CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Bachelor thesis

2019 **Jan Provazník**

CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Jan Provazník

Analysis of Unemployment in Russia: Spatial Analysis on Russian Regions

Bachelor thesis

Prague 2019

Author: Jan Provazník

Supervisor: doc. Ing. Tomáš Cahlík, CSc.

Academic Year: 2018/2019

Bibliographic note

PROVAZNÍK, Jan. *Analysis of Unemployment in Russia: Spatial Analysis on Russian Regions*. 47 p. Bachelor thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Supervisor doc. Ing. Tomáš Cahlík, CSc.

Abstract

This thesis analyses the regional labor markets of Russian Federation in terms of unemployment and Okun's law. Using Hodrick-Prescott filter, the potential GRP and natural rate of unemployment are calculated in order to be used for the estimation of Okun's law coefficient (OLC). The analysis reveals variation in OLC between regions as well as significant spatial correlation thereof (measured by Moran's I). For the examination of Okun's convergence, a new method is developed which consists in estimating a time trend in R-squared obtained by estimating Okun's law cross-sectionally, separately for each year. The results indicate that there is no regional Okun's convergence in Russia.

Abstrakt

Tato práce zkoumá pracovní trhy v regionech Ruské federace s ohledem na nezaměstnanost a Okunův zákon. Pomocí Hodrick-Prescottova filtru jsou spočteny hodnoty potenciálního HRP a přirozené míry nezaměstnanosti a jsou poté použity k odhadnutí Okunova koeficientu. Analýza odhaluje nestejnost Okunova koeficientu pro jednotlivé regiony a významnou prostorovou korelaci v něm (měřenou pomocí Moranova I). Za účelem vyšetření Okunovy konvergence je vyvinuta nová metoda, která spočívá v měření časového trendu v koeficientech determinace obdržených odhadováním Okunova zákona na průřezových datech pro každý rok zvlášť. Výsledky ukazují absenci regionální Okunovy konvergence v Rusku.

Klíčová slova

Nezaměstnanost, Okunův zákon, konvergence, prostorová ekonometrie, regiony Ruska

Keywords

Unemployment, Okun's Law, Convergence, Spatial Econometrics, Russian Regions

Range of thesis: 54 628 symbols.

Declaration of Authorship

1. The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

2. The author hereby declares that all the sources and literature used have been properly cited.

3. The author hereby declares that the thesis has not been used to obtain a different or the same degree.

Prague, 9. 5. 2019 Jan Provazník

Acknowledgments

The author is grateful especially to doc. Cahlík for his advices regarding the subject, structure and methodology of the thesis and for the time he spent discussing the thesis on consultations.

Institute of Economic Studies
Bachelor thesis proposal

Proposed Topic:

Analysis of Unemployment in Russia: Spatial Analysis on Russian Regions

Preliminary scope of work:

Research question and motivation

The entire territory of Russia is divided into 83 administrative regions, most of them called oblasts in Russian. The thesis shall examine the GDP growth and unemployment in these regions and, using the methods of spatial econometrics, detect the spatial effect these regions have upon each other with respect to unemployment, growth of income and their relationship as expressed by Okun's law. More specifically, the task (or one of the tasks) shall be to find out whether the economic development in Russian regions shows convergence, to explore the nature of this convergence and to discuss cases of regions which do not follow the common trend.

Contribution

Unlike the existing literature on the topic which either focuses on spatial analysis of Russian regions[1][3][4][5] or the Okun's law in Russia as a whole[2], the thesis shall combine these two approaches. As there can be expected significant influence which the level of unemployment in one region has on the unemployment in other regions, the spatial effects in this matter seem to be likely. The knowledge of this effects could be useful in policy making as well as in business.

Methodology

The thesis shall work with data from the Russian Federal State Statistics Service and apply several types of spatial weight matrices to discover spatial effects between the regions of Russia, influencing the regional levels of unemployment.

Outline

1) Introduction
2) Literature review
3) Methodology
4) Data
5) Results
6) Conclusion

Bibliography

1. ELENA VYACHESLAVOVNA SEMERIKOVA a OLGA ANATOLYEVNA DEMIDOVA. Interaction of Regional Labour Markets in Russia: Spatial Econometric Analysis. *Prostranstvennaâ Èkonomika, Iss 3, Pp 57-80 (2016)* [online]. 2016, (3), 57-80 [cit. 2018-05-24]. DOI: 10.14530/se.2016.3.057-080. ISSN 18159834.

2. E. VAKULENKO a E. GURVICH. The Relationship of GDP, Unemployment Rate and Employment: In-depth Analysis of Okun's Law for Russia. *VOPROSY ECONOMIKI* [online]. 2015, **3** [cit. 2018-05-24]. ISSN edsrep.

3. DANILENKO, Tatiana, Olga DEMIDOVA a Marcello SIGNORELLI. Unemployment Clubs in Russian Regions. *Emerging Markets Finance* [online]. 2018, **54**(6), 1337-1357 [cit. 2018-05-24]. DOI: 10.1080/1540496X.2017.1281799. ISSN 1540496X.

4. ELENA VYACHESLAVOVNA SEMERIKOVA a OLGA ANATOLYEVNA DEMIDOVA. Analysis of Regional Unemployment in Russia and Germany: Spatial-Econometric Approach. *Prostranstvennaâ Èkonomika, Iss 2, Pp 64-85 (2015)* [online]. 2015, (2), 64-85 [cit. 2018-05-24]. DOI: 10.14530/se.2015.2.064-085. ISSN 18159834.

5. SIMONOVA, Marina V., Larisa V. SANKOVA, Farida I. MIRZABALAEVA, Dina Ye. SHCHIPANOVA a Vladimir E. DOROZHKIN. Assessment Problems and Ensuring of Decent Work in the Russian Regions. *International Journal of Environmental and Science Education* [online]. 2016, **11**(15), 7608-7626 [cit. 2018-05-24]. ISSN 13063065.

Contents

1. Introduction

Among the countless subjects, economic and other, one can inquire into while examining regional data (or country-level data – or any other geographically organized data, for that matter), some are general and come into one's mind almost automatically, the moment one begins to think about regional analysis. Firstly, one may ask: what are the linkages between the regions? How do they interact with each other? How much do the events occurring in one region influence what happens in others? Do regions that are close to each other have more in common than those separated by greater distance? Secondly, are the regions becoming more similar to each other, as time passes, or do they rather walk in different directions? Does there seem to be any unification among them? Upon closer examination, we may notice that these two types of questions are of the same nature, the difference being only the dimension they are concerned with: the first questioning the effect of spatial distance, the other being interested in the effect of time.

The regional labor markets of Russia have been studied from both of these perspectives. While we will discuss in great detail some of the most relevant pieces of literature on this topic in the next section, let us here describe in the most general manner the way these questions are usually addressed. There are several economic variables that can be used for the description of labor market. There is the unemployment and employment rate, wages and various descriptive statistics of it (average wage, median wage), labor productivity, we can also name some of the less easy-to-observe quantities such as human capital. Spatial econometrics provides simple tools for discovering the spatial relations among regions in terms of these variables. Some of those which we shall make extensive use of will be discussed in detail when appropriate. As to the question of convergence, one can, for instance, look at whether the variance of these variables decreases in time. If it does, we say that σ-convergence occurs. There are other definitions of convergence too, such as β-convergence, which can be applied in some cases, while in others (such as in this thesis, as shall become clear later) they do not make much sense.

In this thesis, we shall examine the regional labor markets from a more abstract point of view. Our task will not be to find out whether some quantity, or ratio of two quantities trend toward unification across the regions. The aim of this thesis will be to investigate whether the regional labor markets converge in the dynamic aspect of their nature. As a measure of this, we choose the relationship between unemployment rate and output as expressed by the well-known Okun's law.

The basic idea proposed by Okun (1962) is that in periods of increased growth of output, the unemployment rate decreases, and vice versa. While this relationship appears intuitively clear (almost necessary, one might say) and empirically usually does hold, it is worthwhile to note that not only does the magnitude of the relationship vary, but there can in some instances appear to be no relationship at all (see for example Moosa (1997), examples can also be found in our preliminary analysis in section 5.1 of this thesis). Quantification of this relationship provides description of dynamic behavior of regional markets in a very condensed form. As Perman and Tavéra (2007) explain, it "incorporates several fundamental structural parameters from the firms' optimal demand for labour, the macroeconomic production function and the labour force participation equation. As a result, […] the OLC may be considered as the net effect of several macro-economic structural parameters representative of the macro-economic behaviour of the country under examination and of the characteristics of the adjustment mechanisms lying behind the inverse relationship between output gaps and unemployment gaps over the business cycle."[1] The examples of where Okun's law is used abound. In macroeconomic models, the short run aggregate supply can be derived using the short run Phillips curve and Okun's law. The slope of the short run aggregate supply curve then determines the proportion of the economy's reaction to supply and demand shocks. These can, of course, be caused not only by external factors, but also by government's macroeconomic policies. The relationship's being different across regions would threaten to make the effects of macroeconomic policies different in every region – a situation which might not be fatal, but the expectation of which would no doubt enhance the effectiveness of government's policies.

It is in this context, that we remember what seems to be the main motivation of Perman and Tavéra's research (of which we shall talk in considerable detail later) conducted on countries of European Union – the question of determining the optimal currency area. Although the idea of Russian Federation, as a nation-state, not being a single currency area clearly would not receive any serious political consideration, it can still be interesting as a matter of a purely academic discussion. There are, after all, discussions of whether

[1] PERMAN, Roger & TAVÉRA, Christophe. (2007). "Testing for Convergence of the Okun's Law Coefficient in Europe." Empirica. 34. pp. 45-61. Quotation from page 3 in the paper itself.

the United States constitute an optimal currency area (see for example Kouparitsas, 2001). Also, following the conclusions of previous paragraph and applying elementary macroeconomic theory, the question would arise – in the case of interregional variance and non-convergence of Okun's coefficient – of how the necessary region-specific macroeconomic policies should be organized.

The usual way of estimating Okun's law involves regressing time series of unemployment rate on time series of growth of output. The nature of the available statistical data would allow us only to calculate the Okun's law coefficient for the entire period (Russian official statistics provide yearly regional data for years 1997 to 2016), were we to follow this approach. In fact, we do execute this simple research and include it in subsection Preliminary analysis below. Since our aim is to discover not only the value of the Okun's law coefficient (in fact, this value is not a primary concern of this thesis), but mainly whether there is regional Okun's law convergence, we need to find a way to efficiently use the limited amount of data available in order to answer this question. The solution is to estimate Okun's law in a "perpendicular" dimension: not for each region over time (which would be a classical time-series approach), but for each time period (year in our case) over the cross-section of regions. In this method, newly developed for this thesis, the Okun's convergence is identified with the gradual unification of the regional labor markets along the regression line. If this unification occurs, the R-squared of the regression increases. Thus a statistically significant increasing trend in the obtained time series of R-squared's would be a sign of regional Okun's convergence. The empirical evidence found does not show any signs of such convergence in Russia.

The rest of this thesis is organized in the following way. After summarizing the available literature relevant to the subject of our research in section 2, we explain in section 3 in a general manner the econometrical methods we used, section 4 describes the data sources and makes some comments on the methods described previously when applied in particular cases. In section 5, we discuss the preliminary and final results of the empirical analysis and we summarize and conclude the whole thesis in section 6.

2. Literature Review

2.1 Okun's Law

Arthur M. Okun published his influential paper *Potential GNP: Its Measurement and Significance* in 1962. In it, he specified two versions of the negative relationship between unemployment and output, later named after him as "Okun's Law." Both are in use to this day. The "first differences" approach can be formulated as

$$U_t - U_{t-1} = \beta_0 + \beta_1(\log Y_t - \log Y_{t-1}). \tag{1}$$

where U_t is the unemployment rate in period t and Y_t is the output in period t. The other is called the "gap" version of the Okun's law:

$$U_t = \beta_0 + \beta_1 gapY \tag{2}$$

the "gap" being defined as $gapY = \log Y_t^{potential} - \log Y_t^{actual}$. First of the above specifications clearly has the advantage of only using observed values, whereas the potential output used in the second must be estimated. In this thesis, for reasons explained in section Methodology, we shall use a modified version of the second variant.

There have since been many studies estimating the relationship with varying methodologies and varying results. We will mention only those of them that are most relevant to the goal of our research. Imad A. Moosa (1997) estimates the "gap" version of Okun's law for G7 countries to discover that the Okun's law coefficient (hereafter OLC) differs between the countries (with North American countries having the highest and Japan the lowest values) as well as it changes over time. Setting $U_t^c = U_t^{natural} - U_t^{actual}$, Moosa's model can be written as

$$U_t^c = \beta_0 + \beta_1 U_{t-1}^c + \beta_1 gapY_t + u_t, \tag{3}$$

u_t being an error term.

Sögner and Stiassny (2002) conducted a similar survey for 15 OECD countries and obtained analogous results. The OLC estimated by them differed from country to country as well as in time. Finally, there is a paper by Jim Lee (2000) who investigated the development of Okun's law in 16 OLCD countries in the period 1955-1996. Lee compares the two specifications of Okun's law listed above as well as several ways of obtaining data for the second one and reveals substantial differences between individual countries' OLC.

Analogical findings have also been brought by research on regional level. Blackley (1991) measured the OLC on the state level in USA, Apergis and Rezitis (2003) investigated the OLC in Greek regions and Adanu (2005) estimated the OLC on Canadian provinces. All of them reported varying value of OLC on the regional level.

2.2 Okun's Convergence

The growing literature reporting the spatial and temporal fluctuation of Okun's relationship brought Perman and Tavéra (2005)[2] to the idea of whether the differences in OLC tend to decrease in some areas, that is, to investigate the Okun's convergence. As their paper is one of the very few pieces of academic literature aimed at this question, and as it influenced the content of this thesis more than any other piece of literature, we shall review their method in greater detail now.

Perman and Tavéra estimate the OLC for 17 countries of European Union over the period 1970 to 2002. In order to obtain for each country a time series of OLC that could be tested for convergence to other countries' OLC, the method of rolling regression is used. The Autoregressive Distributed Lag model

$$U_t^c = \sum_{s=0}^{p}(a_{0,s}Y_{t-s}^c) + \sum_{s=1}^{p}(a_{1,s}U_{t-s}^c) + u_t \tag{4}$$

$$b = (a_{0,0} + \cdots + a_{0,p})/(1 - a_{1,1} - \cdots - a_{1,q}) \tag{5}$$

is first estimated for the first 40 observations in their sample (they use semestrial data), then one observation on the beginning of the time series is dropped and one is added on the end and the estimation is made on this new subsample. This procedure is repeated until the last available observation is reached, thus providing a series of OLC estimates. Equation (5) describes the calculation of the "total effect" (the final OLC estimate). Let us now denote the number of a country in the sample as $i \in \{1, \ldots, 17\}$ and the time period

[2] In fact, Perman and Tavéra published two papers of very similar content in 2005 and 2007. We shall refer to the earlier one hereafter throughout the thesis, as it seems to contain everything from the later one too.

as $t \in \{1, \dots, T\}$ (the number T depends on the number of time lags used in the model. Perman and Tavéra eventually never used more than four semestrial lags, which is followed in this thesis, as shall be explained later – in order not to shorten our already short time series, more than for any other reason). A series of cross-regional variances in the OLC estimates is calculated as follows.

$$V_t = \frac{1}{17}\sum_{i=1}^{17}\left(b_{t,i} - \overline{b_t}\right)^2, \ t = 1, \dots, T \tag{6}$$

Finally, this series is examined for the presence of time trend by estimating the model

$$V_t = \alpha_0 + \alpha_1 t + \varepsilon_t, \ t = 1, \dots, T \tag{7}$$

where $\varepsilon_t, t = 1, \dots, T$ is a white noise process. A statistically significant negative estimate of the coefficient α_1 signifies a decreasing trend in cross-sectional variances in the OLC and therefore Okun's convergence. Perman and Tavéra investigated the convergence for the whole group of 17 countries as well as for specific sub-groups of countries associated according to various criteria (convergence clubs). Tre results showed "convergence of the OLC among northern European countries, and among countries with centralized wage bargaining, but an absence of convergence in other country groups."

Besides Perman and Tavéra (2005), there is only one other paper examining the issue of Okun's convergence. Written by Yazgan and Yilmazkuday (2009), it analyzes the OLC convergence in the United States over the period 1978-2002. Using in some respects different methodology from the one of Perman and Tavéra (2005), they find evidence of time-varying OLC, convergence of state-level OLC, the presence of convergence clubs, and the geographical nature of these (as apparent from the previous paragraph, the potential convergence clubs can be based upon other criteria than geographical distance or adjacency).

2.3 *Unemployment and Okun's Law in Russia*

The regional labor markets in Russia have been studied mainly by Elena Semerikova and Olga Demidova. In their first paper, Demidova and Semerikova (2015) analyze the spatial interactions of the regional labor markets in Russia and Germany. The goal of their research was the formulation of a general criterium for choosing an appropriate spatial weight matrix (for definition and discussion of this term, see section Methodology, where it is used extensively). The answer is found in the examination of the statistical significance of spatial correlation coefficients in various spatial models. The results show that the inverse distance matrix is more accurate for Russia than the adjacency matrix. As explained below, the methodology of this thesis prevents the application of this criterium, so we evaluate the spatial weight matrices in a slightly different (but analogous) way.

In their second paper, Demidova and Semerikova (2016) investigate – now in the context of Russian regions exclusively – the effect of unemployment's change in one region upon the change of unemployment in neighboring regions. That is, the study did no longer examine the spatial correlation in the rate of unemployment, but the spatial correlation in its first difference. The results showed positive spatial correlation. Besides that, they also studied other factors that are not very relevant to the subject of this thesis, such as the influence of education on unemployment rate and wages.

To turn our attention to regional convergence of Russian labor market, Vakulenko and Guriev (2012) conducted an extensive study, in which they found no convergence in 1990s but significant convergence in 2000s. The variables in whose terms the convergence took place included unemployment rate, wages, GDP per capita and capital flow. In a subsequent study, Vakulenko (2014) examined the impact of immigration on this regional convergence. Contrary to what might be expected based on a simple economic intuition, Vakulenko found no significant effect of interregional migration on unemployment rate.

Finally, let us briefly comment on the paper whose research field was the closest to the one chosen for this thesis. In a 2015 paper, Elena Vakulenko estimates the Okun's law on Russian regions in the period 1998-2013. The specification she chooses for her research is the first-difference version of the Okun's law. It shall also be noted that she uses spatial econometric methods to account for the mutual influence that the regional labor markets have upon each other. The results revealed substantial interregional spatial effects in terms of unemployment and GDP growth. Vakulenko concludes that without taking these effects into account, the OLC is underestimated.

As far as the Okun's convergence is concerned, no literature inquiring into this field in the context of Russian regions is available. As mentioned above, there are only two papers in the world studying this subject. After providing the context of the research conducted in this thesis, we shall now proceed to describe the method used in it.

3. Methodology

As mentioned above, the method which the author of this thesis considers to be the basic approach to the study of Okun's convergence – that is, the examination of a time trend in the cross sectional variances in the OLC's obtained by rolling regression, which was performed by Perman and Tavéra (2005) – cannot be used in the case of Russian regions. To compare once more our situation with the one in which the Perman and Tavéra's research was done, let us repeat that the sub-samples on which the OLC was estimated comprised 40 observations and there was 21 of them, rendering the series of cross-country variances upon which the trend estimation was performed 21 elements long. Our dataset, on the other hand, only consists of 20 time periods (thus we could only for example run the rolling regression on 10 observations and produce 11 estimates of OLC) which would seriously limit the robustness of our estimation. Therefore we instead choose to exploit the cross sectional dimension in which our data set is relatively much richer (there are 79 regions which were subjected to the analysis).

The conceptual construction of our approach shall be as follows. Looking at the regional markets in one particular year, we should observe that the regions whose gross regional product (GRP hereafter) is below the value of potential output, show unemployment rate above the natural rate of unemployment, and vice versa. By the Okun's law, we would also expect the relationship to be more or less linear: the farther the region is below the potential level of output, the farther it should be above the natural rate of unemployment. If the relationship between output and unemployment is similar or the same for every region, the regions should be very densely gathered around the imaginary regression line, as illustrated in Figure 1 below. If the dynamics of each region's labor market is different, the scatter plot should look as in Figure 2. The Okun's convergence would then consist in a gradual shift from Figure 2-like situation towards the pattern illustrated in Figure 1. Mathematically, the difference is captured by an increase in the regression's R-squared.

Figure 1.

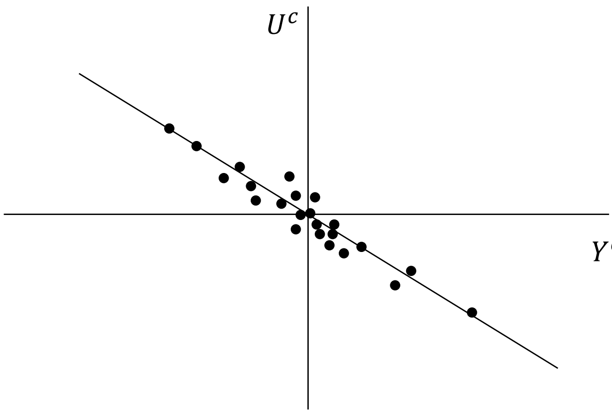
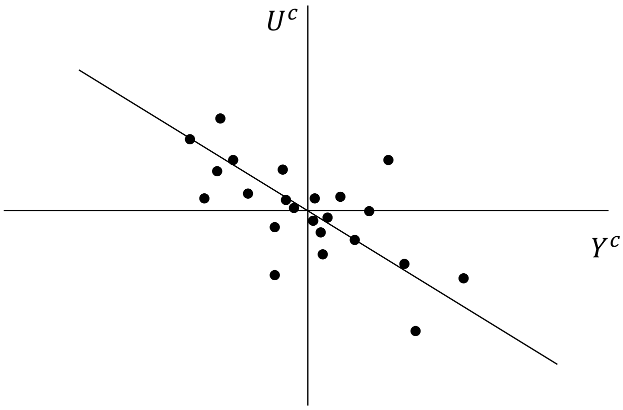


Figure 2.



As explained above, there are two specifications of the relationship between unemployment and output formulated by Okun (1962). Throughout this thesis, we shall use the "gap" version, which can be formulated as

$$U^c = aY^c \tag{8}$$

where U^c is the cyclical component of unemployment and Y^c is the percentage deviation of output from its potential value. In other words, let U be the observed unemployment rate, U^* be the natural rate of unemployment, Y the observed value of output, and let Y^* be the potential product. Then

$$U^c = U - U^*, \tag{9}$$

$$Y^c = 100(\ln Y - \ln Y^*). \tag{10}$$

We call the coefficient a in equation (1) the Okun's Law coefficient. Despite having the disadvantage of using unobserved values (natural rate of unemployment and potential product), the "gap" version of Okun's law was chosen for this thesis, since it allows the regression to be made "through origin", that is, with the intercept parameter being assumed equal to zero. The theoretical justification of such an assumption follows from the idea that an economy whose output is momentarily at its potential level should be at its natural rate of unemployment (i.e. if $Y^c = 0$ then $U^c = 0$). As shown in section Preliminary analysis, the intercept in Okun's model estimated for Russia as a whole over the period 1997-2016 is close to zero and statistically insignificant.

The first and simplest model we estimate is a static one:

Model 1 $$U_i^c = \beta_1 Y_i^c + u_i$$

where i is the number of region and u is an error term. As Perman and Tavéra (2005) explain, such model might be mistaken in its assumption that the relationship is completely contemporaneous. Therefore, we shall also use the Autoregressive Distributed Lag model:

Model 2 $$U_i^c = \beta_1 lagU_i^c + \beta_2 Y_i^c + \beta_3 lagY_i^c + u_i$$

where $lagU^c$ is the cyclical unemployment rate in previous year and $lagY^c$ is the cyclical element of output in previous year. This model seems more reasonable, since it allows the cyclical unemployment to be influenced by the conditions on the labor market in the previous year. In order not to shorten the resulting time series, we do not include higher-order time lags. We can, however, utilize the geographic information in the data to investigate the spatial effects the regional labor markets have upon each other. Before we formulate the spatial models used in this thesis, let us define the concept of spatial weight matrix, which is central to the field of spatial econometrics.

The spatial weight matrix stores the geographical information about the regions in the following way. Let $n \in \mathbb{N}$ be the number of regions in the sample. For every $i, j \in \{1, \dots, n\}$, the element $\omega_{i,j}$ represents the degree of geographical association between regions i, j defined by a specific criterium. One of the common types of spatial weight matrices, for instance, is one in which each element $\omega_{i,j}$ is equal to 1 if regions i, j are contiguous (share common border) and 0 otherwise. The elements on the diagonal are always equal to zero, i.e. a region is not considered to be a neighbor of itself, regardless the type of spatial weight matrix. Before the use in regression analysis, the matrix is row standardized by dividing every row by the sum of the elements in it. This row standardization assures that the sum of every row is equal to one, and therefore the matrix serves as a storage of weights for the calculation of a weighted average of a given variable's values in the neighboring regions. In this thesis, six different spatial matrices are used. For their detailed description, see section "Data".

The first spatial model we use in this thesis is the so called "Spatially Lagged X's" (SLX) model

Model 3 $$U_i^c = \beta_1 Y_i^c + \theta \sum_{j=1}^{n} (\omega_{i,j} Y_j^c) + u_i$$

where n is the number of regions in the sample, the rest of the notation remaining the same as above. Since, as already stated, the matrix is row-standardized, the parameter θ measures the marginal effect of the weighted arithmetic average of the values of output in neighboring regions, the weights being defined by the criteria according to which the spatial weight matrix is constructed.

A logical extension of this model is to include time-lagged independent variable, thus obtaining something that could be called "Temporally and Spatially Lagged X's" (TSLX) model

Model 4

$$U_i^c = \beta_1 Y_i^c + \beta_2 lagY_i^c + \theta_1 \sum_{j=1}^{n} (\omega_{i,j} Y_j^c) + \theta_2 \sum_{j=1}^{n} (\omega_{i,j} lagY_j^c) + u_i$$

where the inclusion of output of neighboring regions one year previously ($lagY^c$) is especially natural, intuitively, since the economic conditions of neighboring regions are likely to influence the region in question with some delay.

Another possible choice of spatial model would be the Spatial Autoregressive (SAR) model which could be specified as follows.

$$U_i^c = \beta_0 + \rho \sum_{j=1}^{n} (\omega_{i,j} U_j^c) + \beta_1 Y_i^c + u_i \tag{11}$$

This model, however popular it is in spatial econometric studies, cannot be used for the purpose of this thesis. As we allow the dependent variable, cyclical unemployment in region i, to be influenced by the cyclical unemployment in neighboring regions, we also allow the cyclical unemployment in neighboring regions to be influenced by the cyclical unemployment in region i, thus creating an endless loop of mutual effects. That is why the SAR model is sometimes called global, as opposed to the SLX model, which is regarded as local. This model, as well as the Spatial Error Model (SEM)

$$U_i^c = \beta_0 + \beta_1 Y_i^c + u_i \tag{12}$$

$$u_i = \lambda \sum_{j=1}^{n} (\omega_{i,j} u_j) + \varepsilon_i \tag{13}$$

whose spatial aspect lies in the error term, is estimated by the maximum likelihood estimation, for which the R-squared is not defined. The same applies to all possible combinations of the spatial models in which the spatial effect is allowed in the dependent variable or in the error term.

A clarifying remark seems appropriate at this point. Despite the time lagged variables being included, all of the models 1 through 4 are cross-sectional. None of them is actually the primary concern of our analysis, but only an intermediary step in it. We do not aim at estimating the Okun's law coefficient itself, but the convergence of the Russian regions with respect to it. Therefore, at the center of our attention lie not the coefficients of the models estimated, but above all the goodness-of-fit of each model, measured by the R-squared.

We estimate models 1 through 6 for each of the 20 time periods separately (19 in case of models 2 and 4 which include lagged variables) and then examine the development of R-squared over time. If there is an increasing time trend, we can conclude that the regional labor markets head towards the same relationship between cyclical unemployment and cyclical output. This is easy to investigate. We need simply to regress the R-squared time series, generated by each model, on time:

Convergence Model $\qquad R_t^2 = \beta_0 + \beta_1 t + u_t.$

Here the null hypothesis of non-convergence can be formulated as

$$H_0 : \beta_1 = 0$$

with a two-sided alternative $H_1 : |\beta_1| > 0.$

If the null hypothesis is rejected, we can subsequently turn our attention to whether the coefficient β_1 is positive or negative. Positive β_1 would mean an increasing trend, signifying Okun's law convergence, negative β_1 would signify an Okun's law divergence, that is, that the regional labor markets tend to be gradually more diverse rather than similar.

The problem with the "gap" version of Okun's law is that it uses potential income and natural rate of unemployment, variables which cannot be observed and have to be estimated. There are multiple possibilities how this can be done. This thesis chooses a method popularized in economics by Hodrick and Prescottt (1997) and named after them as the Hodrick-Prescottt filter (hereafter referred to as HP filter).

Let us have a time series $\{y_t, t = 1, \ldots, T\}$. HP filter decomposes the time series to its trend element τ_t and cyclical element c_t, satisfying $y_t = c_t + \tau_t$. The trend series is found as the solution of

$$\min_{\tau} \left(\sum_{t=1}^{T} c_t + \lambda \sum_{t=3}^{T} \left((\tau_t - \tau_{t-1}) - (\tau_{t-1} - \tau_{t-2}) \right)^2 \right).$$

The first summation in the above formula penalizes the deviations from the trend, the second penalizes the variation in the trend, measured by its second difference. The parameter λ determines the degree to which the variation of the trend is penalized. For λ approaching infinity the method is identical with estimating a linear time trend using OLS. For $\lambda = 0$, obviously, the trend series would be identical to the original y_t.

For the examination of spatial correlation between regions in terms of unemployment and OLC, we use a statistic devised by Patrick Alfred Pierce Moran (1950) and subsequently named after him as Moran's I. The Moran's I is defined as

$$I = \frac{n}{W} \frac{\sum_i \sum_j \omega_{i,j} (x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \tag{14}$$

where $W = \sum_i \sum_j \omega_{i,j}$ is the sum of all elements of the spatial weight matrix. As all of our spatial weight matrices are row-standardized, the sum of all their elements always equals the number of regions, and therefore the first factor equals one. The numerator in the second factor represents what could be called a geographically weighted covariance; the denominator standardizes. Apparently, the Moran's I has a lot in common with the coefficient of correlation. Under the null hypothesis of no spatial correlation, the expected value of Moran's I is

$$E(I) = \frac{-1}{n-1}$$

where n is the number of regions, so we can see that for large samples the expected value converges to zero. We test the significance of each obtained value of Moran's I against the one-sided alternative $I > 0$ and report the resulting p-value. This statistic can also be used for the evaluation of individual spatial weight matrices as follows. Let I_1, I_2 be the values of Moran's I obtained by using matrices W_1, W_2 and let $I_1 > I_2 > 0$. Then we can say that the spatial autocorrelation obtained from matrix W_1 is stronger than the one obtained from matrix W_2, and therefore matrix W_1 captures the spatial relationships between the regions better.

4. Data

This thesis uses annual data on regional unemployment rate and growth of regional product for the period from 1997 to 2016 provided by Federal State Statistics Service of Russian Federation. Russia's territory is divided into 85 federal subjects. Three of these were excluded from the analysis, since they are regarded as a part of another subject[3] and

[3] Khanty-Mansi Autonomous Okrug and Yamalo-Nenets Autonomous Okrug are parts of Tyumen Oblast, and Nenets Autonomous Okrug is contained in Arkhangelsk Oblast.

thus they are accounted for in the data. Another three subjects were discarded because of lack of data: for the Chechen Republic the data is available only for the years following 2005, the Republic of Crimea and the city of Sevastopol (which constitutes a separate federal subject) only became part of Russian Federation (and therefore of Russian statistics too) in 2014. The analysis was conducted on the remaining 79 subjects.

The growth of regional product was measured as index of the physical volume of production. In this respect, we follow here the work of Vakulenko (2015). In order to be able to apply the HP filter on the time series and thus obtain the percentage deviation of real product from the potential, the data was subjected to the following procedure. Originally, we have for each region a time series of physical volume of production indices, each expressed as percentage of the previous year (data available for 1998-2016). We divide the whole series by 100 and then multiply each element in it by all the preceding elements. Subsequently, we can include 1 on the beginning of the series as the value for year 1997. Thus, we recover the volume of each year's production, divided by the volume of production for the year 1997. This division does not affect our analysis, since we are only interested in the percentage, not absolute, deviation from the potential GRP and thus do not need to know the absolute values of physical product. That is a positive, simplifying circumstance, since the units of measurement of physical production would no doubt constitute a very nontrivial issue.

The HP filter is used to estimate the potential GRP and natural rate of unemployment. Following Ravn and Uhlig (2002), the value of the smoothing parameter was set to $\boldsymbol{\lambda} = \mathbf{6.25}$, since we are dealing with annual data. Figures 3 and 4 bellow demonstrate the decomposition on the example of Adygey Republic.

Figure 3. Observed and Potential GRP
Adygey Republic

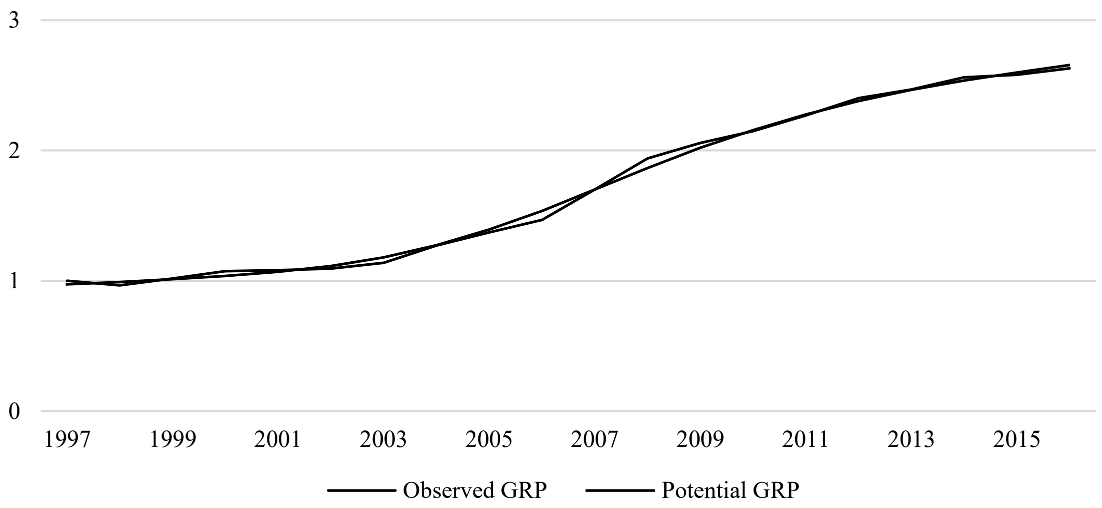
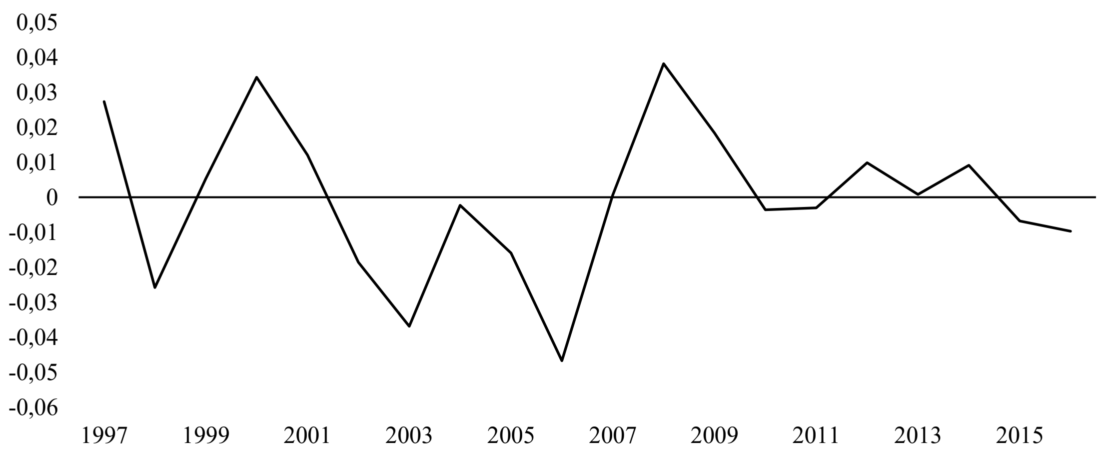


Figure 4. Cyclical GRP
Adygey Republic



The spatial weight matrices were generated in R, using package `shapefiles` created by Ben Stabler (2013). The shapefiles from which the matrices were created were provided by Center for Spatial Sciences, University of California, Davis. Spatial models (models 3 and 4) were estimated with six different spatial weight matrices defined as follows.

Matrix 1. Element $\omega_{i,j}$ equals 1, if regions i,j share common border, otherwise $\omega_{i,j} = 0$.

Matrix 2. Element $\omega_{i,j}$ equals 1, either if regions i,j share common border, or if they are "second order neighbors", i.e. regions i,j do not share common border and there is some region k which shares border with both i and j. Zero otherwise.

Matrix 3. Element $\omega_{i,j}$ equals 1, if region j belongs to 8 nearest neighbors of region i, zero otherwise.

Matrix 4. Let $d_{i,j}$ be the distance between regions i,j in kilometers. If $d_{i,j} \leq 1220$, then $\omega_{i,j} = 1$, otherwise $\omega_{i,j} = 0$.

Matrix 5. Inverse distance matrix. $\omega_{i,j} = d_{i,j}^{-1}$.

Matrix 6. "Gravity" matrix. $\omega_{i,j} = d_{i,j}^{-2}$.

Since matrices 1 and 2 are adjacency matrices, the question arises as to how islands and enclaves should be treated. There is one island and one enclave in our sample. Kaliningrad Oblast, the enclave, was excluded from the analysis, because the nearest regions to it would only be its third-order neighbors (were Lithuania considered as its first-order neighbor, and Latvia and Belarus as second-order). Sakhalin Oblast, comprising Sakhalin island and a number of smaller islands, was considered a first-order neighbor to Khabarovsk Kray, Kamchatka Kray, and Primor'ye, since it is separated from the continent only by 7.3 km wide channel. Both Kaliningrad and Sakhalin were included when matrices 3 through 6 were used. The distance of 1220 km in matrix 4 was set in order that each region can have at least one neighbor (the minimum distance satisfying this requirement being 822.48 km). Setting the distance any higher could seem inappropriate since some regions already have 35 neighbors within the distance of 1220 km. In fact, we should not expect much from Matrix 4, because of the unevenness in the assigned number of neighbors. Looking at the map of Russian federation, we can see that the regions in the European part of it are much smaller in terms of area than the ones in the north of Asia. Were we to divide Russia into the European part and the Asian part and apply different radius in each for the construction of such matrix, the results might have been much more meaningful. The distances between the regions are measured as a great-circle distance between their geographic centers of weight. A matrix can be imagined in which the distances are measured between the centers of economic activity or between the capital cities, but such a matrix was not available for this thesis and a construction of one would no doubt be a large project. The author of this thesis leaves this task to future researchers.

5. Empirical Results

5.1 Preliminary Analysis

This section is intended to provide the reader with the necessary context of this thesis' research, so that the natural expectations of what the results should look like can be made. The thesis aims to examine the development of the relationship between regional unemployment and gross regional product in Russia, therefore it is logical to look at these variables first. Figure 5 plots the overall level of unemployment in Russia during the period of years 1997 through 2016. We can see that the initial level of unemployment in the late 1990s was above 10 %, reaching its peak in 1998 at 13.2 % and gradually decreasing from then until the beginning of global economic crisis in 2008. Having reached its second local maximum in 2009 at 8.4 %, the unemployment decreased in Russia until 2012 and seems to stay at about 5.5 % from then to the end of the available time series.

Figure 5. Unemployment in Russia (1997-2016)

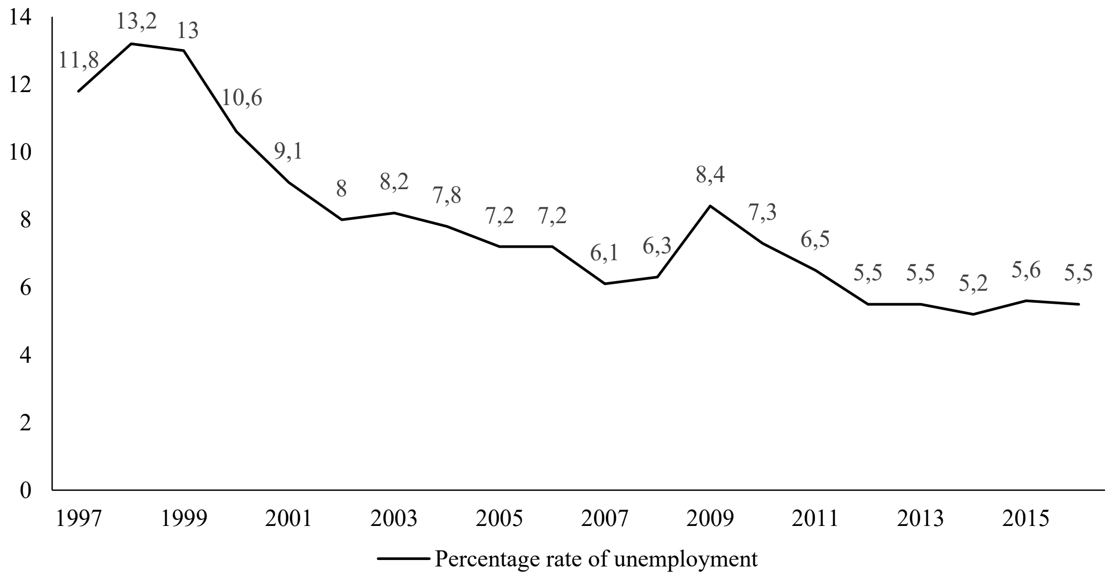
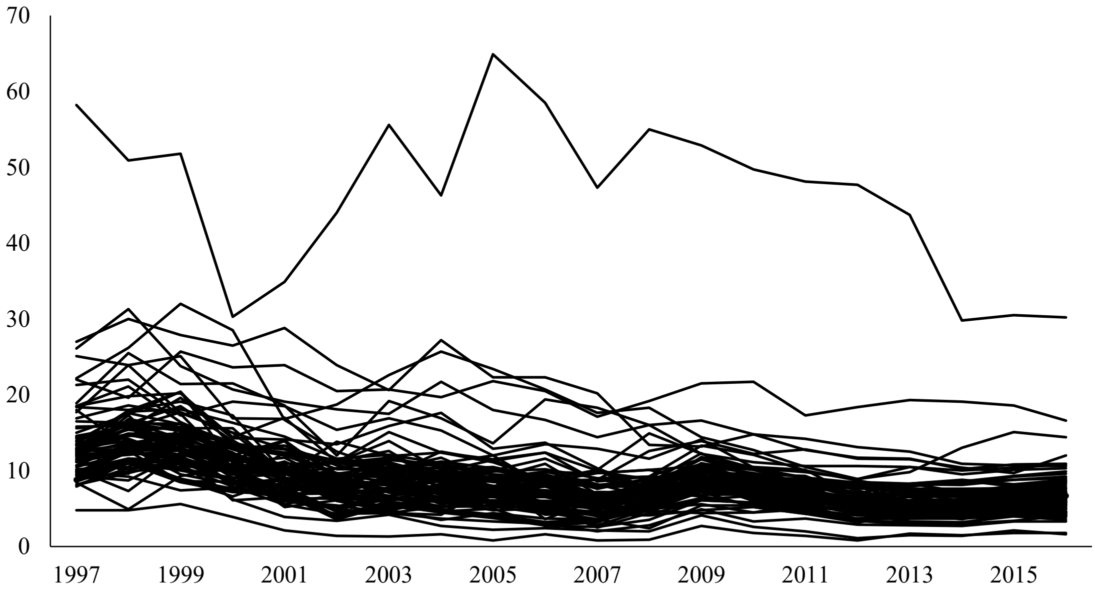


Figure 6 examines the same variable, except now on the regional level. The overall decreasing trend is apparent in the main cluster of regions, as it gradually moves from the 10-15 % range at the beginning to the 5-10 % interval around 2007, then temporarily increasing to the 2009 maximum, to return back to sub-10 % levels in the final third of the time series. Some tendency towards convergence can also be suspected from this figure, as the less-dense “cloud” surrounding the main cluster of regions almost disappears in the final years of the observed period. Finally, we shall mention the case of Ingushetia Republic which exceeds by far the level of unemployment in other regions, reaching 64.9 % unemployment in 2005 and staying above 40 % for most of the examined period, to show some signs of optimism only in the last three years with unemployment about 30 %. At the other end of the spectrum, we can notice the conversely extreme case of Moscow City with unemployment below 5 % for most of the period, to be joined, in the last years, by the Sankt Petersburg City.

Figure 6. Regional unemployment in Russia (1997-2016)



This observation is in agreement with Guriev and Vakulenko (2012) who find regional convergence in unemployment during the first decade of this century. We shall also employ the Moran’s I to investigate the spatial correlation in regional unemployment (see section Methodology for definition and properties).

Figure 7. Moran's I for Regional Unemployment

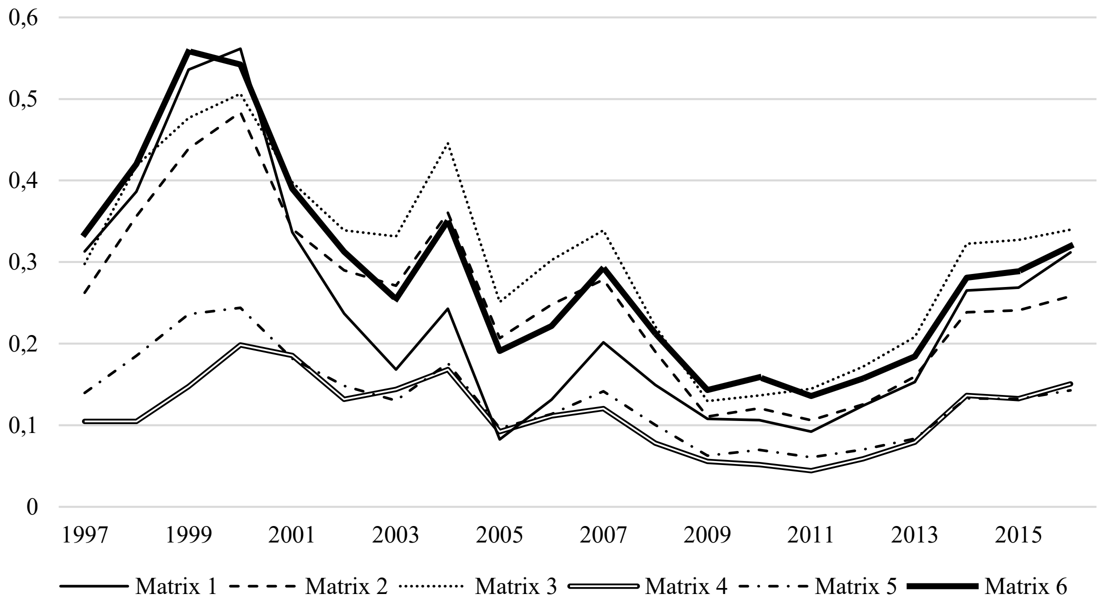


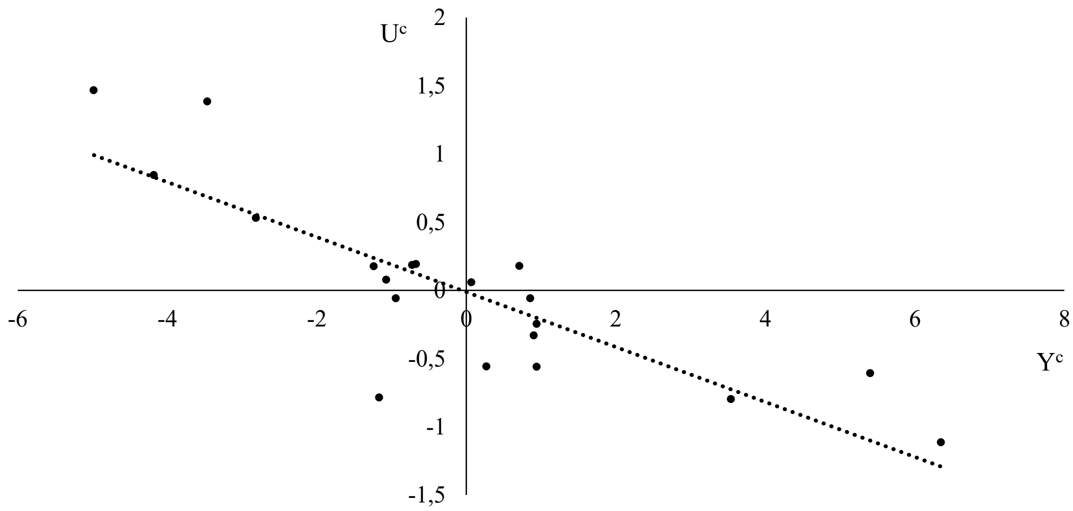
Figure 7 shows the development in Moran's I in regional unemployment as calculated with the use of the six spatial weight matrices defined in the previous subsection. We can see that all six matrices show positive spatial correlation in every year, with an apparent decreasing trend in the pre-2009 years and a converse trend thereafter. For exact values of Moran's I in regional unemployment, see Appendix 1. For the respective p-values, see Appendix 2 – as seen from it, even Matrix 4 produces significant spatial correlations. Judging by the average p-values for each matrix, we can tell that Matrix 5 produces the most significant spatial correlations, matrices 2, 3, and 6 being comparably appropriate. Matrix 1, however, brings substantially smaller significance of spatial correlation. So does Matrix 4, for which, considering it's poor suitability for Russia, it is not surprising. Vakulenko (2015) used primarily Matrix 5 in her analysis, relying on Semerikova and Demidova's (2015) finding that the inverse distance matrix is preferable for Russia to the adjacency matrix. Based solely on this simple application of ours, we are led to agree with such claim. We shall return to the evaluation of spatial weight matrices in the next section.

Let us now investigate the basic specification of the relationship of unemployment and output. First, we shall estimate the Okun's law coefficient for the Russian Federation as a whole, using the simplest model

Model 5 $$U_t^c = \beta_0 + \beta_1 Y_t^c + u_t, t = 1, \dots ,20$$

which is identical with Model 1 defined in section Methodology, except this time it is time-series rather than cross-sectional. The results are shown in Figure 8 below. The estimated Okun's law coefficient for the period 1997-2016 is -0.2, which means that for every one percent the domestic product deviated from the potential product, the unemployment rate deviates 0.2 % from the natural rate of unemployment in the opposite direction. In conformity with our expectation, the intercept is very close to zero and statistically insignificant. We shall perform an analogous analysis for each region separately, only this time running the regression through origin, that is, in Model 5 we assume $\beta_0 = 0$. Regression through the origin is motivated by the fact that it allows the Okun's relationship to be represented by a single real number, making it easy for us to calculate descriptive statistics such as interregional variance and Moran's I, and justified A. by the theoretically-intuitive assumption that an economy with output at its potential level should have unemployment equal to the natural rate of unemployment and B. by the empirical observation of the origin being indeed statistically zero in the previous estimation.

Figure 8. Okun's Coefficient in Russia



Coefficients:

	Estimate	Std. Error	t-value	p-value	
(Intercept)	-0.01221	0.08534	-0.143	0.888	
Y^c	-0.20152	0.03075	-6.554	$3.7 \cdot 10^{-6}$	***
	R-squared:	0.7047	DF:	18	

Table 1 below summarizes the results. We can see that the average Okun's law coefficient is similar to the one we obtained earlier for Russia as a whole, although the p-value being

0.11 signifies that the relationship is not strong in every region. Most of the regional OLC's, however, are negative, with no region's OLC being lower than -1. There are regions, such as – most notably – Primor'ye, Sakha and Zabaykal'ye, whose OLC is positive although not at all significant. The regressions R-squared for these regions is generally very low. We can conclude that for these regions the pattern predicted by the Okun's law does not occur. This observation together with the stated fact of OLC varying from region to region is consistent with the findings of the above-listed literature on the cross-country and cross-regional variance of OLC. The lower part of Table 1 lists the values of Moran's I for regional OLC's obtained by using different spatial weight matrices. Once more can we see that matrices 1, 5 and 6 capture the nature of the spatial dependence the best, matrices 2 and 3 being considerably less appropriate, and Matrix 4 showing barely any spatial correlation at all. For region-specific values of the OLC's and the respective statistics pertaining to them, see Appendix 3.

Table 1. Regional Okun's Law Coefficients Summary

	OLC	SE	t-value	p-value	R-squared	
Avg. values	-0.17337	0.07233	2.727647	0.11613	0.280088	
	Matrix 1	Matrix 2	Matrix 3	Matrix 4	Matrix 5	Matrix 6
Moran's I	0.287107	0.105019	0.136734	0.019483	0.093577	0.240141
p-value	$4.73*10^{-5}$	0.0025	0.0010	0.2272	$1.40*10^{-6}$	$1.65*10^{-6}$

After having been acquainted with the Russian labor market in some elementary manner, having discussed only those of its aspects immediately connected to our research, we are prepared to discuss the results of the main part of our analysis. Let us, however, not confuse the reader by creating the illusion of having listed facts of trivial nature. Although Vakulenko (2015) studies the spatial effects between regional labor markets of Russia using the Okun's model, she never estimates the OLC's for individual regions. With her paper being the closest to the regional analysis of Okun's law in Russia, there is no academic literature on this subject, save this thesis.

5.2 Regional Okun's Law Convergence

Models 1 through 4 were estimated by OLS. Table 2 stores the values of R-squared obtained by the estimation. We shall first make a brief comparison of the four models. In Table 3, we have the average value of R-squared for each model over the whole period.

Table 2. R-squared from Estimated Models

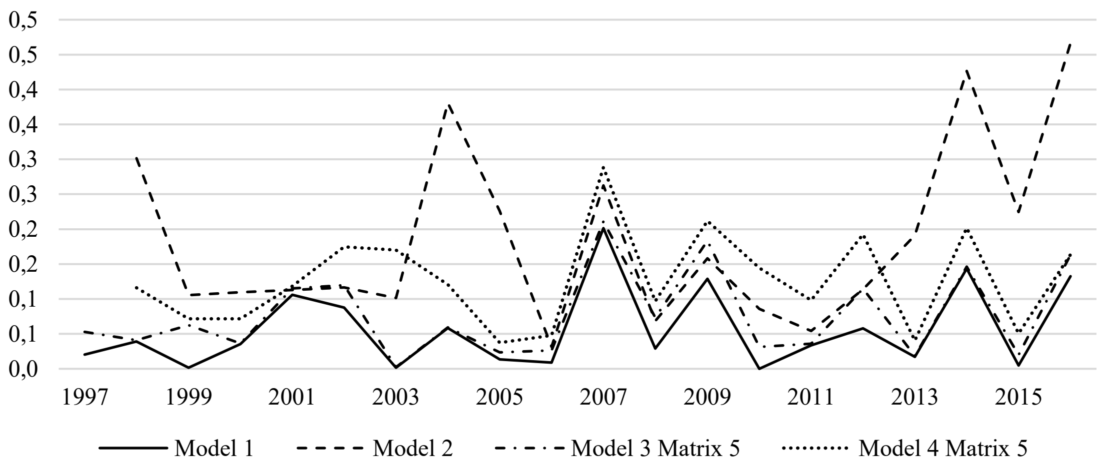
Year	Model 1	Model 3 *Matrix 1*	Model 3 *Matrix 2*	Model 3 *Matrix 3*	Model 3 *Matrix 4*	Model 3 *Matrix 5*	Model 3 *Matrix 6*
1997	0.0204	0.0775	0.0368	0.0255	0.0204	0.0527	0.0783
1998	0.0392	0.0399	0.0459	0.0464	0.0508	0.0407	0.0396
1999	0.0014	0.0015	0.0342	0.0438	0.0246	0.0623	0.0276
2000	0.0353	0.0514	0.0551	0.0408	0.0493	0.0365	0.0354
2001	0.1058	0.1100	0.1300	0.1133	0.1122	0.1151	0.1089
2002	0.0875	0.1073	0.2410	0.1320	0.1564	0.1195	0.0993
2003	0.0017	0.0342	0.0062	0.0063	0.0077	0.0023	0.0136
2004	0.0583	0.0691	0.0982	0.0595	0.0593	0.0595	0.0601
2005	0.0136	0.0229	0.0134	0.0136	0.0137	0.0235	0.0269
2006	0.0087	0.0542	0.0369	0.0192	0.0323	0.0262	0.0292
2007	0.2014	0.1836	0.1860	0.2142	0.2018	0.2109	0.2021
2008	0.0291	0.0960	0.1017	0.0726	0.0663	0.0736	0.0850
2009	0.1288	0.1813	0.1735	0.2145	0.1445	0.1832	0.1770
2010	0.0000	0.0321	0.0576	0.0317	0.0008	0.0312	0.0424
2011	0.0335	0.0360	0.0318	0.0349	0.0390	0.0360	0.0337
2012	0.0579	0.0858	0.0898	0.0672	0.0847	0.1149	0.1123
2013	0.0170	0.0247	0.0209	0.0172	0.0210	0.0184	0.0189
2014	0.1436	0.1449	0.1417	0.1438	0.1448	0.1461	0.1451
2015	0.0046	0.0426	0.0196	0.0089	0.0048	0.0198	0.0601
2016	0.1325	0.1808	0.1750	0.1375	0.1411	0.1633	0.1839
Year	Model 2	Model 4 *Matrix 1*	Model 4 *Matrix 2*	Model 4 *Matrix 3*	Model 4 *Matrix 4*	Model 4 *Matrix 5*	Model 4 *Matrix 6*
1997	-	-	-	-	-	-	-
1998	0.3016	0.1132	0.1272	0.1296	0.1085	0.1159	0.1023
1999	0.1054	0.0067	0.0404	0.0535	0.0418	0.0715	0.0547
2000	0.1095	0.1083	0.0882	0.0766	0.1118	0.0716	0.0871
2001	0.1130	0.1370	0.1355	0.1175	0.1176	0.1185	0.1132
2002	0.1166	0.1099	0.2626	0.1837	0.2452	0.1746	0.1587
2003	0.1016	0.1350	0.1119	0.1289	0.0766	0.1699	0.1867
2004	0.3808	0.1587	0.1408	0.1334	0.0878	0.1205	0.1395
2005	0.2261	0.0508	0.0136	0.0279	0.0152	0.0374	0.0459
2006	0.0315	0.1057	0.0495	0.0441	0.0480	0.0480	0.0479
2007	0.2624	0.2563	0.2387	0.2607	0.2546	0.2884	0.2922
2008	0.0686	0.1250	0.1067	0.0974	0.0877	0.0970	0.1056
2009	0.1584	0.1869	0.1790	0.2269	0.1551	0.2118	0.1877
2010	0.0859	0.1143	0.1655	0.1477	0.1284	0.1442	0.1490
2011	0.0543	0.0745	0.0991	0.0963	0.0815	0.0980	0.1067
2012	0.1136	0.1644	0.1700	0.1319	0.1774	0.1927	0.2108
2013	0.1911	0.0554	0.0883	0.0319	0.0245	0.0405	0.0388
2014	0.4267	0.2006	0.2119	0.2130	0.2012	0.2015	0.2024
2015	0.2244	0.0825	0.0675	0.0418	0.0428	0.0504	0.0863
2016	0.4664	0.1843	0.1760	0.1662	0.1441	0.1635	0.1857

Table 3. Average R-squared

	Model 1	Model 3 *Matrix 1*	Model 3 *Matrix 2*	Model 3 *Matrix 3*	Model 3 *Matrix 4*	Model 3 *Matrix 5*	Model 3 *Matrix 6*
Average R-squared	0.0560	0.0788	0.0848	0.0721	0.0688	0.0768	0.0790
	Model 2	Model 4 *Matrix 1*	Model 4 *Matrix 2*	Model 4 *Matrix 3*	Model 4 *Matrix 4*	Model 4 *Matrix 5*	Model 4 *Matrix 6*
Average R-squared	0.1862	0.1247	0.1301	0.1215	0.1131	0.1272	0.1316

Models 2 through 4 constitute various extensions of Model 1, so it is no surprise that their R-squared are higher. Model 2, however, apparently reaches higher values than the other models, despite its not including the spatially lagged variables. This higher goodness-of-fit can be explained by the explanatory power of the lagged dependent variable. The unemployment of the previous year is observed to be correlated with the unemployment of current year. We are therefore led to judge Model 2 to be superior to the other estimated models. It is in this context that the author of the thesis must admit that the inclusion of temporally lagged unemployment rate in models 3 and 4 would be a better choice. There is, however, little to no reason to reckon that such improvement would significantly change the final result, as far as the Okun's convergence is concerned. Figure 9 reveals that the higher average R-squared obtained for Model 2 is on account of two occasions in the first half of the observed period, which can as well be a random coincidence, and the last four years in which Model 2 seems to be superior consistently. We shall also mention that we choose Matrix 5 to be presented in said figure, as it seems to have proven its alleged preferability over the other matrices in the preliminary analysis.

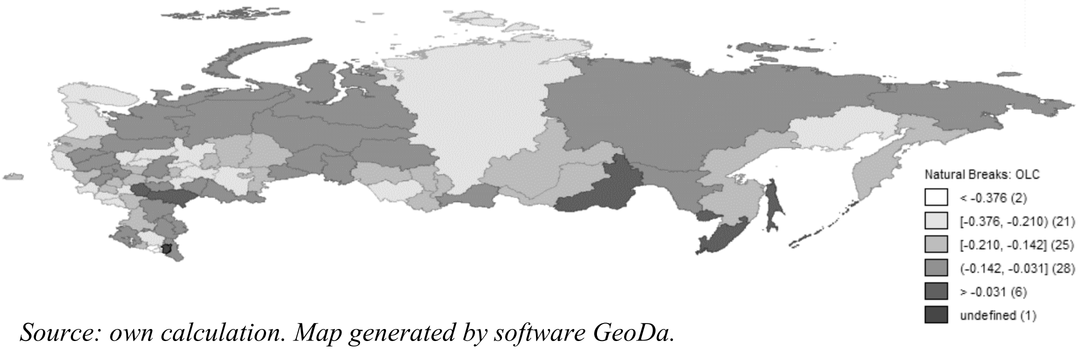
Figure 9. R-squared over Time



After examining the plotted development of R-squared over the observed period, we can essentially guess the final results of the analysis: Russian regions do not seem to converge in terms of Okun's law. Table 4 reports the results of the regression of R-squared on time in detail. None of the time series exhibits a significant trend. The null hypothesis could not be rejected in any of the estimated models. The estimated parameter β_1 from the Convergence Model is positive for every model, however, so at least some faint convergence below the sensitivity of our method is still more probable than divergence. We should not be confused by the magnitude of the estimates, since R-squared takes on only values between zero and one. We can notice that the highest estimate is reported for Model 2, which corresponds well with our observation that this model generated unprecedentedly high values of R-squared in the last four years. Even this, however, does not suffice for a significant increasing trend. Baring in mind the lack of degrees of freedom in our data series, we do not calculate heteroskedasticity- or serial correlation-robust standard errors. Those would, after all, only diminish the statistical significance of the estimates, and thus would not influence the conclusion drawn.

Table 4.	β_1	Standard error	t-value	p-value
Model 1	0.002260	0.002249	1.005012	0.328211
Model 2	0.006908	0.005282	1.307882	0.208324
Model 3, *Matrix 1*	0.002856	0.002149	1.329275	0.200361
Model 3, *Matrix 2*	0.001634	0.002702	0.604838	0.552836
Model 3, *Matrix 3*	0.001976	0.002552	0.774170	0.448880
Model 3, *Matrix 4*	0.001724	0.002372	0.726718	0.476744
Model 3, *Matrix 5*	0.002382	0.002382	0.999901	0.330612
Model 3, *Matrix 6*	0.003371	0.002207	1.527164	0.144101
Model 4, *Matrix 1*	0.002793	0.002460	1.135624	0.271872
Model 4, *Matrix 2*	0.002082	0.002845	0.731784	0.474269
Model 4, *Matrix 3*	0.001369	0.002904	0.471618	0.643196
Model 4, *Matrix 4*	0.000906	0.003000	0.302007	0.766310
Model 4, *Matrix 5*	0.001706	0.002925	0.583249	0.567383
Model 4, *Matrix 6*	0.002939	0.002833	1.037079	0.314226

Figure 10. OLC of Russian Regions



Source: own calculation. Map generated by software GeoDa.

This conclusion should also be put into the context of varying regional OLC’s in Russian regions, shown for this purpose on a map in Figure 10. Although we can see that spatial clusters do appear, the variation is obvious[4]. It is beyond the scope of this work to discuss the possible arrangements that should be done in order that the macroeconomic policies of the Russian government are optimal for each region, but their potential usefulness seems to be clearly implied by the above-presented results. As the regional OLC’s differ, so do, consequently, differ the regional short run aggregate supply functions, and therefore the regions should be expected to react to supply shocks diversely. The resulting different inflationary pressures that can be expected as a consequence might form a powerful argument against Russia being an optimal currency area.

Let us consider the appropriateness of the used spatial weight matrices one last time. We can see from Table 4 that the choice of matrix does not influence the final results. In case of Model 4, the Matrix 4 produces even less significant estimate of time trend than is obtained by other models. This result, however, has no analogy in Model 3, Matrix 4, where the magnitude and significance of the estimate is similar to those obtained in other matrices and other models. The average values of R-squared reported in Table 3 do not seem to be sensitive to the choice of matrix either. For the commentary on the matrices based on our research to be complete, we include in Appendix 4 two figures plotting the values of R-squared obtained by Model 3 and Model 4 respectively, each comparing the results generated by all six matrices. We can see that the choice of matrix does not cause

[4] Let us once more clarify that the regions of Crimea and Sevastopol which are new to the Russian Federation were not included to the analysis nor to the map. As it would not be logical to exclude Chechen Republic from the map, it is designated as “undefined.”

a dramatic difference in either of the models. The common trend, or more precisely, the correlation between the results produced by different matrices is apparent. There also seems to be more similarity between the results of different languages in the latter half f the studied period than in the former half. This phenomenon could be interesting to analyze, but it exceeds the range of this text. The interpretation that offers itself is that the choice of matrix becomes less and less significant with respect to the results.

6. Conclusion

In this thesis, we studied the properties of regional labor markets of Russian Federation. The ultimate goal of our analysis was the estimation of the Okun's model for the Russian regions, the examination of the inter-regional variance of the Okun's coefficient and the investigation of a inter-regional convergence in terms of the Okun's coefficient. We applied Hodrick-Prescott filter to obtain potential values of regional output and natural rate of unemployment, which is necessary for the "gap" version of Okun's law to be estimated. We estimated a simple linear model which produced negative Okun's coefficient for most of the regions and considerable cross-regional variance thereof.

As there is no literature examining the Okun's law for Russian regions, we provide in Appendix 3 a table of Okun's coefficients estimated over the period of 1997-2016 for each of the 79 Russian regions that were part of the analysis separately. The Okun's law relationship between unemployment and regional product appears not to be statistically significant in most of the regions, the average p-value being 0.1161. There are even cases in which the estimated Okun's coefficient is positive. In most of the regions, however, the relationship is negative (average estimated value of regional Okun's coefficient equals -0.173).

Since the available statistical data is insufficient for the standard method of examining Okun's convergence (that is, by means of rolling regression obtaining time series of OLC estimates and then examine the presence of a trend in the cross-sectional variances, see Perman and Tavéra (2005) for details) to be applied, we develop our own method which utilizes the cross-sectional dimension of the data. The cyclical element of unemployment is regressed on the cyclical element of regional product over the whole population of 79 regions, separately for every year. The obtained R-squared of each regression is saved to form a time series of goodness-of-fit measurements which, if a positive time trend is detected in such a series, signifies "gathering" of the regions more tightly around the regression line which is equivalent with the standard notion of $\boldsymbol{\sigma}$-convergence.

We estimated four different models of Okun's law:

1. A static linear model
2. An Autoregressive Distributed Lag model
3. A Spatial Lag model
4. A Spatial and Temporal Lag model

None of the models produces a trending time series of R-squared, which signifies an absence of Okun's convergence.

For the purposes of spatial models, we constructed six different spatial weight matrices:

1. An adjacency of order 1 matrix
2. An adjacency of order 2 matrix
3. A matrix of 8 nearest neighbors
4. A distance bound matrix
5. An inverse distance matrix
6. An inverse square distance matrix

The choice of spatial weight matrix did not influence the results of the convergence analysis. Examination of spatial correlation between regions in terms of unemployment and long-run (20 years) Okun's coefficient did, however, support the findings of Semerikova and Demidova (2015) that the inverse distance matrix is more appropriate for Russian Federation than the adjacency matrices. Matrix of 8 nearest neighbors and inverse square distance matrix capture the spatial correlation comparably well, while the distance bound matrix proves inappropriate for Russia. All matrices produced positive and significant spatial correlation in unemployment and Okun's coefficient.

To conclude, this thesis joins the literature supporting the empirical existence of inverse relationship between unemployment and product formulated by Okun (1962). On the example of Russian regions, we demonstrate the variability of the relationship which has also been observed by many researchers on other countries' contexts, as well as the existence of cases where the relationship does not occur at all. Despite the observed convergence of Russian regions in terms of unemployment, the convergence in terms of

Okun’s law does not seem to take place. This finding is consistent with Perman and Tavéras (2005) paper’s discovering absence of Okun’s convergence between countries of European Union (except for two different convergence clubs – see section Literature review for details).

As the question of Okun’s convergence has only been studied twice yet and never in the context of Russia, the thesis has only employed basic tools that sufficed for the convergence analysis. There remain many opportunities for future research. As far as the author is aware, there is still no definite criterion for evaluation of spatial weight matrices. As some of the matrices used in this thesis proved inadequate, further research in this field might bring useful information. Furthermore, the possibility of extending the method of estimating Okun’s convergence used here to other spatial models, for which the R-squared is not defined, could be inspected. The author, however, finds the conventional approach used by Perman and Tavéra (2005) preferable, although demanding long time series. There is also a possibility of formulating a spatial model in which multiple spatial weight matrices are used, as well as many other model specifications that could be more appropriate than those used here could probably be found.

As our method had already exploited the cross-sectional dimension of the data panel, it was impossible to examine the convergence in case of various convergence clubs (i.e. groups of regions associated by a predefined criterion in terms of which they show similarity). Yet, it is only in the case of convergence clubs that Perman and Tavéra (2005) are able to find Okun’s convergence in European Union. There are many possible ways how such clubs could be defined in the context of Russia. One of the original intentions of the author of this thesis was to investigate whether the grouping of Russian regions into federal districts (federalnyye okruga) is based on (or at least correlates with) real similarities in terms of various economic factors.

7. Bibliography

1. ADANU, Kwami. (2002) “A Cross-Province Comparison of Okun's Coefficient for Canada.” Econometrics Working Papers [online]. 2002 [cit. 2019-05-06]. doi:10.1080/0003684042000201848.
2. APERGIS, Nicholas, & REZITIS, Anthony. (2003) “An examination of Okun's law: evidence from regional areas in Greece.” Applied Economics [online]. 2003, 35(10), 1147 [cit. 2019-05-06]. doi:10.1080/0003684032000066787.
3. BLACKLEY, P. (1991). “The measurement and determination of Okun's Law: Evidence from state economies.” Journal of Macroeconomics [online]. 1991, 13(4), 641 - 656 [cit. 2019-05-06]. doi:10.1016/S0164-0704(05)80017-5. ISSN 01640704.
4. DANILENKO, Tatiana, DEMIDOVA, Olga, & SIGNORELLI, Marcello. (2018). “Unemployment Clubs in Russian Regions. Emerging Markets Finance.” 2018, 54(6), 1337-1357 doi:10.1080/1540496X.2017.1281799. ISSN 1540496X.
5. GURIEV, Sergei & VAKULENKO, Elena. (2012). “Convergence between Russian regions,” Working Papers w0180, Center for Economic and Financial Research (CEFIR).
6. HODRICK, Robert J., & PRESCOTTT, Edward C. (1997). “Postwar U.S. Business Cycles: An Empirical Investigation.” Journal of Money, Credit and Banking, 29(1), pp. 1-16. doi:10.2307/2953682
7. KOUPARITSAS, Michael A. (2001). “Is the United States an Optimum Currency Area? An Empirical Analysis of Regional Business Cycles.” (December 2001). FRB of Chicago Working Paper No. wp2001-21. Available at SSRN: https://ssrn.com/abstract=295566 or http://dx.doi.org/10.2139/ssrn.295566
8. MOOSA, Imad A. (1997). “A Cross-Country Comparison of Okun's Coefficient.” Journal of Comparative Economics [online]. 24(3), 335 - 356 [cit. 2019-05-01]. doi:10.1006/jcec.1997.1433. ISSN 01475967.
9. MOOSA, Imad A. (1999). “Cyclical output, cyclical unemployment, and Okun's coefficient A structural time series approach.” International Review of Economics & Finance, 8, issue 3, pp. 293-304. doi:10.1016/s1059-0560(99)00028-3
10. MORAN, Patrick Alfred Pierce. (1950). “Notes on Continuous Stochastic Phenomena.” Biometrika. 37 (1): 17–23. doi:10.2307/2332142. JSTOR 2332142.
11. OKUN, Arthur M. (1962). “Potential GNP: Its measurement and significance.” Proceedings of the Business and Economic Statistics Section of the American Statistical Association pp. 89–104.

12. PERMAN, Roger & TAVÉRA, Christophe. (2005). "Testing for Convergence of the Okun's Law Coefficient in Europe." Applied Economics, 2005, 37, 2501–2513. ISSN 0003–6846 print/ISSN 1466–4283 online. http://www.tandf.co.uk/journals doi:10.1080/00036840500366395.
13. PERRET, Jens K. (2011). "A Proposal for an Alternative Spatial Weight Matrix under Consideration of the Distribution of Economic Activity." Schumpeter Discussion Papers sdp11002, Universitätsbibliothek Wuppertal, University Library.
14. RAVN, Morten O., & UHLIG, Harald. (2002). "On Adjusting the Hodrick-Prescottt Filter for the Frequency of Observations." Review of Economics and Statistics, 84(2), pp. 371–376. doi:10.1162/003465302317411604
15. SEMERIKOVA, Elena Vyacheslavovna, & DEMIDOVA, Olga Anatolyevna. (2016). "Interaction of Regional Labour Markets in Russia: Spatial Econometric Analysis." Prostranstvennaâ Èkonomika, Iss 3, pp. 57-80 doi:10.14530/se.2016.3.057-080. ISSN 18159834.
16. SEMERIKOVA, Elena Vyacheslavovna, & DEMIDOVA, Olga Anatolyevna. (2015). "Analysis of Regional Unemployment in Russia and Germany: Spatial-Econometric Approach." Prostranstvennaâ Èkonomika, Iss 2, pp. 64-85. 2015, (2). doi:10.14530/se.2015.2.064-085. ISSN 18159834.
17. SIMIONESCU, Mihaela. (2014). "Testing Sigma Convergence Across EU-28." Economics & Sociology, Vol. 7, No 1, 2014, pp. 48-60. doi:10.14254/2071-789X.2014/7-1/5.
18. SIMONOVA, Marina V., SANKOVA, Larisa V., MIRZABALAEVA, Farida I., SHCHIPANOVA, Dina Ye., & DOROZHKIN, Vladimir E. (2016). "Assessment Problems and Ensuring of Decent Work in the Russian Regions." International Journal of Environmental and Science Education. 2016, 11(15), 7608-7626. ISSN 13063065.
19. SÖGNER, Leopold, & STIASSNY, Alfred. (2002) "An analysis on the structural stability of Okun's law – a cross-country study." Applied Econometrics, 2002, 14, 1775-1787. doi:10.1080/00036840210124180.
20. STABLER, Ben. (2013). "shapefiles: Read and Write ESRI Shapefiles." R package version 0.7. https://CRAN.R-project.org/package=shapefiles
21. TIEFELSDORF, M. & BOOTS, B. (1995) "The exact distribution of Moran's I." Environment [online]. 1995, 27(6), 985-999 [cit. 2019-05-08]. doi: 10.1068/a270985. ISSN 0308518X.

22. VAKULENKO, Elena, & E. GURVICH. (2015). “The Relationship of GDP, Unemployment Rate and Employment: In-depth Analysis of Okun’s Law for Russia.” VOPROSY ECONOMIKI 2015, vol. 3.
23. VAKULENKO, Elena. (2014). “Does migration lead to regional convergence in Russia?” No WP BRP 53/EC/2014, HSE Working papers, National Research University Higher School of Economics.
24. VAKULENKO, Elena. (2015). “Analysis of the relationship between regional labour markets in Russia using Okun’s model.” Applied Econometrics, Publishing House "SINERGIA PRESS", vol. 40(4), pp. 28-48.
25. WEBER, Christian. (1995). “Cyclical Output, Cyclical Unemployment, and Okun's Coefficient: A New Approach.” Journal of Applied Econometrics, 10, issue 4, pp. 433-45. doi: 10.1002/jae.3950100407
26. YILMAZKUDAY, Hakan, & YAZGAN, Mustafa Ege. (2009). “Okun's Convergence within the U.S.” (May 1, 2009). Letters in Spatial and Resource Sciences, Vol. 2, Nos. 2-3, pp. 109-122, 2009. Available at SSRN: https://ssrn.com/abstract=995543
27. YOUNG, Andrew T., HIGGINS, Matthew J., & LEVY, Daniel. (2008). “Sigma Convergence versus Beta Convergence: Evidence from U.S. County-Level Data.” Journal of Money, Credit and Banking, 40(5), 1083–1093. doi:10.1111/j.1538-4616.2008.00148.x

8. List of appendices

9. Appendices

Appendix 1:

Moran's I in Regional Unemployment						
	Matrix 1	Matrix 2	Matrix 3	Matrix 4	Matrix 5	Matrix 6
1997	0.313151	0.262541	0.297622	0.104586	0.139616	0.334707
1998	0.386493	0.356312	0.418035	0.104777	0.185048	0.420236
1999	0.536000	0.439186	0.476513	0.147344	0.236132	0.558591
2000	0.561582	0.483702	0.506399	0.198460	0.244005	0.542036
2001	0.336485	0.340656	0.397769	0.185722	0.181459	0.390028
2002	0.237024	0.290045	0.338858	0.131642	0.148360	0.312675
2003	0.168297	0.270935	0.331587	0.143894	0.130293	0.254510
2004	0.242821	0.360427	0.445813	0.168237	0.176063	0.350853
2005	0.082668	0.206634	0.251143	0.092169	0.096340	0.190983
2006	0.131513	0.248117	0.302542	0.111514	0.113639	0.221712
2007	0.201681	0.278753	0.338876	0.120279	0.141541	0.293022
2008	0.149481	0.190247	0.221151	0.078163	0.100431	0.212523
2009	0.108016	0.110724	0.129666	0.055778	0.062976	0.143068
2010	0.106426	0.120718	0.136600	0.051690	0.069917	0.158887
2011	0.092292	0.106178	0.144952	0.044336	0.060705	0.135851
2012	0.124135	0.125572	0.171862	0.058799	0.070275	0.157325
2013	0.153179	0.159595	0.208173	0.079397	0.083468	0.184434
2014	0.265070	0.238363	0.322344	0.136561	0.132813	0.280754
2015	0.268600	0.241035	0.327069	0.132764	0.132331	0.288773
2016	0.311747	0.258043	0.339777	0.150702	0.142884	0.319751

Appendix 2:

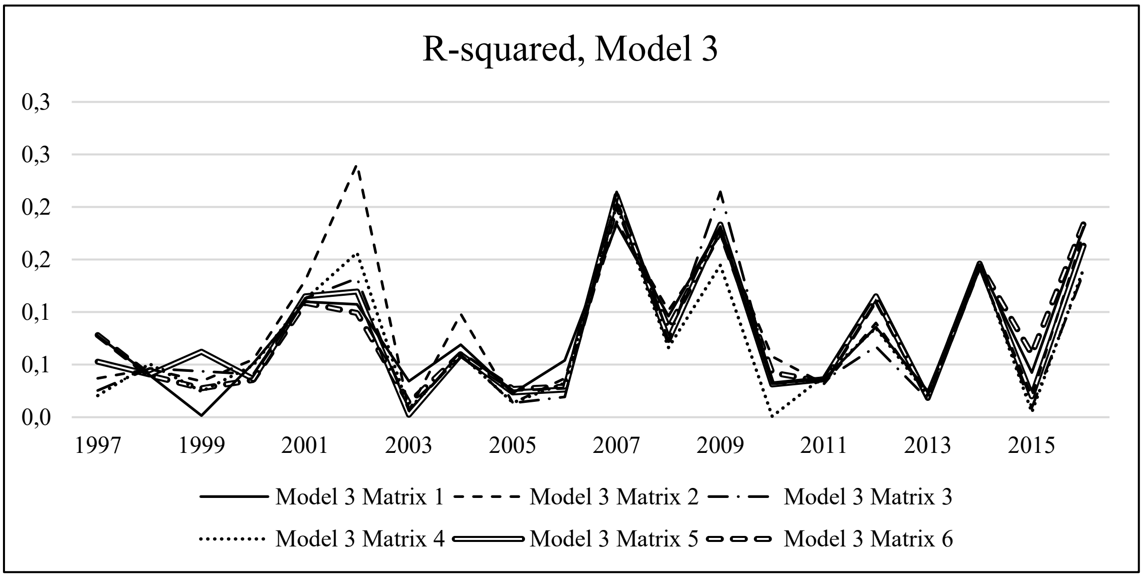
p-values of Moran's I in Regional Unemployment						
	Matrix 1	Matrix 2	Matrix 3	Matrix 4	Matrix 5	Matrix 6
1997	1,79E-07	2,52E-15	8,18E-15	5,55E-04	5,37E-16	8,97E-15
1998	1,44E-08	4,9E-21	1,26E-21	0,001824	8,04E-21	9,63E-18
1999	3,83E-15	1,09E-31	2,42E-28	2,71E-05	5,42E-33	1,43E-30
2000	2,81E-14	1,34E-32	2,79E-27	4,39E-07	3,36E-30	5,94E-25
2001	1,41E-06	3,08E-18	1,35E-18	1,07E-06	6,32E-19	1,15E-14
2002	9,08E-05	7,41E-17	3,66E-17	5,92E-05	1,96E-16	2,87E-12
2003	0,00164	2,68E-17	4,49E-19	3,07E-06	2,8E-15	4,53E-10
2004	0,00012	8,42E-23	8,75E-26	1,86E-06	2,66E-20	8,03E-14
2005	0,038685	8,34E-14	6,06E-15	0,000284	7,21E-12	5,37E-08
2006	0,007434	5,93E-16	2,15E-17	9,89E-05	4,19E-13	1,24E-08
2007	0,000302	1,12E-17	3,29E-19	7,98E-05	5,76E-17	2,81E-12
2008	0,001601	1,01E-11	9,19E-12	0,001673	2,83E-12	3,91E-09
2009	0,011021	1,04E-05	1,01E-05	0,010605	8,12E-07	1,59E-05
2010	0,012631	2,62E-06	4,92E-06	0,01604	1,11E-07	2,92E-06
2011	0,017526	7,26E-06	3,07E-07	0,021256	4,67E-07	1,38E-05
2012	0,003328	3,12E-07	3,73E-09	0,005917	1,99E-08	1,01E-06
2013	0,001124	3,97E-09	6,16E-11	0,001282	1,46E-09	1,49E-07
2014	1,82E-05	5,43E-12	1,52E-15	3,91E-05	1,43E-13	3,56E-10
2015	1,27E-05	2,33E-12	3,9E-16	5,29E-05	1,18E-13	9,07E-11
2016	7,91E-07	1,52E-13	7,18E-17	8,38E-06	3,91E-15	1,83E-12
mean	0,004777	1,03E-06	7,68E-07	0,00299	7,05E-08	1,69E-06

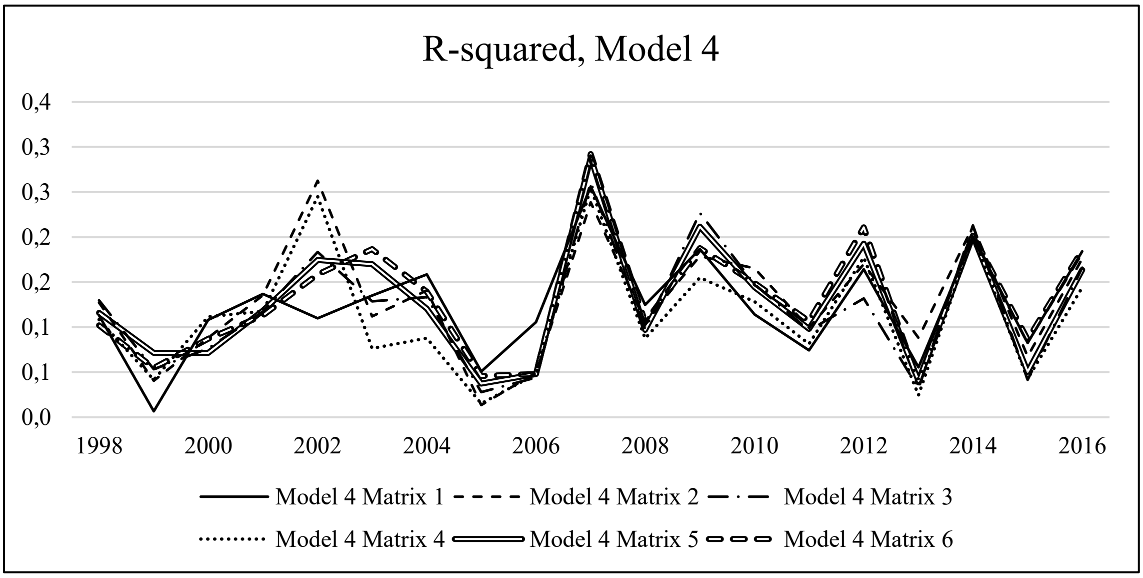
Appendix 3:

	OLC	Std. error	t-value	p-value	R-squared
Adygey	-0.2794	0.1464	1.9080	0.0716	0.1608
Altay	-0.2227	0.1047	2.1268	0.0468	0.1923
Amur	-0.1169	0.0429	2.7231	0.0135	0.2807
Arkhangel'sk	-0.1354	0.0614	2.2060	0.0399	0.2039
Astrakhan'	-0.1273	0.0385	3.3080	0.0037	0.3655
Bashkortostan	-0.2657	0.0565	4.7016	0.0002	0.5378
Belgorod	-0.2320	0.0717	3.2371	0.0043	0.3555
Bryansk	-0.2501	0.0568	4.4042	0.0003	0.5052
Buryat	-0.1561	0.0951	1.6420	0.1170	0.1243
St. Petersburg City	-0.1835	0.0258	7.1217	0.0000	0.7275
Dagestan	-0.0311	0.0199	1.5678	0.1334	0.1145
Gorno-Altay	-0.1564	0.0363	4.3047	0.0004	0.4937
Chelyabinsk	-0.1417	0.0601	2.3570	0.0293	0.2262
Chukot	-0.0987	0.1043	0.9465	0.3558	0.0450
Chuvash	-0.1984	0.0699	2.8381	0.0105	0.2977
Ingush	-0.2653	0.1230	2.1562	0.0441	0.1966
Irkutsk	-0.1886	0.0541	3.4846	0.0025	0.3899
Ivanovo	-0.2814	0.0784	3.5914	0.0019	0.4044
Kabardin-Balkar	-0.6270	0.1185	5.2914	0.0000	0.5957
Kaliningrad	-0.2104	0.0799	2.6326	0.0164	0.2673
Kalmyk	-0.0664	0.0330	2.0143	0.0584	0.1760
Kaluga	-0.1208	0.0328	3.6845	0.0016	0.4167
Kamchatka	-0.1473	0.1014	1.4525	0.1627	0.0999
Karachay-Cherkess	-0.1895	0.1663	1.1398	0.2685	0.0640
Karelia	-0.2564	0.0779	3.2916	0.0038	0.3632
Kemerovo	-0.1839	0.0682	2.6952	0.0143	0.2766
Khabarovsk	-0.1714	0.0699	2.4513	0.0241	0.2403
Khakass	-0.1502	0.1326	1.1324	0.2715	0.0632
Kirov	-0.2770	0.0638	4.3450	0.0003	0.4984
Komi	-0.1044	0.1348	0.7745	0.4482	0.0306
Kostroma	-0.2424	0.0588	4.1259	0.0006	0.4726
Krasnodar	-0.1183	0.0552	2.1451	0.0451	0.1950
Krasnoyarsk	-0.3755	0.0978	3.8381	0.0011	0.4367
Kurgan	-0.1338	0.0842	1.5878	0.1288	0.1171
Kursk	-0.2329	0.0730	3.1894	0.0048	0.3487
Leningrad	-0.1574	0.0795	1.9797	0.0624	0.1710
Lipetsk	-0.2436	0.0735	3.3140	0.0036	0.3663
Maga Buryatdan	-0.3392	0.0869	3.9040	0.0010	0.4451
Mariy-El	-0.2698	0.1135	2.3760	0.0282	0.2291
Mordovia	-0.1793	0.0332	5.3957	0.0000	0.6051
Moscow City	-0.0678	0.0268	2.5321	0.0203	0.2523
Moskva	-0.1204	0.0486	2.4789	0.0227	0.2444
Murmansk	-0.2813	0.0898	3.1318	0.0055	0.3405
Nizhegorod	-0.0890	0.0420	2.1185	0.0475	0.1911
North Ossetia	-0.4701	0.2163	2.1730	0.0426	0.1991
Novgorod	-0.0985	0.1271	0.7746	0.4481	0.0306

Novosibirsk	-0.1764	0.0377	4.6803	0.0002	0.5355
Omsk	-0.0755	0.0322	2.3437	0.0301	0.2243
Orel	-0.1608	0.0534	3.0110	0.0072	0.3230
Orenburg	-0.1357	0.0660	2.0548	0.0539	0.1818
Penza	-0.1354	0.0845	1.6011	0.1259	0.1189
Perm'	-0.1974	0.0474	4.1618	0.0005	0.4769
Primor'ye	0.0225	0.0768	0.2928	0.7728	0.0045
Pskov	-0.2567	0.0615	4.1731	0.0005	0.4782
Rostov	-0.1989	0.0488	4.0768	0.0006	0.4666
Ryazan'	-0.1958	0.0891	2.1982	0.0405	0.2028
Sakha	0.0931	0.0695	1.3399	0.1961	0.0863
Sakhalin	-0.0695	0.0911	0.7629	0.4549	0.0297
Samara	-0.1125	0.0460	2.4475	0.0243	0.2397
Saratov	-0.0261	0.0631	0.4136	0.6838	0.0089
Smolensk	-0.1696	0.0398	4.2643	0.0004	0.4890
Stavropol'	-0.3127	0.1063	2.9409	0.0084	0.3128
Sverdlovsk	-0.1874	0.0394	4.7583	0.0001	0.5437
Tambov	-0.0176	0.0626	0.2805	0.7821	0.0041
Tatarstan	-0.3074	0.0566	5.4299	0.0000	0.6081
Tomsk	-0.0911	0.0784	1.1615	0.2598	0.0663
Tula	-0.2027	0.0640	3.1685	0.0051	0.3457
Tuva	-0.0806	0.1561	0.5164	0.6115	0.0138
Tver'	-0.1194	0.0706	1.6928	0.1068	0.1311
Tyumen'	-0.0725	0.0487	1.4907	0.1525	0.1047
Udmurt	-0.1570	0.0368	4.2707	0.0004	0.4898
Ul'yanovsk	-0.2842	0.0530	5.3578	0.0000	0.6017
Vladimir	-0.2033	0.0599	3.3921	0.0031	0.3772
Volgograd	-0.0865	0.0524	1.6508	0.1152	0.1254
Vologda	-0.1299	0.0466	2.7881	0.0117	0.2903
Voronezh	-0.1577	0.0467	3.3791	0.0031	0.3754
Yaroslavl'	-0.2277	0.0515	4.4215	0.0003	0.5071
Yevrey	-0.0140	0.0435	0.3207	0.7519	0.0054
Zabaykal'ye	0.0049	0.1028	0.0474	0.9627	0.0001

Appendix 4:


R-squared, Model 3
0,3
0,3
0,2
0,2
0,1
0,1
0,0
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
Model 3 Matrix 1
Model 3 Matrix 2
Model 3 Matrix 3
Model 3 Matrix 4
Model 3 Matrix 5
Model 3 Matrix 6


R-squared, Model 4
0,4
0,3
0,3
0,2
0,2
0,1
0,1
0,0
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
Model 4 Matrix 1
Model 4 Matrix 2
Model 4 Matrix 3
Model 4 Matrix 4
Model 4 Matrix 5
Model 4 Matrix 6