Bank of Russia

LABOUR MARKET AND INFLATION RELATIONSHIP INDICATOR

WORKING PAPER SERIES

No. 96 / September 2022

Dmitry Orlov and Evgeny Postnikov

Dmitry Orlov

Bank of Russia, Monetary Policy Department

E-mail: dmitryorlo888@gmail.ru

Evgeny Postnikov

Bank of Russia, Ural Main Branch, Chelyabinsk Division

E-mail: postnikovea@yandex.ru

The authors are grateful to Henry Penikas for his helpful comments and suggestions.

The Bank of Russia's Working Paper Series is subject to blind review by members Bank of Russia Advisory Board members and external reviewers.

Cover photo: national park "Taganay" (Evgeny Postnikov, Bank of Russia)

Address: 12 Neglinnaya Street Moscow 107016

Telephone: +7 495 771-91-00, +7 495 621-64-65 (fax)

Bank of Russia website: www.cbr.ru

Contents

ABSTRACT

The labour market is closely connected to inflation processes, and is therefore a key factor to consider in monetary policy decisions. Russian regions differ substantially in terms of employment, wages, migration flows and the age structure of their population. Therefore, the effects of regional changes in the labour market on prices may be different. Since the Central bank's inflation targeting policy is pursued nationwide, it is important for a regulator to factor in regional heterogeneity when assessing the impact of changes in the labour market on inflation growth.

This paper brings forward a composite indicator of the contribution of labour market changes to inflation increase – the Labour Market Indicator (LMI). To capture regional heterogeneity in terms of market labour indicators, regions are grouped into four clusters with different social, demographic and economic characteristics. We make the case that the impact of unemployment on inflation can be described as slight or moderate in Russia. The calculated quarterly LMI values are overall consistent with the actual effect of the labour market on inflation processes over the entire time horizon under study, which suggests that the estimates are reliable. The important benefit of the LMI is that it is possible to interpret and allows to assess the future impact of labour market on inflation one quarter ahead of available statistical data – which helps make better informed monetary policy decisions.

Key words: impact of the labour market on inflation, regional heterogeneity, clustering, principal component analysis, unemployment, wages, regression analysis.

JEL-classification: C32, C38, E24, E31.

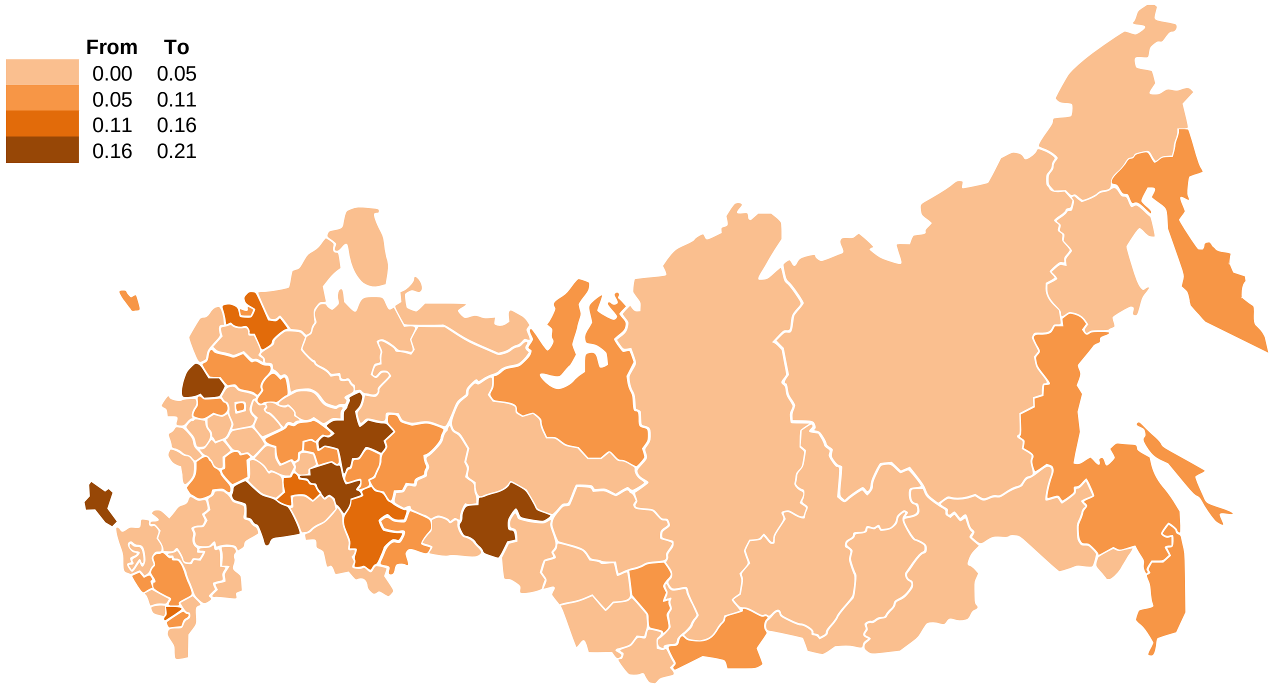
1. Introduction

The indicators of movements in labour market condition are of particular interest. There was a surge in unemployment over a short period of time in 2020–2021 which was followed by a similarly rapid drop. Coronavirus-linked restrictions drove a significant reduction in labour migration flows. The resulting labour shortages made the wage growth uneven in certain industries. The unbalanced labour markets had a strong impact on inflation processes, and it is likely to be a long time before labour flows return to normal.

The Russian Government's and regional authorities' anti-pandemic response was aimed at supporting employment and household incomes. Among such measures were increased amounts of unemployment benefits, the promotion of self-employment, temporary public works and the employment of graduates, young people and people with disabilities, as well as arrangements to expand employment opportunities through remote interactions between employers or employment agencies and applicants, and one-off payments to some categories of citizens.

Inflationary pressure from the labour market is driven by various interrelated factors, such as unemployment and wage movements, the balance of labour supply and demand, labour productivity and skills. The recent considerable redistribution of labour demand and supply (especially in low-skilled industries) makes it more complicated to assess the impact of social and demographic processes on the national economy. It should be noted that, on the one hand, the impact of the labour market on inflation is heterogeneous across regions (Figure 1), and on the other hand, the impact of a uniform monetary policy on regions may also be heterogeneous. Specifically, regions with the strongest effect of the labour market on inflation include the Kirov, Saratov, Smolensk, Tyumen Regions, and the Republic of Tatarstan. The group with a slightly weaker impact includes the Leningrad and Ulyanovsk Regions, and the Republics of Bashkortostan and Ingushetia. Importantly, regions within one cluster may differ by unemployment, migration flows and wages. This is why it is important to take into account the regional heterogeneity as the structural differences largely determine the sensitivity of inflation processes to changes in the labour market.

Figure 1. Inflation elasticity by unemployment gap in Russian regions, 2011–2020, %



Source: Authors' calculations.

However, the basic tool to assess the labour market and inflation relationship is the Phillips curve, a relationship between the output gap and the unemployment gap. However, the Russian labour market is known to adjust to economic fluctuations more through the changes in the price of labour (alignment of wages and bonuses) rather than through the changes in employment. This reduces the ambit of the Phillips curve and calls for a more nuanced approach.

This is why it becomes necessary to find an indicator that would holistically cover all various labour market measures and help quantify the impact of key indicators of the labour market on inflation. Such composite indicator would measure the aggregate impact of the labour market on inflation subject to regional heterogeneity, and may enhance the quality of economic analysis to enable better informed monetary policy decisions.

This paper proposes an indicator describing the inflationary or disinflationary influence of the labour market since 2014, a year that marked the beginning of inflation targeting in Russia. The results are consistent with actual relationships observed throughout the period under study, and therefore provide valuable insights at this moment in time, when labour markets are unbalanced.

The structure of this paper is as follows. Section 2 provides an overview of indicators used to analyse the labour market in foreign and domestic practice, as well as the literature exploring the impact of the labour market on price growth in Russia. Section 3 describes the methodology for calculating the labour market indicator and its step-by-step procedure. The impact of the labour market on inflation processes based on the obtained indicator calculations is analysed in Section 4. The conclusions offer a brief summary of the findings.

2. Review of the literature on labour market indicators

Considering that it is a challenging task to compare social and economic processes against individual indicators, it is necessary to calculate indicators that would simultaneously combine several measurements. Such indicators may differ significantly from each other in terms of value, units of measure and contents. Various indices were also developed to analyse the labour market at different times and for different purposes.

For example, in 2014–2017, the US Fed calculated the Labour Market Conditions Index (LMCI). It was based on a dynamic factor model of 19 US labour market indicators covering such broad categories as unemployment, underemployment, employment, workweeks, wages, vacancies, hiring, firing, resigned employees, well as surveys of consumers' and businesses' perceptions (Chung et al., 2014). The LMCI was intended to support the Fed's decision-making in the context of its key goal of maximum employment. However, given the high correlation of the LMCI with the unemployment rate and the absence of a significant link between labour market conditions and wage growth, the Fed decided to stop updating the index in August 2017.

Gallup, a US analytics and advisory company, brought forward its own standard for measuring employment based on annually updated and cross-country comparable indicators statistically related to per capita GDP (Clifton, Marlar, 2011). *The good job employment indicator* (a full-time job for, at least, 30 hours a week) includes indicators such as full-time employment and self-employment, part-time employees willing to work full-time, and unemployment.

The Ivanov Consumer Index, or Sberbank's Consumer Confidence Index,[1] calculated since 2013, reflects current needs and expectations of middle-income Russians, taking into account consumer spending, savings, consumer behaviour and the overall level of consumer confidence.

[1] SberCIB Investment Research Portal: [http://research.sberbank-cib.com]

Also, joint efforts of the Russian Union of Industrialists and Entrepreneurs, PRIME Economic Information Agency, HeadHunter Group and IBS Group, resulted in 2013 in a prototype index to analyse the Moscow labour market from a macro-business perspective, covering multiple industries.[2] However, this index saw no further development and was never put into use.

The labour market index for Russian regions, annually calculated by the RIA[3], aggregates eight indicators capturing employment, wages, labour market conditions and capacity. It shows the attractiveness of a particular region as a potential employer.

However, what central banks are primarily interested in is the relationship between the labour market and inflation, with existing labour market indices mostly uncovering certain areas without a link with inflation processes. Also, there are a number of studies focused on the impact of certain labour market indicators on price growth in Russia. For instance, the neo-Keynesian Phillips curve provides deep insights into the impact of the unemployment gap on inflation, and this area has been studied by Russian researchers (Bragin, Osakovsky, 2004; Gafarov, 2011; Sokolova, 2014; Orlov, Postnikov, 2020; etc.). In this area, the most interesting studies of foreign researchers are papers seeking to measure the NAIRU and the unemployment gap based on country data (Gordon, 2013; Rusticelli, 2015), as well as to estimate jointly the Phillips curve and Okun's law, providing a rationale for unemployment and output gaps (Chow, 2011). A decline in unemployment below its non-accelerating inflation rate (NAIRU) marks a positive output gap (that is, actual output exceeds the potential one), which makes prices rise under the pressure of demand in core markets. Conversely, a positive unemployment gap is associated with recession, i.e. a decrease in output below its potential level and a drop in prices driven by weak demand.

In terms of the cost-plus inflation model, wage growth triggers a rise in input costs, and consequently, product costs. This relationship also works in reverse, that is, price growth spells the need to raise wages (demand-pull inflation). Advanced economies (e.g. the US) tend to exhibit a one-way connection under the demand-pull inflation (Gurvich, Vakulenko, 2018). Russia is characterised by a two-way relationship between inflation and wages; however, as the Bank of Russia has been applying inflation targeting since 2014,

---

[2] Indeks rynka truda: kratkoe opisanie metodiki [Labour market index: short methodology description]. An RSPP, AEI PRIME, Headhunter Group and IBS Group project [in Russian]. 2013. 20 c. [https://1prime.ru/files/pdf/methodology.pdf]
[3] The labour market index for Russian regions / RIA Rating website: [https://riarating.ru/infografika/20210906/630207557.html]

while changing the inflationary mechanisms, it facilitates a switch to the demand-pull inflation model, among others (Ivanova, 2016).

Domestic academic literature lacks any insights into what is the relationship between migration and inflation. The first attempt in this research area is the paper by Kudaeva and Redozubov (2021). The authors conclude that the statistically significant impact of migration is seen only in several economic and labour market indicators: GDP and wages. However, global studies have provided evidence that migration processes do not have a noticeable effect on inflation. This follows, for example, from studies by the Central Bank of Norway (Furlanetto and Robstad, 2016), the Reserve Bank of New Zealand (Smith and Thoenissen, 2018), as well as analysis of migration in Canada (Dungan, Fang and Gunderson, 2012) and in the US (Weiske, 2019). Specifically, the migration ratio in Norway, although far above that in Russia, does not have any significant impact on inflation processes. This further suggests that the impact of migration on inflation in Russia is negligible.

Annex 1 offers a brief review of the literature on the impact of labour market indicators on inflation.

Therefore, the labour market influences inflation through a number of indicators. The nature of this influence varies across Russian regions. We propose the composite indicator – LMI – to capture the labour market's aggregate effect on inflation.

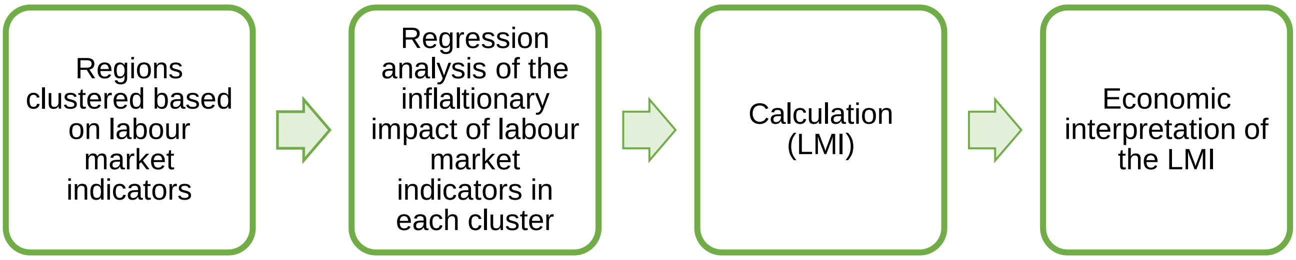
3. Constructing the labour market indicator

3.1. Methodology

Given the substantial heterogeneity of regional labour markets, some regions are marked by a weak relationship between inflation and their labour markets, and others are marked by a stronger relationship. Consequently, the response of the labour market to monetary policy varies across regions. This is why calculating the proposed indicator involves cluster analysis in the first stage, which is intended to identify homogeneous groups of regions (Figure 2). Clustering regions by regional labour market will bring us groups of territories with similar characteristics and a similar impact on inflation processes; we will also be able to rectify the problem of observations having superfluous highly correlated features and reduce the number of regressions from 85 (consistent with the number of regions) to 4 (consistent with the number of regional clusters). While doing so, we will be able to keep 95% of explanation for the variance in all indicators.

An inflation model is thereafter built for each cluster subject to labour market indicators, to identify their contribution to inflation trends of the regional cluster.

Figure 2. LMI methodology in brief



The labour market indicator is then calculated as a weighted average value of unemployment and wage based on each cluster's contribution to inflation. Finally, a detailed economic description of this indicator is available for the current quarter.

3.2. Regions clustered based on labour market indicators

In order to come to an appropriate breakdown of regions into clusters, it is necessary to correctly identify determinant indicators to be used for classifying and structuring territories. With our objective being analysis of the labour market and its impact on monetary policy, we identify the following key social and economic indicators for regional labour markets:

1. Unemployment rate, %
2. Labour shortage rate reported to state employment agencies, people/thousand people
3. Labour force participation rate, %
4. Migration growth rate per 10,000 people
5. Natural growth rate per 1,000 people
6. Under working age population, % of the total population
7. Working age population, % of the total population
8. Above working age population, % of the total population
9. Average monthly nominal wage, rubles
10. Graduates of higher education institutions, % of the total population.

We took Rosstat statistics on all 85 Russian regions and 10 indicators as of the end of 2019.

All cluster analysis stages were performed with the use of Python 3. In addition to standard programming tasks (*matplotlib, pandas, numpy, seaborn*) we used *scikit-learn,* one of the most popular and accessible open-source machine learning libraries.[4]

For the computational algorithms of clustering to take on a commensurable view and ensure their correct operation, the initial data were normalised by the root of sum- of-squares (Euclidean norm). This method scales each data point so that the feature vector has a Euclidean length.

The correlation analysis of the ten labour market characteristics revealed closely related indicators (Chart 3).

Specifically, there was a high linear correlation between:

- the natural growth rate, the share of under working age population and the share of above working age population;
- the average monthly nominal wage, labour force participation rate and the share of working-age population.

A moderate linear correlation was found between:

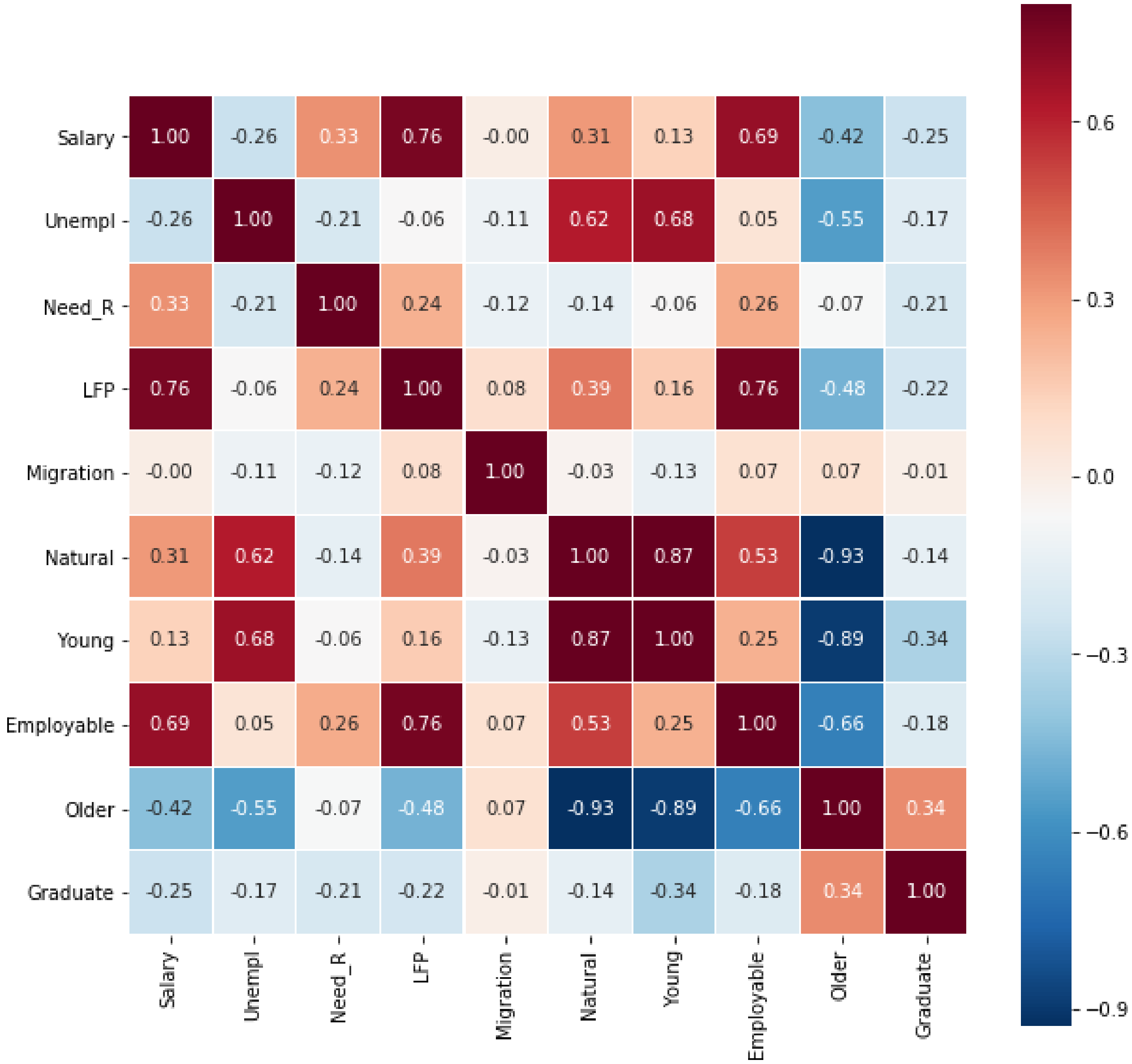
- the unemployment rate and the natural growth rate, the share of under working age population and the share of above working age population;
- the share of above working age population and the average monthly nominal wage, the unemployment rate, labour force participation rates, and the working age population;
- the natural growth rate and the share of working age population.

Since the features are greatly interconnected, their simultaneous presence is excessive. When the problem of redundant highly correlated features in observations was eliminated, it was possible to solve the problem of reducing the dimensionality of data through the principal component method.

In practice, researchers choose as many principal components as it is necessary to keep 90% of the variance of initial data. The calculation of explained variation for each component shows that the identification of three principal components (that is, reduction in dimensionality from 10 to 3) will explain more than 90% of the variation of all indicators (Figure 4). Yet, for more accurate results, we conducted cluster analysis for three, four and five principal components, as well as for all the ten labour market indicators without a reduction in dimensionality.

[4] Scikit-learn. Python machine learning. [https://scikit-learn.org/stable/index.html]

Figure 3. Correlation matrix of labour market indicators

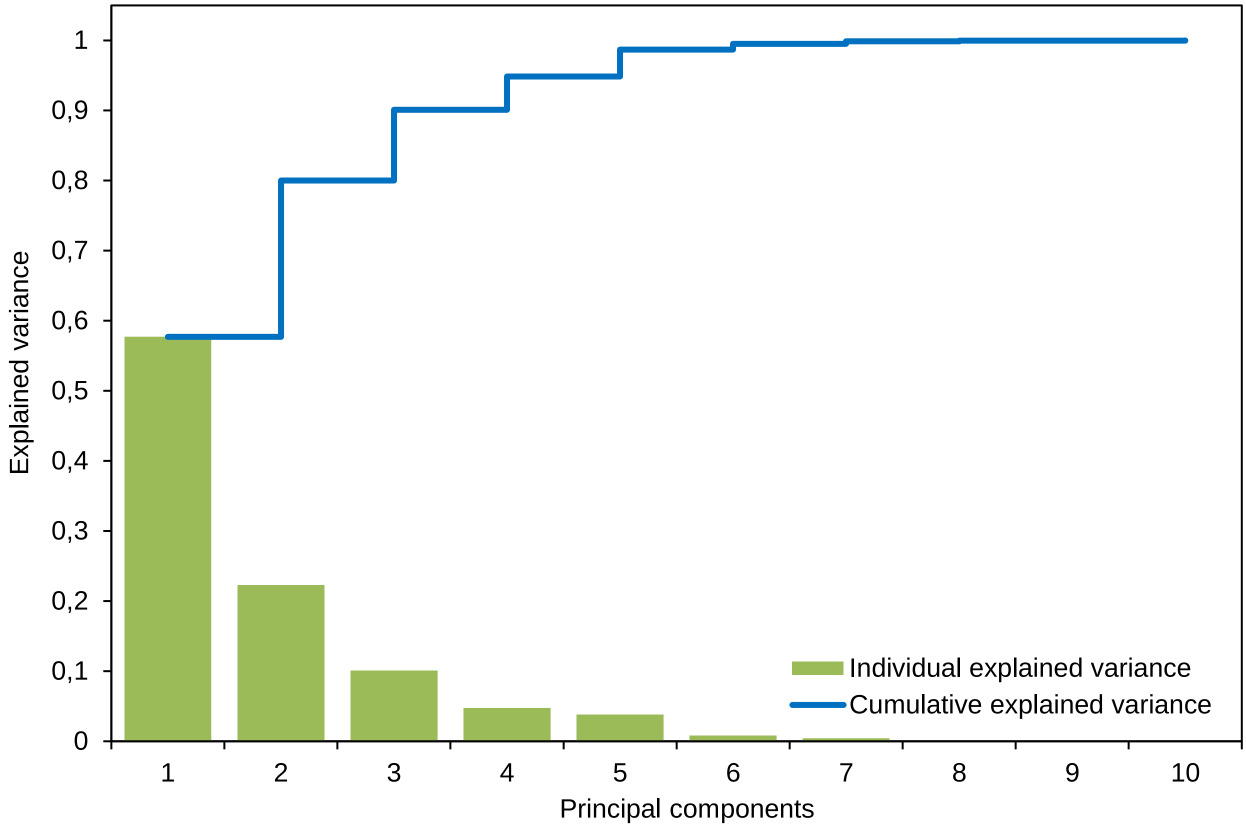


Sources: Rosstat, Authors' calculations.

Names of variables:
Salary (here stands for 'wage') – average monthly nominal wage; *Unempl* – unemployment rate; *Need*_R is the labour shortage rate*; LFP* – the labour force participation rate*; Migration* – the migration growth rate*; Natural* – the natural growth rate*; Young* – the share of under working age population*; Employable* – the share of working age population*; Older* – the share of above working age population*; Graduate* – share of university graduates in the total population.

We applied the most advanced clustering approaches, such as k-means, Agglomerative Clustering, Affinity Propagation, Spectral Clustering, DBSCAN, OPTICS (Vorontsov, 2010; Muller, Guido, 2017; Rashka, 2017). A brief summary of current clustering methods is offered in Annex 2.

Figure 4. Explained variation for principal components



Source: *Authors' calculations.*

Key parameters were changed for each clustering method (Table 1), and the results were compared by internal quality metrics. These metrics assess the clustering quality (the optimal number of clusters) based only on a set of available data and do not use any external information. One of these criteria is *the Silhouette Coefficient*, which is a tool for interpreting and verifying consistency in data clusters. The technique gives a brief graphical representation of how well each object has been classified (Rousseeuw, 1987). The silhouette value is a measure of how well an object matches its own cluster relative to other clusters and varies within [-1; + 1]; the higher its value, the better the object is aligned with its own cluster and the worse with neighbouring ones.

Table 1. Key parameters of clustering methods

Method	Key parameters
k-means	Cluster number
Hierarchical clustering	Cluster distance
DBSCAN	Minimum number of points and the radius of surroundings
OPTICS	Minimum number of points and the radius of surroundings
Affinity Propagation	Damping factor, preferred point
Spectral clustering	Cluster number

However, in the course of analysis, the DBSCAN and OPTICS methods nearly always identified one cluster with several regions as noise. This prevented us from solving the problem of regional labour market clustering, which is why we did not use these methods afterwards.

Table 2 shows the optimal number of clusters for each method and different numbers of principal components.

Table 2. Selecting the best clustering method

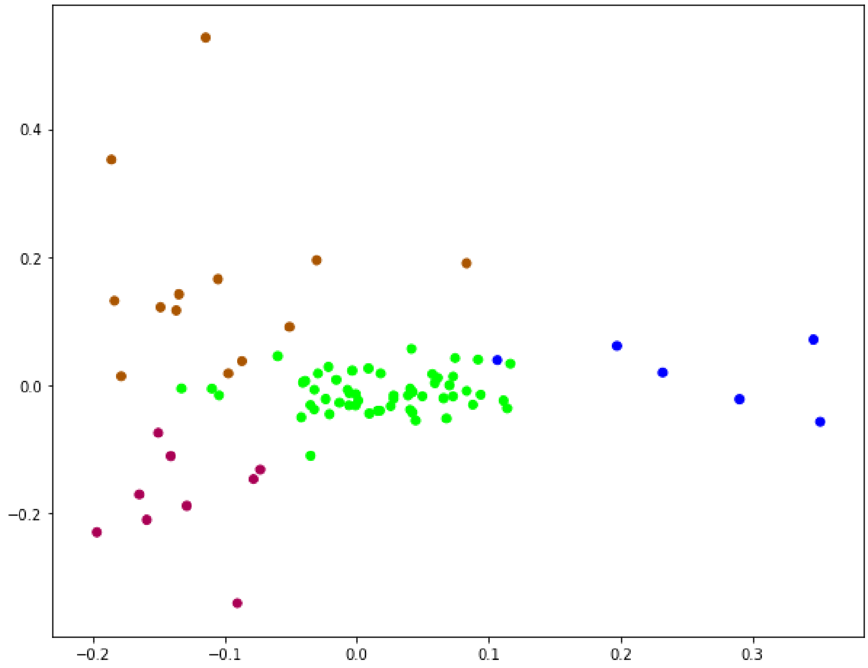
Method	Number of principal components			
	Without components	3	4	5
k-means	3 (0.44)*	4 (0.49)	4 (0.40)	5 (0.42)
Affinity Propagation	3 (0.43)	4 (0.49)	4 (0.37)	5 (0.37)
Agglomerative Clustering	5 (0.38)	4 (0.48)	4 (0.39)	6 (0.44)
Spectral Clustering	3 (0.42)	4 (0.45)	4 (0.39)	5 (0.40)

* – the Silhouette is shown in brackets (internal clustering quality metric). The higher the Silhouette, the better the object is aligned with its own cluster versus other clusters.

Source: Authors' calculations.

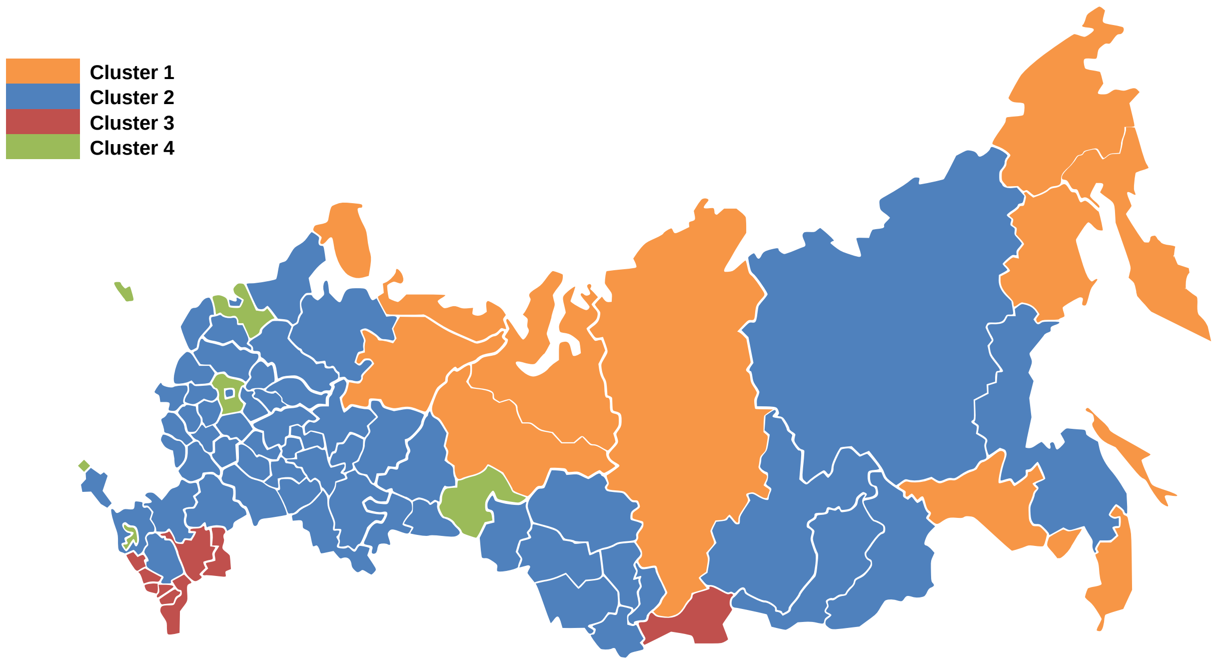
Having compared the results with due regard to the specifics of individual regions, we selected the final clustering for three principal components out of ten key labour market indicators using Affinity Propagation. Figure 5 visualises the final clustering with a fairly clear breakdown into groups.

Figure 5. Final clustering by Affinity Propagation, visualised



As a result, given the heterogeneity of regional labour markets, we identified four groups of regions that are similar to each other within one cluster, but differ from regions in other clusters (Figure 6, Table 3). All clusters, except the fourth, have certain geographical characteristics. For example, the first group of regions mainly includes northern and far eastern border areas, the third is made up of mainly North Caucasus republics, and the second includes almost all remaining regions.

Figure 6. Geographical map of regional labour market clusters



Source: Authors' calculations.

Table 3. Composition of regional labour market clusters

Cluster number	Regions	Features
1 (13 regions)	Amur, Magadan, Murmansk, Sakhalin Regions Jewish Autonomous Region Kamchatka, Krasnoyarsk, Primorsky Territory Republic of Komi Nenets, Khanty-Mansi, Chukotsk, Yamal-Nenets Autonomous Districts	Labour market tightness is low. Labour demand, labour force participation rate and wages are high. The regions are attractive investment wise. Companies in all regions have positive P&L. Nearly a third in the GRP structure is mining. Trade and services are poorly developed in general.
2 (57 regions)	Moscow, St. Petersburg Arkhangelsk, Belgorod, Bryansk, Vladimir, Volgograd, Vologda, Voronezh, Ivanovo, Irkutsk, Kaluga, Kemerovo, Kirov, Kostroma, Kurgan, Kursk, Lipetsk, Nizhny Novgorod, Novgorod, Novosibirsk, Omsk, Orenburg, Oryol, Penza, Pskov, Rostov, Ryazan, Samara, Saratov, Sverdlovsk, Smolensk, Tambov, Tver, Tomsk, Tula, Ulyanovsk, Chelyabinsk, Yaroslavl Regions Altay, Zabaikalye, Krasnodar, Perm, Stavropol, Khabarovsk Territories Republic of Altai, Bashkortostan, Buryatia, Karelia, Crimea, Mari El, Mordovia, Sakha (Yakutia), Tatarstan, Khakassia, Udmurt Republic, Chuvash Republic	There is a high natural population decline, a large share of above working age population, and the smallest share of under working age and employable population.
3 (9 regions)	Astrakhan Region Republic of Daghestan, Ingushetia, Kabardino-Balkaria, Kalmykia, Karachayevo-Cherkessia, Alania, Tuva, Chechnya	Labour market tightness is high. Unemployment, migration outflow and natural population growth are high. Labor demand and remuneration are low. There is a large share of under working age population and a small share of above working age population. The share of agriculture, construction and state participation in GRP is sizeable. There is a very small number of companies. More than half of the regions are loss-making.
4 (6 regions)	Sevastopol Kaliningrad, Leningrad, Moscow, Tyumen Regions Republic of Adygeya	The unemployment rate and migration inflow are low. Wages are high. Companies in all regions have positive P&L. The regions are attractive investment wise. The share of agriculture and public administration in the GRP is the smallest. Trade and real estate transactions are the most developed in the service sector.

The third cluster was the smallest by population (about 6% of the Russian population) with even less weight in total inflation; the second one was the largest (78.1%). The fourth cluster by appeal to migrant workers that records the lowest unemployment and high salaries is populated by almost a tenth of the total population, and its weight in total inflation is 11% (Table 4).

Table 4. Breakdown of clusters by population

Clusters	Number of regions	Population, thousand people	Share in Russia by population	Share in Russia by labour force	Share in Russia by Number of unemployed	Weight in the CPI*
1	13	10,539.9	7.2%	7.5%	7.3%	7.6%
2	57	114,570.0	78.1%	78.1%	71.4%	77.4%
3	9	8,707.2	5.9%	5.3%	14.6%	4.1%
4	6	12,947.6	8.8%	9.1%	6.7%	10.9%
Total	85	146,764.7	100%	100%	100%	100%

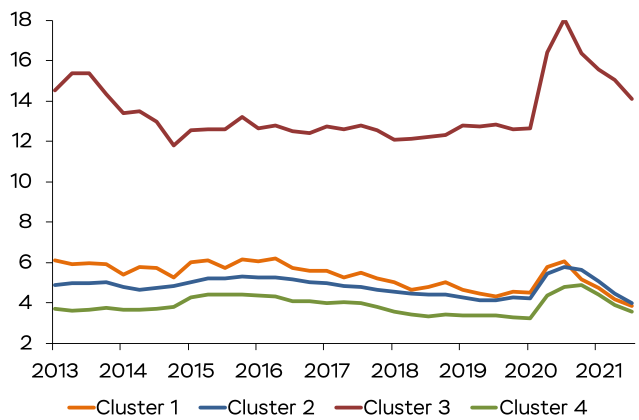
* CPI – consumer price index.

Sources: Rosstat, authors' calculations.

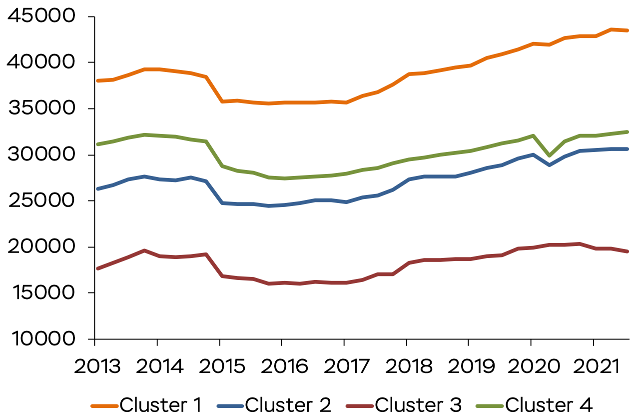
Labour market indicators for each cluster vividly show a distribution of both unemployment and wages for each regional group throughout the period under study (Figure 7).

Figure 7. Labour market indicators for clusters

Unemployment rate (%, SA)



Average monthly real wage (RUB, SA)



Sources: Rosstat, authors' calculations.

Comparative analysis of regional clusters based on socio-economic indicators (Table 5) enabled us to formulate the following distinctive features.

The *first cluster* brings together regions with the least tight labour markets (1.1 applicants/vacancy) and the highest labour demand, labour force participation rate (66%) and wages (60 thousand rubles/month). Mining accounts for almost a third (31.9%) of GRP, while the share of trade is almost twice lower than the national average (8.4% against 15.8%). The cluster has the least developed service sector (at 49.7% of GRP compared to the national average of 62.9%), positive P&L of all companies in its regions (as well as the fourth cluster), and a higher volume of per capita capital investment (2.5 times the national average).

Table 5. Social and economic characteristics of clusters in 2019

Indicator		Cluster			
		1	2	3	4
Unemployment rate, %		4.5	4.3	12.7	3.2
Market tightness ratio, applicants/vacancy		1.1	2.1	91.5	2.0
Labour shortage rate reported to state employment agencies, employees/thousand employees		27.1	10.5	2.5	9.5
Labour force participation rate, %		66.0	61.7	61.3	65.9
Migration growth rate per 10,000 people		-18.7	11.3	-23.4	127.1
Natural growth rate per 1,000 people		-0.5	-3.1	7.8	-1.0
Under working age population,% of the total population		20.2	18.1	22.6	18.9
Working age population,% of the total population		58.8	55.8	58.1	58.3
Above working age population,% of the total population		21.0	26.1	16.4	22.8
Average monthly nominal wage accrued, rubles		60,053	43,611	28,947	55,513
University graduates, % of the total population		0.4	0.7	0.5	0.3
Share in GRP,%	Agriculture, forestry, hunting, fishing and fish-breeding	4.8	6.6	14.4	2.5
	Mining and quarrying	31.9	7.4	7.8	16.3
	Manufacturing sectors	13.6	21.6	5.4	17.5
	Services	49.7	64.4	72.4	63.6
Capital investment per capita, thousands of rubles per capita		331.4	110.1	64.7	251.9
Number of companies per 1,000 people (as of 01.01.2019)		23.2	27.4	11.1	25.5
Share of regions where companies have positive P&L, %		100.0	96.5	44.4	100.0
Share of profitable companies in the total number of companies, %		66.5	71.2	67.9	70.1

Sources: Rosstat, authors' calculations.

The *second cluster*, which is the largest one, is generally close to the average national measures both in terms of the labour market and other social and economic indicators. The cluster is distinguished by a higher natural decline of the population and a larger share of above working age population (26.1% against 16.4%, 21%, and 22.8% in other clusters). So, the shares of under working age and working age population are the lowest. Companies with negative P&L are seen only in the Rostov and Tver Regions. Based on all characteristics of the labour market (labour demand, migration growth, age structure of the population, education level) Moscow and St. Petersburg are also close to the second cluster.

The *third cluster* comprises the border areas of the North Caucasus region and the Republic of Tuva, which shares borders with Mongolia. All these regions have very tight labour markets (91.5 applicants/vacancy vs 2.1 applicants/vacancy across Russia) and a high unemployment rate (the group average is 12.7%). Other specific features of this cluster include a very low labour shortage rate reported to state employment agencies (the group average is 2.5 vacancies per 1,000 people – against the national average of 11.1) and, as a result, the lowest level of wages (29,000 rubles a month on average). The cluster shows the largest migration outflow and – unlike other clusters – positive natural population growth.

In the age structure, the cluster is notable for its high share of under working age population (25.6% vs 18.7% across Russia), and in contrast, a low share of the above working age population (16.4% vs 25% across Russia).

As for GRP structure, this group of regions has an outstanding share of agriculture, forestry and hunting (14.4% against 4.3% across Russia) with a very low share of the manufacturing sector (5.4% vs 18% across Russia). An important feature is a significant share of state participation in the regional economies (the share of public administration and defence in GRP is 11.1% vs 4.5% across Russia). There is the highest share of construction in GRP – 12.1% (vs the 5.6% national average).

Another important feature of the cluster is a very small number of companies (11.1 per 1,000 people vs 26.1 per 1,000 people across Russia). Also, companies have negative P&L in five out of nine regions.

The *fourth cluster* has the smallest number of regions but they are significant. It enjoys the lowest unemployment rate (3.2% vs 4.6% across Russia). It is also characterised by significant migration inflows (127.1 per 10,000 people), wages above the national average (55,500 vs 47,500), and companies with positive P&L in all its regions, as well as two-fold capital investment per capita vs the country average.

Also, the regions have the smallest share of agriculture, forestry and hunting (2.5%). Top business lines in the service sector are retail and real estate, with minimum involvement of public administration.

Therefore, the cluster analysis of all Russian regions based on the key labour market indicators identified four main regional clusters with different unemployment, wages, migration, and natural growth rates as well as several other indicators.

3.3. Assessing the impact of labour market indicators on inflation in clusters

To measure the contribution of wage changes and unemployment to inflation, regression models were estimated for each cluster (Table 7). Importantly, the models have virtually identical specifications to ensure further comparability of parameter estimates and calculation of the aggregate labour market indicator.

The input data are quarterly regional statistics for labour market indicators and inflation, as well as the ruble exchange rate for 2013 Q1–2021 Q2.

The consumer price index (CPI, QoQ) for each cluster was calculated in the following procedure. First, we recalculated the weight of each region by its contribution to the CPI of Russia[5] based on its contribution to the cluster (Annex 3). Second, we calculated the average CPI of regions within one cluster based on their recalculated weights. Wages were also weighted by the number of employees in the regions. The unemployment rate was calculated as the ratio of the total number of unemployed to the number of workforce in the cluster, in accordance with ILO methodology.

Below is the resulting model specification for all clusters:

$$\Delta CPI_t = \beta_0 + \beta_1 \Delta CPI_{t-1} + \beta_2 Neer_t + \beta_3 \Delta Wage_{t-4} + \beta_4 \Delta Unempl_{t-1} + \\ + \beta_5 D_{15q1} + \beta_6 D_{15q2} + \varepsilon_t, \quad (1)$$

where *t* is the time index (quarters), Δ is the first difference operator, CPI_t is the CPI (% on the preceding period), $Neer_t$ is the index of the nominal effective exchange rate of the ruble to foreign currencies (% of growth on the previous period), $Wage_t$ is average monthly real wage accrued (% on the previous period), $Unempl_t$ is the unemployment rate (%), D_{15q1} and D_{15q2} are dummy variables equal to '1' for 2015 Q1 and Q2 respectively and to '0' for the rest, and ε_t is the random component. Statistical characteristics of the variables for the entire Russia and each cluster are shown in Annex 4.

All figures except the exchange rate were seasonally adjusted. The quantitative variables in all the models were tested and reduced to a stationary form if necessary. Also, the constructed models were tested for multicollinearity, heteroscedasticity, autocorrelation

[5] According to Rosstat

in balances, with adjustments made where necessary. The inclusion of the nominal effective exchange rate of the ruble against foreign currencies (*Neer*) in the index model at the same time with the CPI is attributable to the negligible impact of this variable on CPI growth in lags. This can be explained as follows. Since the data are quarterly-based, a change in the exchange rate at the beginning of the quarter bring about changes in inflation as early as by the end of the quarter.

The results of model estimates for clusters, as well as those for Russia as a whole, stripping out regional heterogeneity, are shown in Table 6.

In the obtained models, labour market indicators make a strong impact on inflation (except for unemployment in the 4th cluster); at the same time, it should be noted that the quality of model adjustment to input data is high (R^2 more than 0.93). The high R^2 is explained by the great contribution of the dummy variables to the quality of equation fitting, providing further confirmation of a low but meaningful impact of labour on inflation.

As follows from the decomposition of inflation factors (Annex 5), the contribution of unemployment to inflation prevails over that of real wages in the first cluster (unlike the other clusters and Russia as a whole). This is due to the fact that the first cluster regions are mainly engaged in extraction, that is, they are the most attractive investment wise with the highest labour force participation rate and a rate of inflation considerably lower than across Russia. Therefore, a slight change in employment may have a meaningful impact on the economy and exert further inflationary pressure.

The rest of the clusters are comparable with the national average in terms of contributions of inflation drivers but differ in the sensitivity of these drivers. This is primarily due to component volatility. For example, the third cluster is distinguished by the highest unemployment (volatility of the unemployment gap), tight labour market, migration outflows and the lowest wages. All this increases the weight of labour market in price movements.

The lowest sensitivity of factors in the fourth cluster is explained by their lowest unemployment rate, as well as a steady labour market, explained by significant migration inflows and high wages.

Table 6. Regression models of the impact of the labour market on inflation in regional clusters

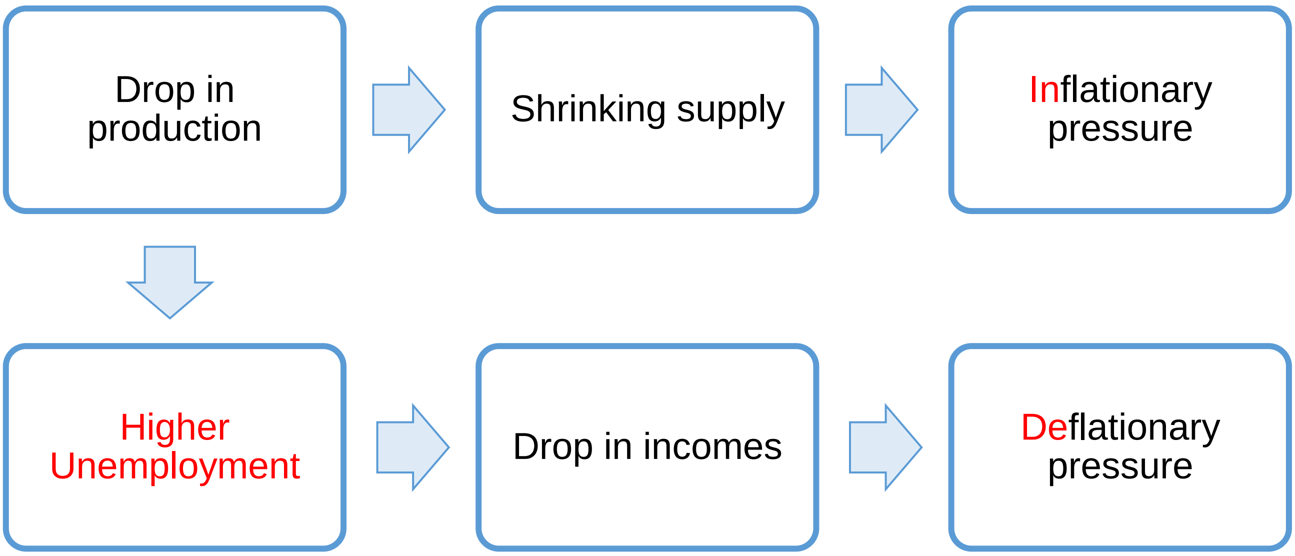
Variable	Ratios				
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Russia
β_0	-0.06 (0.06)*	-0.01 (0.03)	0.03 (0.07)	0.02 (0.14)	-0.03 (0.04)
ΔCPI_{t-1}	0.23 (0.07)	0.15 (0.04)	0.19 (0.08)	0.11 (0.06)	0.16 (0.06)
$Neer_t$	-0.06 (0.01)	-0.04 (0.01)	-0.06 (0.02)	-0.02 (0.01)	-0.05 (0.01)
$\Delta Wage_{t-4}$	0.06 (0.03)	0.03 (0.02)	0.09 (0.03)	0.04 (0.02)	0.09 (0.03)
$\Delta Unempl_{t-1}$	-0.88 (0.19)	-0.25 (0.12)	-0.19 (0.1)	-0.27 (0.18)	-0.36 (0.18)
D_{15q1}	2.16 (0.39)	4.46 (0.3)	4.89 (0.76)	3.71 (0.25)	3.73 (0.47)
D_{15q2}	-3.55 (0.43)	-5.60 (0.27)	-6.78 (0.59)	-4.13 (0.50)	-5.12 (0.39)
D_{17q3}	–	-0.58 (0.21)	–	–	–
D_{20q1}	0.64 (0.31)	–	–	–	–
$AR(1)$	–	-0.54 (0.17)	-0.65 (0.21)	0.69 (0.19)	-0.76 (0.19)
R^2_{adj}	0.95	0.98	0.93	0.96	0.95

* Standard errors are shown in brackets.
Negligible estimates (α > 0.1) are highlighted in orange.
Source: Authors' calculations.

The endogeneity problem in models of the unemployment and inflation relationship

Unemployment impacts inflation as follows. An increase in unemployment makes it easier for employers to hire people, which in turn has a positive effect on costs and ultimately slows down inflation. We therefore expect that the relationship between unemployment and inflation is reverse; yet, when supply-side shocks emerge, both prices and unemployment may simultaneously rise (Figure 8). This co-directional change in unemployment and inflation under the influence of third factors may bring about the endogeneity problem in linear models – that do not take into account the influence of these factors – to the effect that estimates become biased and unreliable.

Figure 8. Unemployment and inflation relationship



This necessitated the search for an instrumental variable for unemployment rate, which would not correlate with ε_t but would highly correlate with $Unempl_t$. In the course of our research, no significant tools were found for the models for clusters, individual regions and entire Russia, other than regressions for Moscow and St. Petersburg. We managed to select high-quality tools for these two cities and confirm endogeneity, but the labour market factors become insignificant and have incorrect signs in economic sense.

3.4 Calculating the Labour Market Indicator (LMI)

To measure the labour market and inflation relationship indicator, it is proposed to calculate the total contribution of unemployment and wages to inflation for each cluster; then, to evaluate the total contribution of the labour market to countrywide inflation given the weight of each cluster in the CPI.

The contribution of the i-th factor to the CPI of the k-th cluster in the t-th period of time is determined by the following formula:

$$\gamma_{i_t}^k = \frac{\hat{\beta}_0 \cdot \left|\hat{\beta}_i x_{i_t}\right|}{\sum_{j=1}^{n} \hat{\beta}_j x_{j_t}} + \hat{\beta}_i x_{i_t}, \tag{2}$$

where $\gamma_{i_t}^k$ is the contribution *of the i-th* factor to the CPI *of the k-th* cluster in the *t-th* period of time (quarter), $\hat{\beta}_i$ is the *i-th* factor estimate, and x_{i_t}, x_{j_t} *is the i-th* and *j-th* factor respectively.

The final indicator is the sum of average contributions of unemployment and wages weighted by the contribution of each cluster to the CPI.

$$LMI_t = \sum_{k=1}^{4} \pi^k \cdot \left(\gamma_{wage_t}^k + \gamma_{un_t}^k\right), \tag{3}$$

Where π^k is the weight in the CPI *of the k-th* cluster, $\gamma_{wage_t}^k$, $\gamma_{un_t}^k$ is respective contributions of wages and unemployment to the CPI of the *k-th* cluster.

Calculations show the following cluster weights in the CPI:

$$\pi^1 = 0.0755,$$
$$\pi^2 = 0.7742,$$
$$\pi^3 = 0.0411,$$
$$\pi^4 = 0.1092.$$

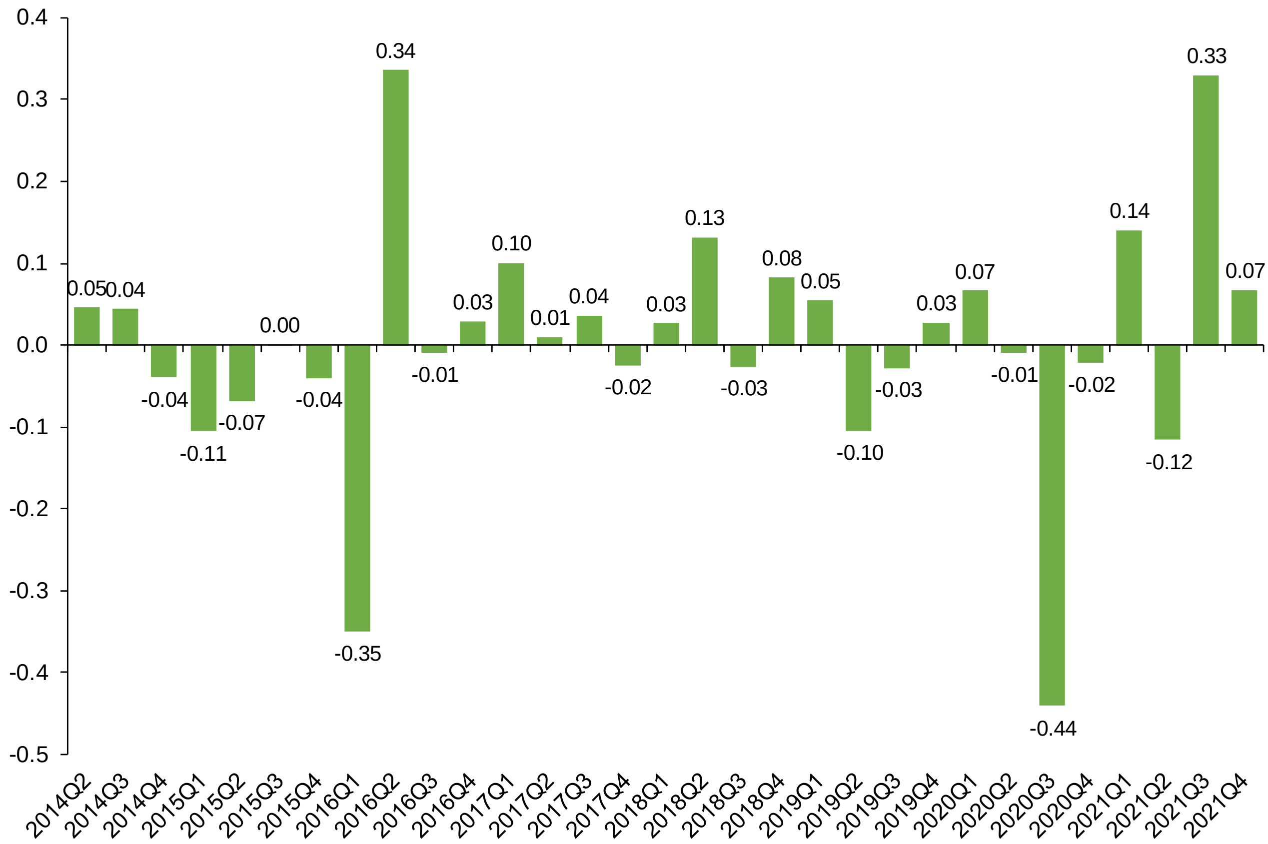
LMI values for 2014 Q2–2021 Q3 are shown in Figure 9.

Indicator values describe the contribution of changes in the labour market to inflation growth in pp.

Given that real wages are incorporated in the index with a four-quarter lag, the current LMI value is determined by wage changes in the previous year; similarly, the index captures unemployment with a one-quarter lag. This enables us to estimate the indicator for one quarter ahead of available statistics. The positive value of the indicator signals an inflationary effect of the labour market, and the negative value suggests a disinflationary one.

In particular, a noticeable LMI decline in 2020 Q3 was due to a sharp increase in Q2 unemployment, with relatively steady real wage growth one year before, resulting in a marked disinflationary effect of the labour market (-0.44pp).

Figure 9. Labour market (inflationary pressure) indicator (LMI), pp



Source: Authors' calculations.

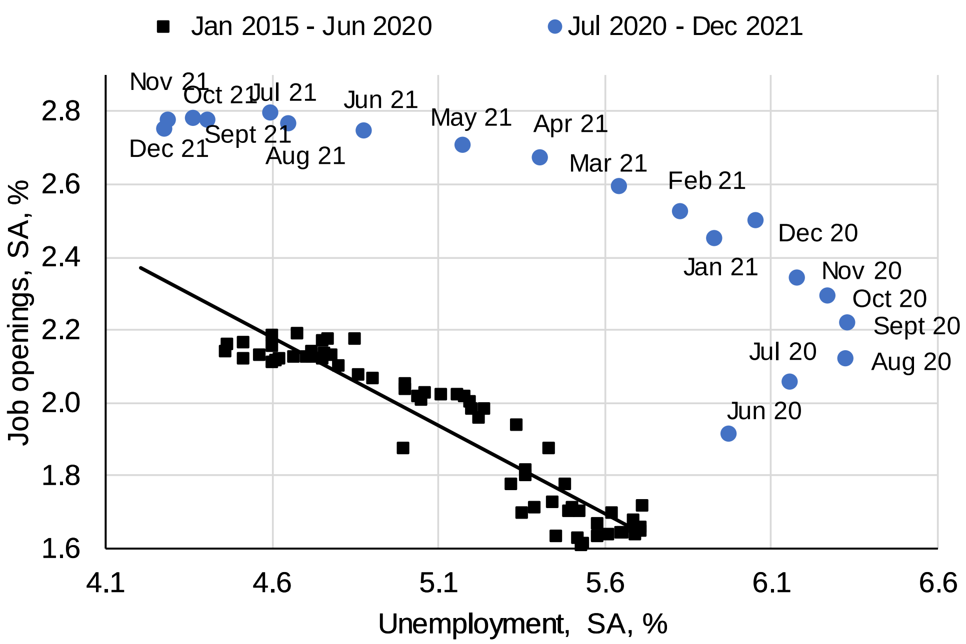
A noticeable increase in real wages at the end of 2019 significantly weakened the disinflationary effect of the labour market in 2020 Q4.

Although unemployment rate remained rather high in 2021 Q1, at the end of 2020, unemployment began to decline noticeably, triggering higher inflationary pressure from labour market. Labour market started to recover actively, with growing demand for labour resources. The contribution to inflation growth was +0.14pp.

In 2021 Q2, the indicator signalled a disinflationary effect of the labour market due to a significant drop in household incomes during the lockdown (2020 Q2) though there was a rising negative unemployment gap. Moreover, labour demand reached a six-year high. The scarcity of labour migrants combined with rising demand for employees in individual industries made the labour shortage more acute.

In 2021 Q3, there was a noticeable inflationary effect from labour market prompted by a recovery of household incomes after the lockdown and a considerably accelerated decline in unemployment in the previous quarter. From the second half of the year to the present, the Beveridge curve has shifted to the right and upwards, signalling the pronounced significant structural shifts in labour demand and supply, specifically, a rising natural unemployment rate (Figure 10). This is particularly evident in low-skilled industries. In the context of labour shortages in several industries, competition for labour resources tightened at both cross-regional and cross-sectoral levels, labour intensity increased triggering wage growth.

Figure 10. Beveridge curve for Russia



Sources: Rosstat, authors' calculations.

Below are key reasons for the existing labour shortage.

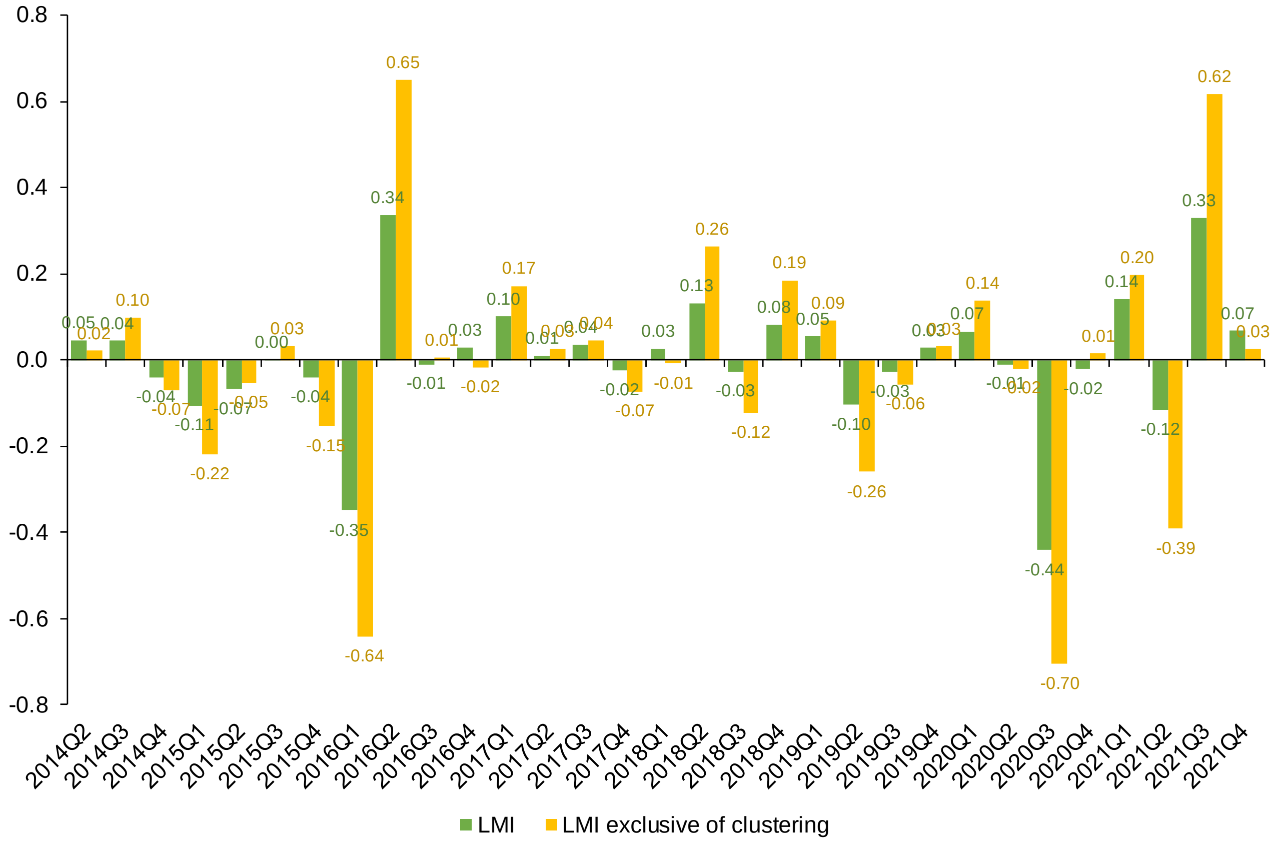
1. A demographic pitfall: young people who were born in the 1990s when birth rates were low are entering the labour market.
2. The shortage of blue-collar workers due to the problems in vocational education (which was not popular with students).
3. HR policies during the crisis period: in 2020, employers laid off people or kept vacancies open, cut down wages and employee benefits to cut costs.
4. Applicants became more selective: The 2020 layoffs made them value their current jobs more.
5. Shortage of labour migrants (restrictions, increasing entry costs).

Specifically, agricultural companies noted that it was more feasible to recruit employees through recruitment agencies, rather than to make arrangements for migrants to come to the region. Industrialists and builders highlighted the unpopularity of blue-collar jobs among the local population at the moment, while migrants increasingly preferred to work in the service sector.

In 2021 Q4, the inflationary impact of the labour market (+0.10 pp) grew weaker as the decline in unemployment slowed down after hitting the historic low with a slight drop in and real wages registered at the end of 2020. However, labour demand was at an all-time high because of labour shortages mainly in construction, industrial production, transport, and logistics. Foreign labour migration continued to recover significantly lagging behind the 2018–2019 levels.

Also, based on the model for entire Russia, we calculated the labour market indicator exclusive regional clustering for comparison purposes (Figure 11).

Figure 11. Comparison of labour market indicators inclusive and exclusive of regional clustering



Source: Authors' calculations.

We note its noticeably greater volatility in most periods and a stronger response to shocks. This suggests that the indicator estimates are somewhat tilted due to the fact that regional labour market specifics were out of scope.

4. Conclusion

Our study puts forward a composite indicator that makes it possible to measure the aggregate *quantitative* contribution of changes in the labour market to inflation growth both for the entire country and regional clusters.

Our analysis enabled us to classify regional labour markets (with regions having been grouped into four clusters) and identify their key distinctive social and demographic features in conjunction with economic development. Specifically, the first cluster has high investment appeal and the highest labour force participation rate and wages. However, its sectorial structure (mining) and underdeveloped services result in population outflows. The second cluster faces the problems of natural decline and ageing of the population. The third cluster has poor social and economic environment: low employment, labour demand and wages come together with loss-making status of its regions and their low investment appeal. The cluster is focused on trade and agriculture. Finally, the fourth cluster is attractive to investment; it posts high economic performance and boasts a mature service sector. As a result, the household economic activity is high, demographic trends are positive. These specific features may be captured in the forecast models of regional inflation.

Labour market plays an important role in inflation processes in many countries. We expected Russia to have the same situation. However, given the regional specific features, our results show that labour market has a statistically significant yet small effect on inflation in Russia.

Most regions are marked by a weak impact of unemployment on inflation, which may suggest that the Phillips curve is horizontal. At the same time, price growth in the northern mining regions is more sensitive to changes in unemployment in contrast to the entrenched country-wide trend of the labour market to adjust to economic fluctuations mainly through wage adjustment. This sensitive link between the labour market and inflation for low-unemployment regions may be the subject of a separate study.

Yet another advantage of the LMI is its interpretability (it allows to quantify the contribution of the labour market to inflation). An individual value of the indicator is a measure of pro-/disinflationary impact in a given period. This helps make better informed monetary policy decisions.

As the obtained models show the labour cost pass-through into price growth takes from one to four quarters, the indicator makes it possible to assess the future impact of the labour market on inflation (one quarter ahead of available statistical data).

Moving forward, the search for and subsequent use of high quality tools to include the endogeneity in models may well improve the accuracy of the labour market indicator, with the proposed methodology facilitating its calculation for individual regions. The LMI could also be used in economic models as an independent factor.

Literature

1. Ankerst, M., Breunig, M., Kriegel, H., Sander, J. (1999). OPTICS: Ordering points to identify the clustering structure. *ACM SIGMOD International Conference on Management of Data.* ACM Press, pp. 49–60.
2. Arias-Castro, E., Chen, G., Lerman, G. (2011). 'Spectral clustering based on local linear approximations.' *Electronic Journal of Statistics, Vol.* 5. https://doi.org/10.1214/11-EJS651
3. Chow W. (2011). A state space model of NAIRU, 16 March. https://www.hkeconomy.gov.hk/en/pdf/wp/structural_unemp.pdf
4. Chung, H.T., Fallick, B., Nekarda, C.J., Ratner, D.D. (2014). 'Assessing the Change in Labor Market Conditions (2014)'. *Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board*, Washington, D.C.
5. Clifton, J., Marlar, J. (2011). Good Jobs: The New Global Standard. Gallup, Inc. p. 20 https://www.gallup.com
6. Dungan, P., Fang, T., Gunderson, M. (2012). Macroeconomic Impacts of Canadian Immigration: Results from a Macro-Model. Institute for the Study of Labor. IZA DP No. 6743. p. 36.
7. Ester, M., Kriegel, H., Sander, J., Xu, X. (1996). 'A density-based algorithm for discovering clusters in large spatial databases with noise'. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*, pp. 226–231.
8. Frey, B., Dueck, D. (2007). 'Clustering by passing messages between data points.' *Science*, *Vol.* 315. https://doi.org/10.1126/science.1136800
9. Furlanetto, F., Robstad, O. (2016). 'Immigration and the macroeconomy: some new empirical evidence'. Norges Bank Working Papers No. 18. 42 p.
10. Gordon R.J. (2013). The Phillips curve is alive and well: inflation and the NAIRU during the slow recovery. NBER Working Paper, No. 19390. https://doi.org/10.3386/w19390
11. Rusticelli E. (2014). Rescuing the Phillips curve: making use of long-term unemployment in the measurement of the NAIRU. OECD Journal: Economic Studies, Vol. 2014/1 pp. 109–125. http://dx.doi.org/10.1787/eco_studies-2014-5jxrcm2cdff6
12. Rousseeuw, P. (1987). Silhouettes: 'A graphical aid to the interpretation and validation of cluster analysis'. Journal of Computational and Applied Mathematics. Vol. 51(20), pp. 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

13. Smith, C., Thoenissen, C. (2018). 'Migration and Business Cycle Dynamics'. Discussion Paper Series No. 7. p. 47
14. Steinhaus, H. (1956). 'Sur la division des corps materiels en parties'. *Bulletin de l'academie Polonaise des Sciences, Cl.* III – Vol. No. 12.
15. Weiske, S. (2019). On the Macroeconomic Effects of Immigration: A VAR Analysis for the US. German Council of Economic Experts. p. 29.
16. Bragin, V., Osakovsky, V. (2004). 'Otsenka estesvennogo urovnya bezrabotitsy v Rossii v 1994–2003 gg.: empiricheskii analiz [Measuring the natural unemployment rate in Russia in 1994–2003]'. *Voprosy Ekonomiki* [in Russian]. No. 3, p. 95.
17. Vorontsov, K.V. (2010). Lektsii po algoritmam klasterizatsii i mnogomernogo shkalirovaniya. [Lectures in clustering and multi-dimensional scaling algorithms] [in Russian]. Moscow, Moscow State University.
18. Gafarov, B.N. (2011). 'Krivaya Fillipsa i stanovlenie rynka truda v Rossii [The Phillips curve and the evolvement of the Russian labour market]'. *HSE Economic Journal* [in Russian]. Vol. 2, pp. 155–176.
19. Gurvich, E., Vakulenko E. (2018). 'Issledovaniya rossiyskogo rynka truda i ekonomocheskaya politika [Studies of the Russian labour market and economic policy]' [in Russian]. *Journal of the New Economic Association* [in Russian]. Vol. 1 (37), pp. 203–212.
20. Ivanova, M. (2016). 'Analiz kharaktera prichinno-sledstvennoi svyazi mezhdu inflyatsiei i zarabotnoi platoi v Rossii [Analysing the nature of the cause-effect relationship between inflation and wages in Russia]'. *Studies on Russian Economic Development* [in Russian]. Vol. 5 (158), pp. 119–132.
21. Indeks rynka truda: kratkoe opisanie metodiki [Labour market index: short methodology description]. An RSPP, AEI PRIME, Headhunter Group and IBS Group project [in Russian], 2013, p. 20.
22. Indeks rynka truda v regionakh Rossii [Labour Market Index in Russian Regions. Website of RIA Rating [in Russian]. [https://riarating.ru/infografika/20210906/630207557.html]
23. Kudayeva, M., Redozubov, I. (2021). 'Vliyanie migratsionnykh potokov na ekonomicheskuyu askivnost i rynok truda v tselom i regionalnom aspekte [Impact of migration flows on the economic activity and labour market of Russia as a whole and regionally'. Bank of Russia Working Papers, p. 45,

http://www.cbr.ru/StaticHtml/File/131869/wp_khab_dec.pdf

24. Muller, A., Guido, S. (2017). Introduction to machine learning with Python: a guide for data scientists [Translation into Russian]. Moscow. Williams.
25. Orlov, D., Postnikov, E. (2020). ‘Krivaya Fillipsa: inflyatsia i NAIRU v rossiiskikh regionakh [Phillips curve: inflation and NAIRU in Russian regions]’. Bank of Russia Working Papers [in Russian], p. 31, http://cbr.ru/Content/Document/File/115690/pc_reg.pdf
26. Sberbank Investment Research Portal. http://research.sberbank-cib.com
27. Rashcka, S. (2017). Python and machine learning. [Translated from English by A.V. Logunov]. Moscow.: DMK Press.
28. ‘Rossiyskiy rynok truda: tendentsii, instituty, strukturnye izmeneniya [Russian Labour Market: Trends, Institutions, Structural Changes]’ *Report of the Centre for Labour Research and the Labour Market Research Laboratory of HSE*. Edited by Gimpelson, V., Kapelyushnikova, R., and Roshchina, S., Moscow, 2017, p. 145
29. Sokolova, A.V. (2014). ‘Inflyatsionnye ozhidaniya i krivaya Fillipsa: otsenka na rossiyskikh dannykh [Inflation expectations and the Phillips curve: estimates based on Russian data]’. *Russian Journal of Money and Finance*. Vol. 11., pp. 61–67.

Annex 1. Brief review of the literature on the impact of labour market indicators on inflation

indicators	Authors	Summary
Unemployment / unemployment gap	Bragin, Osakovsky (2004), Paliy (2006), Gafarov B.N. (2011), Gurvich, Vakulenko (2016, 2018)	estimates the natural unemployment rate (including the NAIRU) and the unemployment gap in Russia; the case made for an inverse relationship between inflation and unemployment
	Gordon (2013), Rusticelli (2015)	estimates NAIRU and unemployment gap based on country data
	Orlov, Postnikov (2020)	estimates NAIRU for Russian regions using state-space models (Kalman filter)
Unemployment gap / output gap	Chow (2011)	estimates the Phillips curve and Okun's law; provides rationale for the unemployment gap and output gap relationship and measures the degree of inflationary pressure
	Sokolova (2014)	estimates the hybrid Phillips curve with an inflation expectations component
Wages	Ivanova (2016), Gurvich, Vakulenko (2018)	shows a two-way relationship between inflation and wages in Russia with a shift to inflation targeting facilitating a switch to the demand-pull inflation model
Migration	Dungan at el. (2012), Furlanetto (2016), Smith, Thoenissen (2018), Weiske (2019), Kudaeva, Redozubov (2021)	shows that migration processes have no significant effect on inflation

Annex 2. Brief summary of current clustering methods

k-means clustering is one of the simplest commonly used clustering algorithms (Steinhaus, 1956). The k-means method works well when the clusters are compact 'clouds', considerably divided, and their structure is hyperspherical. It delivers good results in processing large data volumes, but it is not applicable for detecting non-convex clusters or clusters of very different sizes. Moreover, the method is very sensitive to noise and isolated points of space, since even a small number of such points can significantly influence the calculation of a cluster's centre of mass.

A special feature of the method is the need to know the number of clusters in advance. For the k-means method, the optimal number of clusters is commonly estimated with the following criterion: the sum of squares of the distance from points to the centroids of the clusters to which they belong:

$$J(C) = \sum_{k=1}^{K} \sum_{i \in C_k} \|x_i - \mu_k\|^2 \to \min_{C},$$

where *C is* a set of *K* power clusters, μ_k is the centroid of cluster C_k.

It is necessary to select the number of clusters starting from which the functionality *J(C*) decreases the slowest, or

$$D(k) = \frac{|J(C_k) - J(C_{k+1})|}{|J(C_{k-1}) - J(C_k)|} \to \min_{k}$$

Spectral clustering techniques use the spectrum (own values) of the data affinity matrix in order to reduce the dimensionality before clustering in smaller dimension spaces (Arias-Castro et al., 2011). The affinity matrix is presented as an input and consists of quantitative estimates of the relative similarity of each pair of points in the data.

Spectral clustering is quite closely related to the k-means method, and therefore it is necessary to know the number of clusters in advance. The optimum number of clusters are estimated in a fashion similar to the k-means method. The method works well for a small number of clusters and is not recommended for multiple clusters.

Agglomerative (hierarchical) clustering is an algorithm without a fixed number of clusters that allows to build a cluster gluing tree – a dendrogram. Based on its view, we can determine the best time to stop the algorithm.

Hierarchical clustering is less sensitive to noisy data, but performs worse in clustering big data relative to the *k-means* method, with the time complexity of the algorithm being quadratic for the hierarchical clustering method.

Density-based spatial clustering with noise (DBSCAN) is a popular density-based clustering algorithm used in data analysis as an alternative to the k-means method (Ester et al., 1996). If a set of points is given in a space, the algorithm groups together closely located points (those with many close neighbours), marking as outliers those points that are lonely in low density areas (whose nearest neighbours lie far).

To apply the method, two parameters must be configured: the maximum distance between adjacent points and the minimum number of points in the neighbourhood (the number of neighbours), when it is possible to conclude that these copies of data make up one cluster. The determined DBSCAN clusters can have any form.

OPTICS (Ordering Points to Identify the Clustering Structure) is another density-based algorithm to find clusters in spatial data (Ankerst et al., 1999). It is meant to eliminate a weak point of the DBSCAN algorithm – the problem of detecting content clusters in data that have different densities. This is achieved by ordering database points (linearly) so that spatially close points become adjacent. In addition, for each point a special distance is stored; it represents the density the cluster should take on to assign the points to one and the same cluster.

The basic idea of the **affinity propagation** method involves clustering observations in groups based on how they ‘communicate’, or how similar they are to each other (Frey, Dueck, 2007). This algorithm performs well when the proximity function is known in advance and many clusters of various forms are expected to appear with a slightly varying number of elements.

Annex 3. Regions' weighs by their contribution to the CPI

Region	Cluster	Region's weight in the Russian Federation	Region's weight in its cluster	Cluster's weight in the Russian Federation
Amur Region	1	0.0048	0.0630	0.0755
Jewish Autonomous Region		0.0009	0.0121	
Kamchatka Territory		0.0035	0.0460	
Krasnoyarsk Territory		0.0204	0.2708	
Magadan Region		0.0015	0.0198	
Murmansk Region		0.0073	0.0968	
Nenets Autonomous Area		0.0003	0.0043	
Primorye Territory		0.0152	0.2019	
Republic of Komi		0.0053	0.0700	
Sakhalin Region		0.0051	0.0669	
Nenets Autonomous Area		0.0081	0.1076	
Chukotka Autonomous Area		0.0004	0.0059	
Yamalo-Nenets Autonomous Area		0.0026	0.0350	
Altai Territory	2	0.0114	0.0147	0.7742
Arkhangelsk Region without the autonomous district		0.0085	0.0110	
Belgorod Region		0.0102	0.0131	
Bryansk Region		0.0062	0.0080	
Vladimir Region		0.0079	0.0102	
Volgograd Region		0.0140	0.0181	
Vologda Region		0.0074	0.0096	
Voronezh Region		0.0133	0.0172	
Moscow		0.1574	0.2033	
St. Petersburg		0.0526	0.0679	
Zabaikalye Territory		0.0058	0.0075	
Ivanovo Region		0.0065	0.0084	
Irkutsk Region		0.0132	0.0171	
Kaluga Region		0.0066	0.0085	
Kemerovo Region		0.0138	0.0179	
Kirov Region		0.0072	0.0093	
Kostroma Region		0.0037	0.0048	
Krasnodar Territory		0.0361	0.0467	
Kurgan Region		0.0041	0.0053	
Kursk Region		0.0063	0.0081	
Lipetsk Region		0.0068	0.0087	
Nizhny Novgorod Region		0.0210	0.0271	
Novgorod Region		0.0032	0.0042	
Novosibirsk Region		0.0152	0.0196	
Omsk Region		0.0105	0.0135	
Orenburg Region		0.0101	0.0130	
Orel Region		0.0039	0.0050	
Penza Region		0.0065	0.0083	
Perm Territory		0.0162	0.0209	
Pskov Region		0.0031	0.0041	
Altai Republic		0.0010	0.0012	

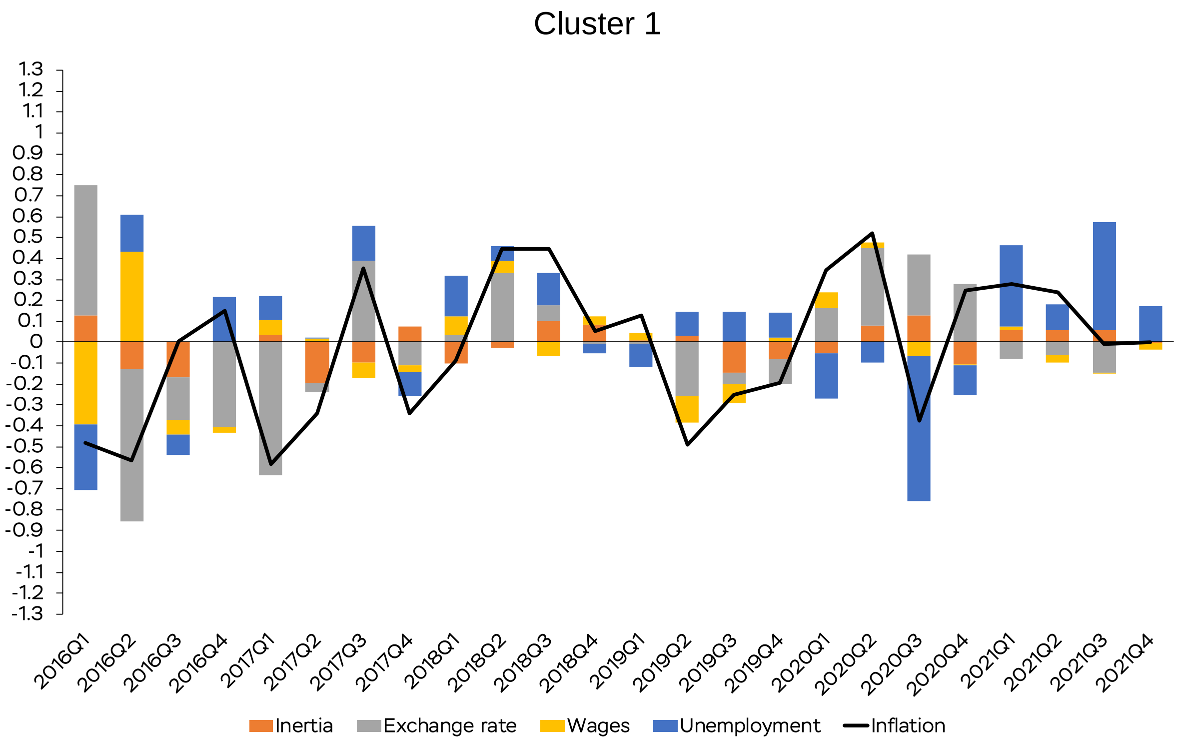
Region	Cluster	Region's weight in the Russian Federation	Region's weight in its cluster	Cluster's weight in the Russian Federation
Republic of Bashkortostan		0.0255	0.0329	
Republic of Buryatia		0.0045	0.0058	
Republic of Karelia		0.0042	0.0055	
Republic of Crimea		0.0095	0.0123	
Mari El Republic		0.0029	0.0037	
Republic of Mordovia		0.0034	0.0044	
Republic of Sakha		0.0086	0.0111	
Republic of Tatarstan		0.0272	0.0351	
Republic of Khakassia		0.0029	0.0037	
Rostov Region		0.0259	0.0335	
Ryazan Region		0.0048	0.0062	
Samara Region		0.0204	0.0263	
Saratov Region		0.0115	0.0148	
Sverdlovsk Region		0.0284	0.0367	
Smolensk Region		0.0050	0.0064	
Stavropol Territory		0.0138	0.0178	
Tambov Region		0.0046	0.0060	
Tver Region		0.0071	0.0091	
Tomsk Region		0.0069	0.0090	
Tula Region		0.0087	0.0112	
Udmurt Republic		0.0092	0.0119	
Ulyanovsk Region		0.0060	0.0078	
Khabarovsk Territory		0.0110	0.0142	
Chelyabinsk Region		0.0203	0.0262	
Chuvash Republic		0.0049	0.0063	
Yaroslavl Region		0.0074	0.0096	
Astrakhan Region	3	0.0063	0.1535	0.0411
Kabardino-Balkar Republic		0.0040	0.0967	
Karachay-Cherkess Republic		0.0021	0.0503	
Republic of Dagestan		0.0152	0.3702	
Republic of Ingushetia		0.0017	0.0419	
Republic of Kalmykia		0.0012	0.0280	
Republic of North Ossetia–Alania		0.0033	0.0809	
Republic of Tuva		0.0012	0.0292	
Chechen Republic		0.0061	0.1493	
Sevastopol	4	0.0026	0.0239	0.1092
Kaliningrad Region		0.0056	0.0513	
Leningrad Region		0.0128	0.1170	
Moscow Region		0.0675	0.6182	
Republic of Adygeya		0.0026	0.0235	
Tyumen Region without autonomous districts		0.0181	0.1661	

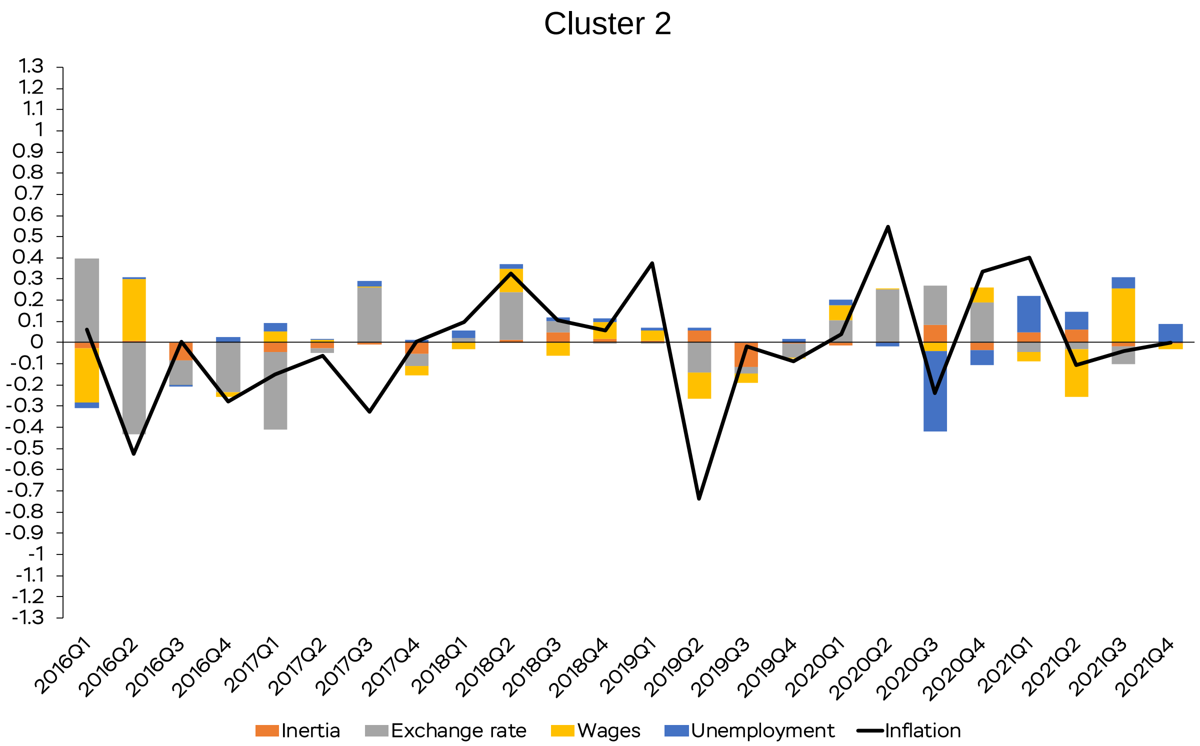
Annex 4. Descriptive statistics of regression variables

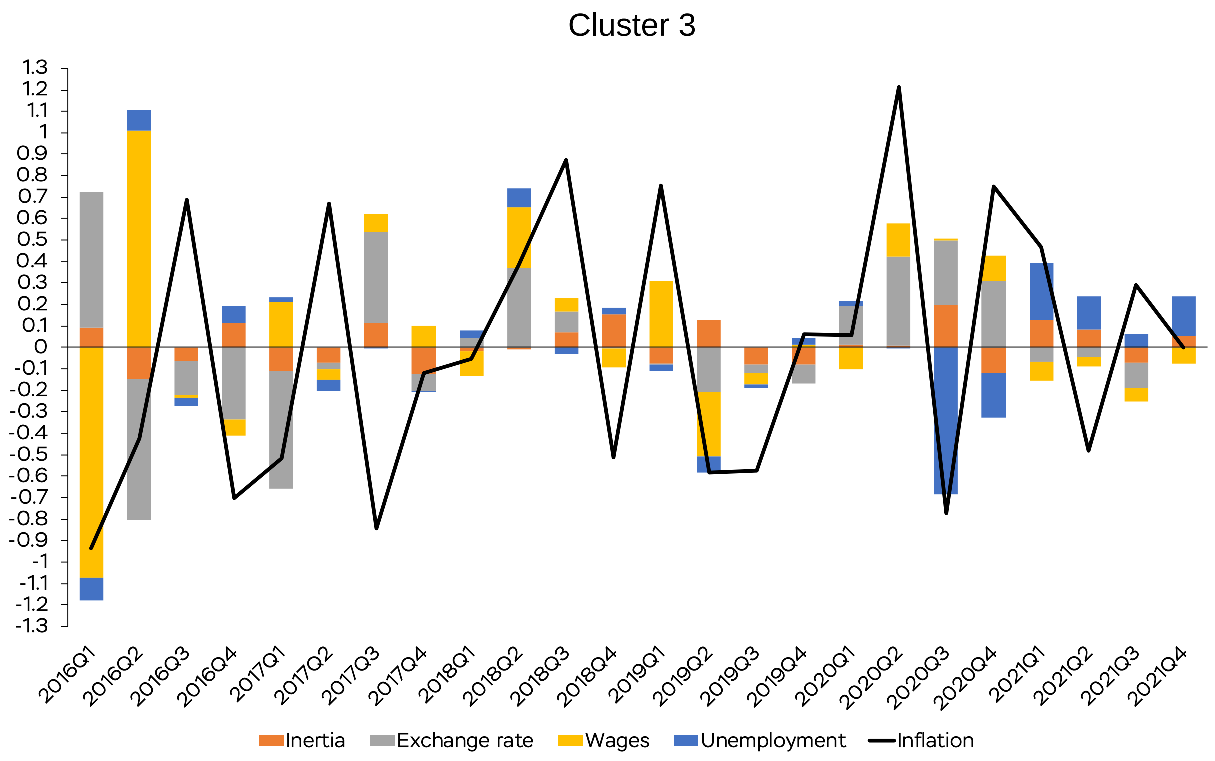
No	Variable	Description	Observations	Average	Standard deviation	Minimum	Maximum
Russia-wide							
1.	*CPI*	Consumer price index (% on previous period)	34	101.57	1.36	100.12	108.1
2.	*Neer*	Index of the nominal effective exchange rate of the ruble to foreign currencies (% growth on previous period)	34	-1.37	7.35	-18.8	20
3.	*Wage*	Average monthly real wages accrued (% on previous period)	33	100.38	8.76	82.09	113.12
4.	Unempl	Unemployment rate (%)	33	5.28	0.49	4.41	6.35
5.	D_{15q1}	Dummy variable for 2015 Q1	34	0.03	0.17	0	1
6.	D_{15q2}	Dummy variable for 2015 Q2	34	0.03	0.17	0	1
7.	D_{17q3}	Dummy variable for 2017 Q3	34	0.03	0.17	0	1
8.	D_{20q1}	Dummy variable for 2020 Q1	34	0.03	0.17	0	1
Cluster 1							
9.	*CPI*	Consumer price index (% on previous period)	34	101.45	1.19	100.26	107.19
10.	*Wage*	Average monthly real wages accrued (% on previous period)	33	100.35	9.66	84.73	113.28
11.	*Unempl*	Unemployment rate (%)	33	5.43	0.62	4.25	6.50
Cluster 2							
12.	*CPI*	Consumer price index (% on previous period)	34	101.58	1.39	100.16	108.28
13.	*Wage*	Average monthly real wages accrued (% on previous period)	33	100.43	8.82	81.69	113.62
14.	*Unempl*	Unemployment rate (%)	33	4.87	0.46	4.04	5.84
Cluster 3							
15.	*CPI*	Consumer price index (% on previous period)	34	101.55	1.72	99.29	109.65
16.	*Wage*	Average monthly real wages accrued (% on previous period)	33	100.62	9.44	79.52	115.70
17.	*Unempl*	Unemployment rate (%)	33	13.42	1.51	11.52	16.92
Cluster 4							
18.	*CPI*	Consumer price index (% on previous period)	34	101.68	1.37	100.07	107.56
19.	*Wage*	Average monthly real wages accrued (% on previous period)	33	99.99	7.62	83.12	112.05
20.	*Unempl*	Unemployment rate (%)	33	3.90	0.47	3.22	4.85

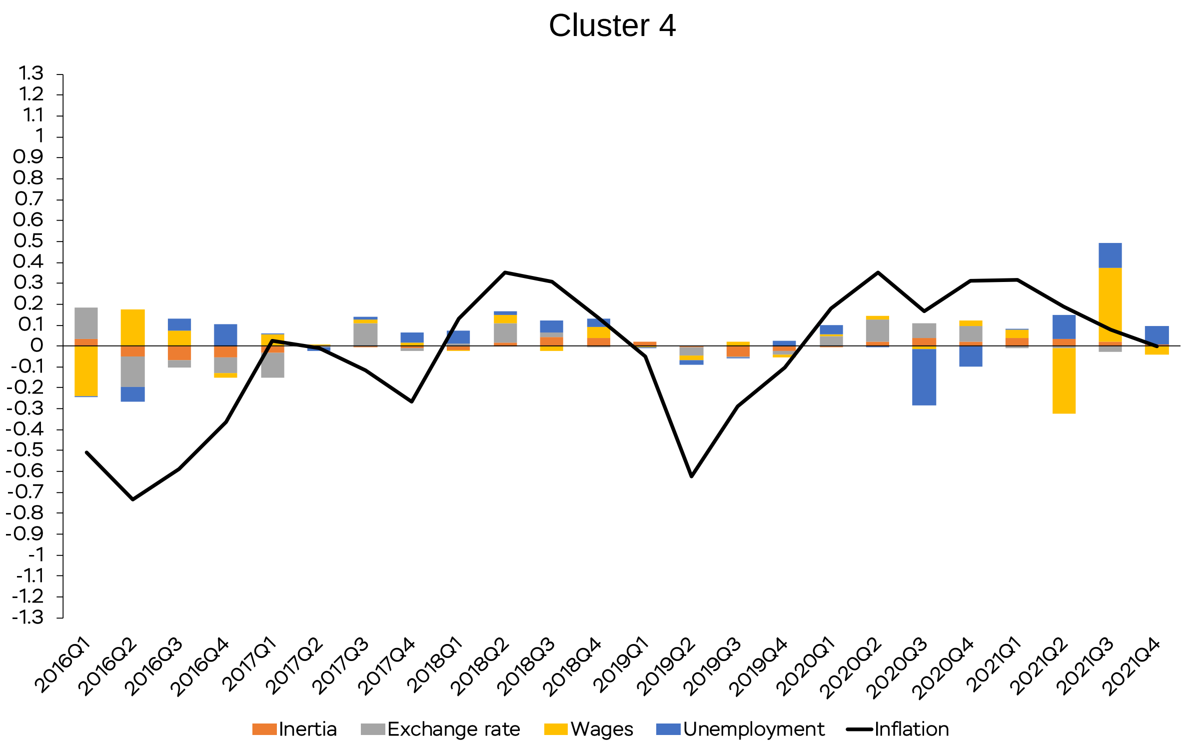
Annex 5. Decomposition of inflation factors in regression models

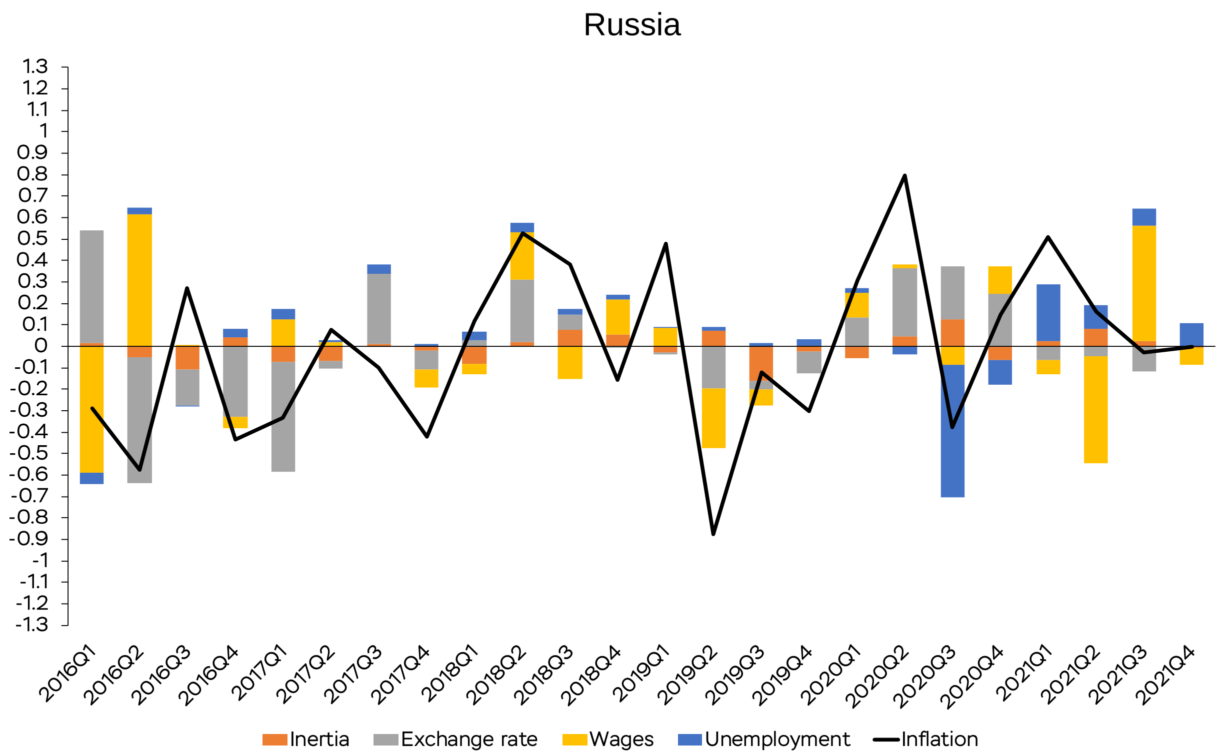
* The charts exclude dummy variables, the autoregression component and the impact of unconsidered factors to better visualise key factors.






Cluster 3
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
-1
-1.1
-1.2
-1.3
2016Q1
2016Q2
2016Q3
2016Q4
2017Q1
2017Q2
2017Q3
2017Q4
2018Q1
2018Q2
2018Q3
2018Q4
2019Q1
2019Q2
2019Q3
2019Q4
2020Q1
2020Q2
2020Q3
2020Q4
2021Q1
2021Q2
2021Q3
2021Q4
Inertia
Exchange rate
Wages
Unemployment
Inflation


Cluster 4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
-1
-1.1
-1.2
-1.3
2016Q1
2016Q2
2016Q3
2016Q4
2017Q1
2017Q2
2017Q3
2017Q4
2018Q1
2018Q2
2018Q3
2018Q4
2019Q1
2019Q2
2019Q3
2019Q4
2020Q1
2020Q2
2020Q3
2020Q4
2021Q1
2021Q2
2021Q3
2021Q4
Inertia
Exchange rate
Wages
Unemployment
Inflation


Russia
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.7
-0.8
-0.9
-1
-1.1
-1.2
-1.3
2016Q1
2016Q2
2016Q3
2016Q4
2017Q1
2017Q2
2017Q3
2017Q4
2018Q1
2018Q2
2018Q3
2018Q4
2019Q1
2019Q2
2019Q3
2019Q4
2020Q1
2020Q2
2020Q3
2020Q4
2021Q1
2021Q2
2021Q3
2021Q4
Inertia
Exchange rate
Wages
Unemployment
Inflation