NATIONAL BANK OF KAZAKHSTAN

FORECASTING AND POLICY ANALYSIS SYSTEM IN THE NATIONAL BANK OF KAZAKHSTAN

Contents

1. Forecasting and Policy Analysis System at the National Bank of the Republic of Kazakhstan....3
1.1 About the Forecasting System at the National Bank of the Republic of Kazakhstan 3
1.2 Conducting Forecast Rounds at the National Bank....3
2. Short-Term Forecasting....5
2.1 Inflation Forecasting System. System of Selective Combined Inflation Forecast (SSCIF)....5
2.2 The KOMPAS Model for Short-Term Forecasting of GDP by the Expenditure Approach....10
2.3 Short-Term Forecasting of GDP by the Production Approach. System for Comprehensive Analysis and Forecasting of the Economy of Kazakhstan. 13
3. Medium-Term Forecasting....17
3.1 Phillips Curve Block....17
3.2 IS Curve Block....18
3.3 Monetary Policy Block....19
3.4 Exogenous Blocks....20
4. External Sector Models....20
4.1 Forecasting the Current Account of the Balance of Payments....20
5. Improvement and Development of the Forecasting and Analytical System 22
5.1 Research Activities at the National Bank....22
5.2 Models and Tools under Development....23
Recommended Reading....24

1. FORECASTING AND POLICY ANALYSIS SYSTEM AT THE NATIONAL BANK OF THE REPUBLIC OF KAZAKHSTAN

1.1 About the Forecasting System at the National Bank of the Republic of Kazakhstan

In 2014, as part of the transition to an inflation targeting regime, the National Bank of the Republic of Kazakhstan (NBK) began developing and implementing a **Forecasting and Policy Analysis System** (FPAS).

The forecasting and analysis system provides information and analytical support for monetary policy decision-making. The system was built in line with international best practice, with technical assistance from the International Monetary Fund (IMF) and the Czech National Bank, which is one of Europe's leading central banks in the field of modelling.

Within the FPAS, short-term and medium-term forecasts of key macroeconomic indicators are produced for the coming three years, and monetary policy decisions, including on the level of the base rate, are taken on their basis. Forecasts are produced four times a year, on a quarterly basis, within Forecast Rounds.

A forecast round comprises several stages, during which the forecasting techniques are checked, the current macroeconomic situation is comprehensively analysed, and the forecast assumptions are formulated, the short-term and medium-term forecasts are produced, and the monetary policy path over the forecast horizon is assessed.

At the end of the forecast round, the forecast prepared serves as the basis for the Monetary Policy Committee's discussion and decision on the base rate. At the same time, during the discussion the forecast may be refined to take into account new information received after the completion of the forecast round, significant changes in external and domestic macroeconomic conditions, the materialisation of previously identified risks or the emergence of new ones, changes in the balance of risks to inflation, and other factors capable of having a significant impact on the projected path of inflation and the base rate. Subsequently, the forecasts and possible risks are communicated, with the aim of conveying and explaining the decision taken to market participants and the public in a timely manner.

1.2 Conducting Forecast Rounds at the National Bank

The forecast round at the NBK is governed by the internal document “Procedure for Conducting Forecast Rounds at the National Bank of the Republic of Kazakhstan”, and its timing is set in accordance with the NBK Forecast Round Calendar.

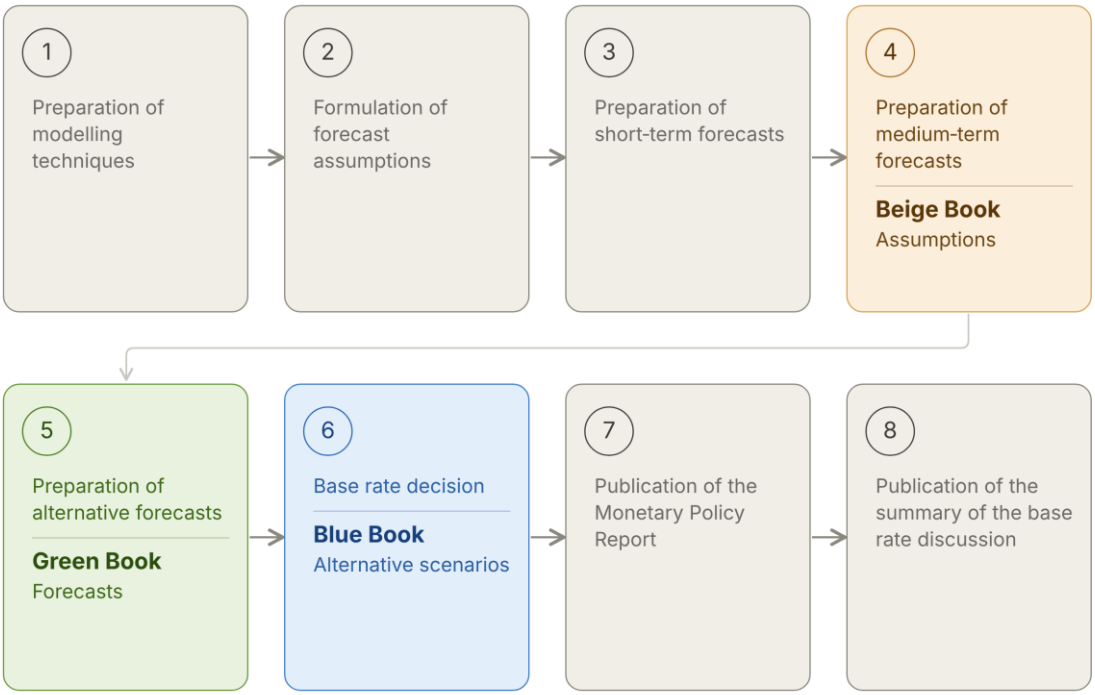
The initial stage includes an analysis of the current macroeconomic situation in Kazakhstan, as well as of external conditions in trading partner countries, on global commodity markets and in the global economy. Based on the information obtained, the

domestic and external forecast assumptions are formulated, including scenario oil prices, fiscal and quasi-fiscal policy parameters, GDP and inflation dynamics in partner countries, and prices on global food markets.

Current conditions are analysed in the following areas:

1. Real sector of the economy – sectoral analysis, producer price dynamics, profitability and other indicators;
2. Inflation processes and inflation expectations;
3. Domestic demand – household consumption, investment, changes in inventories, household income and expenditure;
4. Financial sector – money and foreign exchange markets, lending, deposits, the government securities market;
5. Global commodity markets – oil, metals, food;
6. Current account of the balance of payments;
7. Macroeconomic situation in partner countries and the global economy – GDP, inflation and exchange rate dynamics (including China, Russia and the EU), as well as the policies of the leading central banks;
8. Fiscal policy – budget policy parameters and the quasi-fiscal sector.

Figure 1. Forecast round calendar by week



Based on this analysis, the forecast assumptions are formulated, after which the short-term forecasting stage is carried out over a horizon of up to three quarters (the current quarter and the two following ones). The short-term baseline forecast is produced using econometric methods and includes estimates of GDP (by the expenditure and production approaches), inflation and its components.

The results of short-term forecasting are used at the next stage – medium-term forecasting. The medium-term baseline forecast is produced for the subsequent 4–8

quarters, which together gives an overall forecast horizon of up to three years. The medium-term forecast is based on the Quarterly Projection Model (QPM).

In addition to the baseline forecast, alternative scenarios are developed at the request of the Monetary Policy Committee, covering different paths of macroeconomic indicators and considering possible risks and ways in which the economy may develop that could cause inflation to deviate from the target.

The final stage of the Forecast Round is communication work to convey and explain the Committee's decision on the base rate. It includes holding a press briefing and publishing a press release on the base rate, holding a meeting with experts, publishing the Monetary Policy Report and preparing other accompanying materials.

The forecast round lasts eight weeks and is a comprehensive, multi-stage process based on close coordination and interlinked work of various departments (Figure 1). Each stage is carried out within strictly set deadlines in accordance with the approved Forecast Round Calendar. The process is accompanied by continuous monitoring and coordination, which ensures the transparency, consistency and high quality of the preparation of monetary policy decisions.

2. SHORT-TERM FORECASTING

2.1 Inflation Forecasting System. System of Selective Combined Inflation Forecast (SSCIF)

2.1.1. System description

The System of Selective Combined Inflation Forecast (SSCIF) is a specialised modelling platform for short-term inflation forecasting used by the NBK within its monetary policy analysis and forecasting system. The system has been in use at the NBK since 2016 and serves as the main tool for producing short-term estimates of inflation dynamics.

SSCIF was developed and implemented in EViews as the result of the gradual evolution of the econometric and statistical inflation forecasting models used during Kazakhstan's transition to an inflation targeting regime. The methodological architecture of the system is based on a combination of different classes of models and algorithms for combining them, which makes it possible to consider both the inertial properties of inflation processes and the impact of macroeconomic factors and structural changes in the economy.

Within SSCIF, forecasting is carried out not at the level of aggregate inflation but at the level of its disaggregated components, which makes it possible to capture in greater detail the specifics of price formation in different groups and classes of the consumer market. The system uses price dynamics for the main groups and classes of goods and services that make up the consumer price index.

Thus, inflation is forecast on a bottom-up basis: forecasts are first produced for individual price sub-indices, then aggregated into forecasts of food, non-food and services inflation, after which the final forecast of the overall consumer price index is

calculated using the official weights of the consumer basket.

This approach significantly improves forecast accuracy, since the dynamics of individual inflation components can differ substantially in their sensitivity to macroeconomic and external factors.

The SSCIF architecture is based on several classes of models that differ in the type of relationships used and in estimation methods. This multi-model approach makes it possible to capture different aspects of inflation dynamics and reduces the risks associated with relying on a single model specification.

2.1.2. Data

CPI modelling and forecasting is based on an analysis of disaggregated indicators (classes) formed in accordance with the Nomenclature of Commodity Items for the Classification of Individual Consumption According to Purpose (COICOP) used by the Bureau of National Statistics (hereinafter – the BNS). Thus, in SSCIF the dependent variables are the various classes of goods and services that together make up Kazakhstan's CPI. The time series used in the system comprise monthly data from January 2011 to the present. The forecast is produced for the current quarter and the two following quarters.

The exogenous factors explaining the dynamics of these sub-indices are macroeconomic and external indicators, including:

- the exchange rate of the tenge against foreign currencies;
- inflation in the main trading partner countries;
- industrial producer price indices;
- agricultural producer price indices;
- money market indicators;
- real money income of the population;
- world prices of key commodities.

All variables are monthly seasonally adjusted time series. To ensure stationarity and comparability of the time series, the variables are transformed into log differences, which is consistent with standard practice in macroeconometric modelling.

2.1.3. Methodology and modelling techniques

The system uses the following main classes of models:

- Autoregressive models are used to describe the inertial properties of inflation processes. These models are based on the assumption that current price dynamics are partly determined by past dynamics. This approach captures the effect of inflation inertia and price stickiness typical of many categories of goods and services.

Within SSCIF, separate equations of this type are estimated for each CPI class:

$$\Delta y_t^k = \beta_0 + \beta_1 \Delta y_{t-1}^k + \beta_1 trend^k + \varepsilon_t^k, \quad (1)$$
$$k \in \overline{1, n\ classes}$$

where Δy_t^k is the log difference of CPI class k; Δy_{t-1}^k is the log difference of CPI class k in the previous month; $trend^k$ is the trend component of CPI class k; ε_t^k are the random errors specific to the equation for CPI class k.

The model class presented also corresponds to the classical approach to describing inflation dynamics through the mechanism of adaptive expectations, in which economic agents form expectations on the basis of past price dynamics, without explicitly including forward-looking components.

- Random walk models are used to model the stochastic component of price dynamics that cannot be explained by observable macroeconomic factors. Using this class of models makes it possible to capture random price fluctuations and the uncertainty of inflation processes.

$$\Delta\left(\Delta y_t^k\right)^j = \beta_0 + u_t^{kj}, \quad (2)$$
$$u_t^{kj} \in N(0,1),$$
$$k \in \overline{1, n\ classes},$$
$$j \in \overline{1,500};$$

where $\Delta\left(\Delta y_t^k\right)^j$ is the first difference of the log difference of CPI class k in random walk simulation number j; u_t^{kj} is a random error with a standard normal distribution for CPI class k in random walk simulation number j.

To model the dynamics of each CPI class, 500 random walk models are estimated. This is done in order to generate 500 draws of a random variable from the standard normal distribution, which serve as the explanatory factors in each model j for the dynamics of the dependent variable – CPI class k. The final random walk forecast for CPI class k is taken as the arithmetic mean of the forecasts from the 500 estimated models.

Random walk models, loosely imitating the Monte Carlo method, make it possible to obtain an asymptotic estimate of the stochastic fluctuations in the prices of goods and services that cannot be captured and described by the explanatory factors of inflation.

- Multiple regression models make it possible to quantify the impact of macroeconomic factors on the dynamics of individual inflation sub-indices. In these models, inflation is treated as a function of a set of exogenous variables reflecting the state of the domestic and external economic environment.

For each CPI class k:

$$\Delta y_t^k = \beta_0 + A(L^l)Z_t^k + \varepsilon_t^k, \quad (3)$$
$$k \in \overline{1, n\ classes},$$
$$l \in \{1, 2, \dots, 6\};$$

where Δy_t^k is the log difference of CPI class k, $\left(L^l\right)Z_t^k$ is the vector of explanatory factors for CPI class k with lag l, and ε_t^k are the random errors specific to the equation for CPI class k.

- Bayesian vector autoregressive models are used to model the relationships between the various components of inflation and macroeconomic variables. The

Bayesian approach makes it possible to incorporate prior information on the structure of economic relationships and to obtain more robust and economically interpretable parameter estimates.

To estimate the BVAR models, the CPI classes were divided into groups by expert judgement. It is assumed that a common group includes the prices of classes of goods and services that may influence one another through the microeconomic properties of substitutability and complementarity of market goods and services. Within SSCIF, BVAR models are estimated on the basis of the Litterman/Minnesota prior:

$$Y_t^g = X_t^g A + Z_t^g B + \varepsilon_t, \quad (4)$$
$$g \in \overline{1, n\ groups};$$

where Y_t^g are the endogenous factors – the vector of log differences of the CPI classes included in group g; X_t^g is the matrix of lagged values of the endogenous factors; Z_t^g is the vector of exogenous explanatory factors specific to group g; ε_t is the vector of random errors.

In addition to traditional econometric models, SSCIF applies machine learning methods, in particular regularised regression algorithms, including Ridge Regression, Lasso Regression and Elastic Net. These algorithms make it possible to work effectively with a large number of explanatory variables and to address typical problems of macroeconomic modelling such as multicollinearity and overfitting. Regularisation methods provide automatic selection of significant factors and make it possible to identify more complex relationships between economic variables. The integration of machine learning methods expands the analytical capabilities of the system and increases its ability to adapt to the changing structure of macroeconomic relationships.

Traditional regression methods assume that the explanatory variables are independent, but in reality these variables are often interrelated. This reduces the accuracy of the estimates and makes the model unstable. Ridge Regression minimises the impact of multicollinearity through regularisation, adding a penalty on the size of the regression coefficients and thereby stabilising the model.

$$L_{Ridge}(\beta) = \sum_{i=1}^{n}(Y_i - X_i\beta)^2 + \lambda \sum_{j=1}^{p} \beta_j^2, \quad (5)$$

where $L_{Ridge}(\beta)$ is the loss function that the algorithm minimises to find the optimal regression coefficients, Y is the dependent variable, X is the matrix of explanatory variables, β are the model coefficients to be estimated, and λ is the regularisation coefficient controlling the size of the (L2) "penalty" on large coefficient values.

- In turn, Lasso Regression not only estimates the model coefficients but also automatically excludes insignificant variables. The basic principle of Lasso is regularisation using the L1 norm, which adds a penalty on the absolute values of the coefficients. This changes the standard minimisation function as follows:

$$L_{Lasso}(\beta) = \sum_{i=1}^{n}(Y_i - X_i\beta)^2 + \lambda \sum_{j=1}^{p} |\beta_j|, \quad (6)$$

where $L_{Lasso}(\beta)$ is the loss function that the algorithm minimises to find the optimal regression coefficients, Y is the dependent variable, X is the matrix of explanatory variables, β are the model

coefficients to be estimated, and λ is the regularisation coefficient controlling the size of the "penalty".

- Elastic Net combines elements of two regularisation methods – Lasso and Ridge – which makes it possible both to deal with multicollinearity and to select the most significant variables.

Mathematically, Elastic Net is an extension of linear regression with two penalty functions added:

$$L_{Enet}(\beta) = \sum_{i=1}^{n}(Y_i - X_i\beta)^2 + \lambda_1 \sum_{j=1}^{p}|\beta_j| + \lambda_2 \sum_{j=1}^{p}\beta_j^2, \quad (7)$$

where $L_{Enet}(\beta)$ is the loss function that the algorithm minimises to find the optimal regression coefficients, Y is the dependent variable, X is the matrix of explanatory variables, and β are the model coefficients to be estimated.

In formula (7), λ_1 is responsible for L1 regularisation (as in Lasso), which sets irrelevant coefficients to zero, while λ_2 is responsible for L2 regularisation (as in Ridge), which shrinks all coefficients and stabilises the model in the presence of correlated variables.

2.1.4. Forecast combination

SSCIF uses a selective combination algorithm in which the model weights are calculated dynamically on the basis of their local accuracy over different segments of the test sample. The test sample is formed as the last *h* observations before the current point in time and corresponds to the forecast horizon.

The algorithm comprises several stages.

- At the first stage, all models are trained on the training sample of time series for each inflation sub-index.
- At the second stage, recursive test forecasts are produced, making it possible to assess the accuracy of each model over different segments of the time series. Forecast accuracy is measured using the root mean squared error (RMSE).

On the basis of the accuracy estimates obtained, recursive combination weights of the models are calculated for each point in time. Unlike static weights, which reflect the average accuracy of the models over the entire test period, recursive weights make it possible to take into account changes in the models' predictive performance over time.

Thus, each time a forecast is produced, the system assigns a higher weight to those models that show higher accuracy over the corresponding segment of the time series.

The final inflation forecast is produced in several stages.

First, forecasts for each consumer price sub-index are calculated through selective combination. The forecasts obtained are then aggregated by group, producing forecasts of food, non-food and services inflation.

At the final stage, these forecasts are aggregated using the consumer basket weights, which gives the final forecast of the overall consumer price index.

Thus, SSCIF is a comprehensive system for short-term inflation forecasting based on a combination of disaggregated modelling, a multi-model approach and adaptive forecast combination methods. The use of selective combination and recursive weighting of models substantially improves the robustness and accuracy of inflation forecasts in a volatile macroeconomic environment.

2.2 The KOMPAS Model for Short-Term Forecasting of GDP by the Expenditure Approach

2.2.1. System description

To forecast gross domestic product by the expenditure approach (hereinafter – GDP-E), the NBK uses the KOMPAS system (from the Russian acronym for "Comprehensive General Model for Forecasting Aggregate Demand"; hereinafter – KOMPAS). This approach involves producing a demand-side GDP forecast by modelling each of its components separately using different modelling techniques. KOMPAS has been in use since 2021 and is the result of improving the methodologies previously used for short-term GDP-E forecasting and transforming them into a single modelling framework.

KOMPAS comprises about 20 thousand equations and is implemented in EViews. The forecasting equations are estimated by program code that re-estimates the model on a quarterly basis as new statistical data are published. The modelling system as a whole is endogenous, with the exception of a number of exogenously set parameters used to produce forecasts under three scenarios (baseline, optimistic and pessimistic).

The final GDP forecast is produced by aggregating the forecast values of the individual demand components using their nominal weights for the corresponding period of the previous year. Within KOMPAS, different modelling methods are used to forecast the components, and these may show different predictive power over different time intervals.

The resulting short-term forecasts of GDP and its components are used as inputs to the QPM in producing the medium-term macroeconomic forecast.

2.2.2. Data

The time series used in the system comprise quarterly and monthly data from 2005 to the present. All input data, including the GDP-E components, are used in real terms and are seasonally adjusted beforehand.

The explanatory variables are mainly high-frequency statistical indicators (predominantly monthly) describing the dynamics of the real sector of the economy and selected financial sector indicators, as well as budget parameters and external macroeconomic indicators.

The Brent oil price and the nominal US dollar exchange rate are used as exogenous variables in producing the scenario forecasts. In addition, the exogenous assumptions include forecasts of macroeconomic indicators for trading partner countries (GDP growth rates and inflation), as well as projected oil production volumes in Kazakhstan.

2.2.3. Methodology and modelling techniques

KOMPAS produces a GDP forecast four quarters ahead. The forecasting process is conventionally divided into two stages: the nowcast (which also includes the backcast) and the near-term forecast (hereinafter – NTF). This structure reflects the existing lag in the publication of statistics on GDP by the expenditure approach, which exceeds 90 days, as well as differences in the availability of other statistical data at different forecast horizons. In particular, at the nowcast stage, models are used that can take account of partially observed timely statistics. At the NTF stage, the forecast is produced in the absence of observed values for the periods being forecast, which implies a significant role for model relationships and exogenously set scenario assumptions. Consequently, the forecasting techniques for the nowcast and NTF stages differ somewhat.

Nowcast

At the nowcast stage, various approaches are used, including MIDAS models, random walk models, multiple regression models, ARIMA models and regularisation (shrinkage) models.

To forecast quarterly GDP dynamics, mixed-frequency models are used, which make it possible to include indicators published at a higher frequency in the estimation. In particular, KOMPAS uses MIDAS models, which produce a quarterly forecast of GDP and its components on the basis of timely monthly data from the Bureau of National Statistics (hereinafter – the BNS) and other sources. The MIDAS specification makes it possible to use the high-frequency information available at the time of forecasting without aggregating it. Aggregating high-frequency data leads to a loss of information on intra-period dynamics, which can reduce their predictive content when modelling quarterly indicators.

In general form, the model can be represented as a relationship between a low-frequency indicator and a weighted sum of lags of a high-frequency variable, with the effect of the lags given by a parameterised weighting function. The equation has the following specification:

$$y_t = \beta_0 + \sum_{j=0}^{K} w(j;\theta) x_{t-j/m} + \varepsilon_t, \quad (8)$$

y_t is the quarterly dependent variable (a GDP-E component), $x_{t-j/m}$ is the high-frequency (monthly) indicator with lag j, $w(j;\theta)$ is the lag weighting function defined by the parameters θ, K is the maximum number of lags, m is the frequency ratio (for example, 3 months in a quarter), and ε_t is the random error.

For each GDP component, several alternative specifications of MIDAS models are estimated, differing in the way the lag weights are defined. On the basis of the estimated MIDAS specifications, forecasts are produced, which are then combined.

Multiple regression models are used to assess the impact of various macroeconomic factors on the dynamics of GDP components. The multiple regression models have a standard linear specification:

$$y_t = \alpha + \sum_{i=1}^{k} \beta_i x_{i,t} + \varepsilon_t, \quad (9)$$

where y_t is the indicator being forecast (a GDP-E component) in period t, $x_{i,t}$ is the set of explanatory variables (i) observed in period t, β_i are the estimated parameters/coefficients on explanatory variable i, ε_t is the random error in period t, and k is the number of explanatory variables.

To capture the inertial properties of the time series and random fluctuations, ARIMA models and a random walk model are used. ARIMA models make it possible to model inertial processes in which, in the absence of significant shocks, economic indicators maintain their established dynamics for a certain period. In general form, an ARIMA model can be represented as follows:

$$\varphi(L)(1-L)^d y_t = \theta(L)\varepsilon_t, \quad (10)$$

where y_t is a GDP-E component, L is the lag operator, $\varphi(L)$ and $\theta(L)$ are lag polynomials, d is the order of differencing, and ε_t is the random error.

The random walk model is used to account for unpredictable changes and shocks that cannot be incorporated into the models in advance.

In addition, to reduce the risk of overfitting when a large number of explanatory variables is used, regularisation methods (shrinkage models) such as Ridge and Lasso regressions are applied. These methods automatically limit the influence of less significant factors and thereby improve the quality of the forecast estimates.

The estimates obtained are then used as inputs for producing the near-term forecast (NTF) for the following two quarters.

Near-Term Forecast

The near-term forecast is produced under a scenario approach, in which the inputs include exogenous parameters set outside KOMPAS. These comprise the key external and scenario assumptions.

At this stage, vector autoregressions (VAR) are used instead of MIDAS models, making it possible to capture the endogenous relationships between GDP components. In general form, the VAR models are represented as follows:

$$Y_t = c + \sum_{l=1}^{p} A_l Y_{t-l} + BZ_t + \varepsilon_t, \quad (11)$$

where Y_t is the vector of endogenous GDP-E components, c is the vector of constants, Y_{t-l} is the vector of lagged values of the endogenous GDP-E components (l is the lag, p is the maximum number of lags), A_l is the matrix of coefficients,

Z_t is the vector of exogenous variables, including scenario variables, B is the vector of coefficients on the exogenous variables, and ε_t is the vector of random errors.

Automated ARIMA models and random walk models are also used to produce the forecasts, making it possible to account for uncertainty and possible unpredictable changes in the dynamics of the indicators.

The forecasts of each component obtained using the different modelling methods are then combined.

2.2.4. Forecast combination

The estimates of the GDP components are combined using weighting by root mean squared errors (RMSE), according to the equation:

$$Y^{f,i}_{combined,t_k} = \sum_{j=1}^{n} \omega_{j,k} \hat{y}^{i,j}_{t_k} \text{ , (12)}$$

where $Y^{f,i}_{combined,t_k}$ is the final combined forecast for demand component i for period t_k; $\hat{y}^{i,j}_{t_k}$ is the out-of-sample forecast obtained from model j for demand component i for period t_k; $\omega_{j,k}$ is the weight of the forecast of model j in the combined forecast for period t_k, calculated as:

$$\omega_{j,k} = \frac{1/RMSE_{j,k}}{\sum_{m=1}^{n} 1/RMSE_{m,k}}, \text{ (13)}$$

n is the number of models in the forecast combination procedure.

All models involved in the forecast receive weights that are re-estimated in each forecast round.

The method is based on an expanding window: for each forecast horizon, the model weights are calculated on the basis of their errors (RMSE) over a pseudo-out-of-sample segment of the data. As the data are updated, the window expands and the RMSE is recalculated for each new period. This approach makes the combination adaptive: the weights are redistributed in favour of the models showing the best forecast accuracy in the current forecast round.

As a result, a short-term forecast is produced under three scenarios (baseline, optimistic and pessimistic), which is used as an input to the QPM in producing the medium-term macroeconomic forecast.

2.3 Short-Term Forecasting of GDP by the Production Approach. System for Comprehensive Analysis and Forecasting of the Economy of Kazakhstan.

2.3.1. System description

Quarterly data on GDP by the production approach are published on the 45th day after the end of the reporting period. At the same time, the structure of GDP by the production approach, consisting of a set of sectoral components, complicates the use of traditional econometric approaches because of the high interdependence of the indicators and the large number of variables.

To produce the short-term forecast of GDP by the production approach, the System for Comprehensive Analysis and Forecasting of the Economy of Kazakhstan (SCAFEK) is used. The system is based on the decomposition of GDP into sectoral and sub-sectoral components and provides for the sequential aggregation of forecast estimates in line with the structure of the economy. This approach ensures that the forecast calculations are consistent with the statistical structure of GDP and makes it possible to take into account the contribution of individual economic activities to the final estimate of economic growth.

Forecast estimates of the individual components are produced using various quantitative methods, including autoregressive models, linear regressions, mixed-frequency (MIDAS) models and dynamic factor models (DFM). Alongside the model-based estimates, the system can take into account expert assumptions and additional sectoral information affecting the expected dynamics of individual indicators.

The forecast estimates are aggregated using the structure of GDP and the previous year's weights, which ensures the comparability of the sectoral contributions and the final estimate of economic growth. The decomposition structure of the system also ensures the transparency of the forecast calculations and makes it possible to analyse the influence of individual sectors and sub-sectors on the GDP forecast.

2.3.2. Data

SCAFEK includes a detailed structure of GDP by the production approach, consisting of the main sectors of the economy and their sub-sectoral components. For each component, growth rates and nominal volumes are recorded and used for the subsequent aggregation of the forecast estimates. This structure covers the production of goods and services and makes it possible to produce the GDP forecast on the basis of the dynamics of individual economic activities.

A broad panel of monthly and quarterly indicators reflecting key aspects of economic activity is used to construct the model estimates. The sample includes indicators of the real sector, price dynamics, monetary conditions, fiscal policy, corporate reporting and external economic conditions. The use of different blocks of indicators provides a comprehensive picture of the interaction between demand and supply, as well as of the influence of domestic and external factors on business activity.

At the data preparation and preliminary modelling stage, some of the initial indicators may be excluded because of low variability, insufficient information content or a limited contribution to the quality of the model estimates.

2.3.3. Methodology and modelling techniques

The time series are brought to a comparable form using standard procedures. Seasonal adjustment is carried out. Taking logarithms and differencing are applied to achieve stationarity, and standardisation removes differences in scale. For individual

indicators with pronounced autocorrelation, a procedure may be applied to first separate out the autocorrelated component so that the cleaned residuals can be used.

The dynamic factor model (DFM) is used to reduce the dimensionality of the information set and to extract the common factors driving the co-movement of macroeconomic indicators. In a broad panel of monthly macroeconomic indicators $x_{i,t}$, $i = 1, \dots, N$ denotes the variable and $t = 1, \dots, T$ the time period. The main assumption of the dynamic factor model is that the co-movement of these series can be described by a small number of unobserved factors $f_t \in R^r$. Formally, the measurement equation takes the form:

$$x_t = \Lambda f_t + \varepsilon_t, \quad (14)$$

where x_t is the vector of observed variables of dimension $N \times 1$, Λ is the factor loading matrix of dimension $N \times r$, and ε_t is the vector of idiosyncratic errors. The model assumes a diagonal covariance structure of the errors.

The dynamics of the latent factors f_t are restricted to a set of independent first-order autoregressions. Formally, for each factor $j = 1, \dots, r$:

$$f_{j,t} = \rho_j f_{j,t-1} + \eta_{j,t}, \ \ \eta_{j,t} \sim N(0, \sigma_j^2), \quad (15)$$

where ρ_j characterises the persistence of the factor and $\eta_{j,t}$ is an innovation with variance σ_j^2.

This specification helps to preserve persistence in the common component, while the transition equation remains fairly simple.

In addition to the dynamics of the common factors, the idiosyncratic components $\varepsilon_{i,t}$ corresponding to each observed series are also allowed to follow independent AR(1) processes:

$$\varepsilon_{i,t} = \phi_i \varepsilon_{i,t-1} + \xi_{i,t}, \ \ \xi_{i,t} \sim N(0, \psi_i^2), \quad (16)$$

where ϕ_i reflects series-specific persistence and $\xi_{i,t}$ is white noise.

By imposing a diagonal covariance structure on the idiosyncratic errors, all systematic co-movement in the data is attributed to the factors rather than to correlated measurement noise.

The combination of the AR(1) dynamics of the factors and of the idiosyncratic components forms a state-space system that is computationally convenient and robust in limited samples. Estimation is carried out by maximum likelihood using the Kalman filter, which makes it possible to determine the model parameters and the paths of the latent factors from the observed data panel.

Once the factor time series have been obtained, the forecasting equation for quarterly GDP can be estimated by weighted least squares with exponentially decaying weights. The aim is to reduce the influence of older observations reflecting economic trends that may not be characteristic of current macroeconomic dynamics. Each point in time t receives the weight:

$$w_t = \gamma^{T-t}, \ \ 0 < \gamma < 1, \quad (17)$$

where T is the last available point in the sample.

Under this approach, the most recent observations retain their full influence, while the contribution of earlier data declines smoothly and exponentially. An intuitive measure of the sensitivity of the weighting function is its half-life, that is, the number of periods after which the weight is halved:

$$h = \frac{\ln(0.5)}{\ln(\gamma)}. \quad (18)$$

Exponential discounting of observations is applied directly at the stage of estimating the regression equation of GDP on the factors, through the weighting of observations in the objective function:

$$y = \alpha + \beta' f_t + u_t, \quad (19)$$

where the contribution of each residual u_t is scaled by the weight w_t.

This approach assumes gradual structural changes and provides a controlled influence of the most recent observations while taking into account all historical information.

To make the forecasting equation more robust, ridge regression – a regularised version of the least squares method – may additionally be applied. This approach introduces a penalty on the size of the coefficients, which stabilises the estimates under strong multicollinearity or with an expanded set of regressors when additional variables are included. The model minimises a modified quadratic error function:

$$\beta, \alpha = \arg\min_{\alpha,\beta} \left[\frac{1}{T}\sum_{t=1}^{T}(y_t - \alpha - \beta' f_t)^2 + \lambda_R \beta'\beta\right], \quad (20)$$

where $\lambda_R \geq 0$ is the regularisation parameter.

With $\lambda_R = 0$, the model coincides with the least squares method, whereas for positive values of the penalty the coefficient estimates become smoother and less sensitive to random fluctuations in the data. This is particularly important in applied macroeconomic panels, where time series often show substantial mutual correlation and the inclusion of additional variables can lead to unstable or excessively variable coefficients.

The two regularisation methods address different aspects of the problem of instability of the forecasting equation. Using these methods together makes it possible to take into account changes in economic relationships over time, as well as redundancy and collinearity of the predictors.

3. MEDIUM-TERM FORECASTING

For medium-term forecasting, an approach different from the short-term econometric techniques is used – the Quarterly Projection Model (QPM). The QPM for Kazakhstan is a semi-structural model methodologically based on New Keynesian macroeconomic theory. The model reflects the main channels of monetary policy transmission and takes into account the key structural features of Kazakhstan's economy, including the significant role of fiscal policy and the oil sector. The model is built around a system of behavioural equations describing the relationships between inflation, economic activity, exchange rate dynamics, the interest rate and external sector variables.

The QPM is built using gap variables. A gap is the relative deviation of the actual value of an indicator from its potential level. The potential level (trend) is a long-term tendency determined by fundamental factors. In the model, potential values are generally specified as autoregressive processes converging to their long-run steady states. Trend variables are unobserved and are estimated using a multivariate Kalman filter. Their values change smoothly and gradually, and the corresponding equations are characterised by high persistence and low shock variance.

The basic structure of the model comprises three endogenous blocks: the Phillips curve, the investment–savings (IS) curve and the monetary policy block. In addition, exogenous blocks are built into the model – the fiscal and oil blocks, as well as the external sector.

3.1 Phillips Curve Block

Inflation dynamics are described by a hybrid New Keynesian Phillips curve. The Phillips curve block is split into food and non-food inflation components (21). The non-food inflation equation combines the services and non-food components.

$$PI = \omega * PI_F + (1 - \omega) * PI_{NF} + \varepsilon_{PI}, \quad (21)$$

where PI is core inflation (q/q), PI_F is food inflation (q/q), PI_{NF} is non-food inflation (q/q), excluding prices of fuels and lubricants and regulated housing and utility services, and ε_{PI} are weights errors.

$$PI_F = a_{1_{F_1}} * PI_{F\,exp} + \left(1 - a_{1_{F_1}}\right) * PI_F\{-1\} + a_{1_{F_2}} * \left[a_{1_{F_3}} * Y_{gap}\{-2\} + a_{1_{F_4}} * REER^F_{GAP}\{-2\}\right] + \varepsilon_{PI_F}, \quad (22)$$

where $PI_{F\,exp}$ is food inflation expected one quarter ahead, $PI_F\{-1\}$ is food inflation in the previous quarter, Y_{gap} is the output gap, $REER^F_{GAP}$ is the real effective exchange rate (REER) gap for food, and ε_{PI_F} are supply shocks on food markets.

$$PI_{NF} = a_{1_{NF_1}} * PI_{NF\,exp} + \left(1 - a_{1_{NF_1}}\right) * PI_{NF}\{-1\} + a_{1_{NF_2}} * \left[a_{1_{NF_3}} * Y_{gap}\{-1\} + a_{1_{NF_4}} * REER^{NF}_{GAP}\{-1\}\right] + \varepsilon_{PI_{NF}}, \quad (23)$$

where $PI_{NF_{exp}}$ is non-food inflation expected one quarter ahead, $PI_{NF}\{-1\}$ is non-food inflation in the previous quarter, Y_{gap} is the output gap, $REER_{GAP}^{NF}$ is the REER gap for non-food inflation, and ε_{PI_F} are supply shocks on non-food and services markets.

The Phillips curve equation shows that in equilibrium (the output gap and the REER gap equal zero and there are no supply shocks), actual inflation will correspond to a weighted average of future and past values of food and non-food inflation, and, once the target is reached and expectations are anchored accordingly, it will remain at the target level.

3.2 IS Curve Block

The domestic economy is described by the IS curve; demand-side GDP dynamics are characterised by equation (24):

$$Y_{GAP} = sh_{dd} * DD_{GAP} + sh_G * G_{GAP} + sh_X * X_{GAP} - sh_I * M_{GAP} + \varepsilon_{Y_{GAP}}, \quad (24)$$

where DD_{GAP} is the domestic demand gap, G_{GAP} is the government consumption gap, X_{GAP} is the export gap, M_{GAP} is the import gap, and $\varepsilon_{Y_{GAP}}$ is the combined shock of changes in weights.

Domestic demand in the QPM combines household consumption and gross fixed capital formation.

$$DD_{GAP} = b_1 * DD_{GAP}\{-1\} + b_2 * RIR_{GAP} + b_3 * PO_{GAP}\{-1\} + b_4 * FiscalIndicator + \varepsilon_{DD_{GAP}}, \quad (25)$$

where $DD_{GAP}\{-1\}$ is the inertia of the domestic demand gap, RIR_{GAP} is the real interest rate gap, PO_{GAP} is the oil price gap, $FiscalIndicator$ is the fiscal policy indicator, and $\varepsilon_{DD_{GAP}}$ are domestic demand shocks.

$$G_{GAP} = \alpha_1 * G_{GAP}\{-1\} + \alpha_2 * [e_1 * PO_{GAP} + (1 - e_1) * PO_{GAP}\{-1\}] + \alpha_3 * TR_{GAP} + \varepsilon_{G_{GAP}}, \quad (26)$$

where $G_{GAP}\{-1\}$ is the inertia of the government consumption gap, TR_{GAP} is the gap in transfers from the National Fund of the Republic of Kazakhstan, and $\varepsilon_{G_{GAP}}$ are government consumption shocks.

The export equation consists of non-oil ($EX_{GAP}^{Non-oil}$) and oil exports (EX_{GAP}^{Oil}), as well as a total export shock:

$$X_{GAP} = sh_{OX} * EX_{GAP}^{Oil} + (1 - sh_{OX}) * EX_{GAP}^{Non-oil} + \varepsilon_{EX}, \quad (27)$$

Non-oil exports are given by the equation:

$$EX_{GAP}^{Non-oil} = \delta_1 * EX_{GAP}^{Non-oil}\{-1\} + \delta_2 * REER_{GAP}^{F}\{-1\} + \delta_3 * REER_{GAP}^{NF}\{-1\} + \delta_4 * YNO_{GAP}^{*} + \varepsilon_{EX_{GAP}^{Non-oil}}, \quad (28)$$

where $EX_{GAP}^{Non-oil}\{-1\}$ is the inertia of non-oil exports, YNO_{GAP}^{*} is the effective foreign non-oil output gap, and $\varepsilon_{EX_{GAP}^{Non-oil}}$ is the non-oil export shock.

YNO^{*}_{GAP} characterises external demand, which is approximated by the business cycle phases of the main partners – China, the European Union and the Russian Federation. The country weights are determined by their shares in non-oil exports.

The import equation is a function of domestic demand and also takes into account the effect of the exchange rate channel.

$$M_{GAP} = x_1 * [\theta_1 * DD_{GAP}\{-1\} + (1-\theta_1) * DD_{GAP}\{-2\}] - x_2 * REER^{NF}_{GAP}\{-3\} + \varepsilon_{Im_{GAP}}, \quad (29)$$

3.3 Monetary Policy Block

The monetary policy block consists of the exchange rate equation and the Taylor rule.

Exchange rate equation

In the standard version, under perfect capital mobility and developed financial markets, the change in the exchange rate depends directly on the interest rate differential and the risk premium. For emerging economies, including Kazakhstan, it is assumed that the exchange rate is affected through two channels:

$$\Delta S = (1 - capmob) * \Delta S_1 + capmob * \Delta S_2 + \varepsilon_{\Delta S}, \quad (30)$$

where $capmob$ is capital mobility and $\varepsilon_{\Delta S}$ is the shock of the deviation of actual capital mobility $(capmob)$ from that in the model.

Equation (31) describes the trade channel (BoP channel) and includes ΔS_{EQ} – the equilibrium change in the nominal exchange rate – and BoP_{GAP} – the trade balance gap adjusted for the terms of trade ($BoP_{GAP} = TOT_{GAP} + X_{GAP} - M_{GAP}$).

$$\Delta S_1 = \Delta S_{EQ} - \rho_1 * BoP_{GAP}, \quad (31)$$

Equation (32) reflects the effect of the uncovered interest parity channel, or the financial channel (UIP channel).

$$\Delta S_2 = \rho_2 * \Delta S_2\{1\} + (1-\rho_2) * \Delta S\{-1\} - (IR - IR^{*} - prem_{SS}) - \rho_3 * BoP_{GAP} - \rho_4 TR_{GAP}, \quad (32)$$

where $\Delta S_2\{1\}$ is the change in the nominal exchange rate expected one quarter ahead through the UIP channel, IR is the NBK's domestic interest rate, IR^{*} is the foreign nominal interest rate, represented by the US Federal Reserve rate, $prem_{SS}$ is the risk premium, and TR_{GAP} is the gap in transfers from the National Fund of the Republic of Kazakhstan.

Taylor rule

$$IR = (1 - capmob) * IR_1 + capmob * IR_2 + \varepsilon_{IR}, \quad (33)$$

where IR is the NBK's nominal interest rate and ε_{IR} is the monetary policy shock.

The monetary policy rule is a modified Taylor rule that responds to the deviation of the inflation forecast from the target and to the output gap, and additionally to the deviation of the change in the exchange rate from its equilibrium change.

$$IR_1 = \beta_1 * IR\{-1\} + (1-\beta_1) * \left(RIR_{EQ} + PI_{TAR}\right) + \mu * (\Delta S - \Delta S_{EQ}), \quad (34)$$

where $IR\{-1\}$ is the NBK's nominal interest rate in the previous quarter, RIR_{EQ} is the real neutral interest rate, and PI_{TAR} is the inflation target.

$$IR_2 = \beta_1 * IR\{-1\} + (1-\beta_1) * \left(RIR_{EQ} + PI_{TAR}\right) + \varphi * (PI_4\{4\} - PI_{TAR}) + \epsilon Y_{GAP}, \quad (35)$$

where $PI_4\{4\}$ is expected annual core inflation one year ahead.

3.4 Exogenous Blocks

The exogenous blocks in the QPM comprise the external sector, the oil sector and fiscal policy indicators. In the model they are represented by autoregressive processes.

The external sector includes the foreign non-oil output gap (YNO^*_{GAP}), foreign aggregate inflation (PI^*), foreign aggregate food inflation (PI^*_F), the foreign nominal interest rate (IR^*) and the oil price gap (PO_{GAP}).

The oil sector represents the volumes of real exports of oil and gas condensate. The fiscal policy indicator is an aggregate measure that takes into account the non-oil budget deficit and quasi-fiscal operations, adjusted for the stage of the business cycle.

4. EXTERNAL SECTOR MODELS

4.1 Forecasting the Current Account of the Balance of Payments

The forecast of the current account of Kazakhstan's balance of payments includes estimates of the three main components of the current account – the trade balance, the balance of services and the income balance. The final current account forecast is produced by aggregating the forecasts for these components according to the formula:

$$CA_t = TB_t + BS_t + IB_t = (Ex_G_t - Im_G_t) + (Ex_S_t - Im_S_t) + (I_R_t - I_P_t), \quad (36)$$

where CA_t is the current account, TB_t is the trade balance, BS_t is the balance of services, IB_t is the income balance, Ex_G_t is exports of goods, Im_G_t is imports of goods, Ex_S_t is exports of services, Im_S_t is imports of services, I_R_t is income receivable, I_P_t is income payable, and t is the forecast period.

4.1.1. Trade balance

The technique for forecasting **exports of goods** is based on the following formula:

$$Ex_G_t = \sum (P_{it} * Q_{it}), \quad (37)$$

where Ex_G_t is the value of exports of goods, P_{it} is the price of good i, Q_{it} is the volume of exports of good i, and t is the forecast period.

The forecast of exports of goods is produced for nine main commodity groups: oil and gas condensate, petroleum products, non-ferrous metals, ferrous metals, gas, coal, uranium, grains and other exports.

The forecast prices of exported goods are determined on the basis of NBK scenarios, as well as information from the international analytical agency Consensus Economics and other sources. The forecast export volumes are based on data from sectoral government bodies and leading industry companies, open sources of information and, where necessary, expert judgement.

The technique for forecasting imports of goods includes a forecast of official imports of goods and of adjustments under the balance of payments methodology, and is calculated using the following formula:

$$Im_G_t = \sum Im_G_official_{it} + \sum BoP_est_{jt}, \quad (38)$$

where Im_G_t is the value of imports of goods, $Im_G_official_{it}$ is official imports of commodity group i, BoP_est_{jt} is balance of payments methodology adjustment j, and t is the forecast period.

Official imports of goods are forecast for two commodity groups – consumer and investment goods. Dynamic regression models with exogenous variables and an inertial process are used for the forecast, and expert assumptions are made where necessary.

In forecasting official imports of goods, macroeconomic indicators are analysed such as Kazakhstan's real and nominal GDP, the nominal exchange rate, the REER and NEER of the tenge, fixed capital investment, foreign direct investment, loans issued in various breakdowns, real and nominal money income of the population, wages, and government spending.

The balance of payments methodology adjustments to imports of goods include freight, shuttle trade and e-commerce. The freight forecast is produced by applying actual estimated ratios to the forecast volume of official imports of goods, and the shuttle trade forecast by applying estimated ratios to the number of residents departing from Kazakhstan. E-commerce is forecast using an econometric method – an autoregressive distributed lag (ARDL) model.

4.1.2. Balance of services

The forecast of exports of services is produced for the main components: transport services, travel and other services. The forecast of imports of services is developed for transport services, travel, construction, business and other services.

Econometric models (ARIMA, VECM and/or others), calculated ratios and expert judgement are used in the forecasting process. The choice of forecasting method is determined by the specifics of each forecast component of the balance of services.

4.1.3. Income balance

The income balance forecast is produced separately for income receivable and income payable, which include primary and secondary income.

Econometric time series models and regression models (OLS, ARDL, ARMA CLS and others) are used to forecast the components of the income balance, and expert adjustments are allowed. For some less significant components of the income balance, calculation methods are used, including averaging of historical values.

The forecasting approaches and models, data sources and set of expert assumptions used in forecasting the current account of the balance of payments may be adjusted and refined as new statistical information becomes available, external economic conditions change, structural shifts are identified or the quality of the forecast estimates deteriorates. These changes are made in order to improve forecast accuracy and are introduced following a comprehensive review and expert discussion of the available information at each stage of their approval. The models are recalibrated quarterly within the NBK's forecast rounds.

5. IMPROVEMENT AND DEVELOPMENT OF THE FORECASTING AND ANALYTICAL SYSTEM

5.1 Research Activities at the National Bank

The Forecasting and Policy Analysis System (FPAS) involves not only the use of existing models and tools but also their continuous development and improvement. In a changing economic environment and with the emergence of new sources of uncertainty, there is a growing need to regularly update the methodologies and approaches used, which calls for continuous research.

Research work is aimed at improving the quality of forecasts, expanding the analytical capabilities of the system and adapting the models to structural changes in the economy. This work involves reviewing existing models, calibrating and upgrading them, and developing new tools for analysis and forecasting.

An important area is the creation and development of so-called “satellite” models, which complement the core forecasting tools and allow individual sectors and factors to be analysed in greater detail. Research and development work is under way on assessing the impact of fiscal policy, oil price dynamics and the exchange rate on macroeconomic indicators.

Thus, research is an integral part of the FPAS and ensures its flexibility, relevance and ability to respond effectively to new challenges and changes in the economic environment.

Research is published on the NBK website and in other print publications, and is presented at various domestic and international conferences, expert meetings, etc.

5.2 Models and Tools under Development

As a continuation of the research carried out within the FPAS, new models and analytical tools are being developed and implemented to expand the capabilities of macroeconomic analysis and forecasting.

One of the key areas is the development of a DSGE (Dynamic Stochastic General Equilibrium) model of a small open economy with an oil sector. This class of models is based on microeconomic foundations of the behaviour of economic agents and makes it possible to form a coherent picture of the functioning of the economy, taking into account domestic and external shocks. The DSGE approach ensures the consistency of macroeconomic relationships and makes it possible to carry out scenario analysis and assess the consequences of monetary and fiscal policy.

The model can be used for in-depth analysis of key macroeconomic trends and transmission channels, including the impact of oil prices, external conditions and the exchange rate. Once testing and validation have been fully completed, the model may be integrated into the forecasting techniques and used as an additional forecasting tool alongside the QPM.

In addition, models aimed at forecasting and analysing the financial sector are being developed. These tools focus on a more detailed assessment of financial conditions, lending dynamics, interest rates and the monetary policy stance, which will strengthen the analytical component of the FPAS.

Recommended reading:

1. Chernyavskiy D. O. Quarterly Projection Model of the Republic of Kazakhstan and Its Role in Monetary Policy Decision-Making. May 2017. – URL: https://nationalbank.kz/file/download/9032

2. Chernyavskiy D. O., Mukanov N. S. Introducing a Monetary Policy Rule into the Quarterly Projection Model of Kazakhstan // Money and Credit. 2017, No. 5.

3. Tuleuov O. A. System of Selective Combined Inflation Forecast (SSCIF): Choosing the Optimal Technique for Forecasting Consumer Price Dynamics under a Structural Shock (the Case of Kazakhstan). National Bank of the Republic of Kazakhstan. 2017. – URL: https://nationalbank.kz/file/download/9019

4. Samat M., Bukenov A. & Taibekova A. The System for Short-Term Forecasting of GDP by the Expenditure Approach at the National Bank of the Republic of Kazakhstan. National Bank of the Republic of Kazakhstan. 2022. – URL: https://nationalbank.kz/file/download/85926

5. Adilkhanova Z., Yerzhan I. Selective-Combined Inflation Forecasting System. National Bank of the Republic of Kazakhstan. 2024. – URL: Selective-Combined Inflation Forecasting System

6. Akhmet A. Forecasting Kazakhstan's GDP Based on a Dynamic Factor Model with Regularization. National Bank of the Republic of Kazakhstan. 2026. – URL: https://nationalbank.kz/file/download/119724