EVALUATION OF MACROECONOMIC DYNAMICS USING GENERALIZED MODIFIED PRINCIPAL COMPONENTS ANALYSIS

Alexei Verenikin – Anna Verenikina

Abstract

We propose a composite indicator of social and economic development alternative to GDP. It is based on the means of generalized modified principal component analysis (GMPCA) applied to temporal structure of various indicators of national social and economic activities. The weights of subindices that reflect various aspects macroeconomic pattern and constitute the aggregate indicator are generated in an automatic mode by GMPCA. This approach doesn't require any expert assessments, so it is a kind of objective estimation the determinants of social and economic development. The proposed methodology is applied to evaluation of the overall performance of Russian economy using global competitiveness indicators for the period from 2007 to 2017. It allows to reveal the factors which affect macrodynamics and consolidate them into the groups with positive or negative impact. The obtained composite indicator reveals cyclical fluctuations along the trend of economic development. The decomposition of the aggregate indicator shows the structural shifts in the Russian macroeconomic performance in the past decade.

Key words: composite indicator of social and economic development, principal components analysis

JEL Code: C38, E66, O47

Methodology, data and analysis

GDP has proved to be an inappropriate measure of economic development (Stiglitz, Sen, and Fitoussi, 2009; Noll, 2011; Giannetti, et. al, 2015; Veenhoven, 1996). National social and economic performance is a multidimensional characteristic that comprises a variety of economic indicators $X = \{x_i\}_{i=1}^{n}$. There is an acute need for a composite measure of a national social and economic performance that will not be sensitive to subjective preferences concerning the relative significance of specific features of economic activity. The question is how to aggregate partial measures using appropriate weighting coefficients that will not rely on subjective judgments.

Principal component analysis (PCA) is a well known multivariate statistical technique that is widely used in social indicators research (Böhringer and Jochem, 2007; Booysen, 2002; Ram, 1982; Slottje, 1991; Mazziotta and Pareto, 2019; Fan, Wang and Zhu, 2013; Doukas, et. al., 2012). It is the means to transform a set of original correlated variables $X = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix}$, where each, j-th $(j = 1, \dots, m)$ column represents the data vector that corresponds to the particular year, into a set of artificial uncorrelated variables $Z_1, Z_2, \dots, Z_n$, where $Z_k = (z_{k1}, \dots, z_{km})$ is the k-th principal component vector, $k = 1, \dots, n$. Principal components form an orthogonal normalized system of linear combinations of original statistical variables $Z = \begin{pmatrix} z_{11} & \cdots & z_{1m} \\ \vdots & \ddots & \vdots \\ z_{n1} & \cdots & z_{nm} \end{pmatrix} = \begin{pmatrix} l_{11} & \cdots & l_{1n} \\ \vdots & \ddots & \vdots \\ l_{n1} & \cdots & l_{nn} \end{pmatrix} \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix}$ that retains their total variation. Principal components are ranked in accordance with the share of comprised variance of the available data.

Principal component loadings are determined as eigenvectors of the covariance matrix $\Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix}$ of the original data, where $\sigma_{pq} = \sum_{j=1}^{m} (x_{pj} - \bar{x}_p)(x_{qj} - \bar{x}_q)$ is the covariance of the p-th and q-th environmental impact indicators. The eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$ of the matrix Σ are equal to the variance of the corresponding principal component scores $Z_1, Z_2, \dots, Z_n$.

Total variance of principal component scores equals to the total variance of primary data, thus $\rho_k = \lambda_k \Big/ \sum_{k=1}^{n} \lambda_k$ is the share of data variance explained by the k-th principal component.

The first principal component score $z_{1j} = \sum_{i=1}^{n} l_{1i} x_{ij}$ is known to be used as an aggregate indicator of economic performance (corresponding to the j-th period). This approach is characterized by transparent and accurate methodology that is based solely on statistical data, excludes any expert estimates (Booysen, 2002) and guarantees reliability of the resulting indicators.

Unfortunately the first principal component is not an appropriate measure in a typical situation when the proportion of total factors' variance explained by the first principal component is not sufficiently large which is particularly the case of the companies' environmental impact. Thus in this case the first principal component approach implies significant loss in variance of initial factors that are taken into consideration.

The research is based on the methodology verified in previous paper of the authors (Verenikin and Verenikina, 2018). We use the generalized modified principal component approach that consists in calculation of an aggregate measure of national economic activity as a weighted sum of all its principal component scores:

The principal component scores as constituting elements of the composite index are weighted by the corresponding shares of explained variance ρ_k (Ram, 1982). There is no loss in variance of the considered data. The explaining capability of the proposed indicator is extended to the total variance of initial variables.

Unlike other studies (Ram, 1982; Doukas, et. al., 2012) we use modified principal components $y_{1j} = \sum_{i=1}^{n} l_{1i}^2 x_{ij}$ instead of original principal components z_{1j} as constituting elements of the aggregate indicator of economic activity (Aivazian, Stepanov and Kozlova, 2006). This approach is based on normalized property of component loadings: $l_1 l_1^T = \sum_{i=1}^{n} l_{1i}^2 = 1$ (Aivazian and Mkhitarian, 2001; Ram, 1982). It allows to treat the weighting coefficients l_{1i}^2 as shares reflecting the impact of a primary variable x_i on the resulting integrate score. This indicator retains units of measure of initial variables x_{ij}.

Summing up, we use the following expression for the composite indicator of economic activity:

$$I_j = \sum_{k=1}^{n} \rho_k y_{kj} = \sum_{k=1}^{n} \rho_k \sum_{i=1}^{n} l_{ki}^2 x_{ij} = \frac{\sum_{k=1}^{n} \left(\lambda_k \sum_{i=1}^{n} l_{ki}^2 x_{ij} \right)}{\sum_{k=1}^{n} \lambda_k}.$$

The proposed methodology is applied to evaluate the overall Russian national social and economic performance and to have a glimpse of its dynamics in the past decade. Despite there exists an opinion that PCA is not applicable to time series (Mazziotta and Pareto, 2019) we argue that the methodology of our concern is not subject to any particular assumptions that could prohibit construction an aggregate indicator using temporal variations of original data. This point of view is supported by a number of studies (see, for instance, Fan, Wang and Zhu, 2013).

We use 96 indicators of the GCI database (Global Competitiveness Report, 2017) that characterize performance of the Russian economy for the years 2007-2017 (j=11) – those which are compatible within this time horizon. There is an ambivalent representation of the data. A group of 81 indicators which represent mainly WEF's Executive Opinion Survey data are scaled from 1 to 7 (best). These indicators reflect qualitative aspects of competitiveness, such as: property rights; quality of roads; internet access in schools; effectiveness of anti-monopoly policy; reliance on professional management; quality of scientific research institutions and so on. There are two 2 complex indicators (domestic market and foreign market size indices) among them (Global Competitiveness Report, 2017-2018).

The other 28 indicators have different dimensions (e.g. strength of investor protection (0–10 (best)); fixed telephone lines/100 pop.; life expectancy (years); tertiary education enrollment (%) and so on). To make them compatible with the previous group we normalize them into the 1-7 range.

There are two opportunities: if an indicator corresponds to the case "the more the better" (e.g. country credit rating, 0–100 (best), education enrollment, %; mobile broadband subscriptions/100 pop) then we adjust it to 1-7 ranking scale in the following way:

$$x_{ij}^{n} = 1 + 6\left(\frac{x_{ij} - x_{ij}^{\min}}{x_{ij}^{\max} - x_{ij}^{\min}}\right),$$

where x_{ij}^{n} is a normalized variable, $x_{ij}^{\max}$ and $x_{ij}^{\min}$ are correspondingly the "best" and the "worst" value of initial indicator x_{ij}.

If the case is less the better (e.g. general government debt, % GDP, no. procedures to start a business, trade tariffs, % duty) then the following normalizing transformation is applied:

$$x_{ij}^{n} = 1 + 6\left(\frac{x_{ij} - x_{ij}^{\max}}{x_{ij}^{\min} - x_{ij}^{\max}}\right),$$

where $x_{ij}^{\max}$ and $x_{ij}^{\min}$ are correspondingly the "worst" and the "best" value of initial indicator x_{ij}.

Imagine now that we handle with such normalized variables x_{ij}^{n} instead of x_{ij} in all the expressions above.

Results and discussion

We use Gretl to calculate principle component loadings for 96 initial variables. There are 10 principal component vectors with positive eigenvalues. So the PCA actually yields reduction of factor space dimension.

For every year we calculate the composite macroindicator I_j by summing up modified principal component scores y_{kj} weighted by the corresponding share of explained variance ρ_k.

The dynamics of generalized modified principal component (GMPC) composite macroindicator is represented by tab. 1. It can be hardly compared with the country's GCI dynamics because of the shifts in methodology applied by WEF in calculation of national competitiveness raitings.

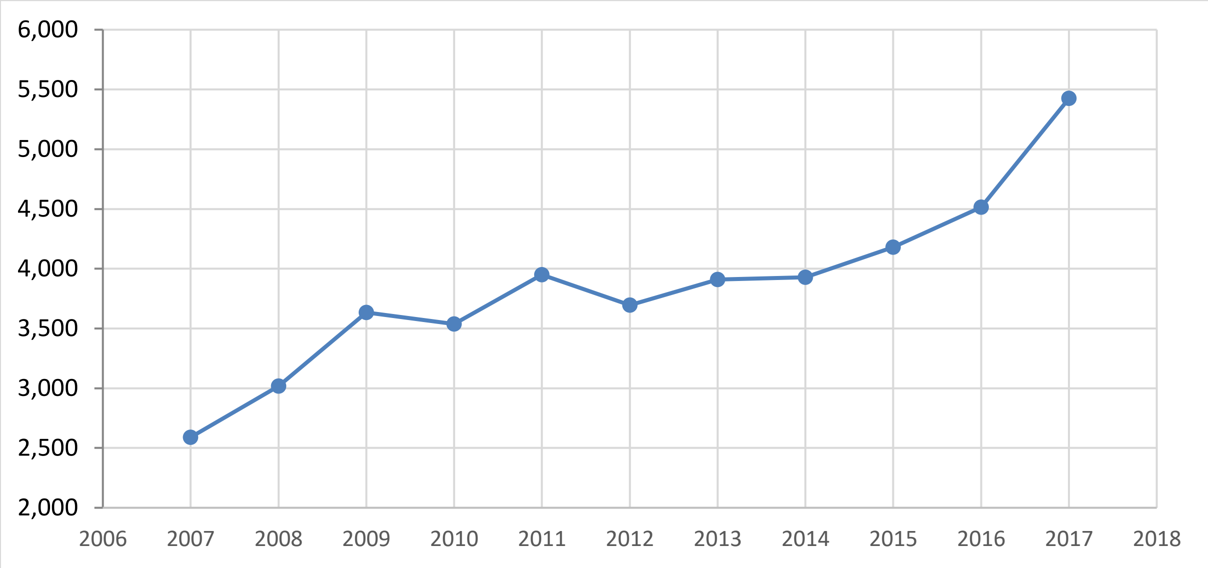
Tab. 1. Dynamics of Russian composite macroeconomic indicator and its components

Year	Total	Institutions	Infrastructure	Macroeconomic environment	Health and primary education	Higher education and training	Goods market efficiency	Labor market efficiency	Financial market development	Technological readiness	Market size	Business sophistication	R&D innovation
2007	2.590	0.159	0.213	0.754	0.236	0.176	0.505	0.359	0.067	0.090	0.001	0.019	0.012
2008	3.018	0.117	0.372	0.720	0.385	0.124	0.457	0.627	0.067	0.114	0.001	0.020	0.013
2009	3.633	0.117	0.387	0.726	0.628	0.153	0.680	0.707	0.064	0.137	0.001	0.019	0.012
2010	3.538	0.117	0.506	0.398	0.630	0.187	0.600	0.696	0.063	0.310	0.001	0.019	0.012
2011	3.950	0.115	0.581	0.536	0.782	0.187	0.583	0.734	0.064	0.338	0.001	0.018	0.011
2012	3.696	0.071	0.617	0.649	0.904	0.213	0.538	0.234	0.063	0.378	0.001	0.018	0.011
2013	3.910	0.074	0.637	0.664	0.924	0.215	0.638	0.235	0.064	0.427	0.001	0.020	0.011
2014	3.930	0.076	0.576	0.500	0.842	0.291	0.824	0.232	0.064	0.489	0.001	0.021	0.012
2015	4.179	0.134	0.560	0.341	0.885	0.291	1.072	0.161	0.164	0.537	0.001	0.021	0.012
2016	4.514	0.225	0.490	0.195	0.878	0.375	1.260	0.124	0.364	0.567	0.001	0.022	0.012
2017	5.424	0.270	0.493	0.433	1.290	0.435	1.328	0.185	0.364	0.589	0.001	0.023	0.013

Source: Own elaboration.

The obtained composite indicator reveals cyclical fluctuations along the trend of economic development (see fig. 1). As one could expect there is a fall-down in 2010 as a result of the world recession of 2008-2010. The national economy contracts again in 2012 after the compensating growth in the year before. This reveals the volatile situation in the Russian economy placed in unstable world environment. There is an obvious stagnation in 2014 as a result of economic sanctions imposed on Russian Federation. Nevertheless the economy copes with the puzzles and the growth resumes from 2015 onwards. One can observe that the composite macroindicator is to some extent rigid with respect to the ordinary economic cycle. The fluctuations of output (GDP) are smoothed out by nonmonetary factors of economic development.

Fig. 1: Russian composite macrodynamics by GMPCA



Source: Own elaboration.

Usually original data can be grouped into a number (θ) of subsets or pillars that reflect definite attributes of national social and economic performance:

$$X = \begin{pmatrix} \tilde{X}_1 \\ \vdots \\ \tilde{X}_\theta \end{pmatrix}, \text{ where } \tilde{X}_\alpha = \begin{pmatrix} X_{n_{\alpha-1}+1} \\ \vdots \\ X_{n_\alpha} \end{pmatrix} = \begin{pmatrix} x_{n_{\alpha-1}+1,1} & \cdots & x_{n_{\alpha-1}+1,m} \\ \vdots & \ddots & \vdots \\ x_{n_\alpha,1} & \cdots & x_{n_\alpha,m} \end{pmatrix}, \; 1 \le \alpha \le \theta,\; 1 \le \theta \le n.$$

For instance, in our case the indicators of competitiveness in (Global Competitiveness Report, 2018) are grouped into 12 categories, the pillars of competitiveness (θ=12), that represent three key determinants of development (subindices): fundamental factors (institutions $(\alpha = 1)$, infrastructure $(\alpha = 2)$, macroeconomic environment $(\alpha = 3)$, health and primary education $(\alpha = 4)$), factors of efficiency (higher education $(\alpha = 5)$, goods $(\alpha = 6)$, labor $(\alpha = 7)$ and financial $(\alpha = 8)$ markets efficiency, technological readiness $(\alpha = 9)$ and market size $(\alpha = 10)$) and innovation factors (business sophistication $(\alpha = 11)$ and R&D innovation $(\alpha = 12)$).

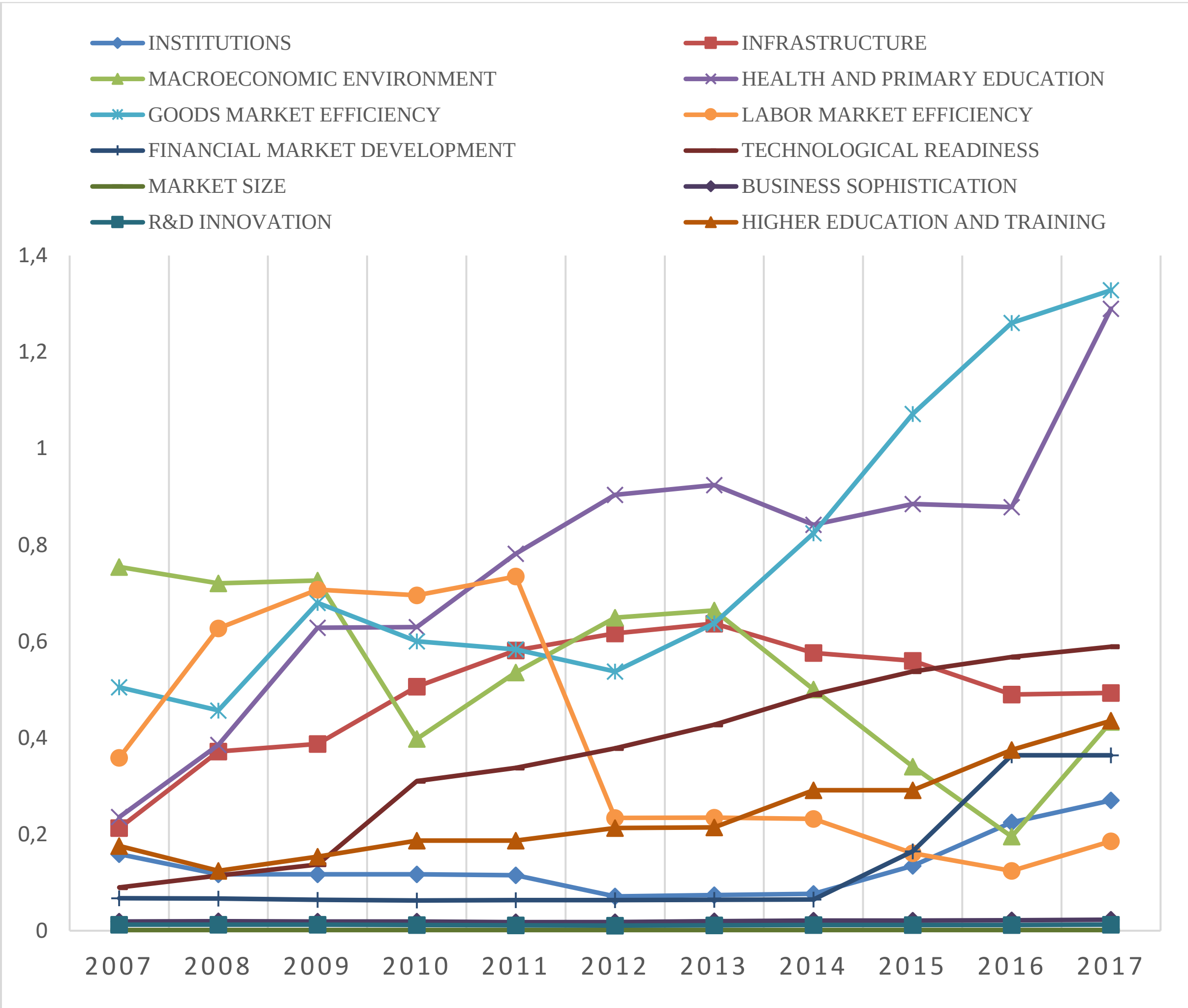
The aggregate index I_j can be decomposed into a sum of partial indicators:

$$I_{j\alpha} = \sum_{i=n_{\alpha-1}+1}^{n_\alpha} \sum_{k=1}^{n} \lambda_k l_{ki}^2 x_{ij} \Big/ \sum_{k=1}^{n} \lambda_k \,.$$

We obtain partial indices as the sums of weighted modified principal component scores for each of 12 data pillars. The 12 pillar sub-indices sum up to the country's overall macroindicator $I_j = \sum_{\alpha=1}^{\theta} I_{j\alpha}$. These sub-indices demonstrate positive or negative impact of the particular pillar on the overall indicator (see fig. 2). Russian economic development is supported by the positive dynamics of goods market efficiency, financial market development, technological readiness, human capital (health, education and training). The progressing (since 2014) decline in infrastructure, macroeconomic environment (with two definite pits – in 2010 – dating back to the global recession – and in 2014-2016 – corresponding to exogenous macroeconomic pressure due to sanctions confrontation) hinder economic progress. The influence of institutions is somewhat ambiguous. They slump in 2012-2014 and continue enhancing afterwards.

Thus the decomposition of the aggregate indicator provides a glimpse of the factors of macro performance and of the potential to improve it.

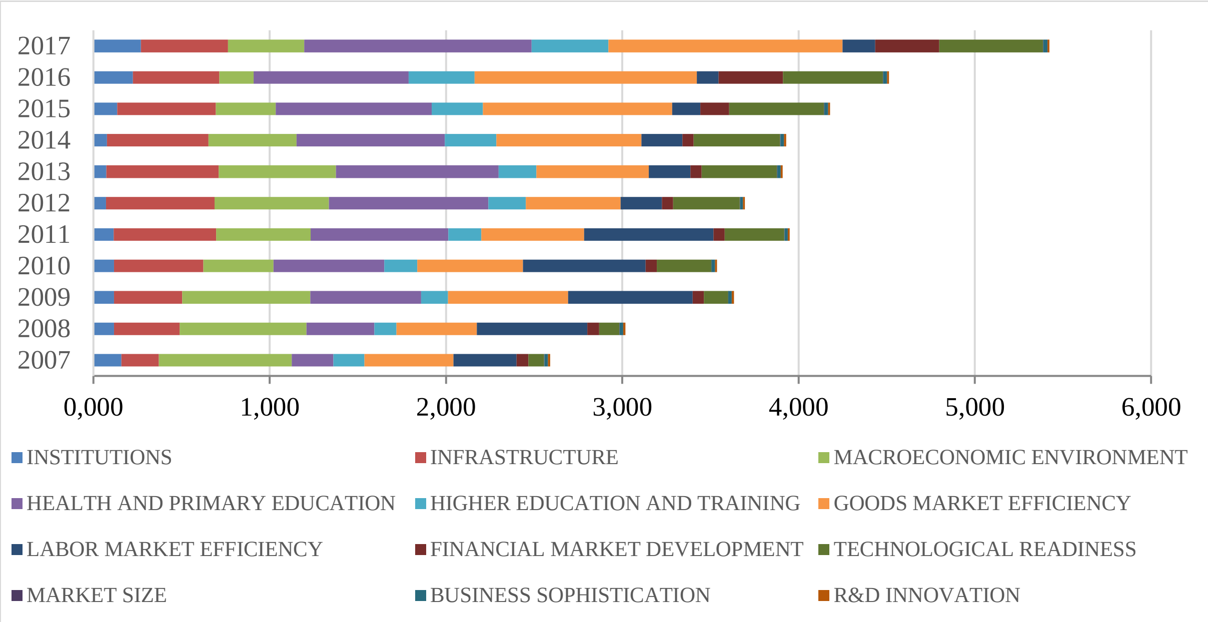
Fig. 2: Dynamics of partial macroindicators by GMPCA



Source: Own elaboration.

At the same time decomposition of the aggregate indicator shows the structural shifts in the Russian macroeconomic performance in the past decade (see fig. 3). Health and primary education, goods market efficiency, financial market development as well as technological readiness exhibit increasing input into economic development. Meanwhile macroeconomic environment and labor market efficiency become less significant for economic progress.

Fig. 3: Decomposition of Russian macrodynamics by GMPCA



Source: Own elaboration.

Conclusion

Measurement of national social and economic development is an ambiguous task because it takes into account a large number of indicators which are hard to compare. GDP is not an appropriate overall measure here.

We propose a kind of "natural" measure of national social and economic performance as the sum of modified principal component scores weighted the shares of explained variance of original data. The typical features of this approach are twofold. Firstly, we do not impose any subjective weights to the factors that influence economic competitiveness as opposed, for instance, to GCI. Secondly, unlike ordinary PCA there is no neglect to any residual variance of original data. The generalized modified principal component analysis regards the overall data scatter.

The proposed methodology is applied to construction of an overall economic indicator that reflects various qualitative aspects of national economic performance and can be treated as an alternative to GDP.

Further research can be devoted to application of this method of competitiveness assessment to calculations of economic effect of the Euroasian or other regional integration – at macro level of economic research and estimation of national companies success by evaluation of KPI or other indicators of economic activities - at micro level of research.

Acknowledgment

The study was supported by RFBR, research project No.18-010-00974A “Developing the model of management of territorial resource potential”.

The publication was prepared with the support of the «RUDN University Program 5-100».

References

Aivazian, S.A. & Mkhitarian, V.S. (2001). *Applied Statistics. Basics of Econometrics.* (in Russian). Moscow: UNITI-DANA.

Aivazian, S.A., Stepanov, V.S., & Kozlova, M.I. (2006). Measurement of Synthetic Categories of Quality of Life of the Regional Population and Identification of the Key Directions to Improve Regional Policy (in Russian). *Applied Econometrics*, *2*(2), 18-84.

Böhringer, C. & Jochem, P.E.P. (2007) Measuring the Immeasurable – A Survey of Sustainability Indices. *Ecological Economics, 63,* 1–8.

Booysen, F. (2002). An Overview and Evaluation of Composite Indices of Development. *Social Indicators Research, 59,* 115–151.

Doukas, H., Papadopoulou, A., Savvakis, N., Tsoutsos, Th. & Psarras J. (2012) Assessing Energy Sustainability of Rural Communities Using Principal Component Analysis. *Renewable and Sustainable Energy Reviews, 16,* 1949–1957.

Fan, Q., Wang, Y. & Zhu, Li. (2013) Complexity Analysis of Spatial-temporal Precipitation System by PCA and SDLE. *Applied Mathematical Modelling, 37,* 4059–4066.

Giannetti, B.F., Agostinho, F., Almeida, C.M.V.B. & Huisingh D. (2015) A Review of Limitations of GDP and Alternative Indices to Monitor Human Wellbeing and to Manage Eco-system Functionality. *Journal of Cleaner Production, 87,* 11–25.

Mazziotta, M. & Pareto, A. (2019) Use and Misuse of PCA for Measuring Well‐Being. *Social Indicators Research,* 142, 451–476.

Noll, H.-H. (2011). The Stiglitz-Sen-Fitoussi-Report: Old Wine in New Skins? Views from a Social Indicators Perspective. *Social Indicators Research, 102*, 111–116.

Ram, R. (1982). Composite Indices of Physical Quality of Life, Basic Needs Fulfilment, and Income. A "Principal Component" Representation. *Journal of Development Economics, 11*, 227–247.

Slottje, D.J. (1991). Measuring the Quality of Life Across Countries. *Review of Economics and Statistics, 73*, 684–693.

Stiglitz, J. E., Sen, A. & Fitoussi, J.-P. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Paris.

Veenhoven, R. (1996). Happy Life Expectancy. A Comprehensive Measure of Quality-of-Life in Nations. *Social Indicators Research, 39*, 1–58.

Verenikin, A.O. & Verenikina, A.Y. (2018). An Application of Principal Component Analysis to International Comparison of Economic Activities. *The 12th International Days of Statistics and Economics. Conference Proceedings.* Prague: Libuše Macáková, Melandrium, 1830–1841.

The Global Competitiveness Report 2017-2018. (2017). Geneva: World Economic Forum.

Contact

Alexei Verenikin
Lomonosov Moscow State University
Leninskie gory, 3rd new educational building, Moscow, 119991, Russian Federation
verenikin@mail.ru

Anna Verenikina
Peoples' Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya Str., Moscow, 117198, Russian Federation
verenikina_ayu@pfur.ru