Determinants of price differences in Russian regions

1. Introduction

The problem of price differences between regions of one country is typical not only for Russia but also for other countries (US, euro zone both between countries and within countries). In the Russian economic literature this question is considered by Glushchenko (2010). Understanding the causes of price level differences in the Russian regions can help both in the modeling of inflation and the development of economic policy measures aimed at alignment of prices in the Russian Federation.

We want to determine the factors responsible for the differences in price levels and inflation rates between the regions of the Russian Federation.

This paper is organized as follows. Section 2 reviews theoretical literature. Section 3 describes data and variables. Section 4 presents empirical results. Section 5 concludes.

2. Theoretical background. Factors of the regional price differences

According to the law of one price the same goods in different regions should be sold at the same price. However, there are deviations from this law, almost in regions of the same country. Possible causes of the violation of this law are represented in numerous theoretical and empirical studies.

The main theoretical foundation to explain the regional differences in the aggregate price level is the Balassa-Samuelson effect (Balassa, 1964), (Samuelson, 1964) without taking into account the nominal exchange rate, because it is the same for all regions. Under this model, the explanation of price differences in the regions with the same currency is based on the discrepancy of the relative prices of non-tradable goods in these regions.

The mechanism of Balassa-Samuelson effect for the regional differences in the aggregate price level can be described as follows. Technological progress is developing actively in the tradable sector due to inter-regional competition and the rapid modernization of the production of tradable goods. If there is a positive shock in the tradable sector performance in one of the regions, the wages in this sector increased. Following this, wages and production in non-tradable sector also increase, so that not all employees are moved to the tradable sector. Wage growth will be accompanied by a rise in prices. So price of non-tradable goods rise, which will lead to an increase in the aggregate price level in the region where there was a shock performance.

Balassa-Samuelson effect explains why overall price levels in the highly developed regions with a high level of performance, higher than in the relatively less successful.

There are many other factors leading to price differentiation in regions. In (Duarte, Wolman, 2008) the price differences between regions are explained by differences in fiscal policy. The overall price level will be higher in the region in which the regional fiscal policy stimulates aggregate demand.

Different phases of the business cycle is an important reason for price differentiation.

Another factor of different levels of consumer prices is the heterogeneity of consumption of traded goods in different regions (see Altissimo et al., 2005). The share of consumption of tradable goods in the total consumption basket may vary between regions. This will lead to regional price differences in response to the common macroeconomic shocks.

Other factors of regional price differences are rigidities in the labor market: limited labor mobility, differences in the level of wages and unemployment benefits; and imperfections in the retail sector. The less would be the level of competition in the retail sector in some regions, the greater the incentives for sellers to inflate trade margins of final goods. It will increase the aggregate price level in these regions.

In (Marques et al., 2014) the main factor of regional price differences is the transport costs. Change in transportation costs will lead to an asymmetric change in prices in the regions, if one of them produced most of the goods. This is because the inhabitants of the other regions will have to pay for the increased cost of transporting goods traded. Thus, in the model (Marques et al., 2014) regional differences in overall price levels are explained by the uneven distribution of final goods producers across the country, and as a consequence, arise differences in cost of transportation of goods.

Finally, regional price levels can be affected by factors such as inter-regional trade barriers: in the regions, which imposed restrictions on import or export of goods, prices will be higher than in others with no restriction on trade. In addition, various prices of non-tradable resources used in the production of final goods may be one of these factors.

Table 1 summarizes determinants of regional price level differentials. These factors are divided into two groups: operating on the aggregate demand and supply side.

Table 1. Determinants of regional price level

Aggregate demand (AD)	Aggregate supply (AS)
Balassa-Samuelson effect *Altissimo et al. (2005)*	Transportaition costs *Marqueas et al. (2014)*
Asynchronous business cycles *Andersson et al. (2009)*	Competition in retail sector *Andress et al. (2008)*
Regional fiscal policy *Duarte, Wolman (2008)*	Labor market imperfections *Campolmi, Faia (2011)*
	Regional input costs *Crucini et al. (2005)*
	Regional economic structure *Krugman, Venables (1996)*

3. Data description

Before we analyze the reasons for differences in the level of consumer prices in the Russian regions we need to understand the scale of the phenomenon under investigation. Data of annual growth rate of the consumer price index (CPI) has been collected from the Rosstat[1] by region from 1992 to 2015. Then we composed the base CPI. The base year was 1991, in which the prices were equal for all regions. Aggregate price level rose about 34 thousand times during 1992-2015.

Changes in prices during this period was very heterogeneous, the differences in the accumulated CPI between the most "expensive" and the "cheapest" regions reached 10 times in 2015. Possible reasons are:

a) in 1991 the prices in the regions were not equal, so it is incorrect to use the same initial value of the price index in all the regions;

b) there were errors in the measurement of CPI especially in the early stages, as a result of the accumulation of errors the real picture can significantly distorts;

c) in the regions prices are rising at different rates in the long term, which corresponds to the different average inflation rate.

Another way to measure the differences in price levels between regions is to calculate the cost of a fixed basket of goods and services, conducted by the State Statistics Service. Data on this indicator are available at the regional level since 2000. According to information for 2015 this basket included the set

[1] www.gks.ru

of 30 food, 42 non-food goods and 12 types of services. On average in Russia the cost of fixed set of goods and services was 2254 rubles in 2000, and 13404 rubles in 2015, which corresponds to an increase in about 6 times.

According to this indicator, the level of prices between the most expensive and the cheapest region of the Russian Federation in 2015 was distinguished by two times, which more corresponds to real picture. In addition, the use of the cost of a fixed basket of goods and services to solve the problem of the initial value of the underlying index for the regions.

In our empirical analysis we will consider dynamic of the relative aggregate price level by region (measured as the ratio of base CPI in the region to the base Russian CPI) in the period from 2000 to 2015. The choice of this period can be explained that in the mid-1990s, there were significant changes in prices caused by the transition from planned to market economy, which greatly complicates the correct evaluation of the impact of the factors identified in the theoretical papers. Period 1998-1999 is omitted as after a default in August 1998 there was a significant depreciation of the national currency. As a result of the pass-through effect the price level doubled during 1998-1999. Thus, we can assume that there was a structural change in the statistics due to the 1998 crisis. If we include this period in our sample, along with a relatively stable 2000s interval, it may cause bias in the coefficient estimates.

In econometric modeling a problem arises in the selection of appropriate variables which can approximate the selected factors. As dependent variable we will use the ratio of base CPI in the region to the Russian price index. This variable will be denoted in the future by p. 2000 is selected as the base year. However, it is clear that price levels in 2000 differed between the Russian regions, so not quite correct to normalize the overall price level to a unit for each region in 2000. We use a different base value for the different regions to resolve this problem. The initial value was determined based on the ratio of a fixed basket of goods and services in the region in 2000 compared to the same indicator for the Russian economy as a whole. For example, the cost of a fixed basket of goods and services in the Vladimir region was 1788 rubles in 2000, and in the Russian economy - 2254, so the initial value of the relative aggregate price level for the Vladimir region is 1788/2254 = 0.79, instead of 1.

All explanatory variables are also presented as ratio of regional value to the Russian average. Thus, they show how the value of the indicator in a separate region is distinguished from the mean value of the Russian economy at any given time. Therefore, logarithm of the variable multiplied by 100% will characterize the percentage deviation of the regional factor from the Russian average.

The Balassa-Samuelson effect we approximate by per capita income in the region. The logic of using this indicator as follows. If the regional income is higher than the national average, then price level in non-tradable sector is higher in this region. As a result, we observe the excess of the price level in this

region over the national average. We assume that the coefficient of this variable, which is denoted *inc*, will be positive.

To account for differences in the phases of the business cycle was used index of physical volume of gross regional product (GRP), hereinafter designated *gr*. It is assumed that, if the growth rate of GRP exceeds the rate of growth of the Russian GDP, due to a higher level of demand in the region, then the regional price level will be higher than national.

We use the ratio of retail trade turnover via street markets to total retail trade volume as another indicator. This variable can help to approximate the level of competition in the sector of tradable goods, denoted by *comp*. It is assumed that if in some of the regions the proportion of retail sales by markets and fairs is higher than in the Russian average, the price level in the region is lower than the average in the Russian economy.

To account for regional differences in the prices of labor resources we use the unemployment rate in the region, *un*. If the unemployment rate is higher in some of the regions compared with the Russian level, it can lead to a slowdown in wage growth in the region, leading to a reduction in regional price levels as compared to the national.

Prices of non-tradable intermediate goods used in the manufacture of the final product, we approximate by the producer price index of electricity, gas and water, designated *inp*. The rise in price of electricity, gas and water in region, on large compared to the average level value should lead to a more rapid increase in the price level in the region.

The differences in regional economy structure can be approximated by using the share of the services sector in the GRP, in the future will be referred to *serv*.

Transport costs are usually approximated by the distance between regions. In our study *normdist* variable will be used, which is calculated as follows. We determine the distance from the administrative center of the region to the other regional capitals by paved road and calculate a simple average. The obtained value is divided by the average of all regional distances. It is assumed that the region that is located farther from the others has the higher transportation costs, and hence higher prices. With the growth of variable *normdist* relative price should increase. It is worth noting that this variable takes the same values at each time point.

Legend of the original variables and their brief description are represented in Table 2.

Table 2. Variables legend

Notation	Description
p	The ratio of base consumer price index in the j-th region to the Russian average CPI, 2000 as the base, the base for each region adjusted for the cost of a fixed basket of goods and services in 2000
inc	The ratio of per capita income in the region to the Russian average level
comp	The ratio of the proportion of retail sales in the markets and fairs in the region similar to the average figure
un	The ratio of unemployment in the region to the average indicator
inp	The ratio of the base of the index of producer prices of electricity, gas and water in the same average figure region
serv	The ratio of the share of the services sector in the region's GRP in the share of services sector in Russian GDP
normdist	The ratio of the average distance from the center of the region to the rest of the centers of the RF subjects by road to the average distance

4. Empirical results

In the next phase of the study, we will try to answer the question of what factors influence the differences in regional price levels at a certain time. For this purpose, we estimate the following equation (1) on the cross-section data (by region):

$$\log(p_i) = \beta_0 + \beta_1 \log(inc_i) + \beta_2 \log(un_i) + \beta_3 \log(comp_i) + \beta_4 \log(serv_i) + \\ + \beta_5 \log(normdist_i) + \beta_6 \log(serv_i) + \varepsilon_i, i = 1, \dots 77. \quad (1)$$

Estimation is carried out for each year, so it turns 16 estimated equations. Brief results for each year are shown in Tables 3-5

Table 3. The estimation results for 2000-2005 years

	y_2000	y_2001	y_2002	y_2003	y_2004	y_2005
(Intercept)	0.044***	0.049***	0.049***	0.060***	0.069***	0.071***
	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.012)
log(inc)	0.245***	0.241***	0.267***	0.262***	0.272***	0.266***
	(0.019)	(0.020)	(0.025)	(0.024)	(0.025)	(0.031)
log(un)	0.005	-0.005	-0.006	-0.019	0.001	-0.003
	(0.025)	(0.021)	(0.024)	(0.022)	(0.021)	(0.025)
log(comp)	-0.042*	-0.031	-0.029	-0.022	-0.024	-0.017
	(0.017)	(0.017)	(0.020)	(0.019)	(0.018)	(0.018)
log(serv)	0.021	0.026	0.034	0.002	0.032	0.036
	(0.029)	(0.035)	(0.045)	(0.042)	(0.035)	(0.040)
log(normdist)	0.188***	0.185***	0.158***	0.170***	0.149***	0.172***
	(0.024)	(0.024)	(0.028)	(0.027)	(0.025)	(0.029)
log(inp)			-0.365	-0.392*	-0.383*	-0.407*
			(0.194)	(0.192)	(0.184)	(0.196)
R-squared	0.9	0.8	0.8	0.8	0.8	0.8
adj. R-squared	0.9	0.8	0.8	0.8	0.8	0.8
F	88.4	76.1	51.7	59.5	56.4	51.0
p	0.0	0.0	0.0	0.0	0.0	0.0
N	77	77	74	74	74	74

As expected at each time relative income and transportation costs have a significant positive coefficient. Table 3 shows that in 2000-2005, the coefficient of the variable *inc* varies in a very narrow range of 0.24 to 0.27, which confirms the hypothesis about the stability of relationship between Balassa-Samuelson effect and price level. The coefficient of the variable *normdist* of the region varied from 0.15 in 2004 to 0.19 in 2000. The coefficient of determination over time falls in regression equations. This may indicate that the regional differences in the prices begin to occur under the influence of new, unrecorded in the regression equation factors. The variable level of competition in the retail sector is significant at the 5% in only one regression equation for the year 2000 and has a negative sign as predicted by theoretical models

Table 4. The estimation results for 2006-2011 years

	y_2006	y_2007	y_2008	y_2009	y_2010	y_2011
(Intercept)	0.075***	0.071***	0.068***	0.078***	0.083***	0.091***
	(0.012)	(0.012)	(0.013)	(0.013)	(0.014)	(0.015)
log(inc)	0.298***	0.272***	0.292***	0.320***	0.354***	0.384***
	(0.030)	(0.032)	(0.040)	(0.041)	(0.044)	(0.048)
log(un)	0.007	-0.008	-0.015	-0.016	0.030	0.049
	(0.022)	(0.021)	(0.026)	(0.038)	(0.035)	(0.037)
log(inp)	-0.399	-0.325	-0.192	-0.170	-0.269*	-0.295*
	(0.200)	(0.181)	(0.137)	(0.128)	(0.130)	(0.129)
log(comp)	-0.012	-0.012	-0.007	-0.007	-0.007	-0.008
	(0.019)	(0.016)	(0.018)	(0.016)	(0.016)	(0.015)
log(serv)	0.044	0.023	0.041	0.031	0.017	0.043
	(0.039)	(0.041)	(0.046)	(0.047)	(0.048)	(0.046)
log(normdist)	0.170***	0.161***	0.176***	0.172***	0.161***	0.168***
	(0.027)	(0.029)	(0.033)	(0.030)	(0.030)	(0.031)
R-squared	0.8	0.8	0.7	0.8	0.8	0.8
adj. R-squared	0.8	0.8	0.7	0.8	0.7	0.7

F	49.9	40.6	32.3	38.4	34.6	34.5
p	0.0	0.0	0.0	0.0	0.0	0.0
N	74	74	74	74	74	74

Estimation of cross-sectional regression for the 2006-2011 period (the results are presented in Table 4) fully confirm the findings previously obtained. Noteworthy coefficient of variable *inc* increased from 0.27 in 2007 to 0.38 in 2011. At the same time, there is a decrease the range of fluctuations of the coefficient of the variable *normdist*. The coefficient of determination largely stable and stands at 80% with the exception of the crisis of 2008. It is said that with the help of used factors can explain about 80% of the variation in the aggregate price level in the Russian regions.

Table 5. The estimation results for 2012-2015 years

	y_2012	y_2013	y_2014	y_2015
(Intercept)	0.088***	0.094***	0.081***	0.065***
	(0.015)	(0.015)	(0.015)	(0.015)
log(inc)	0.362***	0.382***	0.361***	0.349***
	(0.051)	(0.050)	(0.053)	(0.054)
log(un)	0.032	0.052	0.032	0.025
	(0.033)	(0.036)	(0.036)	(0.040)
log(inp)	-0.284*	-0.290**	-0.239*	-0.206*
	(0.120)	(0.109)	(0.100)	(0.097)
log(comp)	0.002	0.003	-0.003	-0.003
	(0.013)	(0.012)	(0.011)	(0.010)
log(serv)	0.023	0.033	0.057	
	(0.047)	(0.045)	(0.045)	
log(normdist)	0.187***	0.177***	0.178***	0.182***
	(0.034)	(0.036)	(0.038)	(0.038)
R-squared	0.7	0.8	0.7	0.7
adj. R-squared	0.7	0.7	0.7	0.7
F	30.9	33.7	29.7	31.6
p	0.0	0.0	0.0	0.0
N	74	74	74	74

Table 5 presents the results of the cross-section regression on the last period. Qualitatively, they are not differ from those obtained with estimation of regional data in previous years. As before, the region with a high level of income per capita demonstrates higher price level. A similar result is observed in the region, which is more distant from the rest. It continues to gradually decrease the coefficient of determination in the regression equation. In 2015, with the help of used variables can explain about 70% of the variation of the aggregate price level between the Russian regions.

Since significant changes in the value of coefficients from year to year is not observed, we can try to combine the cross-section sampling of the various years and to estimate the pool model. It can significantly increase the accuracy of the estimated coefficients. Another way to improve the accuracy of the estimated coefficients is to add a fixed time effects. Their inclusion may be due to the influence of the common factors (variables which change is the same for all regions, for example, the dynamics of the

exchange rate, interest rate or money supply), a change in which regional prices may react differently. We estimate different versions of equation (2):

$$\log(p_{it}) = \beta_0 + \gamma_t + \beta_1 \log(inc_{it}) + \beta_2 \log(un_{it}) + \beta_3 \log(comp_{it}) + \beta_4 \log(serv_{it}) + \\ + \beta_5 \log(normdist_{it}) + \beta_6 \log(serv_{it}) + \varepsilon_{it}, i = 1, \dots 77, \mathrm{t} = 2000, \dots, 2015, \quad (2)$$

To obtain consistent estimates of the specification (2) initial variables must be averaged at each point of time for all regions, we estimate equation (2) in their deviation from the average over time.

The results of estimation pool model and the model with fixed time effects are shown in Table 6.

Table 6. Estimation results of pool and panel model

	pool1	pool2	fix_time1	fix_time2
(Intercept)	0.071***	0.062***	0.071***	0.062***
	(0.003)	(0.003)	(0.003)	(0.003)
log(inc)	0.294***	0.281***	0.295***	0.281***
	(0.009)	(0.008)	(0.009)	(0.008)
log(un)	0.000	-0.009	0.001	-0.008
	(0.007)	(0.006)	(0.007)	(0.007)
log(inp)	-0.236***		-0.240***	
	(0.035)		(0.035)	
log(comp)	-0.009*	-0.014***	-0.009*	-0.014**
	(0.004)	(0.003)	(0.004)	(0.004)
log(serv)	0.024*	0.020*	0.025*	0.020
	(0.011)	(0.010)	(0.012)	(0.010)
log(normdist)	0.175***	0.176***	0.174***	0.175***
	(0.008)	(0.007)	(0.008)	(0.007)
R-squared	0.8	0.8	0.8	0.8
adj. R-squared	0.8	0.8	0.8	0.8
sigma	0.1	0.1	0.1	0.1
F	538.4	772.6	178.3	202.1
p	0.0	0.0	0.0	0.0
Log-likelihood	1158.7	1403.1	1161.5	1406.2
Deviance	5.1	6.0	5.0	5.9
AIC	-2301.4	-2792.2	-2375.3	-2400.3
BIC	-2262.4	-2756.8	-2338.7	-2312.8
N	962	1155	962	1155

Estimation accuracy increased. This led to the fact that all variables became statistically significant. As before, the coefficients of the variables *inc* and *normdist* are statistically significant. Let us give an economic interpretation of the coefficients in front of these variables based on the model with fixed effects for time specifications without variable *inp*.

The coefficient of the variable relative per capita income in the region (*inc*), equal to 0.28, indicates that if the region *j* per capita gross regional product will be higher than in the region *i* 1%, the price level in region *j* will be on 0.28% higher than in region *i*.

The coefficient of the variable remoteness of the region relative to the others (*normdist*), equal to 0.18, indicates that if the region *j* will be more distant on 1%, than the region *i*, the price level in region *j* will be on 0.18% higher than in region *i*.

Thus, inter-regional variation in prices can be explained by two factors: income and distance. At the same time, the regional per capita income can be influenced by means of economic policy: reducing income distribution inequality can achieve greater uniformity in the price level. In addition, on the geographical distance factor to influence directly by means of economic policy is impossible. However, the development of infrastructure and reduce the cost of transport of goods may reduce regional variations in prices.

5. Conclusions

We analyzed the extent of price differences in the Russian regions and the factors causing these differences, it yielded the following conclusions:

a) by 2015, the cost of living in the cheapest and most expensive region has about 2 times;

b) the main reasons causing the inter-regional differencies of the regional prices are differences in per capita income between the regions and the degree of remoteness of the region from the rest; regions with large values of these variables are characterized by higher price levels.

Thus, the price variation between regions can be decreased by means of economic policy aimed at alignment of regional per capita income and infrastructure development providing cost savings for cargo transportation.

6. References

1 Altissimo F., Benigno P., and Rodriguez Palenzuela D., "Long-run determinants of inflation differentials in a monetary union," NBER Working Paper, No. 11473, 2005.

2 Andersson, M., Masuch, K., Schiffbauer, M., "Determinants of inflation and price level differentials across the Euro area countries," ECB Working Paper, No. 1129, December 2009

3 Andrés, Javier, Eva Ortega, and Javier Vallés. "Competition and inflation differentials in EMU." Journal of Economic Dynamics and Control 32.3 (2008): 848-874.

4 Balassa B., "The Purchasing Power Parity Doctrine: A Reappraisal," Journal of Political Economy, Vol. 72, 1964. pp. 584-596.

5 Campolmi, Alessia, and Ester Faia. "Labor market institutions and inflation volatility in the euro area." Journal of Economic Dynamics and Control 35.5 (2011): 793-812.

6 Crucini M., Telmer C., and Zachariadis M., "Understanding European real exchange rates," American Economic Review, Vol. 95, No. 3, 2005. pp. 724-738

7 Duarte M., Wolman A., "Fiscal policy and regional inflation in a currency union," Journal of International Economics, No. 74, 2008. pp. 384–401.

8 Gluschenko, Konstantin. "Anatomy of Russia's market segmentation1." Economics of Transition 18.1 (2010): 27-58.

9 Krugman, Paul, and Anthony Venables. Integration, specialization, and the adjustment. No. w4559. National Bureau of Economic Research, 1993.

10 Marques H., Pino G., and Horrillo J., "Regional inflation dynamics using space–time models," Empirical Economics, Vol. 47, No. 3, 2014. pp. 1147-1172.

11 Samuelson, P. "Theoretical notes on trade problems." The Review of Economics and Statistics (1964): 145-154.