

Beyond Beta-Convergence: Convergence in Differences and its Application to the Russian Regions

Gianni Carvelli¹

¹ University of Brescia, Italy

Correspondence: Gianni Carvelli, Department of Economics and Management, University of Brescia, Italy.
E-mail: g.carvelli@unibs.it

Received: July 23, 2020

Accepted: August 31, 2020

Online Published: September 18, 2020

doi:10.5539/ijef.v12n10p45

URL: <https://doi.org/10.5539/ijef.v12n10p45>

Abstract

The purpose of this paper is to propose a new empirical model capable of highlighting some aspects of cross-economy convergence which cannot be caught by the popular beta-convergence and sigma-convergence models. The idea is to analyse the growth of the economies as a function of the distance between the observed output per capita and the average output per capita within the sample, separating the behaviour of poorest and richest economies. After its specification, I applied the model to the case of the Russian regions over the period 1995-2015 using the fixed-effect estimator. The results show that, although the existence of a significant beta-convergence process, there is a lack of convergence in differences. When the differences between regional and national output per capita are negative, a positive and significant relationship between growth and levels emerges. Such a relationship turns to be negative and non-significant when the differences are positive, therefore denoting weak non-linearity between growth rate and level of output per capita. Similar findings have been found for labor productivity.

Keywords: Beta-convergence, convergence in differences, cross-regional convergence, economic growth, Russian regions

1. Introduction

Nowadays the issue of economic growth and convergence is no more as debated as in the second half of the previous century. It reached the peak of the attention in the economic literature in 80's and 90's. We can distinguish two main groups of economic growth theories: the exogenous and the endogenous growth models. These theories have been largely empirically tested using cross-countries and cross-regional data for several time periods, thus letting the convergence controversy arise (Note 1).

The neoclassical (or exogenous) theories of economic growth, based on the pioneering paper of Solow (1956), predict long-run convergence in levels of output per capita between economies with the same structural characteristics (absolute convergence), thus implying that all the economies within the sample will reach the same steady state level. If the structural characteristics are not homogeneous, the economies will converge towards different steady state levels (conditional convergence) (Note 2) but their growth rates will not differ in the long-run equilibrium. The neoclassical theories of growth are based on two key-assumption: decreasing returns to scale of physical capital and exogenous technology level (Romer, 2006).

A plenty of so called "endogenous growth theories", often referred to as "AK models" have been developed already since 60's (Arrow, 1962; Uzawa, 1965; Nordhaus, 1969; Shell, 1973; Romer, 1986; Lucas, 1988; Barro, 1990; Kremer, 1993; Jones, Manuelli, & Stacchetti, 2000). One of the most important innovation was to allow for the existence of increasing returns to scale of physical capital (Dixit & Stiglitz, 1975; Skiba, 1978; Nishimura, 1983; Romer, 1986).

The most common ways to empirically test the neoclassical growth theories are the so-called "unconditional beta-convergence" and "conditional beta-convergence" models.

The unconditional beta-convergence model, firstly introduced by Baumol (1986), can be represented by the following equation:

$$y_{t_0:t_T,i} = \alpha + \beta x_{t_0,i} + \varepsilon_i \quad (1)$$

where y represents the growth rate of the real GDP per capita between the initial and the final period of the dataset,

x represents the level of the real GDP per capita at the initial period, ε is the error term and i denotes the i -th economy.

The existence of a beta-convergence process implies that the coefficient β is negative and statistically significant. From the economic point of view, the beta-convergence process implies that the poorest economies grow, on average, faster than the richest economies (Note 3).

In order to control the heterogeneity between the economies, many scholars adopted and extended the version of the above-mentioned model, which is called “conditional beta-convergence”. It can be represented as follows:

$$y_{t_0:t_T,i} = \alpha + \beta x_{t_0,i} + \gamma' z_{t_0,i} + \varepsilon_i \quad (2)$$

which is an extension of the equation 1, represented by the vector of variables \mathbf{z} in which are included the conditioning factors (Note 4).

The empirical models testing for the endogenous growth theories are often based on the conditional beta-convergence model, with the inclusion of factors which are considered exogenous by the neoclassical growth theories (Note 5).

Nevertheless, due to the complex nature of the phenomenon under analysis, the literature is still far from providing robust results which can be considered generally valid. Indeed, Mankiw, Phelps, and Romer (1995) state that the payoff resulting from endogenous growth theories lacks of clarity. Furthermore, he says that physical capital is even more relevant than human capital in explaining cross-country differences in output per capita, since knowledge can more easily spread between economies.

From the empirical perspective, several sources of biases in the growth regressions arise (Aghion & Durlauf, 2005; Acemoglu, 2009). Panizza and Persbitero (2013) state that almost the totality of empirical studies in the field of economic growth provides results which could be biased, mainly because of endogeneity.

In the most recent contributions regarding economic growth, the system GMM estimators developed by Arellano and Bond (1991) and Blundell & Bond (1998) have been widely employed in order to address endogeneity. Nevertheless, the GMM approach, although enjoys several advantages compared to other approaches, could lead to non-reliable results due to the problems of “over-instrumentation”, which in turns leads to weakness of instruments (Roodman, 2009).

Furthermore, Roodman (2009) suggests to using a more straightforward fixed-effect estimator when the time series are sufficiently long, since the dynamic panel bias becomes insignificant.

In addition to beta-convergence model, the so-called “sigma-convergence model” has gained relevance over the years. According to the definition of Sala-i-Martin (1996), a group of economies experiences sigma-convergence if the dispersion of their real GDP per capita tends to decrease over the time. He also highlights the direct relationship between beta and sigma-convergence: if poor economies grow faster than rich economies, the dispersion of the real GDP per capita within the sample will decrease (Note 6). Thus, a necessary condition for the existence of sigma-convergence is the existence of beta-convergence, but the existence of beta-convergence does not automatically imply sigma-convergence.

Regardless the choice of the econometric techniques, I believe that both beta and sigma-convergence models, although are capable of providing very useful information, are characterized by some limitations so far ignored in the economic literature.

Firstly, they are very sensitive to the presence of outliers, also because this kind of analysis is generally carried out within a limited sample size (equal to the number of economies, regardless the length of the time span).

Secondly, the average causal effect measured by the coefficient β might led to stating that, within the sample, there is a significant beta-convergence process, thus implying that the economies have been converging toward the same steady state level. This conclusion is often imprecise. The reason is that the behaviour among the economies whose output per capita level is below the national average might considerably differ between them: the beta-convergence process can be totally driven by the economies whose income level is below the national average, but that still are not the poorest. In such a situation, both the processes of sigma-convergence and beta-convergence can coexist with the persistence of differences in terms of output per capita levels.

For the above-mentioned reasons, the purpose of this paper is to propose a new empirical model capable of analysing the process of convergence/divergence in differences in terms of output per capita within a sample of economies. This kind of analysis is not an alternative to the popular beta and sigma-convergence models, but it can be carried out alongside them in order to understand better the convergence dynamics within a sample of economies in a given period of time. The mathematical framework and the empirical construction of the model

will be carried out in the section 2.

After the construction of the model, it will be performed an empirical application to the case of the Russian regions over the period 1995-2015. The process of convergence between the Russian regions represents an interesting case of study since the Russian Federation is known in the literature as one of the countries characterized by the highest internal disparity at regional level (Benini & Czyzewski, 2007; Badunenko & Tochkov, 2010). Furthermore, testing a convergence model to analyse the cross-regional convergence rather than cross-countries gains some advantages in terms of homogeneity of several macroeconomics characteristics, which can be considered region-invariant within the same country (Note 7). The observable region-invariant characteristics should be controlled in cross-country analysis.

The studies focusing on the process of convergence/divergence between the Russian regions are characterized by relevant differences in terms of methodology, time intervals and variable taken into account. Furthermore, the scarcity of variables covering the entire Russian territory and the years of transaction from a socialist to a market economy represented an obstacle to the implementation of deepened empirical analysis. Nevertheless, there exist general robust findings which describe well the evolution of income disparities between the Russian regions, mainly measured by the beta-convergence models. Indeed, there is a clear common evidence for beta-divergence between the Russian regions during the first part of the transition period (Popov, 2001; Dolinskaya, 2002; Fedorov, 2002; Berkovitz & Dejong, 2002; Carluer, 2005; Benini et al., 2007; Badunenko et al., 2010) whereas evidence of beta-convergence has been found in the “Putin era” (Akhmedjonov, Chi, & Izgi, 2013; Oshchepkov, 2015) which in particular increased during the last decade (Vakulenko, 2016; Durand-Lasserve et al., 2018; Kaneva & Untura, 2019).

In a recent contribution, Lehmann, Oshchepkov, and Silvagni (2020), in the spirit of the neoclassical growth theories, estimated growth equations including human capital and migration as conditioning factors. They do find convergence over the period 1996-2017, and a determinant role is played by interregional migration, whereas no significant impact is found for human capital.

2. Convergence in Differences

The idea of strong convergence in differences can be mathematically represented as follows:

$$\lim_{t \rightarrow \infty} y_{i,t} - y_{max,t} = \lim_{t \rightarrow \infty} y_{i,t} - y_{mean,t} = 0 \forall i \quad (3)$$

where y is the real output per capita, i denotes the i -th economy, t denotes the time, max represents the economy with the highest output per capita at time t , y_{mean} is the average output per capita within the sample at time t . The process of strong convergence implies that all the economies within the sample converge to the same steady state level.

The weak convergence in differences can be represented as follows:

$$\lim_{t \rightarrow \infty} g_{i,t} - g_{max,t} = \lim_{t \rightarrow \infty} g_{i,t} - g_{mean,t} = 0 \forall \quad (4)$$

where g represents the growth rate of real output per capita. Weak convergence allows the existence of economy-specific steady state level. Strong convergence implies weak convergence but not vice-versa.

It is interesting to note that neither beta-convergence nor sigma-convergence imply convergence in differences. Indeed, the persistence in differences can coexist both with an above average growth of a subset of the poorest economies and with an aggregate reduction in income dispersion. Instead, the presence of convergence in differences implies both beta and sigma-convergence. Therefore, convergence in differences is the strongest condition of convergence. Although it could be too demanding in real world, its analysis can provide useful insights about the convergence dynamics and the behaviour of the economies on their balanced growth paths.

Empirically, it is not straightforward to test the equations 3 and 4. I would like to propose the following empirical specification:

$$g_{i,t} = \alpha + \phi |DIFF_LEVELS|_{i,t} + \gamma' z_{i,t} + \eta_i + \mu_t + \varepsilon_{i,t} \quad (5)$$

where the variable $|DIFF_LEVELS|$ represents the absolute value of the difference between the observed level of output per capita and the average level of output per capita within the sample, z is a vector of control variables. The dummies η and μ represent, respectively, the time invariant and temporal effects, whereas the indexes i and t denote, respectively, the i -th economy and the t -th period.

The coefficient ϕ will tell us how responsive is the growth to the distance between the average and the observed level of output per capita. Nevertheless, the equation 5 is not our main focus since the variable $|DIFF_LEVELS|$ includes both observations whose output level is below and above the output mean (Note 8). The following two

equations can be more useful.

$$g_{i,t} = \alpha + \lambda |DIFF_LEVELS|_{i,t} * I(y_{i,t} < y_{mean,t}) + \gamma' z_{i,t} + \eta_i + \mu_t + \varepsilon_{i,t} \quad (6)$$

The variables in the equation 6 are the same as the variables in the equation 5, with the difference that here appears the indicator function $I(y_{i,t} < y_{mean,t})$, which takes values equal to one for the observation whose level of output per capita is below the national average in a given year and zero otherwise. The interaction between the variable $|DIFF_LEVELS|$ and the indicator function will provide an estimate λ which will tell us the average change of the growth rate to a unit increase in the absolute value between the observed and average output per capita. Since in the equation 6 are taken into account only the observations for which the output per capita level is less than the average output per capita, we can refer to such equation as *convergence from below*.

$$g_{i,t} = \alpha + \psi DIFF_LEVELS_{i,t} * I(y_{i,t} \geq y_{mean,t}) + \gamma' z_{i,t} + \eta_i + \mu_t + \varepsilon_{i,t} \quad (7)$$

The equation 7 is the same as the equation 6, with the only difference that the indicator function takes value one for the observation whose level of output per capita is above or equal to the average output per capita in a given year and zero otherwise. Furthermore, since the interaction term assures that the differences between observed output per capita and average output per capita is positive, there is no need to employ the absolute value. We can refer to such equation as *convergence from above*.

We can say that the process of convergence in differences is strong if the coefficient λ in the equation 6 is positive and statistically significant and ψ in equation 7 is negative and statistically significant. Indeed, the combination of these two conditions implies that the differences between the average output per capita and the observed output per capita decrease both from below and from above. Such a condition is quite strong since it implies that the economies within the sample converge toward the same or similar steady-state level of output per capita. If only one of the two above-mentioned conditions is met, we can say that the process of convergence in differences is *weak*.

If the parameter λ is negative and statistically significant, it means that only the richer among the poorest economies converge toward the average output per capita. If the parameter ψ is positive and statistically significant, it means that the average growth increases as the output per capita level increases, therefore violating the neoclassical assumptions.

If parameters λ and ψ are not statistically significant, it means either that the economies reached their steady state level or that the output level does not explain variations in the growth rate.

It is necessary to note that the split of the analysis in the estimation of two separate equations, although can provide a more detailed analysis, implies a loss of observation within the two models.

Anyway, an important pros of the convergence in difference model compared to the classical beta-convergence model is that in the first case, taking advantage of the panel context, we have nt observations whereas in the second case we have only n observation. Indeed, in such a context, each observation is made by an economy at a given period, therefore obtaining t observations for each country/region. It further implies that we can control for the behaviour of each economy at each point in the time.

Following the general suggestions of Islam (1995) and Roodman (2009) for the choice of the estimator for the growth regressions, I will estimate the equations 5, 6 and 7 through the fixed-effect estimator.

3. Empirical Application to Russian Regions

In this section, taking advantage of the recent dataset built by Mirkina (2017), I will apply the methodologies described in the previous section to the case of Russian regions in order to analyse the process of convergence between them over the period 1995-2015. The dataset includes time series for 82 Russian regions (namely *Republics, Krai, Oblasts, Federal Cities, Autonomous Oblasts, Autonomous Okrugs*) (Note 9).

After a proper descriptive analysis, I will estimate beta-convergence and convergence in difference models both for Gross Regional Product per capita (GRP) per capita and labor productivity.

For convenience, I will set $|DIFF_LEVELS|_{i,t} * I(y_{i,t} < y_{mean,t}) = |D_GRP_NEGATIVE|$ and $DIFF_LEVELS_{i,t} * I(y_{i,t} \geq y_{mean,t}) = D_GRP_POSITIVE$. I will use the analogue notation when referring to labor productivity.

The definitions of the variables are attached in the Appendix.

Table 1. Descriptive statistics

Variables	Obs	Mean	Std.Dev.	Min	Max
GRP_GROWTH	1620	.198	.171	-.856	1.778
LP_GROWTH	1613	.197	.226	-.842	5.988
GRP_1995	79	.111	.102	.019	.919
D_GRP_95	79	0	.102	-.092	.808
LP_95	79	.231	.101	.073	.712
D_LP_95	79	0	.101	-.158	.481
D_GRP	1702	.812	1.813	0	26.456
D_GRP_NEGATIVE	1238	.558	.554	0	2.802
D_GRP_POSITIVE	464	1.489	3.261	.001	26.456
D_LP	1696	1.337	2.701	0	35.568
D_LP_NEGATIVE	1232	.924	.903	0	4.544
D_LP_POSITIVE	464	2.433	4.783	.005	35.568
POP	1717	1763.927	1606.215	49	12330
CAP_PW	1696	885.367	1449.289	38.16	22319.48
DEPR	901	333.699	624.929	1.677	5778.168

3.1 Beta and Sigma-Convergence

In order to estimate the equation 1 for both output per capita and labor productivity, I will employ as dependent variables not only the initial levels of the variables, but also the absolute value of the difference between regional level and national mean of output per capita ($|D_GRP_95|$) and labor productivity ($|D_LP_95|$) at the year 1995.

Table 2. Unconditional Beta-Convergence over the period 1995-2015 (OLS estimates)

	(1) GRP_GROWTH	(2) GRP_GROWTH	(4) LP_GROWTH	(5) LP_GROWTH
Constant	35.72*** (1.661)	32.24*** (1.702)	37.49*** (2.872)	30.74*** (1.414)
GRP_95	-31.28*** (7.029)			
D_GRP_95		-31.28*** (7.029)		
LP_95			-29.23** (12.24)	
D_LP_95				-29.23** (12.24)
Observations	79	79	79	79
R-squared	0.043	0.043	0.053	0.053

Robust standard errors in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

Consistently with the main literature, the estimates show a significant process of beta-convergence between the Russian regions over the period 1995-2015 both in terms of GRP per capita and labor productivity. Nevertheless, the relatively low R^2 values suggest that the growth of the regions depend on further factors which are not included in the table 1.

The estimates are the same between the initial level and the differences of the variables, with the only small variations in the values of the intercepts.

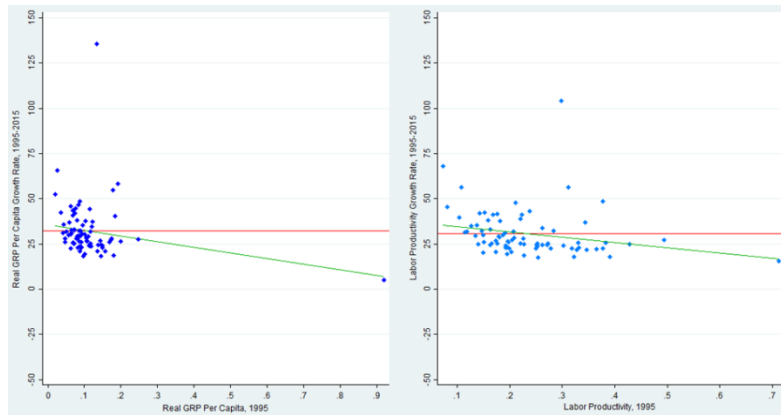


Figure 1. Beta-convergence and sigma-convergence of GRP per capita and labor productivity

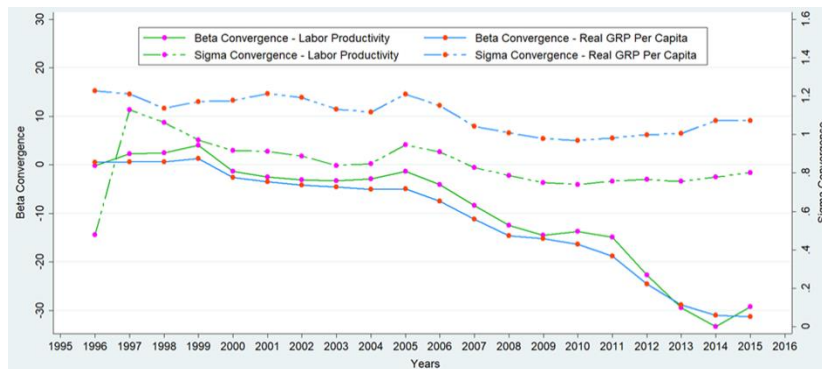


Figure 2. Beta-convergence and sigma-convergence of GRP per capita and labor productivity over the period 1995-2015

In the graph we can observe the dynamics of beta and sigma-convergence of output per capita and labor productivity over the period 1995-2015. In all the cases, the beta-coefficients are estimated taking into account the year 1995 as initial period. We can notice some interesting characteristic: a) GRP per capita and labor productivity follow almost the same dynamics; b) the process of beta-convergence has been intensified since the post-crisis period; c) the intense process of beta-convergence was not accompanied by a process of sigma-convergence, whose levels have been quite persistent since 2007 with a slight increase in the latest three years of the sample.

3.2 Convergence in Differences

In this section I will provide the results of convergence in differences models. First of all, it is worth to analyse graphically the phenomenon in order to perform a useful preliminary analysis.

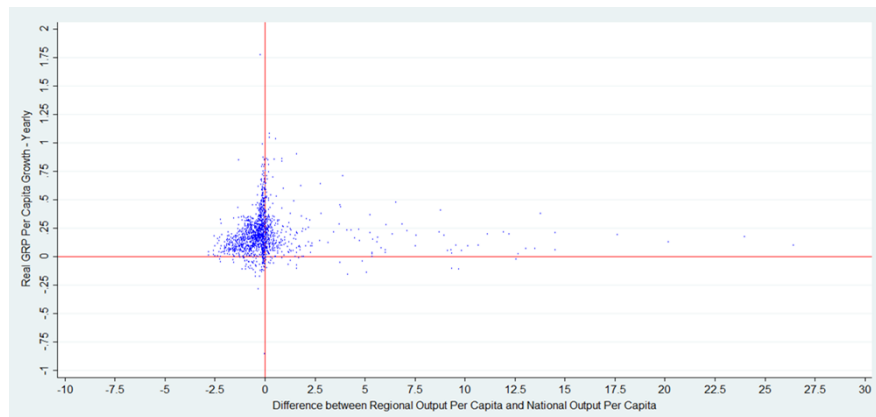


Figure 3. Output per capita growth and differentials between regional and national output per capita (Yearly observation)

In figure 3 we can observe the relationship between the output growth and the distance between regional and national output per capita. The observations located above the horizontal red line experienced a positive growth in a given year, representing the majority of the distribution. The observations at the right (left) of the vertical red line represent the regions whose output per capita is, in a given year, above (below) the national mean.

Four main insights emerge from the graph: a) the majority of the observations lies at the left of the vertical red line, implying that relatively few observations are much above the national average; b) among the observations whose level of output is above the average, there are some outliers which push above the national mean; c) among the observation whose output level is below the national mean, there is a clear positive relationship between output and growth; d) conversely, among the observations whose output level is above the national mean, the relationship between growth and output appears to be negative.

Despite the strong evidence of both beta and sigma-convergence among the Russian regions, it appears that the relationship between growth and output level is non-linear. Indeed, contrary to the predictions of the neoclassical growth models, being among the poorest does not imply having the highest growth rates. In this case, the neoclassical growth hypothesis seems to be verified only among the group of observations whose output levels are above the national average. Therefore, the existence of a significant process of beta-convergence is strongly affected by the presence of outliers whose levels are much above the national mean and whose growth rates are very close to zero.

It is also interesting to notice that the highest growth rates are mostly linked to output values close to the national mean.

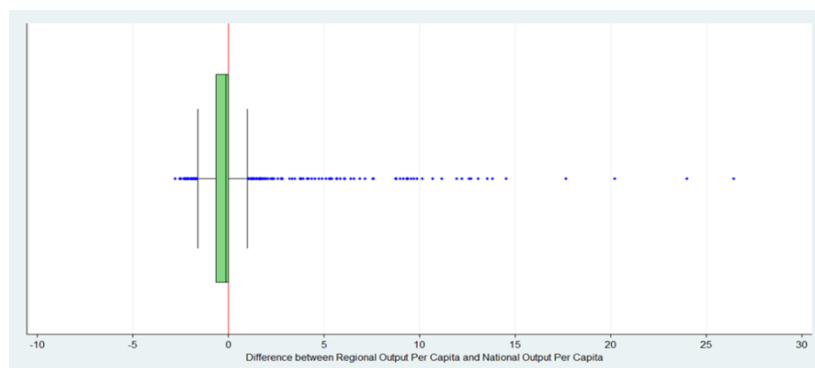


Figure 4. Boxplot of the difference between Regional Output Per Capita and National Output Per Capita

The box-plot in figure 4 highlights the presence of a considerable number of outliers in the right tail of the distribution.

It is now interesting to analyse the estimates obtained through the estimation of the equations 5, 6 and 7, applying it both to output per capita and labor productivity.

Table 3. Convergence in differences. Yearly data. Fixed-effects estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	GRP_GROWTH	GRP_GROWTH	GRP_GROWTH	LP_GROWTH	LP_GROWTH	LP_GROWTH
Constant	0.366*** (0.019)	0.336*** (0.016)	0.448*** (0.055)	0.465*** (0.025)	0.458*** (0.031)	0.489*** (0.048)
D_GRP	0.000 (0.003)					
D_GRP_NEGATIVE		-0.041*** (0.013)				
D_GRP_POSITIVE			0.003 (0.003)			
D_LP				0.000 (0.003)		
D_LP_NEGATIVE					-0.049** (0.022)	
D_LP_POSITIVE						0.004 (0.004)
TIME EFFECTS	YES	YES	YES	YES	YES	YES

R-squared <i>within</i>	0.582	0.656	0.465	0.311	0.293	0.384
R-squared <i>between</i>	0.080	0.494	0.680	0.221	0.224	0.667
R-squared <i>overall</i>	0.573	0.641	0.473	0.306	0.278	0.392
F-statistic	148.77	127.80	52.12	153.48	88.90	77.55
Number of obs.	1620	1185	435	1613	1178	435
Number of regions	82	67	34	82	67	34

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

The estimated coefficients are almost identical between output per capita and labor productivity regressions, but the formers are characterized by lower explanatory power.

The variables $|D_GRP|$ and $|D_LP|$ are not statistically significant, thus highlighting that differences between the regional levels of output per capita and labor productivity and their national averages do not explain variation in the growth rates of output per capita and labor productivity.

The estimated coefficients turn to be negative and statistically significant when the interaction terms $|D_GRP_NEGATIVE|$ and $|D_LP_NEGATIVE|$ are considered: therefore, there is not convergence in differences from below. Indeed, the negative estimated coefficient of the two variable implies that within the group of the poorest and less productive regions, the relationship between levels and growth is positive, therefore suggesting for a process of divergence in differences from below.

When the richer and more productive regions are taken into account, the estimated coefficients are not statistically significant. Economically, this result could be due to the fact that the richest and more productive regions have reached their steady state levels over the years (Note 10).

It is now worth to estimate the model with the addition of the control variables. The control variables I choose approximate the main conditioning factors suggested by the pioneering literature on convergence (Note 11). Even though the model specification differs from the conditional beta-convergence models proposed by the main literature, I believe that the conditioning factors employed in the conditional beta-convergence model are likely to improve the performance of the convergence in differences models.

Table 4. Convergence in differences (with controls). Yearly data. Fixed-effects estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	GRP_GROWTH	GRP_GROWTH	GRP_GROWTH	LP_GROWTH	LP_GROWTH	LP_GROWTH
Constant	0.189** (0.094)	0.459*** (0.080)	0.140 (0.165)	0.163* (0.090)	0.535*** (0.087)	0.127 (0.155)
$ D_GRP $	0.003 (0.010)					
$ D_GRP_NEGATIVE $		-0.149*** (0.020)				
D_GRP_POSITIVE			0.017 (0.013)			
$ D_LP $				-0.001 (0.007)		
$ D_LP_NEGATIVE $					-0.071*** (0.011)	
D_LP_POSITIVE						0.010 (0.010)
POP	-0.003 (0.036)	-0.000** (0.033)	0.023 (0.050)	0.019 (0.035)	-0.000*** (0.034)	0.024 (0.047)
CAP_PW	0.014 (0.010)	-0.043*** (0.010)	-0.022 (0.013)	0.080 (0.008)	-0.042*** (0.012)	-0.007 (0.012)
DEPR	-0.019 (0.018)	-0.000*** (0.023)	-0.030 (0.028)	-0.026 (0.016)	-0.044** (0.021)	-0.037 (0.024)
TIME EFFECTS	YES	YES	YES	YES	YES	YES
R-squared <i>within</i>	0.488	0.615	0.351	0.438	0.560	0.332
R-squared <i>between</i>	0.051	0.000	0.056	0.038	0.006	0.057
R-squared <i>overall</i>	0.461	0.171	0.190	0.414	0.076	0.177
F-statistic	93.99	84.77	49.27	81.60	69.22	62.10
Number of obs.	900	661	239	899	660	239
Number of regions	82	65	27	82	65	27

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

With the addition of control variables, there are not dramatic changes in estimates and their significance. Therefore, the estimates of the variables of our interest seem to be robust.

The significance of the variable $|D_LP_NEGATIVE|$ increases and, in general, the controls are statistically significant only when the poorer and less productive observations are considered.

There are therefore enough elements to state that being among the poorest does not imply performing better. The highest performances occur when the regions have an output level close to the national mean, denoting a possible non-linear impact of the level of output per capita and labor productivity on growth.

4. Concluding Remarks

The methodology implemented in this paper and its application to the case of the Russian regions put in light some interesting insights. Firstly, the process of beta-convergence between the Russian regions over the period 1995-2015 was not accompanied by a process of convergence in differences: conversely, divergence in differences occurred, consistently with the fact that the sigma-convergence process has slightly increased over the latest 5 years of the time interval. Secondly, the relationship between output growth and output level is non-linear: high distances between regional per capita output and average per capita output (both from below and above) are linked to the lowest growth rates, whereas as the distances decrease the growth rates increase. Thirdly, the above-mentioned non-linear relationship between per capita output growth and per capita output level could be due to nonlinearity in the returns to scale of physical capital: low levels of output per capita can be connected to decreasing returns to scale, middle levels of output per capita to increasing returns to scale and high levels of output per capita to decreasing returns to scale, depending on the evolution of the production structure of the regions. Fourthly, once the regional output per capita exceeds the national output per capita, the growth rate does not depend on the level of output, therefore suggesting that in such a situation the regions reached their steady state level.

An important pros of the convergence in differences regressions is the capability of considering the growth performance of each economy in the sample at each point in the time (and, therefore, at each level of output per capita). Indeed, the behaviour of a single economy is likely to differ across the time, and such differences underline interesting growth dynamics.

Thank to this approach, we have detected a non-linear behaviour of output per capita and labor productivity growth in a convergence framework.

The sample size and the way the core independent variables are constructed mitigate the endogeneity bias, therefore justifying the use of the fixed-effect estimator. Such a choice is strengthened by the likely presence of time invariant effects which differ across the groups, considering that Russia, given its high number of regions and its size, is characterized by strong internal structural heterogeneity. Nevertheless, different estimators applied to the equations 5, 6 and 7 can perform well when different space and time features of the sample are taken into account.

References

- Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. NJ: Princeton University Press.
- Aghion, P., & Durlauf, S. (2005). *Handbook of economic growth*. Elsevier. [https://doi.org/10.1016/S1574-0684\(05\)01206-2](https://doi.org/10.1016/S1574-0684(05)01206-2)
- Akhmedjonov, A., Chi Keung Lau, M., & Izgi, B. B. (2013). New Evidence of Regional Income Divergence in Post-Reform Russia. *Applied Economics*, 45(18), 2675-82. <https://doi.org/10.1080/00036846.2012.665600>
- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277-297. <https://doi.org/10.2307/2297968>
- Arrow, K. J. (1962). The Economic Implications of Learning by Doing. *The Review of Economic Studies*, 29(3), 155-73. <https://doi.org/10.2307/2295952>
- Badunenko, O., & Tochkov, K. (2010). Soaring dragons, roaring tigers, growling bears: Determinants of Regional Growth and Convergence in China, India, and Russia. *Economics of Transition*, 18(3), 539-570. <https://doi.org/10.1111/j.1468-0351.2009.00387.x>
- Barro, R. J. (1990). Government Spending in a Simple Model of Endogenous Growth. *Journal of SPolitical Economy*, 98, 103-125. <https://doi.org/10.1086/261726>
- Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics*,

- 106(2), 407-43. <https://doi.org/10.2307/2937943>
- Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy*, 100(2), 223-51. <https://doi.org/10.1086/261816>
- Barro, R. J., & Sala-i-Martin, X. (1995). *Economic Growth*. New York: McGraw- Hill. Retrieved from <https://EconPapers.repec.org/RePEc:eee:dyncon:v:21:y:1997:i:4-5:p:895-898>
- Baumol, W. J. (1986). Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show. *The American Economic Review*, 1072-85. Retrieved from <https://www.jstor.org/stable/1816469>
- Benini, R., & Czyzewski, A. (2007). Regional Disparities and Economic Growth in Russia: Net Growth Patterns and Catching Up. *Economic Change and Restructuring*, 40(1-2), 91-135. <https://doi.org/10.1007/s10644-007-9026-0>
- Berkowitz, D., & Dejong, D. N. (2002). Accounting for Growth in Post-Soviet Russia. *Regional Science and Urban Economics*, 32, 221-39. [https://doi.org/10.1016/S0166-0462\(01\)00077-1](https://doi.org/10.1016/S0166-0462(01)00077-1)
- Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143. [https://doi.org/10.1016/S0304-4076\(98\)00009-8](https://doi.org/10.1016/S0304-4076(98)00009-8)
- Carluer, F. (2005). Dynamics of Russian Regional Clubs: The Time of Divergence. *Regional Studies*, 39(6), 713-726. <https://doi.org/10.1080/00343400500213564>
- Dechert, W. D., & Nishimura, K. (1983). A Complete Characterization of Optimal Growth Paths in an Aggregated Model with a Non-Concave Production Function. *Journal of Economic Theory*, 31, 332-354. https://doi.org/10.1007/978-3-642-22397-6_10
- Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic Competition and Optimum Product Diversity. *American Economic Review*, 67, 297-308.
- Dolinskaya, I. (2002). Transition and Regional Inequality in Russia: Reorganisation or Procrastination? *IMF Technical Report 02/069*, Washington, DC. <https://doi.org/10.5089/9781451858358.001>
- Durand-Lasserve, O., & Blöchliger, H. (2018). Drivers of Regional Growth in Russia: A Baseline Model with Application. *OECD Economics Department Working Papers No. 1523*. <https://doi.org/10.1787/9279f6c3-en>
- Fedorov, L. (2002). Regional Inequality and Regional Polarization in Russia, 1990-99. *World Development*, 30(3), 443-56. [https://doi.org/10.1016/S0305-750X\(01\)00124-3](https://doi.org/10.1016/S0305-750X(01)00124-3)
- Guriev, S., & Vakulenko, E. (2012). Convergence Between Regions. *Cefir Working Paper No.180*. Retrieved from <https://econpapers.repec.org/RePEc:cfr:cefirw:w0180>
- Islam, N. (1995). Growth empirics: A panel data approach. *The Quarterly Journal of Economics*, 110(4), 1127-1170. <https://doi.org/10.2307/2946651>
- Jones, L., Manuelli, R., & Stacchetti, E. (2000). *Technology and Policy Shocks in Models of Endogenous Growth*. Federal Reserve Bank of Minneapolis, Working Paper 281.
- Kaneva, M., & Untura, G. (2019). The impact of R&D and knowledge spillovers on the economic growth of Russian regions. *Growth and Change*, 50(1), 301-334. <https://doi.org/10.1111/grow.12281>
- Kremer, M. (1993). Population Growth and Technological Change: One Million BC to 1990. *The Quarterly Journal of Economics*, 108(3), 681-716. <https://doi.org/10.2307/2118405>
- Lehmann, H., Oshchepkov, A. Y., & Silvagni, M. (2020). Regional Convergence in Russia: Estimating a Neoclassical Growth Model. *IZA Discussion Papers, No. 13039*. <https://doi.org/10.2139/ssrn.3627173>
- Lucas Jr, R. E. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22(1), 3-42. [https://doi.org/10.1016/0304-3932\(88\)90168-7](https://doi.org/10.1016/0304-3932(88)90168-7)
- Mankiw, N. G., Phelps, E. S., & Romer P. M. (1995). The Growth of Nations. *Brookings Papers on Economic Activity 1995*, (1), 275-326. <https://doi.org/10.2307/2534576>
- Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107(2), 407-37. <https://doi.org/10.2307/2118477>
- Mirkina, I. (2017). *Regions of Russia in Comparative Perspective: Introducing a New Dataset*. <http://doi.org/10.2139/ssrn.2969950>
- Nordhaus, W. D. (1969). An Economic Theory of Technological Change. *The American Economic Review*, 59(2), 18-28. Retrieved from <https://econpapers.repec.org/RePEc:aea:aecrev:v:59:y:1969:i:2:p:18-28>

- Oshchepkov, A. (2015). Compensating Wage Differentials Across Russian Regions. *AIEL Series in Labour Economics: Geographical Labor Market Imbalances*. Berlin: Springer. https://doi.org/10.1007/978-3-642-55203-8_4
- Panizza, U., & Presbitero, A. F. (2013). Public debt and economic growth in advanced economies: A survey. *Swiss Journal of Economics and Statistics*, 149(2), 175-204. <https://doi.org/10.1007/BF03399388>
- Popov, V. (2001). Reform Strategies and Economic Performance of Russia's Regions. *World Development*, 29(5), 865-886. [https://doi.org/10.1016/S0305-750X\(01\)00006-7](https://doi.org/10.1016/S0305-750X(01)00006-7)
- Romer, D. (2006). *Advanced Macroeconomics* (3rd ed.). New York: McGraw- Hill.
- Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy*, 94(5), 1002-37. <https://doi.org/10.1086/261420>
- Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, 9(1), 86-136. <https://doi.org/10.1177/1536867X0900900106>
- Sala-i-Martin, X. (1996). The Classical Approach to Convergence Analysis. *The Economic Journal*, 1019-36. <https://doi.org/10.2307/2235375>
- Shell, K. (1973). Inventive Activity, Industrial Organisation and Economic Growth. *Models of Economic Growth*, 77-100. https://doi.org/10.1007/978-1-349-01650-1_4
- Skiba, A. K. (1978). Optimal growth with a convex-concave production function. *Econometrica: Journal of the Econometric Society*, 527-539. <https://doi.org/10.2307/1914229>
- Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics*, 70(1), 65-94. <https://doi.org/10.2307/1884513>
- Uzawa, H. (1965). Optimum Technical Change in an Aggregative Model of Economic Growth. *International Economic Review*, 6(1), 18-31. <https://doi.org/10.2307/2525621>
- Vakulenko, E. (2016). Does Migration Lead to Regional Convergence in Russia? *International Journal of Economic Policy in Emerging Economies*, 9(1), 1-24. <https://doi.org/10.1504/IJEPEE.2016.074943>

Notes

Note 1. Among the most relevant empirical papers we can mention the seminal works of Mankiw, Romer, and Weil (1992) and Barro and Sala-i-Martin (1995).

Note 2. The conditioning factors are generally approximated by saving rate, technology level, human capital and population. See for example Mankiw, Romer, and Weil (1992).

Note 3. This process strongly relies on the assumption of decreasing returns to scale of capital.

Note 4. Generally, the conditioning factors are approximated by the saving rate, the human capital and the population growth rate (Romer, 1986; Lucas, 1988; Mankiw et al., 1992; Barro & Sala-i-Martin, 1995).

Note 5. For example, Sala-i-Martin (1996) emphasized the importance of the technology heterogeneity between countries.

Note 6. In order to measure the dispersion, it is generally employed the coefficient of variation, which is defined

as follows: $CV = \frac{\sigma_{y_t}}{\bar{y}_t} = \frac{\sqrt{\frac{1}{N} \sum_{i=0}^N (y_{it} - \bar{y}_t)^2}}{\bar{y}_t}$ where y is the real output per capita. \bar{y} represents the average output per capita, t and i denote, respectively, the time and the i -th economy.

Note 7. Such characteristics can be, for example, the following: monetary policy, currency, culture, fiscal policy, institutional framework, external economic policy, inflation rate, unobservable national-level heterogeneity etc.

Note 8. It is reasonable to assume that the growth dynamics can dramatically differ between the groups of poorest and richest economies.

Note 9. For the data collection purposes, *Nenets Autonomous Okrug* is treated as part of *Arkhangelsk Oblast*. The *Republic of Crimea* and the city of *Sevastopol* are not included in the dataset. Such dataset includes official statistics from: Federal Statistical Service of Russia (*Rosstat*); Ministry of Interior of Russia; Federal Treasury of Russia; Central Bank (Bank of Russia); Federal Tax Service of Russia; Ministry of Finance of Russia; Ministry for Foreign Affairs of Russia; Federal Customs Service of Russia; Ministry for Civil Defense, Emergencies, and

Elimination of Consequences of Natural Disasters of Russia; Ministry for Regional Development of Russia.

Note 10. Indeed, consistently with the neoclassical theories of growth, in the steady state level the growth rate does not depend on the levels of output per capita.

Note 11. See for example Barro (1991), Mankiw et al. (1992), Barro and Sala-i-Martin (1992).

Appendix

Variables	Label
GRP_GROWTH	Growth rate of real GRP per capita
LP_GROWTH	Growth rate of labor productivity
GRP_1995	Real GRP per capita at year 1995 (millions of rubles)
D_GRP_95	Difference between regional output per capita and national output per capita at year 1995 (absolute value)
LP_95	Labor productivity at year 1995
D_LP_95	Difference between regional labor productivity and national labor productivity at year 1995 (absolute value)
D_GRP	Difference between regional output per capita and national output per capita (absolute value)
D_GRP_NEGATIVE	Interaction term between $ D_GRP $ and the indicator function which takes value one for observation whose level of real GRP per capita is below the national average and zero otherwise
D_GRP_POSITIVE	Interaction term between $ D_GRP $ and the indicator function which takes value one for observation whose level of real GRP per capita is above the national average and zero otherwise
D_LP	Difference between regional labor productivity and national labor productivity (absolute value)
D_LP_NEGATIVE	Interaction term between $ D_LP $ and the indicator function which takes value one for observation whose level of labor productivity is below the national average and zero otherwise
D_LP_POSITIVE	Interaction term between the $ D_LP $ and the indicator function which takes value one for observation whose level of labor productivity is above the national average and zero otherwise
POP	Population (thousands)
CAP_PW	Stock of physical capital per worker (millions of rubles)
DEPR	Depreciation of physical capital (millions of rubles)

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).