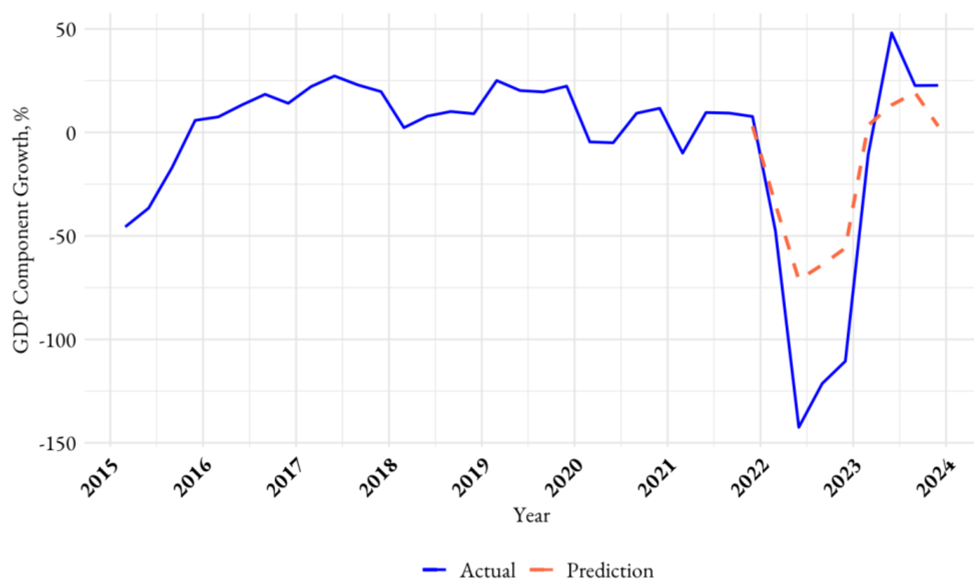

Figure 3. The actual and fitted values of the GDP construction growth rate

Figure 3 illustrates that the model partially captures the general trend of the economic downturn and subsequent recovery but fails to predict the magnitude of these fluctuations accurately. The predictions show a smoothing effect, where extreme lows and highs are moderated. However, the drastic drop in construction growth during 2022 was so significant that the model could not capture it. This suggests that while the model is generally well-calibrated to the average historical data, it may underreact to sudden economic shifts and extreme values, which the first year of the war is characterized by.

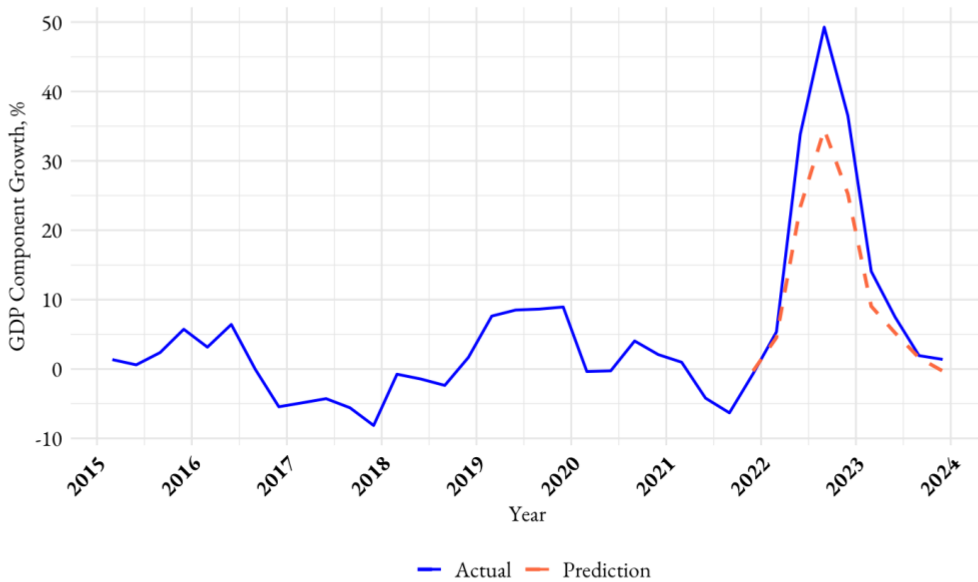

Figure 4. The actual and fitted values of the GDP public administration growth rate

One might expect the situation with public administration predictions to be more favorable. Figure 4 shows that the model effectively predicts the sharp rise and fall. However, similar to the previous case, the model struggles with the extremities of fluctuations. The RMSE value of 7.4% in predictions is quite significant, but given that the changes include a massive spike up to nearly 50%, an RMSE value suggests that the model has a reasonable degree of accuracy. It manages to capture the general movements well, which is challenging in economic series with such unusual fluctuations.

The models for estimating all components were built based on the estimated monthly related series, which incorporated Google Trends categories. Table 5 provides an overview of the hard indicators and associated categories selected by the LASSO method.

Table 5. Hard Indicators and Selected Google Trend Categories

Hard Indicator	**Google Trend Category**
Mining Production	Metals & Mining, Industrial Materials, Nuclear Energy, Manufacturing
Industrial Production	Automotive Industry, Manufacturing
Energy Supply	HVAC & Climate Control, Energy & Utilities, Consumer Electronics
Water Supply	Energy & Utilities, Nuclear Energy, Crops & Seeds
Construction Production	Construction & Maintenance, Gardening, Property Development, Apartments
Retail Turnover	Food & Grocery Retailers, Retail Trade, Shopping Portals
Transport	Bus & Rail, Trucks & SUVs, Transportation
Price Indices of Agricultural Products	Chemicals Industry, Crops & Seeds, Agricultural Equipment
Government Spending	Local Government, Army, Politics, Multilateral Organizations
VAT	Business Finance, Mass Merchants, Shopping Portals

The LASSO model's effectiveness in variable selection is evident from the intuitive association of the chosen Google Trends categories with the corresponding economic indicators, reflecting the anticipated economic activities. Categories like "Metals & Mining" and "Industrial Materials" are directly related to mining activities, suggesting a strong predictive relevance for changes in mining production levels. The selection of "Automotive Industry" and "Manufacturing" aligns well with the broader industrial production sector, indicating that trends in these areas are closely tied to industrial output. "HVAC & Climate Control" and "Energy & Utilities" reflect consumer and industrial demand for energy, which correlates with overall energy supply dynamics. The choice of "Construction & Maintenance" and "Gardening" reflects the broader range of construction activities, from large-scale developments to residential and small-scale construction. "Food & Grocery Retailers" and "Shopping Portals" indicate consumer spending patterns, which are a key component of retail economic activity. The association with "Local Government" and "Multilateral Organizations" indicates an interest in public sector activities and expenditures, which are directly related to government spending levels. "Business Finance" and "Mass Merchants" suggest that VAT-related searches could be linked to business transactions and larger retail operations, impacting VAT collections.

The model's selection of categories, such as "Nuclear Energy" for "Water Supply" and "Gardening" for "Construction Production" might initially seem unconventional. However, these selections can be insightful.

"Nuclear Energy" in the context of "Water Supply" could reflect the significant role of energy in water treatment and distribution facilities, many of which are energy-intensive operations. Similarly, the "Gardening" category for "Construction Production" might capture consumer trends related to home improvement and small-scale construction projects.

Now we are to examine the forecasts of hard indicators, which served as the foundation for the subsequent GDP forecasts. We will proceed with a comparison of the construction and public administration sectors. Figures 5 and 6 present the twelve-month forecasts for construction production and government spending, respectively. All estimation results and forecast graphs for the hard indicators are provided in Appendix B.

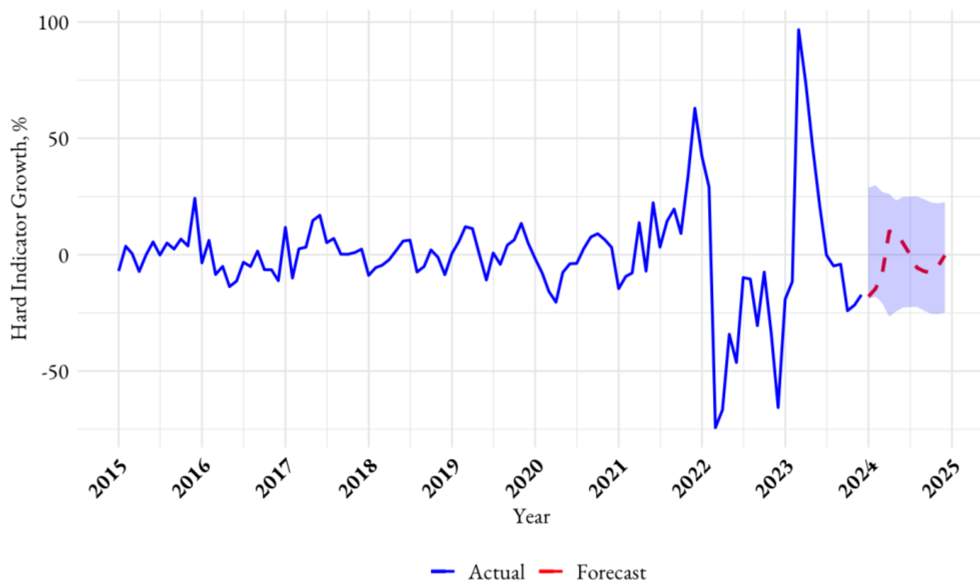

Figure 5. The forecast of the monthly construction production

Figure 5 demonstrates that the model captures the general trend of contraction of construction growth following a peak period. This could be interpreted as the model predicting a market correction or a normalization of construction activity after an unsustainable growth spurt in 2023. The predicted values indicate a sharp decline in early 2024, followed by stabilization. The uncertainty ranges from -25% to +25% growth.

The actual historical data presented in Figure 6 highlights the inherent volatility in the public administration and defence sector over the past decade. In contrast, the twelve-month forecast indicates a more stable growth trajectory. However, the confidence intervals reveal a wide uncertainty around the forecast, with a maximum drop of 25%. This uncertainty can be attributed to the unstable societal situation caused by the unpredictability on the battlefield. This, in turn, is reflected in the Google queries. One of the reasons is that Ukraine is still at war, and the actions on the battlefield depend on external support and the invaders' decisions, whether they will advance more actively or whether the Ukrainian army will counterattack. The situation within the country is also contingent on the enemy's actions. Government spending can be affected by the shelling of cities, upcoming reforms, etc. Consequently, the Google Trends data should be updated monthly, and forecasts should be made for a shorter period, such as approximately three months.

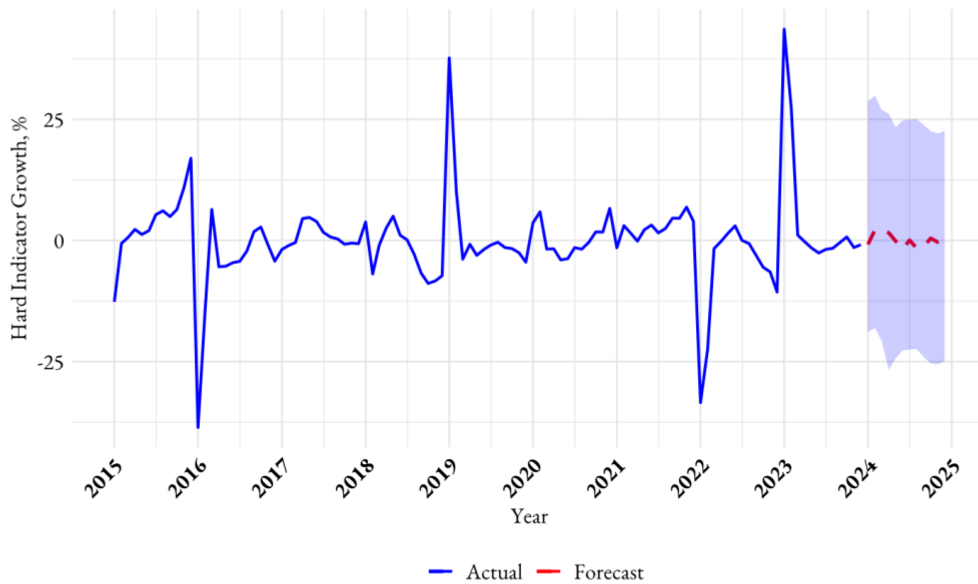

Figure 6. The forecast of the monthly government spending

Finally, we construct the forecasts of GDP components based on the aggregated estimations and forecasts of hard indicators. Figure 7 displays the actual and predicted growth rates for the GDP construction component from 2015 through 2025. The graph shows a drastic decline in construction activity in 2022, corresponding with the full-scale invasion. This downturn is followed by a vigorous recovery in 2023, potentially indicative of reconstruction efforts and increased infrastructure spending as part of recovery initiatives. Starting in 2024, the model predicts a stabilization of growth rates but with a tendency toward decline. This forecast reflects the cautious optimism of gradual recovery tempered by the persistent uncertainties and challenges posed by the war.

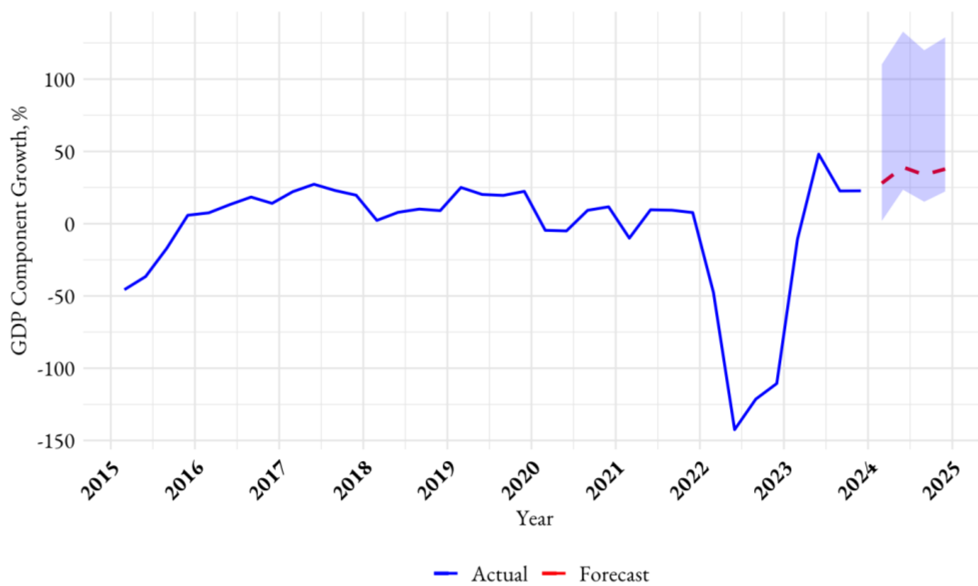

Figure 7. The forecast of the GDP construction component

The actual and forecasted growth rates of the GDP component for public administration and defense in Ukraine from 2015 through 2025 are represented in Figure 8. The actual data indicate considerable fluctuations, with a notable peak in 2022 and a sharp decline in 2023. This significant rise and subsequent fall could be associated with increased government spending during the initial recovery from the full-scale invasion and counteroffensive in the Kyiv and Kharkiv regions, such as enhanced public services and infrastructure investments.

The model forecasts a decrease in growth rates starting in 2024, stabilizing at lower levels into 2025. This projected decline suggests an expected reduction in public administration activities as the intense period of crisis management subsides. The shaded area around the forecast highlights the uncertainty involved, with a relatively wide range suggesting that future growth rates could vary significantly depending on evolving economic conditions and the political landscape. The GDP forecasts and graphs can be found in Appendix C.

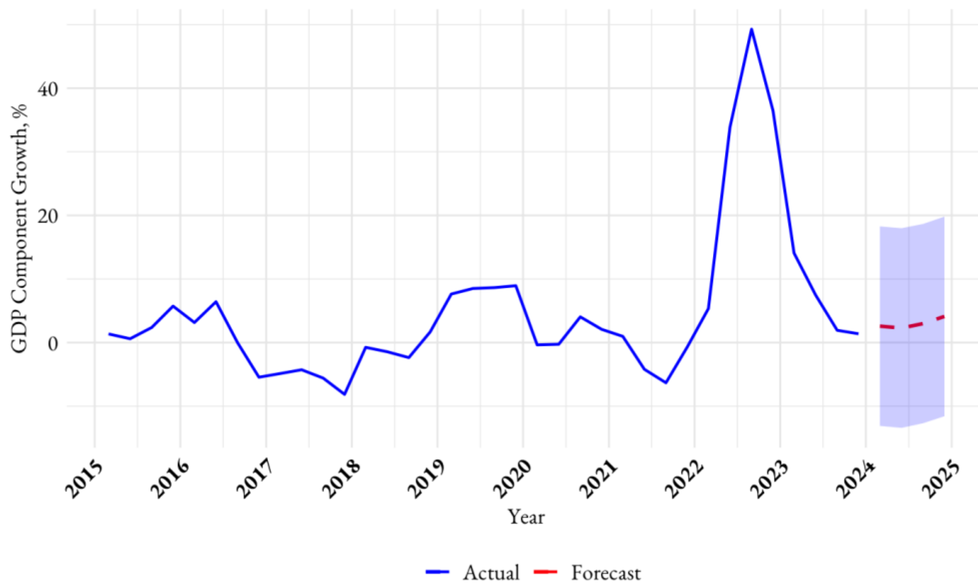

Figure 8. The forecast of the GDP public administration component

In summary, Google Trends demonstrates its usefulness in economic series predicting economic series during times of crisis, as it can capture real-time shifts in public interest and economic activity that traditional economic indicators might miss. The search query categories can predict the dynamics of the official series, but they still struggle with predicting the magnitude accurately. Nevertheless, it is important to remember that incorporating Google Trends data and monthly official series could help only in short-term forecasts, and the data should be updated as soon as it becomes available.

C h a p t e r 6

CONCLUSIONS AND POLICY RECOMMENDATIONS

This thesis evaluated the utility of Google Trends data, categorized more broadly than individual search topics, in enhancing the short-term forecasting of Ukraine's GDP using a mixed-frequency model. The analysis concentrated on ten GDP components, demonstrating that when integrated with traditional economic statistics, high-frequency digital data offers a nuanced understanding of economic dynamics in real time, which is particularly valuable in a crisis context like that in Ukraine.

The integration of Google Trends has proven to be a valuable addition to the economic nowcasting toolkit, providing timely insights into shifts in economic activities that are not immediately apparent through conventional data sources. The findings underscore the potential of search data to act as an early indicator of economic trends, particularly in a highly volatile environment where traditional data may lag or be unavailable.

Key conclusions drawn from the thesis include the predictive relevance of Google Trends, which showed significant potential in predicting the short-term movements of GDP components. These high-frequency digital indicators are especially useful in capturing immediate economic sentiments and activities, which are crucial during periods of uncertainty and rapid change. The BEM employed in this study effectively utilized the high-frequency nature of Google Trends data alongside traditional economic measures, allowing for more responsive nowcasting of GDP components, reflecting both sudden economic shifts and longer-term trends. While the models provided valuable forecasts, they also highlighted the inherent volatility and uncertainty in economic data during crises.

Based on the findings, we can conclude that the Google Trends categories should be used carefully alongside the main predictors, not as the main players. Moreover, there are limitations on using categories instead of specific search query topics, as there is a high chance of misclassifying searches because of wrong translation or interpretation. Therefore, only in times of high uncertainty and the unavailability of the official data, the Google Trends can be used as a proxy for economic activity in main sectors of the country's economy.

The findings from this research open several avenues for further research that could deepen our understanding of high-frequency data's role in economic forecasting. One promising area is extending the analysis to include a regional breakdown. This approach would allow researchers to explore the specificity and unique economic patterns of different regions within Ukraine, providing more localized insights that could enhance targeted economic policies and responses.

Additionally, investigating the region-specific effectiveness of Google Trends data could uncover how regional variations in internet usage and search behavior impact the accuracy of nowcasting models. This could help in customizing models to better fit the economic realities of different regions, considering local industries, consumer behavior, and economic conditions.

Further research could also explore the integration of regional high-frequency data with traditional economic indicators to assess whether this combined approach provides a more robust and nuanced picture of regional economies. This could be particularly valuable in a country like Ukraine, where economic conditions can vary significantly from one region to another due to factors like industrial diversity, war impacts, and regional economic policies.

Such studies would not only validate the effectiveness of Google Trends data across different contexts but also potentially lead to the development of forecasting models that can dynamically adapt to the economic conditions of specific regions.

Overall, this research contributes to the growing body of knowledge on the integration of digital data into economic forecasting, highlighting its potential to enhance the timeliness and accuracy of economic indicators, especially in challenging environments.

WORKS CITED

Ashwin, Julian, Eleni Kalamara, and Lorena Saiz. 2021. "Nowcasting Euro Area GDP With News Sentiment: A Tale of Two Crises." *Social Science Research Network*. https://doi.org/10.2139/ssrn.3971974.

Banbura, Martha, Domenico Giannone, Michèle Modugno, and Lucrezia Reichlin. 2013. "Now-Casting and the Real-Time Data Flow." In *Handbook of Economic Forecasting*, 195–237. doi:10.1016/b978-0-444-53683-9.00004-9.

Bantis, Evripidis, Michael P. Clements, and Andrew Urquhart. 2021. "Forecasting GDP Growth Rates Using Google Trends in the United States and Brazil." *Social Science Research Network*, January. doi:10.2139/ssrn.3860194.

Basselier, Raïsa, David De Antonio Liedo, and Geert Langenus. 2018. "Nowcasting Real Economic Activity in the Euro Area: Assessing the Impact of Qualitative Surveys." *Journal of Business Cycle Research* 14 (1): 1–46. doi:10.1007/s41549-017-0022-9.

Carrière-Swallow, Yan, and Felipe Labbé. 2013. "Nowcasting with Google Trends in an Emerging Market." *Journal of Forecasting* 32 (4): 289–98. doi:10.1002/for.1252.

Constantinescu, Mihnea. 2023. "Sparse Warcasting." *Social Science Research Network*, January. doi:10.2139/ssrn.4388757.

D'Amuri, Francesco, and Juri Marcucci. 2017. "The Predictive Power of Google Searches in Forecasting US Unemployment." *International Journal of Forecasting* 33 (4): 801–16. doi:10.1016/j.ijforecast.2017.03.004.

Ferrara, Laurent, and Anna Simoni. 2019. "When Are Google Data Useful to Nowcast GDP? An Approach via Pre-Selection and Shrinkage." *Social Science Research Network*, January. doi:10.2139/ssrn.3370917.

Giannone, Domenico & Reichlin Lucrezia & Small, David. 2008. "Nowcasting: The Real-Time Informational Content of Macroeconomic Data." *Ideas.Repec.Org*. https://ideas.repec.org/a/eee/moneco/v55y2008i4p665-676.html.

Götz, Thomas, and Thomas A. Knetsch. 2019. "Google Data in Bridge Equation Models for German GDP." *International Journal of Forecasting* 35 (1): 45–66. doi:10.1016/j.ijforecast.2018.08.001.

Klein, Lawrence R., and E. Sojo. 1989. "Combinations of High and Low Frequency Data in Macroeconometric Models." In *Advanced Studies in Theoretical and Applied Econometrics*, 3–16. doi:10.1007/978-94-009-0463-7_1.

Lysenko, Roman, and Nataliia Kolesnichenko. 2016. "Nowcasting of Economic Development Indicators Using the NBU's Business Survey Results." *Visnyk of the National Bank of Ukraine*, no. 235 (March): 43–56. doi:10.26531/vnbu2016.235.043.

Schumacher, Christian. 2014. "Midas and Bridge Equations." *Social Science Research Network*, January. doi:10.2139/ssrn.2797010.

Shi, Yuying and Stephen L. France. 2018. "Aggregating Google Trends: Multivariate Testing and Analysis." *Ideas.Repec.Org*, March. https://ideas.repec.org/p/arx/papers/1712.03152.html.

Smith, Paul A. 2014. "Nowcasting UK GDP During the Depression: Model Comparisons." *Social Science Research Network*, January. doi:10.2139/ssrn.2507971.

Thorsrud, Leif Anders. 2016. "Nowcasting Using News Topics. Big Data versus Big Bank." *Social Science Research Network*, January. doi:10.2139/ssrn.2901450.

Vosen, Simeon, and Torsten Schmidt. 2011. "Forecasting Private Consumption: Survey-Based Indicators vs. Google Trends." *Journal of Forecasting* 30 (6): 565–78. doi:10.1002/for.1213.

APPENDIX A

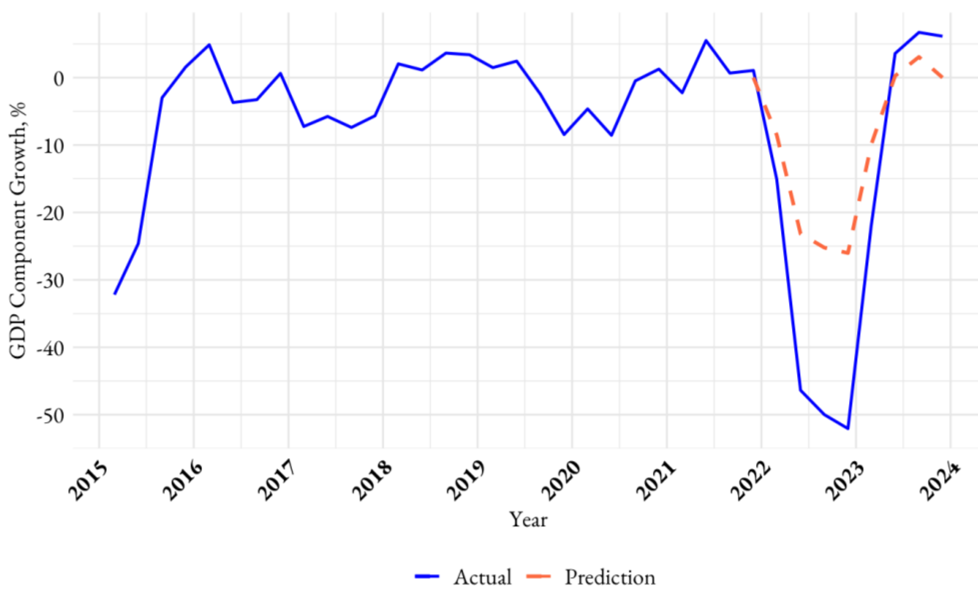

Figure 9. In-sample prediction of the GDP mining and quarrying component

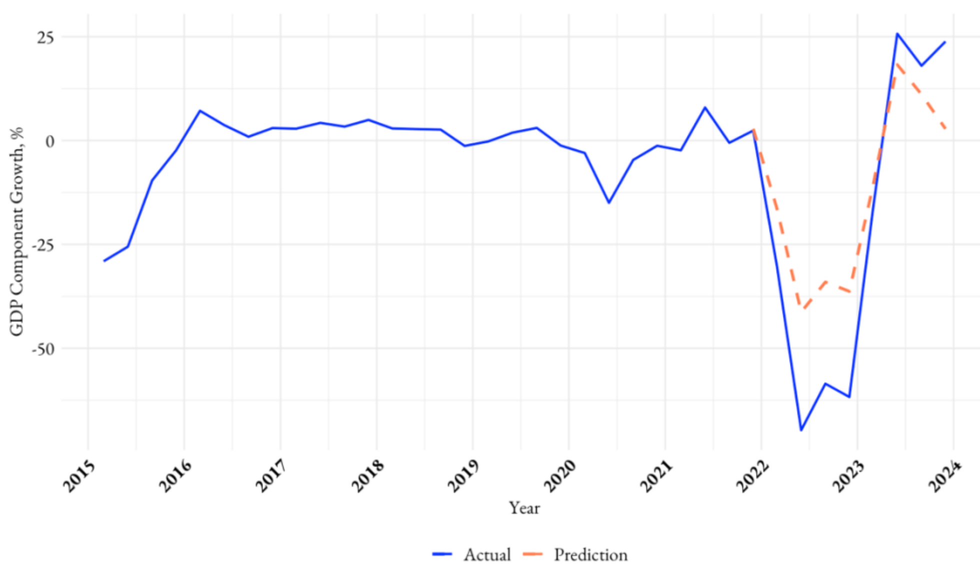

Figure 10. In-sample prediction of the GDP manufacturing component

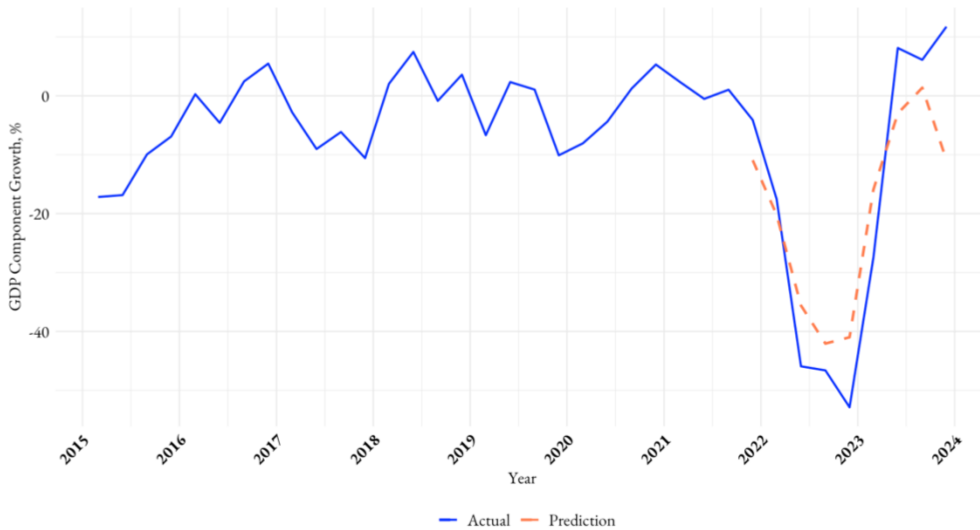

Figure 11. In-sample prediction of the GDP energy supply component

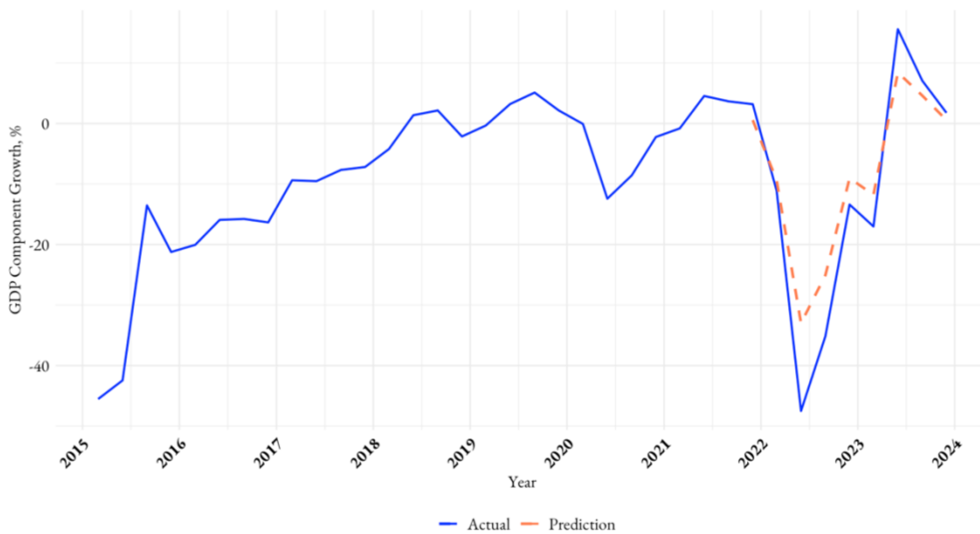

Figure 12. In-sample prediction of the GDP water supply component

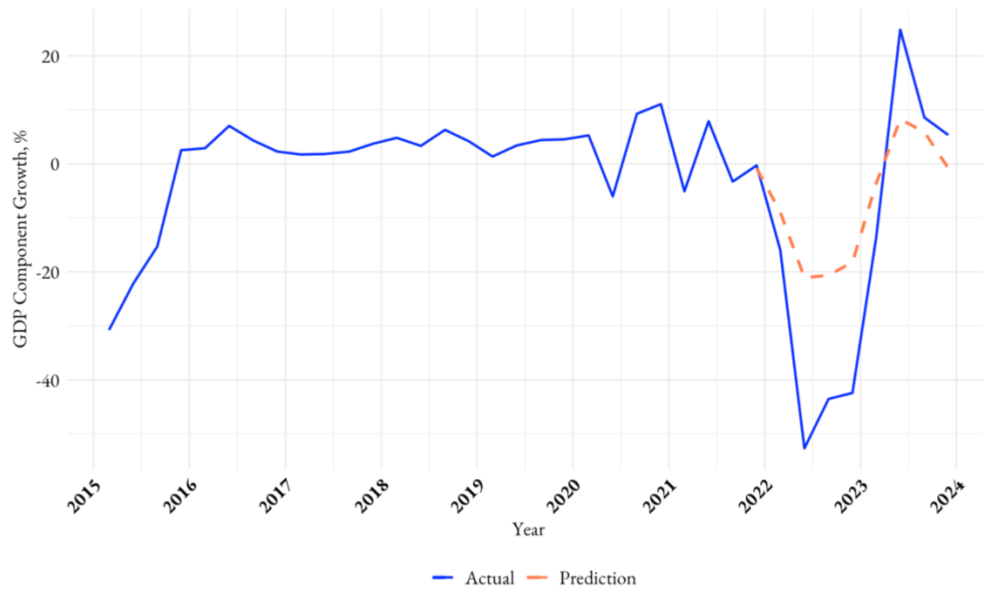

Figure 13. In-sample prediction of the GDP retail sales component

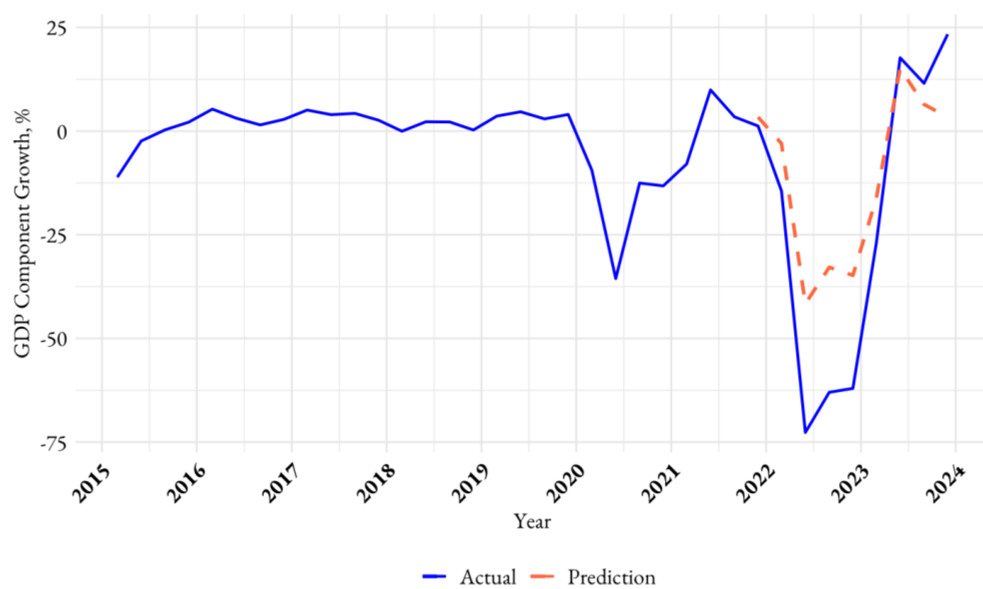

Figure 14. In-sample prediction of the GDP transport component

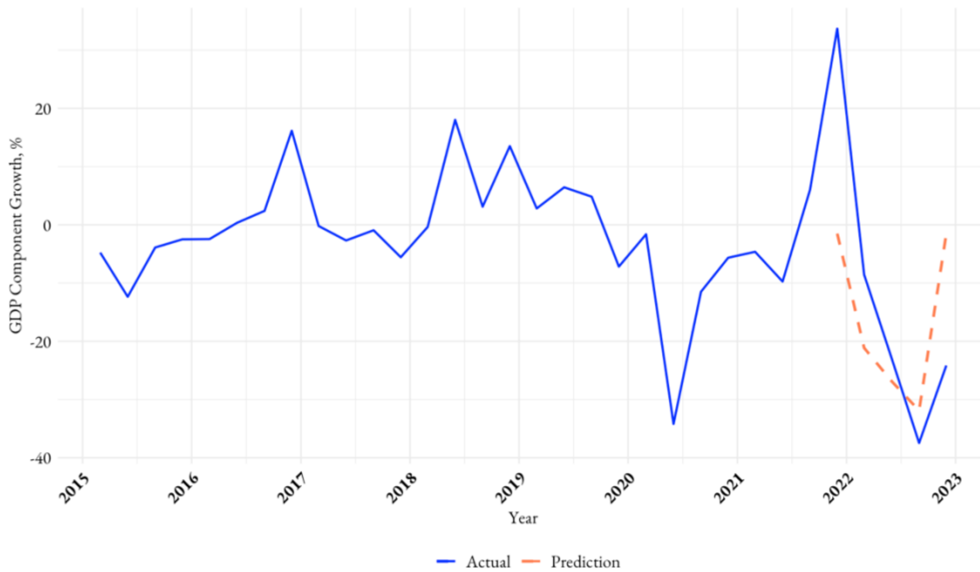

Figure 15. In-sample prediction of the GDP agriculture component

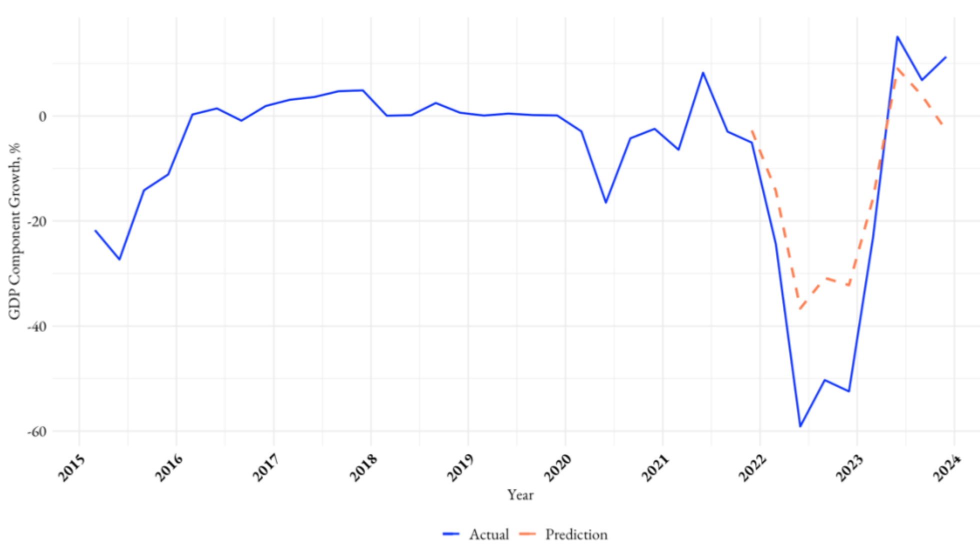

Figure 16. In-sample prediction of the GDP VAT component

APPENDIX B

Table 6. ADL Model Estimation Results for Mining Production

	Dependent variable
	ΔLog(Mining Production)
ΔLog(Mining Production) t-1	0.433***
	(0.088)
ΔLog(Metals & Mining) t-1	0.400
	(0.395)
ΔLog(Industrial Materials) t-1	-0.508**
	(0.214)
ΔLog(Nuclear Energy) t-1	-0.349*
	(0.198)
ΔLog(Manufacturing) t-1	-0.465
	(0.483)
Constant	-0.001
	(0.011)
Observations	107
R2	0.299
Adjusted R2	0.264
Residual Std. Error	0.118
F-Statistic	8.611***
AIC	-145.653
BIC	-126.943
*Note: *p<0.1; **p<0.05; ***p<0.01*	

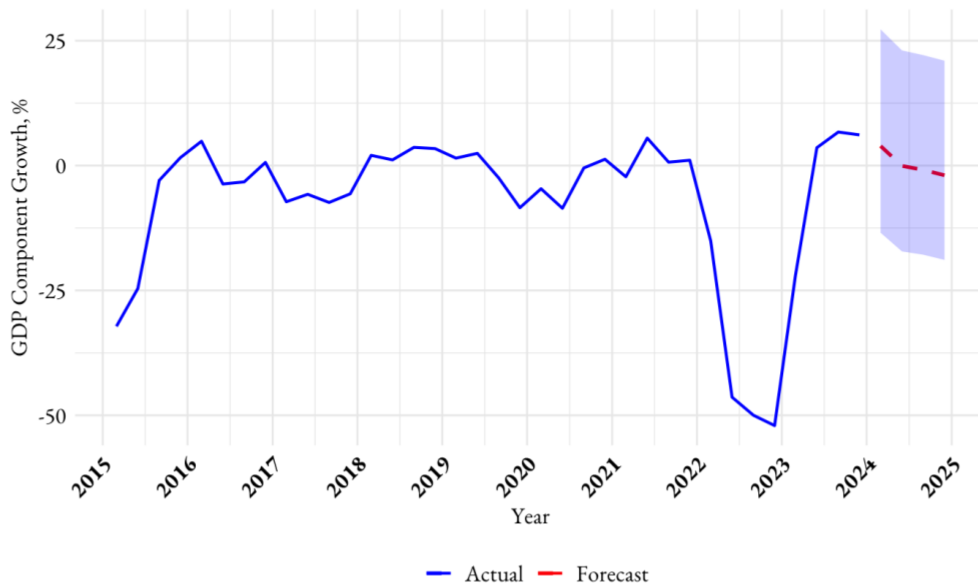

Figure 17. The forecast of the monthly mining production

Table 7. ADL Model Estimation Results for Industrial Production

	Dependent variable
	ΔLog(Industrial Production)
ΔLog(Industrial Production) t-1	0.534***
	(0.084)
ΔLog(Automotive Industry) t-1	-0.426
	(0.310)
ΔLog(Manufacturing) t-1	-0.590**
	(0.247)
Constant	-0.0002
	(0.013)
Observations	107
R2	0.362
Adjusted R2	0.343
Residual Std. Error	0.132
F-Statistic	19.440***
AIC	-123.748
BIC	-110.384
Note: *p<0.1; **p<0.05; ***p<0.01	

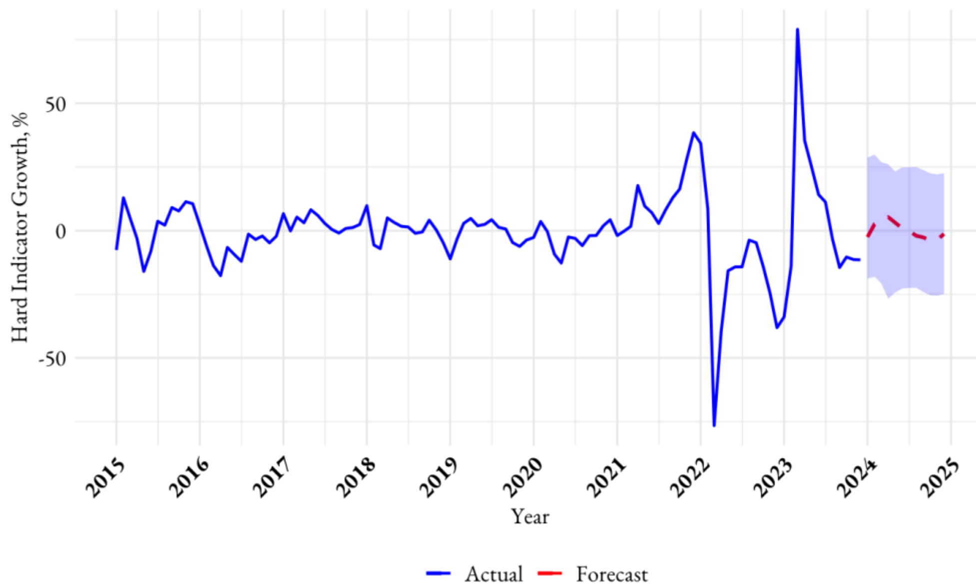

Figure 18. The forecast of the monthly industrial production

Table 8. ADL Model Estimation Results for Energy Supply

	Dependent variable
	ΔLog(Energy Supply)
ΔLog(Energy Supply) t-1	0.664***
	(0.074)
ΔLog(HVAC & Climate Control) t-1	-1.507***
	(0.488)
ΔLog(Consumer Electronics) t-1	-0.772*
	(0.435)
ΔLog(Energy & Utilities) t-1	-0.500
	(0.486)
Constant	-0.001
	(0.011)
Observations	107
R2	0.491
Adjusted R2	0.471
Residual Std. Error	0.112
F-Statistic	24.635***
AIC	-158.482
BIC	-142.445
*Note: *p<0.1; **p<0.05; ***p<0.01*	

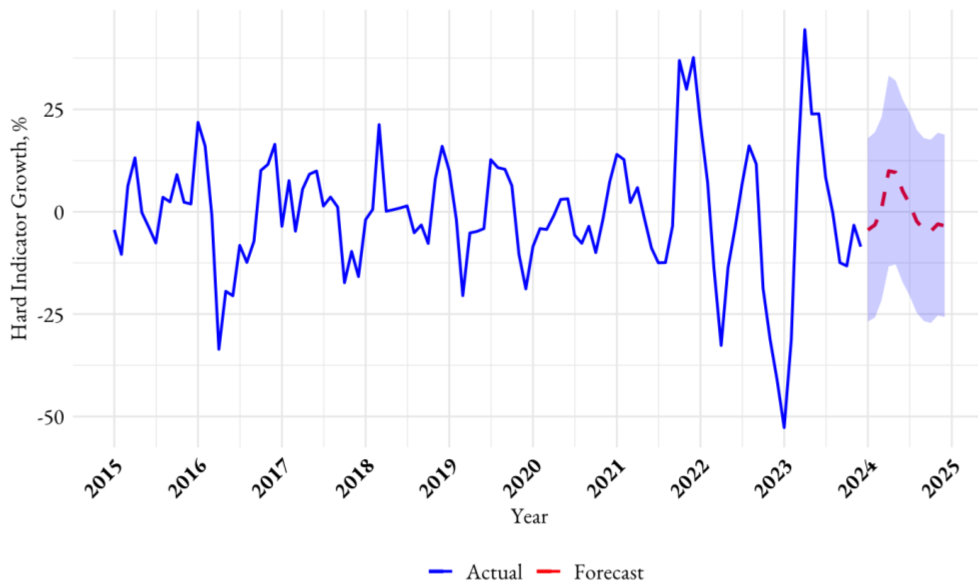

Figure 19. The forecast of the monthly energy supply

Table 9. ADL Model Estimation Results for Water Supply

	Dependent variable
	ΔLog(Water Supply)
ΔLog(Water Supply) t-1	0.398***
	(0.087)
ΔLog(Energy & Utilities) t-1	-0.767**
	(0.343)
ΔLog(Nuclear Energy) t-1	-0.289**
	(0.133)
ΔLog(Crops & Seeds) t-1	-0.342
	(0.226)
Constant	-0.001
	(0.008)
Observations	107
R2	0.317
Adjusted R2	0.290
Residual Std. Error	0.078
F-Statistic	11.824***
AIC	-236.339
BIC	-220.302

*Note: *p<0.1; **p<0.05; ***p<0.01*

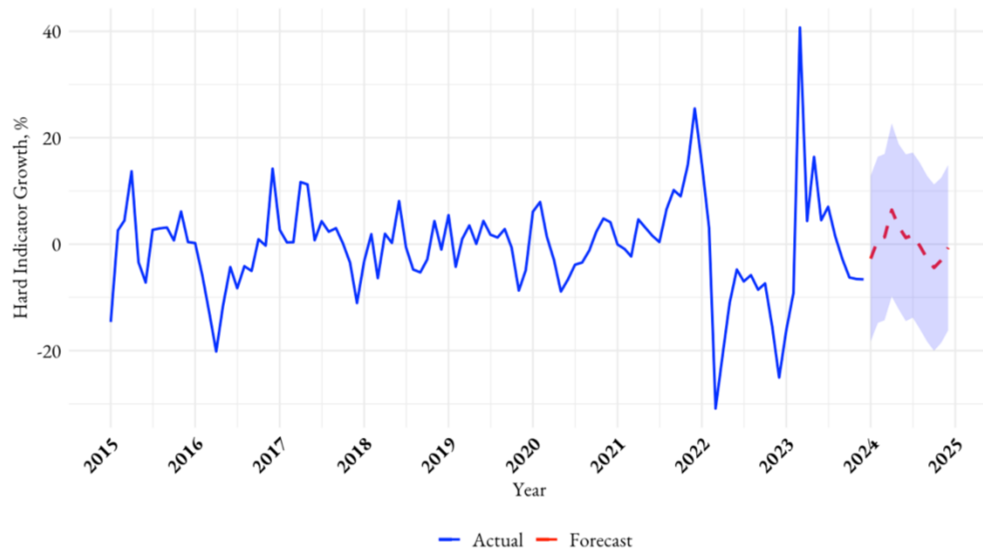

Figure 20. The forecast of the monthly water supply

Table 10. ADL Model Estimation Results for Transport

	Dependent variable
	ΔLog(Transport)
ΔLog(Transport) t-1	0.542***
	(0.075)
ΔLog(Bus & Rail) t-1	-0.219
	(0.139)
ΔLog(Trucks & SUVs) t-1	-0.836***
	(0.287)
ΔLog(Transportation) t-1	-0.656*
	(0.388)
Constant	-0.001
	(0.009)
Observations	107
R2	0.445
Adjusted R2	0.423
Residual Std. Error	0.089
F-Statistic	20.464***
AIC	-206.914
BIC	-190.877

*Note: *p<0.1; **p<0.05; ***p<0.01*

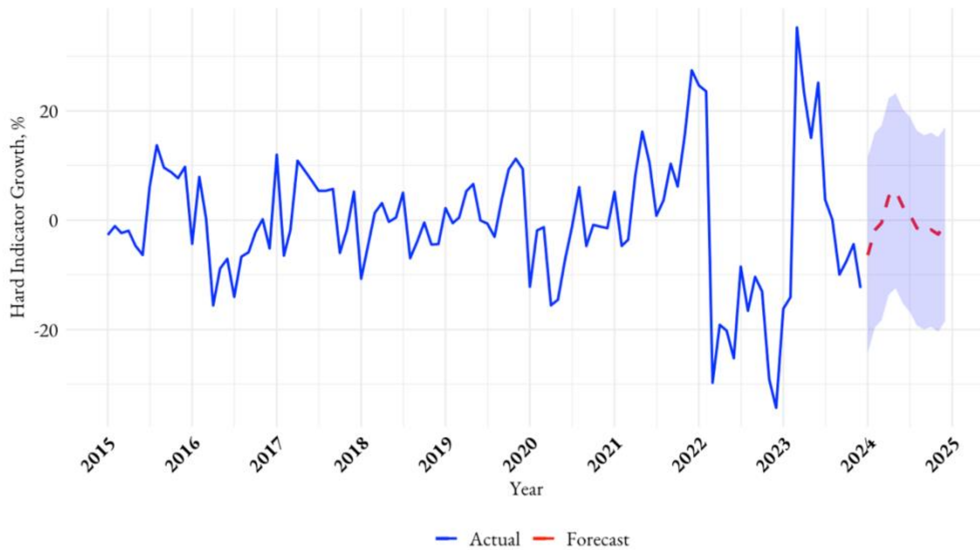

Figure 21. The forecast of the monthly transport

Table 11. ADL Model Estimation Results for Construction Production

	Dependent variable
	ΔLog(Construction Production)
ΔLog(Construction Production) t-1	0.678***
	(0.082)
ΔLog(Construction & Maintanance) t-1	-0.225
	(0.618)
ΔLog(Gardening) t-1	-0.276
	(0.548)
ΔLog(Property Development) t-1	-0.582
	(0.718)
ΔLog(Apartments) t-1	-0.631
	(0.562)
Constant	-0.001
	(0.017)
Observations	107
R2	0.408
Adjusted R2	0.379
Residual Std. Error	0.176
F-Statistic	13.939***
AIC	-59.694
BIC	-40.984

*Note: *p<0.1; **p<0.05; ***p<0.01*

Table 12. ADL Model Estimation Results for Retail Turnover

	Dependent variable
	ΔLog(Retail Turnover)
ΔLog(Retail Turnover) t-1	0.299***
	(0.095)
ΔLog(Food & Grocery) t-1	-0.523*
	(0.271)
ΔLog(Retail Trade) t-1	0.256
	(0.221)
ΔLog(Shopping Portals) t-1	0.584*
	(0.311)
Constant	-0.0005
	(0.012)
Observations	107
R2	0.150
Adjusted R2	0.116
Residual Std. Error	0.123
F-Statistic	4.490***
AIC	-137.826
BIC	-121.789

*Note: *p<0.1; **p<0.05; ***p<0.01*

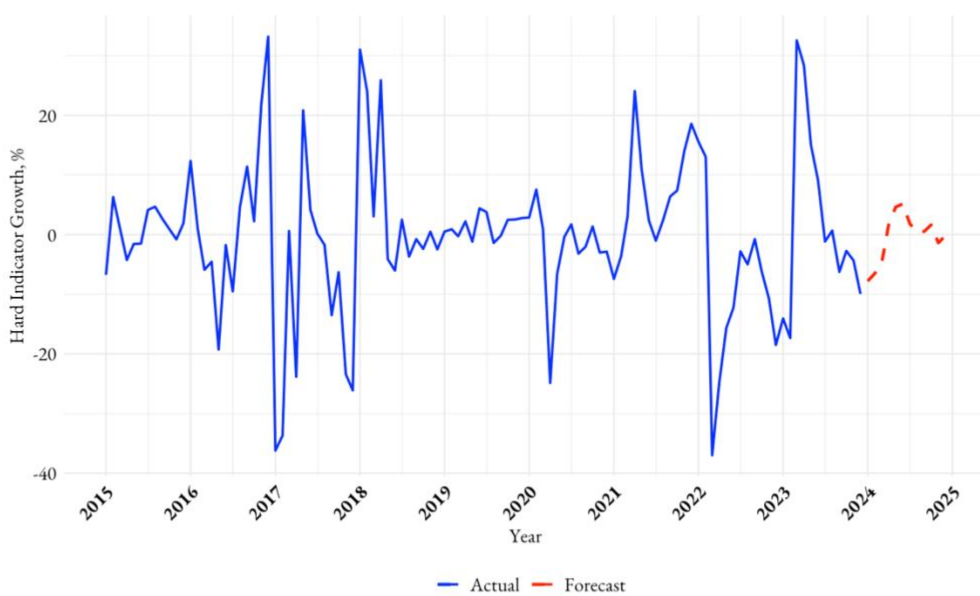

Figure 22. The forecast of the monthly retail turnover

Table 13. ADL Model Estimation Results for Price Index of Agricultural Products

	Dependent variable
	ΔLog(Price Index of Agricultural Products)
ΔLog(Price Index of Agricultural Products) t-1	0.594***
	(0.085)
ΔLog(Chemicals Industry) t-1	-0.274
	(0.487)
ΔLog(Crops & Seeds) t-1	-0.759*
	(0.414)
ΔLog(Agricultural Equipment) t-1	0.639**
	(0.315)
Constant	-0.003
	(0.010)
Observations	95
R2	0.380
Adjusted R2	0.353
Residual Std. Error	0.096
F-Statistic	13.797***
AIC	-167.878
BIC	-152.554

*Note: *p<0.1; **p<0.05; ***p<0.01*

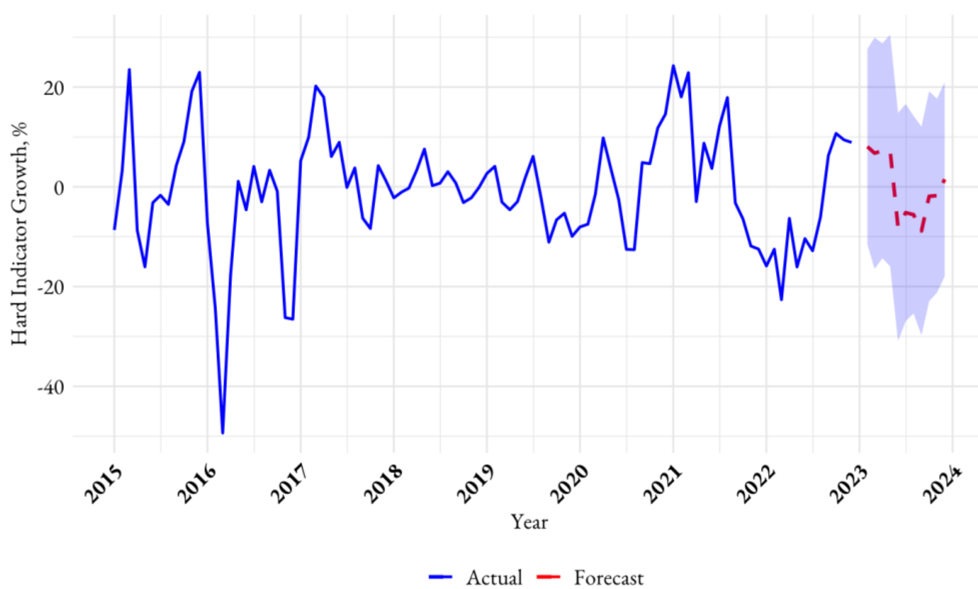

Figure 23. The forecast of the monthly price index of agricultural products

Table 14. ADL Model Estimation Results for Price Index of Taxes on Products

	Dependent variable
	ΔLog(Taxes on Products)
ΔLog(Taxes on Products) t-1	0.619***
	(0.075)
ΔLog(Business Finance) t-1	-0.338*
	(0.200)
ΔLog(Mass Merchants) t-1	-0.229
	(0.159)
ΔLog(Shopping Portals) t-1	0.134
	(0.139)
Constant	0.001
	(0.005)
Observations	107
R2	0.434
Adjusted R2	0.412
Residual Std. Error	0.054
F-Statistic	19.571***

*Note: *p<0.1; **p<0.05; ***p<0.01*

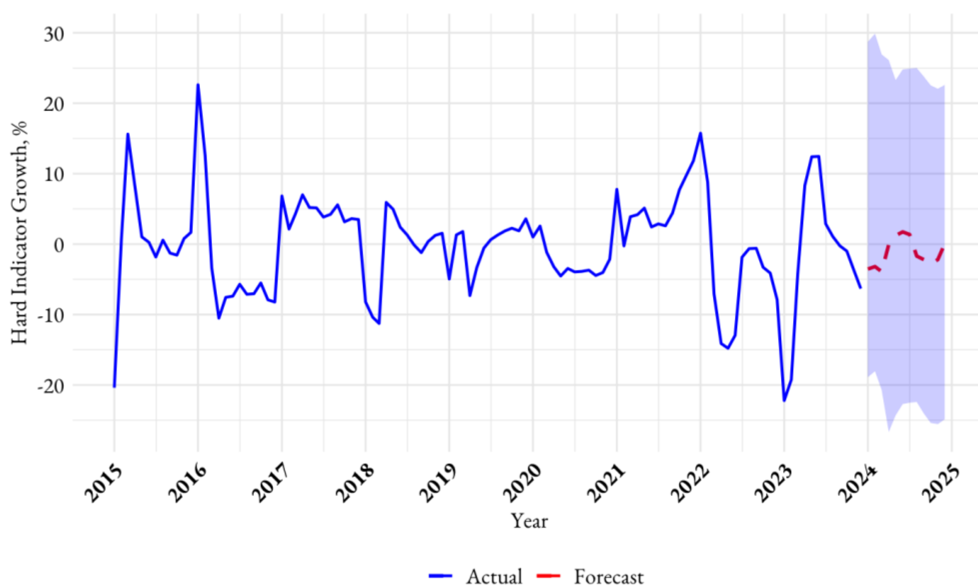

Figure 24. The forecast of the monthly price index of taxes on products

Table 15. ADL Model Estimation Results for Price Index of Government Spending

	Dependent variable
	ΔLog(Government Spending)
ΔLog(Government Spending) t-1	0.263***
	(0.090)
ΔLog(Local Government) t-1	-0.484***
	(0.175)
ΔLog(Army) t-1	-0.510**
	(0.220)
ΔLog(Politics) t-1	0.310**
	(0.138)
ΔLog(Politics) t-1	0.390*
	(0.209)
Constant	0.001
	(0.008)
Observations	107
R2	0.213
Adjusted R2	0.174
Residual Std. Error	0.085
F-Statistic	5.472***
AIC	-216.988
BIC	-198.278

Note: *p<0.1; **p<0.05; ***p<0.01

APPENDIX C

Table 16. Forecasts of GDP components

	Period						
GDP Component	**t-2**	**t-1**	**t**	**t+1**	**t+2**	**t+3**	**t+4**
Mining and quarrying	0.036	0.067	0.061	0.069	0.029	0.021	0.011
Manufacturing	0.257	0.180	0.238	0.199	0.183	0.149	0.163
Electricity Supply	0.081	0.061	0.117	0.161	0.159	0.084	0.086
Water Supply	0.156	0.071	0.018	0.060	0.033	-0.025	-0.028
Construction	0.481	0.226	0.227	0.019	0.235	0.152	0.223
Retail Sales	0.248	0.086	0.053	0.147	0.240	0.139	0.104
Transport	0.177	0.115	0.234	0.281	0.313	0.259	0.266
Agriculture	-0.229	-0.375	-0.241	-0.177	-0.124	-0.104	-0.097
Public Administration	0.075	0.019	0.014	0.026	0.023	0.029	0.04
VAT	0.151	0.069	0.113	0.103	0.137	0.102	0.096

The t index is the end of the available data, 2023Q4. The t+1, .., t+4 refer to quarters of 2024.

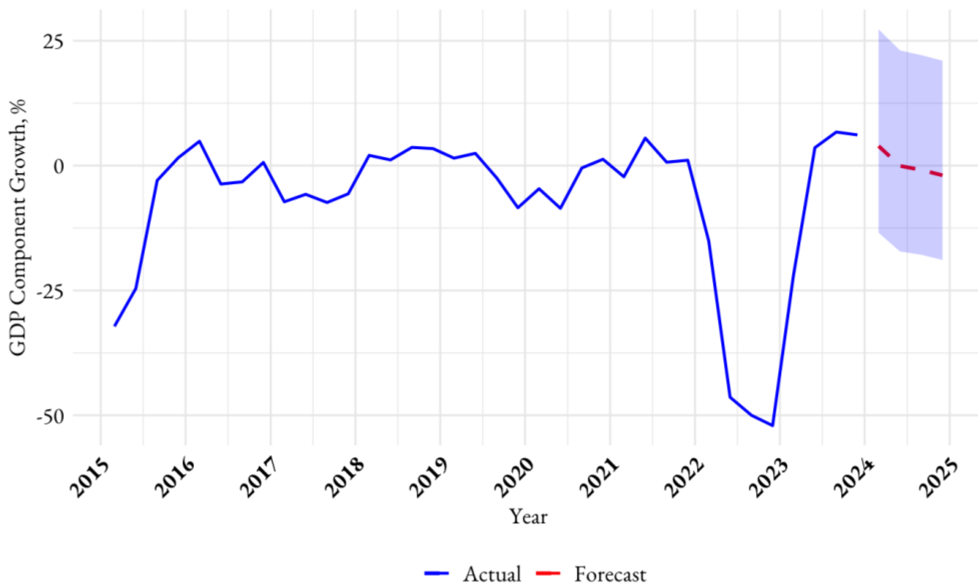

Figure 25. The forecast of the mining GDP component

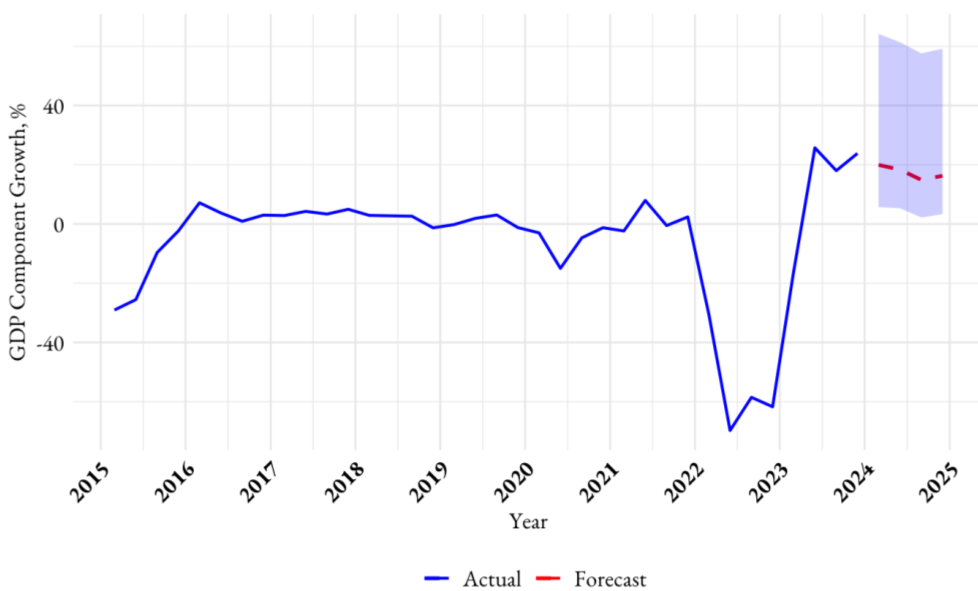

Figure 26. The forecast of the manufacturing GDP component

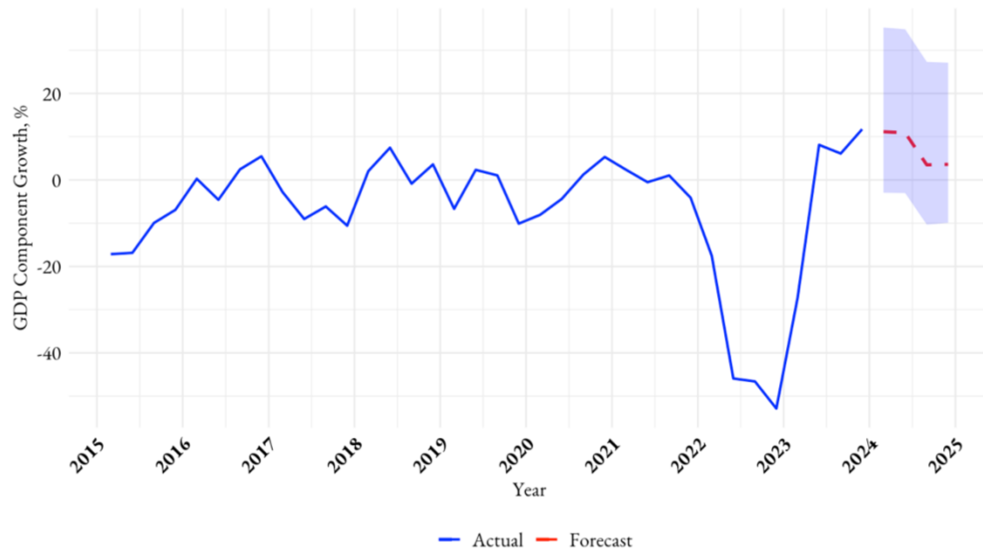

Figure 27. The forecast of the electricity supply GDP component

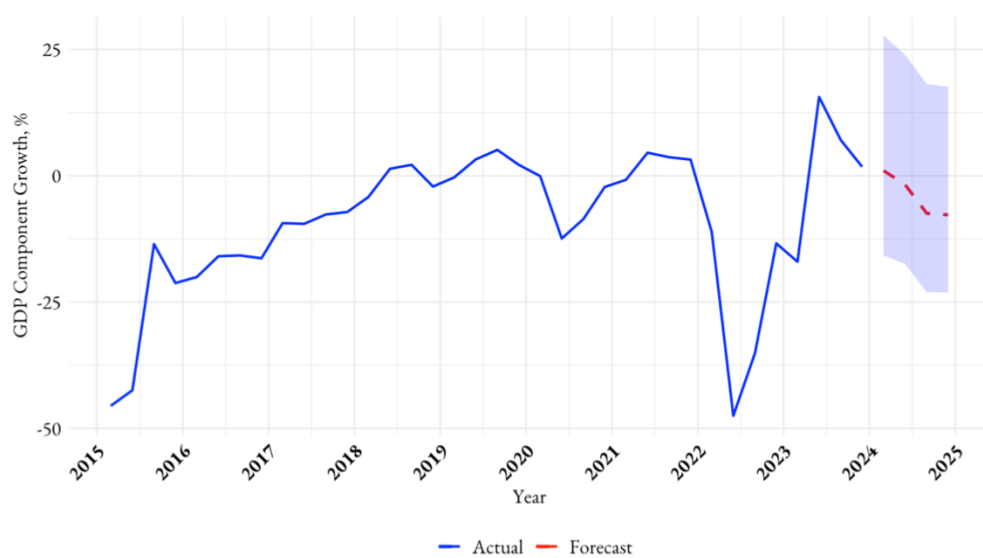

Figure 28. The forecast of the water supply GDP component

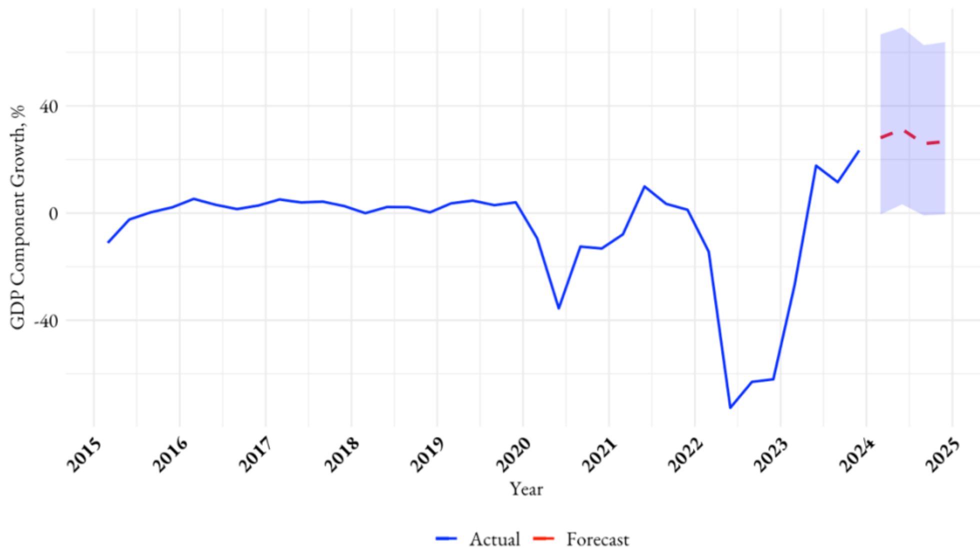

Figure 29. The forecast of the transport GDP component

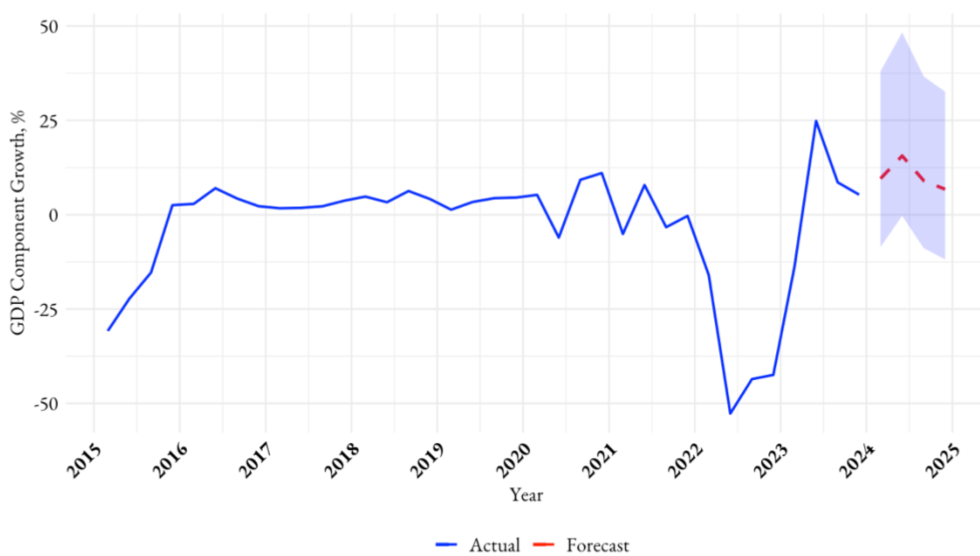

Figure 30. The forecast of the retail sales GDP component

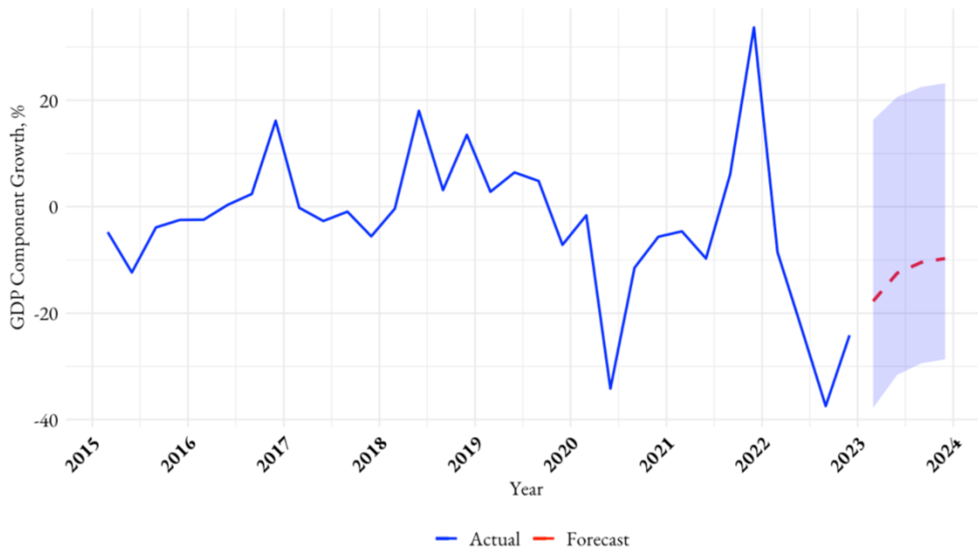

Figure 31. The forecast of the agriculture GDP component

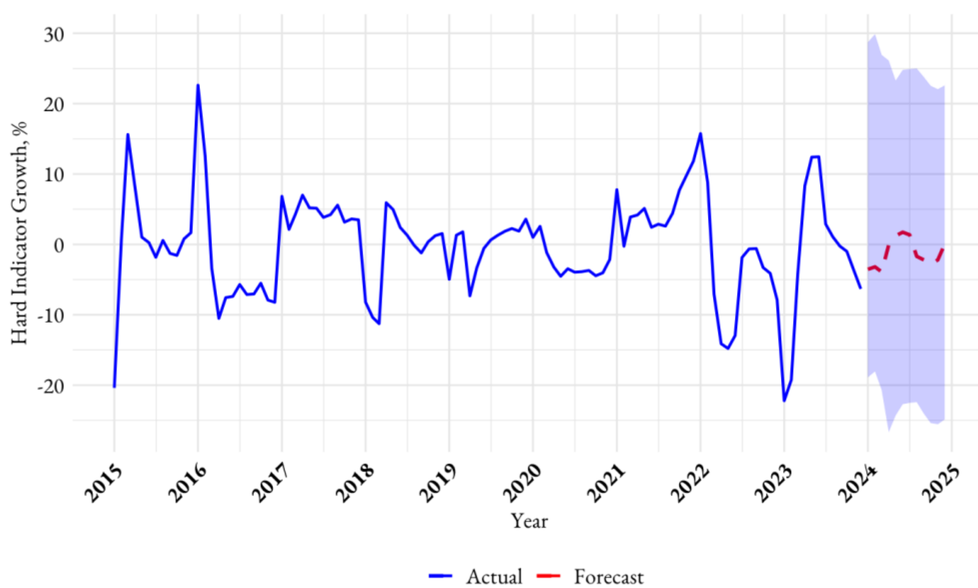

Figure 32. The forecast of the VAT GDP component

APPENDIX D

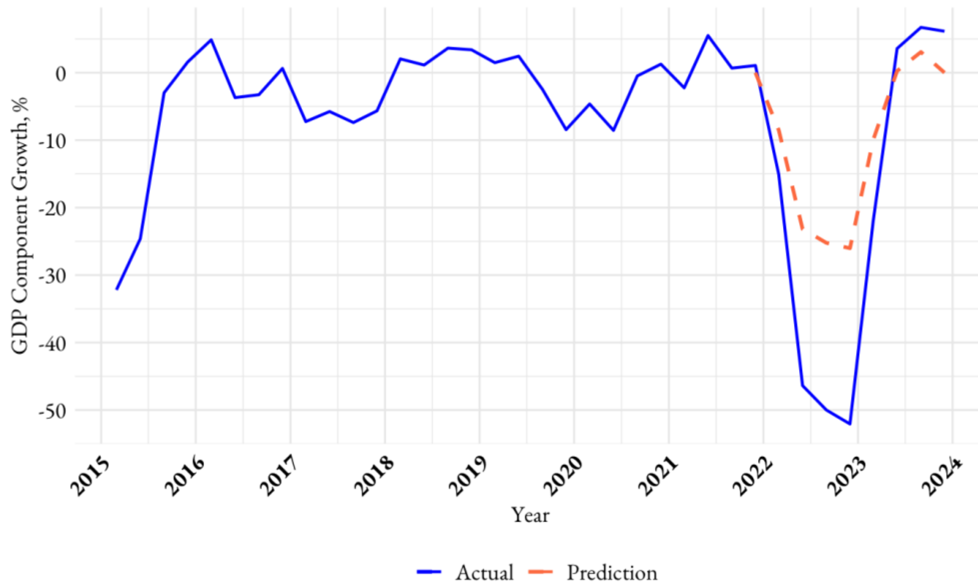

Figure 33. The prediction of the mining GDP component

Table 17. Estimation of Mining GDP component

	Dependent variable
	ΔLog(Mining)
ΔLog(Mining Production) x-1	-0.560*
	(0.301)
ΔLog(Mining) y-1	0.845***
	(0.120)
Constant	-0.003
	(0.018)
Observations	35
R2	0.621
Adjusted R2	0.597
Residual Std. Error	0.096
F-Statistic	26.233***

Note: *$p<0.1$; **$p<0.05$; ***$p<0.01$

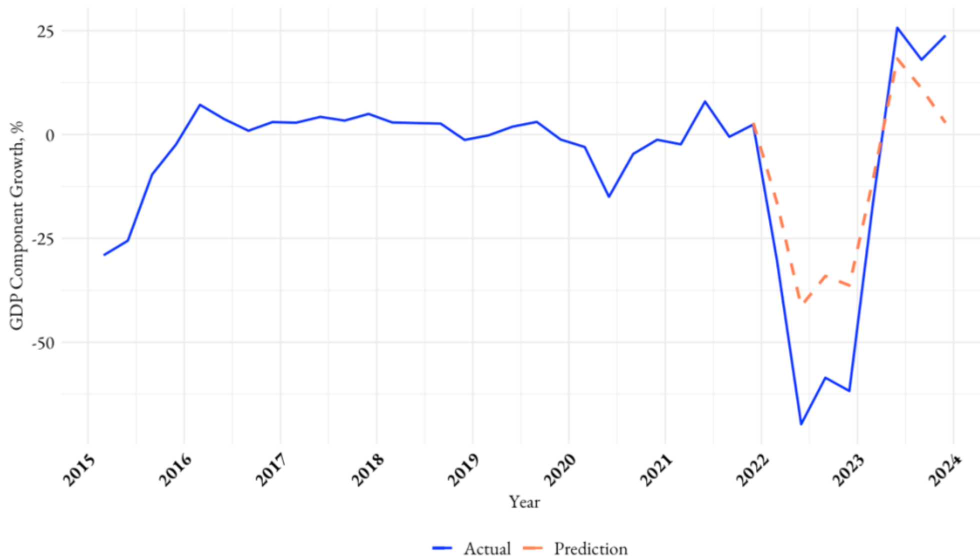

Figure 34. The prediction of the manufacturing GDP component

Table 18. Estimation of Manufacturing GDP Component

	Dependent variable
	ΔLog(Manufacturing)
ΔLog(Industrial Production) x-1	-0.996***
	(0.323)
ΔLog(Manufacturing) y-1	0.955***
	(0.122)
Constant	0.012
	(0.023)
Observations	35
R2	0.661
Adjusted R2	0.639
Residual Std. Error	0.127
F-Statistic	31.129***
*Note: *p<0.1; **p<0.05; ***p<0.01*	

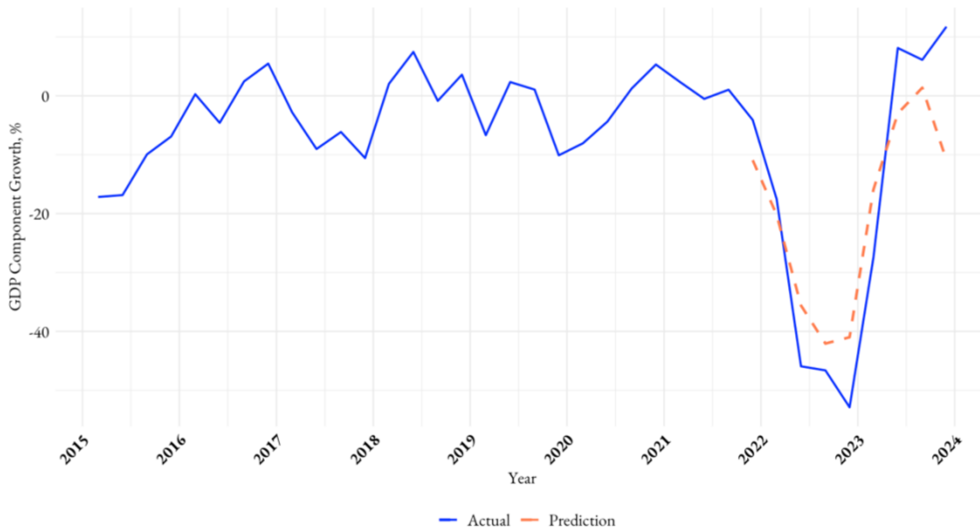

Figure 35. The prediction of the energy supply GDP component

Table 19. Estimation of Energy Supply GDP Component

	Dependent variable
	ΔLog(Energy Supply)
ΔLog(Electricity Supply) x-1	-0.689***
	(0.186)
ΔLog(Energy Supply) y-1	0.893***
	(0.107)
Constant	-0.001
	(0.017)
Observations	35
R2	0.689
Adjusted R2	0.670
Residual Std. Error	0.088
F-Statistic	35.469***

*Note: *p<0.1; **p<0.05; ***p<0.01*

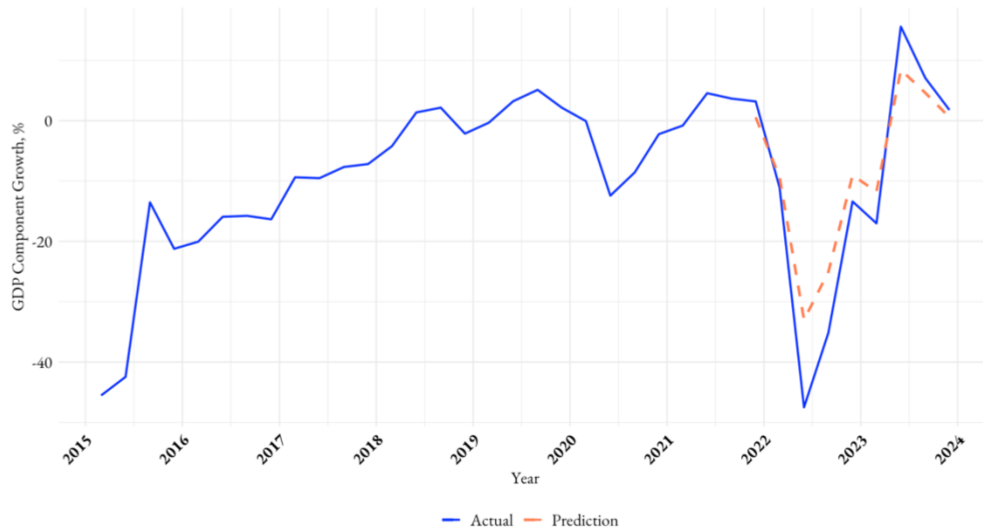

Figure 36. The prediction of the water supply GDP component

Table 20. Estimation of Water Supply GDP Component

	Dependent variable
	ΔLog(Water Supply)
ΔLog(Water Supply) x-1	-0.942*
	(0.477)
ΔLog(Water Supply) y-1	0.737***
	(0.129)
Constant	-0.014
	(0.020)
Observations	35
R2	0.512
Adjusted R2	0.482
Residual Std. Error	0.098
F-Statistic	16.813***
*Note: *p<0.1; **p<0.05; ***p<0.01*	

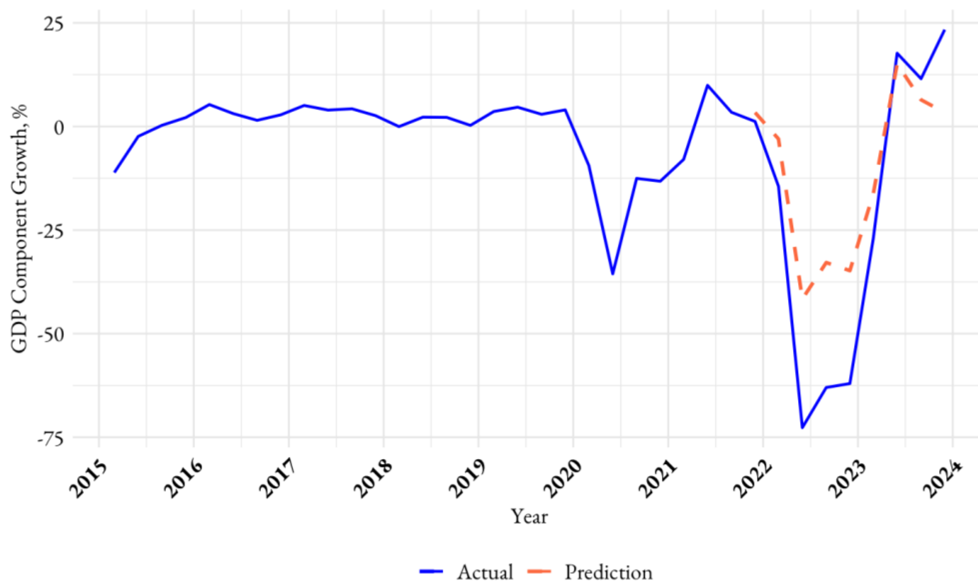

Figure 37. The prediction of the transport GDP component

Table 21. Estimation of Water Supply GDP Component

	Dependent variable
	ΔLog(Transport)
ΔLog(Transport) x-1	-1.210**
	(0.485)
ΔLog(Transport) y-1	0.972***
	(0.153)
Constant	0.006
	(0.026)
Observations	35
R2	0.583
Adjusted R2	0.557
Residual Std. Error	0.144
F-Statistic	22.380***
*Note: *p<0.1; **p<0.05; ***p<0.01*	

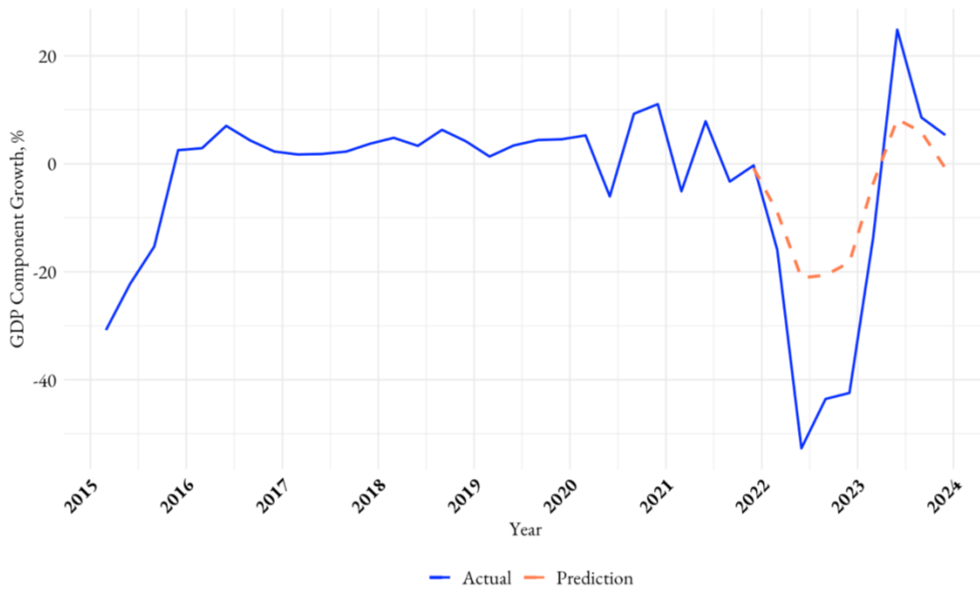

Figure 38. The prediction of the sales GDP component

Table 22. Estimation of Sales GDP Component

	Dependent variable
	ΔLog(Sales)
ΔLog(Retail Sales) x-1	-1.762***
	(0.587)
ΔLog(Sales) y-1	0.852***
	(0.125)
Constant	0.002
	(0.018)
Observations	35
R2	0.592
Adjusted R2	0.566
Residual Std. Error	0.106
F-Statistic	23.189***

*Note: *p<0.1; **p<0.05; ***p<0.01*

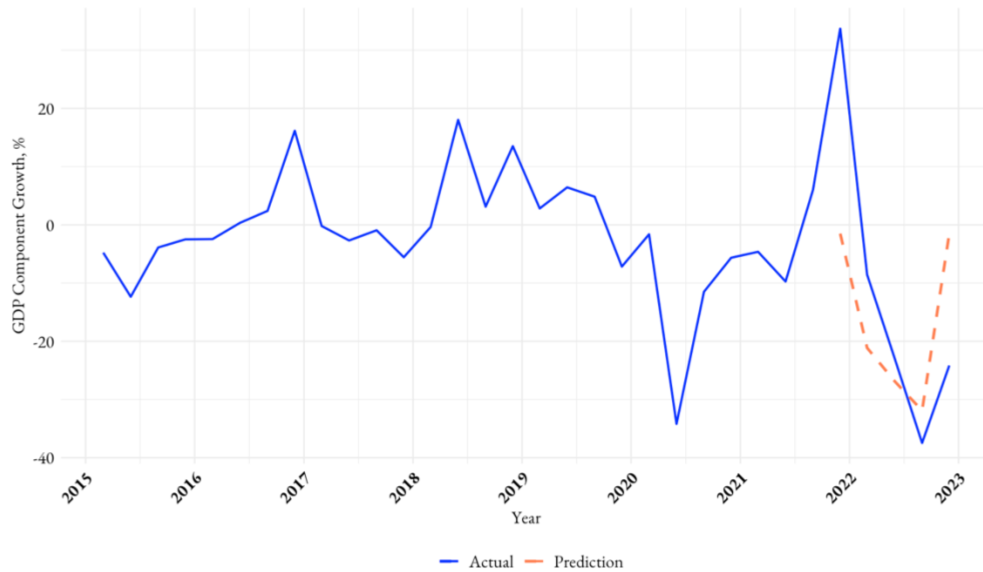

Figure 39. The prediction of the agriculture GDP component

Table 23. Estimation of Agriculture GDP Component

	Dependent variable
	ΔLog(Agriculture)
ΔLog(Price Index of Agriculture Products) x-1	0.833*
	(0.425)
ΔLog(Agriculture) y-1	0.393**
	(0.169)
Constant	-0.013
	(0.023)
Observations	31
R2	0.284
Adjusted R2	0.233
Residual Std. Error	0.125
F-Statistic	5.552***
*Note: *p<0.1; **p<0.05; ***p<0.01*	

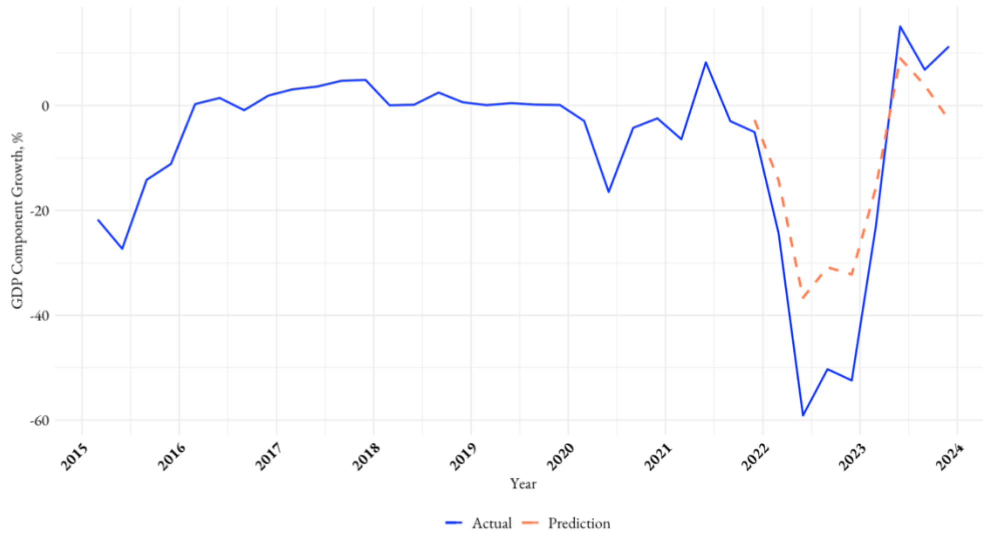

Figure 40. The prediction of the VAT GDP component

Table 24. Estimation of VAT GDP Component

	Dependent variable
	ΔLog(VAT)
ΔLog(Taxes on Products) x-1	-1.457***
	(0.486)
ΔLog(VAT) y-1	0.885***
	(0.113)
Constant	-0.0001
	(0.020)
Observations	35
R2	0.657
Adjusted R2	0.635
Residual Std. Error	0.105
F-Statistic	30.607***

*Note: *p<0.1; **p<0.05; ***p<0.01*

Table 25. Estimation of Construction GDP Component

	Dependent variable
	ΔLog(Construction)
ΔLog(Construction Production) x-1	-1.492***
	(0.405)
ΔLog(Construction) y-1	0.994***
	(0.118)
Constant	0.021
	(0.041)
Observations	35
R2	0.694
Adjusted R2	0.675
Residual Std. Error	0.240
F-Statistic	36.231***

*Note: *p<0.1; **p<0.05; ***p<0.01*

Table 26. Estimation of Public Administration GDP Component

	Dependent variable
	ΔLog(Public Administration)
ΔLog(Government Spending) x-1	-0.704
	(0.450)
ΔLog(Public Administration) y-1	0.777***
	(0.107)
Constant	0.011
	(0.014)
Observations	35
R2	0.636
Adjusted R2	0.613
Residual Std. Error	0.076
F-Statistic	27.941***

*Note: *p<0.1; **p<0.05; ***p<0.01*