ECONOMETRIC MODELING OF REGIONAL ECONOMIC GROWTH IN DIGITAL TRANSFORMATION

(Based on the integration of Big Data and ML methods)

Kuchaboyev Ruslan Tellayevich
University of Economics and Pedagogy
Senior lecturer of the Department of "Computer Systems"
https://doi.org/10.5281/zenodo.18360270

Abstract. *In the context of the emergence of a digital economy, regional economic development is increasingly dependent on the level of digitization, the use of big data and modern analytical tools. This article explores methods for econometric modeling of regional economic growth based on traditional panel models and their expansion with Big Data technologies and machine learning methods. The aim is to quantitatively assess the impact of digital factors on regional economic growth and compare the effectiveness of classical econometric models with models enhanced on the basis of ML.*

Keywords: *economic growth, regional economics, econometrics, digital economics, Big Data, machine learning, panel data.*

Mazmunnoma: *Raqamli iqtisodiyot shakllanayotgan sharoitda mintaqaviy iqtisodiy rivojlanish tobora raqamlashtirish darajasi, katta hajmdagi ma'lumotlar va zamonaviy tahliliy vositalardan foydalanishga bog'liq bo'lib bormoqda. Ushbu maqolada mintaqaviy iqtisodiy o'sishni ekonometrik modellashtirish usullari an'anaviy panel modellari asosida hamda ularni Big Data texnologiyalari va mashinaviy o'rganish usullari bilan kengaytirish orqali tadqiq etiladi. Raqamli omillarning mintaqaviy iqtisodiy o'sishga ta'sirini miqdoriy baholash va klassik ekonometrik modellar samaradorligini ML asosida kuchaytirilgan modellar bilan solishtirishdan iborat.*

Kalit so'zlar: *iqtisodiy o'sish, mintaqaviy iqtisodiyot, ekonometrika, raqamli iqtisodiyot, Big Data, mashinaviy o'rganish, panel ma'lumotlar.*

Аннотация. *В условиях становления цифровой экономики региональное экономическое развитие все больше зависит от уровня цифровизации, использования больших данных и современных аналитических инструментов. В данной статье рассматриваются методы эконометрического моделирования регионального экономического роста на основе традиционных панельных моделей и их расширение с использованием технологий больших данных и методов машинного обучения. Цель состоит в количественной оценке влияния цифровых факторов на региональный экономический рост и сравнении эффективности классических эконометрических моделей с моделями, усовершенствованными на основе машинного обучения.*

Ключевые слова: *экономический рост, региональная экономика, эконометрика, цифровая экономика, Big Data, машинное обучение, панельные данные.*

INTRUDUCTION

Regional economic growth is one of the main categories of modern economic science, since it is at the regional level that disparities in economic structures, institutional differences, and the

impact of economic policies are most clearly manifested. In recent years, digitalization processes have significantly changed the mechanisms of economic growth, which requires a revision of traditional methods of its analysis and forecasting.

The development of the digital economy is characterized by a sharp increase in the volume of economic data, the diversity of information sources and the expansion of computational capabilities. In this regard, Big Data and machine learning methods are becoming an important addition to classical econometric tools. However, despite the active introduction of new technologies, econometric modeling remains the methodological basis of empirical economic research and ensures the statistical validity and interpretation of the results.

The relevance of this study is determined by the need to integrate econometric models with modern IT tools in order to more accurately analyze regional economic growth. The study focuses on their rational combination, not on replacing econometrics with machine learning.

LITERATURE ANALYSIS

The scientific study of the regional economy and its development has been the focus of attention of economists. In particular, the research of the following scientists has served as an important theoretical and practical basis for this topic. British economist A. Venables has deeply analyzed the impact of agglomeration processes, transport costs and international trade on regional growth in regional economic development. His work has served as a theoretical foundation for determining the competitiveness of regions. [4]

P. Krugman developed the theory of the "new economic geography". He substantiated the economic centralization, trade flows and the interdependence between regions through mathematical models. These studies determined the main directions of the methodology of regional economic modeling. [5]

Japanese scholar M. Fujita, in collaboration with P. Krugman, developed theories of regional development and urbanization. He studied the relationship between regional clustering, the location of production centers, and economic development using econometric methods. [6]

H. Glenn's work explores the impact of innovation factors, the knowledge economy, and human capital on regional development. He particularly emphasizes the role of technological factors and innovation policies in economic growth. [7]

The problem of regional economic growth is central to neoclassical and institutional economic theory. The work of Solow and Romer provided a theoretical framework for analyzing the factors of economic growth, including capital, labor, and technological progress. These models were later adapted to the regional level using panel data.

RESEARCH METHODOLOGY

Fixed and random effects models, as well as gradient boosting and random forest algorithms, were used as the methodological basis. The research results show that the integration of digital data and machine learning tools increases the forecast accuracy and explanatory power of econometric models. The practical significance of the research is determined by the possibility of applying the proposed approach in the development of regional economic policy in the context of digital transformation.

DISCUSSION AND ANALYSIS

Modern research emphasizes the increasing role of digitalization as a factor of economic growth. A number of authors note that the development of digital infrastructure, the expansion of Internet access, and the use of information technologies have a positive impact on labor productivity

and investment activity in regions. At the same time, empirical assessment of these effects requires an expansion of the set of traditional econometric tools.

In recent years, the direction of applying Big Data and machine learning methods in economic research has been rapidly developing. Although machine learning methods demonstrate a high level of predictive ability, they are often criticized for their insufficient interpretability. Therefore, interest in hybrid models that combine econometric rigor and the computational power of IT tools is growing in the scientific literature.

However, despite the large number of existing publications, the issue of integrating machine learning into the process of econometric modeling of regional economic growth has not been sufficiently systematized. This fact determines the scientific novelty of this study.

Integrating Big Data and machine learning. The second stage of the study uses Random Forest and Gradient Boosting algorithms. These methods allow us to identify complex nonlinear relationships between variables and increase forecast accuracy.

Big Data used in the form of expanded digital indicators, which include data on internet activity, the density of digital services, and the use of electronic platforms. In this case, the econometric model remains the main tool of analysis, while the results of machine learning are used to conduct comparative analysis and validate the conclusions drawn.

The empirical basis of the study is formed on the basis of regional statistical data over many years. The main sources of information are national statistical services and open databases of socio-economic indicators.

Econometric model the following panel regression is used as:

$$GDP_{it} = \alpha + \beta_1 Inv_{it} + \beta_2 Lab_{it} + \beta_3 Dig_{it} + \mu_i + \varepsilon_{it}$$

here:

GDP_{it}- gross regional product of region i in period t;
Inv_{it}- investments in fixed capital;
Lab_{it}- employment or labor resources;
Dig_{it}- digitization index of the region;
μ_i- individual regional effects;
ε_{it}- random error.

The model is estimated based on fixed (FE) and random (RE) effects, the selection of which is carried out using the Hausman test.

The results of the panel model evaluation showed that investment and employment have a statistically significant impact on regional economic growth. The digitalization variable also shows a positive and significant coefficient, confirming the hypothesis regarding the role of digital factors in shaping the economic growth process.

The comparative analysis showed that fixed effects models had a higher explanatory power than random effects models. The use of machine learning methods made it possible to identify nonlinear effects of digitalization, but these effects are not fully reflected within the framework of linear econometric models.

Models supplemented with machine learning methods showed higher results in terms of forecast accuracy, but econometric models provided a more accurate and understandable interpretation of the impact of individual factors.

The results obtained are consistent with the conclusions of modern research on the positive impact of digitalization on economic growth. Unlike purely computational approaches, the proposed

model allows us to preserve the economic interpretation of the coefficients, which is of great importance for the development of economic policy.

Integrating Big Data and machine learning methods does not replace econometrics, but rather extends its analytical capabilities. This approach is particularly promising for regional studies where data are highly asymmetric.

Today, when forecasting regional changes, factors such as macroeconomic indicators (GDP growth rate, investment volume, employment rate), social indicators (education and healthcare coverage, per capita income), as well as digital infrastructure and environmental sustainability criteria (Internet coverage, integration of green technologies) are studied in a comprehensive manner.

Table 1

System of endogenous and exogenous indicators of regional economic development

Indicator type	**Direction**	**Main indicators**
Endogenous indicators	Result indicators	- GDP (GRP, GRDP): - regional gross product volume - volume of industrial production - per capita income - employment rate / number of new jobs - export volume or foreign trade balance - regional development index (composite index)
Exogenous indicators	Economic factors	Investment volume (INV), including FDI: - diversification of industrial enterprises (HHI, share of sectors) - transport and logistics indicators (cargo turnover, length of roads, transport costs) - sales volume and retail turnover
	social factors	- population size and demographic composition - human capital (quality of education, number of qualified personnel, health indicators) - employment and unemployment rates - tourism flows and service sector indicators
	innovative and digital factors	- research and development expenses (R&D) - share of digital services (IT sector, e-commerce volume) - number of techno parks and startups - export of innovative products
	Ecological and resource factors	- energy supply and consumption - share of renewable energy - waste recycling rate - the volume of use of natural resources
	Institutional factors	Legislative Quality and Regulation Index: - regional budget expenditures - number of regional cooperation and integration projects - business environment and Doing Business indicators

The close interdependence and dynamics of these factors provide the basis for developing forecast models necessary for reducing interregional disparities, effectively allocating resources, and implementing targeted regional policies.

CONCLUSION

The article develops and practically tests an approach to econometric modeling of regional economic growth in the context of a digital economy. It is proven that digital factors have a statistically significant impact on the economic development of regions. The integration of Big Data and machine learning methods increases the accuracy of econometric analysis and forecasting, while preserving the methodological foundations of econometrics.

The results of the study can be used in the process of formulating regional development strategies, as well as in further scientific research in the fields of econometrics and digital economics.

REFERENCES

1. Mullajonov, A.M. (2021). *Raqamli iqtisodiyot: nazariya va amaliyot.* Toshkent: Iqtisodiyot nashriyoti.
2. Abdurahmonov, Q.X. (2020). *Iqtisodiy o‘sish va mintaqaviy rivojlanish.* Toshkent: Fan va texnologiya.
3. Айвазян, С. А. (2020). *Эконометрика*. Москва: Магистр; ИНФРА-М.
4. Venables A.J. Trade and trade policy with imperfect competition: the theory of the second best. Journal of International Economics, 1990.
5. Krugman P. Geography and Trade. MIT Press, Cambridge, 1991.
6. Fujita M., Krugman P., Venables A. The Spatial Economy: Cities, Regions and International Trade. MIT Press, Cambridge, 1999.
7. Glenn H. Regional Innovation Systems and Economic Development. Routledge, London, 2007.
8. Магнус, Я. Р., Катышев, П. К., & Пересецкий, А. А. (2019). *Эконометрика. Начальный курс.* Москва: Дело.
9. Клейнер, Г. Б. (2018). Цифровая экономика и системная трансформация экономики. *Вопросы экономики*, (6), 71-86.
10. Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach.* Boston: Cengage Learning.
11. Baltagi, B. H. (2021). *Econometric analysis of panel data.* Cham: Springer.