Convergence and spatial interactions: evidence from Russian regions

Vera Ivanova

NRU Higher School of Economics

Motivation

Purpose of the Study:

- to **reveal the crucial spatial factors** affecting growth of Russian regions;
- to measure the **impact of spatial factors** on growth of Russian regions;
- to estimate spatial externalities of regional distribution of economic activities in Russia and to disclose **regions with distinguished growth paths**.

Motivation:

- existing empirical works conducted on Russian data:
  - cover a short time period;
  - estimated spatial models principally based on cross-sectional data.

Russia


Норвежское море
СЕВЕРНЫЙ ЛЕДОВИТЫЙ ОКЕАН
Восточно–сибирское море
Берингово море
НОРВЕГИЯ
ШВЕЦИЯ
ФИНЛЯНДИЯ
Балтийское море
Баренцево море
Карское море
Море Лаптевых
Республика Саха
(Якутия)
Охотское море
ЭСТОНИЯ
ЛАТВИЯ
ЛИТВА
БЕЛОРУССИЯ
Москва
КАЗАХСТАН
МОНГОЛИЯ
КИТАЙ
ЯПОНИЯ
Японское море
ГРУЗИЯ
АРМЕНИЯ
АЗЕРБАЙДЖАН
Каспийское море

Empirical papers on Russian data

Buccellato T. (2007):

- conditional convergence, data: 1999-2004

Carluer F. (2010):

- convergence clubs, data: 1985–1999

Solanko L. (2003):

- conditional convergence, data: 1992–2001

Kolomak E., Zverev D. (2010):

- conditional convergence, data: 1995–2006

Lugovoy O. et al. (2007):

- conditional convergence, data: 1998–2004

Research questions

Question 1.
Are spatial externalities at work? If yes, are they rather positive or negative?

Question 2.
Is it true that urbanization is associated with the regional growth? Do Russian agglomerations matter?

Question 3.
Do cross-sectional and dynamic market size variations impact on economic performance of the Russian regions?

Data description

79 Russian regions (Chechen Republic is excluded; composite regions – Tyumen and Arkhangelsk oblasts – are considered as single regions). Years: 1996–2010.

Sources:

- indicators of the economic and social development: statistical yearbooks “Russian regions”, “Social development and quality of life in Russia”: http://www.gks.ru;
- distances between regional centers: K.Glushchenko’s web-site: http://econom.nsu.ru/staff/chair_et/gluschenko/index.htm

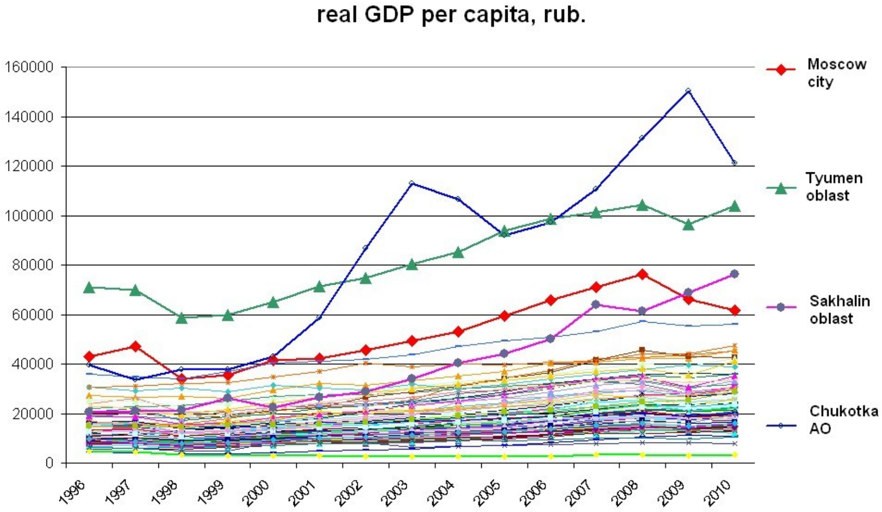
Methodology

Convergence equation (Barro regression)

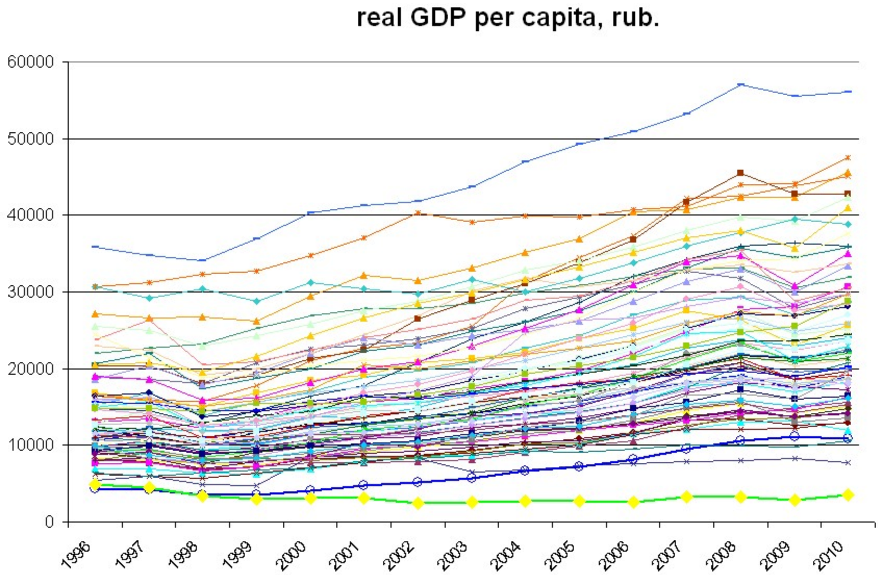
$$\frac{1}{T}\ln\frac{y_{Ti}}{y_{0i}} = a + b\ln y_{0i} + \varepsilon_i,$$

- i – region, $i = 1, 2, ..., n$,
- T – time (years),
- y_{0i} – initial key indicator (e.g. GDP per capita),
- y_{Ti} – key indicator in the year T,
- ε_i – errors, $\varepsilon_i \sim iid$ with 0 mean and finite 2nd moment.

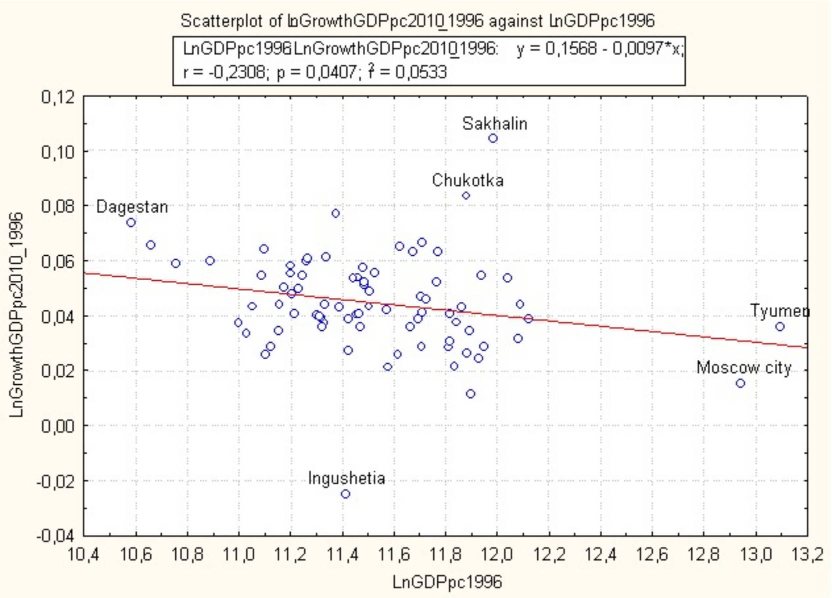
Real GDP per capita corrected by the PPP



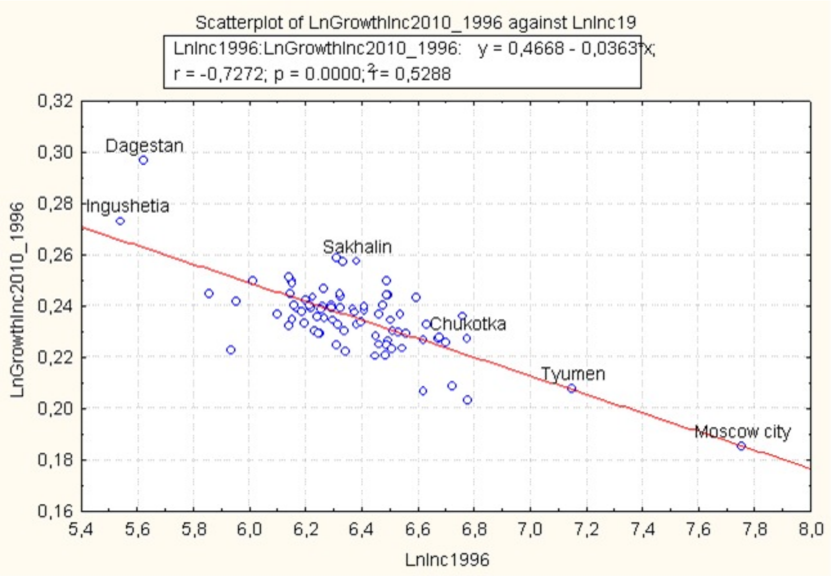
Real GDP per capita



GDP per capita



Income per capita



Methodology. Spatial models

Conditional convergence equation. Spatial error model (SEM), (Anselin, Hudak, 1992)

$$\frac{1}{T}\ln\frac{y_{(t+1),i}}{y_{ti}} = a + b\ln y_{ti} + c' X_{ti} + u_{ti},$$

$$u_{ti} = \lambda W u_{ti} + \varepsilon_{ti},$$

where

- y_{ti} – a key indicator of output (e.g. GRP per capita) in the i-th region at time t, $t = 0, ..., T$,
- X_{ti} – a vector of control variables,
- W – a matrix of spatial weights,
- λ – a spatial autoregressive parameter,
- ε_{ti} – a vector of homoskedastic and uncorrelated errors.

Testing hypothesis 1

Methods:

- to add regional dummies in regression models:
  - for regions with high or low growth rates;
  - border dummy (for regions bordering with foreign countries);
  - sea dummy (for regions with a navigable non-freezing sea port);
- to plot regional spatial interaction using matrix of weights; to measure spatial interaction of regions by means of spatial autocorrelation (Moran's I).

Spatial autocorrelation

Moran's I:

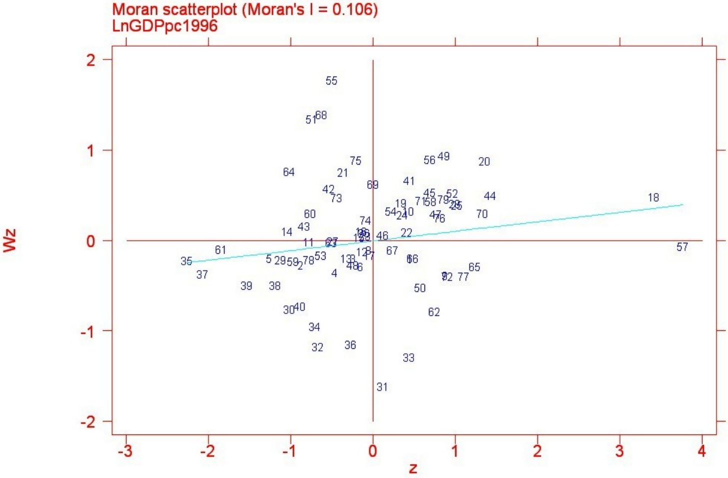
$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2}$$

Different types of spatial matrices w_{ij}:

- matrix of neighbors ($w_{ij} = 1$, if regions i and j have a common border, $w_{ij} = 0$, otherwise);
- Kaliningrad-modified matrix of neighbors
- Moscow-modified matrix of neighbors
- inverse railway-distance matrix ($w_{ij} = \frac{1}{dist_{ij}^{\gamma}}$), $dist_{ij}$ – railway distance between regional centers i and j (thousands km).

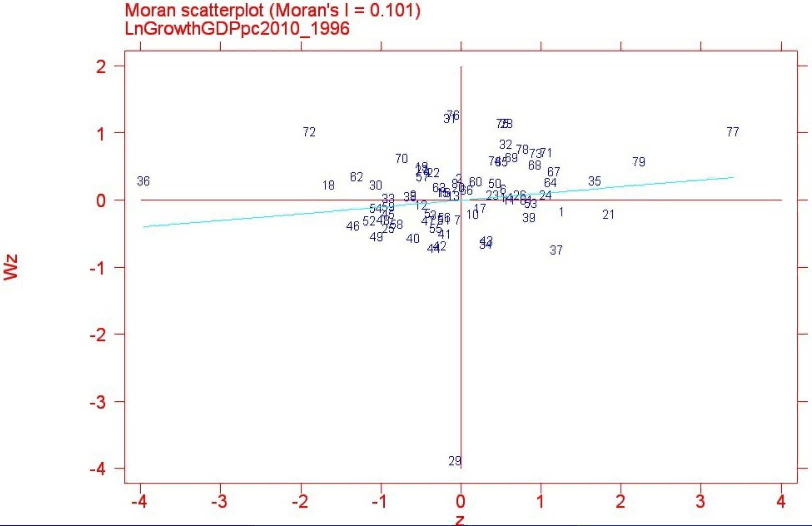
Variable: **GDP per capita 1996 (log).**

W – matrix of neighbors. N=79.



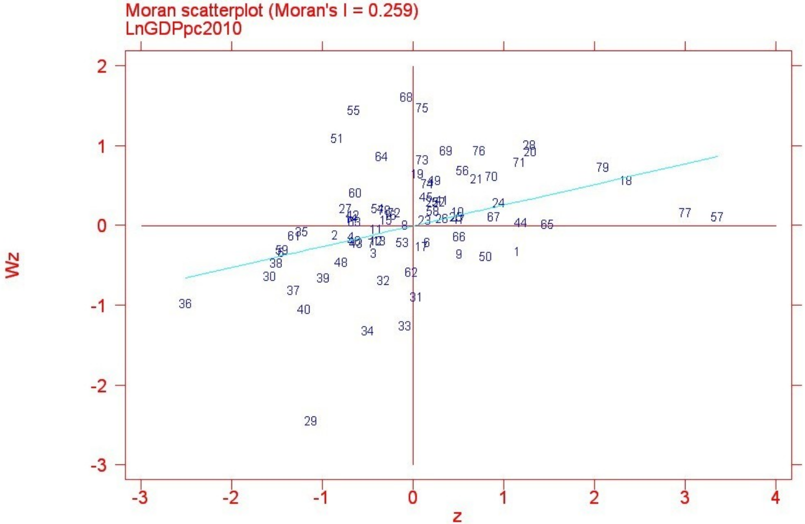
Moran's scatterplot

Variable: **annual GDP per capita growth (log).**
W – matrix of neighbors. N=79.



Moran's scatterplot

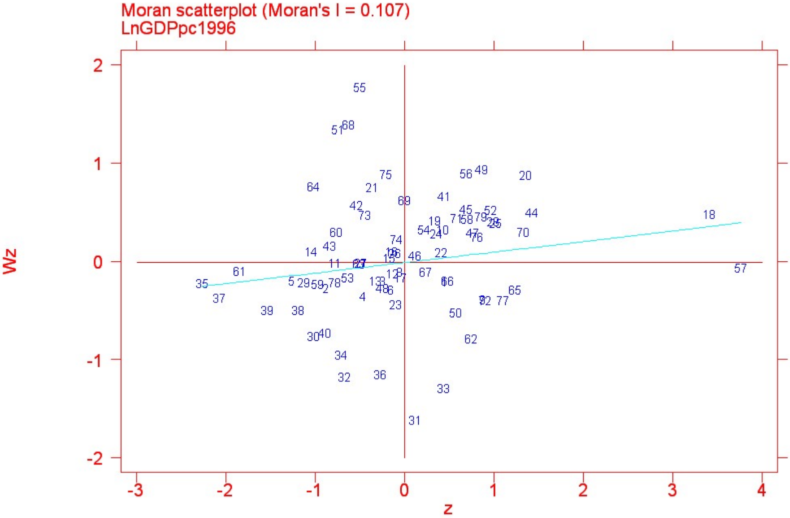
Variable: **GDP per capita 2010 (log).**
W – matrix of neighbors. N=79.



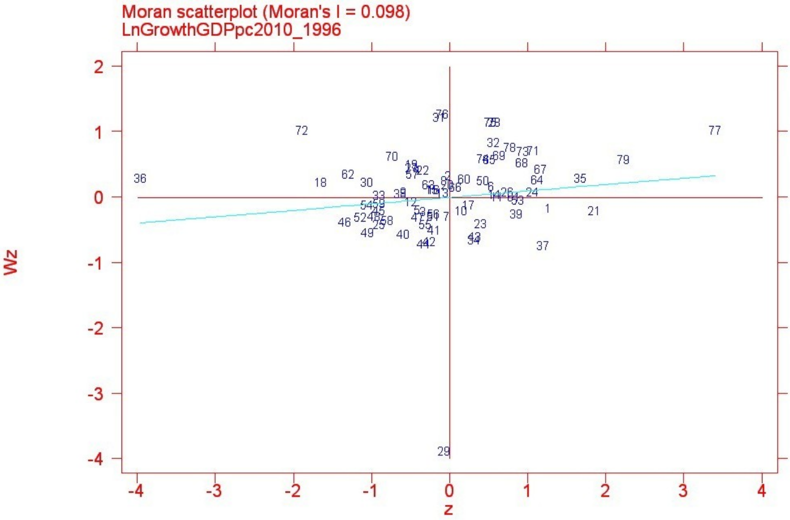
Moran's scatterplot

Variable: **GDP per capita 1996 (log).**

W – matrix of neighbors (Kaliningrad-modified). N=79.

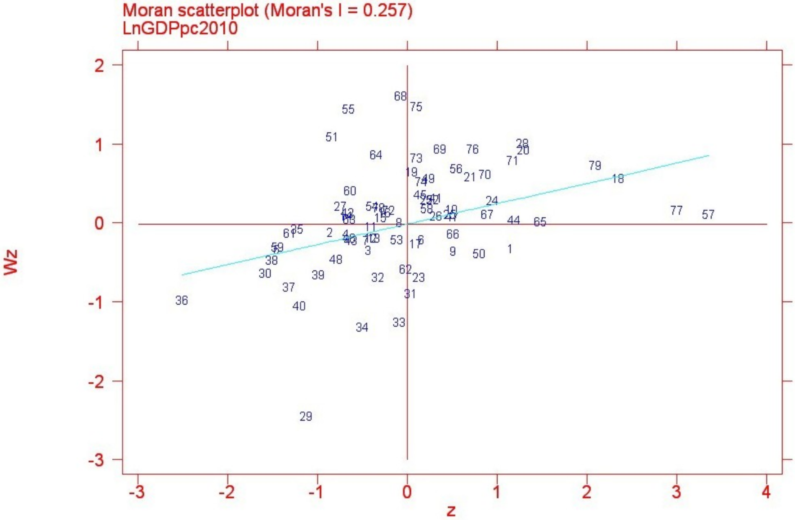


Variable: **annual GDP per capita growth (log).**
W – matrix of neighbors (Kaliningrad-modified). N=79.



Moran's scatterplot

Variable: **GDP per capita 2010 (log)**.
W – matrix of neighbors (Kaliningrad-modified). N=79.



Moran's I

Variable: **GDP per capita (log), 1996–2010.**
W – matrix of neighbors (Kaliningrad-modified). N=79.

Moran's I

Variables	I	E(I)	sd(I)	z	p-value*
LnGDPpc1996	0.107	-0.013	0.078	1.535	0.125
LnGDPpc1997	0.148	-0.013	0.078	2.059	0.040
LnGDPpc1998	0.167	-0.013	0.079	2.284	0.022
LnGDPpc1999	0.201	-0.013	0.079	2.709	0.007
LnGDPpc2000	0.179	-0.013	0.079	2.438	0.015
LnGDPpc2001	0.224	-0.013	0.079	3.006	0.003
LnGDPpc2002	0.265	-0.013	0.079	3.527	0.000
LnGDPpc2003	0.268	-0.013	0.079	3.568	0.000
LnGDPpc2004	0.265	-0.013	0.079	3.528	0.000
LnGDPpc2005	0.272	-0.013	0.079	3.627	0.000
LnGDPpc2006	0.280	-0.013	0.078	3.734	0.000
LnGDPpc2007	0.255	-0.013	0.079	3.408	0.001
LnGDPpc2008	0.260	-0.013	0.078	3.481	0.000
LnGDPpc2009	0.253	-0.013	0.078	3.386	0.001
LnGDPpc2010	0.257	-0.013	0.079	3.438	0.001

*2-tail test

Moran's I

Variable: **GDP per capita (log), 1996–2010.**
W – matrix of neighbors (Moscow-modified). N=79.

Moran's I

Variables	I	E(I)	sd(I)	z	p-value*
LnGDPpc1996	0.027	-0.013	0.059	0.675	0.500
LnGDPpc1997	0.033	-0.013	0.060	0.766	0.443
LnGDPpc1998	0.074	-0.013	0.060	1.445	0.148
LnGDPpc1999	0.096	-0.013	0.060	1.823	0.068
LnGDPpc2000	0.083	-0.013	0.060	1.602	0.109
LnGDPpc2001	0.117	-0.013	0.060	2.166	0.030
LnGDPpc2002	0.153	-0.013	0.060	2.762	0.006
LnGDPpc2003	0.146	-0.013	0.060	2.654	0.008
LnGDPpc2004	0.144	-0.013	0.060	2.625	0.009
LnGDPpc2005	0.136	-0.013	0.060	2.486	0.013
LnGDPpc2006	0.143	-0.013	0.060	2.609	0.009
LnGDPpc2007	0.149	-0.013	0.060	2.711	0.007
LnGDPpc2008	0.150	-0.013	0.060	2.721	0.007
LnGDPpc2009	0.166	-0.013	0.060	2.997	0.003
LnGDPpc2010	0.177	-0.013	0.060	3.174	0.002

*2-tail test

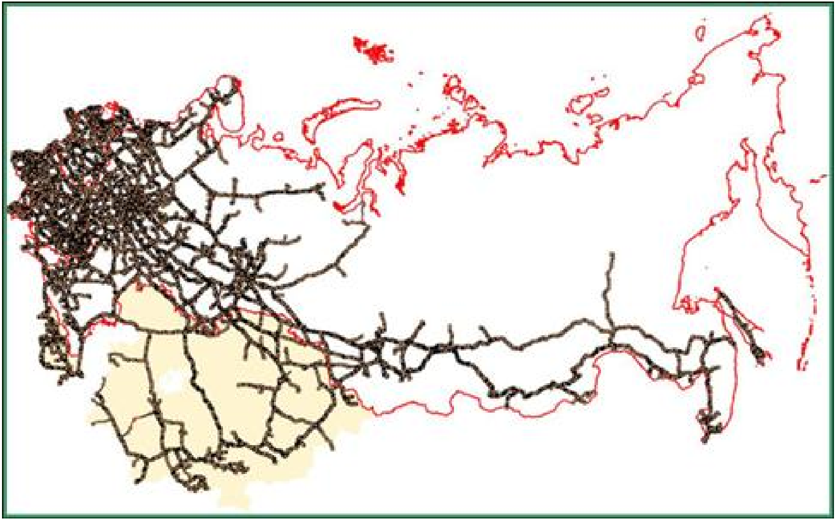
Does distance matter? Moran's I

Variable: **GDP per capita (log), 1996–2010.**
W – inverse distance matrix ($\gamma = 1, 2, 3$). N=79.

Moran's I (inverse distance matrix)

Variables	I(gamma=1)	I(gamma=2)	I(gamma=3)
LnGDPpc1996	0.050	0.110	0.128
LnGDPpc1997	0.057	0.128	0.150
LnGDPpc1998	0.069	0.148	0.174
LnGDPpc1999	0.087	0.186	0.225
LnGDPpc2000	0.071	0.157	0.191
LnGDPpc2001	0.077	0.168	0.203
LnGDPpc2002	0.081	0.176	0.215
LnGDPpc2003	0.083	0.182	0.224
LnGDPpc2004	0.085	0.194	0.245
LnGDPpc2005	0.080	0.186	0.237
LnGDPpc2006	0.078	0.182	0.232
LnGDPpc2007	0.070	0.163	0.207
LnGDPpc2008	0.070	0.168	0.213
LnGDPpc2009	0.061	0.150	0.191
LnGDPpc2010	0.071	0.164	0.211

Russian railways system



The role of transportation system. SEM

Variables	Dependent variable: $\frac{1}{14}\ln\frac{GRPpc_{2010}}{GRPpc_{1996}}$		
	(1)	(2)	(3)
constant	0,205*** (0,055)	0,209*** (0,037)	0,212*** (0,036)
$\ln GRPpc_{1996}$	-0,014* (0,004)	-0,014*** (0,003)	-0,015*** (0,003)
dummySakhalin	–	0,064*** (0,011)	0,064*** (0,011)
dummyChukotka	–	0,039*** (0,011)	0,046*** (0,012)
dummyDagestan	–	0,016 (0,012)	0,016 (0,011)
dummyIngushetia	–	-0,068*** (0,011)	-0,067*** (0,000)
$\ln RailWayDensity$	–	–	0,0007* (0,0004)
Variance ratio	0,108	0,509	0,519
Speed of convergence, %	1,53	1,58	1,65
Half-level convergence, years	45,2	43,8	42,1

The role of transportation system. SEM

	(1)	(2)	(3)	(4)
Variance ratio	0,108	0,509	0,517	0,519
Squared corr,	0,053	0,526	0,540	0,526
Sigma	0,02	0,01	0,01	0,01
Log likelihood	211,54	239,73	239,94	241,15
λ	0,370* (0,193)	0,426** (0,188)	0,353 (0,226)	0,497** (0,176)
Wald test of λ =0: $\chi^2(1)$	3,675 (p=0,055)	5,171 (p=0,023)	2,444 (p=0,118)	8,010 (p=0,005)
Likelihood ratio test of λ =0: $\chi^2(1)$	3,236 (p=0,072)	4,405 (p=0,036)	2,216 (p=0,137)	6,197 (p=0,013)
Lagrange multiplier test of λ =0: $\chi^2(1)$	2,438 (p=0,118)	3,285 (0,070)	1,297 (p=0,255)	4,416 (p=0,036)

The role of population growth. SEM

Spatial error model (inverse distance matrix) Variables	Dependent variable: $\frac{1}{14}\ln\frac{Income_{2010}}{Income_{1996}}$				
	(1)	(2)	(3)	(4)	(5)
constant	0.463*** (0.024)	0.458*** (0.024)	0.393*** (0.031)	.460*** (0.023)	0.401*** (0.031)
$\ln Income_{1996}$	-0.036*** (0.004)	-0.035*** (0.004)	-0.024*** (0.004)	-0.035*** (0.003)	-0.026*** (0.005)
LnPopulationGrowth	–	.281** (0.136)	.518*** (0.177)	–	–
LnLaborGrowth	–	–	–	0.406*** (0.132)	0.374*** (0.133)
dummyMoskva	–	–	-0.027** (0.011)	–	-0.019* (0.011)
dummyTumen	–	–	-0.017* (0.009)	–	-0.011 (0.009)
dummyChukotka	–	–	0.016* (0.009)	–	0.009 (0.008)
dummySakhalin	–	–	0.027*** (0.008)	–	0.022*** (0.008)
dummyDagestan	–	–	0.029*** (0.009)	–	0.032*** (0.009)
dummyIngushetia	–	–	-0.0006 (0.009)	–	0.008 (0.008)
Variance ratio	0.511	0.509	0.716	0.611	0.718
Speed of convergence. %	5.01	4.81	2.92	4.81	3.23
Half-level convergence. years	13.8	14.4	23.7	14.4	21.4

The role of population growth. SEM

	(1)	(2)	(3)	(4)	(5)
Variance ratio	0.511	0.561	0.716	0.611	0.718
Squared corr.	0.529	0.552	0.673	0.581	0.682
Sigma	0.01	0.01	0.01	0.01	0.01
Log likelihood	252.98	255.054	268.77	257.41	268.47
λ	0.268* (0.159)	0.28* (0.16)	.35** (0.15)	.245 (0.164)	.255 (0.16)
Wald test of λ =0: $\chi^2(1)$	2.821 (p=0.093)	3.028 (p=0.082)	5.282 (p=0.022)	2.237 (p=0.135)	2.498 (p=0.114)
Likelihood ratio test of λ =0: $\chi^2(1)$	2.626 (p=0.105)	2.795 (p=0.095)	4.486 (p=0.034)	2.076 (p=0.150)	2.293 (p=0.130)
Lagrange multiplier test of λ =0: $\chi^2(1)$	2.438 (p=0.118)	2.338 (0.126)	3.463 (p=0.063)	1.682 (p=0.195)	1.877 (p=0.171)

Panel data. Do regions converge by income?

Variables	Dependent variable: $\ln \frac{Income_{t+1}}{Income_t}$		
	Fixed effects model	**Random-effects model**	**Between-effects model**
constant	0.507*** (0.005)	0.508*** (0.006)	0.626*** (0.056)
$\ln Income_t$	-0.039*** (0.001)	-0.039*** (0.001)	-0.053*** (0.007)
R^2 within	0.769	0.769	0.769
R^2 between	0.426	0.426	0.425
R^2 overall	0.708	0.708	0.708
Statistics	F(78, 1026) = 6.71	Wald chi2(1) = 3450.22	F(1,77) = 57.07
corr(u_i, Xb)	0.0538	0 (assumed)	sd(u_i + avg(e_i.))= .016
sigma_u	0.017	0.015	
sigma_e	0.024	0.024	
rho	0.324	0.28	
Speed of convergence. %	5.64	5.64	9.68
Half-level convergence. years	12.3	12.3	7.2

Panel data. Do regions converge by GDPpc?

Variables	Dependent variable: $\ln \frac{GDPpc_{t+1}}{GDPpc_t}$	
constant	-0.940*** (0.083)	0.663*** (0.006)
$\ln GDPpc_t$	0.083*** (0.111)	-0.048*** (0.007)
Dummy for crisis years (1998, 2008)	no	yes
R^2 within	0.1169	0.7236
R^2 between	0.0165	0.0165
R^2 overall	0.0477	0.7057
Speed of convergence. %		5.64
Half-level convergence. years		12.3

Testing hypothesis 2

$$\frac{1}{T}\ln\frac{y_{(t+1),i}}{y_{ti}} = a + b\ln y_{ti} + c'X_{ti} + u_{ti},$$

$$u_{ti} = \lambda W u_{ti} + \varepsilon_{ti},$$

- y_{ti} – GRP-by-industry per capita,
- $X_{ti} = (Aggl_{ti}, H_{ti}, Innov_{ti}, Edu_{ti}, MP_{ti})'$ – a vector of control variables;
  - $Aggl_{ti}$ – agglomeration indicator of the i-th region at time t ;
  - H_{ti} – indicator of sectoral diversity (Herfindhal-Hirshman index);
  - $Innov_{ti}$ – measure of own innovative activity;
  - Edu_{ti} – share of skilled workers;
  - $MP_{ti} = \sum\limits_{j\neq i} \frac{GDPind_{t_{ij}}}{Dist_{ij}}$ – real market potential, $Dist_{ij}$ – distance between regional centers.

Testing hypothesis 2. Preliminary results

GDP-by-industry is available for the period 2005–2009 (15 sectors).
Simple regressions show unconditional convergence by GDPpc.

Testing hypothesis 3

$$\frac{1}{T}\ln\frac{y_{(t+1),i}}{y_{ti}} = a + b\ln y_{ti} + c_1\ln MarketSize_{ti} + c_2\ln Migration_{ti} + u_{ti},$$

$$u_{ti} = \lambda W u_{ti} + \varepsilon_{ti},$$

$$MarketSize_{ti} = \Omega_i + v_{ti},$$

$$v_{t+1,i} = \rho v_{ti} + \xi_{t+1,i},\ 0 \le \rho < 1,$$

- $MarketSize$ – population, Ω_i – a stable target size, v_{ti} – a possibly persistent random shock, ξ_i – iid random shocks,
- $Migration_{ti}$ – in-migration.

Determinants for Russian in-migration (Yu.Andrienko, S. Guriev, 2003):

- initial personal income in the region i ;
- big cities;
- resource potential;
- climate;
- public goods.

Endogeneity of Market Size

As market size is clearly endogenous it should be instrumented.
Instruments:

- data of population census in Russian Empire in 1897,
- data of population census in USSR (1926, 1937, 1939, 1959, 1970, 1979, 1989),
- data of population census in Russian Federation (2002, 2010).

Conclusions

Spatial externalities exist.

Russian regions interactions highly depend on accessibilitiy of railways system. Interaction almost disappears while distance between regions becomes large.

Regions with an advantageous geographic position (e.g. bordering foreign countries, having sea-ports) grow faster.

There are regions with distinguished growth paths: Moscow city, Tyumen oblast, Sakhalin oblast, Chukotka AO, Dagestan Republic, Ingushetia Republic.

Further plans

- Russian regions are highly heterogeneous, so it's reasonable to get access to more disaggregate data in order to redo the estimates, which may allow to catch the desirable effects in a more subtle way.

Thanks

Thank you for your attention!