Nonlinear Models of Convergence

Gluschenko, Konstantin

Novosibirsk State University, Institute of Economics and Industrial Engineering of the Siberian Branch of the Russian Academy of Sciences

28 March 2020

Nonlinear Models of Convergence

Konstantin Gluschenko[1, 2]

[1] Novosibirsk State University, Novosibirsk, Russia
[2] Institute of Economics and Industrial Engineering of the SB RAS, Novosibirsk, Russia
glu@nsu.ru

Abstract. A sufficient issue in studies of economic development is whether economies (countries, regions of a country, etc.) converge to one another in terms of per capita income. In this paper, nonlinear asymptotically subsiding trends of income gap in a pair of economies model the convergence process. A few specific forms of such trends are proposed: log-exponential trend, exponential trend, and fractional trend. A pair of economies is deemed converging if time series of their income gap is stationary about any of these trends. To test for stationarity, standard unit root tests are applied with non-standard test statistics that are estimated for each kind of the trends.

Keywords: Income Convergence, Time Series Econometrics, Nonlinear Time-Series Model, Unit Root.

1 Introduction

A sufficient issue in studies of economic development is whether economies (countries, regions of a country, cities, etc.) converge to one another in terms of per capita income. There are a number of methodologies to test for the convergence hypothesis. The most widespread one in the literature is the analysis of a negative cross-section correlation between initial per capita income and its growth, the so-called beta-convergence (see, e.g., [1]). An alternative methodology is the distribution dynamics analysis that explores the evolution of cross-economy income distribution [2]. Both approaches provide only an aggregated characterization of convergence. If the whole set of economies under consideration is found to converge, it is not possible to reveal economies with a deviant behavior (e.g., diverging or randomly walking). On the other hand, if the convergence hypothesis is rejected, it is not able to detect a subset (or subsets) of converging economies.

Methodologies based on time-series analysis make it possible to overcome this problem. They consider time series of the income gap, i.e., the difference of logarithms of per capita incomes in a pair of economies r and s, $y_{rst} = y_{rt} - y_{st} = \ln(Y_{rt}/Y_{st})$, t denoting time. To discriminate between logarithmic and real (e.g., percentage) terms, $Y_{rt}/Y_{st} - 1$ is called income disparity. One element of the pair can be an aggregate, for instance, the national economy when economies under consideration are country's regions.

Bernard and Durlauf [3] have put forward a formal definition of convergence: economies r and s converge if the long-term forecasts of per capita income (conditionally on information available by the moment of the forecast, I) for both economies are equal, that is

$$\lim_{t\to\infty} \mathrm{E}\big(y_{rst} \,|\, I\big) = 0. \tag{1}$$

Despite this definition of convergence is general, procedures of testing for convergence applied in [3] in fact detect only a particular class of processes satisfying (1), namely, stationary processes with no trend (implying that y_{rt} and y_{st} have a common trend). Thus, such procedures are not able to classify the most interesting case of catching-up as convergence.

As a way out, [4] proposes to model the (square of) income gap by a trend $h(t)$ of an a priory unknown form, approximating it by a power series of degree k. The respective econometric model looks like (ε_t denotes residuals with standard properties, α_i is a coefficient to be estimated):

$$y_{rst}^2 = h(t;k) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_k t^k + \varepsilon_t \; (t = 1, \ldots, T). \tag{2}$$

Albeit the trend may be nonlinear, equation (2) is linear with respect to coefficients. Convergence takes place if $dh/dt < 0$ holds for all t. This condition is supposed to be equivalent to negativity of the time average of $dh(t)/dt$:

$$\frac{1}{T}\sum_{t=1}^{T}\frac{dh}{dt} = \sum_{i=1}^{k}\alpha_i \frac{i}{T}\sum_{t=1}^{T} t^{i-1} < 0. \tag{3}$$

However, the equivalence is not the fact. It is obvious, considering a continuous-time counterpart of (3):

$$\frac{1}{T}\int_1^T \frac{dh}{dt}dt = \frac{1}{T}\big(h(T) - h(1)\big) < 0.$$

Hence, the mere fact that $h(T) < h(1)$ suffices to accept the convergence hypothesis. In the general case, this does not evidence convergence. For instance, a U-shape path of the income gap may satisfy (3). Moreover, even if $dh/dt < 0$ is true for every $t = 1,\ldots T$, condition (1) knowingly does not hold, as $h(\infty; k) = \pm\infty$ for any finite k.

Thus, there is a want of developing an alternative methodology. This paper puts forward such a methodology, namely, modeling the convergence process by asymptotically subsiding trends. This leads to nonlinear econometric models that need nonstandard distributions of test statistics to test models for unit roots.

2 Modeling Convergence

Actual convergence processes are in fact a superposition of two processes that can be called long-run, or deterministic, convergence, and stochastic, or short-run, convergence. Long-run convergence is a deterministic path of the income gap y_{rst} that tends to zero over time: $y_{rst}^* = h(t), h(t) \underset{t\to\infty}{\longrightarrow} 0$. In [4], only this process is considered (albeit with no latter condition). Short-run convergence is an autocorrelated stochastic process containing no unit root (i.e., a stationary process), $\nu_t = \rho\nu_{t-1} + \varepsilon_t$, where ρ is the autocorrelation coefficient, $\rho < 1$, and $\varepsilon_t \sim N(0, \sigma^2)$ with finite σ. Intuitively, short-run convergence characterizes the behavior of transient random shocks. A unit shock deviates the income gap from its long-run path, dying out over time with half-life θ = ln(0.5)/ln(ρ), so that the income gap eventually returns to its long-run path. Only such processes are considered in [3] (assuming $y^*_{rst} = 0$).

Superposition of these two processes gives a process that is stationary around an asymptotically subsiding trend h(t). That is, albeit random shocks force the process to deviate from the trend, it permanently tends to return to the trend, thus satisfying (1). The following econometric model of the class AR(1) describes such a process:

$$y_{rst} = h(t) + \nu_t \; (t = 0, \dots, T-1),\, \nu_t = \rho\nu_{t-1} + \varepsilon_t \; (t = 1, \dots, T-1;\; \nu_0 = \varepsilon_0).$$

Applying the Cochrane-Orcutt transformation to this equation, the following model is arrived at:

$$\Delta y_{rst} = h(t) - (\lambda + 1)h(t-1) + \lambda y_{rs,t-1} + \varepsilon_t \; (t = 1, \dots, T-1), \tag{4}$$

where $\Delta y_{rst} = y_{rst} - y_{rs,t-1}$ and $\lambda = \rho - 1$.

To make model (4) operational, a specific function $h(t)$ has to be taken from the class of asymptotically subsiding functions. A few such functions are preferable in order to model more adequately properties of a process under consideration. The following three functions seem convenient from the practical viewpoint: log-exponential trend $h(t) = \ln(1 + \gamma e^{\delta t})$, $\delta < 0$, exponential trend $h(t) = \gamma e^{\delta t}$, $\delta < 0$, and fractional trend $h(t) = \gamma/(1 + \delta t)$, $\delta > 0$. The respective models are nonlinear with respect to coefficients, having the forms:

$$\Delta y_{rst} = \ln\left(1 + \gamma e^{\delta t}\right) - (\lambda + 1)\ln(1 + \gamma e^{\delta(t-1)}) + \lambda y_{rs,t-1} + \varepsilon_t\,; \tag{4a}$$

$$\Delta y_{rst} = \gamma e^{\delta t} - (\lambda + 1)\gamma e^{\delta(t-1)} + \lambda y_{rs,t-1} + \varepsilon_t\,; \tag{4b}$$

$$\Delta y_{rst} = \frac{\gamma}{1+\delta t} - (\lambda + 1)\frac{\gamma}{1+\delta(t-1)} + \lambda y_{rs,t-1} + \varepsilon_t\,. \tag{4c}$$

An advantage of the log-exponential trend is the ease of interpretation. Parameter γ is the initial (at $t = 0$) income disparity. Parameter δ characterizes the convergence rate which can be simply expressed in terms of the half-life time of the (deterministic) income disparity, i.e., the time the disparity takes to halve: $\Theta = \ln(0.5)/\delta$. A shortcoming of this trend is in that is has no symmetry properties with respect to a permutation of the economy indices. Albeit $y_{rst} = -y_{srt}$, the permutation changes absolute values of γ and δ (and may change the estimate of λ in regression (4a)).

Contrastingly, the exponential and fractional trends have symmetry properties. A permutation of r and s changes only the sign of γ, leaving its absolute value and the value of δ (as well as λ in (4b), (4c)) intact. However, while is the initial income disparity can be easily calculated from γ, equaling $e^{\gamma} - 1$ in both trends, the half-life of the deterministic income gap involves a mixture of γ and δ. This results in hardly interpretable expressions. For the exponential trend, $\Theta = \frac{1}{\delta}\ln(\frac{\ln(0.5(e^{\gamma}+1))}{\gamma})$; for the fractional trend, $\Theta = \frac{1}{\delta}(\frac{\gamma}{\ln(0.5(e^{\gamma}+1))} - 1)$.

Models (4a)–(4c) are also applicable to the case of deterministic divergence. It takes place if $\delta > 0$ in the log-exponential and exponential trends, or $\delta < 0$ in the fractional trend. The time the (deterministic) income disparity takes to double can characterize the divergence rate.

Model (4) encompasses two particular cases. With $h(t) = 0$, which corresponds to $\gamma = 0$ in (4a)–(4c), it degenerates to ordinary AR(1) model with no constant:

$$\Delta y_{rst} = \lambda y_{rs,t-1} + \varepsilon_t\,. \tag{5}$$

This implies that series y_{rt} and y_{st} are cointegrated with cointegrating vector [1, –1], i.e., they have the same trend. Intuitively, this means that convergence as such, i.e., catching-up, has completed by $t = 0$ (if it had occurred before). In the further dynamics, per capita incomes in economies r and s are equal up to random shocks (hence, only stochastic convergence takes place).

With $h(t)$ = const, which corresponds to $\delta = 0$ in (4a)–(4c), model (4) degenerates to ordinary AR(1) model with a constant:

$$\Delta y_{rst} = \alpha + y_{rs,t-1} + \varepsilon_t\,. \tag{6}$$

This implies that series y_{rt} and y_{st} are cointegrated with cointegrating vector $[1, -\gamma]$, i.e., they have a common trend: $h_s(t) = \gamma + h_r(t)$, $\gamma = -\alpha/\lambda$. In other words, the income gap is constant (up to random shocks); y_{rt} and y_{st} move parallel to each other with distance between their paths equaling γ. Again, only stochastic convergence takes place here. Just models (5) and (6) are considered in [3] (albeit within a more evolved framework).

Having estimated parameters of a specific model of the form (4), we need to check its adequacy. First of all, the question is whether y_{rst} is indeed stationary around the given trend (y_{rst} has no unit root). There are a number of tests for unit root (testing hypothesis $\lambda = 0$ against $\lambda < 0$, or $\lambda < 0$ against $\lambda = 0$). Most of them use t-ratio of λ, $\tau = \lambda/\sigma_\lambda$, as the test statistic. In the case of testing for unit roots, it has non-standard distributions, differing from the t-distribution (that is why it is designated τ, and not t). Such distributions (named the Dickey-Fuller distributions) are tabulated for AR(1) models with no constant, with a constant, and with a linear and quadratic trends, but not for models with proposed nonlinear trends. To estimate them, τ in every model with a specific trend was estimated for each of 1 million generated random walks $y_t = y_{t-1} + \varepsilon_t$. Table 1 reports some values of the τ-statistic from the obtained distributions for sample size $T = 204$ (used in the empirical analysis reported in the next section). Fig. 1 plots the 10-percent tails of the distributions, comparing them with the Dickey-Fuller distributions for the cases of linear and quadratic trend from [6].

Table 1. Selected values of the τ-statistics for models with nonlinear trends, $T = 204$.

Probability	Log-exponential trend (4a)	Exponential trend (4b)	Fractional trend (4c)
1%	–3.841	–3.851	–5.152
5%	–3.220	–3.273	–3.820
10%	–2.898	–2.971	–3.297

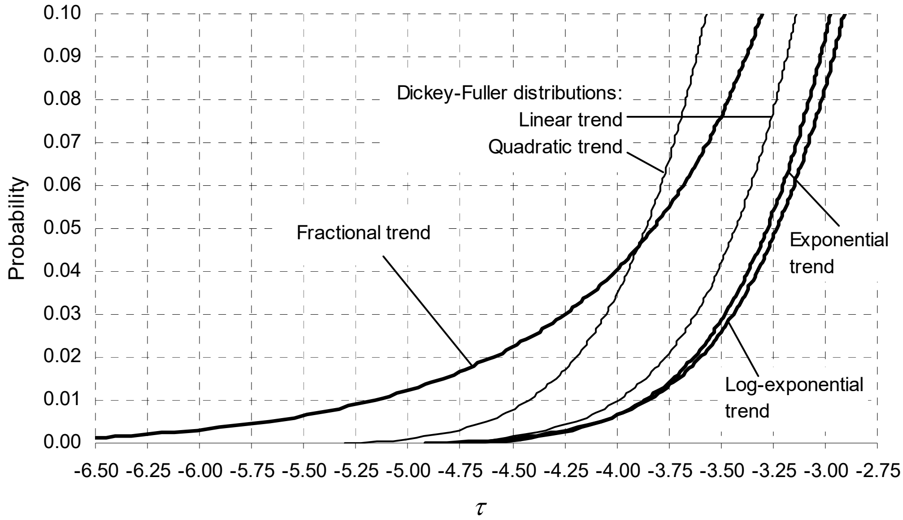

Fig. 1. Distributions of the unit root test τ-statistics for equations (4a)–(4c) and selected Dickey-Fuller distributions; T =204.

If the unit root test rejects the hypothesis of non-stationarity, the ordinary t-test can test parameters γ and δ for statistical significance. Given that there are three versions of model (4), every version is estimated and tested. If they turn out to be completive, the version providing the best fit – namely, the minimal sum of squared residuals (SSR) – is accepted. Note that valid models with the "incorrect" sign of δ suggest deterministic divergence. The rejection of all versions because of presence of unit root or insignificance of γ or δ evidences the absence of (deterministic) convergence as well. If statistical reasons for no-convergence are of interest, we can estimate and test regression (6) and then, if it is rejected, regression (5). In this case, we find whether no-convergence is due to coinciding or "parallel" dynamics of per capita incomes in a pair of economies under consideration (the same or common trend), or – if both models are rejected – it is due to a random walk.

3 Empirical Application

This section provides an illustration of empirical application of the proposed methodology to analyzing convergence of regional incomes per capita in Russia. The time span covers January 2002 through December 2018 with a monthly frequency (204 months). The indicator under consideration is real personal income per capita by region. The term "real" means that the income is adjusted to the respective regional price level. The cost of the fixed basket of goods and services for cross-region comparison of population's purchasing capacity serves as an indicator of regional price level. The official statistical data on nominal incomes and the costs of the fixed basket come from [6–8].

Convergence is considered with respect to the national income per capita. Thus, index s is fixed, denoting Russia as a whole; then y_{rst} is the gap between regional and national incomes. To test models for unit roots, the Phillips-Perron test (PP test) is applied with modifications proposed in [9, 10].

Since the whole set of results is cumbersome (involving 79 regions), this sections gives them only partially for illustrative purposes. It presents examples of qualitatively different cases discussed in the previous section. Table 2 reports these.

Table 2. Selected results of analyzing regional convergence in Russia.

Model	λ	PP-test p-value	γ / α in (6)	p-value of γ/α	δ	p-value of δ	SSR
			Kursk Oblast				
(4a)	–0.484 (0.061)	0.000	–0.354 (0.018)	0.000	–0.011 (0.001)	0.000	0.550
(4b)	–0.496 (0.062)	0.000	–0.430 (0.025)	0.000	–0.013 (0.001)	0.000	0.546
(4c)	–0.361 (0.054)	0.000	–0.493 (0.066)	0.000	0.029 (0.008)	0.000	0.592
			Republic of Karelia				
(4a)	–0.457 (0.059)	0.000	–0.100 (0.012)	0.000	0.005 (0.001)	0.000	0.680
(4b)	–0.462 (0.059)	0.000	–0.103 (0.013)	0.000	0.005 (0.001)	0.000	0.679
(4c)	–0.423 (0.057)	0.000	–0.122 (0.013)	0.000	–0.003 (0.000)	0.000	0.695
			Saint Petersburg City				
(4a)	–0.427 (0.058)	0.000	0.236 (0.035)	0.000	–0.001 (0.001)	0.287	
(4b)	–0.427 (0.058)	0.000	0.212 (0.028)	0.000	–0.001 (0.001)	0.288	
(4c)	–0.427 (0.058)	0.000	0.212 (0.030)	0.000	0.001 (0.002)	0.365	
(6)	–0.419 (0.057)	0.000	0.078 (0.012)	0.000			
			Republic of Bashkortostan				
(4a)	–0.359 (0.053)	0.000	0.018 (0.033)	0.576	0.000 (0.015)	0.976	
(4b)	–0.359 (0.053)	0.000	0.018 (0.032)	0.573	0.000 (0.015)	0.976	
(4c)	–0.359 (0.053)	0.003	0.019 (0.032)	0.564	0.000 (0.014)	0.985	
(6)	–0.359 (0.053)	0.000	0.006 (0.005)	0.249			
(5)	–0.317 (0.052)	0.000					
			Moscow Oblast				
(4a)	–0.211 (0.043)	0.076	0.018 (0.016)	0.264	0.013 (0.005)	0.012	
(4b)	–0.209 (0.043)	0.094	0.019 (0.016)	0.246	0.013 (0.005)	0.013	
(4c)	–0.180 (0.040)	0.262	0.041 (0.019)	0.029	–0.004 (0.001)	0.000	
(6)	–0.125 (0.034)	0.357	0.010 (0.005)	0.043			
(5)	–0.091 (0.030)	0.116					

Standard errors are in parentheses.

Convergence manifests itself in the Kursk Oblast. All three versions of the trend model can be accepted, suggesting fast convergence. Choosing model (4b) as providing the best fit, the half-life time of the income gap equals 5.3 years (65.3 months). Fig. 2(a) plots the path of the actual income gap and its estimated exponential trend. According to this trend, income per capita in the Kursk Oblast was below the national level by 35% at the beginning of the time span under consideration and by only 3% by its end. The log-exponential and fractional trends suggest even faster convergence with half-live times 5.1 and 3.6 years, respectively.

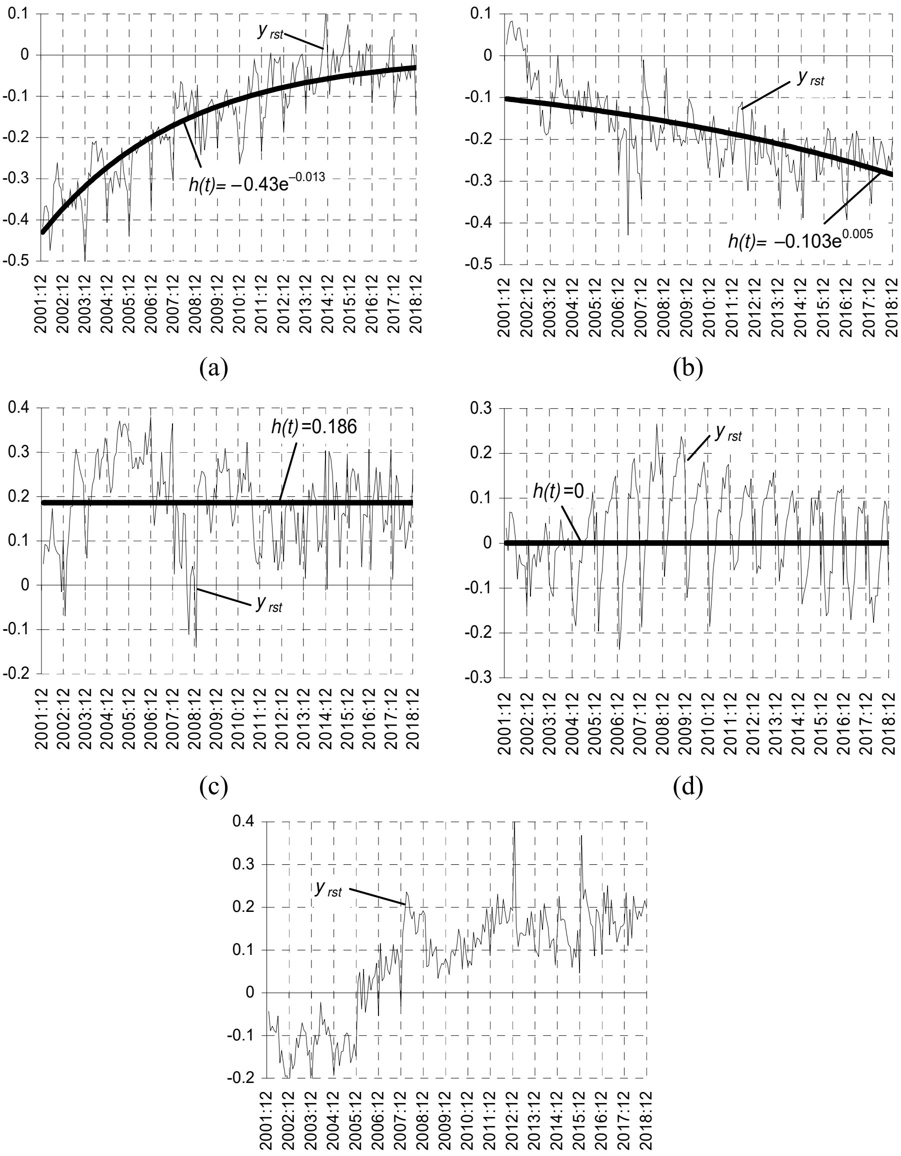

Fig. 2. Different cases of behavior of the income gap: (a) convergence (the Kursk Oblast); (b) divergence (Republic of Karelia); (c) a constant income gap (Saint Petersburg City); (d) no income gap (Republic of Bashkortostan); (d) random walking of income gap (the Moscow Oblast).

Divergence occurs in the Republic of Karelia. Again, all three versions of the trend model can be accepted. Model (4b) seems preferable, albeit its SSR differs from the SSR in model (4a) only slightly. Fig. 2(b) depicts the dynamical pattern. The income gap rises, doubling every 10.4 years. Income per capita in this region was 9% below the national level in January 2002 and 28% in December 2018.

The case of Saint Petersburg City (which is a separate administrative-territorial unit considered as a region) illustrates the absence of convergence that is due to “parallel” dynamics of the national and regional incomes per capita. Fig. 2(c) shows this case. Although the unit root test rejects the hypothesis of nonstationarity with confidence in all trend models, high *p*-values of δ in all of suggest the absence of a trend. Model (6) proves to be valid, implying the income gap to be time-invariant. It equals 0.186 ($= -\alpha/\lambda$); in other words, real income per capita in Saint Petersburg City remains on average constant, being 20.5% above the national level.

The Republic of Bashkortostan demonstrates a similar pattern, Fig. 2(d), with difference that there is no income gap; real income per capita here remains on average equal to the national per capita income (in fact, the regional income fluctuates around the national level). In all trend models, *p*-values of both γ and δ are high, thus implying rejection of these models. The constant in model (6) has high *p*-value as well, which leads to model (5). It proves to be valid, the unit root hypothesis being rejected with confidence.

At last, no one model seems to describe the behavior of the income gap in the Moscow Oblast, Fig. 2(e). We can reject models (4a) and (4b) because of high *p*-value of γ, and models (6) and (5) because of non-rejection of a unit root. The conclusion may be that non-convergence here is due to a random walk of the income gap.

Briefly summing up the results of the full analysis of income convergence in Russia, convergence takes place in the whole of Russia, as the Gini index decreases over time. Analysis by region yields the “anatomy” of convergence. Among all 79 regions in the spatial sample, 44 regions (55.7%) are converging. In 16 regions (20.3%), non-convergence is due to common trends with the national income per capita (in three cases, regional trends coincide with the national trend). An unpleasant feature of the pattern obtained is a considerable number of diverging regions; there are 17 of them (21.5%). Besides, random walks are peculiar to two regions.

4 Conclusion

This paper develops a methodology of modeling convergence by asymptotically subsiding trends of income gap in a pair of economies. This way conforms to the theoretical definition of convergence. Three specific kinds of such trends are proposed, namely, log-exponential trend, exponential trend, and fractional trend. This makes it possible to select a specific model that most adequately describes properties of an actual dynamics.

Transformation to testable versions generates nonlinear econometric models that represent a superposition of stochastic and deterministic convergence. Such models need additional efforts: the application of methods for estimation of nonlinear regressions and estimating distributions of the unit root test statistics for every specific trend.

However, these efforts are repaid, providing a theoretically adequate and practically fairly flexible and helpful tool for studying processes of convergence between countries, regions within a country, regions of different countries (e.g., in the European Union), etc.

The reported examples of applying the proposed methodology to the empirical analysis of convergence of real incomes per capita between Russian regions show that the results obtained look reasonable and correspond to economic intuition. As regards the whole analysis, it has yielded an interesting pattern. In spite of the fact that convergence occurs in Russia as a whole, a deviant dynamics is peculiar to a number of regions: almost a quarter of regions are found to diverge, either deterministically or stochastically.

References

1. Barro, R.J., Sala-i-Martin, X.: Convergence. Journal of Political Economy 100 (2), 223–251 (1992).
2. Quah D.: Galton's fallacy and tests of the convergence hypothesis. Scandinavian Journal of Economics 95 (4), 427–443 (1993).
3. Bernard, A. B., Durlauf, S. N.: Convergence in international output. Journal of Applied Econometrics 10 (2), 97–108 (1995).
4. Nahar, S., Inder, B.: Testing convergence in economic growth for OECD countries. Applied Economics 34 (16), 2011–2022 (2002).
5. MacKinnon, J. G.: Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics*, 11(6), 601–618 (1996).
6. Monthly bulletin "Socio-Economic situation of Russia", Rosstat, Moscow (2002–2007, various issues). [In Russian].
7. Rosstat Homepage, http:// www.gks.ru/free_doc/new_site/population/urov/urov_11sub2008.xls, http://www.gks.ru/free_doc/new_site/population/urov/urov_11sub09-14.xls, http://www.gks.ru/free_doc/new_site/population/urov/2013-2015kv.xls, last accessed 2016/11/10. [In Russian].
8. EMISS Homepage, https://fedstat.ru/indicator/57039, https://fedstat.ru/indicator/31052, last accessed 2020/01/26. [In Russian].
9. Perron, P., Ng, S.: Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. Review of Economic Studies, 63(3), 435–463 (1996).
10. Ng, S., Perron, P.: Lag length selection and the construction of unit root tests with good size and power. Econometrica, 69(6), 1519–1554 (2001).

Prepared for the International Conference
"Mathematical Optimization Theory and Operations Research"
(MOTOR 2020)
July 6-10, 2020, Novosibirsk, Russia