REGIONAL LABOR PRODUCTIVITY DISPARITIES IN UKRAINE: MAIN CAUSES AND SPATIAL PATTERNS

by

Oleksandr Shkurpat

A thesis submitted in partial fulfillment of
the requirements for the degree of

Master of Arts in Economics

National University "Kyiv-Mohyla Academy"
Economics Education and Research Consortium
Master's Program in Economics

2006

Approved by ____
Ms. Serhiy Korablin (Head of the State Examination Committee)

____

____

____

Program Authorized
to Offer Degree Master's Program in Economics, NaUKMA

Date ____

National University "Kyiv-Mohyla Academy"

Abstract

REGIONAL LABOR PRODUCTIVITY DISPARITIES IN UKRAINE: MAIN CAUSES AND SPATIAL PATTERNS

by Oleksandr Shkurpat

Head of the State Examination Committee: Mr. Serhiy Korablin,
Economist, National Bank of Ukraine

This work covers the issues of interregional labor productivity disparities in Ukraine. It was motivated by the current statistics on regional labor productivity in the country. One of the main goals of the research is to disclose the causes of differences in aggregate productivity per worker using the shift-share approach. It was found that labor productivity disparities across Ukrainian regions are mostly attributable to the productivity differences across regions that in fact are uniform among types of economic activities within each region. To lesser extent labor productivity disparities can be ascribed to the regional specialization. Another important issue that was covered in this research is a detection of regional clustering within the Ukrainian borders with respect to labor productivity. According to the obtained results there is a strong evidence of the presence of two clusters: low productivity cluster including three western regions and high productivity cluster that consist of the three regions situated in the eastern part of Ukraine.

TABLE OF CONTENTS

LIST OF TABLES

Number *Page*

ACKNOWLEDGMENTS

I thank my family for finding the right words of encouragement and support during my study at EERC

I thank my thesis advisor Prof. Tatyana Zabotina for her valuable suggestions and critical comments concerning this thesis. Also for invaluable knowledge that she share with the students while teaching courses at EERC.

I want to thank Dunn, Esteban and Anselin for their work on inventing shift-share analysis and development of the spatial econometrics technique.

I want to express my gratitude for Nadiya Klos, Olga Zayets, Inna Bisovetska, Kateryna Demjanchuk, Oleksandr Slobodyanyk, Yuriy Kurganov for their support and being good friends during the study at EERC.

I want to thank Tom Coupe, Olesya Verchenko, Sergiy Malyar, Volodymyr Bilotkach for being the best professors at EERC.

GLOSSARY

Word.

GDP – Gross Domestic Product

GVA – Gross Value Added

LMERROR – Lagrange Multiplier Test for Spatial Error Model

LMLAG – Lagrange Multiplier Test for Spatial Lag Model

MLE – Maximum Likelihood Estimation

OLS – Ordinary Least Squares

SCS – State Committee of Statistics

SUR – Seemingly Unrelated Regression

ULFS – Ukrainian Longitudinal Firm Survey

C h a p t e r 1

INTRODUCTION

Beginning from the year 2000 we observed the significant growth of GDP in Ukraine. As possible explanation for this fact we can mention the growth in productivity. However, quit recently we encounter with the slowdown of GDP growth. In the following table we would like to present the figures which shows the aggregate productivity per worker in the regional aspect for the three consecutive years (limited due to availability of the data) starting from the year 2001:

Table 1. Aggregate productivity per worker in the regional aspect

	2001		2002		2003	
Top 5	11572	Dniprop	9812	Dniprop	11907	Donetsk
	10663	Donetsk	9805	Volyn	11497	Dniprop
	10596	Volyn	9494	Donetsk	11450	Volyn
	9909	Odesa	8813	Poltava	10422	Odesa
	9041	Poltava	8471	Odesa	10080	Poltava
……………	…….	……….	……...	……….	…….	……….
	…….	……….	…….	……….	…….	……….
Bottom 5	4756	Ivano-fr	4070	Zakarp	5279	Ternopil
	4536	Zhytomyr	4059	Vinnytsia	5152	Zakarp
	4062	Khmelnyt	3998	Ternopil	4953	Vinnytsia
	3882	Ternopil	3986	chernivtsi	4855	chernivtsi
	3571	chernivtsi	3300	Khmelnyt	4622	Khmelnyt
Ratio of the best to the worst	**3,24**		**2,97**		**2,58**	

In this table (table for all 25 regions is given in the Appendix A) we present the top 5 performing regions and bottom 5 sorted by the labor productivity, which is measured as the ratio of gross value added in the region to the employment in that region. In the table we also present the ratio of the best to the worst performing regions. Though, we can see that this ratio is declining, it is still remains high – 2.58. Therefore, we can see a clear evidence of the labor productivity disparities across Ukrainian regions.

Another important fact to be mentioned is the composition of the two opposite "fives". In the top 5 we can see the regions that are situated in the East of Ukraine, whereas in the bottom 5 we can see the regions from the West of Ukraine.

These two facts will be the subject of our analysis. Basically, we would like to answer the following questions:

1. What are the main reasons of the interregional labor productivity disparities in Ukraine?

2. Is there any statistical evidence of the regional clustering with respect to aggregate productivity per worker?

To address the first question we will use the so-called shift-share analysis, which was introduced for this particular issue by (Esteban, 1972), and was further developed in the works of (Dickerson et al., 2000), (Esteban, 1972), (Kamarianakis and Le Gallo, 2003), (Dall'erba et al., 2003), (Benito and Ezcurra, 2005) and others.

To answer the second question we will use the global and local Moran's I statistic together with Moran scatter plots, and to visualize regional clusters we will use Moran significance maps. Methodological notes for this issue are presented in (Kamarianakis and Le Gallo, 2003) and in more details described in methodology and estimation results sections.

Further, we will follow the next structure: in the literature review section we will briefly discuss and summarize the existing studies devoted to the studies of productivity, and in more details, those that use the shift-share analysis for detecting the main causes of the regional labor productivity disparities. In the next chapter we will present the methodological issues of the shift-share approach, introducing the recent developments in this field of study. Chapter four contains the brief description of the datasets used in research and major drawbacks of the datasets exploited in research. Empirical results obtained during the research will be presented in chapter 5. And in the final section of our research we will present concluding remarks for the obtained results and give policy recommendations together with suggestions for further research.

Chapter 2

LITERATURE REVIEW

Productivity is one of the most important concepts in the framework of theories concerning the economic growth and development. That is why it receives considerable attention of the researchers, especially after introduction of new growth theories that empirically prove the importance of the productivity for economic growth.

In the framework of estimation of the major factors influencing productivity various setups were used, including cross-country comparisons (specifically, absolute and relative productivity convergence between developed and developing countries); estimation of factors that influence within country productivity growth, using various definitions of productivity: total factor productivity (TFP), multifactor productivity (MFP), and also single factor productivity (labor productivity, technical productivity). Considerable attention was also dedicated to the problem of productivity differential across regions of the countries and unions.

For better comprehension of the literature related to the topic of my research it is structured as follows: at first we will review the papers that initiate the researches on productivity and provide the basic definitions of the main concepts in productivity studies, then we will look at papers that introduce the shift-share analysis for exploration of the factors that determine interregional productivity differentials as well as examining the spatial patterns of the productivity differentials. And finally we will analyze the papers describing the econometric techniques that are used in the estimation of interregional productivity disparities.

2.1 Initiating researches

Studies of the economic growth and main determinants of it were always in the center of attention of economists. The starting point of the modern neo-classical approach can be ascribed to the fundamental works of (Solow, 1956) and (Swan, 1956). They constructed a model in which growth along the transition to steady-state is a function of the capital stock, exogenously determined saving rate, population growth, and depreciation. According to the empirical study made by (Solow, 1957), which considered developments in the USA for the period 1909-1949, "87.5 % of the increase attributable to technological change and the remaining 12.5% to increased use of capital". So, the importance of productivity growth was outlined in the earliest neo-classical growth models. Determined factors of economic growth are mostly applicable for the country level and may explain differences in economic growth between countries. However, such economic variables as saving rates, capital stock and population growth may not significantly vary within the country.

2.2 Basic studies on the shift-share analysis

As suggested by the data used in various researches there are quite large differences in levels of labor productivity even within borders of the country. However, not that much attention was given to the disclosing of possible sources of differences in productivity that was proven to be one of the major causes of the distinct economic growth across countries. In my thesis I will concentrate on the determinants of productivity differences between regions in Ukraine. That may be very important issue in developing regional policies and harmonization of the regional development.

Among research studies dedicated to the issue of interregional differences in labor productivity I would like to outline papers written by (Esteban, 1972) (Esteban, 2000). In these papers author tried to find the main causes of the interregional disparities in labor productivity. He was the first to apply the shift-share analysis that was introduced by (Dunn, 1960) for the estimation of interregional labor productivity differences. He proposed decomposition of the labor productivity disparities into three components: industry-mix, productivity differential and allocative. Presented econometric methodology allows measuring the influence of each of the three components on the labor productivity disparities between regions. Despite the obvious simplicity, this methodology gives reliable assessment of the sources of interregional labor productivity disparities. Many authors applied Esteban's approach in estimation productivity disparities especially for European Union.

(Dickerson et al., 2000) estimated the influence of each of the three shift-share components on the productivity disparities within the UK using the data for the period of ten years (1992-2002). They found that interregional variance in the aggregate productivity per worker can be attributed to the productivity differential component. They also show that productivity within regions of UK for the period of time under study was fairly invariable across sectors.

(Dall'erba et al., 2003) used the shift-share analysis for investigating the reasons of the labor productivity disparities in three candidate countries for EU accession (Poland, Czech Republic, Poland, and Hungary). They measure labor productivity disparities as the deviation of labor productivity of each of the region in three countries from EU average. And again as in most researches their main finding was that the major source for productivity disparities is the region-specific factors (productivity differential component).

(Benito and Ezcurra, 2005) extended the shift-share analysis in order to use non-parametric technique. Authors estimated the change in entire distribution of regional productivity in EU for the time period of 1977-1999. They confirm the presence of the spatial association of neighboring regions. Concerning the impact of the shift-share components they also came to the conclusion that the interregional labor productivity disparities can be attributed to the intrinsic differences across regions.

(Kamarianakis and Le Gallo, 2003) estimated the impact of each of the three shift-share components on the labor productivity differentials across EU regions using the time period of 1975-2000. Again the main conclusion of this paper is that productivity differentials are attributed mostly to the region specific factors (productivity differential component). Besides, the authors shown the presence of the spatial clustering of European regions with high and low productivity, they also outlined the regions with atypical spatial association (regions with high (low) productivity per worker surrounded by the regions with low (high) labor productivity). It is also worth to mention a few useful extensions of shift-share analysis introduced in this paper. First, is the usage of SUR estimation instead of separate OLS regressions, and the second, is the incorporating of the spatial regression techniques, including spatial SUR estimation.

Authors of this particular paper provided rather comprehensive study on the sources of the productivity differentials. That is why I will use proposed methodology for studying the causes of the labor productivity disparities in Ukraine.

2.3 Methodological issues of spatial regressions estimation

According to the findings of the (Benito and Ezcurra, 2005) and (Kamarianakis and Le Gallo, 2003) one of the issues in the interregional studies that should not be ignored in the shift-share analysis is the presence of spatial autocorrelation. Econometric methods that introduce the spatial effects into the common estimation approaches (OLS, SUR) were described in (Cliff and Ord, 1972), (Ord, 1975), (Anselin, 1980, 1988, 1999).

In order to determine the presence of spatial autocorrelation in the data various tests could be used: Moran I statistic, LMLAG and LMERROR, and also robust LMLAG and LMERROR tests. The decision rule on the type of the spatial dependence, and henceforth the type of the model to use was proposed by (Anselin and Florax, 1995), which is as follows: if LMERROR

(LMLAG) is more significant, than LMLAG (LMERROR) and robust LMERROR (LMLAG) is significant whereas robust LMLAG (LMERROR) is not, then the most suitable model is the spatial error model (spatial autoregressive model). After specification of the model the following adjustments should be done for the models:

- In case of spatial error model the most appropriate specification for error terms would be:

$$\varepsilon = \lambda W \varepsilon + u \text{, with } u \sim N(0, \sigma_u^2 I)$$

and λ is a coefficient that measures the extent of spatial autocorrelation between the residuals;

- In case of spatial lag model specification should be as follows:

$$y = \rho W y + X\beta + \varepsilon \text{,}$$

  but there should be noted that unlike the case of time series, even if the error terms are *i.i.d.*, nevertheless, spatial lag term is correlated with the error terms (Anselin, 1999).

For both specifications OLS estimated parameters will be biased and inconsistent (Anselin, 1999), therefore another estimation methods should be used. Appropriate estimation methods for spatial regression models could be MLE introduced by (Ord, 1975), spatial two stage least squares described in (Anselin, 1980), and finally Method of Moments estimators developed by (Kelejian and Prucha, 1999). Due to available software I will use in my research Maximum Likelihood Estimation method. The details on estimation will be described in the methodology section.

2.4 Conclusions

As can be seen from the mentioned above literature, the field of regional productivity disparities is rather developed for the moment, as in terms of different countries and time datasets, so in terms of estimation techniques. However, for the Ukraine that is one of the ex-soviet countries, described methodology has not been hitherto used. So, in order to fulfill the existing gaps we will cover the following issues:

- Apply the shift-share analysis for interregional labor productivity disparities estimation using the available data for Ukraine;

- Examine the presence of the spatial patterns of labor productivity disparities within Ukraine;

- Provide the empirical evidence for the policy makers concerning the implementation of the specific regional policies.

C h a p t e r 3

METHODOLOGY

The first stage of our research is dedicated to shift-share analysis. The main idea of the shift-share analysis developed by (Esteban, 1972, 2000) is to separate the influence of the three sources of interregional productivity disparities. These are composition of industries represented in each region, differences in productivity of the same industries in different regions, and the differences between regions in reallocation of the labor from least to most efficient industries.

Before going further into computation details of the components of productivity disparities we need to introduce the following variables:

l_i^j - is the share of employment in region ***i*** involved in production of ***j***'s industry. Besides, the following condition should hold $\sum_j l_i^j = 1$ for all regions ***i.*** Then by l_{Ua}^j will be denoted employment share of ***j***'s industry on Ukrainian level. Hence forth the condition $\sum_j l_{Ua}^j = 1$ should also hold.

In the similar fashion will be denoted labor productivity:

x_i^j - is the labor productivity in region ***i*** and industry ***j.*** Similarly, x_{Ua}^j is the labor productivity at the Ukrainian level.

In order to compute the deviation of the regional labor productivity from the Ukrainian average we need to compute regional productivity - x_i and average Ukrainian - x_{Ua}. According to the previous notions these can be computed as follows:

$$x_i = \sum_j l_i^j x_i^j \quad \text{and} \quad x_{Ua} = \sum_j l_{Ua}^j x_{Ua}^j .$$

Then the deviation of regional productivity (weighted average of industrial productivities) from the average Ukrainian productivity can be computed as:

$$y_i = x_i - x_{Ua} \tag{1}$$

For the sources of interregional labor productivity disparities I will use the notions originally proposed by (Esteban, 1972). So, he showed that there are three components of regional differences in labor productivity. These are *the industry-mix component*, *the productivity differential component*, and *the allocative component.*

The first of the three – *industry-mix component* - m_i of the ***i***'s region shows the deviation of the labor productivity in region ***i*** from average of Ukraine due to the specific industrial composition of this region. Hence, if there are mostly high productive industries situated in the region ***i***, then value of the industry-mix component will be positive and negative otherwise. The highest value will be in region that specializes in the most productive industries.

What is basically done is that labor productivity held at the country average level, while allowing labor shares to deviate:

$$m_i = \sum_j (l_i^j - l_{Ua}^j) x_{Ua}^j \tag{2}$$

The second – *productivity differential component* p_i - is constructed to capture the industry by industry differences between region ***i*** and national average labor productivity. Thus, the highest values will be for regions in which industrial labor productivities are higher than the Ukrainian average. In principle to compute productivity differential component labor shares for industries are fixed at the Ukrainian average level, whereas labor productivities are taken for each industry ***j*** at region ***i:***

$$p_i = \sum_j (x_i^j - x_{Ua}^j) l_{Ua}^j \qquad (3)$$

The third – *allocative component a_i* - can be interpreted as the measure of regional efficiency in reallocating labor from least to most productive industries. The values of this component are positive if region specialized, in comparison to Ukrainian average, in industries which has higher than Ukrainian average labor productivity. It is basically a combination of the previous two components:

$$a_i = \sum (l_i^j - l_{Ua}^j)(x_i^j - x_{Ua}^j) \qquad (4)$$

Therefore, we can measure the discrepancy between regional labor productivity and average of the country as the sum of the three above mentioned components:

$$y_i = x_i - x_{Ua} = m_i + p_i + a_i$$

In order to show whether interregional differences in labor productivity can be explained solely by one of the three components (Esteban, 2000) proposed to run the following three OLS regressions:

$$y_i = \alpha_m + \beta_m m_i + \varepsilon_{m,i}\text{, } \boldsymbol{i}=1,\ldots N \qquad (5)$$

$$y_i = \alpha_p + \beta_p m_i + \varepsilon_{p,i}\text{, } \boldsymbol{i}=1,\ldots N \qquad (6)$$

$$y_i = \alpha_a + \beta_a m_i + \varepsilon_{a,i}\text{, } \boldsymbol{i}=1,\ldots N \qquad (7)$$

Where N is the number of regions, and error terms in each of the three regressions have the usual properties (normally and identically distributed: $\varepsilon_m, \varepsilon_p, \varepsilon_a \sim Nid(0, \sigma^2)$).

Since I want not only to identify the major source of the interregional productivity disparities, but also to trace the change of impact of each of the three components over time, then I will perform the shift-share analysis for several consecutive years.

The second part of the research is devoted to the testing of the data I use for sift-share analysis on the presence of the spatial autocorrelation. It is worth to mention that in (Anselin, 1988) it is shown that if in regressions such as used in shift-share analysis spatial autocorrelation is not taken into account, then this can result in model misspecification. So, I will follow the approach proposed by (Kamarianakis and Le Gallo, 2003) in order to identify spatial autocorrelation in my dataset and introduce it explicitly into the model.

First of all, there should be clearly defined the definition of spatial autocorrelation: that is coincidence of value similarity with location similarity (Anselin, 2001). The data exhibit positive spatial autocorrelation if high or low values are clustered in space, and negative if there are dissimilar values concentrated in neighboring locations. We can expect positive spatial autocorrelation within Ukrainian borders since, there is industrialized east and central regions and less industrialized northern and western regions.

Global spatial autocorrelation is usually measured by Moran's ***I*** statistic (Cliff and Ord, 1981). In matrix notation it can be written as:

$I_t = \frac{n}{S_0} \times \frac{z_t' W z_t}{z_t' z_t}$, index *t* means that statistic is computed for each year

where

$\boldsymbol{n}$ is the number of regions

$\boldsymbol{z_t}$ is a vector of n observations (in deviation from the mean value) for year t

$\boldsymbol{S_0}$ is the global sum of weights

$\boldsymbol{W}$ is the matrix of spatial weights.

This statistic shows the degree of linear association between observed values $\boldsymbol{z_t}$ and spatially weighted averages $\boldsymbol{Wz_t}$ of neighboring values.

Before proceeding further, I would like to define how the spatial weight matrix $\boldsymbol{W}$ is constructed.

There are many different spatial weights matrices considered in the literature, but some general conditions should hold for such matrices. Namely, each region in this matrix should be connected to the set of the neighboring regions by means of purely spatial pattern. Diagonal elements $\boldsymbol{w_{ii}}$ of such matrices takes zero values whereas the other values $\boldsymbol{w_{ij}}$ shows the spatial dependence between region i and j. Those elements should also be non-stochastic, non-negative and finite.

In order to maintain exogeneity of the model we need to introduce strictly exogenous weights. For that purpose (Kamarianakis and Le Gallo, 2003) used great circle distance between regional centroids. Due to the absence of the data on the geographical location of regional centroids of Ukrainian regions (oblast) I will use the great circle distance between administrative centers of Ukrainian regions (unfortunately this limitation will not allow to separate values of Kyiv city from values of Kyiv region).

According to (Kamarianakis and Le Gallo, 2003) the elements of non standardized spatial weight matrix are defined as follows:

$$\begin{cases} w_{ij}^{u} = 0 & if \quad i = j \\ w_{ij}^{u} = \dfrac{1}{d_{ij}^{2}} & if \quad d_{ij} \le D(1) \\ w_{ij}^{u} = 0 & if \quad d_{ij} \ge D(1) \end{cases}$$

where $\boldsymbol{d_{ij}}$ is the great circle distance between administrative centers of regions $\boldsymbol{i}$ and $\boldsymbol{j}$, $\boldsymbol{D(1)}$ is the first quartile of the great circle distance distribution, that actually serves as the cutoff parameter above which the dependence between regions is considered to be negligible.

Finally, in order to obtain not absolute, but relative distances the obtained matrix is standardized by replacing each element using the following relation:

$$w_{ij} = \frac{w_{ij}^{u}}{\sum_{j} w_{ij}^{u}}$$

Therefore, the sum of the elements in each row of the standardized matrix is equal to one.

Now turning back to the discussion of global spatial correlation it is worth to mention that Moran's $\boldsymbol{I}$ statistic gives the answer on the question whether the values are distributed dependently in space (shows the presence of clustering).

One of the very useful tools for the assessing the extent of regional clustering within the borders of the country is the Moran scatter plot. It actually helps to visualize the spatial autocorrelation, and serve as the complement in estimating Local Indicators of Spatial Association (LISA) (Anselin, 1995). So, first we need to compute LISA:

$$I_{i,t} = \frac{(x_{i,t} - \mu_t)}{m_0} \cdot \sum_{j} w_{i,j} (x_{j,t} - \mu_t)$$, where $$m_0 = \sum_{i} (x_{i,t} - \mu_t)^2 \Big/ n$$

In the expression above $x_{i,t}$ is the observation for region i for period t, μ_t is the mean of observations of variable x across all regions for the period t. Summation over j is such that only nearest (D(1)) regions are included.

Scatter plot is drawn on the basis of LISA computations. According to the (Anselin, 2002) this can be described as "the spatial lag (Wz_t) of the variable on the vertical axis and the original variable (z_t) on the horizontal axis". Here z_t stands for normalized variable x. normalization is done as follows:

$z_t = \frac{x_{i,t} - \mu_t}{\sqrt{\sigma^2_{x,t}}}$, where $\sqrt{\sigma^2_{x,t}}$ is the standard deviation of the variable x in period t.

Positive values of computed LISA – $I_{i,t}$ indicate clustering of the similar values, whereas the negative values of $I_{i,t}$ reveal the clustering of dissimilar values over the region and its neighbors.

The next step is to draw the significance map. From the scatter plot we can divide regions for which computed values of $I_{i,t}$ are significant into four categories:

Category	**Scatter plot quadrant**	$I_{i,t}$	**Interpretation**
High-high	upper right	positive	Region with high value of x surrounded by regions with high values of x
High-low	lower right	negative	Region with high value of x surrounded by regions with low values of x
Low-low	lower left	positive	Region with low value of x surrounded by regions with low values of x
Low-high	upper left	negative	Region with low value of x surrounded by regions with high values of x

Then we can assign for each of the category appropriate color code and draw the map with regions filled with respect to the assigned color codes. This will give the visual picture of the clustering and atypical spatial association.

Finally, according to the suggestions made by (Kamarianakis and Le Gallo, 2003), in order to obtain unbiased and consistent estimators we need to use spatial SUR estimation instead of OLS that will allow to account for temporal and spatial interdependence.

As it was mentioned earlier we need to test our regressions (5),(6),(7) on the presence of spatial dependence. This can be done using mentioned in (Anselin, 1999) LMLAG and LMERROR tests. After identifying the appropriate type of spatial regression model we need to estimate the unknown parameters using maximum likelihood estimation. Appropriate log likelihood functions for estimation of two possible specifications will be as follows:

1. For spatial lag model:

$y = \rho Wy + X\beta + \varepsilon$ or in the reduced form $y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon$

From the last equation it is clear that the spatial lag term *Wy* is correlated with the disturbances.

As it was previously mentioned consistent and unbiased estimators can be obtained from the explicit maximization of the following log likelihood function:

$$\ln L = -(N/2)\ln(2\pi) - (N/2)\ln\sigma^2 + \ln|I - \rho W| -$$
$$-(1/2\sigma^2)(y - \rho Wy - X\beta)'(y - \rho Wy - X\beta)$$

2. For spatial error model:

It is worth to mention, first, that in log likelihood function disturbances should be of the multivariate normal case (Magnus, 1978). Therefore, using expression for error terms $\varepsilon = y - X\beta$ and variance-covariance matrix $\Sigma = \sigma^2[(I - \lambda W)'(I - \lambda W)]^{-1}$ the following log likelihood function will result:

$$\ln L = -(N/2)\ln(2\pi) - (N/2)\ln\sigma^2 + \ln\left|I - \lambda W\right| -$$
$$- (1/2\sigma^2)(y - X\beta)'(I - \lambda W)'(I - \lambda W)(y - X\beta)$$

And finally we will run spatial SUR in order to get consistent estimates of the parameters.

Chapter 4

DATA DESCRIPTION

In my research I will use two different datasets for computing the three components of the labor productivities disparities between Ukrainian regions. Labor productivity is computed as the ratio between output measure to labor involved in production of that output. First of the two datasets is based on the ULFS that provides the data on regional industrial production (value of manufactured commercial products in thousand UAH (in current wholesale prices)) and the data on the quantity of workers involved in production of each industry for each region. Data is available for 6 consecutive years from 1998 to 2003. This data will yield 150 observations for estimation.

Brief description of the general parameters of the dataset obtained from ULFS dataset for further estimation of shift-share components can be summarized in the following table:

Table 2. General description of the data for shift-share analysis from ULFS dataset

<table>
<tr><td>General description of the dataset 1</td><td colspan="6">Enterprise level data from ULFS is used for computing industrial productivities and shares of labor involved in each industry production, covers:
• 16 industries
• 25 regions
• Time period – 1998-2003</td></tr>
<tr><td rowspan="5">Descriptive statistics of the variables</td><td>Variable</td><td>Obs.</td><td>Mean</td><td>St. dev.</td><td>Max</td><td>Min</td></tr>
<tr><td>y_i</td><td>150</td><td>-6,77</td><td>19,80</td><td>62,03</td><td>-46,48</td></tr>
<tr><td>m_i</td><td>150</td><td>-5,05</td><td>11,02</td><td>36,01</td><td>-24,48</td></tr>
<tr><td>p_i</td><td>150</td><td>-13,09</td><td>16,70</td><td>28,73</td><td>-59,94</td></tr>
<tr><td>a_i</td><td>150</td><td>9,36</td><td>11,35</td><td>38,48</td><td>-46,70</td></tr>
</table>

The other dataset is obtained from State statistical committee of Ukraine (SCS) and it includes the data on gross value added in Ukrainian regions for types of economic activity and also labor force involvement by types of economic activity across regions. This dataset covers the period 2001-2003 (3 years). Descriptive statistics of the second dataset is summarized in the table below:

Table 3. General description of the data for shift-share analysis from dataset of the State Statistical Committee

<table>
<tr><td>General description of the dataset 2</td><td colspan="6">Data from State Committee of Statistics on GVA and number of workers by type of economic activities includes:
• 11 types of economic activities
• 25 regions
• Time period covered 2001-2003</td></tr>
<tr><td rowspan="5">Descriptive statistics of the variables</td><td>Variable</td><td>Obs.</td><td>Mean</td><td>St. dev.</td><td>Max</td><td>Min</td></tr>
<tr><td>y_i</td><td>75</td><td>-731,4</td><td>2201,4</td><td>4130,1</td><td>-3871,0</td></tr>
<tr><td>m_i</td><td>75</td><td>-453,8</td><td>1136,3</td><td>2282,7</td><td>-2511,1</td></tr>
<tr><td>p_i</td><td>75</td><td>-302,4</td><td>2023,7</td><td>7927,8</td><td>-2455,9</td></tr>
<tr><td>a_i</td><td>75</td><td>24,8</td><td>649,3</td><td>653,1</td><td>-3161,8</td></tr>
</table>

For further analysis I need to construct spatial weight matrix. I will use the data on distances between administrative centers of Ukrainian regions. This data is available online from various sources. I will use the 25×25 matrix from the http://www.brama.com/travel/distanc.html.

Two datasets were used because of some disadvantages related to the data from ULFS. These shortcomings are as follows:

- In the ULFS dataset only the industrial output is covered, whereas in dataset from State Committee of Statistics all types of economic activities are covered;

- The number of enterprises is considerably changing over time in ULFS dataset, that influences on the consistency of the obtained variables;

- GVA is more reliable indicator in measuring labor productivity, than enterprises output. Moreover, data from SCS is obtained from an official reports of enterprises;

The only disadvantage of the dataset based on the SCS's data is the lower number of observations (only 3 years are covered). This deprives my research from the possibility to analyze labor productivity disparities over time.

C h a p t e r   5

MODEL SPECIFICATION AND EMPIRICAL RESULTS

Following the described methodology the first step is to estimate three regressions separately for each year:

$$(1) y_i = \alpha_m + \beta_m m_i + \varepsilon_m$$

$$(2) y_i = \alpha_p + \beta_p p_i + \varepsilon_p$$

$$(3) y_i = \alpha_a + \beta_a a_i + \varepsilon_a$$

For the first, ULFS dataset I obtain the following results:

Table 4. OLS estimation results for ULFS dataset 1.

Year	Model (industry mix component)			Model (productivity differential component)			Model (allocative component)		
	a	b	R^2	a	b	R^2	a	b	R^2
1998	-2,09 (-1.17)	1,14 (3.63)	**0.3637**	1,42 (0.97)	1,00 (6.95)	**0.6776**	-4,87 (-2.02)	0,01 (0.03)	**0.0000**
1999	-3,16 (-1.37)	1,15 (4.22)	**0.4359**	1,09 (0.50)	0,91 (6.19)	**0.6248**	-6,65 (-2.02)	-0,16 (-0.49)	**0.0102**
2000	-3,20 (-0.87)	1,11 (3.92)	**0.4002**	2,81 (5.97)	0,97 (0.83)	**0.5906**	-10,60 (-1.93)	0,03 (0.09)	**0.0003**
2001	0,69 (0.22)	-1,14 (-4.05)	**0.4164**	-2,52 (-0.63)	-0,58 (-3.23)	**0.3118**	5,51 (1.22)	-0,12 (-0.44)	**0.0085**
2002	-1,98 (-0.59)	1,20 (4.32)	**0.4484**	5,97 (2.10)	1,02 (7.71)	**0.7211**	-4,34 (-0.67)	-0,39 (-0.85)	**0.0305**
2003	-2,47 (-0.59)	1,30 (5.64)	**0.5803**	11,22 (2.78)	1,10 (8.53)	**0.7597**	1,53 (0.15)	-0,87 (-1.76)	**0.1183**

Estimated slope coefficients for different years for each of the components do not have stable signs, though they are statistically significant according to the values of p-value. Specifically, we obtain that except for the year

2001 signs of the slope coefficients in the regressions on industry-mix and productivity differential components are positive. For the regression on allocative component we get that except for the years 1998 and 2000 slope coefficients are negative. Due to mentioned in data description section shortcomings of the ULFS dataset 1, we may not treat those results as reliable. Moreover, obtained results are not consistent with the results of the other researchers, which use the same methodological approach.

Another important assumption about the residuals normality can also be proved by performing tests on residuals' skewness and kurtosis.Since probability is higher than 0.05, then in each of the OLS regression we can not reject the hypothesis of normally distributed residuals.

Therefore, we need to refer to the dataset 2. And first we run OLS regressions:

Table 5. OLS estimation results for State Statistical Committee dataset 2.

Year	Model (industry mix component)			Model (productivity differential component)			Model (allocative component)		
	a	b	R^2	a	b	R^2	a	b	R^2
2001	-76,32 (-0,22)	1,39 (5,31)	**0.5506**	-501,11 (-1,74)	0,86 (6,44)	**0.6436**	-702,16 (-1,67)	-1,57 (-2,60)	**0.2268**
2002	-49,42 (-0,16)	1,41 (5,24)	**0.5442**	-442,03 (-1,81)	0,87 (6,6)	**0.6546**	-632,73 (-1,74)	-1,61 (-2,54)	**0.2191**
2003	-84,40 (-0,26)	1,58 (5,9)	**0.6019**	-446,59 (-1,64)	0,92 (7,03)	**0.6825**	-744,44 (-1,74)	-1,48 (-2,29)	**0.1856**

From the table above we can see that slope coefficients have the same signs for time period 2001-2003, besides they have comparable values, slightly increasing over time. For three years the highest R^2 was obtained for the productivity differential component, it increases from 0.6436 to 0.6825. However, for industry-mix component we get comparable values of R^2 (increases from 0.5506 to 0.6019). Signs of the slope coefficients are consistent with the similar researches: positive for industry-mix and productivity differential components, and negative for the allocative components.

The residual normality check for the SCS's dataset 2 gives the satisfactory results. Thus, for dataset 2 we can not reject hypothesis of the residuals' normality, since probability is higher than 0.05 (results of the residuals' normality test are in Appendix B)

The next step in the research is to test variables on the presence of spatial autocorrelation. For this purpose we need to construct the matrix of spatial weights. As it was discussed in the methodology section we will use the distances between administrative centres of the Ukrainian regions to construct matrix of weights (See Appendix C).

At first, we compute the values of the global Moran's I statistic. Using to different datasets we obtain the following results:

Table 6. Global Moran's I-statistic (ULFS dataset 1)

Year	Yi			Mi			Pi			Ai		
	I	z	p-value	I	z	p-value	I	z	p-value	I	z	p-value
1998	0,228	2,668	0,004	0,049	0,895	0,185	0,160	1,947	0,026	-0,069	-0,280	0,390
1999	0,337	3,715	0,000	0,080	1,203	0,115	0,369	3,954	0,000	0,062	1,022	0,153
2000	0,222	2,552	0,005	0,134	1,746	0,040	0,065	1,015	0,155	-0,088	-0,444	0,329
2001	0,166	1,997	0,023	0,109	1,503	0,066	-0,136	-0,904	0,183	-0,144	-1,120	0,131
2002	0,135	1,812	0,035	0,154	1,995	0,023	0,200	2,322	0,010	0,171	2,053	0,020
2003	0,083	1,239	0,108	0,146	2,008	0,027	0,167	1,926	0,022	0,217	1,239	0,005

From this table we see that no clear inferences can be made about spatial association between regions with respect to each of the variable over the six year period. For all four variables computed values of global Moran's I statistic do not reveal stable significance. However, for the second dataset 2 we obtain more sound results:

Table 7. Global Moran's I-statistic (SCS dataset 2)

Year	Yi			Mi			Pi			Ai		
	I	z	p-value	I	z	p-value	I	z	p-value	I	z	p-value
2001	0,215	2,235	0,013	0,556	5,269	0,000	-0,118	-0,816	0,207	-0,126	-1,506	0,066
2002	0,160	1,754	0,040	0,391	3,787	0,000	-0,096	-0,569	0,285	-0,127	-1,300	0,097
2003	0,210	2,183	0,015	0,396	3,846	0,000	-0,057	-0,160	0,436	-0,109	-1,209	0,113

These results provide more consistent evidence of spatial autocorrelation over time. Now we can make inferences about the presence of spatial autocorrelation between Ukrainian regions with respect to productivity differences (Yi) and Industry-mix component (Mi), basing on the Global Moran's I statistic. Since computed values of Moran's I statistic are not significant for the rest two shift-share components we can not accept the hypothesis of the presence of spatial autocorrelation between Ukrainian regions with respect to productivity differential (Pi) and allocative (Ai) components.

Now we can proceed straight to the issue of regional clustering with respect to the productivity disparities. For this purpose we can use the Moran scatter plot tool. This involves computation of the local indicators of spatial association (LISA) (see appendix 4). Due to determined inconsistency of the ULFS dataset 1, for this particular issue we will use SCS dataset 2. Regions for which computed values of the LISA were significant, according to the rule described in methodology section can be placed within four categories in the following way:

Table 8. Categorization of regions with respect to productivity disparities

Year	Category	Regions
2001	High-High	Zaporizska, Kharkivska, Donetska
	High-Low	Volynska
	Low-Low	Khmelnytska, Rivnenska, Vinnytska, Ternopilska
2002	High-High	Zaporizska, Kharkivska, Donetska
	High-Low	Volynska
	Low-Low	Rivnenska, Vinnytska, Ternopilska
2003	High-High	Zaporizska, Kharkivska, Donetska
	High-Low	Volynska
	Low-Low	Rivnenska, Vinnytska, Ternopilska

Now we can make a clear inference about spatial pattern of the distribution of labor productivity disparities across Ukrainian regions. There are to clusters with respect to the labor productivity disparities within the borders of the country: western cluster with low productivity per worker (Rivnenska, Vinnytska, Ternopilska and Khmelnytska regions) and eastern cluster with high labor productivity (Zaporizska, Kharkivska and Donetska regions). We also obtain one significant atypical spatial association. Specifically, Volynska region having high labor productivity is surrounded by regions with low productivity. In order to have clear visualization of the clusters we use Moran significance maps (see Appendix D)

The last part of empirical analysis is dedicated to the specification of the most suitable model, and subsequent estimation of its parameters. At first we check the appropriateness of SURE instead of separate OLS regressions. Separate regressions do not take into account the possibility of intertemporal correlations between the residuals of the estimated equations (5), (6), (7), whereas in SURE this possibility is explicitly introduced in the estimation. For comparison between the two estimation techniques we use the Breush-Pagan test for independence of the residuals of the three regressions. On the basis of the mentioned test we can not accept the hypothesis of independence of the residuals:

Table 9. Breush-Pagan test for independence of the residuals

Correlation matrix of residuals:			
	yi	2yi	3yi
yi	1.0000		
2yi	-0.1755	1.0000	
3yi	0.3883	0.8360	1.0000
Breusch-Pagan test of independence: **chi2(3) = 66.037, Pr = 0.0000**			

We can also mention here that standard deviations of the estimated parameters by SURE is lower, than that of OLS, therefore the usage of

SURE is more appropriate here, than the estimation by three separate OLS regressions.

Further specification search goes in line with detecting of the spatial autocorrelations, and if its significance will be proven then, appropriate corrections and estimation techniques will be used.

First of all we will check the form of spatial dependence if any is present in the given framework. Tests and the decision rules are described in the methodology section. The results of the performed tests are as follows:

Table 10. Results of the spatial regression specification tests

	Model					
	Industry-mix		Productivity differential		Allocative	
Test	Statistic	p-value	Statistic	p-value	Statistic	p-value
Spatial error:						
Lagrange multiplier	0.392	**0.531**	35.205	**0.000**	30.006	**0.000**
Robust Lagrange multiplier	0.218	**0.641**	10.427	**0.001**	5.601	**0.018**
Spatial lag:						
Lagrange multiplier	0.188	**0.664**	41.905	**0.000**	24.763	**0.000**
Robust Lagrange multiplier	0.014	**0.907**	3.727	**0.054**	0.358	**0.550**

Therefore, only for productivity differential and allocative models we should account for spatial dependence, which takes the form of spatial error.

In order to get consistent and unbiased estimator of parameter λ we need to perform an explicit maximization of the described in the methodology section log likelihood function. Maximization yields the following values of the parameters:

- For the productivity differential model:

Table 11. Lambda (λ_p) coefficient for productivity differential component model

	value	t-stat	p-value
lambda	**0.6639907**	**8.18**	**0.000**
Tests:			
Wald test of lambda=0:	chi2(1) = 66.963 (0.000)		
Likelihood ratio test of lambda=0:	chi2(1) = 36.056 (0.000)		
Lagrange multiplier test of lambda=0:	chi2(1) = 35.205 (0.000)		
Acceptable range for lambda: -1.375 < lambda < 1.000			

- For the allocative component model:

Table 12. Lambda (λ_a) coefficient for allocative component model

	value	t-stat	p-value
lambda	**0.5706268**	**6.13**	**0.000**
Tests:			
Wald test of lambda=0:	chi2(1) = 37.517 (0.000)		
Likelihood ratio test of lambda=0:	chi2(1) = 25.488 (0.000)		
Lagrange multiplier test of lambda=0:	chi2(1) = 30.006 (0.000)		
Acceptable range for lambda: -1.375 < lambda < 1.000			

Further estimation of spatial SUR specification needs the following adjustment of the variables, so-called spatial filtering:

- For productivity differential component

$p_f = (I - \lambda_p W) p_i$;

- For allocative component

$a_f = (I - \lambda_a W) a_i$, where

I – identity matrix

W – matrix of weights

λ_p, λ_a - parameters obtained in the previous step

Finally, we estimate the spatial SUR model:

Table 13. Estimation results for spatial SUR

Spatial SUR estimation					
Equation		Obs	Parms	R^2	
Industry-mix component		75	1	**0.5626**	
Productivity differential component		75	2	**0.8089**	
Allocative component		75	2	**0.4007**	
		Coef.	Std.Err.	z	P>z
1	yi mi	**1.557655**	0.081781	19.05	0.000
2	yi pi	**0.6780605**	0.0245831	27.58	0.000
3	yi ai	**-1.111799**	0.1092339	-10.18	0.000

On the basis of information criteria test (see Appendix E) our inferences should be based upon the estimation results of the spatial SUR model.

First of all we again look at the R^2 of each of the three components. According to this criteria labor productivity differentials across the Ukrainian regions can be attributed mostly to the productivity differential component (R^2 = 0.8089). The slope coefficient is equal to 0.6781 we can interpret as the magnitude of labor productivity disparities growth if the productivity differentials across regions will increase by one unit. Loosely speaking, with uniform increase of labor productivity in retarded regions we will get an increase in difference of that regions' productivity from average productivity in Ukraine. This finding coincides with the conclusion of the other researchers that use the same methodological approach for other countries. However, for Ukraine we find that the variation in labor productivity disparities can also be attributed to the industry-mix component (R^2 = 0.5626). And it is also has positive relation with the dependent variable (slope coefficient is equal to 1.557655). Slope coefficient for the third, allocative component, has negative sign and also the least explanatory power of labor productivity disparities across Ukrainian regions (R^2 = 0.4007).

C h a p t e r 6

CONCLUSIONS AND POLICY RECOMMENDATIONS

The data on labor productivity across Ukrainian regions presented in the introduction section reveals considerable differences in productivity levels within the borders of Ukraine. The results of this research based on the shift-share analysis disclose the sources of such labor productivity inequalities across Ukraine. Moreover, we show the statistical evidence of the presence of regional clustering with respect to the labor productivity levels.

So first of all we may outline that the labor productivity disparities are mainly attributable to the uniform productivity differences across regions and to the smaller extent they can be ascribed to the differences in types of economic activities prevailing in creation of value added across Ukrainian regions (industry-mix). These findings are basically consistent with the similar researches undertaken mostly for regions of European Union. However, what was not typical for European regions is that in Ukraine productivity disparities can also be attributed to the industry-mix in each region.

The second important conclusion is the presence of two productivity clusters in Ukraine. Three western regions Rivnenska, Ternopilska, and Vinytska form the low productive cluster and three eastern regions Donetska, Kharkivska and Zaporizhska form the high productive one. We also find that Volynska oblast fall into the atypical category of spatial association (region with high labor productivity surrounded by regions with low productivity).

On the basis of the results we obtain in this research the following policy recommendations can be made, in order to improve existing situation:

1. For the defined low productive cluster should be implemented such policy that will result in labor productivity growth in all sectors that produce the major share of regions' value added.

2. Policy actions that will stimulate regional specialization in highly productive sectors should also be undertaken in order to improve the regional aggregate productivity.

There should also be noted that this paper provides the suggestion for general direction of the policy actions. And further research in the field of productivity study can disclose the factors that will cause the growth of productivity in the retarded Ukrainian regions, as well as in the whole country.

BIBLIOGRAPHY

Anselin, Luc. *Estimation methods for spatial autoregressive structures.* Regional Science Dissertation and Monograph Series 8. Field of Regional Science, Cornell University, Ithaca, N.Y., 1980.

Anselin, Luc. *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers: Dordrecht (NL), 1988.

Anselin, Luc. *Spatial econometrics* Bruton Center School of Social Sciences University of Texas at Dallas Richardson, 1999.

Anselin, Luc. *Spatial econometrics. In Companion to econometrics*, edited by B. Baltagi. Basil Blackwell, Oxford, 2001.

Anselin L., and Florax R, *New Directions in Spatial Econometrics*, pp. 3–18. Berlin: Springer-Verlag, 1995.

Benito, J.M., Ezcurra, R. *Spatial Disparities in Productivity and Industry Mix: the Case of the European Regions*, European Urban and Regional Studies, 12(2), 177-194, 2005.

Cliff AD, and Ord JK *Testing for spatial autocorrelation among regression residuals.* Geographical Analysis 4, 267–284, 1972.

Cliff AD, and Ord JK *Spatial processes: models and applications.* Pion, London, 1981.

Cuadrado-Roura JR *Regional convergence in the European Union: from hypothesis to the actual trends.* Annals of Regional Science 35:333-356, 2001.

Dall'erba S *Distribution of regional income and regional funds in Europe 1989-1999: an exploratory spatial data analysis.* Annals of Regional Science, 2003.

Dickerson A. *Sectoral productivity differences across the UK*, A report prepared for the: Sector Skills Development Agency, 2005.

Dunn Edgar S. Jr. *A statistical and analytical technique for regional analysis. Papers and* Proceedings of the Regional Science Association 6:97-112, 1960.

Esteban J *A reinterpretation of shift-share analysis.* Regional and Urban Economics, 2:249-261, 1972.

Esteban J *Regional convergence in Europe and the industry mix: A shift-*

share analysis. Regional Science and Urban Economics 30:353-364, 2000.

Kamarianakis, Yiannis and Le Gallo, Julie *The evolution of regional productivity disparities in the European Union, 1975-2000* Cahiers du GRES (Groupement de Recherches Economiques et Sociales) , 2003.

Kelejian H., and Prucha I, *A generalized moments estimator for the autoregressive parameter in a spatial model.* International Economic Review, 1999.

Magnus J., *Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix,* Journal of Econometrics, 7:281–312. Corrigenda, Journal of Econometrics 10, 261, 1978.

Ord J.K, *Estimation methods for models of spatial interaction.,* Journal of the American Statistical Association 70, 120–126, 1975.

Solow, Robert M. *A Contribution to the Theory of Economic Growth.* Quarterly Journal of Economics, 70:65-94, 1956

Solow, Robert M. *Technical Change and the Aggregate Production Function.* Review of Economics and Statistics 39: 312—320, 1957

Swan, Trevor *Economic Growth and Capital Accumulation.* Economic Record, Vol. 32 (2), p.334-61, 1956.

Tibshirani R *Estimating optimal transformations for regression. Journal of the.* American Statistical Association 83:394-405, 1987.

http://www.brama.com/travel/distanc.html

APPENDIX A. AGGREGATE PRODUCTIVITY PER WORKER IN THE REGIONAL ASPECT

2001		2002		2003	
11572	Dniprop	9812	Dniprop	11907	Donetsk
10663	Donetsk	9805	Volyn	11497	Dniprop
10596	Volyn	9494	Donetsk	11450	Volyn
9909	Odesa	8813	Poltava	10422	Odesa
9041	Poltava	8471	Odesa	10080	Poltava
8871	Zaporizh	7843	Zaporizh	10052	Zaporizh
8456	kharkiv	7764	Lugansk	9031	kharkiv
7989	Lugansk	6819	kharkiv	9015	Lugansk
6908	Sumy	6202	crimea	7725	crimea
6775	Mykolayiv	6080	Lviv	7612	Lviv
6703	crimea	5746	Kyiv_city	7478	Kyiv_city
6449	Lviv	5649	Mykolayiv	7199	Ivano-fr
6052	Kyiv	5402	Ivano-fr	6881	Mykolayiv
5639	chernigiv	5338	Sumy	6347	Kirovograd
5382	Kirovograd	5068	Kirovograd	6082	Kherson
5290	Kherson	4770	chernigiv	5930	cherkasy
5036	Vinnytsia	4622	Zhytomyr	5895	Sumy
4955	Rivne	4587	Kherson	5748	Zhytomyr
4836	Zakarp	4586	cherkasy	5739	chernigiv
4817	cherkasy	4246	Rivne	5294	Rivne
4756	Ivano-fr	4070	Zakarp	5279	Ternopil
4536	Zhytomyr	4059	Vinnytsia	5152	Zakarp
4062	Khmelnyt	3998	Ternopil	4953	Vinnytsia
3882	Ternopil	3986	chernivtsi	4855	chernivtsi
3571	chernivtsi	3300	Khmelnyt	4622	Khmelnyt

APPENDIX B. RESULTS OF THE RESIDUALS NORMALITY TEST

Table 1. Skewness/Kurtosis tests for normality of residuals (ULFS-dataset1)

Year	Model	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	Prob>chi2
1998	Industry-mix	0.082	0.145	5.05	0.0799
	Productivity differential	0.442	0.029	5.22	0.0735
	Allocative	0.032	0.291	5.48	0.0646
1999	Industry-mix	0.165	0.700	2.29	0.3178
	Productivity differential	0.967	0.281	1.26	0.5326
	Allocative	0.026	0.353	5.52	0.0632
2000	Industry-mix	0.331	0.236	2.60	0.2725
	Productivity differential	0.505	0.484	1.00	0.6062
	Allocative	0.227	0.807	1.66	0.4357
2001	Industry-mix	0.907	0.098	3.06	0.2160
	Productivity differential	0.272	0.263	2.73	0.2555
	Allocative	0.360	0.481	1.45	0.4839
2002	Industry-mix	0.214	0.892	1.71	0.4256
	Productivity differential	0.394	0.157	3.03	0.2195
	Allocative	0.150	0.306	5.42	0.0611
2003	Industry-mix	0.232	0.305	2.75	0.2531
	Productivity differential	0.646	0.290	1.45	0.4855
	Allocative	0.086	0.499	3.73	0.1548

Table 2. Skewness/Kurtosis tests for normality of residuals (SSC-dataset)

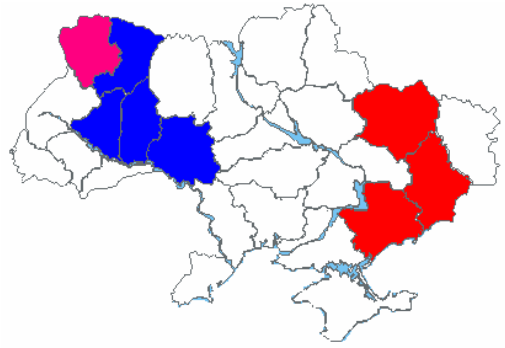
Year	Regression	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	Prob>chi2
2001	Industry-mix	0.089	0.156	3.97	0.1224
	Productivity differential	0.476	0.362	1.46	0.4827
	Allocative	0.103	0.186	3.73	0.1251
2002	Industry-mix	0.094	0.176	3.75	0.1157
	Productivity differential	0.333	0.812	1.07	0.5865
	Allocative	0.026	0.678	5.05	0.0801
2003	Industry-mix	0.091	0.194	3.45	0.1358
	Productivity differential	0.772	0.516	0.53	0.7688
	Allocative	0.037	0.924	4.46	0.1076

APPENDIX C. RESULTS OF THE RESIDUALS NORMALITY TEST

	Cherkasy	Chernihiv	Chernivtsi	Crimea	Dnipropetrovsk	Donetsk	Ivano-Frankivsk	Kharkiv	Kherson	Khmelnytskyj	Kirovohrad	Kyiv	Luhansk	Lviv	Mykolayiv	Odessa	Poltava	Rivne	Sumy	Ternopil	Vinnytsya	Lutsk	Uzhorod	Zaporizhia	Zhytomyr
Cherkasy	0	0,14	0	0	0,04	0	0	0	0	0	0,35	0,14	0	0	0	0	0,09	0	0,1	0	0,04	0	0	0	0
Chernihiv	0,07	0	0	0	0	0	0	0	0	0	0	0,26	0	0	0	0	0	0	0,1	0	0	0	0	0	0,06
Chernivtsi	0	0	0	0	0	0	0,23	0	0	0,11	0	0	0	0,06	0	0	0	0,03	0	0,1	0,05	0,03	0	0	0
Crimea	0	0	0	0	0	0	0	0	0,05	0	0	0	0	0	0,03	0	0	0	0	0	0	0	0	0	0
Dnipro-petrovsk	0,07	0	0	0	0	0,18	0	0,16	0,04	0	0,09	0	0	0	0,03	0	0,19	0	0	0	0	0	0	0,67	0
Donetsk	0	0	0	0	0,07	0	0	0,1	0	0	0	0	0,83	0	0	0	0	0	0	0	0	0	0	0,09	0
Ivano-Frankivsk	0	0	0,33	0	0	0	0	0	0	0,06	0	0	0	0,24	0	0	0	0,03	0	0,16	0	0,04	0,46	0	0
Kharkiv	0	0	0	0	0,09	0,14	0	0	0	0	0	0	0,17	0	0	0	0,3	0	0,34	0	0	0	0	0,06	0
Kherson	0	0	0	0,6	0,04	0	0	0	0	0	0,09	0	0	0	0,65	0,27	0	0	0	0	0	0	0	0,06	0
Khmelnytskyj	0	0	0,18	0	0	0	0,08	0	0	0	0	0	0	0,07	0	0	0	0,07	0	0,22	0,33	0,04	0	0	0,11
Kirovohrad	0,46	0	0	0	0,07	0	0	0	0,06	0	0	0,07	0	0	0,1	0,1	0,1	0	0	0	0,05	0	0	0,05	0
Kyiv	0,18	0,59	0	0	0	0	0	0	0	0	0,06	0	0	0	0	0	0,05	0,03	0,1	0	0,07	0	0	0	0,25
Luhansk	0	0	0	0	0	0,49	0	0,07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lviv	0	0	0,09	0	0	0	0,26	0	0	0,06	0	0	0	0	0	0	0	0,06	0	0,19	0	0,13	0,54	0	0
Mykolayiv	0	0	0	0,4	0,04	0	0	0	0,73	0	0,17	0	0	0	0	0,63	0	0	0	0	0	0	0	0	0
Odessa	0	0	0	0	0	0	0	0	0,09	0	0,05	0	0	0	0,18	0	0	0	0	0	0	0	0	0	0
Poltava	0,1	0	0	0	0,13	0	0	0,37	0	0	0,09	0,05	0	0	0	0	0	0	0,35	0	0	0	0	0,07	0
Rivne	0	0	0,06	0	0	0	0,06	0	0	0,1	0	0,06	0	0,1	0	0	0	0	0	0,12	0,05	0,61	0	0	0,14
Sumy	0,06	0,12	0	0	0	0	0	0,22	0	0	0	0,05	0	0	0	0	0,19	0	0	0	0	0	0	0	0
Ternopil	0	0	0,22	0	0	0	0,25	0	0	0,28	0	0	0	0,28	0	0	0	0,11	0	0	0,09	0,11	0	0	0,05
Vinnytsya	0,06	0	0,07	0	0	0	0	0	0	0,25	0,05	0,08	0	0	0	0	0	0,03	0	0,05	0	0	0	0	0,31
Lutsk	0	0	0,06	0	0	0	0,06	0	0	0,05	0	0	0	0,19	0	0	0	0,56	0	0,11	0	0	0	0	0,07
Uzhorod	0	0	0	0	0	0	0,05	0	0	0	0	0	0	0,06	0	0	0	0	0	0	0	0	0	0	0
Zaporizhia	0	0	0	0	0,53	0,19	0	0,08	0,04	0	0,06	0	0	0	0	0	0,08	0	0	0	0	0	0	0	0
Zhytomyr	0	0,15	0	0	0	0	0	0	0	0,09	0	0,3	0	0	0	0	0	0,08	0	0,03	0,31	0,04	0	0	0

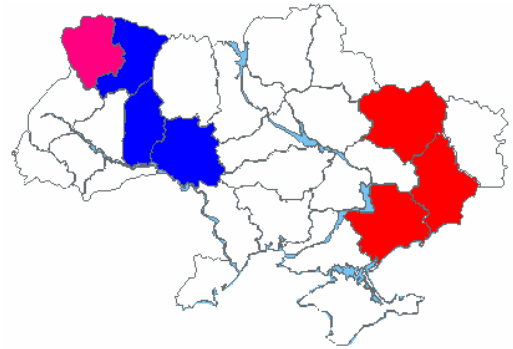
APPENDIX D. MORAN SIGNIFICANCE MAP FOR LABOR PRODUCTIVITY DISPARITIES

2001



Category	Regions
High-High	Zaporizka Kharkivska Donetska
Low-Low	Khmelnytska Rivnenska Vinnytska Ternopilska
High-Low	Volynska

2003



Category	Regions
High-High	Zaporizka Kharkivska Donetska
Low-Low	Rivnenska Vinnytska Ternopilska
High-Low	Volynska

APPENDIX E. INFORMATION CRITERIA TEST FOR SUR AND SPATIAL SUR

Model	Obs	ll(model)	AIC	BIC
SUR	75	-1477.65	2967.3	2981.205
Model	Obs	ll(model)	AIC	BIC
Spatial SUR	75	-1798.585	3613.169	3631.709