Consortium for Economic Policy Research and Advice

Canadian
International
Development
Agency

Association
of Universities and
Colleges
of Canada

Institute
for the Economy
in Transition

WCER

Working
Center
for Economic
Reform

Academy
of National
Economy

Analysis of Economic Growth in Regions: Geographical and Institutional Aspect

Moscow
IET
2007

Authors: Chapter 1 – O. Lugovoy, V. Dashkeyev, I. Mazayev, D. Fomchenko, E. Polyakov; Chapter 2 – A. Hecht.

This study explores disparity in regional development in Russia and in Canada and role of geography in their development. In the first chapter analysis of the role of geographic, economic, and institutional factors in economic growth over 1996–2004 is presented. Additionally the issue of the interregional spatial interaction is analyzed, which is studied in the framework of the new economic geography. The second chapter devoted to the development of Canadian provinces and analyzes factors of provincial differences along with factors of migration in Canada over the period of 1991–2001.

Table of Contents

Chapter 1. Analysis of Economic Growth in Russian Regions: Geographical and Institutional Aspect

Introduction

The role of geography and institutions in the long-run economic development is being extensively discussed in the recent economic literature. Whether prospects for the future economic growth are the same for countries with different climate, natural resource endowment, and geographic location or not, is a question that can be applied to regions of a single country, especially to such a vast federation like Russia that consists of 89 administrative regions. Economic growth in Russia, which commenced soon after financial crisis of August 1998, is spatially heterogeneous and depends on numerous factors with distinct geographic features.

This paper is a sequel to the previous one, published as a part of IET & CIDA research project[1], which made emphasis on analysis of convergence of income among Russian regions and in which an attempt of growth accounting was made. With this paper we intended, first, to investigate how post-crisis growth in Russia influenced the relative development of distinct regions, and whether there was evidence of diminishing regional disparity during this period. Second, to quantitatively estimate influence of different external factors (i.e. specific local conditions) on economic growth in regions and on direct factors of production (i.e. investment and population migration). This study of regional development of Russian regions differs from the preceding one by the adjustment of GRP to take account of regional price levels and by utilization of the spatial econometrics methods to take account of spatial autocorrelation of control variables.

The first section reviews recent literature on growth, paying special attention to the role of geographic factors, and forms base for the following empirical research. We treat geographic features of Russian regions – geographic location, climate, natural resource endowment, settlement

---

[1] *Drobyshevsky S., Lugovoy O., Astafyeva E. et al. (2005).*

patterns, development of telecommunications and transport infrastructure – as determinants of economic growth.

In the second section we analyze economic growth in Russian regions over the period of 1996–2004 and find statistically significant evidence neither for sigma convergence nor for unconditional beta convergence, while utilization of spatial framework and estimation of conditional beta convergence model with spatial lags yielded statistically significant results. We argue for the spatial spill-over effect, i.e., economic growth of a given region is significantly related to economic growth of its neighboring regions, and that the closer they are located and the larger they are as measured by per capita GRP, the greater is their influence. Next, we estimate a simultaneous equations model to explain accumulation of factors of growth, thus, growth itself by interregional differences in geography, institutions, and human capital in Russian regions by 3SLS and FIML (to take account of spatial dependence). We model growth of per capita GRP with equations for investment and migration flows. Estimation results are used for decomposition of economic growth and construction of 2 topologies of growth factors. In the concluding section we present implications for regional policy.

Determinants of Economic Growth: Recent Research Trends

Discussion about causes of differences in international development has extensive history, for example, one can recall classic paper of Adam Smith "Inquiry into the nature and causes of the wealth of nations". Traditional neoclassic growth theory treats accumulation of labor and capital (growth of population, employment, investment) and innovations as fundamental determinants of growth. School of new institutional economics asserts that accumulation of growth factors, education, and innovations

constitute the growth itself[2]. In its view, the underlying factor of growth is the institutional factor, i.e. rules and traditions, maintained by the society, which stimulate (or prevent) accumulation of factors, generation and application of new ideas.

An idea that institutions are endogenous by themselves spread fast recently[3]. *D.Acemoglu, S.Johnson, J.Robinson (2005)* in their well known paper single out three possible factors of institutions: rules, established by people, geography, and culture.

In modern empiric research the quantity and types of factors that potentially influence economic growth outnumber by a significant margin those traditionally employed by neoclassical theory and include some fields that were traditionally out of economists' interests[4]. *Robert Barro (1997)* demonstrates statistical significance of numerous factors in explanation of levels and rates of economic growth.

Growth determinants can be divided into several groups. One variant of possible classifications is presented in Dani Rodrik's work[5] (*Fig. 1.1*). This classification divides factors of growth into 2 groups: proximate and deep. Into proximate group Rodrik places traditional factors of production (augmentation of physical capital, accumulation of human capital) and productivity growth including total factor productivity. Growth accounting gained in popularity in 1960s and 1970s after the seminal papers of *R.Solow (1957)*, *E.Denison (1962, 1967)*, *D.Jorgenson and Z.Griliches (1967)*. The abovementioned can be referred to as to the first stage of growth factors analysis, i.e. at the level of factors utilization. What is the underlying reason behind changes in proximate factors is still under discussion.

In the group of deep factors Dani Rodrik places three subgroups of factors: external trade, institutions (partially endogenous factor), and geography (absolutely exogenous factor). Under this scheme, deep factors are the main determinants of economic growth and differences among countries in levels of development.

---

[2] *North, Thomas (1973).*
[3] see, e.g., *Rodrik D. (2003); Acemoglu D., S. Johnson, J. Robinson (2005).*
[4] *Sala-I-Martin X (1997).*
[5] *Rodrik D., (2003), p. 1–19.*
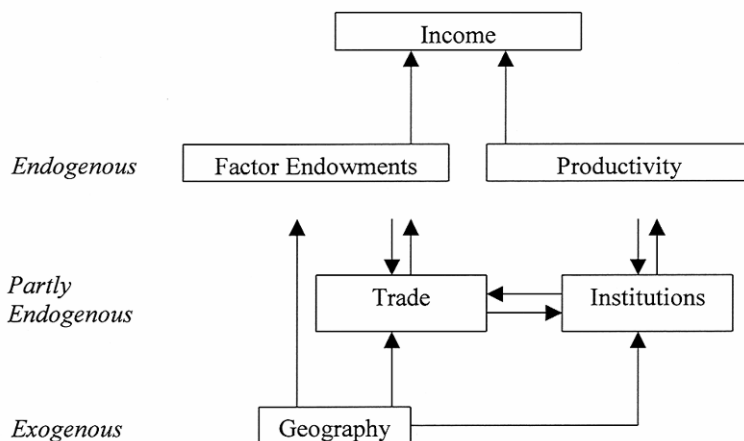
Fig. 1.1. Dani Rodrik's Classification of Factors of Growth

Econometric analysis of international growth was stimulated by publication of Alan Heston and Robert Summers' tables with statistical data, also known as Penn World Tables[6]. These tables made international comparisons of development feasible[7]. Notwithstanding popularity of such research, some economists are skeptical about results of international comparisons and policy recommendations drawn on their basis, since, first, existing differences among countries and recipes for future economic policy depend on the current nation-specific policies and circumstances, second, the same recommendations will result in different outcomes if blindly applied to countries that differ in levels of technological development.

Jonathan Temple (1999) reviews advantages and disadvantages of either approach in his survey of empiric research of economic growth and makes a notice that different approaches are complementary, rather than

---

[6] See http://pwt.econ.upenn.edu.
[7] Among which the most prominent are *Mankiw, Romer, Weil (1992); Barro R.. (1997); Barro R.., X. Sala-i-Martin (2003).*

mutually exclusive. *Dani Rodrik (2003)* also stresses importance of combination of econometric country comparisons with conclusions form country case studies (so called analytical regional geography).

Specific Features of Cross-Regional Growth Studies

Differences among different countries and among regions of a country are not quite the same. A country has a number of common characteristics. For example, each country can be characterized by common macroeconomic conditions; absence of formal barriers to movement of labor, capital; and common system of formal institutions (institutional differences between regions are usually significantly lower than those between countries); normally all regions of one country are affected by changes in national policy, exchange rate dynamics, external shocks etc.

In market economy administrative borders have a very limited influence on the life of society. Economic regions are usually not attached to the carcass of administrative division of the country's territory[8].

It is also true that labor and capital mobility among regions is ceteris paribus higher than among countries. Common language and culture, absence of formal barriers to migration and investments lead to higher integration within states. Of course different types of resources have different mobility, e.g. farming land or mineral resources are absolutely immobile. People may differ in their desire to migrate, which depends not only on differences in wages, but also on their cultural values and attachment to place[9]. If the mobility is high, regional development becomes divergent: regions with higher labor productivity attract more labor and capital from disadvantaged regions. Growth of productivity in such regions cannot fully compensate for decline in income that arises from decline in output. On the other part, regions that attract migrants are growing faster: growth of employment in such regions occurs along with growth of output. This process leads to divergence of development levels: high-high and low-

---

[8] United States represents a good example of such differences. See *Smirnyagin L.V. (1989).*

[9] See, e.g., *Tuan Yi Fu (1998).*

low clustering in space, i.e. wealthier regions abut on wealthier regions while poorer regions abut on poorer regions[10].

Geography and Economic Growth

The question of geography's influence on human development has always been of great interest to scientists. Ancient Greek and Roman philosophers tried to explain differences in character and mentality of people, technical and military success and failure of states by differences in climate, relief, their coastal or inland location and other features of geography. In the last centuries geographical determinism (nature determines development) and indeterminism (geography has no impact on human development) in turn became dominant paradigms.

For a rather long while economy and geography had common fields of study. *Regional economics* was dealing with modeling of economic processes (location theory, regional analysis, and urban economics) and *economic geography* specialized on description of existing spatial structures – distribution in space of regions, cities, industrial networks etc.

From the outset, neoclassical economic theory didn't pay much attention to geography of the countries, whose growth it tried to explain. This situation changed in recent decades when many articles on geography and growth were published. On the one hand, a big advance in modeling of spatial dynamics was made by Paul Krugman et al., who modeled development under assumption of positive returns to scale in industries and took account of agglomeration effect (see next part 'New economic geography'). On the other hand, geographical features of the countries (climatic conditions, latitude, coastal or landlocked location and many others) were included in econometric models as exogenous determinants of growth[11]. For example *Gallup & Sachs (1998)* show significant negative dependence between latitude and PPP-adjusted GDP per capita, which they try to explain it via different factors (harsh climate, spread of tropical diseases, colonial history and others). In these studies it's also shown that mineral resource abundant countries tend growth at a slower pace than others.

---

[10] *Baumont et al. (2003).*
[11] *Gallup et al. (2002); Hausmann (2001); Gallup, Sachs (1998).*

An appearance of geography in growth econometrics (alongside with institutions) was a natural reaction of science to incapacity of endogenous growth models in explanation of development differences across countries. In fact, very few developing countries managed to sustain growth after World War II despite optimistic forecasts of scientists and reasonable financial support from international organizations[12].

In our study geographical differences between regions are presented by variables that represent physical features (climatic conditions, mineral fuel deposits) and economic geography (access to seacoast, quality of infrastructure, geographical proximity to markets etc).

Differences between regions of Russia are significant in several aspects: physical geography features differ radically across Russia; its regions have different climatic conditions, different types of soils; vegetation, fauna and mineral resources are spread extremely irregularly across territory.

Historically (in general terms) the colonization of modern Russia's territory was shifting from comfortable to extreme conditions of living. Different time of settlement is a major factor of contemporary differences in specialization, level of development, and institutional features.

All elements of economic landscape – pattern of settlement, location of industries, objects of infrastructure (railroads and other networks) – are highly path-dependent categories.

Obviously, one can't say that geography is the only factor that determines economic development, but some geographic factors may be significant growth determinants. Also some geographical variables can be useful as proxy variables for some other factors that are hard to measure directly (for example, institutions). Here we assume that geography matters for growth along with many other factors and may constitute comparative advantages or disadvantages of regions.

New Economic Geography

In recent years models of new economic geography (NEG)[13] have gained in importance. These types of models deal with increasing returns

---

[12] *Easterly W. (2002).*
[13] *Fujita et al. (1999); Fingleton (2002).*

that are fueled by Dixit-Stiglitz monopolistic competition mechanism[14]. This mechanism is linked to transport costs and economic-geographical factors. Under assumption of imperfect competition firms' decisions about geographical location become significant. Ceteris paribus firms will try to decrease their transport costs by locating their factories closer to big markets with higher potential demand on their products.

The concept of market potential, which represents the power of economic interaction between a given region and other regions, is crucial for new economic geography. The concept is based on a gravity model that assumes that trade (and intensity of economic interactions) between two regions is positively correlated with their gross regional products (that represents economic 'mass' of regions) and negatively correlated with the distance between them. For each region its market potential can be written as follows (*Isard, (1960)*):

$$MP_i = \sum_{j \neq i} \frac{GRP_j}{d_{ij}^2}.$$

In other words, market potential is an aggregate demand of all other regions for goods produced in a given region if transport costs were taken into account. Theoretical implications of NEG predict emergence of the agglomeration effect and core-periphery type of development. In terms of spatial specification of convergence model this means that steady-state growth rate of a given region depends on the spatial cluster (core or periphery) where this region is located.

Theoretically the effect of spatial clustering may be explained by (1) market potential and transport costs, or (2) agglomeration and spillover effects (innovation diffusion and migration of qualified labor).

The concept of market potential assumes that regions get positive externality from growth of demand in their neighborhood – this leads to growth of trade and overall economic growth. The scale of this external effect correlates positively with neighbors' growth rates and negatively with distance that increases transport costs.

---

[14] *Dixit, Stiglitz (1977).*

In some recent studies NEG models are used together with endogenous growth models in order to explore interdependence between agglomeration and growth[15]. These studies show that agglomeration and growth have reinforcing impact on each other, that leads to faster growth of agglomerations and metropolitan areas and fosters overall economic growth. The main feature of this impact is innovation diffusion and migration of qualified workers. All innovations (including technologies, models of business development, institutional changes etc.) are subject to diffusion in space. Spillovers allow regions to take advantage of more efficient technology (institution) and increase their factor productivity.

The process of core-periphery divergence can lead to faster growth and higher differences in level of income at the same time. In the long run such model of growth is beneficial for *all* regions. Regions of the "core" are growing faster because of agglomeration effect and specialization on "new" goods. At the same time peripheral regions get extra benefits from increase in demand for their "traditional" goods thanks to overall economic growth. Thus regions of periphery are better-off compared to situation without agglomeration effect in place. In this paper we use NEG models for study of convergence, and agglomeration is one of economic-geographical variables in our growth regressions.

Determinants of Economic Growth in Russian Regions: Empirical Analysis

This part presents empirical study of growth factors in Russian regions. The main objective of this research is to determine so-called "deep" determinants of economic growth in accordance with described in theoretical part logic. Below influence of available to us geographic, infrastructural and institutional variables on income, growth rates, convergence of income across regions, and their steady-state levels is tested.

It's important to notice that the time period covered by our study is rather short for identification of long-run trends: 1997–2004 (1998–2004

---

[15] *Fujita, Thisse (2002); Martin, Ottaviano (1999, 2001); Baldwin, Martin, Ottaviano (2001).*

for convergence)[16]. Thus, acquired results are more likely to describe growth determinants of the recovery period, which commenced after 1998 crisis.

In this chapter we also formulate some main assumptions about the influence of deep determinants on growth and convergence.

Data and Methodology

Russian Federation had 89 regions during the period under consideration[17]. These regions are not equal in some sense. Some autonomous okrugs are at the same time parts of other regions (for example, Hanty-Mansiyskiy and Yamalo-Nenetskiy autonomous okrugs are included in Tumen Oblast) and for such regions (Autonomous okrugs) most of statistical data are not available. Alongside with autonomous okrugs, three more regions were excluded: Chechen Republic (no data available), Ingush Republic(unreliable data), and Chukotka (unreliable data)[18]. Thus, here we use a panel of 77 regions.

Since determination of common for regions of Russia growth factors was main objective of our study, we excluded regions that were caused by exogenous shocks, don't follow general trend, and thus can't be described by our model.

As for intra-regional convergence of income levels, we must take into account fact that purchasing power of ruble differs substantially over regions. In international comparisons this problem is resolved by estimation of Purchasing Power Parity GDP (PPP GDP). Some regional studies of Russia use index of survival minimum (published by Rosstat), but this index is calculated on basis of a consumer goods basket that is limited in number of included goods and not representative for total GRP. Official estimates of PPP for Russian regions are absent, hence we employed in-

---

[16] Data end in 2004 since Rosstat makes critical for our analysis data (such as GRP) available with a 2-year lag.

[17] At the time of this work's publication, as a consequence of the ongoing administrative division reform, Russia had only 85 regions.

[18] Data for Chukotka's gross regional product are not reliable because Gazprom's taxes were paid in this region for some years; Ingush Republic (Chechen Republic's neighbor) had a vast inflow of refugees that distorted data.

14

dexes, published by *Granberg and Zaitseva (2003)* for 1999. Then we deflated PPP-adjusted GRP with implicit deflator.

The official population data were adjusted to be comparable with 2002 census results.

Spatial Econometrics Methods

Estimations involving regional data cause a number of problems studied in the framework of spatial econometrics. From the economic point of view in the context of a simple – unconditional or conditional – convergence model the possibility of spatial interactions is ignored as long as regions under consideration are implicitly assumed to be independent geographical units. Such factors as capital mobility, mobility of human capital and labor force, diffusion of knowledge and technology, transportation costs substantially influence regional interactions and, therefore, basic economic activities of regions and their growth rates. It's plausible that closely located regions are usually more integrated between each other than those separated by a large distance. So the basic prerequisite of our spatial analysis is assumption, according which regional economic data could be autocorrelated in space, i.e. economic activity is spatially clustered to a greater extent than under a random assignment.

From the econometric point of view presence of spatial correlation violates assumptions of Gauss-Markov theorem, namely independence and non-correlatedness of errors in model. The errors can be correlated with explanatory variables as well as spatially correlated with each other. In this case usage of ordinary least squares estimator can lead to biased, inefficient, or inconsistent estimates. Another problem is that estimates can be biased due to omitted variables if model doesn't take into account of spatial lags, which could be, in fact, significant.

From methodological point of view endogenous and exogenous variables of an econometric model of regional growth are subject to test for possible spatial correlation. Such a testing is carried out by means of special statistics of global spatial autocorrelation (Moran's *I*, Geary) and Moran scatterplot. Next, on the basis of formulated spatial hypotheses as well as spatial diagnostics tests of the errors estimated by ordinary least squares estimator, in order to take into account revealed spatial correla-

tion, one employs spatial lag (endogenous and/or exogenous) model or spatial error model (see *Fingleton et al. (2003)*; *Getis et al. (2004)*).

The key role in spatial analysis plays spatial weight matrix, which is designed in any way possible, e.g. using a measure of distance between regions (proximity), to specify for each region weights that reflect influence of all other regions on this particular region.

In spatial analysis as a measure of distance between regions we use 1) **the shortest distance** by motor roads between regional centers and 2) **minimal traveling time** required to cover the distance between regional centers by motor roads. To estimate the distances and traveling time we use special distance calculation software[19]. The basic assumptions on average speed by types of motor roads and traveling delays time used in computation of traveling time are presented in *Table 1.1*.

Table 1.1

Average Speed and Traveling Delays

Type of road	Speed (km/h)
Major highway	80
Highways	70
Major connecting	60
Connecting	50
Other roads	40
Winter road	30
Ferryboat	20
Settlement	**Traveling delays (min.)**
less than 10,000	5
from 10,000 to 50,000	10
from 50,000 to 100,000	15
from 100,000 to 500,000	30
from 500,000 to 1,000,000	60
1,000,000 and more	90
Loading/unloading time	120

Source: AutoTransInfo, http://www.autotransinfo.ru.

---

[19] Internet-system AutoTransInfo, http://www.autotransinfo.ru.

The concept of minimal traveling time is more preferable from the economic point of view since it represents the shortest "economic" distance, reflecting minimal time costs to cover a distance (not necessarily the shortest in geographic sense) between regional centers by motor roads. Traveling time is more preferable also from the point of view of results interpretation since it is controllable factor that can be changed by development of transportation infrastructure. Other things being equal, construction of modern road junctions and detour roads, improvement of pavement quality in big cities and settlements increase average speed and decrease traveling delays, thus result in smaller transportation cost and favor higher proximity between regions. Therefore, as a measure of distance for spatial econometric analysis minimal traveling time was preferred for computation of geographical weight matrix.

Application of Mundlak's Specification to Simultaneous Equations Model of Regional Growth

For estimation of the simultaneous equations model (SEM) we used panel data, which allowed to increase sample size significantly (77 regions over 8 years yielded 616 observations). Besides, panel econometrics methods are more flexible in analysis and interpretation. We had choice of several models: fixed effect (FE), random effect (RE), between estimator (BE), and traditional OLS pool estimator. The latter is the most simple one, but ignores time and space structure of panel data, and thus requires strong *ceteris paribus* assumption about main links of the model.

Mundlak's specification[20] in our case allowed to analyze relations both in space (BE, differences among variables) and over time (FE, regional changes over time). This specification requests creation of 2 new variables for each variable ($\bar{x}_{it}$): time demeaned values ($\bar{x}_i$) and deviations from the mean $(x_{it} - \bar{x}_i)$.

In other words, if a simple linear model can be presented as follows:

$$y_{it} = a_0 + a_1 \cdot x_{it} + \varepsilon_{it},$$

---

[20] *Mundlak (1978, 1981).*

then Mundlak's simple model can be presented as follows:

$$y_{it} = a_0 + \alpha_1 \cdot (x_{it} - \bar{x}_i) + \alpha_1^* \cdot \bar{x}_i + \mu_{it},$$

where

α_1 – FE-estimator,

α_1^* – BE-estimator,

μ_{it} – residual.

Analysis of Spatial Autocorrelation of Economic Development

Below we present the results of statistical tests for presence of global spatial autocorrelation in log average GRP per capita growth rates and log initial GRP per capita levels. The tests were carried out by means of Moran's I statistics, the most widely known measure of global spatial clustering (see *Appendix 1* for more), and two types of exogenous spatial weight matrices.

The first matrix is based on minimal traveling time between regional centers and characterizes existing regional infrastructure. The weights of the matrix were calculated as inverse squared minimal time to cover distance between regional centers by motor roads.

The second matrix besides traveling time contains information about GRP values of other regions and as a result is called market potential matrix. The spatial weights for particular region were calculated as inverse squared traveling time times GRP of each neighbor in 1998.

The values of Moran's I statistics and corresponding p-values for average growth rates and initial GRP per capita level for both matrixes are presented in *Tables 1.2* and *1.3*[21].

---

[21] The estimates of Moran's I coefficient for sample of 79 regions were statistically insignificant at 90% confidence level and are not presented therefore. In *tables 1.2* and *1.3* we present estimation results of Moran's I statistics on subsample of 77 regions, excluding Chukotka AD and Ingush Republic – obvious "outliers" by average GRP per capita growth rates. See below for more.

Table 1.2

Moran's *I* (Traveling Time Matrix)

Variables	*I*	*E(I)*	*sd(I)*	*z*	*p*-value*
Log. Average GRP per capita growth rates over 1998–2004	0.093	–0.013	0.053	2.019	0.043
Log. GRP per capita in 1998 (corrected on PPP)	0.103	–0.013	0.053	2.21	0.027

* – two-sided test.

Table 1.3

Moran's *I* (Market Potential Matrix)

Variables	*I*	*E(I)*	*sd(I)*	*z*	*p*-value*
Log. Average GRP per capita growth rates over 1998–2004	0.104	–0.013	0.053	2.225	0.026
Log. GRP per capita in 1998 (corrected on PPP)	0.089	–0.013	0.053	1.947	0.05

* – two-sided test.

According to the above results, the null hypothesis of no spatial auto-correlation is not accepted for both variables at 5% significance level. With some certainty one can say that levels of GRP per capita and their growth rates are positively spatially clustered. In other words regions with relatively high values of GRP per capita average growth rates are surrounded on average by relatively fast-growing neighbors, as well as relatively wealthy regions are surrounded by relatively wealthy (as of 1998).

The same conclusion can be drawn from the analysis of the Moran scatterplot. This plot displays standardized values of the spatial lag of the variable (ordinate axis) against standardized values of the variable itself (abscissa axis), as well as the regression line obtained by regressing the lag on its variable with the slope equal to the estimated value of Moran's *I* (see *Appendix 1* for more details).

Fig. 1.2 shows the Moran scatterplot (using traveling time spatial matrix) for GRP per capita average growth rates. The first and the third quadrants are characterized by positive spatial autocorrelation. The top-right (bottom-left) quadrant reflects spatial clustering of regions with relatively high (low) values of average growth rates surrounded by also relatively fast-growing (slow-growing) neighbors. It is important to stress

that observed spatial clustering of growth rates across regions is quite moderate due to a substantial number of regions located in top-left and bottom-right quadrants of the diagram corresponding to negative spatial correlation. These quadrants represent in the former case clusters of regions with relatively low values of average growth rates surrounded by relatively fast-growing neighbors, and, conversely, clusters of regions with relatively high values surrounded by relatively low values – in the latter. A substantial number of regions with negative spatial autocorrelation signifies that, to our mind, we are short to conclude about long-run tendencies in spatial clustering of regions in average growth rates. At the same time, even over the relatively short time interval under consideration one can talk about significant spatial heterogeneity of Russian regions economic development, which, apparently, should be taken into account in empirical studies of regional growth.
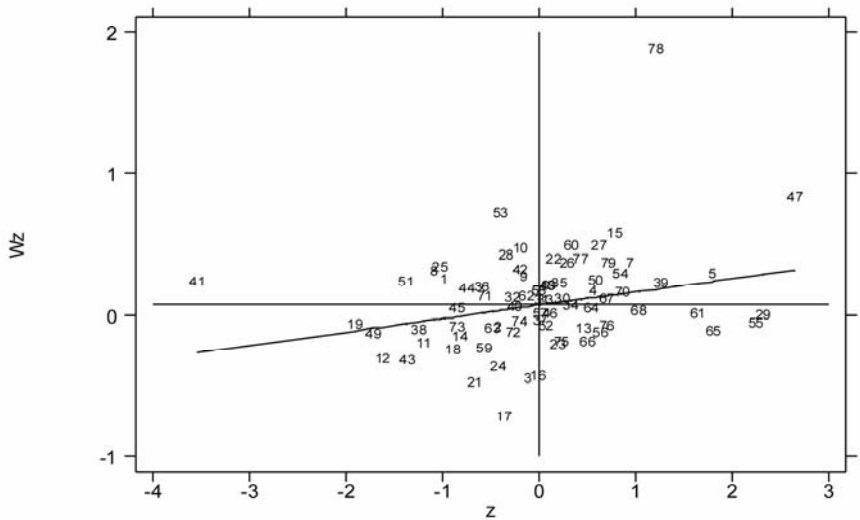
Fig. 1.2. The Moran Scatterplot for Log. Average Growth Rates of GRP per capita over 1998–2004 (Using Traveling Time Geographical Weights Matrix)

In recent papers on European regional growth (see, for example, *Fingleton et al. (2003); Baumont, Ertur, Le Gallo (2002)*) the Moran scatterplot is used for revealing so-called spatial convergence-clubs. Clustering of the clubs is carried out by initial per capita income. One chooses regions getting to the first and the third quadrants, next, the hypothesis of unconditional beta convergence is tested for each of the clubs: relatively poor surrounded by relatively poor as well as relatively wealthy surrounded by relatively wealthy regions.

Moran scatterplot (Moran's $I = 0.103$)
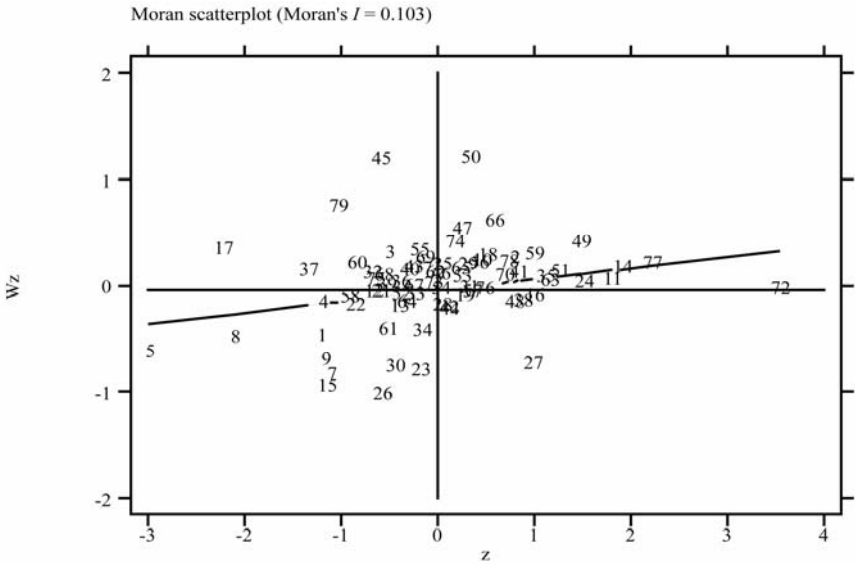
Fig. 1.3. The Moran Scatterplot for Log. GRP per capita in 1998
(Using Traveling Time Geographical Weight Matrix)

As it can be seen from the Moran scatterplot for per capita GRP in 1998 (see *Fig. 1.3*), pronounced "clubs" are hard to observe, and Russian regions can be hardly divided into groups by this figure. Moreover, substantial number of regions can be found in the second quadrant of the diagram corresponding to negative spatial correlation. Therefore we test the hypotheses of unconditional and minimal-conditional beta conver-

gence on the subsample of 77 regions (Ingush Republic and Chukotka AD excluded) without dividing regions into clusters by this figure.

In spite of substantial number of regions characterized by negative spatial correlation, the tests presented above show statistically significant spatial correlation on the subsample of 77 regions. This result shows that, on the whole, Russian regions are not homogeneously distributed in space as measured by employed data. To the contrary, regions are spatially clustered both in average growth rates of GRP per capita and GRP per capita levels in 1998. In other words, relatively fast-growing regions are surrounded on average by relatively fast-growing neighbors, as well as relatively wealthy regions are surrounded by relatively wealthy neighbors.

From the position of both new economic geography and analysis of convergence particular regions with relatively high values of initial GRP per capita and at the same time relatively high values of average growth rates are of interest. On the Moran scatterplots above such regions are positioned in the first or in the fourth quadrants on each of the diagrams. As it is seen from the Moran scatterplots and corresponding maps (see *Appendix 3*, *Maps 10* and *11*) these regions are mainly regions of the European part (St. Petersburg city (78), Leningrad Oblast (47), Moscow city (77), Moscow Oblast (50), Arkhangelsk Oblast (29), Vologda Oblast (35), Yaroslavl Oblast (76), Belgorod Oblast (31), Kurgan Oblast (46), Orel Oblast (57), Lipetsk Oblast (48)), as well as some regions of the Urals (Orenburg Oblast (56) and Sverdlovsk Oblast (66)), South of the Western Siberia (Novosibirsk Oblast (54) and Tomsk Oblast (70)) and South of the Far East (Khabarovsk Krai (27), Sakhalin Oblast (65)). Note, that mentioned regions, except for Orenburg Oblast and Sverdlovsk Oblast, form moderate in size groups of regions with common borders and at the same time such groups are characterized by both relatively high average growth rates and average GRP per capita values in 1998 (see also *Maps 2* and *3* in *Appendix 3*).

Note also, that computation of spatial correlation statistics is just a preliminary step in spatial econometric analysis. These statistics provide some evidence but do not explain the reasons for clustering of regions in space (see, for instance, *Anselin (1988)*). In order to test the hypotheses

about the reasons for such clustering, explained in general by new economic geography and theories of endogenous growth (*Fujita, Krugman, Venables (1999)*; *Fujita, Thisse (2002)*)) as well as the hypotheses about influence of spatial heterogeneity on economic development dynamics, one should use spatial econometric models.

Description of the Main Hypotheses

Below we present description of variables, chosen for estimation of deep factors of growth, as well as hypotheses of their relation to growth, migration, and investment. The variables were selected to reflect interregional differences in physical geography (average January temperature, degree of permafrost prevalence, availability of natural resources and their extraction), geo-economical specificity (access to sea trade, population of the main city, some variables characterizing infrastructure development), peculiarities of the earlier period of development (reverse migration), institutions (legislative risk, corruption indexes), and some other regional features.

Average temperature of January and permafrost. To take account of climate specificity of regions we employ 2 variables: average temperature of January and prevalence of permafrost. The latter is a dummy variable that has 3 gradations: continuous, partial, and rare. Distribution of permafrost and geography of average January temperatures are presented in *Appendix 3*, *Map 4*.

The above variables describe degree of climate severity and its appropriateness for human's habitation and economic activity. According to Dani Rodrik's logic, physical geography influences economic growth directly by channels productivity, accumulation of capital and labor (through migration flows and direction of investment), and indirectly – be means of trade and institutional development[22]. We suppose that severity of climate is one of the principal factors of the settlement pattern and directions of regional institutional development.

These variables are certainly exogenous, yet in some cases their interpretation can be complicated because of their significant correlation with

---

[22] For influence of geography on formation of distinct institutional systems see, e.g., *Acemoglu, Johnson, Robinson (2001); Engerman, Sokoloff (2002).*

other variables. For example, distance from north and harshness of climate is the main determinant of the cost of living and entrepreneurship in any region. Thus, regions of Russia with harsh climate were developed only recently, compared to regions with milder climate; therefore, in former regions infrastructure and transport communications are less developed (see *Appendix 3*, *Maps 6* and *7*).

The chosen climate variables can be correlated with structure of industrial output, owing to extractive specialization of the regions with harsh climate. In such regions share of extractive industries and raw materials intensive industries is higher than country average.

Correlation of age of settlement with January temperature permits to use the climate variable as a proxy for "institutional age" of regions. However, presence of such correlations does not interfere with general idea of our research, in which geography is a strongly exogenous factor influencing, in particular, institutional development.

Access to sea routes. In recent research[23] negative influence of landlocked location on economic development of countries was underlined. Absence of a sea shore or of a navigable river makes national trade dependent on political relations with bordering countries and quality of their transport infrastructure. Costs of other means of transportation are normally higher as well. Thus, countries that possess access to a sea, on the contrary, have comparative advantage of preferable geographic location.

Presence of sea ports in regions can be viewed as additional possibility for accelerated development. On the one hand, a sea port is an important communication link between national and foreign markets, while regions with ports have an opportunity for growth of service and trade sector. Acting as intermediaries, such regions undoubtedly benefit from growth of foreign trade and integration of Russian economy with world market.

On the other hand, regions that have sea ports on their territory have capability for integration with global market and participation in world trade on their own[24]. As compared to land-locked regions trade flows of

---

[23] *See, e.g., Mellinger A., Sachs J, Gallup J. (1999); Gallup J., Gaviria A., Lora E. (2003).*

[24] Process of regionalization on the global level (direct "embedding" of sub-national actors into international market) is collinear to the contemporary trends of world economy

such regions are normally more diversified, besides risk of territorial changes of the neighboring countries is almost negligible, while break of bonds with traditional continental neighbors usually poses a serious threat to trade.

There are 9 regions with unfreezing ports in Russia: Krasnodarskiy Kray, Primorskiy Kray, Khabarovskiy Kray, Rostovskaya Oblast, Leningradskaya Oblast, Murmanskaya Oblast, Kaliningradskaya Oblast, Arkhangelskaya Oblast, Sakhalinskaya Oblast, and city of St.Petersburg. These regions belong to those settled a while ago, i.e. densely populated, with developed communications. Navigation is possible 12 months a year, though for some regions ice-breaker fleet is necessary for the coldest periods.

Consequently, regions that have access to sea routes ceteris paribus will grow faster than others as a result of higher productivity, or more attractive investment climate. Besides being a reflection of geographic location of region, a sea port is a result of human activities and can be referred to as to an economic variable. We threat sea port variable as an exogenous, what is suitable for the relatively short period under consideration, since obviously port and corresponding infrastructure is a result of long term development.

Agglomerations. As aforementioned, agglomeration influences growth, since it is related to spatial location of industries[25]. Positive scale effect is an innate advantage of larger cities for attracting migrant inflows and investment and it facilitates accelerated economic growth. Some research[26] demonstrates, that general trend of recent Russian development is characterized by comparative success of larger cities and of their bordering areas contrasted by degradation of rural areas not involved in agglomeration networks.

Since Russian statistical agency does not provide data on metropolitan areas, which is the best way to measure agglomeration effect, hence we

---

development and in most cases promoted regional development in many countries. See *Scott A. J. (1998)*.

[25] *Fujita M., Krugman P., Veanbles A. (2002)*.

[26] See, e.g., *Nefedova (2004)*.

use population of the largest city of the region. We hypothesize that larger agglomerations have higher rate of growth.

Though it was already mentioned, we'd like to reiterate that under the centrally planned economic system of the USSR administrative centers of all Russian regions were at the same time the largest cities, except for 2 regions: Kemerovskaya Oblast (largest city – Novokuznetsk) and Vologodskaya Oblast (largest city – Cherepovets).

Settlement patterns are quite rigid; therefore we treat agglomeration, which reflects settlement patterns and spatial location of industries inherited by Russia from the USSR, as an exogenous variable.

Reverse migration. During the Soviet times huge areas with harsh climate were developed by exploitation of prisoners, deportation of people from other regions of the USSR, and by attraction of workers with economic stimuli, such as wage rise for arduous working conditions (one of rare exceptions from the traditional homogenous payment principle of the Soviet system). The mechanism of this process is described in detail in work of Fiona Hill and Clifford Gaddy[27], who also provided statistical evidence of significant decrease of average temperature for areas inhabited in the USSR in the XX century.

Afterward collapse of the plan system, weakening of the migration barriers and real wage cuts facilitated a reverse process – migration of population from recently developed regions with harsh climate back to the "domestic" or regions of early development.

We hypothesize that regions that experienced substantial migration inflows during the Soviet period would experience most significant migration outflows during the period under consideration.

In our case reverse migration can hardly be referred to any of Rodrik's groups of determinants. This variable describes transition from plan economic system to market, where different economic mechanisms and incentives are put in place. However, process of reverse migration could take place under market economy as well, for instance, in case of exhaustion of natural resources in such harsh climate regions, invention of less labor-intensive technology in extractive industries, decrease of demand

---

[27] *Fiona Hill, Clifford Gaddy (2003).*

on extracted resources etc. Thus, effects captured by this variable can be described as structural changes caused by exogenous changes.

In another sense reverse migration variable is a proxy for partial settlement of regions and can indirectly describe "institutional age" of regions (see below).

Extractive specialization. Extraction of resources is a regional specialization caused by physical geography – obviously, resource endowment is an absolutely immobile factor. Extraction of natural resources is a regional comparative advantage in productivity per unit of labor and capital costs, since it creates additional income owing to natural rent.

We use two variables to take account of extractive specialization. First, share of extractive industries output in total industrial output and, second, output of extractive industries per capita (see *Appendix 3*, *Map 5*).

This variables are from the geographic group, but it's difficult do disentangle physical and economic geography, as in the case of sea port variable. Regions with developed extractive industry normally both have natural resources (physical geography) and extract them (economic geography).

In international comparisons share of mineral resources in export often claimed to be negatively related with average annual growth rates over long periods of time[28], reason for which lies in specifics of budgetary finance of resource abundant countries, which in turn stems from reliance on generally unstable prices of natural resources, Dutch disease, and erosion of democratic institutions aggravated by resource nationalism policy[29].

In our study extractive industries are likely to be positively related to growth, since extractive industries are more competitive than majority of other Russian industries and the period under consideration is a period of high prices on most goods of Russian resource export.

Human capital. We hypothesize that regions possessing more educated and qualified labor have better opportunities for growth. First of all,

---

[28] *Sachs J., Warner A. (1997).*
[29] *Karl T. L. (2002).*

this supposition is applicable to regions with developed manufacturing and service sector.

To test the hypothesis we use share of population with higher education and number of postgraduate students per 10,000 inhabitants. Though the latter variable has certain disadvantages (for instance, substantial share of male youth in Russia uses institute of postgraduate education to avoid military draft), it is appropriate to differentiate regions by their R&D potential.

Telecommunications and transport infrastructure. We employ several Telecommunications and transport infrastructure variables to take account of development level and quality of regional infrastructure. In long run infrastructure is an endogenous factor of growth and develops in response to increasing demand. In our case, because of short period of research, it's unlikely for infrastructure to be an endogenous factor of growth. Nevertheless, to decrease acuteness of endogeneity problem we use initial values of variables: transportation of passengers by railway per capita (see *Appendix 3*, *Map 6*) and relative quantity of stationary phones (see *Appendix 3*, *Map 7*).

Economic institutions. Institutions, accepted and maintained by society rules and traditions, define degree of risk and uncertainty, and, thus, transactional costs of economic activity, consequently, influencing decision-making process of economic agents[30].

Although they are an important factor of growth, institutions are difficult to be quantitatively assessed. In empiric papers indirect measurements are usually used, such as estimates based on surveys, monitoring, and/or indexes of rating agencies or independent institutions. We employed results of small and medium enterprises survey administered by "OPORA-VTsIOM", investment rating and its components (Expert RA), rating of regional officials credibility (Standard and Poor's), index of residents credence towards regional authorities (INDEM foundation).

As up to date of the research, only 15 regions were assigned rating by Standard and Poor's, including a few withdrawn ratings. Availability of the rating was used as a dummy variable (1 – credit rating was obtained

---

[30] See *North D. (1990).*

by a region, 0 – credit rating was not obtained) reflecting overall quality of regional administration and economic framework: first, authorities of ranked regions were at least capable of providing necessary and satisfactory information for the agency, hence availability of the rating is a sign of governance publicity; second, authorities of such regions are likely to draw long-term credits for regional development[31].

A drawback of such data is their timing, for example, survey of OPORA-VTsIOM was conducted in 2005, i.e. in year that is not covered by our statistical data. Besides, institutional variables are dependent on growth and income, so endogeneity problem arises. This issue could be avoided by using lagged institutional variables. Since this approach is infeasible, below we use instrumental variables method to get rid of endogeneity.

Some of the above-mentioned variables can be treated as proxies for regional institutions. For instance, Mancur Olson in his research of distributive coalitions and economic growth employed age of USA states, measured by their settlement timing, as an indicator of distributive coalitions burden on the state economy[32]. A.N.Pilyasov suggests notion of institutional age for Russian regions[33]: inequality of regional development and different timing of settlement is a reliable way to differentiate regions and at the same time to escape endogeneity problem.

Age of settlement and corresponding type of institutions of Russian regions are closely related to climatic conditions. Settlement patterns were directed mainly from more to less favorable for habitation areas, as aforementioned.

Analysis of Convergence across Russian Regions with Respect to Spatial Dependence

Presented in this section results of the analysis of per capita income convergence in Russian regions continue research started in the previous

---

[31] Financing of any large-scale long-term project, especially of those involving foreign borrowing, requests a rating issued by an international credit agency.

[32] According to Olson's theory, number and, thus, burden of such coalitions normally grows over time *(Olson M. (1983))*.

[33] *Pilyasov A.N. (2003).*

IET-CEPRA study of regional growth[34]. This study differs in its empirical testing of the conditional convergence among Russian regions, which is conducted from the perspective of new economic geography using spatial econometric methods. Moreover, unlike in the previous paper, convergence hypotheses are tested for regional income levels corrected to take account for interregional differences in price levels.

Concepts and Models of Convergence[35]

For convergence of income levels analysis one normally uses two related but not identical ideas of convergence: β convergence and σ convergence. The former assumes accelerated development of poor regions resulting in gradual smoothing of regional differences, thus, β convergence model assumes existence of long-run tendency towards smoothing of economic development levels. The latter, σ convergence, assumes decrease of regional disparities in GRP per capita levels or other income indicators.

A starting point in convergence analysis – so-called **unconditional** β convergence model based on the neoclassical growth theory (*Solow (1956); Swan (1956)*). According to this model, rates of economic growth are positively correlated with the starting gap between the initial per capita income of a region and the steady-state per capita income level, which is the same for all regions. Regions on the steady growth trajectory are characterized by constant growth rates of per capita income. According to the model, poor regions should grow at higher pace than wealthy regions, so that the long-run perspective should tend to smooth regional differences in economic development.

Unconditional convergence model can formally be represented as:

$$g_T = \alpha + \beta y_0 + \varepsilon, \qquad \varepsilon \sim N(0, \sigma^2 I) \tag{1.1}$$

where

---

[34] See *Drobyshevsky S., Lugovoy O., Astafyeva E. et al. (2005)*.
[35] A detailed survey of convergence models can be found in *Drobyshevsky S., Lugovoy O., Astafyeva E. et al. (2005)*.

g_T – log average growth rates over the period of length T,

y_0 – log initial per capita income,

α – parameter that takes account for rate of technological progress and steady-state level of per capita income,

β – convergence coefficient,

ε – random component.

Convergence process is usually characterized by **convergence speed** and **half-life** time period, which can be computed using the estimate of the convergence coefficient β as $\hat{b} = -\ln(1 - T\hat{\beta})/T$ and $hl = \ln(2)/\hat{b}$, respectively.

In the unconditional convergence model, in accordance with neoclassical growth theory, hypothesis of negative correlation between average growth rates and initial per capita income level is tested[36]. At the same time, theory assumes that regions tend to the same steady-state growth trajectory.

Note that this model makes assumption of homogeneous regional structure of economy while differences in GRP per capita levels are temporary and caused by initial level of income only. In presence of substantial economic, institutional, and geographic differences among regions this assumption is too strong and hardly realistic for Russian regions. Therefore, it would be plausible to assume, that different regions have distinct steady-state growth trajectories and, therefore, distinct long-run rates of growth. In that case elimination of differences of economic development levels may not occur. In that case state regional policy objective should be the increase of the steady-state growth levels of underdeveloped regions.

The assumption of distinct regional steady-state growth trajectories is formalized in the framework of **conditional** β convergence model as follows:

---

[36] See, e.g., *Barro, R., X. Sala-i-Martin.* Economic Growth. (2nd edition) The MIT Press, 2003.

$$g_T = \alpha + \beta y_0 + Z\phi + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I), \qquad\qquad (1.2)$$

where Z – matrix of regional growth factors controlling for differences in steady-state levels of each region.

Thus, under conditional convergence model framework the hypothesis of negative correlation between average growth rates and initial per capita income is tested while additional factors controlling for regional differences in steady-state levels are taken into account.

In spatial analysis additionally model of β convergence with minimal conditions or the **minimal conditional convergence** model is considered, which allows for regions on distinct steady-state trajectories, since in this framework dynamics of economic development of a particular region may be conditioned on the dynamics and/or the level of economic development of its neighbors. The minimal conditional convergence model can be represented as follows:

$$g_T = f(y_0, W, Wg_T, Wy_0, \varepsilon) \qquad \varepsilon \sim N(0, \sigma^2 I), \qquad\qquad (1.3)$$

where

W – geographical weight matrix,

Wg_T – endogenous spatial lag of log average growth rates,

Wy_0 – exogenous spatial lag of log initial per capita GRP.

For these models in addition to the conditional convergence hypothesis two basic spatial hypotheses are tested:

- Spatial clustering by average growth rates through the endogenous spatial lag on average growth rates of GRP per capita,
- Spatial clustering by average growth rates through the exogenous spatial lag on the initial values of GRP per capita.

These spatial hypotheses rest on assumption of economic interactions, which strength decreases with increase of distance between regions under consideration. At the same time interactions between regions can be both direct (trade of goods and services, population and labor force migration) and more underlying (diffusion of knowledge and information, spreading of innovations, institutional and social interactions).

The spatial conditional β convergence model in general case can be stated formally as follows:

$$g_T = f(y_0, Z, W, Wg_T, Wy_0, WZ, \varepsilon) \qquad \varepsilon \sim N(0, \sigma^2 I), \quad (1.4)$$

where WZ – matrix of exogenous spatial lags on controlling growth factors.

Thus, average growth rates depend not only on the initial level of GRP per capita, but also on the dynamics and/or the level of economic performance of neighboring regions as well as additional growth controlling factors and spatial lags on these factors.

In present paper to take account of the regional steady-state levels in conditional convergence models we consider indicators of population mobility (migration), infrastructure (passengers departures by railways), dependence of region on the Federal Government (financial aid), geographical location (availability of a see port), raw-materials specialization (share of fuel industry in industrial output), human capital stock (postgraduate students per 10,000 inhabitants, share of population with higher and postgraduate education) as well as indicators of spatial interaction between the regions.

Results of the Empirical Analysis of Convergence

Below we present the results of the analysis of σ and β convergence of per capita income (income is measured as gross regional product, GRP, per capita) for Russian regions. As it has been mentioned, the analysis is carried out over 79 and in some cases over 77 Russian regions. As a measure of distance for spatial analysis we use data on minimal traveling time required for the freight motor vehicle to cover the distance between two regional centers (given assumptions on the average speed by types of motor roads, see above).

Sigma Convergence (1996–2004)

Tests of the hypothesis of σ Convergence of per capita GRP over the period 1996–2004 were carried out on the basis of estimation of four

measures of inequality and dispersion[37]: coefficient of variation (which is the ratio of the standard deviation over the mean), Gini coefficient (measuring the level of inequality), inter-quartile range (IQR, the difference between the upper and the lower quartiles of the log per capita GRP) and range (the difference between maximal and minimal values of the log per capita GRP). The values of these indicators for each year are presented in *Table 1.4*. The Table shows that all these indicators move in sync and provide no evidence of decreasing disparity of per capita regional income.

Table 1.4

Measures of Inequality and Dispersion for GRP per capita, over 79 Regions

Year	Coefficient of variance	Gini coefficient	IQR	Range
1996	0.4869	0.2305	0.5255	2.6110
1997	0.4977	0.2368	0.5044	2.6247
1998	0.5036	0.2361	0.4825	2.7027
1999	0.4928	0.2356	0.4604	2.7269
2000	0.4910	0.2330	0.4710	2.6906
2001	0.4923	0.2328	0.4690	2.6302
2002	0.5028	0.2396	0.4918	2.8949
2003	0.5153	0.2442	0.5279	2.9314
2004	0.5038	0.2417	0.5475	2.9451

The dynamics of the coefficient of variation is shown on *Fig. 1.4*. This figure illustrates the increasing disparity of GRP per capita over the whole period under consideration, though reduction of the coefficient of variation took place in 1999, 2000, and 2004.

---

[37] We use the same measures of inequality as in *Fingleton B. (2003)*.
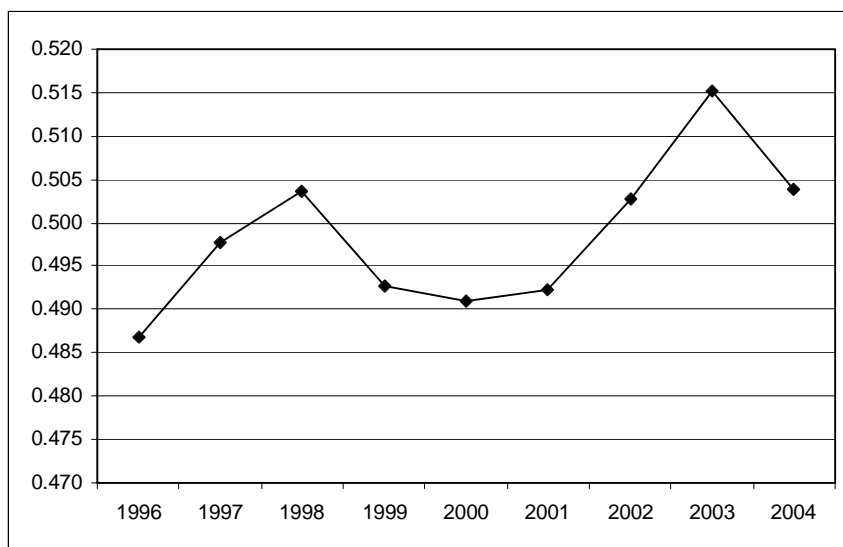
Fig. 1.4. The Coefficient of Variation for GRP per capita,
Corrected Data, over 79 Regions

However, the magnitude of the dispersion is quite moderate. In order to test significance of the coefficient of variation change, we further perform the Likelihood Ratio Test of the Equality of the Coefficients of Variation of *k* Normally Distributed Populations for GRP per capita (for the 10 percent significance level), see the table below[38].

---

[38] The test is based on *Verrill, Johnson (2005)* and can be found at: http://www1.fpl.fs.fed.us/covtestk.html. Note that income (GRP per capita) is log-normally distributed and, therefore, taking logs of GRP per capita we can obtain normal distribution. This test is based on Likelihood Ratio and is robust to log-normal distributions if the number of distributions tested is less or equal 50. In our case the condition is satisfied, since we test the hypothesis only for 2 particular years under consideration. We ran the test for both for GRP per capita and its logarithm, with similar results, and so we present the results of the test only for GRP per capita.

Table 1.5

**The Likelihood Ratio Test of the Equality
of CV of GRP per capita, 79 Regions**

Year-to-year	*p*-value*
1996–1997	0.8732
1997–1998	0.9528
1998–1999	0.7897
1999–2000	0.9267
2000–2001	0.9753
2001–2002	0.8796
2002–2003	0.8543
2003–2004	0.7840
Subperiods	
1996–1998	0.7825
1998–2000	0.7626
2000–2003	0.6750
1996–2003	0.6143

* The *p*-values are from the asymptotic likelihood ratio test.

High *p*-values mean that we cannot reject the hypothesis of the equality of coefficients of variation across time with the 10 percent significance level. Hence the annual changes of the coefficients of variation of GRP per capita for different years appear to be statistically insignificant. This result holds for the whole period under consideration (1996–2004) as well as for the following subperiods: 1996–1998, 1998–2000, and 2000–2003. In other words, the above trends in the coefficient of variation (see *Fig. 1.4*) – increasing divergence, convergence, and divergence again – are statistically insignificant.

If we exclude Chukotka AD and Ingush Republic from the sample, which are obvious outliers by GRP per capita average growth rates over 1998–2004, see below, the pattern slightly changes. Coefficient of varia-

tion and range decreased over the whole period of 1996–2004 (*Table 1.6*), while Gini coefficient increased, though changed slightly over the whole period (only in third character). The rise of the difference between the upper and the lower quartiles of the log per capita GRP (IQR) shows that the inequality of per capita income measured using the interquartile range is not sensitive to excluded regions mentioned above.

Table 1.6

Measures of Inequality and Dispersion for GRP per capita, over 77 Regions (Chukotka AD and Ingush Republic Excluded)

Year	Coefficient of variance	Gini coefficient	IQR	Range
.1996	0.4859	0.2276	0.4764	2.6110
1997	0.4977	0.2348	0.4734	2.6247
1998	0.5020	0.2336	0.4432	2.7027
1999	0.4890	0.2322	0.4465	2.7269
2000	0.4862	0.2288	0.4503	2.6906
2001	0.4877	0.2273	0.4614	2.6302
2002	0.4879	0.2277	0.4712	2.5849
2003	0.4891	0.2283	0.4822	2.5371
2004	0.4855	0.2285	0.5140	2.4558

Fig. 1.5 presents the dynamics of the coefficient of variation for the reduced sample: dispersion of per capita income increased until 1998, while signs of sigma convergence appeared over the post-crisis period 1999–2004. Note also a small increase of disparity over 2001–2003.
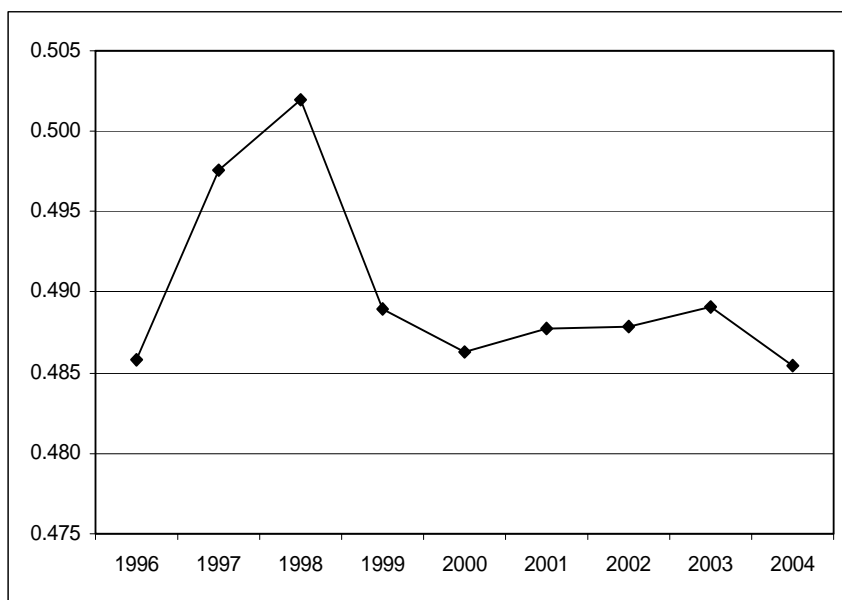
Fig. 1.5. The Coefficient of Variation of GRP per capita,
Corrected Data, over 77 Regions (Chukotka AD
and Ingush Republic Excluded)

Nevertheless, changes in the coefficient of variation for the reduced sample are statistically insignificant (the same as for the full sample of 79 regions) relying on the p-values of the likelihood ratio test of the equality of the coefficients of variation (see *Table 1.7*).

Therefore, we conclude that the changes of regional income disparity measured as the coefficient of variation of per capita GPR are statistically insignificant, so we can't reject either hypothesis of sigma convergence or hypothesis of sigma divergence.

Table 1.7

**The Likelihood Ratio Test of the Equality
of CV of GRP per capita, 77 Regions**

Year-to-year	*p*-values*
1996–1997	0.8597
1997–1998	0.9784
1998–1999	0.7564
1999–2000	0.9114
2000–2001	0.9779
2001–2002	0.9568
2002–2003	0.9728
2003–2004	0.8974
Subperiods	
1996–1998	0.7936
1998–2000	0.7162
2000–2003	0.9990
1996–2003	0.9947

* The *p*-values are from the asymptotic likelihood ratio test.

Beta Convergence (1998–2004)

In this section we test the hypothesis of beta convergence among Russian regions. The analysis is based on three types of beta convergence models: the unconditional, with minimal conditions or the minimal conditional, and the conditional beta convergence.

Estimation of the unconditional beta convergence model (Eq. 1.1) on the full sample of 79 regions does not provide statistically significant estimate of the convergence coefficient[39]. We can't reject either hypothesis of regional convergence or hypothesis of regional divergence.

---

[39] The estimates over 79 regions are not presented here due to the statistical insignificance. Below we present results for 77 regions.

The following diagram (*Fig. 1.6*) presents the scatterplot of the log GRP per capita average growth rates over 1998–2004 versus the log per capita GRP in 1998. On this diagram one can easily notice two "outliers" – Chukotka AD (87) and Ingush Republic (6). Chukotka AD has had extremely high average growth rates (seemingly, related to the re-registration of oil-traders and large investments in AD by Sibneft). Ingush Republic, on the contrary, is the region with the most negative average growth rates of GRP per capita (about – 4%, which is not surprising due to political tension and economic instability in adjacent Chechen Republic). This clearly signifies existence of other factors that influence economic performance of these two regions. The only way to take such factors into account in the regression analysis is to use dummy variables, what is equivalent to excluding them from the sample (except for change in degrees of freedom); thus, we carry out further analysis over the sample of 77 regions, excluding Chukotka AD and Ingush Republic.
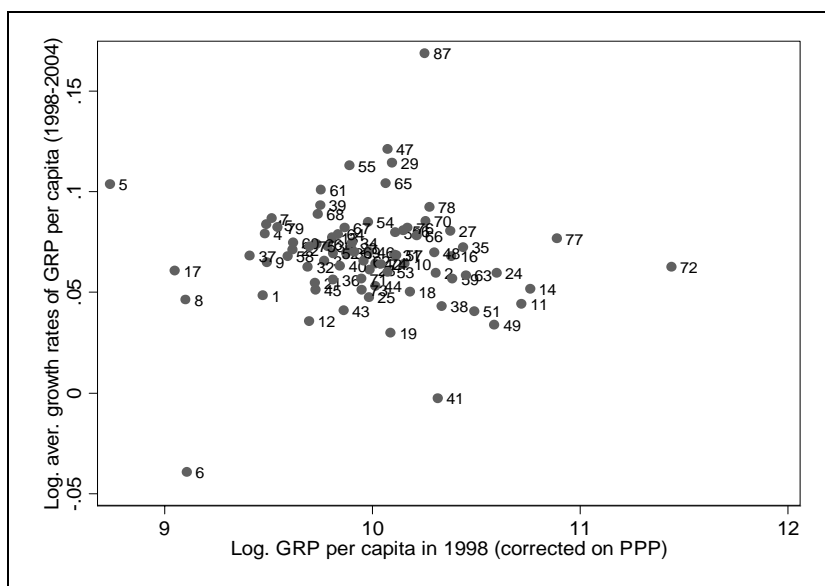
Fig. 1.6. The Log Average Growth Rates of GRP per capita over 1998–2004 versus the Log Initial GRP per capita 1998, 79 Regions

The Unconditional Beta Convergence Model

Table 1.8 below presents the estimates of the unconditional beta convergence model (Eq. 1.1) by means of ordinary least squares estimator on the sample of 77 regions (excluded Chukotka AD and Ingush Republic).

The Unconditional Beta Convergence Model, OLS

Log average growth rates of GRP per capita over 1998–2004	Coef.	Std. Err.	t	$P>|t|$
Log GRP per capita (corrected on PPP) in 1998	–0.0098	0.0054	–1.8	0.075
Intercept	0.1660	0.0543	3.06	0.003
Convergence speed, %	1.02		**LIK**	194.16
Half-life, years	68		**AIC**	–4.991
Number of observation	77		**BIC**	–714.10
F(1, 75)	3.25		**White**	0.32
Prob > F	0.0754			(0.854)
R-squared	0.0416			
Adj R-squared	0.0288			
Root MSE	0.0197			

As can be seen from the table, the estimated convergence coefficient is negative and statistically significant at 90% confidence level; however it is insignificant at 95% level. The estimated convergence speed (if it exists) is quite low, 1% per year, which corresponds to half-life time period (time needed to clear half of the distance to the steady-state GRP per capita level) equal to 68 years. Note that 90% confidence level is fairly low, so that we can make a mistake in 10% cases. At a higher confidence level (95% or 99%) the null hypothesis of absence of convergence is not rejected.

Note also that the unconditional model, in fact, can be misspecified due to spatial autocorrelation of the errors. The results of spatial diagnostics tests of the residuals estimated by means of OLS estimator are pre-

sented in *Table 1.9*. The Moran's *I* statistic computed with two weight matrices is significant at 5% significance level.

Spatial Diagnostic Tests of OLS-Residuals in the Unconditional Beta Convergence Model

Weight matrix	Traveling time		Market potential	
Diagnostics				
Test	**statistic**	**p-value**	**statistic**	**p-value**
Spatial error:				
Moran's *I*	2.041	0.041	2.204	0.028
Lagrange multiplier	2.711	0.1	2.799	0.094
Robust Lagrange multiplier	0.036	0.849	0.016	0.899
Spatial lag:				
Lagrange multiplier	2.68	0.102	2.975	0.085
Robust Lagrange multiplier	0.005	0.943	0.192	0.661

At the same time spatial diagnostic tests reveal more significant spatial correlation of the residuals with the market potential weight matrix as well as provide for this matrix some more evidence in favor of the spatial lag model as compared to the spatial error[40].

The Minimal Conditional Beta Convergence Model

Consider now the minimal conditional convergence model (Eq. 1.3) in the spatial lag specification:

$$g_T = \alpha + \beta y_0 + \rho W g_T + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I), \qquad (1.6)$$

in which one can tackle a problem of spatial autocorrelation of the errors considering endogenous spatial lag on the log GRP per capita average

---

[40] See *Appendix 1* about spatial econometric models. The procedure of searching the final specification is described, for instance, in *Florax R. et al. (2003)*; see also *Anselin L. (2003)*.

growth rates, Wg_T, as an explanatory variable. The lag was computed using the market potential weight matrix.

The model is estimated by means of the maximum likelihood estimator since the OLS estimator would produce inconsistent estimates due to the presence of the stochastic regressor (see *Anselin (1988)*). *Table 10* presents the estimation results of the spatial lag model (Eq. 1.6).

Table 10

**The Minimal Conditional Beta Convergence Model,
the Spatial lag Model, Maximum Likelihood Estimator**

Log average growth rates of GRP per capita over 1998–2004	Coef.	Std. Err.	z	P>|z|
Log GRP per capita (corrected on PPP) in 1998	–0.0094	0.0052	–1.82	0.069
Intercept	0.1289	0.0544	2.37	0.018
Spatial lag: log average growth rates of GRP per capita	**0.4613**	**0.2151**	**2.15**	**0.032**
Convergence speed, %	**0.97**			
Half-life, years	**71**			
Number of observation	**77**			
Variance ratio	**0.066**			
Squared corr.	**0.117**			
Sigma	**0.02**			
Log likelihood	**196.134**			

As may be seen from the table, addition of the endogenous spatial lag somewhat improves significance of the convergence coefficient; however, it is still within 10% significance level. Thus, the minimal conditional beta convergence hypothesis (i.e. that regional steady-state levels differ only due to the spatial clustering in GRP per capita growth rates) is not rejected at 5% significance level.

Another important result of estimation of the minimal conditional convergence model is the evidence of the *spatial spill-over* effect: the estimated spatial lag coefficient of 0.46 means that in the model considered above economic growth of a given region is significantly related to economic growth of the other neighboring regions, and that the closer they are located and the larger in economic sense they are, the greater is their influence on the given region (in accordance with the market potential weight matrix).

According to the econometric results, the average growth rates of a given region are positively correlated at 5% significance level with the average growth rates of neighboring regions by means of the endogenous spatial lag (Rho coefficient).

Note that the log initial GRP per capita and the endogenous spatial lag considered in the model explain up to 11.7 % of variation of the log GRP per capita average growth rates. Low explained variation value indicates that probably some other factors exist, which are not considered in the model and which affect regional steady-state levels. Further, we consider various specifications of the conditional convergence model.

The Conditional Convergence Model

Below we present the results of the conditional convergence analysis for Russian regions, which is a modification of the regional convergence in EU analysis carried out by Fingleton in 2004[41]. First, we estimate the conditional convergence model of the following form:

$$g_T = \alpha + \beta y_0 + \phi_1 fapc98 + \phi_2 sh\_fuel98 + \phi_3 t2 + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I), (1.7)$$

where

*fapc*98 – financial aid of the Federal Government to the regional budgets[42] in 1998 calculated on per capita basis;

*sh_fuel*98 – share of fuel industry in industrial output in 1998;

*t*2 – dummy-variable for depressed regions.

---

[41] *Fingleton B. (2004).*

[42] The main objective of these funds was to provide territorial justice in public spending. We calculated Federal Government assistance to the regions as the sum of five elements: The Fund of the Financial Support of the Regions (FFPR), mutual settlements, subsidies, subventions, and budgetary loans.

These variables were chosen to take account of wealthy natural resource abundant regions, the most indigent regions-recipients of financial aid from the Federal Government, and a group of regions which we refer to as "depressed regions". This group consists of 13 regions that experienced greater than 40% decline in industrial output in 2004 as compared to 1990 and in whose share of industry in GRP was greater than 20% in 2003[43].

Tables 1.11 and *1.12* represent model estimation results by means of OLS estimator and spatial diagnostic tests of the residuals, respectively.

<div align="right">

Table 1.11
</div>

<div align="center">

The Conditional Beta Convergence Model, OLS
</div>

Log average growth rates of GRP per capita over 1998–2004	Coef.	Std. Err.	*t*	*P>|t|*
Log GRP per capita (corrected on PPP) in 1998	–0.0235	0.0056	–4.22	0.000
Financial aid to the regional budgets per capita in 1998	–0.0199	0.0043	–4.66	0.000
Share of fuel industry in industrial output in 1998	0.0002	0.0001	1.59	0.116
Dummy-variable on depressed regions	–0.0133	0.0056	–2.36	0.021
Intercept	0.3134	0.0563	5.57	0.000
Convergence speed, %	**2.57**		LIK	206.901
Half-life, years	**27**		AIC	–5.244
Number of observation	77		BIC	–726.556
F(4, 72)	8.15		White	8.85
Prob > F	0.0000			(0.7842)
R-squared	0.3116			
Adj *R*-squared	0.2734			
Root MSE	0.0170			

---

[43] This group comprised of the following regions: Briansk Oblast, Voronezh Oblast, Ivanovo Oblast, Pskov Oblast, Dagestan Republic, Karachaevo-Cherkess Republic, Chuvash Republic, Kirov Oblast, Penza Oblast, Kurgan Oblast, Altai Krai, Primorskii Krai, Amur Oblast.

Table 1.12

Spatial Diagnostic Tests of OLS-Residuals in the Conditional Beta Convergence Model

Weight matrix	Traveling time		Market potential	
Diagnostics				
Test	**statistic**	***p*-value**	**statistic**	***p*-value**
Spatial Error:				
Moran's *I*	3.804	0.000	3.116	0.002
Lagrange multiplier	10.457	0.001	5.54	0.019
Robust Lagrange multiplier	6.741	0.009	2.148	0.143
Spatial Lag:				
Lagrange multiplier	6.32	0.012	3.805	0.051
Robust Lagrange multiplier	2.604	0.107	0.412	0.521

Note that the convergence coefficient is negative and statistically significant at 1% level. The variables characterizing low-income (aid recipients variable) and depressed regions are significant as well. Negative coefficients at these variables provide evidence of a lag in growth of these regions compared to the others. At the same time the variable characterizing regions' resource orientation is statistically insignificant.

Spatial diagnostic tests of the model (see *Table 1.12*) point to the existence of strong spatial correlation in estimated regression residuals for both spatial matrices (with more significant Moran's *I* value for traveling time matrix). Moreover, the spatial lag model is preferred to the spatial error model.

Conditional convergence model (Eq. 1.7) in spatial error specification can be represented as follows:

$$g_T = \alpha + \beta y_0 + \phi_1 fapc98 + \phi_2 sh\_fuel98 + \phi_3 t2 + u$$
$$u = \lambda Wu + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I)$$
. (1.8)

From the position of convergence analysis this model assumes that the log average growth rates of a region are explained by the log initial GRP

per capita and the set of region's exogenous factors, however, stochastic errors follow first order spatial autoregressive process. Thus, in the framework of this model one tests the hypothesis that the log average growth rates are influenced not only by random shocks in this region, but also by random shocks in all other regions.

The spatial error model is also estimated by maximum likelihood method since OLS yields in this case inefficient parameter estimates. The estimation results (with traveling time matrix) are presented in *Table 1.13*.

<div align="right">*Table 1.13*</div>

The Conditional Beta Convergence Model in Spatial Error Specification (Traveling Time Matrix), Maximum Likelihood Estimator

Log average growth rates of GRP per capita over 1998–2004	Coef.	Std. Err.	z	P>|z|
Log GRP per capita (corrected on PPP) in 1998	—0.0262	0.0053	–4.95	0.000
Financial aid to the regional budgets per capita in 1998	–0.0217	0.0041	–5.31	0.000
Share of fuel industry in industrial output in 1998	0.0003	0.0001	2.37	0.018
Dummy-variable on depressed regions	–0.0135	0.0048	–2.84	0.004
Intercept	0.3401	0.0536	6.35	0.000
Spatial error autoregressive coefficient	0.5699	0.1650	3.45	0.001
Convergence speed, %	**2.89**			
Half-life, years	**24**			
Number of observation	77			
Variance ratio	0.382			
Squared corr.	0.31			
Sigma	0.02			
Log likelihood	211.4355			

The estimation results of the model considering spatial interactions are analogous to the previous one; however statistical significance of the coefficients is notably higher, what can be related to more efficient estimates of the maximum likelihood estimator. Note that OLS-estimates of the model that ignores spatial autocorrelation of the residuals (see *Table 1.10*) underestimate the convergence speed. In the model

estimated by means of the maximum likelihood estimator the half-life time period is 24 years versus 27 based on OLS-estimates. That means that the average statistical region needs 24 (27) years to cover a half of the distance to its steady state growth trajectory.

Other coefficients reflect differences in regional steady-state levels. So, higher level of the Federal Government spending on assistance to the regions, lower share of fuel industry in industrial output in 1998, as well as belonging to the group of depressed regions correlates with lower average growth rates over the period under consideration. In other words, over the time interval considered there was observed convergence of regional income (GRP per capita) levels, not to the same but to an individual level depending on region's resource endowment, financial assistance status, and economic growth record.

Regions with higher share of fuel industry in industrial output have higher steady-state levels in contrast to those with the receiving status in 1998 and the depressed – those with the largest decline in industrial output during the first part of transitional period.

Statistical significance of the level of financial aid from the Federal Government in 1998 (the starting moment of the period under consideration) requires paying special attention. Presence and amount of financial assistance to a region depend on the level of its own income (GRP), which has been already considered in the model. In fact, that means that for two regions with the same GRP level and different financial assistance from the Federal Government the region with higher level of financial aid in 1998 has lower growth rates later on. Such situation is possible, from one hand, if financial aid was provided indeed to those regions that were more indigent, since further they had lower growth rates. From the other hand, this may indicate negative influence of financial assistance on regional economic growth. In this paper we don't make our goal to estimate the efficiency of financial assistance to the regions. We can just say that recipient regions have essentially lower estimated steady-state growth levels than other regions (at least over the period under consideration).

The estimate of spatial spill-over effect is represented by spatial coefficient lambda. Random component follows first order spatial autoregres-

48

sive process with the estimated coefficient equal to 0.57. Thus, regional average growth rates are substantially influenced by random shocks in neighboring regions besides random shocks in a region itself; the strength of such influence increases as distance between such regions decreases.

From position of conditional convergence analysis as well as new economic geography it's important to explain reasons for positive external effects of spatial interaction; thus, the alternative approach that allows to model spatial effects directly and, therefore, test hypotheses about the sources of such effects, seems more preferable[44].

Hypotheses about possible influence of endogenous and exogenous spatial lags on average growth rates can be tested in the context of alternative specification of the spatial error model – the spatial Durbin model.

Expressing a random component in (Eq. 1.8), $u = (I - \lambda W)^{-1} \varepsilon$, the spatial error model can be rewritten as follows:

$$
\begin{aligned}
g_T = & \delta + \beta y_0 + \phi_1 fapc98 + \phi_2 sh\_fuel98 + \phi_3 t2 + \\
& + \gamma_0 W y_0 + \gamma_1 W fapc98 + \gamma_2 W sh\_fuel98 + \gamma_3 W t2 + \quad (1.9) \\
& + \lambda W g_T + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I)
\end{aligned}
$$

with nonlinear constraints on the coefficients $\gamma + \lambda \phi = 0$, where $\gamma = (\gamma_0, \gamma_1, \gamma_2, \gamma_3)^T$ and $\phi = (\beta, \phi_1, \phi_2, \phi_3)^T$ (see also *Appendix 1*).

In spatial Durbin specification of the conditional convergence model log average growth rates of GRP per capita are explained by (1) the log initial GRP per capita level and the set of exogenous factors reflecting the hypotheses about the nature of economic growth over the period under consideration; (2) spatial lags on all exogenous variables; as well as (3) endogenous spatial lag on the log average growth rates of GRP per capita.

Estimation of the model is carried out by means of the maximum likelihood estimator, after that the equivalence of the initial spatial error model (Eq. 1.8) and the Durbin model (Eq. 1.9) is tested. Using the like-

---

[44] For more discussion of that problem see, e.g., *Fingleton B., Lopez-Bazo E. (2006).*

lihood ratio test the common factor hypothesis of nonlinear constraints on the coefficients in the Durbin model is tested.

The results of the Durbin model estimation and the likelihood ratio test are presented in *Tables 1.14* and *1.15*, respectively.

<div align="right">*Table 1.14*</div>

The Conditional Beta Convergence Model, the Spatial Durbin Model (Traveling Time Weight Matrix), Maximum Likelihood Estimator

Log average growth rates of GRP per capita over 1998–2004	Coef.	Std. Err.	z	$P>|z|$
Log GRP per capita (corrected on PPP) in 1998	–0.0273	0.0053	–5.17	0.000
Financial aid to the regional budgets per capita in 1998	–0.0220	0.0042	–5.29	0.000
Share of fuel industry in industrial output in 1998	0.0003	0.0001	2.43	0.015
Dummy-variable on depressed regions	–0.0163	0.0055	–2.97	0.003
Spatial lag: log GRP per capita (corrected on PPP) in 1998	0.0303	0.0161	1.88	0.060
Spatial lag: Financial aid to the regional budgets per capita in 1998	0.0228	0.0126	1.81	0.071
Spatial lag: Share of fuel industry in industrial output in 1998	–0.0007	0.0004	–1.62	0.105
Spatial lag: Dummy-variable on depressed regions	–0.0093	0.0229	–0.41	0.685
Intercept	0.0114	0.1547	0.07	0.941
Spatial lag: log average growth rates of GRP per capita over 1998–2004	**0.4993**	**0.1797**	**2.78**	**0.005**
Convergence speed, %	**3.03**			
Half-life, years	**23**			
Number of observation	77			
Variance ratio	0.386			
Squared corr.	0.431			
Sigma	0.01			
Log likelihood	212.8239			

50

Table 1.15

**Likelihood Ratio Test for Spatial Durbin Model
and Spatial Error Model**

Likelihood ratio	2.7767
Degrees of freedom	4
p-value	0.5959
sample estimates:	
Log likelihood of Spatial Durbin model	212.8239
Log likelihood of Spatial Error model	211.4355

Likelihood ratio test shows that we can not reject common factor hypothesis (nonlinear constraints on the coefficients) in the Durbin model (*p*-value = 0.5959). Hence, spatial error model (Eq. 1.8) can be represented as equivalent model which includes spatial lags on endogenous and all exogenous variables. The Durbin model (Eq. 1.9) is the final spatial specification of conditional convergence model (Eq. 1.7).

As follows from *Table 1.14*, conclusions about influence of exogenous factors hold for the Durbin model as well. It will take about 23 years for the economy of the average statistical region to cover a half of the distance to its steady state growth rate, what corresponds to the convergence speed approximately 3% per year. Nevertheless, regions notably differ in their steady state levels as it follows from the estimated coefficients on controlling growth factors. As in previously estimated models, lower spending of the Federal Government on assistance to the regions and higher share of fuel industry in industrial output correlate with higher average growth rates, while relatively lower growth rates of per capita GRP were observed in depressed regions.

Spatial lags on exogenous factors are insignificant at 5% significance level. In particular, the hypothesis that average growth rates correlate with geographically weighted initial GRP per capita levels of neighboring regions is not confirmed at this level of confidence. If this is the case,

then regional economic dynamics is unrelated to the *level* of economic performance of neighboring regions.

Regional steady states also essentially differ depending on what spatial growth cluster they belong to. The estimate of geographical coefficient at endogenous spatial lag on average growth rates ($\hat{\lambda} \cong 0.5$) points to the existence of substantial external effect of spatial dependence related to the *dynamics* of economic development of neighboring regions. Accordingly, for the average statistical region up to a half of the observed average (over the period under consideration) GRP per capita growth rates of the neighbors is transmitted to a given region. In other words 1% increase in spatially weighted (i.e. taking distance into account) average growth rates of neighboring regions results in a 0.5% increase in average growth rates of a given region.

Relatively higher spatially determined growth rates (see *Map 13* in *Appendix 3*) were observed in three macro regions: (1) on the whole in the European part of Russia, (2) regions of the South of Siberia (Novosibirsk Oblast, Tomsk Oblast, Kemerovo Oblast, Atai Krai, Altai Republic) and (3) regions of the South of Far East (Amur Oblast, Khabarovsk Krai, Evrei autonomous Oblast, Primorskii Krai)

As it has been mentioned above, from the standpoint of new economic geography endogenous spatial lag is determined by regional spill-over effect. The key role in explanation of spatially determined growth plays market potential concept according to which it's assumed that a given region takes an advantage from increasing demand in neighboring regions on goods and services produced in this region itself. The strength of such interaction among the regions involved is negatively correlated with the distance between them, since it approximates transport costs.

As aforementioned, as a measure of distance we use minimal traveling time to cover distance between the two regional centers by motor roads, while spatial weights were computed as inverse squared traveling time. Thus, freight transportation costs depend on the level of development and configuration of existing regional transport infrastructure. Other things being equal, more developed infrastructure leads to smaller traveling time and, therefore, smaller transport costs, which are reflected in their turn in higher spatially determined growth rates.

Besides market potential and infrastructure development, geography in general, human capital stock and migration are also essential factors explaining spatial spill-over effect.

In order to test the hypotheses about possible sources of spatially determined growth, we estimate the conditional convergence model with additional controlling factors. Formally, the model can be written as:

$$g_T = \alpha + \beta y_0 + \phi_1 fapc98 + \phi_2 sh\_fuel98 + \phi_3 t2 +$$
$$+ \phi_4 port + \phi_5 prpc98 + \phi_6 postgrads + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I), \quad (1.10)$$

where

*prpc*98 – regional passengers departures by railways calculated on per capita basis (see also *Appendix 3*, *Map 6*); this variable reflects regional transportation activity and, indirectly, connection of this region with the rest of the country, transport availability, and to some extent region's population mobility[45];

port – dummy-variable on availability of a seaport, that reflects, from one hand, geographical openness of a region to external markets, and from the other – higher infrastructure development level;

postgrads – postgraduate students per 10,000 inhabitants (average over 1998–2003). This variable reflects human capital stock and regional scientific research potential, as well as to some extent agglomeration factor: leading universities and research centers are located in big cities and metropolitan agglomerations, on the whole, characterized by higher share of persons with higher and postgraduate education[46].

---

[45] In some specifications we also used migration (per capita) data in 1998 and in average over 1998–2004 to test the hypothesis that region's labor supply dynamics positively correlates with GRP per capita average growth rates. Nevertheless, migration variables turned out to be insignificant (at 5% level), what can points to the presence in such specification, first, endogeneity problem – migration has influence on growth, what in its turn attracts to a region additional labor force, and, second, multicollinearity – migration can be positively correlated with the initial GRP per capita level. We examine this question in more detail in the next section.

[46] In alternative specification we used share of persons with higher and postgraduate degrees in economically active population variables and population of the main city vari-

Note that infrastructure and human capital, on one hand, serve as direct growth factors affecting the dynamics of economic development. On the other hand, the set of these variables reflects geographical proximity, region's population mobility and its scientific research potential, as well as the rate of agglomeration – in other words, factors favoring rise of positive external effect of spatial dependence serving, in its turn, as additional factor of growth.

The results of estimation of the model (Eq. 1.10) and spatial diagnostic tests of OLS-residuals are presented in *Tables 1.16* and *1.17*, respectively.

Table 1.16

Conditional Beta Convergence Model with Infrastructure and Human Capital Factors, OLS

Log average growth rates of GRP per capita over 1998–2004	Coef.	Std. Err.	*t*	*P>|t|*
Log GRP per capita (corrected on PPP) in 1998	–0.0280	0.0050	–5.67	0.000
Financial aid to the regional budgets per capita in 1998	–0.0149	0.0040	–3.77	0.000
Share of fuel industry in industrial output in 1998	0.0003	0.0001	2.53	0.014
Dummy-variable on depressed regions	–0.0115	0.0049	–2.37	0.021
Dummy-variable on availability of a seaport	0.0184	0.0051	3.61	0.001
Passengers departures by railways per capita in 1998	0.0006	0.0002	2.3	0.024
Postgraduate students per 10,000 inhabitants (average over 1998–2003)	0.0070	0.0034	2.05	0.045
Intercept	0.3442	0.0492	6.99	0.000
Convergence speed, %	**3.12**		LIK	220.179
Half-life, years	**22**		AIC	–5.511
Number of observation	77		BIC	–740.081
F(7, 69)	10.36		White	33.33
Prob > F	0.0000			(0.4514)
R-squared	0.5124			
Adj *R*-squared	0.4629			
Root MSE	0.0147			

able. Nevertheless, these variables were highly correlated with passenger departures, which resulted in their low statistical significance.

Estimated coefficients point in favor of the fact that regions with more developed infrastructure and higher human capital stock were characterized by relatively higher steady-state levels over the time interval under consideration.

Spatial Diagnostic Tests of OLS-Residuals in Conditional Convergence Model with Infrastructure and Human Capital Factors

Weight matrix	Traveling time		Market potential	
Test	statistic	*p*-value	statistic	*p*-value
Spatial Error:				
Moran's *I*	1.718	0.086	0.925	0.355
Lagrange multiplier	1.338	0.247	0.125	0.724
Robust Lagrange multiplier	1.342	0.247	0.028	0.867
Spatial Lag:				
Lagrange multiplier	0.46	0.498	0.101	0.75
Robust Lagrange multiplier	0.464	0.496	0.004	0.948

Addition to the model of transport infrastructure and human capital stock (and agglomeration rate as well) factors leads to insignificance of spatial correlation in the model's residuals (Moran's *I* statistic is insignificant at 5% level of significance). This remarkable result shows that in many respects spatial dependence may be carried out by means of these variables.

Relative impacts of exogenous factors on average growth rates are presented in *Table 1.18*. Values in the last column show by how many percent points average growth rates of per capita GRP would change in response to one standard deviation change of an exogenous factor[47]

---

[47] Analysis of a dependent variable change in response to a one standard deviation change in the exogenous variable is most informative in our case, since it allows to demonstrate the borders of this factor's influence on the dependent variable by taking into account its variation (3 standard deviations cover 99% of variable's variation). Such approach is used, e.g., in *Da Mata D. et al. (2005)*.

As follows from the table, the most important factor affecting growth is the log GRP per capita in 1998. Its influence is negative: for a typical region one standard deviation increase of the initial level of per capita income would lead to a 1.25% decrease of average growth rates. The most substantial among controlling growth factors is per capita financial aid to the regions (with negative sign): one standard deviation rise in this factor corresponds to a 0.76% decline in regional growth rates. The second important factor by its relative impact is availability of a seaport (increases GDP per capita growth rate by 0.67%). The third most important factor is share of fuel industry (0.51%). The least important by their impact, but still significant, are the variables for depressed regions (–0.46%), passenger transportation by railways (0.46%), and number of postgraduate students per 10,000 inhabitants (0.4%).

Table 1.18

Relative Impact of Exogenous Factors of the Conditional Convergence Model

Growth factors	Estimated coefficient times st. dev. of the factor	GRP per capita growth rate change, % change
Log GRP per capita (corrected on PPP) in 1998	–0.0117	–1.25
Financial aid to the regional budgets per capita in 1998	–0.0071	–0.76
Share of fuel industry in industrial output in 1998	0.0047	0.51
Dummy-variable on depressed regions	–0.0043	–0.46
Dummy-variable on availability of a seaport	0.0062	0.67
Passengers departures by railways per capita in 1998	0.0043	0.46
Postgraduate students per 10,000 inhabitants (average over 1998–2003)	0.0037	0.40

The results obtained above, in general, point that more developed infrastructure and population mobility, better transport availability as well as higher human capital stock lead to higher economic-geographic proximity and favor to diminish spatial barriers to growth, what in its turn positively correlates with higher average growth rates of GRP per capita.

New economic geography predicts in the long-run formation of, first, relatively wealthy and at the same time fast growing macro regions and, second, relatively lagging (underdeveloped) periphery. It would be short to draw such conclusion from our study since the time interval under consideration is quite small (7 years). Nevertheless, the above analysis shows that even over the considered time period one can talk about formation of distinct clusters of regions with high spatially determined growth rates of GRP per capita. In general in such macro regions leading roles have several relatively wealthy and fast-growing regions characterized by greater openness to foreign markets and favorable economic-geographic location, as well as by developed infrastructure and relatively higher human capital stock. Economic growth in such regions through the spatial spill-over effect (due to market potential and diffusion in broad sense of new technologies and information as well) spreads to neighboring regions, resulting in higher steady-state GRP per capita levels of the latter and leads to accelerated economic growth.

Note also that the case of low spatially determined growth rates in some region does not necessarily mean that such regions experienced low actual growth rates of GRP per capita. Such regions' neighbors' impact may in some cases be negligible due to its unfavorable geographical location. In general case, relatively higher actual regional growth rates (except for spatial spill-over factor) correlate with lower financial aid of the Federal Government, higher level of fuel industry and, on the whole, with more developed industrial sector.

Conclusions

1. Regional disparity indicators for per capita GRP over the 1996–2004 period do not show statistically significant changes. The above analysis points to the lack of Sigma Convergence or divergence of regional income. Actual changes of the coefficient of variation are statistically

insignificant and can not serve as a basis to draw conclusion about increasing or decreasing tendencies of regional disparity.

2. Estimation of unconditional beta convergence model does not provide statistically significant convergence coefficient. We can confirm neither unconditional convergence nor unconditional divergence of regional economic performance levels.

3. Average growth rates of per capita GRP are positively spatially clustered, which emphasizes the importance of spatial differences of regional growth. Some regions tend to form distinct groups with similar growth regimes: relatively fast-growing regions generally surrounded by relatively fast-growing neighbors. Even during a relatively short period under consideration (1998–2004) one can talk about significant spatial heterogeneity in economic development of Russian regions, which obviously should be taken into account in empirical studies of regional growth.

4. Estimation of minimal convergence model also provides statistically insignificant convergence coefficient. In other words we can not confirm the minimal conditional convergence hypothesis.

5. The conditional convergence hypothesis is not rejected. Other things being equal, regions with low initial GRP per capita had higher average growth rates over the period under consideration. It takes 22–24 years for an average Russian region to cover half of the distance to its steady-state growth level, which corresponds to the convergence speed of approximately 2.9–3.1% per year.

6. Regions substantially differ in their steady state levels as follows from the estimation results:

   - As in previous paper under IET-CEPRA project on regional development (*Drobyshevsky, Lugovoy, Astafyeva et. al. (2005)*) we do not reject hypothesis that financial assistance to the regions did not contribute to higher GRP per capita growth. Regions with higher values of financial aid (calculated on per capita basis) in starting 1998 were characterized by lower average growth rates of per capita GRP over the whole period under consideration.

   - Relatively higher steady-state growth trajectories were observed in regions with higher share of fuel industry in industrial output

in 1998. Accelerated growth in these regions took place in part due to accelerated growth in the industry itself but mainly because of income effect owing to high world market prices of energy resources.

- Lower growth rates were observed in underdeveloped regions of the depressed group. Regions with greater than 40% decline in industrial output in 2004 (compared to 1990 level, with the industry share in GRP not less than 20% in 2003) had relatively low steady-state levels over the period under consideration. Nevertheless, regions of this group are the regions with low staring conditions and considerable under-utilized industrial capacities, which could argue for unused potential for growth.
- Higher average GRP per capita growth is typical to regions with seaports, which reflects positive influence on the steady-state levels, on one hand, of higher infrastructure development level and, on the other hand, of higher proximity and openness to foreign markets: regions benefit from favorable economic-geographic location and take additional advantages from international trade.
- Passenger transportation (per inhabitant) by railways correlates with relatively higher growth rates. This variable reflects configuration of existing infrastructure and shows that regions with higher transport accessibility and, in part, higher population mobility have higher steady-state levels of GRP per capita.
- Finally, higher growth rates were observed in regions with higher number of postgraduate students per 10,000 inhabitants (average over 1998–2003), that reflects the importance for regional growth of such factors as scientific research potential and human capital stock.

7. The most important among controlled growth factors (besides the log GRP per capita in 1998) the most significant factors of regional growth are as follows: financial assistance to the regions (with negative sign), availability of a seaport (positive sign), and share of fuel industry in industrial output (positive sign).

8. Regional steady states substantially differ depending on what spatial cluster these regions belong to. Average growth rates of GRP per capita are positively and significantly correlated with average growth rates of neighboring regions through the endogenous spatial lag. The geographical coefficient estimate points to the existence of notable spatial spill-over effect: for the average statistical region 1% increase in spatially weighted (i.e. taking distance into account) average growth rates of neighboring regions leads to a half percent increase in average growth rates of a given region.

9. Spatial lags on exogenous variables turn out to be insignificant; in particular, regional growth rates correlate with the growth rates but not the initial levels of GRP per capita in neighboring regions. In other words, regional economic dynamics is spatially related to the *dynamics* but not the *levels* of economic performance of other regions.

10. GRP per capita average growth rates are positively and significantly correlated with the level of infrastructure development, i.e. with presence of a sea port, substantial passenger transportation per inhabitant, number of postgraduate students per 10,000 inhabitants. Developed infrastructure reflects higher population mobility and smaller transport costs, which lead to decrease of spatial barriers for growth.

11. Over the period under consideration we can note the tendency of formation of separate clusters of regions with high spatially determined growth rates. These are regions of (1) the European part as a whole, (2) the South of Western Siberia and (3) the South of Far East. In such macro regions most important are several relatively wealthy and fast-growing regions characterized by greater openness, developed infrastructure, and relatively higher human capital stock. In accordance with the predictions of new economic geography economic growth in leading regions is transmitted to neighboring regions through spatial spill-over effect and promotes accelerated economic development of the latter.

Model of Economic Growth in Russian Regions

In this section we discuss model of economic growth in Russian regions and present its estimation results. Following Dani Rodrik's scheme

(see *Fig. 1.1* above) we attempted to disentangle influence of proximate and deep factors of growth by utilizing a system of 3 simultaneous equations (SEM). The first equation describes regional GRP growth and uses proximate factors (labor, capital, and productivity growth). As was aforesaid, proximate factors are themselves endogenous, therefore, the other 2 equations of the SEM define accumulation of direct growth factors in regions by using proxies of deep determinants: institutional factors, factors of natural geography, and factors of economic geography of Russian regions.

SEM was estimated in two alternative ways: first, by 3-step least squares method (3SLS), second, by full information maximum likelihood method (FIML). In the former case estimation follows Mundlak's model specification (see description above); in the latter case we use spatial weight matrix in our system. Each approach rests on its own set of assumptions that are discussed below.

Output Growth Equation

The model rests on conventional production function, augmented to include natural resources:

$$Y = AK^{\alpha}(HL)^{\beta}(1+R),\qquad\qquad(1.11)$$

where

A – total factor productivity, TFP, (not determined by extraction of natural resources);

L – labor input;

K – capital input;

H – human capital;

R – natural resources input.

Introduction of natural resources to the production function was designed to take account of resource rent, acquired by some regions. Such regions have an apparent advantage: their output can be substantially greater than that of other regions with the same amount of labor and capital.

Dividing Eq. (1.11) by labor and taking logarithms we arrive at:

$$y = a + r + \alpha k + \beta h + \gamma l, \qquad\qquad (1.12)$$

where

$r = \ln(1 + R)$,

$\gamma = \beta + \alpha - 1$ (for constant return on scale $\gamma = 0, \beta = 1 - \alpha$),

$l = \ln(L)$,

y, k, h – logs of corresponding variables (Y, K, H) per capita.

Rewriting Eq. (12) in growth terms we arrive at following:

$$\dot{y} = \dot{a} + \dot{r} + \alpha\dot{k} + \beta\dot{h} + \gamma\dot{l}. \qquad\qquad (1.13)$$

In this specification growth of TFP ($\dot{a}$) refers to growth of output resulting from sources other from growth of capital and labor inputs and natural resources extraction ($\dot{r}$).

Equation (1.14) is a general specification of the growth equation estimated under SEM framework:

$$\dot{y}_{it} = b_0 + b_1 \cdot inv_{it} + b_2 \cdot m_{it} + b_3 \cdot pgrs_i + b_4 \cdot rawind_{it} +$$
$$+ b_5 \cdot port_i + b_6 \cdot y_{i0} + \varepsilon_{it} \qquad (1.14)$$

where

$\dot{y}_{it}$ – log of GRP per capita growth rate in region i in year t;

inv_{it} – log of investment per capita in region i in year t; defines capital accumulation in the region ($\dot{k}$ in Eq. 1.13);

m_{it} – growth rate of population of region i in year t; dependent on migration; describes growth of labor input in the region ($\dot{l}$ in Eq. 1.13);

$pgrs_i$ – period average number of postgraduate students in region i per 10,000 inhabitants ($\dot{h}$ in Eq. 1.13);

$rawind_{it}$ – share of extractive industry in total industrial output of region i in year t; describes resource intensiveness of the regional industrial output (r in Eq. 1.13);

$port_i$ – presence of a navigable non-freezing sea port in region i dummy variable (equal to 1 if a port is present, equal to 0 if a port is absent); describes resource endowment (r in Eq. 1.13);

y_{i0} – per capita GRP in region i at the start of the period;

ε_{it} – residual.

Notice that Eq. (1.14) contains two geographic variables: share of extractive industry in total industrial output ($rawind_{it}$) and non-freezing sea port dummy variable ($port_i$). The former variable describes degree of extractive specialization of region and can act as a rough proxy for natural resource rent. The latter, on one hand, takes account of economic activity (construction of the port); on the other hand, port is an element of transport infrastructure, prerequisite for which is a number of geographic advantages of the region (geographic location, climate etc).

Note another possibility, i.e. a region that has all necessary geographic prerequisites for construction of a sea port so far has not constructed one – in this case a proxy variable of access to sea routes would be preferred. Taking into account short time period of our research data, we consider construction of new ports or expansion of existing ones.

We'd like to point out that inputs of labor and capital are endogenous, and thus estimation of Eq. (1.13) by OLS would yield biased results. We specify SEM equations for capital and labor and describe them below.

Labor Factors Input

In the long run supply of labor is determined by natural growth rate of population (birth and mortality rates) and migration. In the short run rates of birth and mortality change slightly, therefore, the principal source of population growth and, thus, of labor supply is migration.

In empirical literature one can find various variables that are argued to influence migration. Among the most widespread factors are income per capita, climate, and population density (*Barro (2003, p.483)*).

Exploring migration in Russia from historical point of view, one can notice centrally planned migration flows directed to northern and under-developed territories during Soviet period. After collapse of the USSR, such centralized programs were suspended. Since then reverse migration, i.e. from developing regions to areas of old settlement, commenced. As other significant factors of migration, we considered such determinants as population of the largest regional city, transportation of passengers by railway per capita, and quality of regional institutions (corruption in police administration index). Former two variables describe degree of agglomeration and degree of transport infrastructure development, while the latter can reflect, first, presence of informal barriers to migration, second, additional expenses for non-residents of the region, third, quality of the law and order service in the region. Such characteristics can influence population's will to reside in a given region, i.e. migration flows.

Thus, we arrive at the migration regression, which is as follows:

$$m_{it} = c_0 + c_1 \cdot \dot{y}_i + c_2 \cdot tjan_i + c_3 \cdot gpop_i + c_4 \cdot lcity_i +$$
$$+ c_5 \cdot rpass_i + c_6 \cdot cpol_i + \qquad\qquad , \qquad (1.15)$$
$$+ c_7 \cdot rurdens_i + \eta_{it}$$

where

m_{it} – migration balance in region i in year t (number of persons arrived minus number of persons left);

$\dot{y}_i$ – period average GRP per capita growth rate in region i;

$gpop_i$ – population growth rate in region i, 1926–1989;

$tjan_i$ – period average January temperature in region i;

$lcity_i$ – period average population of the largest city in region i;

$rpass_i$ –number of passenger trips by railway per capita in region i in 1995 (equal to 0 for regions without railway roads);

$cpol_i$ – index of corruption in police administration in region i (the greater the magnitude of the index, the less corrupted is the regional police; source: «OPORA-VTsIOM» 2005);

$rurdens_i$ – log of rural population density;

η_{it} – residual.

Capital Input Factors

Capital input in regions is determined by the current stock of capital and new investment that in turn are determined by the expected revenue from new investment and risk of doing business.

Investment model employed in the SEM is as follows:

$$inv_{it} = d_0 + d_1 \cdot y_i + c_2 \cdot tjan_i + c_3 \cdot pm_i +$$
$$+ c_4 \cdot fuel_{it} + c_5 \cdot phone_i + c_6 \cdot inst_i + v_{it}$$

$$\tag{1.16}$$

where

inv_{it} – investment per capita in region i in period t;

y_i – period average GRP per capita in region i;

pm_i – dummy for permafrost in region i;

$fuel_{it}$ – fuel industry output in region i in period t;

$phone_i$ – quantity of fixed phones per capita in region i;

$inst_i$ – institutional variable (differed by specification):

   - officials corruption index in region I (the greater the magnitude of the index, the less corrupted are the regional officials; source: «OPORA-VTsIOM» 2005);
   - dummy for 10 regions with highest legislative risk (source: Expert RA rating);
   - dummy for 10 regions with lowest legislative risk (source: Expert RA rating);

- dummy for regions to which Standard-and-Poor's assigned a credit rating;

v_{it} – residual.

Gross Regional Product per Capita

The three equations above explain dynamics of corresponding flows: growth of GRP, investment, labor resources changes. At the same time in equation of GRP growth and investment we utilize level of GRP per capita, which is a stock variable. In the GRP growth equation this variable is included to take account of conditional convergence, which presence is confirmed by the results presented in the preceding part of this paper. In the investment equation level of GRP per capita is a proxy for business activity in the region. We hypothesize that level of investment is positively correlated with level of per capita income.

Logic of Dani Rodrik's scheme (*Fig. 1*) can be applied to level of income as well. GRP per capita is an endogenous variable, since in the short run level of GDP pre capita is not as volatile as its growth rate.

To escape this endogeneity problem we employed 2 different specifications of the simultaneous equations model: one included lagged (predefined) GRP per capita; the other – additional (forth) equation, to explain level of GRP per capita. The former specification was estimated under SEM framework by FIML. The latter, augmented specification was estimated by 3SLS.

The advantage of the augmented approach is in opportunity to analyze factors, related to the level of per capita GRP, thus, while Eq. (1.14) allows estimation of short-run factors of recovery growth, equation for GRP level allows estimation of long-run growth factors.

Specification of the additional equation for level of GRP is similar to the growth equation specification:

$$y_{it} = a_0 + a_1 \cdot inv_{it} + a_3 \cdot pgrs_i + a_4 \cdot rawind_{it} +$$

$$\text{(1.17)}$$

$$a_5 \cdot lcity_i + a_6 \cdot al_i + a_7 \cdot ea_i + \delta_{it}$$

where

y_{it} – log of GRP per capita in region i in year t;

inv_{it} – log of investment per capita in region i in year t;

$pgrs_i$ – period average number of postgraduate students in region i per 10,000 inhabitants;

$rawind_{it}$ – share of extractive industry in total industrial output of region i in year t; describes resource intensiveness of the regional industrial output;

$lcity_i$ – period average population of the largest city in region i;

al_i – log of the share of economically active population share in the total population;

ea_i – log of the share of employed population in the economically active population;

δ_{it} – residual.

Differences between the level of GRP specification and the growth of GRP specification can be summarized as follows:

1) Investment variable is used as a proxy for the utilized capital. Usually stock of capital is used for explanation of GRP level, nevertheless, it's crucial to keep in mind that existing data on capital stock unsatisfactory reflects level of actually utilized capital. In our view, level of investment is a good proxy for the stock of utilized capital unlike the official statistics of the capital stock[48].

2) Input of labor is omitted from equations. In specification (2) right hand side of the equation is supposed to include labor inputs, unless constant return to scale is assumed. However population of region variable reflects size of the region and is strongly correlated with other explanatory variables, especially with population of the largest city of the region.

---

[48] See, e.g. Bessonov (2005). Preference of the investment variable employment is also confirmed in our earlier work: *Drobyshevsky S., Lugovoy O., Astafyeva E. et al. (2005)*.

Thus, to mitigate multicollinearity and endogeneity problems we decided to omit this variable from the equation.

3) Largest city population variable was added to the specification to take account of the agglomeration factor and its influence on the productivity.

4) Sea port variable was excluded due to statistical insignificance in this specification.

5) Model was augmented to include specifics of regional employment. Differences in structure of employed and economically active population among regions is likely to be important factor of productivity per resident. In regions with lower share of employed population one should expect lower productivity per inhabitant.

Both specifications of the SEM and results of their estimation are presented below.

Estimation of the Model with 3SLS

As above mentioned, this model is distinguished by endogenization of GRP per capita level and, accordingly, augmentation of the model to include the additional equation. Estimated system is as follows (Eq. 1.18):

$$
\begin{aligned}
y_{it} &= a_0 + a_1 \cdot inv_{it} + a_3 \cdot pgrs_i + a_4 \cdot rawind_{it} + \\
&+ a_5 \cdot lcity_i + a_6 \cdot al_i + a_7 \cdot ea_i + \delta_{it} \\
\dot{y}_{it} &= b_0 + b_1 \cdot inv_{it} + b_2 \cdot m_{it} + b_3 \cdot pgrs_i + \\
&+ b_4 \cdot rawind_{it} + b_5 \cdot port_i + b_6 \cdot y_{i0} + \varepsilon_{it} \\
m_{it} &= c_0 + c_1 \cdot \dot{y}_i + c_2 \cdot tjan_i + c_3 \cdot gpop_i + \\
&+ c_4 \cdot lcity_i + c_5 \cdot rpass_i + c_6 \cdot cpol_i + c_7 \cdot rurdens_i + \eta_{it} \\
inv_{it} &= d_0 + d_1 \cdot y_i + c_2 \cdot tjan_i + c_3 \cdot pm_i + \\
&+ c_4 \cdot fuel_{it} + c_5 \cdot phone_i + c_6 \cdot inst_i + v_{it}
\end{aligned}
\tag{1.18}
$$

We use the same notations as above.

This system was estimated according to Mundlak's specification: all variables that change in time and among regions were used to generate two new types of variables: averages over time and deviations from these

averages (see details in the corresponding section above). Results are presented in *Table A2-1* of *Appendix 2*. Obtained results allowed us summarize specific properties of different types of regions as follows.

Regions with greater level of income (GRP per capita) are characterized by the following features:
– greater level of investment per capita;
– greater share of extractive and resource-intensive industries in output (including fuel, ferrous, non-ferrous industries, and forestry);
– greater share of economically active population in total population of the region;
– greater share of employed in economically active population;
– greater number of postgraduate students per 10,000 residents;
– greater population of the largest city.

Higher rate of GRP growth is facilitated by:
– greater level of investment per capita;
– growth of investment per capita;
– growth of extractive industries output in the industrial output;
– availability of an unfreezing sea port;
– lower initial GRP level;
– growth of the economically active population share in total population;
– growth of the employed population in economically active population.

Regions with greater inflow of migrants are characterized by the following features:
– higher rate of GRP per capita;
– old settlements (regions with lower growth of population during the Soviet period receive more migrants);
– regions with higher density of rural population are characterized by greater number of migrants;
– less harsh climate as measured by lower average temperature of January;
– more developed transport infrastructure as measured by railroad transport;
– greater population of the largest city of the region.

Regions with higher level of investment are characterized by the following features:

–   greater level of GRP: prosperous regions have higher level of GDP per capita;
–   warmer climate as measured by lower average temperature of January; at the same time regions with permafrost presence have higher level of investment, which can be explained by costs of natural resource extraction in such areas;
–   presence of fuel industry, greater output of fuel industry, and its growth;
–   greater development of communications network as measured by penetration of stationary phones;
–   lower officials corruption index;
–   lower legislative risk; regions with highest legislative risk obtain less investment per capita, for instance, in 10 regions with minimal legislative risk value of investment per capita is 1.5–2% higher compared to Russian average;
–   availability of investment rating; regions to which Standard-and-Poor's assigned a credit rating have statistically significant greater level of investment per capita (about +2% to Russian average).

Impact of different types of growth determinants on regional economic growth

Rodrik's typology of determinants of growth is not easy to apply to individual variables instead of groups of factors. We cannot be sure what effects we observe (as it was above mentioned, the same variables may be of geographic nature but work as proxy for institutions – so we can't say whether it's geography or institutions that attributes to growth). In Rodrik's scheme (see *Fig. 1.1*) these relations are represented by arrows that connect institutions, trade, and geography with each other.

For more accurate attribution of factors to specific groups we use two types of typology of our variables. The first typology divides factors by their nature (physical geography, economic geography, institutions) and the second – by their relation to human activity (i.e. whether a certain advantage or disadvantage is caused mainly by natural conditions or it's a result of development path). The first typology is preferred from theoretical point of view, while the second one is more useful for policy implications.

70

The first typology of growth factors:
- Physical geography – variables that reflect natural conditions (in our list of variables – climate and permafrost);
- Economic geography – variables that reflect differences in regional economic geography (agglomeration, infrastructure, access to seaport etc.);
- Institutions – variables that represent institutional measurements and administrative environment in the region (corruption, trust to government etc.);
- Economic factors – dependent endogenous variables that also have influence on growth (e.g., GDP per capita).

The second typology of growth factors:
- Geographic – factors that reflect use of geographical advantages or disadvantages of the region (natural resources, climatic conditions, access to sea coast etc.);
- Path-dependent – factors that reflect historically formed features of the region. It includes pattern of settlement (rural population density and size of major city), infrastructure (railways, phone prevalence), and total population growth during the Soviet period;
- Institutions (see the first typology);
- Economic factors (see the first typology).

Table 1.19

Input of factors* in GRP growth according to the model, Eq. (1.18)

1st typology		2nd typology	
Group of factors	Input, %	Group of factors	Input, %
Economic geography	48.5	Geographic	45.5
Physical geography	12.8	Path-dependent	36.1
Institutions	11.4	Institutions	11.4
Economic factors	39.9	Economic factors	39.9**
Error	69.6	Error	69.6

Note: * Input is calculated as standard deviation of sum of variables multiplied by corresponding coefficients divided by standard deviation of growth rates (dependent variable).

** Sums in columns are more than 100% because of factors multicollinearity. In other words when we are speaking about input of geography we cannot completely exclude input of institutions from the group, because they are linked by cause-effect relations with each other.

Below we discuss these typologies in more detail.

The first typology of growth factors

Having acquired results of the regional growth model we can analyze relative importance of different growth factors for each region of Russia. *Map 15* (*Appendix 3*) illustrates results of SEM estimation, i.e., comparative impact of factor groups by regions, in which all factors are grouped according to the first typology. The predicted growth rates are based on assumption that influence of each factor is the same for any region of Russia.

The first bar reflects the impact of variables that represent physical environment (mean temperature, permafrost). The map shows that overall contribution of these factors to growth is very low and dependence of growth on these factors is non-linear. Mean January temperature is positively correlated with growth – higher temperatures are associated with higher migration and investment per capita). Yet at the same time permafrost (that is present in regions with colder climate) is also positively correlated with growth. Thus, the most fast growing regions are those of the Southern Russia with warmest climate (Krasnodarskiy Krai, Rostov Oblast and republics of Northern Caucasus) and also those regions for which January temperatures are not extremely low but permafrost is present (Tumen Oblast, Murmansk Oblast, Krasnoyarskiy Krai, and Khabarovskiy Krai). The main implication is that direct influence of physical geography on growth is rather limited, besides its influence can be difficult to disentangle from influence of other factors. Obviously, mechanism of influence of these two factors on growth is different: physical geography factors are proxies for resource-intensive specialization, age of settlement as we mentioned above.

The second bar reflects input of the most broad and heterogeneous group of factors – factors of economic geography, which summarize regional advantages or disadvantages caused by regional location, specialization, and infrastructure. *Map 15* illustrates the following implications. First, resource-intensive specialization is favorable for growth only for those regions that have developed transport infrastructure. For instance, resource abundant regions yet with underdeveloped infrastructure such as Komi Republic, Khakassia, and Sakha have nearly zero or negative influ-

ence of economic geography factors on growth. Next, potential of sea-ports as growth poles is extremely high. All regions with seaports have reasonable positive contribution of economic geography factors on economic geography (Saint-Petersburg, Leningrad Oblast, Kalinigrad Oblast, regions on the Black Sea coast, and in the south of Far East).

Institutional factors were measured by very few direct variables; some institutional features are already captured by geographical variables. We used estimates of corruption rate among executive officials, level of legislative risk (from Expert RA rating) and a dummy for availability of the S&P investment rating. A big gap here is a question about property rights protection and quality of judicial institutions – so far these data for Russian regions were never collected by any think-tank or research group and it was impossible to obtain such data for this study. Nevertheless, the data we employed allows to make some conclusions about institutions and growth. Some regions have substantial comparative advantage that originates from their favorable institutional indicators. The highest institutionally determined growth rates have such regions as Saint-Petersburg City, Leningrad, Moscow, Samara, Vologda, Tumen, Tomsk, Sverdlovsk, Irkutsk Oblasts, Sakha, Tatarstan and Bashkortostan Republics and Yevreyskaya autonomous Oblast.

In a whole we can note forming poles of economic growth in Russian regions. First, these are regions with favorable economic geography characteristics, which would always generate higher growth for such regions, thus, increasing their steady-state growth rates. To such regions belong Moscow and Saint-Petersburg and their agglomerations (Saint-Petersburg is more than twice smaller as measured by population but has an advantage of a seaport). Second, we can expect good growth record from coastal regions specialized on external trade – Khabarovskiy Krai and Primorstkiy Krai in the Far East, Rostov Oblast and Krasnodarskiy Krai in the South, Kaliningrad Oblast in Baltic Region, and to a certain extent Arkhangelsk Oblast in the North of European Russia. Third, these are regions with resource-intensive industry that are specialized on oil and gas extraction (first of all, autonomous okrugs of Tumen Oblast) as well as some resource abundant regions with developed infrastructure and/or

big urban centers (Tatarstan, Bashkortostan, Permskiy Krai, Orenburg Oblast).

The last bar is the unexplained residual – difference between predicted (as if factors had the same influence and coefficients as in the overall country) and actual growth. This bar illustrates which regions' growth dynamics is satisfactory described by our the model and which regions' actual output significantly differs from expected. The best explanation of growth is achieved for such regions as Magadan Oblast, Karelia Republic, Penza, Lipetsk, Kostroma, Kaluga, Ulyanovsk, Chelyabinsk, and Kemerovo Oblasts. These regions are typical from the point of our model, i.e. there growth patterns are well aligned with general trend of Russian Federation, but one should not expect this development trajectory to remain the same in the future.

There is as well a group of regions with significant residuals. Some are growing faster than predicted and some – too slow taking account of their geographical and institutional characteristics. In the latter case one can conclude that some regions don't fully utilize their advantages and their growth potential is much higher than their actual growth, or some significant factors for these regions are omitted from our model. Such regions are: Krasnodarskiy and Stavropolskiy Krai, Saint-Petersburg, Tatarstain, Bashkortostan, Moscow, Irkutsk Oblasts, Primorskiy Krai, and, to a smaller extent, Tumen Oblast.

On the other hand, there is other group of regions (Omsk, Sakhalin, Tambov, Arkhangelsk Oblasts, and Mordovia republic), which is growing faster than expected taking into account their factor endowment, so they are now above their steady-state growth rate as predicted by the model. This either could mean that current growth record of these regions is due to a certain favorable but short-run factor and later they will experience slow-down, or again that our model misses some significant for such regions factors.

The second typology of growth factors

Map 16 (*Appendix 3*) corresponds to the factors grouping according to the second typology. Description of the factors the same, excluding two new groups: geographic factors and factors of path dependence. As shown in the *Table 1.19*, contribution of geographic factors to the growth

is the greatest (46%). The second (by impact value) group is a group of factors of path-dependence (settlement pattern, transport infrastructure, and communications), which explains about 36% of growth. So, development history matters and long-run steady-state growth levels are determined partly by regional past development.

Model of Regional Growth with Spatial Relations

Unlike the previous model, this is a simultaneous equation model (1.19) estimated using pooled data. We do not employ here Mundlak's approach because of technical difficulties of such estimation by FIML (Full Information Maximum Likelihood) that is very demanding to the system specification. Utilization of FIML allows to account for possible spatial relations in the growth regression and to acquire an alternative to 3SLS estimates.

$$
\begin{aligned}
\dot{y}_{it} &= b_0 + b_1 \cdot inv_{it} + b_2 \cdot m_{it} + b_3 \cdot pgrs_i + \\
&+ b_4 \cdot rawind_{it} + b_5 \cdot port_i + b_6 \cdot y_{i0} + \\
&+ b_7 \cdot pgrs_i + \rho \cdot \text{wd}(\dot{y}_{it}) + \varepsilon_{it} \\
m_{it} &= c_0 + c_1 \cdot y_i + c_2 \cdot tjan_i + c_3 \cdot gpop_i + \\
&+ c_4 \cdot lcity_i + c_5 \cdot rpass_i + c_6 \cdot cpol_i + \\
&+ c_7 \cdot rurdens_i + \eta_{it} \\
inv_{it} &= d_0 + d_1 \cdot y_i + c_2 \cdot tjan_i + c_3 \cdot pm_i + \\
&+ c_4 \cdot fuel_{it} + c_5 \cdot phone_i + c_6 \cdot inst_i + v_{it}
\end{aligned}
\qquad (1.19)
$$

where

$\text{wd}(\dot{y}_{it})$ – spatially weighted growth rates in all other regions;

ρ – spatial correlation coefficient;

and other notations are the same as above.

Estimation output of model (9) is presented in *Appendix 2* (*Table A2-2*). Most of the model (1.8) estimation results are confirmed with results of this model with slight differences, which are as follows: average January

temperature variable in investment equation is no longer significant, investment is less significant in the equation of growth rate. On the other hand, number of postgraduate students variable gained in significance, while being insignificant in equation (1.8).

Note statistical significance of the spatial lag in the growth equation. Thus, according to the estimates, GRP growth rate of a given region is related to the GRP growth rate of its neighboring regions. Spatial correlation coefficient is 0.77, which means, that the closer to each other are regions located and the higher are their corresponding GRP levels, the stronger is their mutual influence. To put it another way, 77% of the regional growth is transmitted to the neighbors.

Regional Growth Accounting

The obtained estimates describe accumulation of proximate growth factors, as well as interregional differences of such accumulation, which are in turn explained by the deep regional factors of growth. Therefore, not only do we have a general framework of different factors influencing regional growth, but we can also quantitatively estimate influence of every factor for each of Russian regions. Below we present growth accounting procedure, in which factors of growth are grouped according to the following methodology.

But first we analyze degree of influence of the chosen determinants of growth in borders of the country. In *Table A2-3* (*Appendix 2*) estimates of the factors' contribution to the corresponding variables changes and to the resulting GRP growth through the proximate factors (investment and migration) are presented. Such accounting is made for 2 models: (1.8) and (1.9). The former explains differences in contributions among regions (between effect), while the latter should be interpreted as an average of within (changes over time) and between estimates. Since the employed factors are not orthogonal, their sum is not necessarily equal to 100.

As follows from *Table A2-3*, investment and migration in the 3SLS model (1.8), i.e. proximate factors, explain about 91% and 61%, correspondingly, of interregional growth difference. Availability of a sea port contributes to productivity – this factor's contribution makes up about 32%. Contribution of conditional convergence is tantamount to variation of the interregional growth itself (–100%).

76

Estimation of the model (1.9) by FIML, as aforementioned, is carried out with inclusion of spatial correlation factor. Noteworthy is the fact, that spatial variable becomes the major contributor to the growth. It accounts for more than 57% of the explained growth, thus, share of other factors diminishes drastically: in this case investment and migration explain about 21% and 10% of growth rates variation, correspondingly, convergence – about 18%.

Scenario of Regional Economic Growth based on the SEM-Model

The simultaneous equations model results are relevant for explanation of the recovery growth period tendencies. According to the SEM estimates, geographic factors and built up during the previous years infrastructure are significant factors that explain differences in regional income growth (about half of the growth variance).

Trends of the recovery growth would be better realized if estimated results of the deep determinants of growth and factor accumulation relation are employed for prolongation of the contemporary tendencies under assumption of time-invariant influence of exogenous factors of growth.

Note, that these projections should not be treated as an economic forecast. For a valid forecast, additional assumptions about future changes in exogenous factors and assumption about prevalence of such tendencies over the forecast period should be made. In this paper we did not attempt to forecast exogenous factors. Although it's rather easy to project changes for geographical factors, such as average temperature of January or even presence of a sea port, this task becomes more arduous for variables like share of fuel industry in GRP.

In this paper we have constructed scenario of growth by prolonging current tendencies up to 2050 under premise that all exogenous variables will not change or, as in case of share of fuel industry, will change proportionally to the GDP growth.

According to our estimates, regions of the European part of Russia will be increasing GRP at the highest pace; by 2050 Saint-Petersburg and Leningradskaya Oblast will become leaders in GRP per capita, leaving Moscow and Moscow Oblast behind. Arkhangelsk and Rostov Oblast will have similar levels of GRP per capita. Wealthy regions of Eastern

Siberia and Far East will become regions with average per capita income level by the middle of XXI century, while the most indigenous regions will be Chuvash republic and 3 regions in the south of the Urals mountains: Sverdlovsk, Kurgan, and Chelyabinsk Oblast.

So, ceteris paribus, until 2050 western regions of Russia will continue growth at a substantially higher pace compared to the eastern. These results underline importance of the regional policy, purposed to stimulate growth in the east of Russia.

Major Implications for Regional Policy

Our analysis showed that indicators of regional inequality didn't have any significant trend during the period of study (1996–2004). Changes of the GRP per capita coefficient of variation over time were statistically insignificant, so we could not draw conclusion about presence of sigma convergence or sigma divergence.

The hypothesis of unconditional beta convergence for 1998–2004 period also wasn't confirmed, but there was significant conditional beta convergence if we include in the model exogenous factors that capture differences in steady-state growth rates. An average Russian region is catching up his long run steady-state growth ceteris paribus with the speed of 2.9–3.1% per annum. The half-life of convergence is about 22–24 years. The study shows that Russian regions have significant differences in their long run steady-state level of GRP per capita.

Note, that the model of conditional convergence allows wealthy regions to grow faster than poor regions on the transitional growth paths. Also regions can have the same long run average growth rates but different steady state levels of GRP per capita. This explains that cross-regional inequality will not inevitably decline in the long run. Thus, the task for regional policy is to change long-run steady state growth rates, but as we see from the regression most factors that shape this path are uncontrollable by government, so such programs for regional development have very limited power even if we assume that al resources are allocated efficiently.

Regional steady-state growth rates depend on the spatial cluster in which regions are located (see *Maps 8–13*, *Appendix 3*). Regions with high spatially determined growth rates are located (1) in the European

part of Russia, (2) in the South of Western Siberia and (3) in the South of Far East. Such regional growth clusters are headed by a few wealthy and fast-growing regions that are characterized by good location, developed infrastructure and endowed with human capital. Growth in such regions is translated to its neighborhood by means of spatial spillovers and trade. This shifts neighbors' steady state growth rates to higher positions and leads to spatial clustering of high and low growth rates. The other significant factor of higher steady state growth rates is availability of a seaport. This can be explained by advantages of tighter integration with global markets.

Infrastructure and human capital – unlike geography – are usually claimed to be endogenous factors. Investments in infrastructure development and in human capital foster economic growth in the long run.

From the point of view of new economic geography, development of infrastructure is crucial for transport costs. From this point we can consider two recommendations for regional policy. First, to improve quality of the existing roads and to construct modern traffic interchanges near large cities and metropolitan areas. Second, to make a more intense road network across interregional borders. These measures can make economic space more interconnected and increase market potentials for regions. The latter will have a positive effect on growth.

It's important to mention that low spatially determined growth rates do not inevitably result in low actual growth rates. So, higher growth rates together with low spatially determined growth rates are observed in regions with less financial aid from federal government, higher share of fuel industry in industrial output, and more developed manufacturing sector.

Share of fuel industry in industrial output is significantly and positively correlated with economic growth. Significance of this indicator, which reflects natural advantages of regions highly endowed with mineral resources, shows that such regions have higher long-run steady-state growth rates. This could be explained by high oil and gas prices during the period under consideration.

To the contrary, depressed regions have lower steady-state growth rates. Regions, whose decline of industrial output between 1990 and 2004

was greater than 40%, had lower per capita growth rates during the 1998–2004 period. One of the tasks of federal government is to enhance investment climate and business environment in the regions, make administrative functions more transparent and decrease corruption. Calculations show that good institutions and business environment can overcome disadvantages of location for peripheral regions.

Despite the significance of spatial linkages between regions, intensity of these linkages is lower than in other countries, for instance, countries of Western Europe, as other studies have shown. This can be in part explained by greater distances and poorer infrastructure, and this also means that there are substantial prospects for growth, would spatial connections among regions be strengthened.

Physical geography has a significant impact on economic growth. The influence of geography can be both direct – through migration of labor and capital – and indirect – through other factors (institutions, infrastructure). Regions that are endowed with mineral resources or have a seaport are growing faster than the others. The presence of such natural advantages is an unmanageable factor, but efficient use of natural advantages depends on human activity. That is the main message on geography for future regional and federal policy of development.

Development of infrastructure, communications and pattern of settlement impact growth substantially. These indicators are highly path-dependent and their present state is predetermined by historical patterns of settlement, urban development, industrial location decisions, railroad construction etc.

Quality of administration is also significant for investment and growth. For example, regions with higher legislative risks (Expert RA rating), as well as regions that are characterized by higher corruption among officials ("OPORA-VTsIOM" estimates) have less investment per capita. To the contrary, regions that have credit rating issued by S&P have statistically higher per capita investment levels.

As in the case of cross-country studies, about half of the growth can be accounted by uncontrollable and path-dependent factors that can't be effectively influenced by direct policy measures. Some regions have initial advantages and, thus, are likely to have higher growth rates. There is

a risk that inadequate government policy can become a barrier for growth in potentially prosperous regions. Two main issues for promotion of regional development must become a framework of regional policy: (1) development of infrastructure and (2) creation of a sound institutional environment, independent judicial system, and less corrupt and more transparent public service.

There was a significant unexplained residual for some regions, which means that some regions noticeably differ from the most Russian regions as described by our model.

Bibliography

Acemoglu D., *Simon Johnson, and James Robinson* (2005). Institutions as the Fundamental Cause of Long-Run Growth, Ch. 6, Handbook of Economic Growth.

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

Anselin, L. and A. K. Bera, R. J.G.M. Florax, and M. Yoon (1996). "Simple Diagnostic Tests for Spatial Dependence," Regional Scieence and Urban Economics, 26, 77–104.

Anselin L., Florax R.J.G.M. & Rey S. J. (ed.) (2004). Advances in Spatial Econometrics. Methodology, Tools and Applications. Springer, Berlin.

Anselin L. (2003). An introduction to spatial regression analysis in R, www.sal.uiuc.edu/stuff/stuff-sum/pdf/spdepintro.pdf.

Arthur B. (1989). Competing Technologies, Increasing Returns, and Lock-in by Historical Events, Economic Journal 97.

AutoTransInfo, http://www.autotransinfo.ru.

Baldwin R., Martin P., Ottaviano G. (2001). Global income divergence, trade and industrialization: The geography of growth take-off, Journal of Economic Growth 6, 5–37.

Barro, R. (1997). Determinants of Economic Growth: A Cross-Country Empirical Study, Lionel Robbins Memorial Lectures, MIT Press.

Barro, R., Sala-i-Martin X. Economic Growth. (2003, 2nd edition) The MIT Press.

Baumont C., Ertur C., Le Gallo J. (2002). The European Regional Convergence Process, 1980–1995: Do Spatial Regimes and Spatial Dependence Matter?

Bessonov V.A. (2005). Problemy postroyeniya proizvodstvennyh funktsiy v rossiyskoy perehodnoy ekonomike (Problems of Production Function Estimation for Russian Economy in Transition). Moscow, IET. In Russian.

Da Mata D., Deichmann U., Henderson J. V., Lall S. V., Wang H.G. (2005). Determinants of city growth in Brazil, Policy Research Working Paper Series, No3723. The World Bank.

Durlauf S., Johnson P., Temple J. (2005). Growth Econometrics Ch 08, Handbook Of Economic Growth.

Denison E.F. (1962). The sources of Economic Growth in the United States and the Alternatives before us, Committee for Economic Development. New York.

Denison E.F. (1967). Why growth rates differ: post-war experience in nine western countries, Brooking institutions. Washington.

Drobyshevsky S., Lugovoy O., Astafyeva E., Polevoy D., Kozlovskaya A., Trunin P., Lederman L. (2005) Factors of Economic Growth in Russia's Regions

Easterly W. (2002). The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics, The MIT Press.

Engerman, S. L., Sokoloff K. L. (2002). Factor Endowments, Inequality, and Paths of Development among New World Economies, NBER Working Paper 9259.

Expert RA, Credit Rating of Russian Regions http://www.raexpert.ru/ratings/regions.

Fingleton, B. (2002). Theoretical Economic Geography and Spatial Econometrics: Bridging the Gap between Theory and Evidence, in *Getis A., Mur J., Zoller H.* (Ed.) (2004), Spatial Econometrics and Spatial Statistics, Palgrave, London.

Fingleton B. (ed.) (2003). European Regional Growth, Springer, Berlin.

Fingleton B. (2004). Regional economic growth and convergence: insights from a spatial econometric, in Anselin L., Florax R. and Rey S. (Ed.) Advances in Spatial Econometrics, Springer, Berlin.

Fingleton B., Lopez-Bazo E. (2006). Empirical growth models with spatial effects, Papers in Regional Science, Blackwell Publishing, vol. 85(2), pages 177–198, 06.

Florax R., Folmer H., Rey S. (2003). Specification searches in spatial econometrics: the relevance of Hendry's methodology, Regional Science and Urban Economics 33: 557–579.

Fujita M., P. Krugman and A.J. Venables (1999). The Spatial Economy: Cities, Regions and International Trade, Cambridge (Mass.): MIT Press.

Fujita M., Thisse J.-F. (2002). Agglomeration and growth with migration and knowledge externalities, Kyoto University Institute for Economic Research, WP.

Gallup J., Gaviria A. & Lora E., (2003). Is Geography Destiny? Lessons from Latin America. Stanford: Stanford University Press.

Gallup J., Sachs J. (1998). Geography and Economic Development, Annual World Bank Conference on Development Economics, 1998 http://www.cid.harvard.edu/cidwp/001.htm.

Getis A., Mur J., Zoller H. (ed.) (2004). Spatial Econometrics and Spatial Statistics, Palgrave, London.

Granberg A., Zaitseva Yu. Proizvodstvo i ispolzovanie valovogo regionalnogo produkta: Mezhregionalniye sopostavleniya. Statya 2. Korrektirovki VRP s uchetom territorialnih razlichiy pokupatel'noi sposobnosti deneg. // Rossiyskiy Ekonomicheskiy Jurnal – 2002, No. 11–12, str. 48–70. (Production and Use of Gross regional Product – interregional comparisons. Article 2. Correction of GRP according to regional differences in purchasing power of money). In Russian.

Griliches Z., Jorgenson D.W. (1967). The explanation of productivity change // The Review of Economic Studies. V. 34. No. 3, 249–283.

Hausmann R. (2001). Prisoners of Geography. Foreign Policy 122 (Jan/Feb 2001).

Hill F. and Gaddy C. (2003). The Siberian Curse How Communist Planners Left Russia Out in the Cold, The Brookings Institution Press.

Islam (1995). Growth Empirics: A Panel Data Approach, The Quaterly Journal of Econometrics, 110, 4.

Karl T. L. (1997). The Paradox of Plenty: Oil Boom and Petro-Politics. Berkeley: University of California Press.

Mankiw G., Romer D., Weil D. (1992) A Contribution to the Empirics of Economic Growth, The Quaterly Journal of Econometrics, 107, 2.

Martin P., Ottaviano G. (1999). Growing locations: industry location in a model of endogenous growth, European Economic Review 43, 281–302.

Martin P., Ottaviano G. (2001). Growth and agglomeration, International Economic Review, Vol. 42, No. 4, 947–68.

Mellinger A., Sachs J & Gallup J. (1999). Climate, Water Navigability and Economic Development, CID Working Paper.

Mundlak Y. (1978). On the Pooling of Time Series and Cross Section Data Econometrica, 1978, vol. 46, issue 1, 69–85.

Mundlak Y. and Yahav J.A. (1981). Random Effects, Fixed Effects, Convolution, and Separation Econometrica, vol. 49, issue 6, 1399–1416.

Nefedova T. G. Geograficheskiye variatsii selskogo hozyaistva // Otechestvenniye zapiski, No. 1 2004. (Geographical variations of agriculture). In Russian.

NISP, Socialniy Atls Rossiyskih Regionov, 2006 (Social Atlas of Russian Regions) http://atlas.socpol.ru/index.shtml.

North D. (1990). Institutions, Institutional Change and Economic Performance, Cambridge University Press.

North D., Thomas R.P. (1973). The Rise of the Western World: A New Economic History, Cambridge: Cambridge University Press.

Olson M. (1983). The Rise and Decline of Nations. Economic Growth, Stagflation and Social Rigidities. Yale Press.

OPORA-VTsIOM (2005). Conditions and Factors Affecting the Development of Small Entrepreneurship in the Regions of the Russian Federation.

Penn World Tables, http://pwt.econ.upenn.edu.

Pilyasov A.N. Politicheskiye i ekonomicheskiye factori razvitiya rossiyskih regionov // Voprosi Ekonomiki No. 5, 2003 (Political and economic factors of development of Russian regions). In Russian.

Rodrik D. (1999). "Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses," Journal of Economic Growth, 4, 4, 385–412.

Rodrik D., (ed.) (2003). In Search of Prosperity: Analytical Narratives on Economic Growth. Princeton University Press.

Rosstat, Federalnaya Sluzhba Gosudarstvennoi Statistiki (Federal Statistical Bureau).

Sach J., Warner A. (1997). Natural Resource Abundance and Economic Growth**,** Center for International Development and Harvard Institute for International Development.

Sala-I-Martin X. (1997). 'I just Ran 2 Million Regressions', American Economic Review, Vol. 87 No. 2.

Smirnyagin L. V. Rayoni S.Sh.A. Portret Sovremennoi Ameriki M.: Mysl', 1989 (Regions of the USA. A portrait of contemporary America). In Russian.

Solow R.M. (1956). A Contribution to the Theory of Economic Growth, Quarterly Journal of Economics, Vol. 70 (1), 65–94.

Solow R.M. (1957). Technical Change and the Aggregate Production Function, The Review of economics and Statistics. V. 39 No. 3, 312–320.

Swan T.W. (1956). Economic Growth and Capital Accumulation, Economic Record, Vol. 32 (2), 334–361.

Temple, J (1999). The new growth evidence, Journal of Economic Literature, March, 37(1), 112–156.

Tuan Yi-Fu (1998). Escapism, Johns Hopkins University Press, Baltimore.

Verrill S. (2003). Confidence Bounds for Normal and Lognormal Distribution Coefficients of Variation, Research Paper 609, USDA Forest Products Laboratory, Madison, Wisconsin.

Verrill S., Johnson R. A. (2006). Confidence Bounds and Hypothesis Tests for Normal Distribution Coefficients of Variation, DRAFT Research Paper, USDA Forest Products Laboratory, Madison, Wisconsin.

Appendix 1: Spatial Econometrics

Weight Matrices

Modeling of spatial interactions requires specification of spatial weight matrices. Weight matrix formalizes an assumption that a region is related to a set of neighboring regions. There are several types of these matrices (*Anselin (1988)*; *Fingleton (2003)*): contiguity matrix, matrix of k nearest neighbors, distance-based matrix, as well as distance-based matrix considering a size (output) of a region. All matrices are quadratic with zeros on the diagonal. The latter is important in order to take into account only impacts of neighboring regions; in other words, the influence of a region on itself is excluded, since it is directly considered in a model. Each row of a matrix represents the weights of all other regions impact on a given region. As a rule, weight matrix is raw-standardized (row-sum of the weights equals to one): first, in order to provide well-conditioned property and to avoid problems while inverting matrix in statistical packages and, second, to consider not absolute but relative distances between regions in case of distance-based matrix.

The simplest way to take into account spatial interactions is formalized in binary **contiguity matrix**:

$$w_{ij} = \begin{cases} 0, & \text{if } i = j \\ 1, & \text{if } j \text{ is contiguous with } i \text{,} \\ 0, & \text{otherwise} \end{cases} \qquad (A1\text{-}1)$$

where one assumes that a region is influenced only by its direct neighbors, i.e. by regions adjacent to a given region. Accordingly, influence of regions contiguous to the neighbors (regions of the second, the third circle etc.) is assumed to be negligible, what does not seem realistic from both geographic and economic perspectives.

A more complex idea of neighborhood is realized in binary matrices of k nearest neighbors. This matrix is constructed as follows. At first, for a given region i one computes distances to all other regions d_{ij} and then

selects k shortest distances. The biggest value among these k distances serves as a cutoff distance, so that k regions with $d_{ij} \leq d_i(k)$ are declared to be k nearest neighbors for region i (for region i spatial dependence on those regions do matters, and does not on the others). Each region has exactly k nearest neighbors. Accordingly, if regions are neighbors in the above sense, one puts ones to the cells of corresponding row of a matrix, otherwise – zeros.

$$w_{ij}(k) = \begin{cases} 0, & if \quad i = j \\ 1, & if \quad d_{ij} \leq d_i(k), \\ 0, & if \quad d_{ij} > d_i(k) \end{cases} \qquad \text{(A1-2)}$$

For a purpose of stability of the results in spatial analysis one usually considers matrices of 10, 15, 20, and 25 nearest neighbors.

Note that since this matrix is binary, row-standardization leads to the fact that for a given region influence of its neighbors is taken into account with the same weights. Thus, it is rational to use such matrices when regions of an economic system under consideration are quite homogenous in areas and, therefore, in distances between regional centers, so that the territory coverage, where regions are influenced by their neighbors, is equal from one region to another. Use of equal weights in that case will be well-founded. Nearest neighbors matrices are often used in studies of European economic growth (*Fingleton (2003)*).

Nevertheless, application of these matrices for Russian regions showed that they provide unstable estimation results and work, on the whole, worse than distance-based matrices (see below): spatial correlation statistics computed with the use of nearest neighbors matrices (when they turned out to be significant) were essentially smaller in their values as compared to the same statistics computed for distance-based matrices. One can suppose why it happens: Russian regions are substantially heterogeneous in territory. For instance, in case of 10 nearest neighbors matrix, for regions of the Central European part or the South of Russia critical cutoff distance will be, in average, one thousand kilometers, while for

regions of Siberia and Far East – thousands kilometers. It is clear that in first case neighbors' influence on a given region is underestimated and in the second, conversely, – overestimated. Usage of equal weights is incorrect as well, since the regions substantially differ in size. As a result, estimation of spatial correlation statistics and econometric models was carried out with the use of distance-based matrices.

Spatial weights of a standard **distance matrix** are computed as follows:

$$w_{ij}(q) = \begin{cases} 0, & if \quad i = j \\ 1/d_{ij}^{\gamma}, & if \quad d_{ij} \leq D(q), \\ 0, & if \quad d_{ij} > D(q) \end{cases} \qquad \text{(A1-3)}$$

where d_{ij} – measure of distance between regional centers, and $D(q)$ – distance quartiles, $q = 1, 2, 3, 4$. As a rule, the γ coefficient is considered to be equal to two. In this case the weight coefficient represents the analogue of the coefficient of gravitation: gravity between regions is inversely proportional to the square of distance. The farther geographically the regions are, the lesser they influence each other. In case of $q < 4$ the critical cutoff distance (so that an influence on longer distances is negligible) is the corresponding distance quartile, $D(q)$. In case of $q = 4$ all distances are counted in a matrix, so that zeros are only on the diagonal.

Along with the usual distance matrix, in some papers (see, for example, *Fingleton (2002)*) the market potential matrix or the distance matrix counting the size of neighboring regions is used with the weights computed in accordance with the formula:

$$w_{ij}(q) = \begin{cases} 0, & if \quad i = j \\ A_j/d_{ij}^{\gamma}, & if \quad d_{ij} \leq D(q), \\ 0, & if \quad d_{ij} > D(q) \end{cases} \qquad \text{(A1-4)}$$

where A_j – a size or output of neighboring region j (GRP, industrial output, etc.).

In this paper as a measure of distance we use **minimal traveling time** needed to cover a distance between regional centers by motor roads.

Spatial Correlation Statistics

Next step in spatial analysis is testing of the model variables on possible spatial correlation. With help of the exogenously given weight matrix one estimates **global spatial autocorrelation statistic (Global Moran's I)**:

$$I = \frac{n}{S_0} \frac{\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} w_{ij}(Y_i - \bar{Y})(Y_j - \bar{Y})}{\sum\limits_{i=1}^{n}(Y_i - \bar{Y})^2}, \qquad \text{(A1-5a)}$$

where Y – a variable tested, $S_0 = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} w_{ij}$ – the sum of weights of spatial matrix W. Denoting $Z = Y - \bar{Y}$, deviation of a variable from the mean value $\bar{Y}$, it is convenient to rewrite (A1-5a) in matrix form:

$$I = \frac{n}{S_0} \frac{Z'WZ}{Z'Z} \qquad \text{(A1-5b)}$$

The spatial autocorrelation coefficient shows the degree of linear dependence between the vector Z of normalized values of the variable Y and the vector WZ of spatially weighted normalized values of Y in neighboring regions, which is called the **spatially lagged vector**. The values of the statistic I greater than the mathematical expectation $E(I) = -1/(n-1)$, which is computed in assumption of no global spatial autocorrelation, reflect positive spatial autocorrelation. That is,

higher normalized values of Y correspond to higher values of spatially weighted (with a given matrix) deviations from the mean of Y in neighboring regions. Accordingly, the values of the statistic smaller than the mathematical expectation point to the presence of negative spatial autocorrelation.

Dispersion of values of a variable with regard to its spatial lag is conveniently visualized by means of the **Moran Scatterplot**. This plot displays the values of the vector Wz, i.e. spatially weighted standardized values of a variable (ordinate axis), versus the values of the vector $z = (Y - \bar{Y})/sd(Y)$, i.e. standardized values of a variable itself (abscissa axis). In addition, the Moran scatterplot displays the regression line obtained by regressing the lag Wz on its variable z with the tangent slope equal to the coefficient of global spatial autocorrelation, I (in assumption of row-standardized weight matrix).

The diagram is divided into four quadrants (HH, LH, LL and HL); every quadrant is qualitatively characterized by a certain type of spatial proximity:

- The top-right quadrant of the diagram (HH) shows a group of regions with high values of a variable surrounded by regions with high values as well;
- The top-left (LH) – regions with low values surrounded by regions with high values;
- The bottom-left (LL) – regions with low values surrounded by regions with low values;
- The bottom-right (HL) – regions with high values surrounded by regions with low values.

Quadrants HH and LL are characterized by positive spatial autocorrelation reflecting grouping of regions by similar values. Quadrants LH and HL, conversely, are characterized by negative spatial autocorrelation pointing to the clustering by dissimilar values. Thus, spatial scatterplot may help to reveal atypical regions relating to quadrants LH or HL.

Identification of spatial correlation (clustering of regions) by means of the global spatial autocorrelation coefficient and the Moran scatterplot is the first step in spatial analysis. The Moran's I statistic shows only that

the values of the variable Y are more spatially clustered than under a random assignment, however, it does not explain why it happens. Therefore, the next step of the analysis is test of hypotheses about the nature of spatial dependence.

Consider classical linear multiple regression model:

$$Y = \alpha + X\beta + \varepsilon, \qquad \varepsilon \sim N(0, \sigma^2 I) \qquad\qquad \text{(A1-6)}$$

where $Y - (n \times 1)$ vector of values of the dependent variable, $X - (n \times k)$ matrix of explanatory variables, $\beta - (k \times 1)$ vector of the coefficients. At first model (A1.6) is estimated by means of the least squares estimator. Further, on the basis of spatial diagnostic tests of the OLS-residuals the hypothesis of no spatial autocorrelation of the residuals is either (1) rejected and then relying on the diagnostic tests[49] one chooses a certain specification of a spatial econometric model or (2) not rejected and then (A1.6) is the final specification of a model.

Usually in spatial analysis three types of model are considered: the spatial lag model, the spatial cross regressive model, and the spatial error model. Consider them one after another.

Spatial Lag Model

In matrix form the spatial lag model is written as:

$$Y = \alpha + X\beta + \rho WY + \varepsilon, \qquad \varepsilon \sim N(0, \sigma^2 I) \qquad\qquad \text{(A1-7)}$$

where WY – spatial lag of endogenous variable, ρ – the spatial autoregressive coefficient. Each element of the spatially lagged vector WY is the spatially weighted average (in accordance with the weights of exogenously given matrix W) of values of the variable Y in neighboring regions. In presence of the stochastic regressor WY, which is always correlated with the random component ε, the OLS estimator will provide in-

---

[49] For more details about spatial diagnostic tests see, e.g., *Florax R., Folmer H., Rey S. (2003).*

consistent estimates (*Anselin (1988)*). Therefore, estimation of model (A1-7) is carried out by means of either the maximum likelihood estimator or the instrumental variables method.

A convergence model with spatial lag on endogenous variable is as follows:

$$g_T = \alpha + \beta y_0 + \rho W g_T + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I) \qquad\qquad \text{(A1-8)}$$

The main hypothesis tested in such a model is that the growth rates of a variable under consideration of convergence analysis (per capita income) depend not only on the initial value but on the growth rates of this indicator in the neighboring regions. In other words, regions surrounded by fast-growing regions will grow with higher growth rates. Model (A1-8) can be interpreted as the minimal conditional convergence model.

Spatial Cross Regressive Model

An alternative way of modeling spatial dependence is to consider spatial lag only on explanatory variables. In general a model is written as:

$$Y = \alpha + X\beta + WX\gamma + \varepsilon, \qquad \varepsilon \sim N(0, \sigma^2 I) \qquad\qquad \text{(A1-9)}$$

where $\gamma - (k \times 1)$ vector of spatial coefficients. In presence of exogenous spatial lag a model can be estimated by means of the OLS estimator directly. While analyzing convergence of economic indicators and considering the initial value of an indicator as the only exogenous factor, one can obtain a noteworthy special case of model (A1-9):

$$g_T = \alpha + \beta y_0 + \rho W y_0 + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I) \qquad\qquad \text{(A1-10)}$$

Model (A1-10) represents the minimal conditional convergence model with spatial lag on the initial value. By implication, growth rates of a re-

gion depend, first, on the initial value in a region itself and, second, on the initial values of an economic indicator in neighboring regions.

Spatial Error Model

In spatial error model one assumes that the random component follows first-order spatial autoregressive process:

$$Y = \alpha + X\beta + u \qquad u = \lambda Wu + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I) \qquad \text{(A1-11)}$$

where λ – the coefficient of spatial correlation of the residuals. Estimation of the model by means of least squares estimator leads to inefficient estimates. Moreover the estimates may turn out to be biased due to omitted variables. Therefore, such a model is estimated by the maximum likelihood estimator or general method of moments. Expressing random component as $u = (1 - \lambda W)^{-1}\varepsilon$ and substituting in the initial equation, model (A1-11) can be transformed as follows:

$$Y = (1 - \lambda W)\alpha + X\beta + \lambda WY + WX\gamma + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I) \text{ (A1-12)}$$

with the restriction on the coefficients $\gamma + \lambda\beta = 0$. Model (A1-12) is called the spatial Durbin model and includes spatial lags on both endogenous and all exogenous variables. From the perspective of convergence analysis the model:

$$g_T = (1 - \lambda W)\alpha + \beta y_0 + \lambda W g_T + \gamma W y_0 + \varepsilon \qquad \varepsilon \sim N(0, \sigma^2 I) \text{ (A1-13)}$$

has the following noteworthy interpretation. Regional growth rates may be related to the growth rates in neighboring regions through the endogenous spatial lag and to the initial values of an economic indicator (GRP per capita) in neighboring regions through the exogenous spatial lag. Therefore, in this model one tests at once two basic spatial hypotheses: regions with higher growth rates are surrounded (in a sense of neighborhood given by the weight matrix) as a rule by (1) fast-growing regions (as

a result of the endogenous lag) and (2) wealthy regions (as a result of the exogenous lag).

Appendix 2. Estimation results of Russian regions economic growth model

Estimation results of simultaneous equation system, 3-SLS, 1997–2004, 77 regions

Equation	Obs.	Param.	RMSE	*R*-sq	chi2	*p*-value
GRP per capita, log	616	10	0.1986	0.8331	2920.32	0.0000
GRP per capita growth rates, log	616	8	0.0612	0.2338	257.87	0.0000
Migration, log	616	7	0.0055	0.4535	702.26	0.0000
Investment per capita, log	616	12	0.4440	0.5850	1157.44	0.0000

GRP per capita, log	Coeff.	St. error	*z*-stat.	*p*-value
Investment per capita, log (be)	0.4380	0.0327	13.39	0.000
Investment per capita, log (fe)	0.0661	0.2953	0.22	0.823
Share of extractive and resource-intensive industries, log (be)	0.4005	0.0687	5.83	0.000
Share of extractive and resource-intensive industries, log (fe)	0.2713	0.3420	0.79	0.428
Share of economically active population, log. (be)	1.8576	0.2193	8.47	0.000
Share of economically active population, log. (fe)	3.3464	1.2889	2.60	0.009
Share of employed in economically active population, log (be)	1.1483	0.1244	9.23	0.000
Share of employed in economically active population, log (fe)	3.8457	1.4805	2.60	0.009
Number of postgraduate students per 10,000 residents, log (be)	0.0407	0.0120	3.40	0.001
Population of the largest city, log	0.0602	0.0103	5.85	0.000
Intercept	6.8300	0.3357	20.34	0.000
GRP per capita growth				

rates, log

Investment per capita, log (be)	0.0257	0.0138	1.86	0.063
Investment per capita, log (fe)	0.1298	0.0118	11.03	0.000
Migration, log (be)	1.6119	0.4571	3.53	0.000
Share of extractive and resource-intensive industries, log (be)	0.0174	0.0192	0.90	0.366
Share of extractive and resource-intensive industries, log (fe)	0.3390	0.0699	4.85	0.000
Dummy on availability of a seaport	0.0152	0.0074	2.07	0.038
Number of postgraduate students per 10,000 residents, log (be)	0.0053	0.0041	1.29	0.199
GRP per capita, log (be)	–0.0374	0.0165	–2.27	0.023
Intercept	0.1949	0.0750	2.60	0.009
Migration, log				
GRP per capita growth rates, log (be)	0.1485	0.0213	6.97	0.000
Density of rural population, log (be)	0.0020	0.0007	3.02	0.003
Average temperature of January	0.0001	0.0000	3.40	0.001
Population growth over 1926–1989	–0.0022	0.0002	–9.38	0.000
Passengers departures by railways per capita, log (be)	0.0029	0.0009	3.19	0.001
Population of the largest city, log	0.0017	0.0003	5.35	0.000
Corrupted police administration index, log *	–0.0024	0.0008	–3.03	0.002
Intercept	–0.0074	0.0039	–1.88	0.060
Investment, log				
GRP per capita, log (be)	0.9555	0.0653	14.63	0.000
Average temperature of January	0.0131	0.0033	3.94	0.000
Permafrost dummy (primary)	0.1574	0.0528	2.98	0.003
Permafrost dummy (discontinuous)	0.4423	0.0771	5.74	0.000
Permafrost dummy (complete)	0.4500	0.1087	4.14	0.000
Fuel industry output per capita, log (be)	0.0731	0.0129	5.66	0.000
Fuel industry output per	0.3505	0.0578	6.06	0.000

capita, log (fe)				
Number of stationary phones per 1,000 in 1995, log	0.2781	0.0669	4.16	0.000
Corrupted officials index, log *	0.2573	0.0576	4.47	0.000
Dummy-variable on 10 regions with highest legislative risk**	0.0126	0.0415	0.30	0.762
Dummy-variable on 10 regions with minimal legislative risk **	0.0951	0.0386	2.47	0.014
Dummy-variable on regions with Standard-and-Poor's investment rating	0.1476	0.0399	3.70	0.000
Intercept	—3.5927	0.5860	–6.13	0.000

* *Source*: "OPORA-VTsIOM", 2005; the greater the magnitude of the index, the less corrupted are the regional police administration and the officials.

** *Source*: Standard-and-Poor's investment rating for Russian regions from Expert RA (http://www.raexpert.ru/ratings/regions).

be – between, average over the period.

fe – fixed effect, deviations from the average.

Table A2-2

SEM estimation results, full information maximum likelihood estimator (FIML), 1997–2004, 77 regions FIML estimation of 3-equation SEM:

Model Chisquare = 226.32   Df = 24 Pr(>Chisq) = 0

Chisquare (null model) = 3590.5   Df = 153

Goodness-of-fit index = 0.96392

Adjusted goodness-of-fit index = 0.74291

RMSEA index = 0.11708   90 % CI: (0.10341, 0.13125)

Bentler-Bonnett NFI = 0.93697

Tucker-Lewis NNFI = 0.62479

Bentler CFI = 0.94114

BIC = 2.7895

Normalized Residuals

   Min.   1st Qu.   Median   Mean   3rd Qu.   Max.

–5.42e-01 –3.69e-15 0.00e+00 2.58e-02 0.00e+00 1.16e+00

Iterations = 266

GRP per capita growth rates, log	Coeff.	St. error	z-stat.	p-value
Investment per capita, log	0.0211	0.0121	1.7364	0.0825
Migration, log	0.9571	0.4281	2.2359	0.0254
GRP (corrected) per capita in 1997, log	–0.0301	0.0155	–1.9413	0.0522
Share of extractive and resource-intensive industries, log (fe)	0.1191	0.0594	2.0056	0.0449
Dummy on availability of a seaport	0.0127	0.0063	2.0227	0.0431
Number of postgraduate students per 10,000 residents, log (be)	0.0060	0.0032	1.8910	0.0586
Spatial lag on the log GRP per capita growth rates	0.7695	0.0443	17.3879	0.0000
Intercept	0.1523	0.0745	2.0452	0.0408
Migration, log				
Population revenues in 1997, log	0.0025	0.0011	2.3119	0.0208
Density of rural population in 1997, log	0.0041	0.0008	4.8879	0.0000
Average growth of January	0.0002	0.0000	4.6025	0.0000
Population growth over 1926–1989, log	–0.0022	0.0003	–8.5162	0.0000
Passengers departures by railways per capita, log (be)	0.0028	0.0009	3.1298	0.0017
Population of the largest city, log	0.0022	0.0004	5.8882	0.0000
Intercept	–0.0141	0.0021	–6.6663	0.0000
Investment per capita, log				
GRP (corrected) per capita in 1997, log	0.8497	0.0618	13.7482	0.0000
Average growth of January	0.0054	0.0037	1.4683	0.1420
Permafrost dummy	0.1504	0.0671	2.2417	0.0250
Fuel industry output per capita, log	0.1364	0.0143	9.5138	0.0000
Number of stationary phones per 1,000 in 1995, log	0.4212	0.0789	5.3366	0.0000
Intercept	–2.9651	0.5063	–5.8570	0.0000

Relative impacts on dependent variables of corresponding equations and on GRP growth

Factors grouped by dependent variables (equations)	3SLS estimation (BE)		FIML estimation (pool)	
	Impact on dependent variable of equation, %*	Impact on GRP growth, %**	Impact on dependent variable of equation, %*	Impact on GRP growth, %**
GRP per capita growth rates, log				
Investment per capita, log (be)	91.3		20.8	
Migration, log (be)	60.6		10.1	
Share of extractive and resource-intensive industries, log (be)			6.3	6.3
Dummy on availability of a seaport	31.7	31.7	6.1	6.1
Number of postgraduate students per 10,000 residents, log (be)			6.0	6.0
GRP per capita, log (be)	–100.6		–17.8	–17.8
Spatial lag on the log GRP per capita growth rates			57.7	57.7
Migration, log				
GRP per capita growth rates, log (be)	32.2	19.5		
Population revenues, log			9.2	0.9
Density of rural population, log (be)	16.2	9.8	22.7	2.3
Average growth of January	14.3	8.7	21.6	2.2
Population growth over 1926 – 1989, log	–31.7	–19.2	–28.5	–2.9
Passengers departures by railways per capita, log (be)	11.8	7.2	11.8	1.2
Population of the largest city, log	25.4	15.4	25.9	2.6
Investment per capita, log				
GRP per capita, log (be)	71.7	65.4	50.9	10.6
Average growth of January	13.0	11.8	6.3	1.3
Permafrost dummy (primary)	6.8	6.2		
Permafrost dummy (discontinuous)	17.1	15.6		
Permafrost dummy (complete)	10.5	9.5		

	Col 1	Col 2	Col 3	Col 4
Permafrost dummy (all types)			9.1	1.9
Fuel industry output per capita, log (be)	17.5	16.0	29.9	6.2
Number of stationary phones per 1,000 in 1995, log	11.3	10.3	17.2	3.6
Dummy-variable on 10 regions with highest legislative risk	–0.6	–0.5		
Dummy-variable on 10 regions with minimal legislative risk	4.0	3.7		
Dummy-variable on regions with Standard-and-Poor's investment rating	11.8	10.8		

* The standard deviation of corresponding variable times the estimated coefficient, % of the standard deviation of the dependent variable of a given equation.

** The standard deviation of exogenous variable times the corresponding coefficients, % of the standard deviation of GRP growth rates (computed as impact of the corresponding variable in migration equation or investment equation (previous column) times corresponding impacts of migration or investment on growth rate of GRP).

Appendix 3. Maps

List of the Russian Federation Regions (see map 17)

1	Adigeya Rep.	31	Belgorodskaya oblast	61	Rostovskaya oblast
2	Bashkortostan Rep.	32	Bryanskaya oblast	62	Ryazanskaya oblast
3	Buryatiya Rep.	33	Vladimirskaya oblast	63	Samarskaya oblast
4	Altai Rep.	34	Volgogradskaya oblast	64	Saratovskaya oblast
5	Dagestan Rep.	35	Vologodskaya oblast	65	Sakhalinskaya oblast
6	Ingushetitya Rep.	36	Voronezhskaya oblast	66	Sverdlovskaya oblast
7	Rep. Kabardino-Balkariya	37	Ivanovskaya oblast	67	Smolenskaya oblast
8	Kalmikiya Rep.	38	Irkutskaya oblast	68	Tambovskaya oblast
9	Karachaevo-Cherkesia Rep.	39	Kaliningradskaya oblast	69	Tverskaya oblast
10	Karelia Rep.	40	Kaluzhskaya oblast	70	Tomskaya oblast
11	Komi Rep.	41	Kamchatskaya oblast	71	Tulskaya oblast
12	Maryi-El Rep.	42	Kemeroskaya oblast	72	Tumenskaya oblast
13	Mordovia Rep.	43	Kirovskaya oblast	73	Ulyanovskaya oblast
14	Sakha (Yakutia) Rep.	44	Kostromskaya oblast	74	Chelyabinskaya oblast
15	North Osetia Rep.	45	Kurganskaya oblast	75	Chitinskaya oblast
16	Tatarstan Rep.	46	Kurskaya oblast	76	Yaroslavskaya oblast
17	Tuva Rep.	47	Leningradskaya oblast	77	Moscow City
18	Udmurtiya Rep.	48	Lipetskaya oblast	78	St-Petersburg City
19	Khakassia Rep.	49	Magadanskaya oblast	79	Yevreiskaya Autonomus Obl.
20	Chechen Rep.	50	Moskovskaya oblast	80	Aginskiy Buryatskiy AO*
21	Chuvash Rep.	51	Murmanskaya oblast	81	Komi-Permyatskiy AO
22	Altai Krai	52	Nizhegorodskaya oblast	82	Koryakskiy AO
23	Krasnodar krai	53	Novgorodskaya oblast	83	Nenetskiy AO
24	Krasnoyarsk krai	54	Novosibirskaya oblast	84	Ddlgano-Nenetskiy AO
25	Primorskiy Krai	55	Omskaya oblast	85	Ust-Ordinskiy Buryatskiy AO
26	Stavropol Krai	56	Orenburgskaya oblast	86	Khanti-Mansiyskiy AO
27	Khabarovsk Krai	57	Orlovskaya oblast	87	Chukotskiy AO
28	Amurskaya oblast	58	Penzenskaya oblast	88	Evenkiyskiy AO
29	Arkhangelskaya oblast	59	Permskaya oblast	89	Yamalo-Nenetskiy AO
30	Astrakhanskaya oblast	60	Pskovskaya oblast		

* – AO = autonomous okrug (district).

Chapter 2. Canadian Regional Economic Well-being: a spatial analysis of their differences

Introduction

One of the unique aspects of economic well being and economic growth in a geographically large country like Canada or Russia is that variation tends to be large. As a result, if a federal government has a mandate to provide equal well being for its citizens in all of its regions, it will have to find means to counter the natural forces that create differences like resource industries and create an economic geographic environment that improves the chances of creating more equal economic well being over space. Some of the ways Canada tried to do this in the 1930–1980 period was through subsidization of transport costs, industrial location incentives, labour incentives, special unemployment benefits and, of late, special peripheral resource taxation benefits. One of the more recent regional development endeavours has been to encourage local development based on the local available natural and man endowed resources. The Canadian federal and provincial governments have also decentralised or peripheralised government services locations to help local growth. This process has apparently been quite successful in increasing the local economic well being and improving the local economic multiplier.

What causing economic development in any one region is difficult or next to impossible to say. A number of theories and models have been proposed by academics and practitioners of regional economic well being. Without going into details let me just mention 16 of the ones that might be referred to as the classical ones:

- Growth Pole Theory;
- Staple Growth Theory;
- Comparative Advantage Theory;
- Core-Periphery Theory;
- Myrdal's Cumulative Causation Theory;
- Hirschman's Backward and Forward Linkage Theory;
- Heckscher-Ohlin Trade Theory;

- Economic Base Multiplier Theory;
- Kondratieff Wave Theory;
- Export Base Theory;
- Schumpeterian Growth Theory;
- Productivity Growth Theory;
- Dependency Theory;
- Clark-Fischer Development Model;
- Input Output Model;
- Rostow's Stage Growth Model[50].

The large number of development models and theories suggest that different ones might apply in different regions of the world and in different market systems. Furthermore not all are mutually exclusive. Dependent on one's perspective, different theories and models can be used to explain the total Canadian regional development scene or just regional aspects of it. Early development of Canada and present Alberta as well as many of the northern regions may be explained by the Staple Growth Theory. Southern Ontario and Quebec early growth may be better accounted for by Cumulative Causation and Forward and Backward Linkage or even Growth Pole Theory. Present growth in these regions may be better explained by Export base Theory.

In this paper an attempt will be made to present the variation in the regional population growth between 1991 and 2001 of the 10 Canadian provinces. It is hypothesised that the population growth rate in a given region is probably the best indicator of development prospects in the region as well as an overall attractive well-being situation. Secondly, an attempt will be made to explain, with the help of a multiple regression

[50] No one reference will cover and discuss all of these. A start can be found in J. O. Wheeler's et all book Economic Geography, Wiley, 1996. Readers will also find Peter Dicken and Peter E. Lloyd's classical text Location in Space: Theoretical Perspectives in Economic Geography, Harper, 1990 quite useful. Others of interest are: Paul Krugman, Development, Geography, and Economic Theory, MIT Press, 1995, Klaus Grimm, Theorien der Unterentwicklung und Entwicklungstrategien, Westdeutscher Verlag, 1997, James D. Cockcroft, Andre Gunther Frank and Dale L. Johnson, Dependence and Underdevelopment, Anchor Books, 1972, Paul Knox and John Agnew, The Geography of the World Economy , Edward Arnold, 1989 and L. Needleman (ed) Regional Analysis , Penguin, 1968.

model, the variation in the 293 Canadian Census Divisions using empirical data derived from the 2001 Canadian Census. These divisions are thought to have less intern variation then what is found between them and also more detailed and precise information then what is available for the provinces. The explanatory focus will be on different socio-economic conditions in Canada and related them to potential regional growth causes. In an earlier paper the role of universities as growth poles in Canada was specifically examined (*Meyer, S. and A. Hecht (1996)*).

The Canadian Setting

Just like Russia, Canada is a huge country but with relatively few people and tremendous differences in population densities. It also has many natural resources and the exports of these were very important in the course of Canada's development. They still are very important for export purposes now but are far less significant in the creation of employment now. Canada is also highly urbanized with over 76% of the population living in urban places with about 40% of these residing in the three largest metropolitan regions[51] of Toronto, Montreal and Vancouver. Clearly these centres act as growth poles for the country especially for international immigrants.

The east west dimensions of Canada measure some 5 500 km and the north south dimension are about 4600 km (*Fig. 2.1*). It is not hard to imagine that overcoming such friction of space is expensive whether it is economically, culturally or socially. It is hence not a surprise that a large part of Canada's interactions take place north south, with the US, instead of east west within Canada. Because of the harsh climatic conditions in the North, settlement is restricted to a relative narrow land belt north of the US border. It is therefore not surprising that growth in the US, near the Canadian border, also has a strong influence on growth on the Canadian side.

Little economic activity, relatively speaking, takes place in the North. In addition to that the climatic conditions here are not conducive to an ease of interactions. The fact that more then 40 percent of Canada is

---

[51] http://info.wlu.ca/%7Ewwwgeog/special/vgt/English/can_mod1/unit1.htm.

covered by permafrost, and most of it is found in the three northern territories and in the northern parts of a number of provinces, restricts Canadian interactions overall.

As can be seen Canada is organised politically into 10 provinces and 3 territories. The 10 provinces have equal rights and responsibilities in the Canadian federation which started in 1867. The territories have more limited control and responsibility and come under the operational and administrative responsibility of the Federal government located in Ottawa.
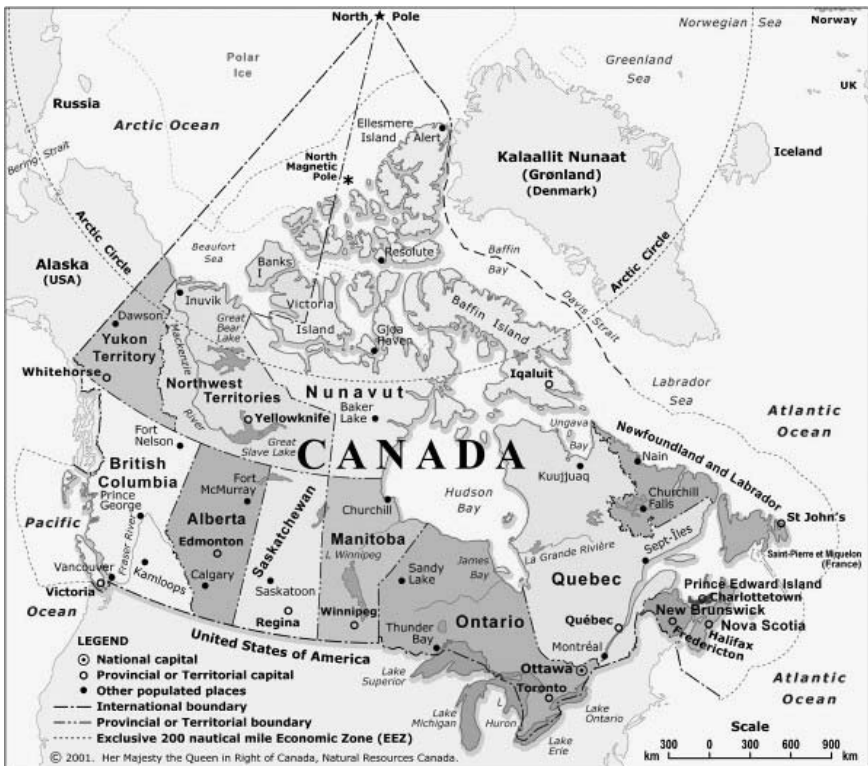
Source: http://atlas.nrcan.gc.ca/site/english/maps/reference/national/can_political_e.

Fig. 2.1. Provincial and Territorial Map of Canada

The Canadian economy of the 21st century is frequently labelled as a post industrial economy. As can be seen from table 1, 68.4% of its GDP in 2005 came from the non goods producing, or broadly speaking, the service sector. Manufacturing contributed only 17.2% and the resources sectors only 6% to the GDP value of the economy. When one looks at employment in the various industrial sectors as shown in table 2, one can see that the service sector is even stronger accounting for 75.2% of the employment. The resource extraction sector, which includes agriculture, directly employs only 4% of total employment. The fact that a smaller percentage of total direct employment in these sectors produces a higher proportion of the GDP clearly points to the higher productivity in these sectors, in large part probably due to higher capital needs then in the service sector.

Table 2.1

Canadian GDP at Basic Prices by Industry, CDN $, 2005

Industrial Sectors:		$ (000,000)	%
All		1,079,142	100
	Agriculture, Forestry, Fishing, and hunting	24,021	2.2
	Mining and oil and Gas extraction	41,353	3.8
	Manufacturing	185,140	17.2
	Construction Industries	63,345	5.9
	Utilities	27,528	2.6
	All Others	737,755	68.4

Source: http://www40.statcan.ca/l01/cst01/econ41.htm.

Table 2.2

Canadian Employment by Industries, 2005

Industrial Sectors:		Emp. in 000	%
All		16,169.7	100
	Goods producing sector	4,002.4	24.8
	Agriculture	343.7	2.1
	Forestry, Fishing, Mining, Oil and Gas	306.4	1.9
	Manufacturing	2,207.4	13.7

Construction Industries	1,019.5	6.3
Utilities	125.3	0.8
All Others	12,167.3	75.2

Source: http://www40.statcan.ca/l01/cst01/econ40.htm.

That the relative low employment in manufacturing has not always been the case in Canada and in the other developed countries can be seen from *Fig. 2.2*. It shows the change of manufacturing employment over time in the major economies of the western world. The Canadian graph is superimposed. It shows the rather strong decline in manufacturing employment in these economies since the end of the 1960s. Although some claim that this decline is due to outsourcing of the manufacturing sector to the developing world, an IMF report claim this is not the case but increased productivity in the sector is the cause. They show that over the 1960 to 1994 period, manufacturing productivity increased on average 3 % per year (*IMF (1996)*). Since population increases were at best about 1% per year, increased output per person increases greatly overall output, which means that far fewer employees are needed each year in this sector. A similar productivity increase story can be told for the resources sectors.
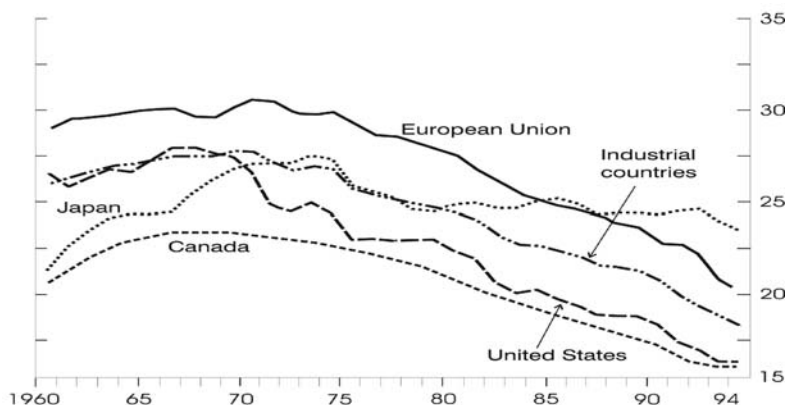
Source: IMF, 1996, Chart 18, p.47.

Fig. 2.2. Manufacturing Employment as a Percent of Total Employment

108

Canadian Development and Regional Well being

It is frequently hypothesized that overall Canadian development is due to the export of resources over time. This theory is known as the Staple Growth Theory, and is presented in graph form in *Fig. 2.3*. Over time Canada has relied very much on the export of raw materials mainly to Europe and the USA. The kind of resources exported has changed over time from fish to furs to lumber to wheat to forest products to minerals and presently to energy. What also changed substantially over time was the place from where in Canada most of these resources came. It started in the east coast and moved inland and westward as time went on. These changes in the spatial importance of resources over time have left a tremendous impact on the local economies which they left and the new economies where they moved to, causing an economic well being downturn and an upturn in the respected areas. As the economy changed in output and in location it left abandoned infrastructure, labour expert and wealth generating institutions behind. The impact of this process can clearly be seen in the present spatial economy. Some regions with increasing important aspects of the economy flourish while others decline. Since Canadian workers are rather mobile, many leave declining regions and move to growing regions causing substantial changes in the settlement pattern and in the relative importance of different regional economies. To keep the declining region in relative good economic health so that the population can receive its basic expected services, the federal government has instituted substantial transfer payments form the 'have' to the 'have not' provinces.

Besides the changing regional structure of the resource economy, the overall significance of the export of resource has substantially declined drastically between 1960 and 2000 as can be seen in *Fig. 2.4* (a self constructed diagram). Clearly Canada has become a more 'advanced developed' country then it was in 1960. In 2005 by far the most important export sector was manufactured goods. Most of these were exported by the auto industry where Canada has had a free trade agreement with the US for some time.

The Staple Growth Theory Revisited ;

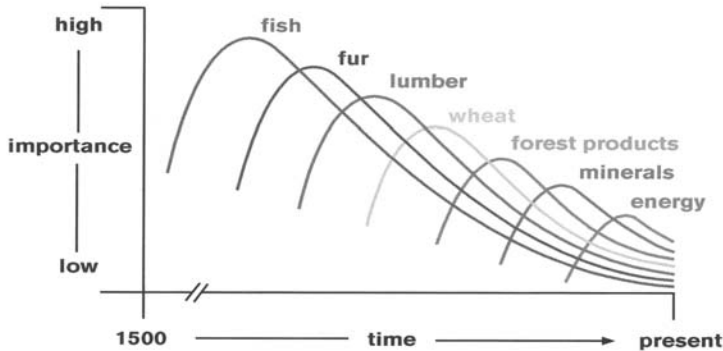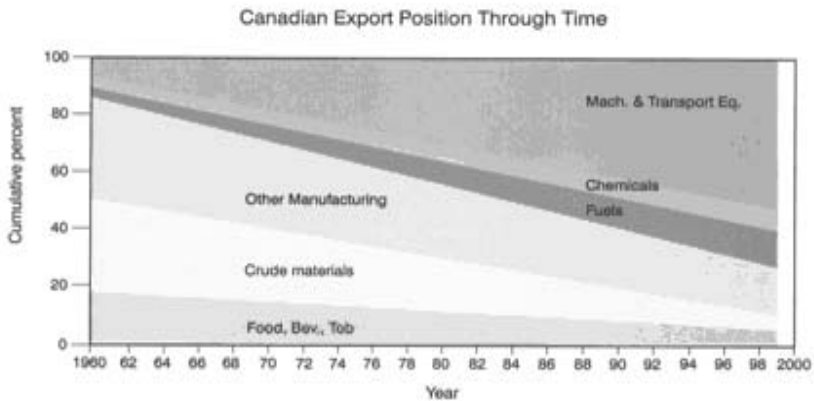the decline in importance of natural resources to the Canadian economy

Fig. 2.3. The Canadian Regional Growth Factors Over Time

Canadian Export Position Through Time

Fig. 2.4. Major Canadian Export Sectors Over Time

What the present situation is can be derived from *Table 2.3*. It shows the proportion of Canadian export in 2005 by major products. In this year the export of all resources; agriculture, fishing, forestry, energy etc accounted for only 34 % of total value of exports. More then half of this came from Energy, in part due to high oil prices in 2005.

Table 2.3

Canadian Exports by Products, 2005

Product Sectors:	$ in 000,000	%
All	453,060.1	100
Agricultural and Fishing	30,217.6	6.7
Energy	86,924.1	19.2
Forestry	36,606.3	8.08
All Others (mainly Manufacturing)	299312.1	66.06

Source: http://www40.statcan.ca/l01/cst01/gblec04.htm.
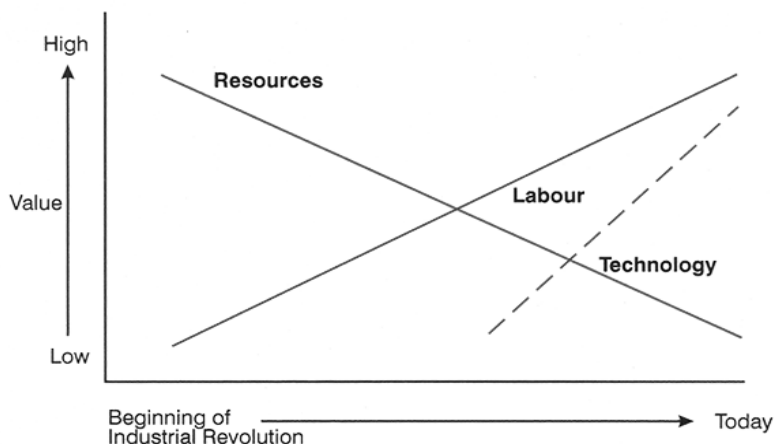
Fig. 2.5. Changing Resource, Labor and Technology Prices over Time

But this shift between economic sectors in Canada is not unique. The world goods production system has seen a major shift in the importance

of resources, labor, capital and technology. As seen in *Fig. 2.5*, a personally constructed figure, the world has experienced a major decline of the value of resources in final products and a corresponding value increase in labor and technology.

But as was alluded to earlier the Canadian economy has shifted from a resource based economy to a modern service based one. This means that the present well being and growth is far more diversified in its cause then the earlier resource economy. A broad set of variable now influence the well being of regions. Our discussion will now turn to describe, examine and analyse these. The reader should be aware that only those variables that are in the final multiple Canada wide regression analysis will be described here.

Provincial Variation in Well Being

Similar to large variation in population size of Russia's 89 regions, the differences between Canadian provinces are tremendous (*Fig. 2.6*). Ontario has nearly 12 Million inhabitants while Prince Edward Island has about 100,000. If one would add the northern territories this variation would be even bigger by a factor of two. The consequence is that a policy for one is not always appropriate for the others. Secondly, except for the three maritime provinces of Nova Scotia, Prince Edward Island and New Brunswick the other provinces tend to have a relative densely settled southern region and a vast lightly settled northern part.
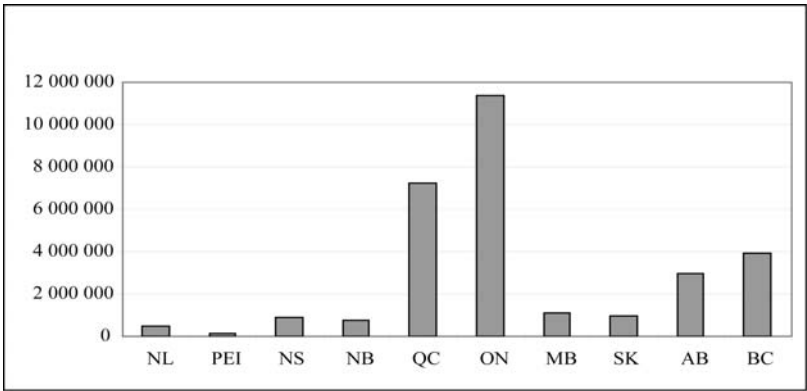

Fig. 2.6. Population of Canadian Provinces, 2001

If one looks at the population growth of the provinces between 1991 and 2001, the last decade between the full censuses, one also finds tremendous differences (*Fig. 2.7*). Clearly only three provinces have had substantial growth, Alberta, Ontario, and British Columbia. Nearly all other provinces had no growth except for Newfoundland and Labrador which even had a decline of 7.0 percent.
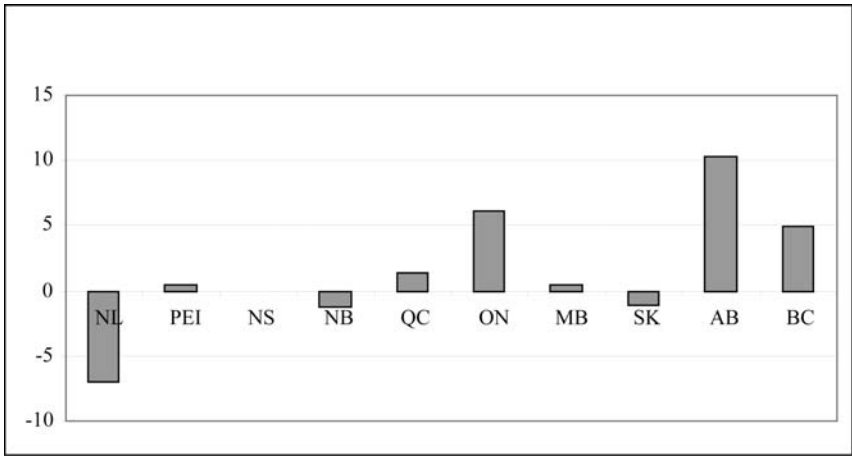

Fig. 2.7. Provincial Population Changes, 1991–2001

If a region grows a lot one would expect that the proportion of younger people is high. *Fig. 2.8* collaborates this very nicely for the Canadian provinces. Except for BC, as one moves westward from the Atlantic provinces the median age of the provinces declines. This corresponds with the provincial growth picture seen in *Fig. 2.7*. The relative high value for BC comes from the fact that BC also is the destination province of many retired people for Canada as a whole.
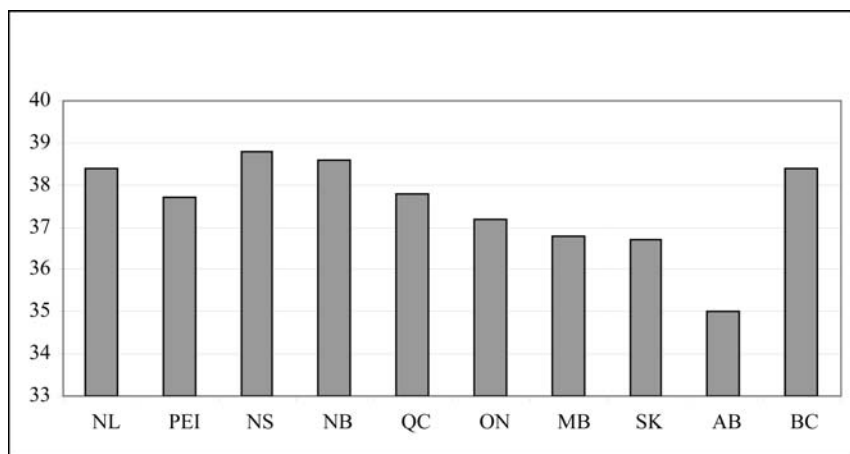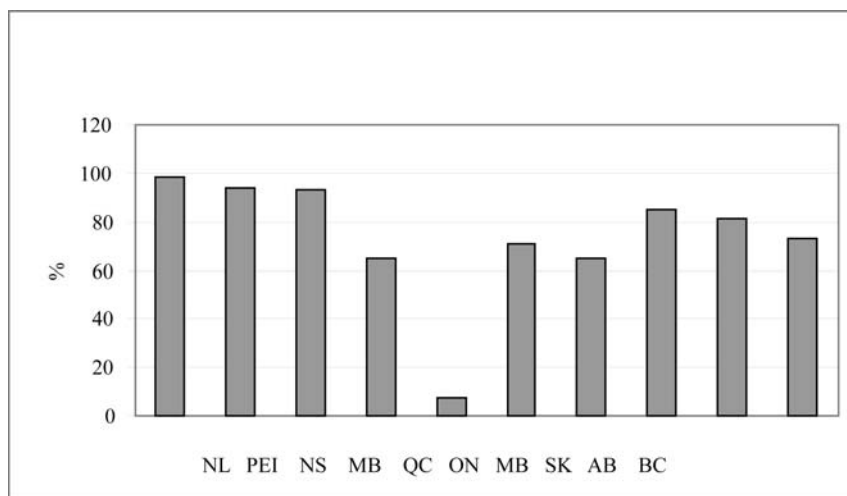
Fig. 2.8. Median Age by Provinces

Fig. 2.9. Native English Speakers

Compared to the Atlantic Provinces, the provinces from Ontario west have relative low proportions of people with English as their mother tongue. The low value in Quebec and New Brunswick comes from the fact that a high proportion of the population here had high percentages French as their mother tongue, 80.9 and 32.9 respectively.

There seems to be an inverse relationship between the presence of English and French as mother tongues and population growth.

If one looks at the proportion of people that had not changed their residence between 1996 and 2001 one can again see a spatial regular picture (*Fig. 2.10*). The East of Canada has higher proportion of non movers then the rest of the country. Alberta in contrast has the lowest. In many instances if people want to improve their well being they have to move. One could also argue that if the conditions of the people improve they will seek better accommodation which results in a change in residence.
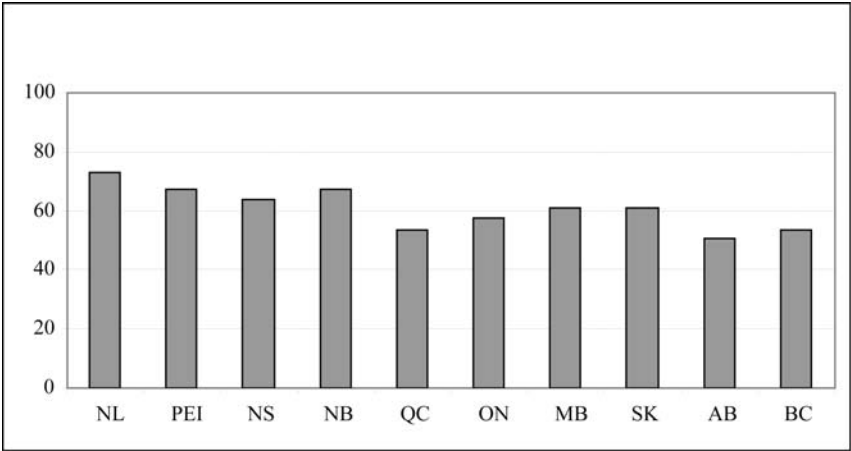

Fig. 2.10. Non Movers in the Last Five Years

The percentage of people with some income who work full time shows a nearly normal picture over the east west spatial dimensions of Canada (*Fig. 1.11*). It is low in the east then increases as one moves westward, reaching a peak in Ontario and slowly declines again to BC. Although the values are relatively high in the west, it is nevertheless a

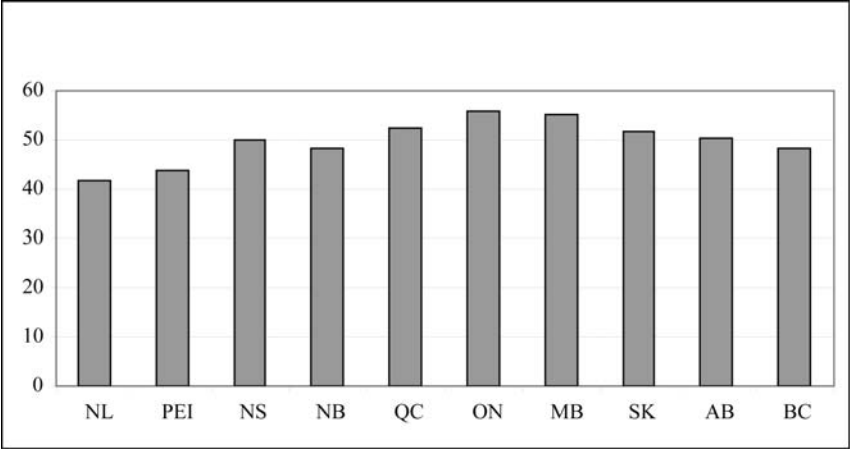surprise that the values in AB and BC are lower then the four central provinces.

Fig. 2.11. Individuals with some Income that are Working Full Time
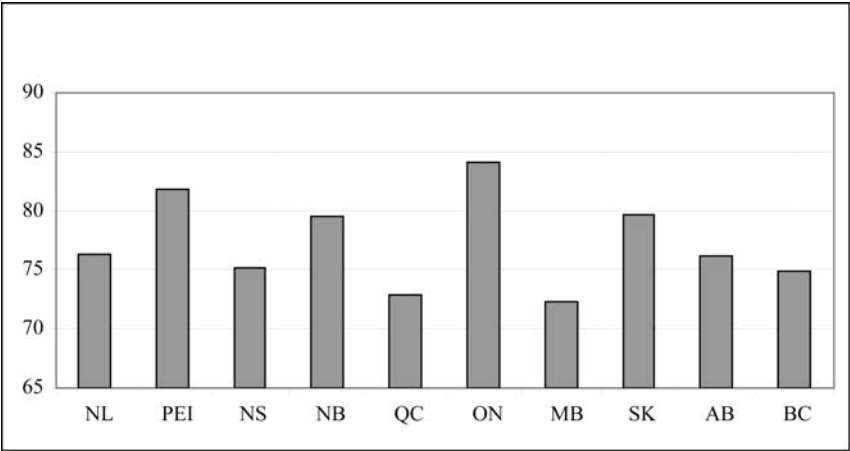

Fig. 2.12. People Going to Work by Car

One of the questions Statistics Canada asks is how an individual gets to work. One possible response is 'by car'. By far the greatest proportions of Canadians go to work by car. It is highest in Ontario, about 84% and lowest in Manitoba about 73%. There seem to be no clear spatial pattern in *Fig. 2.12*.

Fig. 2.13 shows the employment rate of the labour force in each province. From Ontario west the values are higher then from Quebec east. It is especially high for Alberta. Since 2001 economic conditions in Alberta have even proved substantially. The result is that the services sector has a major problem attracting labour and businesses have to close because of the lack of workers. (*Globe and Mail, August 21, 2006, p.1*)
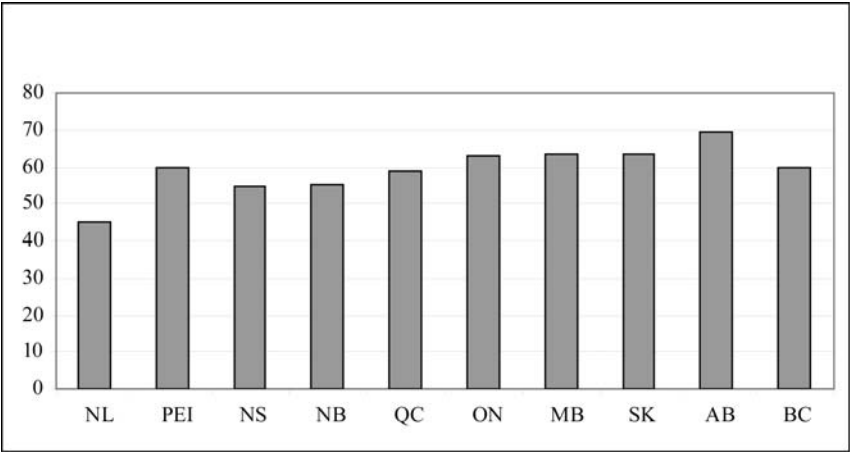

Fig. 2.13. Labour Force Employment Rates

The Canadian Census also asks a person about his or her income for the previous year. The responses are collapsed into three categories; earned income, government source and other income. In *Fig. 2.14* is given the proportion of earned income as part of total.

Alberta's percentage is by far the highest followed by Ontario, British Columbia, Manitoba and Quebec. A clear relationship exists between these values and growth of population in these provinces.
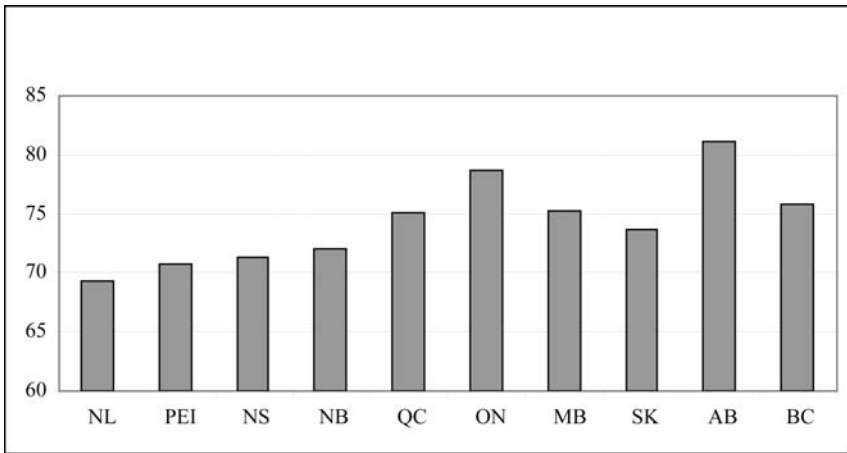
Fig. 2.14. Earned Income as Part of Total

Although the reasons why a region may have a large number of one person households can vary, one reason clearly is that single individuals are attracted to growing areas with jobs. This seems to be the case for Canadian provinces in 2001 (*Fig. 2.15*). The western provinces share is relatively high. The high value in Quebec may be more of a function of its language and culture and it 'hip' live style especially in Montreal and Quebec City then with anything else. Why Alberta is relatively low may be due to the fact that few elderly people live here. As people age a partner may die and the other person then becomes a single member household.
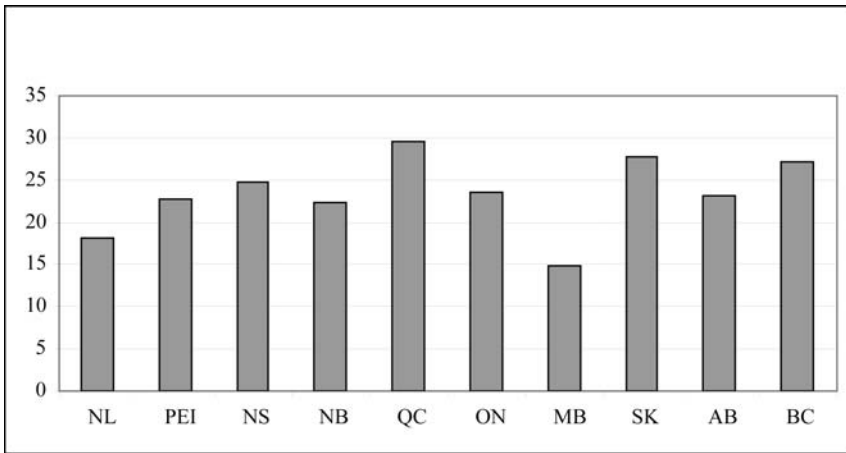
Fig. 2.15. Proportion of People Living in a one Person Household, 2001

As a whole in the English speaking provinces, the proportion of people who own their own dwelling decreases as own moves from east to west. It probably reflects more the price of homes then anything else. Prices are lower in the east then anywhere else in Canada. In the east are also fewer movers as was shown earlier, which means they have had a longer time to purchase their own dwelling. The low value for Quebec again relates to a different value structure towards a dwelling. A house is seemingly not valued as highly as in the rest of Canada and renting is quite an acceptable life style.

Many Canadians see the ownership of their house as a highly desirable goal. In fact many home mortgages are not paid for till the person retires. At this time the life's savings are frequently tied up in the house and, until relatively recently, it was the only major asset an individual had to live off in his/her retirement, after the house was sold, besides the meagre government pension.
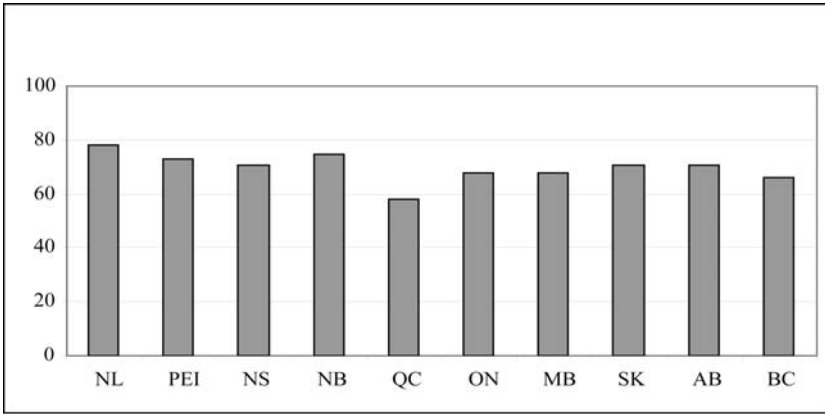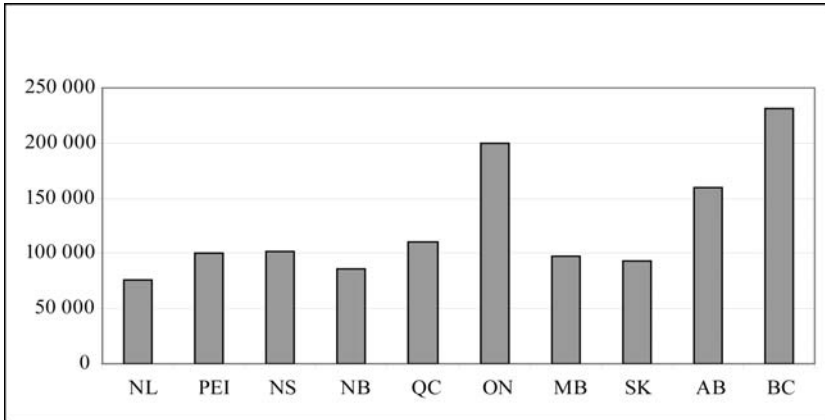
Fig. 2.16. Proportion of Owned Dwellings

Fig. 2.17. Average Value of Owned Dwelling

But as can be seen from *Fig. 2.17* the average price of houses various tremendously between the provinces. By far the highest prices are found in the three major 1991–2001 growth provinces, Ontario, Alberta and British Columbia. Although the provincial data hides it, it is the major cities in these provinces that have the high house prices. Although it is a great challenge for new comers to these provinces to acquire a house,

120

individuals who have lived there for some time and bought houses earlier are now inheriting geographic capital. By the mere fact of living in a particular region before this region experienced boom in house prices, provides them with a 'windfall' of assets which have not been worked for. It is a geographic capital inheritance (*Hecht (1998)*).
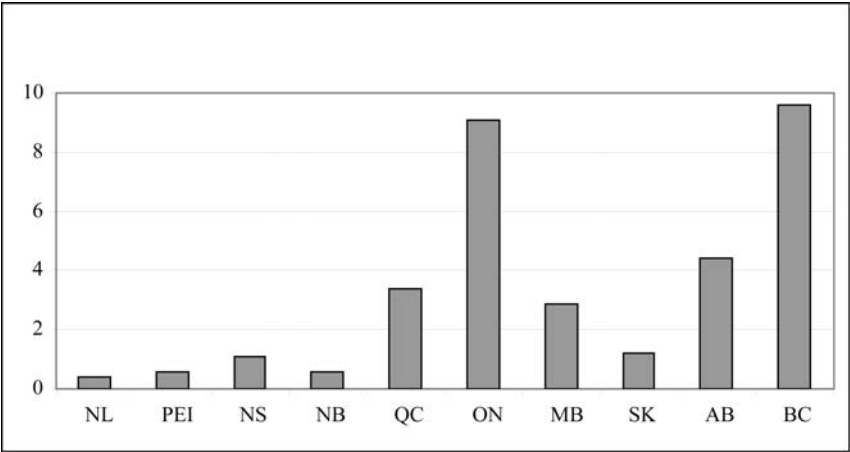

Fig. 2.18. Proportion of Immigrant Population

In Canada immigrants are an important source of population growth. In *Fig. 2.18* the proportion of population in each province that had migrated to Canada during the 1991 to 2001 period is given. The picture is to a large extent similar to the population growth picture. Recent immigrants are highest in the provinces that have had substantial growth in population in the 1991–2001 period namely BC, AB, ON and to a less extend in QC and MB. Immigrants have to state a province that they want to reside in when applying to come to Canada. But after landing in Canada immigrants do not have to stay in that province. In some instances it is easer to migrate to Canada if immigrants are willing to go to a province that wants or needs immigrants like Quebec. But once in Canada, except for the entrepreneurial class of immigrants, they do not have to stay in that province. Freedom to choose where to live, once in Canada, is a highly valued personal right in Canada.

*Analysis of Population Changes in Canada
at the Census Division Level*

One of the major characteristic of a free market economy is that the various economic and social characteristics can change in respond to changes in well being of the regions and the people in them. In Canada people tend to live where they think conditions are best for them. Free mobility is a great privilege. The result of this freedom of movement is that some regions grow in population while others decline. In general terms, people first settled in eastern Canada but the centre of population has been moving west ever since. Although there are major forces during a person's life cycle as to where to live, the most important one probably is where there is an appropriate job for that individual.

In this section of this study certain socio-economic variables from the 293 Canadian Census divisions (*Fig. 2.19*) will be examined to see which and how do they related to the changes in population growth in these regions in the 1991 to 2001. Data was collected from the Canadian Statistics, Community Profile files[52]. These files are made available on the internet by Statistic Canada but must we extracted, examined and modified in order to be used in any analysis. In all some 45 variables where extracted that were thought to relate to economic development in Canada. The list is given in *Table 2.8*. Each of the numeric variables is an absolute number, a percentage or a dummy variable. They measure part of the labour, income, migration, language, ethnicity, demographic, transport, education, occupations, religion and housing aspects. It was hypothesised that specific variables of these general topics would relate to economic well being as measured by the 1991–2001 population increase in the 293 Census division of Canada. The measures were taken in the late spring of 2001 and refer to conditions at that time or in 2000. Most of the values are derived from a 20% sample (long form of the census) of the Canadian population but rounded up to 100% of the population. A few values such as the absolute population counts are derived from the 100% survey (short form of the census).

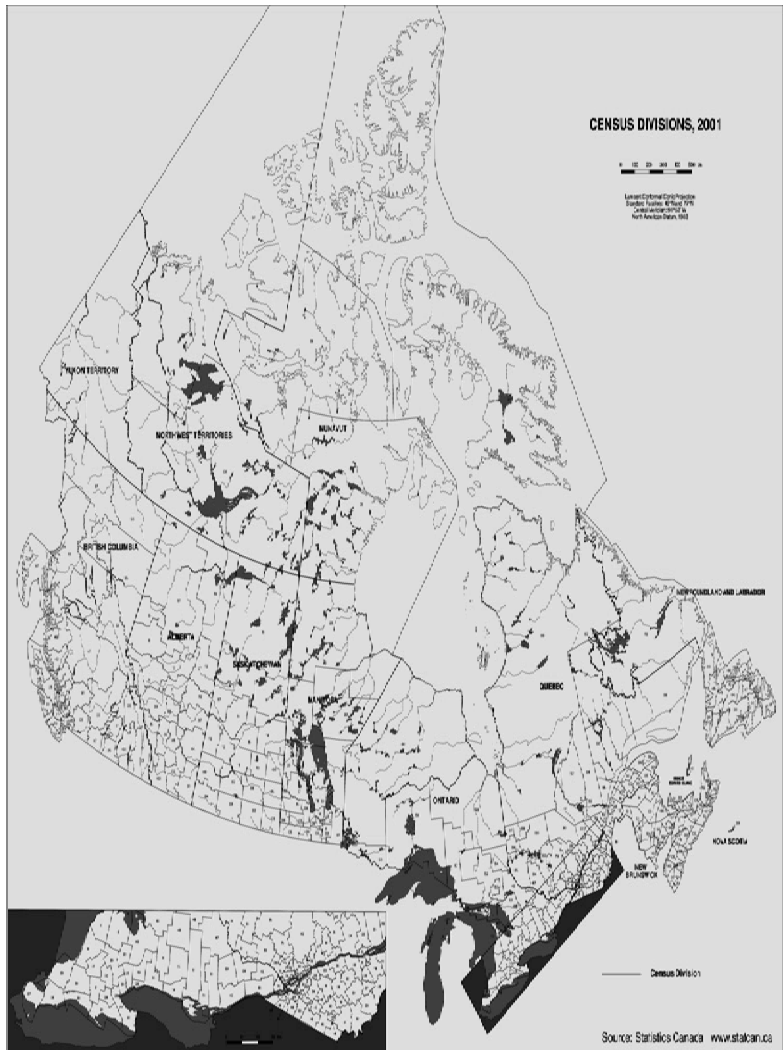[52] See http://www12.statcan.ca/english/profil01/CP01/Index.cfm?Lang=E.

Fig. 2.19. 2001 Canadian Census Divisions

Source: Statistics Canada, http://geodepot.statcan.ca/Diss/Maps/ReferenceMaps/ retrieve_ cd.cfm.

In this paper it is hypothesised that the best indicator of development is the percent growth in population between 1991 and 2001. In a free marker economy people tend to move to place where things are better then were they are. These moves could be international migrants or intra- and inter-provincial migrations. The data however also capture the differences between birth and death. But these rates should not vary substantially over space. The change in population between the 1991 and 2001 census was used as the y variable (dependent). All of other variable listed in *Table 2.8* where entered into a stepwise multiple regression equation as explanatory variables. Some variables were transformed by a log or square root transformation function to create better normality but none of these transformed variables was important enough to enter the equation.

Table 2.8

Definitions of Variables for Canadian Census Divisions, 2001

v1	Ptcode	Provincial or Territorial code
v2	Name	Name of census division
v3	pop01	2001 population
v4	p91-01	Population increase 1991–2001 in %
v5	univ.	University present; 1-yes, 0-no
v6	mage	Median age in the community
v7	emt%	Language first learned and still understood, % English
v8	fmt%	Language first learned and still understood, % French
v9	nm<1y	% of population, one year or older, that did not move in the previous year
v10	m<1y	% of population, 1 year and older, that lived in a different province or country 1 year ago
v11	nm<5y	% of population 5 years or older that did not move in the last 5 years
v12	m<5y	% of population 5 years and older, that lived in a different province or country 5 years ago
v13	i%	1991 to 2001 Census Division in-migration as a % of total population
v14	vm%	% of visible minorities in the community
v15	sft20	Population 20–24 years of age studying full time (%)
v16	d20	Population 20–34 years old with a university certificate, diploma or degree

124

v17	d35	Population 35-44 years old with a university certificate, diploma or degree
v18	d45	Population 45-64 years old with a university certificate, diploma or degree
v19	%ftw	% of population that have some earnings and are working full time
v20	awai$	Average income of all workers
v21	ftwai$	Average income of full time workers only
v22	ef	Language used most often at work – English and/or French (as %)
v23	wup	Worked at usual place (%)
v24	ct	Car transportation to work in %
v25	lfpr	Labour force participation rate– (%)
v26	lfer	Labour force employment rate– (%)
v27	lfur	Labour force unemployment rate (%)
v28	elm	Experienced labour force in manufacturing industries– (%)
v29	elfr	Experienced labour force in fire and real estate industries– (%)
v30	eleh	Experienced labour force in education and health industries– (%)
v31	elbs	Experienced labour force in business services industries– (%)
v32	om	Management occupations (%)
v33	obfa	Business, finance and administration occupations (%)
v34	onas	Natural and applied sciences occupations (%)
v35	oscegr	Social science, education, government service and religion occupations (%)
v36	m$	Median total income of persons 15 years of age and over
v37	ei%	Earned income as % of total
v38	gti%	Income from government transfers as % of total income
v39	oi%	Other money as % of total income
v40	mcfi$	Median income ($) of couple families
v41	oph	% of one-person households
v42	pod	% of owned dwellings
v43	avh$	Average value of owned dwelling ($)
v44	rp%	% of population that are protestants
v45	nra%	% of people with no religious affiliation

Table 2.9

Stepwise Multiple Regression Analysis, Canadian
Census Divisions, 2001

	Un-standardized		Standard-ized	*t*-values	Sig. values
	Coefficients		**Coefficients**		
	B	**Std. Error**	**Beta**		
(Constant)	59.533	6.5542		9.0832	1.90E–17
V43	0.000067	0.00000644	0.5347	10.4095	1.10E–21
V19	0.1456	0.0381	0.1895	3.8216	1.63E–04
V6	–0.8518	0.1095	–0.6126	–7.7766	1.41E–13
V24	0.1057	0.0297	0.1809	3.5555	4.42E–04
V37	–0.6363	0.0626	–0.7547	–10.1594	7.24E–21
V26	0.2394	0.0498	0.3672	4.8091	2.47E–06
V41	–0.2257	0.0674	–0.175	–3.3462	9.30E–04
V11	–0.2559	0.0426	–0.3916	–6.0031	5.94E–09
V7	–0.0659	0.0061	–0.4532	–10.7877	6.08E–23
V42	0.1231	0.0431	0.2287	2.8565	4.60E–03

Dependent Variable: v4

Model Summary

R	Adjusted R Square	Std. Error of Estimate	df2	*F*-value	Sig. Value
0.8652	0.7396	2.8999	282	83.99	0

Variables in the Multiple Regression Equation for Canadian CD's, 2001

V4	Population increase 1991–2001 in %
V6	Median age in the community
V7	Language: English mother tongue as % of total population
V11	% of population 5 years or older that did not move in the last 5 years
V19	% of population that have some earnings and are working full time
V24	% of population that have some earnings and are working full time
V26	Labour force employment rate %
V37	Earned income as % of total
V41	% of one-person households
V42	% of owned dwellings
V43	Average value of owned dwelling ($)

In *Table 2.9* we can see the result of the multiple regression analysis. Overall 10 variables entered the equation as significant variables through a stepwise multiple regression equation. These in total produced a multiple R square (explained variation in y) of .7396, i.e 73.96 of the variation in y (v4) is explained by the variation in these 10 independent x variables and as expected the *F*-value is highly significant. The table also shows the order in which the independent variables entered the equation. Let us now take a closer look at these 10 variables and see how they relate logically to the population growth variable.

It is not surprising that the average value of the dwelling (v43) that individuals live in is related significantly positively ($t=10.41$) to the percentage increase of population in the region. It seems that CD's with high population growth also have high house values. One could ask what comes first; the chicken or the egg. Our logic suggests that high housing values are an indication that economic conditions are good in the CD and this attracts people. Holding other things constant, increases in population increases housing demand which in turn increase prices. Certainly this relationship is known to be true in Canada and holds especially true for some of the bigger cities in the west and for Toronto.

The second variable to enter into the equation to explain the remaining variation of y is v19, 'the % of population that have some earnings and are working full time' with a t value of 3.83. Clearly a region which has a high proportion of its employees working full time is a healthy region. Usually in such regions labour is in short supply and workers can demand higher wages. On the other hand, if they are laid off, the chances are good that they will find another job fast. Furthermore having a full time job also increases the probability of having your own home, which in turn drives up housing prices as was shown above.

'Median age in the community', v6, is the third variable to enter the equation but with a negative sign ($t = -7.78$). It follows from the fact that younger people tend to move more then older people, which tends to bring the age down in the destination area. Furthermore, a dynamic growing region is not very attractive for older people because such regions have not yet build up the infrastructure to look after the elderly.

In the western world private car transportation to work (v24) is frequently presented as a major pollution culprit. But it gives a person freedom of movement which is frequently essential in a growing region. It is therefore not unexpected that there is a significant ($t = 3.56$) relationship with the proportion of people driving to work in a car and the population growth percentage in a particular Census Division. Cars are essential for modern individuals wanting to respond to different spatial opportunities.

An interesting result occurs when variable 37 ''Earned income as % of total'' enters the equation. At first thought one would expect a positive relationship but it is statistically negative with a partial correlation coefficient of – .6363 and a t value of – 10.16. A cursory glance at the simple correlation between v4 and v37 shows it however to be positive and highly significant as one would have expected. But the data show also that the four variables that entered the multiple regression equation earlier all had significant simple positive correlation values with v37 and they probably capture the positive variation in y already.

That jobs are crucial in a region if it wants to grow is brought out by the variable 'Labour force employment rate', v26. It is positively associated with the population growth variable with a partial regression coefficient of .2394, which is significant. When employment rate is high most people have work and when it is low many don't. People are attracted to a region where there are many jobs available.

As in the case of many booming regions, individuals move there first and later move their families and/or establish a family there. Family creation does take place when individuals are relatively young. The above relationship of v6 showed this. The analysis of '% of one family households', v41 is negatively associated with v4 the 1991–2001 population growth rate. The higher the growth rate the lower is the partial regression coefficient. It suggests a booming region is not a good place for a single persons, again holding other things constant. It seems to contradict what is the case in mining towns where many miners are perceived to be young and single. Maybe this is also an old image. But it could also be that in the ten year period from 1991 to 2001 singles had moved there by 2001 when the data was collected and by then they also had married.

It usually thought the stagnant regions; with little in and out movement of population will also not attract new economic activities. This is brought out through the variable v11, '% of population 5 years or older that did (had) not move in the last 5 years'. It is negatively associated, partial regression coefficient of –.2559 and a t value of – 6.0031, with percentage growth of CDs. The reverse is also true. The lower the v11 variable the higher will be the v4 value.

A very interesting relationship is displayed between 'Language first learned and still understood, % English' v7, and 'Population increase 1991–2001 in %' our dependent variable y (v4). The negative relationship with a partial correlation value of –.0659 and a t value of –10.7877 clearly suggest that much of the growth take space in CDs where relative few people have English as the mother tongue. Hence part of the growth is driven by immigrants from not English speaking countries and possibly people from other parts of Canada that were immigrants themselves at one time. In 2001 18.46% of Canada's people did not have English or French as their mother tongue (*Statistic Canada, 2001 Census, Community profile files*). It seems that regional growth is associated with these new Canadians.

The last variable to enter the equation was '% of owned dwellings', v42. The higher the v42 value, the higher the population growth rate. Since owning your own home is valued very highly in Canada this is not surprising. Growing regions with good job prospects also create a tremendous investment opportunity in a private dwelling.

Summary

In summary one would say that holding other things constant regions that grow seem to have good full time job prospects, good housing conditions, tend to have higher proportions of non English non French people living in them, are young communities, use the car to go to work, have many people there that moved there in the last five years and have fewer one family household units. What is somewhat surprising is that none of the profession job categories or employment in the various economic sectors mattered. Nor did the unemployment rate, nor the part time job and earnings conditions have an influence. Although one must keep in mind that this is the case once 10 variables already accounted for the majority

of the variation in y, the 'Population increase 1991–2001 in %'. It should be also pointed out again that this variable is an average number for what happened over the last 10 years while most of the independent variables are snap shots from the year 2001 or 2000.

Bibliography

Boots B., S. Drobyshevsky, O. Kochetkova, G. Matginov, V. Petrov, G. Federov, Al. Hecht, A. Shekhovtsov, A. Yudin. Typology of Russian Regions, Consortium for Economic Policy Research and Advice, CIDA, AUCC, IET, WCER and ANE, Moscow, 2002, 412 pp.

Cockcroft, J. D., A. G. Frank and D. L. Johnson. Dependence and Underdevelopment, Anchor Books, 1972.

Dicken P. and Peter E. Lloyd's. Location in Space: Theoretical Perspectives in Economic Geography, Harper, 1990.

Globe and Mail. «The Downside of Alberta's boom» by Dawn Walton and Katherine Harding, in the series «Short-Handed: Making the most of Canada's labour crunch» August 21, 2006, p. 1.

Grimm K. Theorien der Unterentwicklung und Entwicklungstrategien, Westdeutscher Verlag, 1997.

Meyer S. and A. Hecht. «University growth poles in Canada: An Empirical Assessment», Canadian Journal of Regional Science, Vol. XIX, No. 3, 1996, pp. 263–282.

Hecht A. Review Essay: «Regional Development and Regional Policies»; based on Regional Policy in a Changing World, by N. Hansen, B. Higgins and J. Savoie, New York, Plenum Press, 1990, 311 pp., Canadian Journal of Regional Science, Vol. XIV, No. 3, 1991, pp. 447–453.

Hecht, A. 'A Socioeconomic Comparison of Jews, Mennonites and Ukrainians in Canada, The Prairies and Winnipeg', in Ahornblätter, Schriften der Universitätsbibliothek Marburg Nr. 11, 1998, pp. 45–73.

Boots B. and A Hecht. «Spatial perspectives on Canadian Provincialism and Regionalism», Canadian Journal of Regional Science, Vol. 12, No. 2, 1989, pp. 187–204.

Hecht, A. (ed.). Regional Developments in the Peripheries of Canada and Europe. Winnipeg: Manitoba Geographical Series, Vol. l8, 1983, 255 pp.

International Monetary Fund, 1997, 'Meeting the Challenge of Globalization in the Advanced Economies', World Development Report, April, pp. 45–71.

Knox, P. and J. Agnew. The Geography of the World Economy, Edward Arnold, 1989.

Krugman, P. Development, Geography, and Economic Theory, MIT Press, 1995.

Lander, J. B. and A. Hecht. Regional Development in Ontario – Federal and Provincial Involvement. Heft, 81, Marburger Geographische Schriften, Marburg: Geographisches Institute der Universität Marburg, 1980, 169 pp.

Needleman L. (ed). Regional Analysis, Penguin, 1968.

Statistics Canada, Statcan, Community Profiles, http://www12. statcan.ca/english/profil01/CP01/Index.cfm?Lang=E.

Wheeler J. O., P.O. Muller, G.I. Thrall, and T.J. Fik. Economic Geography, Wiley, 1996.

Zhdanov, V., O. Kuznetsova, V. Mau, V. Plyukhin, S. Prikhodko, M. J. Wojciechowski and A. Hecht. Problems Related to Development of the Kaliningrad Region As an Exclave Territory of the Russian Federation. Consortium for Economic Policy Research and Advice, CIDA, AUCC, IET, WCER and ANE, Moscow, 2002, 249 pp.