**УДК: 378.4**

# ЭКОНОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ И АНАЛИЗ ОБЪЕМА ГОСУДАРСТВЕННЫХ РАСХОДОВ НА ЦИФРОВИЗАЦИЮ ЭКОНОМИКИ

**Слободянюк Г.И. [1], Федотова А.А.[1]**
**Научный руководитель д.э.н., профессор Трегуб И.В.[1]**

Финансовый Университет при Правительстве Российской Федерации, Россия, Москва, e-mail: 173280@edu.fa.ru

**В этой статье рассматривается один из важнейших экономических процессов – процесс формирования государственных затрат, а именно затрат на цифровизацию Российской экономики; предлагается анализ распределения средств государством на преобразование технологической информационной среды России. Выбраны основные переменные, которые гипотетически влияют на объемы затрат государства на цифровизацию, также проведены тесты, которые проверяют построенную гипотезу путем эконометрического анализа независимых переменных, влияющих на формирование объемов затрат на цифровизацию; построенная модель была протестирвоана наличие погрешностей. Был проведен полный комплексный анализ состояния за выбранный период 2000 – 2018 г.г, перспективы и причины также были выяснены; было доказано значительное влияние количества интернет – пользователей в России (млрд. человек) и количества мобильных телефонов (на 100 человек) на объемы государственных затрат на цифровизацию экономики.**

Ключевые слова: цифровизация, государственные затраты, эконометрическое моделирование, экономический анализ, расходы;

# ECONOMETRIC MODELING OF THE VOLUME OF GOVERNMENT SPENDING ON THE DIGITALIZATION OF THE ECONOMY.

**Slobodyanyuk G.I. [1], Fedotova A.A.[1]**

Financial University under the Government of the Russian Federation, Moscow, e-mail: 173280@edu.fa.ru

**This article discusses an important economic process - the process of generating public expenditures, namely, on the digitalization of the Russian economy; an analysis of the distribution of state funds for the transformation of the technological information environment of Russia is proposed. Main variables, which hypothetically affect the amount of government spending on digitalization were selected; the constructed model was tested for errors. A complete comprehensive analysis of the state was carried out for the selected period 2000 - 2018, and the significant effect of the number of Internet users in Russia and the Number of mobile phones on volumes of government spending on digitalization of the economy were proved.**

Keywords: digitalization, government spending, econometric modelling, economic analysis, expenses;

## Introduction

Currently, the global economy is on the verge of a new transformation. Ongoing digitalization leads to a change in the global economy by reducing the cost of collecting, storing, processing data; reduction of production chains, etc. Such changes, of course, affect the requirements that apply to the skill level of workers and to market participants representing business and the state. So, the concept of digitalization is the most popular trend of 21st century, as it has a huge influence on the all the global environment of the economy. Theoretical economists have not yet reached a consensus on how significant the impact of digitalization on productivity, what contribution to GDP growth it can give, whether multiplicative and cross-

sectoral effects are possible. But entrepreneurs and departmental experts, as this study shows, have already felt the scale and inevitability of change in their markets and in their areas of responsibility. In this matter, it is necessary to remember that the most important niche here is shared by small and medium-sized businesses. To his considerations and plans is worth look particularly closely. It is equally important to study the experience of the reaction to the current revolutionary situation of the more technologically advanced States –the work of their experts, their launched programs and initiatives related to digitalization. This will help to avoid the same mistakes when compiling your program and save resources. Finally, it is necessary to analyze our own institutional field in which the digitalization of our economy and public administration will unfold.

The growing interest in this topic is evidenced by the fact that digitalization covers the whole world. Due to this, each country must take some measures in order to remain a reliable and developed partner in the system of international relations. That is, the ability of the country to develop and become more and more flexible to the new world trends is now considered is very important.

In this article we suggest analyzing the process of digitalization of the state, namely the decision of the state to invest budgetary or other investment funds in creating all conditions for the development of the economy and its digitalization.

## Materials and methods

In this article some methods of econometric analysis were introduced. These methods were used to determine the co-dependencies between the main indicators, which were taken as the basis of the analysis. Determination of correlations between selected factors as well as various types of analysis, for example, regression analysis. Also, as the main method for checking the influence of indicators, statistical tests were used to determine the quality of the constructed connection — the model.

Now we are going to introduce the thesis and define the main methods for its consideration. The government of Russian Federation decides to have a specific amount of budget and distributes it between lots of things, which are connected with the wellbeing of the whole country. The state budget is the most important financial document of the country. It is a set of financial estimates of all departments, public services, government programs and others. So, the process of formation of governmental expenses and the decision of the government to spend budgeted money on a particular project, particular economic or social sphere is the main topic in this work. We are going to concentrate on the expenses of the government in the digitalization of the country and define the main factors which may influence the decisions of the government to spend money on the

Firstly, let's consider the data which was taken from the statistical resources on the volumes of government spending on the digitalization for the last 19 years, that is, for the period 2000 – 2018 years.[1]

Next, in the main part of the analysis we may introduce some economic factors, which effect the decision of the government to invest directly in the digitalization of the Russian economy. In order to conduct the analysis and get the appropriate results of the analysis, some tests of the numerical data should be introduced. Let's consider the methodology of the main part of the analysis:

1)We may construct the model, where volumes of government spending on the digitalization – is the observed dependent variable, which may be influenced by some independent variables and may be explained by them. The whole analysis is dependent on the appropriate choice of variables, which my describe, why has the government decided to spend more or less on the digitalization of the economy.

Gauss – Markov Theorem – is the main criteria for the introduction of the analysis: A paired regression model is considered in which observations Y (volumes of government spending on digitalization) linked to X (the chosen independent variables) with the following dependency on the basis of n observations:

$$Y_i = \beta_1 + \beta_2 \cdot X_i + \mu_t.$$

And several criteria are met, such as:

a) The data model is correctly specified;

b)All $X_i$ are determined and not all of them are equal;

c) Errors are not systematic, that is it, $E(\varepsilon_i) = 0 \forall_i$;

d)The variance of errors is the same and equal to $\sigma^2$;

e) $Cov(\varepsilon_i, \varepsilon_j) = 0 \forall_i, j$;

Then under these conditions the least squares estimates are optimal in the class of linear unbiased estimates.

2)In order to check the relations between the variables, we may introduce the regression analysis, which may show the quality of the built model - $R^2 = 1 - \frac{Var(e)}{Var(Y)}$.

3)The next step – to introduce F- test, which shows the quality of the model of the specifications and accepts that previous analysis was not random. The F-test is defined as number of observations minus the number of parameters.

[1] Table 1. Data on the Volumes of Government Spending on the Digitalization

4)T-test, which may accept the main variables and show those, which are not so significant for the analysis.

5) Goldfeld – Quandt test. This test is designed to test the second condition of Gauss – Markov theorem – the equality of variances of random residuals. In order to implement this test, we are going to conduct the following steps:

$$zi = |y_{1i}| + |y_{2i}| + |y_{3i}| + |y_{4i}|,$$

On first n ordered equations of object observations, where n satisfies conditions k+1 < n.

$$RSS_1 = \sum_{i=1}^{n} u^2{}_i = \sum_{i=1}^{n} (y_i - \hat{y}_i) \text{ , where}$$

$u_i -$ GLS – estimate of random disturbance I observation. So, the GQ formula is following: $GQ = \frac{RSS_1}{RSS_2}$,

Random residual of a model is considered to be homoscedastic, when: $\begin{cases} GQ \leq F_{critical} \\ GQ^{-1} \leq F_{critical} \end{cases}$

6)Durbin – Watson statistical test – statistical criterion used to test first-order autocorrelation of elements of the sequence being studied.

## Main Part

Firstly, lets introduce the main independent variables, which may influence decided by the volumes of government spending on digitalization. In this work we may introduce the following variables:

1) Number of internet users in Russia (billion, persons) - $N_t$**.** This is one of the most important variables, which shows the level of digitalization of the overall life of a person and the interconnection of his life and modern technologies. The share of Internet users in Russia is 81% of citizens. Including 65% go online daily. Among Russians from 18 to 24 years, this figure is 97%[2].

2) Rate of rising population in Russia - $R_t$.

3) Number of mobile abonents per 100 people - $A_t$. This variable shows the portion of population, which has already digitalized their day – to - day life – with the integration of mobile connection. The results are: with the development of modern technologies the number of digital opportunities gets bigger and makes the life of people much easier.

4) Variable, which shows the growing importance of the topic of digitalization and the rapid growth of the attention of the whole scientific and public society, is - Publications in the field of Digitalization - $P_t$. In order to identify the growing interest for this topic, look at various scientific works was taken.

[2] https://www.wciom.ru/

Now let's consider the data on these variables and build the model, where $GSD_t$(Volume of government spending on the digitalization) is explained by these chosen variables.

Table 1. Initial data for independent variables

| **Year** | **Number of internet users in Russia (billion persons)** $N_t$ | **Rate of rising population (%)** $R_t$ | **Publications in the field of ICT and Digitalization (pieces)** $P_t$ | **Number of mobile abonents (per 100 people)** $A_t$ |
|---|---|---|---|---|
| 2000 | 3,5 | -0,42 | 1030 | 5,32 |
| 2001 | 4,3 | -0,43 | 1145 | 12,13 |
| 2002 | 9,6 | -0,46 | 1373 | 24,99 |
| 2003 | 11,5 | -0,45 | 1534 | 51,18 |
| 2004 | 17 | -0,52 | 1640 | 83,55 |
| 2005 | 21,7 | -0,49 | 1883 | 105,13 |
| 2006 | 26 | -0,46 | 1889 | 119,59 |
| 2007 | 29,4 | -0,62 | 2000 | 139,44 |
| 2008 | 33,7 | -0,044 | 2010 | 160,77 |
| 2009 | 42 | -0,03 | 2019 | 166,04 |
| 2010 | 46,2 | -0,045 | 2030 | 142,22 |
| 2011 | 54,5 | 0,078 | 2100 | 145,07 |
| 2012 | 59 | 0,168 | 3000 | 152,02 |
| 2013 | 66,6 | 0,213 | 3010 | 153,75 |
| 2014 | 81,4 | 0,218 | 3100 | 157,96 |
| 2015 | 86 | 0,193 | 3678 | 159,15 |
| 2016 | 90 | 0,2 | 3927 | 157,86 |
| 2017 | 95 | 0,18 | 4036 | 158,97 |
| 2018 | 109 | 0,14 | 6020 | 160,5 |

$$\begin{cases} GSD_t = a_0 + a_1 \cdot N_t + a_2 \cdot R_t + a_3 \cdot P_t + a_4 \cdot A_t + \varepsilon_t \\ E(\varepsilon_t) = 0;\ \sigma(\varepsilon_t) = const, \end{cases}$$

And the estimated form of the model:

$$\begin{cases} \widehat{GSD_t} = 117{,}11 + 2{,}88 \cdot N_t + 78{,}22 \cdot R_t + 0{,}12 \cdot P_t - 0{,}4 \cdot A_t \\ \quad (10{,}82) \quad (1{,}70) \quad (8{,}84) \quad (0{,}35) \quad (0{,}63) \\ R^2 = 0{,}973 \quad df = 17; df_1 = 4; df_2 = 13 \quad F = 97{,}134 \\ t = 0{,}05, t_{crit} = 2{,}1604 \quad F_{crit} = 3{,}17912 \end{cases}$$

Where $a_0 = 117{,}11$ with the standard error of 10,82; $a_1$= 2,88, with the standard error of 1,70; $a_2 = 78{,}22$, with the standard error of 8,84; $a_3$ =0,12 with the standard error of 0,35; $a_4 = -0{,}4$, with the standard error of 0,63; All this information is obtained from the Regression analysis table.

Regression analysis

The regression analysis provided in the work with the result that $R^2 = 0{,}98$, which means that 98% in variance of the volume of government spending on digitalization is explained by Number of internet users in Russia, Rate of rising population, Publications in the field of ICT and Digitalization, Number of mobile abonents per 100 people by linear regression model. However, this value may be random. In order to check, whether the value is random, we introduce, firstly, F-test.

F-Test

In the work it equals to 97,13. $F_{crit}$ in the model observation equals 3,18, so the $F_{crit}$ < F. For the consideration of F critical the level of significance 5% (95% of confidence level) was considered and degree of freedom is 13. The quality of the model is high and $R^2$ is not random. So, if the Number of internet users in Russia rise by 1 man per 100 per year, the volume of government spending on digitalization will rise by 2,88 billion rubles. If the Rate of rising population will rise on 1% per year, the volume of government spending on digitalization will rise by 78,22 billion rubles. If the Publications in the field of ICT and Digitalization will grow, the volume of government spending on digitalization will rise by 120 thousand rubles. However, there are variables, by which we get strange coefficients. That is may be connected to the phenomenon of multicollinearity.

T-test

Table 2. The table of t value for variables.

| | $a_4$ | $a_3$ | $a_2$ | $a_1$ | $a_0$ |
|---|---|---|---|---|---|
| **t value** | \|-2,34\| > $t_{crit}$ | \|4,08\| > $t_{crit}$ | \|1,76 \| < $t_{crit}$ | \|2,44\| > $t_{crit}$ | \|3,62\| > $t_{crit}$ |

We may conclude that according to T-test, we consider some of the variables insignificant for the analysis. These variables are: Publications in the field of ICT and Digitalization – Pt. So, we are going to conduct further analysis without these variables. After introducing one more t-test analysis, we find out that the Rate of rising population – Rt is also insignificant for my analysis. So, after this, let's see the results of a new t-test:

Table 3. The table of t values for a new number of variables.

| | T-Test | | |
|---|---|---|---|
| t | -2,29 | 16,92 | 10,89 |
| $t_{critical}$ | 2,13 | | |

Now, as some variables we decided to be insignificant for the analysis of the volume of government spending on digitalization of the economy, we may introduce a new model:

$$\begin{cases} GSD_t = a_0 + a_1 \cdot N_t + a_2 \cdot A_t + \varepsilon_t \\ E(\varepsilon_t) = 0;\ \sigma(\varepsilon_t) = const, df = 14;\ df_1 = 2; df_2 = 15 \end{cases}$$

And now let's see the adjusted form of the new model:

$$\begin{cases} GSD_t = 202{,}18 + 7{,}24 \cdot N_t - 0{,}54 \cdot A_t + \varepsilon_t \\ (14{,}22)\ \ (2{,}7)\ \ (0{,}73) \\ E(\varepsilon_t) = 0;\ \sigma(\varepsilon_t) = const, df_1 = 2; df_2 = 15; df = 17 \\ F_{critical} = 3{,}2;\ t_{critical} = 2{,}13 \\ GQ = 2{,}13.\ DW = 0{,}97 \end{cases}$$

And the new and final regression table for the model is following:

Table 4. Regression Table for the final model.

| | $a_2$ | $a_1$ | $a_0$ |
|---|---|---|---|
| | -0,54 | 7,24 | 202,18 |
| $S_{ai}$ | 0,23 | 0,43 | 18,57 |
| $R^2$ | 0,98 | 32,83 | - |
| | 313,01 | 15 | - |
| | 674737,64 | 16167,16 | - |
| | ESS | RSS | |

GQ – test

We would get $F_{GQ\ constant}$ = 2,13, which may be more than both GQ (0,03) and $GQ^{-1}$ (0,07), that my results with the positive conduction of the GQ – test. The heteroskedasticity is considered to be insignificant. So, the null hypothesis of Gauss - Markov Theorem is rejected.

Durbin – Watson statistical test

In order to come up with the results for this test, we need to get residuals from the Data Analysis folder and download all the information about residuals partly for our data. The first table is for the period 2000 – 2009, so the $RSS_1$ equal to 802,19.

Table 5. Residuals for the period 2000 – 2009.

| *c* | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| Regression | 2 | 32366,12 | 16183,06 | 141,22 | 6,5E-05 |
| Residual | 7 | **802,19** | 114,60 | | |
| Total | 9 | 33168,3 | | | |

Next table is for the period 2009-2018, representing mainly the $RSS_2$, which equals to 23774,93.

Table 6. Residuals for the period 2009 – 2018.

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 2 | 206893,11 | 76095,76 | 30,46 | 2,4E-06 |
| Residual | 7 | **23774,93** | 396,7958 | | |
| Total | 9 | 230668 | | | |

The Durbin-Watson statistic will always have a value between 0 and 4. A value of 2.0 means that there is no autocorrelation detected in the sample. In order to define the values for $d_L$ and $d_U$. For our model we have obtained the following results: DW = 0,97; $d_L$=0,7; $d_U$=1,25. If DW constant lies between dl and du or 4-du and 4-dl, there is no information about autocorrelation. According to the number of n equations and the number of k explanatory variable choose values $d_L$ and $d_U$. Since $d_L$< DW < $d_U$, then there is no a sufficient basis for making decisions about the absence of autocorrelation of random residuals. As result, we have got:

Table 7. DW table for the model.

| | | | | | | |
|---|---|---|---|---|---|---|
| | | DW | 0,97 | | | |
| 0 | dl | du | 2 | 4-du | 4-dl | 4 |
| 0 | 0,7 | 1,25 | 2 | 2,75 | 3,3 | 4 |

## Results

Goals set at the beginning of work were achieved: analysis of the volume of government spending on the digitalization was performed for the period 2000 – 2018. Major indicators of economic performance were regarded.

Instruments for constructing a mathematical model were analyzed, theoretical aspects of model testing such as types of scatter diagrams, Goldfeld - Quandt testing, Durbin – Watson testing, and all the criteria were explained in detail. So, the F – test has showed that the quality of the model is high and $R^2$ is not random. T – test provided the analysis of the significance of the coefficients and concluded that, according to T-test, we consider some of the variables insignificant for the prior analysis. With the t-testing we have obtained a new model. GQ – test conducts, that the test on the heteroskedasticity is considered to be insignificant. So, the null hypothesis of Gauss - Markov Theorem is rejected.

While conducting DW – testing, we have got that according to the number of n equations and the number of k explanatory variable and chosen values $d_L$ and $d_U$, so since $d_L$< DW < $d_U$,

then there is no a sufficient basis for making decisions about the absence of autocorrelation of random residuals.

## Conclusion

Econometric model for the Volume of Governmental Spending on Digitalization was constructed. Full comprehensive analysis of the state for the current period was implemented, trends and prospects were identified, and connections between Number of internet users in Russia (billion, persons) and Number of mobile phones, per 100 people were understood and other variables.

The goal was to determine which variables and in which degree influence on the economic condition mostly and it was achieved.

In compliance with our expectations at the beginning of this work, certain independent variables had to be eliminated from the initial dataset during the process of model construction. This allows us to make conclusion that among various factors chosen, only 2 indicators, which are Number of Internet Users and Number of mobile phones per 100 people have the strongest effect on the Volume of Government Spending on Digitalization.

Analysis of the interdependence of factors showed that many of them have strong correlation dependence, so some factors were excluded and left to those whose correlation with the strongest option. The calculated linear regression model was tested for adequacy. Multicollinearity factors after excluding autocorrelated factors are considered to be insignificant. It does not distort the result of the study. According to records of regression statistics, it is adequate and can be used to predict the real situation for the future periods of making the governmental budgets.

The model is adequate as Y is within the boundaries. According to the results, we see that $N_t$ may increase by 10% - with the grow of $N_t$, $GSD_t$ grows systematically. Change in $A_t$may result with decrease by 4% in neutral and even by 10% in the worst occasion. This may result with 1) growing number of technological perspectives such as free Internet and PC in public places, provided by the government; 2) the growing ability of dashboards and personal computers, ipods and other devices other than mobile phones, which allow to use all technological perspectives.

**References**

1. Accenture // Platform Economy: Technology-driven Business Model Innovation from the Outside In - 2017, Dublin.

2. Aris Pagoropoulos, Daniela C.A. Pigosso, Tim C.McAloone The Emergent Role of Digital Technologies in the Circular Economy: A Review // Procedia CIRP - 2017. №64.

3. Jiaochen Lianga, Stephan J.Goetz Technology intensity and agglomeration economies // Research Policy - 2018. №47.

4. Manyika J. et al // Lions Go Digital: The Internet's Transformative Potential in Africa - 2019. New York, NY: McKinsey Global Institute

5. Manyika J., Roxburgh C. // The Great Transformer: The Impact of the Internet on Economic Growth and Prosperity - 2011. New York, NY: McKinsey Global Institute.

6. Murphy, J.T. and Carmody, P. // Information Revolution: Technical Regimes and Production Networks in South Africa and Tanzania – 2018. Chichester, UK: Wiley. 249 pp. ISBN: 9781118751336

7. Zavadska M., Morales, L., Coughlan, J. // Digitalization and necessary volumes - 2019.

8. Трегуб И.В. Моделирование инфляционных процессов в российской экономике. // Фундаментальные исследования. 2009. № 1. С. 86-87.

9. Трегуб И.В. Особенности инвестирования в инновационные проекты/ // Экономика. Налоги. Право. 2013. № 3. С. 28-32.

10. Трегуб И.В. Моделирование ценообразования на дополнительные услуги сотовой связи на рынке телекоммуникаций/ Диссертация на соискание ученой степени доктора экономических наук / Финансовый университет при Правительстве Российской Федерации. Москва, 2010

11. Трегуб И.В. International Diversification. Москва, 2015.

12. Трегуб И.В. Эконометрика на английском языке / Учебное пособие / - М.: РуСАйс, 2017. 110 с.

13. Tregub I.V. Econometrics. Model of Real System, - M.: PSTM. - 2016. 164 p.