NATIONAL BANK OF KAZAKHSTAN

Dynamic Factor Model of Inflation for Kazakhstan

Monetary Policy Department
Economic Study No.2022-4

Yerzhan I.
Seidakhmetov A.

Dynamic Factor Model of Inflation for Kazakhstan.
NBRK – WP – 2022-4

Dynamic Factor Model of Inflation for Kazakhstan.

Yerzhan Islam[1]
Seidakhmetov Ansar[2]

Abstract

It is extremely important for central banks adhering to the policy of inflation targeting to monitor the current trends in the development of inflationary processes. In addition to the standard analysis of influence by various variables, it is necessary to understand the nature of where the main drivers of inflationary dynamics are stemming from. In this regard, the constant analysis of the impact on inflationary processes by unobserved variables is of particular importance.

As part of this study, the authors have defined unobserved variables and made a quantitative assessment of their effect on the dynamics of inflationary processes.

Key Words: inflation, principal component analysis, factor, an unobserved variable, contribution, decomposition.

JEL-Classification: E31, E37, E39, E52.

[1] Yerzhan Islam – Chief Specialist-Analyst, Macroeconomic Forecasting and Research Division, Monetary Policy Department, National Bank of the Republic of Kazakhstan. E-mail: Islam.Yerzhan@nationalbank.kz

[2] Seidakhmetov Ansar – Leading Specialist-Analyst, Macroeconomic Forecasting and Research Division, Monetary Policy Department, National Bank of the Republic of Kazakhstan. E-mail: Ansar.Seidakhmetov@nationalbank.kz

Contents

1. Preamble

In 2015, the National Bank of Kazakhstan moved to a new monetary policy regime – inflation targeting. The introduction of a system for analyzing and forecasting macroeconomic variables, inflation in particular, is an important aspect of effective inflation targeting. This system involves the construction of structural macroeconomic and econometric toolkit using a wide class of models to support monetary policy decisions. Pursuing a balanced and consistent monetary policy leads to the stabilization and reduction in inflation volatility by increasing confidence in the regulator and diminishing inflationary expectations.

The existence of a considerable number of variables affecting the inflation dynamics complicates the forecasting and analysis process. One has to single out really significant indicators among a large volume of data. One of the methods to solve this problem is the principal component analysis (PCA).

The principal component analysis is a statistical method that allows transforming the inter-correlated variables into a new contracted data set of uncorrelated variables called principal components. In doing so, conversion to a data set with lesser dimensions is accomplished with a minimum loss of general information. The data transformation is defined so that the first principal component has the maximum possible variance of all explanatory variables included in the analysis. Subsequent principal components include the residual variance.

In their study, Stock and Watson (2002) empirically proved that it makes sense to use the principal component analysis in forecasting with a large number of predictors. In forecasting the index of industrial production, the initial data set included 149 explanatory variables of monthly frequency (production, consumption, employment, inflation, interest rates, etc.) and was further predicted by using the principal component analysis. During the study, it turned out that the maximum improvement in forecasts in comparison with naive AR and VAR models is achieved through the use of the first two or three principal components. The inclusion of additional principal components does not lead to a significant improvement in the quality of forecasts.

In addition to the convenience of using a smaller dataset, principal component analysis avoids the curse of dimensionality. Along with this, the principal component analysis is one of the tools that enables to group variables by the nature of their origin in order to obtain a common unobserved variable.

As part of this work, the authors identified unobserved variables that reflect the influence of various aspects (external factors, supply and demand) and made a quantitative assessment of their impact on inflationary processes by building an OLS model. At the same time, as an additional analysis, the authors assessed the contribution of various factors by adding various variables to the OLS model with further grouping by nature of origin. Additionally, a vector autoregression (VAR) model was built to evaluate the "variance decomposition" to interpret the relationships between variables described by the model.

This paper consists of several parts. The second part contains a literature review where similar studies of other authors are discussed. The third section describes the methodology. It is followed by the discussion of outcomes where the authors of this paper share the description of assessment results. Findings and recommendations of this research are the final section in the paper.

2. Literature Review

To obtain a realistic and accurate forecast and analysis of inflation, certain technical requirements have to be met. For example, a sufficiently large amount of data should be available, and appropriate models should be developed to forecast and analyze inflation.

The study of inflation and its sources in the economy is important because it plays a significant role in formulating the monetary policy pursued by the central bank. A study by the International Monetary Fund (Loungani & Swagel, 2001) found that the source of inflation in developing countries is not uniform and varies across continents. In particular, it is noted that inflation factors in African and Asian countries differ, as most of the latter tend to have lower or moderate rates of mean inflation. The study also showed that fiscal factors, which are reflected in the money supply growth, exchange rate peg adjustments in response to oil or non-oil commodity price shocks, make a significant impact.

The most important factor in these countries is the inertial component. This means that anti-inflationary policy in developing countries with moderate or low inflation should focus on structural issues that affect the relationship between past and expected inflation. In countries with higher rates of mean inflation, such as many countries in South America, the importance of fiscal factors prevails and inertial inflation plays a much smaller role.

An article by Feldkircher and Tondl (2020) examined the relationship between inflation and monetary policy, taking into account the influence of global factors. For this purpose, the GVAR model was estimated and a generalized decomposition of the variance of inflation forecast error and short-term interest rates was carried out. The analysis shows that in the countries of Central, Eastern and South-Eastern Europe as well as in the emerging markets, global factors have a moderate effect on inflation. In developed countries, the dynamics of world prices affect inflation stronger, explaining 55 percent of inflation in the long run. Inflationary factors in all sample countries include rising oil prices as well as the exchange rate. In several emerging economies, monetary policy decisions are the main driver of inflation, especially in countries with inflation targeting policies. The regulation of the policy rate has its merits and maintains the price stability in the economy but globalization limits the ability of central banks. In addition, the results show that demand-pull factors are the least important determinant of inflation.

The study entitled “Forecasting Local Inflation with Global Inflation” examines a range of inflation forecasting models and argues that current economic models are poorly suited to inflation forecasts in emerging economies (Duncan & Martínez-García, 2015).

The literature on principal component analysis (PCA) for inflation analysis can be divided into two approaches. The first approach is related to the construction of alternative inflation indicators based on the rates of regional inflation or its individual components (services, industry, food, etc.). It is assumed that information about common factors is contained in the array of regional inflation rates; therefore, the main components are isolated from regional inflation. In the context of the regions of Kazakhstan, a similar study was carried out by employees of the National Bank, which confirmed the hypothesis that there is heterogeneity in the dynamics of inflationary processes across the regions of Kazakhstan (Tuleuov and Seidakhmetova, 2017). Despite the general dependence of inflation on the main factors, the degree and nature of how these factors influence inflation differs depending on individual characteristics of each region.

In their research, Marques et al. (1999, 2000, and 2001) and Abenoja et al. (2017) estimate inflation based on the principal component analysis and also use the first principal component as an indicator of the inflation trend in Portugal and the Philippines, respectively. The resulting new indicator has the desirable properties of a core inflation indicator: it is less volatile than both headline inflation and available core inflation indicators. The PCA-based proxy for core inflation has a lower standard deviation than most of core inflation indicators, even if historically volatile components of the CPI are not excluded from its estimate.

In using the second approach, statistics are used on a large number of time series of macroeconomic indicators common to all regions (growth rates of monetary aggregates, exchange rate changes, GDP growth rates, nominal interest rates, etc.) and the principal components are calculated from this data array. A similar methodology was also used in this study. This method has been previously proposed by Stock and Watson (1998) and Forni et al. (2000, 2004), which explore the properties of generalized dynamic factor models based on Sargent and Sims (1977) and Geweke (1977) models. In a series of papers, Stock and Watson (1998, 1999, and 2002) use factor models to transform a large amount of the US macroeconomic data into separate subsets of factors, and then use the resulting factors to predict future values of macroeconomic series such as GDP and inflation.

Stock and Watson, when comparing forecast mean squared errors – MSE – found out that this two-stage method produces forecasts that are superior to other univariate, bivariate, and multivariate reference models. This method is especially useful in forecasting inflation.

Forni et al. (2001) and Marcellino, Stock, Watson (2003) use factor models to analyze large data volumes of the European Region and Artis, Banerjee, Marcellino (2002) and Matheson (2006) – to forecast economic and financial variables of the United Kingdom and New Zealand.

3. The Data and Methodology Used

In the course of this study, changes in 65 variables to the previous period were used in monthly frequency from January 2011 to September 2021 (Attachment 1).

At the first stage, deseasonalization of all variables was performed. Further, in view of the use of principal component analysis as a tool for determining unobserved variables, all data were normalized and grouped according to the nature of their origin and their influence on inflationary processes.

The principal component analysis (PCA) is a data analysis tool that is commonly used to reduce the dimensionality (number of variables) of a large number of interrelated variables while retaining as much variation as possible (Kunovac, 2007). The PCA computes an uncorrelated set of variables (factors or principal components). These factors are ordered so that the first few retain most of the variance that is present in all of the original variables.

Suppose X is a vector of p random variables. The main idea of the principal component transformation is to find a few (<p) variables that have retained most of the variation given by the variance of p random variables. Let a random vector $\chi' = [X_1, X_2, \dots, X_p]$ have a covariance matrix with eigenvalues $\lambda_1 \geq \lambda_2 \dots \geq \lambda_p \geq 0$.

Let us assume that we have linear combinations $Y_j = \alpha_j X = \alpha_1 X + \alpha_2 X + \cdots + \alpha_p X$ of element X, where α_j – is a vector of p components $\alpha_{j1}, \alpha_{j2}, \dots, \alpha_{jp}$.

Therefore,

$$Var(Y_j) = \alpha_j' \Sigma \alpha_j \qquad \text{where } j=1,2,\dots,p \tag{1}$$

$$Var(Y_j, Y_k) = \alpha_j' \Sigma \alpha_k \qquad \text{where } j,k=1,2,\dots,p \tag{2}$$

Principal components are those uncorrelated linear combinations $Y_j, Y_j, \dots, Y_p$, with a maximum possible variance (1). When finding principal components we look up at their deviations. The first step is to find a linear combination $\alpha_j' X$ with a maximum variance so that:

$$\alpha_1' X = \alpha_{11} X_1 + \alpha_{12} X_2 + \cdots + \alpha_{1p} X_p = \sum_{k=1}^{p} \alpha_{1k} X_k \tag{3}$$

Then we find a linear combination $\alpha_2' X$, uncorrelated with $\alpha_1' X$ and having a maximum variance, and so on, so that on stage k a linear combination $\alpha_k' X$ is obtained, which has the maximum variance and the absence of correlation with $\alpha_1' X, \alpha_2' X, \dots, \alpha_{k-1}' X$. The resulting k variable $\alpha_k' X$ – is a k-th principal component, where p of principal components were found before this component but it is enough to stop after the q-th stage (q ≤ p), where the largest part of variance X was included in q principal components.

- The variance of principal component is equal to its eigenvalue that corresponds to this principal component,

  $$Var(Y_j) = \alpha_j' \Sigma \alpha_j = \lambda_j \qquad \text{where } j=1,2,\dots,p \tag{4}$$

- The total variance of the data is equal to the total variance of its principal components,

  $$\sigma_{11} + \sigma_{22} + \cdots + \sigma_{pp} = \sum_{j=1}^{p} Var(X_j) = \lambda_1 + \lambda_2 + \cdots + \lambda_p = \sum_{j=1}^{p} Var(Y_j) \tag{5}$$

Initially, the data was normalized so that the variables had the same scale, using the standard method of normalizing all data to zero mean and single standard deviation. A random vector $\chi' = [X_1, X_2, \dots, X_p]$ was transformed into corresponding normalized variables

$$Z = \left[Z_j = \frac{X_j - \mu_j}{\sqrt{\sigma_{jj}}}\right] \quad \text{where j=1,2,…,p} \tag{6}$$

In a matrix notation, it can be written as follows:

$$Z = \left(V^{1/2}\right)^{-1}(X - \mu) \tag{7}$$

where $V^{1/2}$ – is a diagonal matrix of standard deviation. Therefore, $Z(E) = 0, Cov(Z) = \rho$.

Principal components Z may be obtained from eigenvectors of a correlated matrix X. All previous properties for X apply to Z as well, so the notation Y_j refers to the j-th principal component, and the values (λ_j, α_j) are a pair of eigenvalue vectors.

j-th principal component of normalized variable $Z' = [Z_1, Z_2, \dots, Z_p]$ with $Cov(Z) = \rho$, is written as:

$$Y_j = \left(V^{1/2}\right)^{-1}(X - \mu), \text{ so that:}$$

$$\sum_{j-1}^{p} Var(Y_j) = \sum_{j-1}^{p} Var(Z_j) = P \quad \text{where j=1, 2,…,p} \tag{8}$$

In this case, $(\lambda_1, \alpha_1), (\lambda_2, \alpha_2), \dots, (\lambda_p, \alpha_p)$ – is a pair of eigenvalue vectors with a condition for p $\lambda_1 \geq \lambda_2 \dots \geq \lambda_p \geq 0$.

The principal components are determined as follows. The load matrix or eigenvectors is a measure of importance of the estimated variable for a given principal component. Assuming all elements are positive, the first component is the weighted average of the variables and is called the inflation trend. Similarly, positive and negative coefficients in the subsequent components can be considered as inflation factors.

The weight of the variance is determined by the principal component that best explains the original variables. A measure of how well the first q principal components explain the variation in Z is presented as:

$$\varpi_q = \frac{\sum_{j=1}^{p} \lambda_j}{P} = \frac{\sum_{j=1}^{p} Var(Z_j)}{P}. \tag{9}$$

After finding the main factors explaining the variation in inflation with the help of PCA, further assessment of the contribution of these factors is carried out.

The method of least squares is used as a tool for determining the degree of influence and contribution of each unobserved variable. This method is used to find the contributions of the obtained q significant factors to the variable of interest Y. Explicitly, a linear dependence on q significant factors $\chi' = [X_1, X_2, \dots, X_q]$ is measured for y_t, and the following equation is estimated by using the OLS method:

$$y_t = \alpha_t + \sum_{i=1}^{q} \beta_i X_{i,t} + \varepsilon_t \quad \text{where } i = 1,2,\dots,q, \tag{10}$$

where the resulting $\hat{\beta}_i$ are the suggested contributions per a unit of change in the corresponding factor X_i. One can find a full explanation of the underlying least squares method in the book by Linnik (1962).

The basics of vector autoregression and subsequent variance decomposition are described in the work of Stock and Watson (2005).

Vector autoregression models have the following form:

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t, \tag{11}$$

where $y_t = (y_{1t}, \dots, y_{Kt})'$ (a prime denotes transposition) – is a vector of K observable variables of interest, A_i' – are the matrices of *(K×K)* parameters, and p – is the lag order. In VAR model (11), all variables are a priori endogenous. The variance decomposition helps interpreting VAR models. It is presented as follows.

The forecast for h forward steps may be obtained from (1) recursively in the form of:

$$y_{t+h} - y_{t+h|t} = A_1 y_{t+h-1|t} + \dots + A_p y_{t-p} + u_t \tag{12}$$

Here, $y_{t+j|t} = y_{t+j}$ for $j \leq 0$. Therefore, the forecast error equals

$$y_{t+h} - y_{t+h|t} = u_{t+h} + \sum_{i=1}^{h-1} \Phi\, u_{t+h-i} \sim \left(0, \Sigma_h = \Sigma_u + \sum_{i=1}^{h-1} \Phi_i\, \Sigma_u \Phi_i'\right), \tag{13}$$

on condition that forecast errors have a zero mean and covariance matrix equal to Σ_h.

The variance of forecast error of k-th forecast element looks as follows:

$$\sigma_{\mathrm{k}}^2(\mathrm{h}) = \sum_{\mathrm{j}=0}^{\mathrm{h}-1} (\theta_{\mathrm{k1,j}}^2 + \dots + \theta_{\mathrm{kK,j}}^2) = \sum_{\mathrm{j}=0}^{\mathrm{h}-1} (\theta_{\mathrm{kj,0}}^2 + \dots + \theta_{\mathrm{kj,h-1}}^2) \tag{14}$$

The term $(\theta_{\mathrm{kj,0}}^2 + \dots + \theta_{\mathrm{kj,h-1}}^2)$ may be interpreted as a contribution of the j-th shock at the h-th step to the variance of the forecast error of variable k. Decomposition of the term by $\sigma_{\mathrm{k}}^2(\mathrm{h})$ gives an interest rate contribution of shock j to the variance of the forecast error of variable k at step h. Such approach was suggested by Sims (1980) and has been frequently used and interpreted for various forecast horizons.

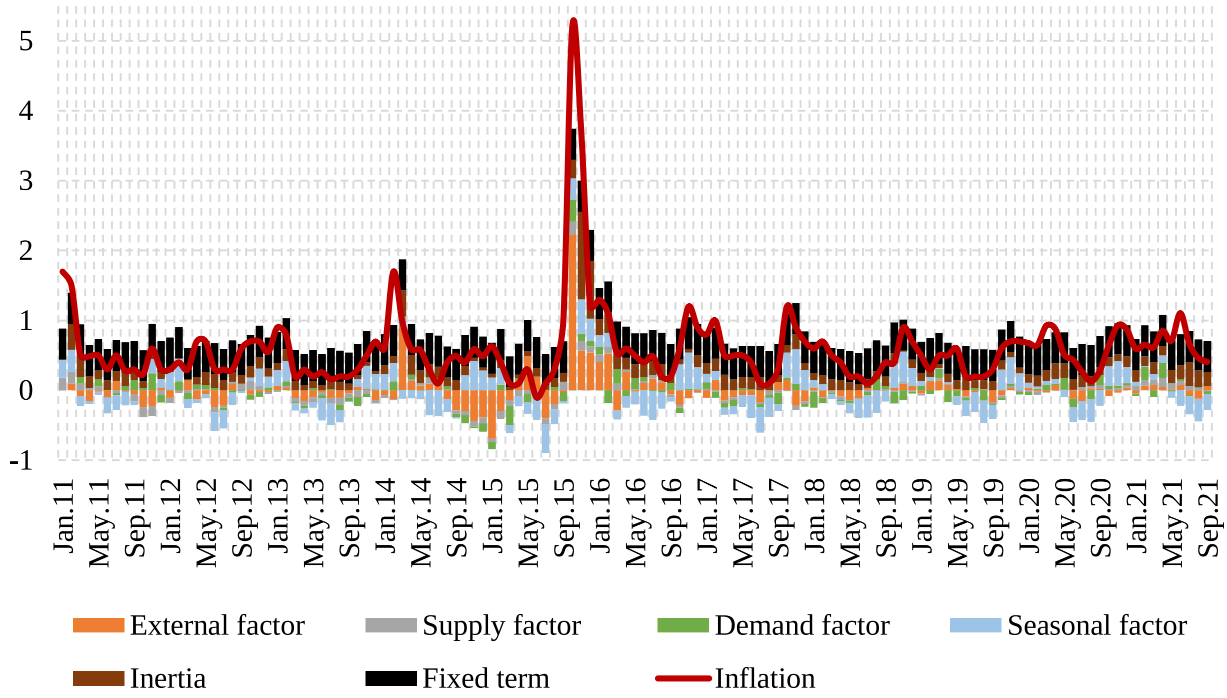
4. Discussion of Outcomes

As part of the study, three unobservable variables were obtained using the principal components analysis: the external factor, the supply and demand-pull factor. At the same time, the demand-pull factor also takes into account the influence of monetary and fiscal policies. The observed variables used to derive the factors were selected based on the principle of maximum correlation with the dependent variable, inflation in this case.

As can be seen from Diagram 1, a significant contribution to inflation in periods without obvious shocks is made by the fixed term, inertia and the seasonal factor. During certain periods of significant acceleration of inflation, namely in February 2014 and October 2015, the main contribution was made by the external factor. Indeed, inflation in these periods accelerated due to the growth of import prices in the context of the tenge exchange rate pass-through onto domestic prices.

Diagram 1

Contributions to the monthly inflation by unobserved variables obtained with the use of the principal component analysis



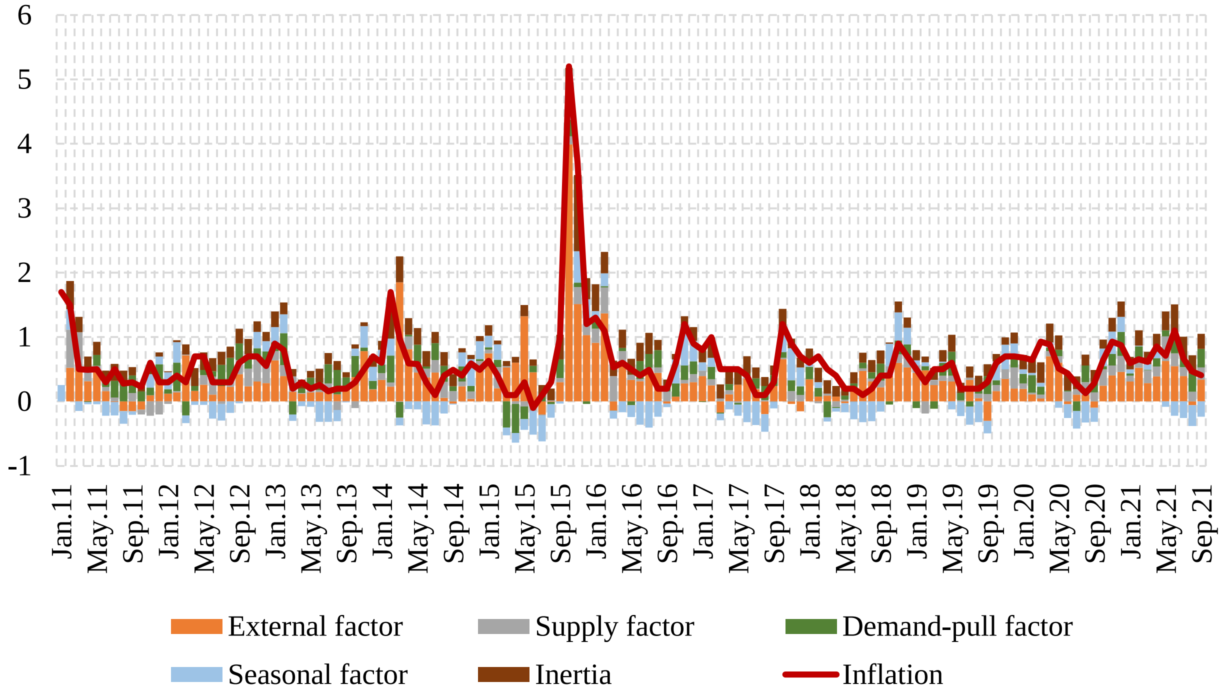
Source: UN FAO, National Bank, BNS ASPR, Central Bank of Russia, authors' computations

The influence of other unobserved variables throughout the reviewed period is moderate. At the same time, analyzing the period from the onset of the pandemic, one may note that the influence of factors, supply and demand has increased compared to other periods. However, it is worth noting that, as shown by this study, the use of the principal component analysis as a tool for identifying and obtaining unobserved variables is not as effective as it might seem at first glance. This is because many observed variables have a divergent impact and at different intervals. To eliminate this shortcoming, a second OLS model was built and the observed variables were directly included into the model and factors were broken down further according to the nature of origin.

The results of this model are somewhat different from the results of the model obtained using the PCA. There is no fixed term in the model that uses the grouping of variables due to the weak significance of coefficients to this variable (Diagram 2).

Diagram 2

Contributions to the monthly inflation by unobserved variables obtained with the use of the variable grouping



Source: UN FAO, National Bank, BNS ASPR, Central Bank of Russia, authors' computations

The main contribution to inflationary processes during the analyzed period is made by an external factor, which is associated with a high share of both finished and intermediate imports. Thus, according to the Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan, in 2020 the share of imports in the consumption of food was 22.1%, and of non-food products - 60.3%. It should be noted that since 2011 the share of imports in the consumer basket has remained the same: about 60% for non-food products and 20% for food products. Most of the contribution by external factors is associated with the effect of imported inflation from the Russian Federation as the main trading partner. This variable has a maximum coefficient and a strong significance in the equation (Attachment 3), which is explained by close economic ties and a high share of Russia in total imports – 41.1% in January-September 2021.

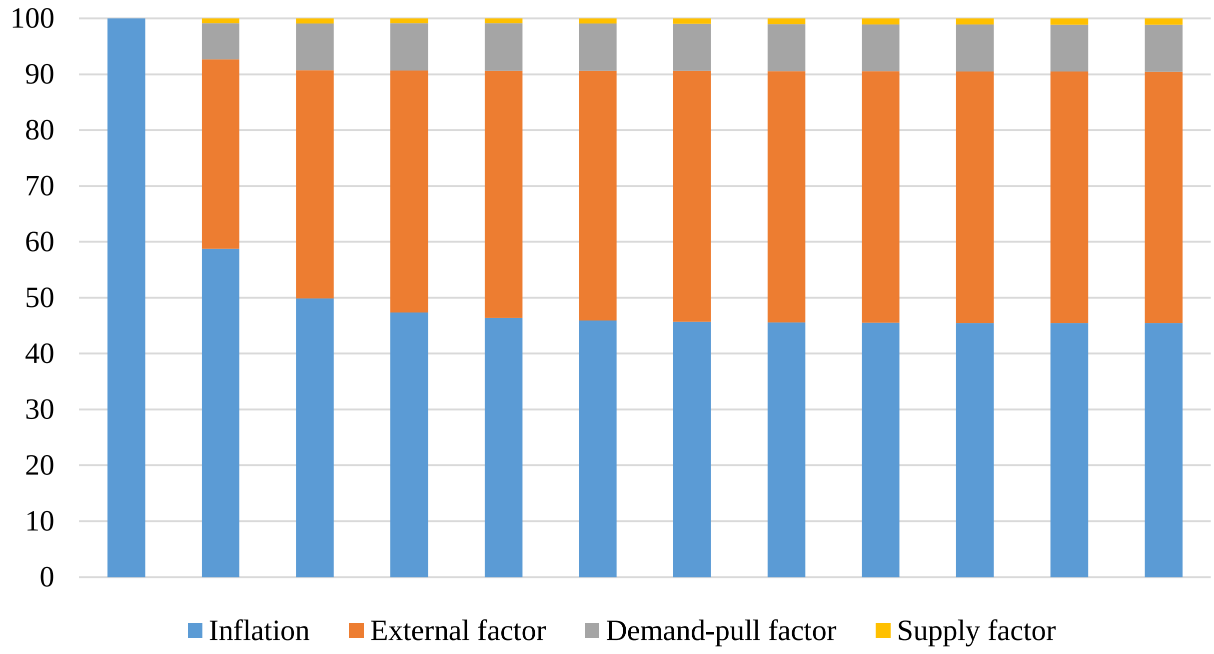
In addition to the external factor, a significant contribution is made by the seasonal factor and inertia factor. A significant contribution of inertia is associated with the heterogeneity of inflationary processes. The magnitude and timing of the price change in response to various shocks vary and depend on the particular region as well as the particular retail outlet. In addition, it is worth mentioning that a significant contribution of the seasonal factor and inertia may indirectly indicate the presence of structural constraints and imbalances in the economy.

In the current variation of the model, in contrast to the model that uses the factors obtained using PCA, there is a more significant contribution of supply and demand factors. In particular, these factors have been contributing greatly since the beginning of the pandemic.

In addition to the two previously described models, the authors of this study built a VAR model in order to estimate the "variance decomposition" to interpret the relationships between variables described by the model (Diagram 3).

Diagram 3

Contributions to the monthly inflation by unobserved variables obtained with the use of the variable grouping



Source: UN FAO, National Bank, BNS ASPR, Central Bank of Russia, authors' computations

This tool "decomposes" the variance of the forecast error into contributions from specific exogenous shocks, and also demonstrates how important a particular shock is for explaining variations in variables over time.

According to the variance decomposition, in the first month, about 100% of the inflation variation is due to inflation shocks. In subsequent months, the contribution of inflation decreases and the contribution of factors to inflation changes increases rather quickly. The system becomes stable in about 6 months after the shocks and the contribution by factors remains at 55%. At the same time, most of the variance (about 44%) is explained by external shocks, which again confirms the strong dependence of the domestic price dynamics on the world market environment.

5. Findings and Recommendations for Future Research

In this study, non-observable factors were determined and a quantitative assessment of their impact on inflationary processes was conducted.

The principal components analysis was used as a tool for obtaining unobserved variables. Next, the least squares (OLS) method was used to determine the degree of influence and contribution of each unobserved variable. Besides, the authors built the second OLS model with direct inclusion of the observed variables into the model and a further breakdown into factors according to the nature of origin. Apart from that, in addition to the two previously described models, the authors of this study built a VAR model in order to estimate the "variance decomposition" for interpreting the relationships between variables described by the model.

According to the results of the study, the main contribution to the dynamics of inflationary processes is made by such factors as the external factor, which is associated with a high share of imports, as well as the seasonal factor and the inertia that indirectly reflects the presence of structural imbalances and restrictions. It should be noted that a significant contribution of inertia according to the study of the International Monetary Fund (Loungani & Swagel, 2001) is typical for many developing countries. To eradicate this problem, anti-inflationary policy should focus on structural issues that affect the relationship of expectations between the past and future inflation.

At the same time, since the onset of the pandemic, two variations of the model have demonstrated the presence of a permanent positive contribution on the demand and supply side.

Summing up, it is worth noting a strong dependence of inflationary processes on external factors, seasonality and inertia. This fact confirms the presence of deep structural problems and insufficient domestic production. At the same time, artificial curbing of inflationary processes through administrative intervention by setting marginal prices and regulating markets only exacerbates the existing imbalances in the markets in the long run.

The existing problems can be addressed by deregulation of markets; indirect support for domestic producers by creating favorable and transparent conditions for doing business; introduction of a countercyclical fiscal rule and synchronization of monetary and fiscal policies.

6. List of References

1. James H. Stock and Mark W. Watson, Forecasting Using Principal Components From a Large Number of Predictors, *Journal of the American Statistical Association*, December 2002;
2. Mr. Prakash Loungani and Mr. Phillip L Swagel, Sources of Inflation in Developing Countries, *IMF Working Papers, 2001;*
3. Feldkircher, M., and Tondl, G., Global Factors Driving Inflation and Monetary Policy: A Global VAR Assessment, *International Advances in Economic Research*, 2020;
4. Duncan, R., and Martínez-García, E., Forecasting local inflation with global inflation: When economic theory meets the facts, *Globalization and Monetary Policy Institute Working Paper*, 2015;
5. Marques, C. R., and Mota, J. M., Using the asymmetric trimmed mean as a core inflation indicator, *Banco de Portugal, Economics Research Department,* 2000;
6. Machado, J. F., Marques, C. R., Neves, P. D., da Silva, A. G., Using the first principal component as a core inflation indicator, *Banco De Portugal*, 2001;
7. Marques, C. R., Neves, P. D., da Silva, A. G., Why should Central Banks avoid the use of the underlying inflation indicator, *Economics Letters*, 2002;
8. Zeno Ronald R. Abenoja, Joselito R. Basilio, Cherrie F. Ramos, and Joan Christine S. Allon., An Alternative Core Inflation Measure for the Philippines using Principal Components Analysis, *Bangko Sentral ng Pilipinas,* 2017;
9. Forni, M., Hallin, M., Lippi, M., & Reichlin, L., The generalized dynamic-factor model: Identification and estimation, *Review of Economics and statistics*, 2000;
10. Forni, M., Hallin, M., Lippi, M. and Reichlin, L., The generalized dynamic factor model consistency and rates, *Journal of Econometrics*, 2004;
11. Stock, J. H., and Watson, M. W., A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series, 1998;
12. Sargent, T. J., and Sims, C. A., Business cycle modeling without pretending to have too much a priori economic theory. New methods in business cycle research, 1977;
13. Geweke, J., The dynamic factor analysis of economic time series. Latent variables in socio-economic models, 1977;
14. Chan, Y. L., Stock, J. H., and Watson, M. W., A dynamic factor model framework for forecast combination, *Spanish Economic Review*, 1999;
15. Forni, M., and Lippi, M., The generalized dynamic factor model: representation theory, *Econometric theory*, 2001;
16. Marcellino, M., Stock, J. H., and Watson, M. W., Macroeconomic forecasting in the euro area: Country specific versus area-wide information, *European Economic Review*, 2003;
17. Artis, M. J., Banerjee, A., and Marcellino M., Factor forecasts for the UK. *Journal of forecasting*, 2005;
18. Matheson, T. D., Factor model forecasts for New Zealand, 2006;
19. Kunovac, D., Factor model forecasting of Inflation in Croatia, *Financial theory and practice*, 2007;

20. Stock, J. H., and Watson, M. W., Implications of dynamic factor models for VAR analysis, 2005;
21. Sims, C. A., Macroeconomics and reality, *Econometrica: Journal Of The Econometric Society*, 1980;
22. Linnik Yu.V., Least Squares Method and the Fundamentals of the Theory of Analysis of Observations, *Fizmatgiz,* 1962;
23. Tuleuov, O. and Seidakhmetova, B., Inflationary Processes in the Regions of Kazakhstan: Analyzing Inhomogeneity of Inflationary Factors and the Disaggregated Inflation Forecasting Model based on the BVAR-Technique. Inflation in the Regions of Kazakhstan, 2017.

7. Attachments

Attachment 1

List of Indicators Used

#	**Indicator**
1	FAO food price index
2	FAO meat price index
3	FAO dairy price index
4	FAO cereal price index
5	FAO vegetable oil price index
6	FAO sugar price index
7	USD/KZT exchange rate
8	RUB/KZT exchange rate
9	Import price index
10	Consumer import price index
11	Food consumer import price index
12	Non-food consumer import price index
13	Consumer price index in Russia
14	Food consumer price index in Russia
15	Non-food consumer price index in Russia
16	Services price index in Russia
17	Average monthly Brent oil price
18	Consumer price index in the EU
19	Consumer price index in China
20	Consumer price index in the USA
21	Personal consumption expenditure index in the USA
22	Quantity of tenge in the economy
23	Money base (reserve money)
24	Narrow money
25	M0 (cash in circulation)
26	M1
27	M2
28	M3 (money supply)
29	Average monthly TONIA
30	Volume of consumer loans
31	Volume of loans to food producers
32	Volume of loans to agricultural producers
33	Volume of loans to the transport sector
34	Volume of loans to the trade sector
35	Consumer product producer price index

36	Food producer price index
37	Beverage producer price index
38	Clothes producer price index
39	Medication producer price index
40	Car producer price index
41	Furniture producer price index
42	Electricity producer price index
43	Gasoline producer price index
44	Diesel producer price index
45	Producer price index in the manufacturing industry
46	Producer price index
47	Producer price index in agriculture
48	Producer price index in plant production
49	Producer price index in animal production
50	Producer price index in the transport sector
51	Physical output index of the manufacturing industry
52	Physical output index in agriculture
53	Turnover in retail trade
54	Index of nominal income
55	Index of real income
56	Real wage index
57	Nominal wage index
58	The numbers of economically active population
59	Budget spending on purchases of goods and services
60	Budget spending on wages
61	Budget spending on individuals
62	Consumer price index
63	Food consumer price index
64	Non-food consumer price index
65	Services price index

Attachment 2

Model Results with the PCA

Sample (adjusted): 2011M11 2021M09
Included observations: 119 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CPI_SA(-1)	0.257043	0.061815	4.158255	0.0001
FACTOR_EXTERNAL(-1)	0.218053	0.027155	8.029799	0.0000
FACTOR_MONETARY(-2)	0.048499	0.022591	2.146823	0.0340
FACTOR_SUPPLY	0.037629	0.022539	1.669525	0.0978
FACTOR_FISCAL(-10)	0.048686	0.026804	1.816360	0.0720
FACTOR_DEMAND(-2)	0.041778	0.024086	1.734519	0.0856
C	0.445164	0.046401	9.593907	0.0000

R-squared	0.681860	Mean dependent var	0.595212
Adjusted R-squared	0.664816	S.D. dependent var	0.525566
S.E. of regression	0.304277	Akaike info criterion	0.515263
Sum squared resid	10.36944	Schwarz criterion	0.678741
Log likelihood	-23.65816	Hannan-Quinn criter.	0.581646
F-statistic	40.00763	Durbin-Watson stat	1.595977
Prob(F-statistic)	0.000000		

Attachment 3

Model Results with the Grouping

Sample (adjusted): 2011M06 2021M09
Included observations: 124 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CPI(-1)	0.367634	0.057527	6.390623	0.0000
D_REAL_INCOME(-2)	0.020297	0.007832	2.591465	0.0108
M_CREDIT_CONSUMER(-3)	0.044291	0.018163	2.438553	0.0163
E_CPI_RU	0.531251	0.067453	7.875904	0.0000
E_IMPORT_CONS_PRICES(-1)	0.040869	0.017994	2.271194	0.0250
E_RUB_KZT(-1)	0.069355	0.010117	6.855165	0.0000
E_USD_KZT(-1)	0.051872	0.011597	4.472701	0.0000
S_PPI_CONSUMER_GOODS(-1)	0.133821	0.051278	2.609722	0.0103
S_PPI_AGRICULTURE(-1)	0.032066	0.023743	1.350541	0.1795
M_MONEY_SUPPLY(-5)	0.024980	0.013434	1.859486	0.0655
M_M2X(-4)	0.026571	0.009562	2.778680	0.0064

R-squared	0.714045	Mean dependent var	0.581831
Adjusted R-squared	0.688740	S.D. dependent var	0.588888
S.E. of regression	0.328545	Akaike info criterion	0.696237
Sum squared resid	12.19739	Schwarz criterion	0.946423
Log likelihood	-32.16668	Hannan-Quinn criter.	0.797868
Durbin-Watson stat	1.744188		