The Interrelation between Digital and Tax Components of Sustainable Regional Development

Natalia Victorova¹*, Elena Vylkova², Vladimir Naumov², Natalia Pokrovskiaia³

¹Peter the Great St. Petersburg Polytechnic University, 29, Politechnicheskaya St., St. Petersburg 195251, Russia
²North-West Institute of Management, branch of Russian Presidential Academy of National Economy and Public Administration, St. Petersburg, 199178, Russia
³Saint Petersburg State University, St. Petersburg, 199034, Russia

Abstract. The paper explores the relationship between digital characteristics and tax potential of the constituent entities of the Russian Federation as factors of sustainable development of territories and economic growth. The purpose of the study involves development and testing of a methodology for assessing the level of digitalization in Russian regions and its relationship with their tax potential, which has not been sufficiently developed in the available scientific research by Russian and foreign authors. For this purpose, cluster and factor analysis were applied with the use of Rstudio, the IBM SPSS statistics package, and the Anaconda Navigator graphical interface. The following data were studied: the number of active subscribers of fixed and mobile access, subscribers of fixed and mobile broadband access as well as mobile communication devices for all constituent entities of the Russian Federation. The authors identified the worst and best regions in terms of mobile and fixed communications. It was concluded that the regions’ readiness for digital transformation is determined by the general level of their economies. Significantly larger tax revenues per capita are generated in regions with a highly developed IT component, which is the basis for solving the problems of sustainable development of such territories. The economic situation in a region, its gross regional product, and its tax potential create the basis for digitalization of each constituent entity of the Russian Federation. Significant tax revenues per capita are the key to the success of the territories in the IT sector. Promising areas for further research are: (1) expansion of the indicators used and time horizons; (2) extrapolation of the results to other countries and groups of countries; (3) use of the methods and models that have proven themselves when working with short series, e.g., autoregressive integrated moving average models.

Keywords: Digital technologies; Region; Russia; Tax potential; Tax sustainable development

1. Introduction

The importance of the problem of national competitiveness on a global scale is increasing in the context of global environmental change, political instability, economic crises, social revolutions, and technological and other challenges. Sustainable regional development is of great importance for the economic growth of the state (Berawi, 2016; Berawi, 2018), so it is comprehensively studied in various scientific publications (Carroll and Stanfield, 2001; Kharlamov and Kharlamova, 2019). It is more flexible and efficient than
national sustainable development, and allows one to capture the features of the economic development (Rodionov et al., 2018) and ecological potential (Mamraeva and Tashenova, 2020) of the territory more sensitively. At the same time, a review of previous studies shows that sustainable regional development is of particular importance for the states whose territories are characterized by significant diversity, e.g., for China (Zheng et al., 2021), Indonesia (Rahma et al., 2019), European countries (Prokop et al., 2019; Davidescu et al., 2020), and Russia (Bubnov and Bubnova, 2018; Kudryavtseva et al., 2020).

Region-specific data studies aim to identify determinants of sustainable regional development that can be extrapolated to other territories. For instance, the team of authors led by Gutman et al. (2018) as well as Leksin and Profiryev (2017) highlight indicators of sustainable development of the northern regions of Russia; Gutman and Rytova (2020) characterize the indicators of sustainable development of smart cities, Zaborovskaya et al. (2019) assess the strategy of sustainable development of Leningrad region, etc. The information and technology component is the basis for sustainable development (Hilty et al., 2005; Jovovic et al., 2017), and it has an effect upon all its other components (Avram et al., 2019).

The state of the information technology (IT) component, the information society, and e-government in international rankings of the development of the information society is usually analyzed through the telecommunication infrastructure index (Park and Oh, 2018; Abu-Shanab and Osmani, 2019). This index takes into account the number of internet users, personal computer users, cellular subscribers, fixed telephone subscribers, etc. This approach takes into consideration only the information infrastructure that is the lower level of the information society. It is advisable to take into account the efforts aimed at creating and implementing innovations to study the upper levels of the information society associated with business processes, information systems, and technologies. These efforts are closely related, on the one hand, and are reflected in the tax capacity, the amount of tax revenues generated, tax collection, tax arrears, effectiveness of tax audits, and other indicators characterizing the state of taxation of the territories, on the other hand. For instance, the study by Victorova et al. (2020b) showed intersections in Russian regions’ clustering based on digital and tax parameters. In this case, digitalization affects the tax system, on the one hand, (Ivanova and Selentyeva, 2018), and the tax system itself undergoes digitalization, on the other hand (Sabitova and Khafizova, 2015). Technological procedures are the focus of attention both within the framework of supranational tax regulation (Bradbury and O’Reilly, 2018) and national tax systems (Petersone and Ketners, 2017). The difference in the level of development of the IT component is largely due to the corresponding attractiveness of the territory, the level of investment (Saksonova, 2014; Sinenko and Mayburov, 2019), and the state policy for the development of innovation clusters (Selentyeva et al., 2018).

The COVID-19 pandemic has been a significant challenge to the sustainable development of the territories (Belov, 2021; Pinskaya et al., 2021). However, the need to mitigate its manifestation and create barriers to its spread has stimulated IT development, which in turn may provide an additional foothold for overcoming its long-term effects.

A review of the scientific literature showed that there is not enough research devoted to a comprehensive study of the relationship between the IT component of a territory and its tax potential. Inter alia, this is relevant in relation to the regions of the Russian Federation. Although assessment of such a relationship may not only be scientific in nature but also the applied one, as it can be in demand in the development of state and regional policies.

Hence, the purpose of this study is to develop and test a methodology for assessing the
The interrelation between digital and tax components of sustainable regional development.

The hypothesis of this study is that regions that are successful in digital development have achieved it due to the fact that they have a greater tax potential than regions lagging behind in the use of IT, since they are able to:

- Generate a larger amount of tax revenues, thereby ensuring the growth of gross regional product;
- Carry out tax control more successfully, accumulating additional funds in the budget system of the Russian Federation.

2. Methods

The present study is a continuation and development based on the results of territories clustering according to digital and tax characteristics (Victorova et al., 2020b). The authors concluded that they did not detect any direct correspondence between tax and digital clustering of Russian regions. This required carrying out deeper studies and improving the methodological approach in analyzing the impact of the digitalization of a region on the basis of specific factors that determine the tax potential, i.e., tax revenue and tax control.

For the purpose of the present study, the authors analyzed the data of Rosstat statistical collections on the state of telecommunications infrastructure in the Russian regions, namely, the number of active subscribers of fixed and mobile access, subscribers of fixed and mobile broadband as well as the number of mobile communication devices for all entities of the Russian Federation. The data analyzed refer to the period of 2016–2018.

The current state of available statistical reporting in Russia, the availability of accounting and analytical systems, big data, and their openness have determined the need to reduce the scale of the problems being examined. Cluster analysis techniques and factor analysis were used for this reduction. The research was carried out with the use of IBM SPSS statistics, Rstudio, and Anaconda Navigator software tools according to the following method.

At the first stage, the regions were clustered based on their information and technological activity. Methods of hierarchical clustering as well as k-means were applied. Each data element was considered as the $i$th example of a dataset. The set of feature values is denoted by the $x_{ij}$ identifier. Since the regions of Russia differ significantly in the size as well as in the number of population living in them, specific values of indicators given to one resident of each region were used for the comparability of the analysis results.

$$x_{ij} = x_{ij} / \text{people}_i$$

where $x_{ij}$ is a per capita indicator in a region; $\text{people}_i$ is population in an $i$th region.

The data normalization procedure was carried out to solve problems with the vector of individual features, taking into account the dynamics of their change over the years of observation.

$$x_{ij}^n = \frac{x_{ij}}{\max_{i} \{ x_{si} \}}$$

(2)

The obtained datasets were analyzed by classical methods of mathematical statistics, which showed that there were anomalous and extreme values that differed from normal values by more than fifty times.

The second stage was related to factor analysis. The principal component analysis was chosen to solve the problem, a fixed number of principal components was set to make it
possible to represent the results of solving the problem of factor analysis in the coordinate system (factor 1, factor 2). Thus, the second stage affected the reduction of the obtained data using factor analysis with a fixed number of factors equal to two as well as with the rotation of the new factor space using the varimax method. Interpretation of the results of factor analysis for the regional IT component was also carried out at this stage. Four groups were formed from the whole set of regions: the worst and the best regions in terms of mobile communications; the worst and the best regions in terms of fixed communications.

At the third stage, factor analysis revealed a statistical relationship between the gross regional product per capita in the region and the IT factor, which made it possible to interpret the correlation of the regional tax potential and the IT status of a region. The indicators of tax revenues, the amount of tax arrears, and additional charges based on the results of tax control for each of the constituent entities in the Russian Federation for 2018 were analyzed in order to assess their tax potential. Absolute figures were taken from the official website of the Federal Tax Service of Russia. Like other factors, these parameters were standardized based on the size of the population, and tax revenues per capita, tax arrears per capita, and value of payments (taxes, penalties and fines) additionally accrued as a result of the field, and office audits per capita were calculated.

The chosen research methodology may be scaled to a larger number of features describing the state of the information society and the state of taxation in the regions.

3. Results and Discussion

Below are the results of the study at each stage as well as their interpretation.

Stage one. Ten out of twenty-six different methods used to analyze the tendency of data to cluster resulted in the recommendation to select four clusters. Visualization of the results of cluster analysis by ordinal methods shown in Figure 1 indicates the existence of a large intracluster dispersion. Clusters intersect, and silhouette values are low. Two regions of Russia do not fall into the 90% confidence interval of the scattering ellipse of the coordinates of cluster objects under the assumption of the normal distribution law. The diagram shows two main components (Dim 1, Dim 2) explaining \( 78.6 + 19.2 = 97.8\% \) of the analyzed variables behavior.

![Cluster Analysis Diagram](image)

**Figure 1** Diagram showing the results of cluster analysis

Stage two. The method of factor analysis with a fixed number of factors equal to two as well as with the rotation of the new factor space by the varimax method was used to solve
the problem of data reduction. The possibility of using factor analysis resulted from the relatively high value of the Kaiser–Mayer–Olkin criterion equal to 0.694 as well as the distribution of variable values close to normal. An example of the histogram of the distribution of standardized values of the number of broadband subscribers confirming this conclusion is shown in Figure 2.

![Figure 2 Histogram of the distribution of the number of broadband subscribers (drawn by the authors)](image)

The results of solving the problem of factor analysis using the method of principal components for the specified parameter values are as follows:

![Figure 3 Results of factor analysis](image)

Two factors (RC 1, RC 2) describe 98% of the original variables’ volatility. The convergence for each variable (h 2) is close to one, which confirms the possibility of replacing individual variables with the factors. In this case, the factor loadings matrix (Figure 3) shows that the first factor describes the use of the mobile and the second one describes the use of the fixed telecommunication technologies. The scatter diagram for observations in the new coordinate system is shown in Figure 4. The diagram shows that the situation is more or less the same for most regions. However, there are leaders and laggards.

The worst regions in terms of mobile communication are Sevastopol (–5.24); Crimea (–3.79); Adygea (–1.66); Karelia (–0.99); Ingushetia (–0.92); Mordovia (–0.87); Tuva (–0.85); the Karachay-Cherkess Republic (–0.85); and Dagestan (–0.82). The best regions in terms of mobile communication are Yamalo-Nenets Autonomous Okrug (2.27); Moscow (2.09); St. Petersburg (1.98); the Krasnodar region (1.82); the Nizhny Novgorod region (1.47); and Khakassia (1.22).
The worst regions in terms of fixed communication are Ingushetia (−2.69); Dagestan (−2.46); Chechnya (−2.30); Altai (−2.09); Tuva (−2.08); the Chukotka Autonomous District (−1.61); the Kabardino-Balkarian Republic (−1.51); and Khakassia (−1.48). The best regions are the Novosibirsk Region (2.95); Karelia (2.34); Moscow (2.06); the Murmansk region (2.02); the Voronezh region (1.30); and the Chelyabinsk region (1.18). Note that all the given values are dimensionless, standardized values. Positive and negative values indicate that the given value is greater or less than the mathematical expectation, respectively.

In general, the IT component is worst developed in the city of Sevastopol and in the Republic of Crimea as well as in the Republic of Ingushetia, the Republic of Dagestan, the Republic of Adygea, and it is best developed in the city of St. Petersburg, Moscow, Novosibirsk region, the Republic of Karelia, and Murmansk region.

Stage three. The study also found that there is a statistical relationship between the gross regional product per capita in a region and the IT factor, when solving the problem of factor analysis for one factor. The pair correlation coefficient between them is 0.50, and it significantly differs from zero at the 0.05 level. Therefore, we can conclude that the regions’ readiness for digital transformation is determined by the general level of their economy. This single factor is chosen because it explains 79% of the variables’ variation.

The growth of the gross regional product of a constituent entity of the Federation indicates the growth of tax potential, namely, the high level of generated tax revenues and additional charges per capita based on the results of tax control.

Substantially higher tax revenues per capita are generated in regions with a highly developed IT component (see Figure 5).

The efficiency of the tax control in most of the leading regions is also higher in the field of IT technologies than in the regions with a less developed IT component. The average value of additionally accrued payments as a result of field and office audits per capita (see Table 1) in the regions with a highly developed IT component is three times higher than the corresponding value in the regions with a poorly developed IT component. This value amounted to 8,358 thousand rubles per capita in Moscow, to 3,227 thousand rubles in St. Petersburg, to 1,306 thousand rubles in the Murmansk region, and to 1,271 thousand rubles in Karelia. However, the corresponding value was 565 thousand rubles in Adygea, 666

Figure 4 Scatter plot of observations in the new coordinate system
The interrelation between digital and tax components of sustainable regional development

Thousand rubles in Crimea, 1,017 thousand rubles in Sevastopol, 1,043.5 thousand rubles in Dagestan; although the amount of payments accrued as a result of the field and cameral tax audits in Ingushetia amounted to 2,042 thousand rubles per capita, that is close to the values of the regions with a highly developed IT component. Thus, it proved impossible to unambiguously identify the connection between digital development and the effectiveness of the tax control in Russian regions at this stage of the study.

![Figure 5 Tax revenues in 2018](image)

**Table 1** Results of tax audits in 2018 per capita, thousand rubles

<table>
<thead>
<tr>
<th>Constituent entity of the Federation</th>
<th>Additional charges, including tax sanctions and penalties as results of:</th>
<th>Tax audits</th>
<th>Cameral tax audits</th>
<th>Field tax audits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moscow</td>
<td>8,357.8</td>
<td>1,317.3</td>
<td></td>
<td>7,040.4</td>
</tr>
<tr>
<td>St. Petersburg</td>
<td>3,227.5</td>
<td>345.4</td>
<td></td>
<td>2,882.1</td>
</tr>
<tr>
<td>Murmansk region</td>
<td>2,384.0</td>
<td>1,080.2</td>
<td></td>
<td>1,303.8</td>
</tr>
<tr>
<td>Novosibirsk region</td>
<td>1,305.8</td>
<td>328.5</td>
<td></td>
<td>977.3</td>
</tr>
<tr>
<td>Karelia</td>
<td>1,271.2</td>
<td>444.7</td>
<td></td>
<td>826.5</td>
</tr>
<tr>
<td>Adygea</td>
<td>564.9</td>
<td>173.2</td>
<td></td>
<td>391.7</td>
</tr>
<tr>
<td>Sevastopol</td>
<td>1,017.4</td>
<td>71.9</td>
<td></td>
<td>945.5</td>
</tr>
<tr>
<td>Crimea</td>
<td>666.3</td>
<td>164.7</td>
<td></td>
<td>501.6</td>
</tr>
<tr>
<td>Dagestan</td>
<td>1,043.5</td>
<td>109.3</td>
<td></td>
<td>934.1</td>
</tr>
<tr>
<td>Ingushetia</td>
<td>2,042.3</td>
<td>553.7</td>
<td></td>
<td>1,488.6</td>
</tr>
</tbody>
</table>

It should be noted that the results of the study fully confirmed the correlation between the receipt of taxes and the level of the IT infrastructure development in the region. This is due to the active use of cloud-based software products by the taxpayer in the performance of their duties and the availability of the internet in the corresponding territory. At the same time, it should be noted that modern technologies are used to quite a high extent in tax procedures in the country as a whole. However, the influence of a human factor or other (nontechnological) factors on the results of tax audits also remains significant, which was confirmed by the results obtained in four regions.

It is also important to emphasize that the results are not entirely consistent with the Olatunji and Ayodele (2017) study, which confirmed an unambiguous positive link between tax productivity and IT development for Nigeria as well as the work of Hodzic (2019) on Croatia, showing the negative consequences of underdeveloped IT infrastructure for tax procedures. This can be explained both by the choice of the analyzed parameters.
characterizing the informational and technological as well as tax components of sustainable regional development and also by the specifics of the national experience in the development of the IT and tax administration. In particular, previous studies have shown that tax capacity factors differ not only by country, but even by federal districts of the Russian Federation (Victorova et al., 2020a). This is also true for the IT characteristics (Victorova et al., 2020b).

We believe that weak connection between the results of our study and previous works on other countries leads to the need for additional research on the factors of the IT component of the sustainable regional development as well as the specifics of taxation in the territories of Russia. This is due to the fact that wide disparities among the country territories create the need for the balanced, sustainable regional development throughout the country in general, and its IT and tax components in particular.

4. Conclusions

The study largely confirmed the hypothesis that had been put forward. The level of development of the region’s economy, its gross regional product, and tax potential create the basis for the digital development of a constituent entity of the Russian Federation. Significant tax revenues per capita are the key to success of the territories in the IT sphere, which confirms the first part of the hypothesis. The second part of the hypothesis has not received a clear confirmation, so it needs to be tested in dynamics over a longer time interval.

Summing up the study, we can draw the following conclusions.

The level of society informatization should be assessed not only on the basis of data on development of the telecommunications infrastructure, but efforts aimed at creating and introducing innovations should also be taken into account.

The study may be continued in the future in terms of expanding the time horizons and indicators used.

The experience of assessing the link between the IT and tax component of Russian regions may be extrapolated to other countries with the federal structure.

Despite the fact that the existing statistical compilations represent a large collection of open datasets of the Russian Federation, they contain “raw data” that need to be processed. The number of observations is not enough. Therefore, it is necessary to use proven methods and models when working with short rows, e.g., with autoregressive integrated moving average (ARIMA) models, in the long run.

Acknowledgements

The research is partially funded by the Ministry of Science and Higher Education of the Russian Federation under the strategic academic leadership program “Priority 2030” (Agreement 075-15-2021-1333 dated 30.09.2021).

References


The Interrelation between Digital and Tax Components of Sustainable Regional Development


