Econometric Modeling of the Use of Digital Technologies in Improving the Infrastructure of the Railways of Uzbekistan

Shohruh Ergashev[1,2], Oybek Achilov[1,2], Marguba Agzamova[1,2] and Muhabbat Kurbanova[1,2]
[1]*Departments of Business Management and Accounting, Tashkent State Transport University, Temiryolchilar Str. 1, 100167 Tashkent, Uzbekistan*
[2]*University of Diyala, 32009 Baqubah, Diyala, Iraq*
shoxruhergashev1993@gmail.com, achilovoybek88@gmail.com, mmarguba77777@gmail.com, kmukhabbat2507@gmail.com

Keywords: Digital Economy, Railway Transport, Infrastructure, Passenger Transport, Multifactor Econometric Model, Breush- Pagan Test.

Abstract: This study examines the growing global trend of integrating digital technologies into various economic sectors, with a specific focus on the Republic of Uzbekistan. Following Presidential Decision No. PQ-4699, adopted on April 28, 2020, significant steps have been taken to promote the digital economy and implement electronic government solutions. In this context, the railway transport sector, particularly JSC "Uzbekistan Railways," has become a key area for digital transformation. The research highlights the strategic importance of introducing digital software to address infrastructure challenges and enhance operational efficiency. Drawing on international perspectives, the study identifies key benefits of digitalization in railway systems, including cost reduction, improved management, and business model innovation. An econometric model was developed to assess the relationship between the adoption of digital technologies and economic outcomes in the railway sector. The analysis demonstrates a strong positive impact, where a 1% increase in investment in digital software corresponds to a 2.53% rise in passenger traffic. Key indicators used in the model include total revenue from railway transport (d), total freight transported (freight), and the number of passengers. The variable (dtx) was used to represent product costs associated with digital supply development and infrastructure management. The findings affirm the critical role of digital technologies in advancing railway infrastructure and support broader economic growth initiatives in Uzbekistan.

1 INTRODUCTION

The main difference between the current economy and the digital economy is that in the former, GDP is the main indicator of its efficiency. Businesses first produce products, after which they look for sales markets. Forecasting plays a key role in the digital economy: the first the demand forecast is determined, then the supply is formed. Digitization of the economy of control systems allows more informed business entities to make the right decisions. Today's realities clearly show the increasing value of accurate forecasting based on mathematical models based on large volumes of data. Thus, these capabilities are used, for example, to adjust the balance of supply and demand and to obtain detailed information about reliability to reduce sales and logistics costs. The following should be noted: the closer the date of the predicted event, the higher the probability that it will maximally correspond to the obtained forecast. The issue of modernization of the infrastructure of the transport sector takes a serious place in many foreign strategies for the development of the digital economy [1], [2], [3].

The implementation of digital technologies in the infrastructure of railway transport is a current topic that is being considered by many scientists and researchers, and the general conclusions in this regard are as follows: reducing costs, increasing the ability to manage the system, changing the existing business process model of railway transport operations through a digital transformation program, and justifying opportunities [4], [5].

To date, in the scientific work of local and foreign scientists, great attention is paid to the use of mathematical methods to improve the efficiency of transport management, the use and application of information and digital technologies in transport systems, including the organization and regulation of railway traffic flows.

In the conducted studies [1], issues of optimization of management in the transport system using digital technologies were considered. Methods of mathematical description of railway transport management processes have been developed.

Dimitrov I.D. as researched, digital control methods were proposed to allow effective use of existing transport resources [2].

A group of authors [3] considered the organization of rapid data exchange and information transfer between management subjects, as well as the evaluation of the effectiveness of the use of information technology in management.

2 ANALYZING THE N-INDEX AND LINEAR REGRESSION MODELS

In the Resolution of the President of the Republic of Uzbekistan dated April 28, 2020 "On measures for the widespread introduction of digital economy and electronic government" No. PQ-4699, a number of effective results were achieved based on the tasks set in the framework of the wider involvement of digital technologies in the activities of JSC "Uzbekistan Temir Yollari" [6]. The author's hypothesis was formed for the purpose of econometric modeling of the use of digital technologies in improving the infrastructure of the railways of Uzbekistan. According to it, the costs incurred for the development and improvement of software aimed at the wide implementation of digital technologies (dtx) were chosen as a factor indicator. The total revenue of the railway transport industry (d), the total volume of freight transported by railway (freight), and the number of passengers using the railway were selected as the outcome indicators (Fig. 1).

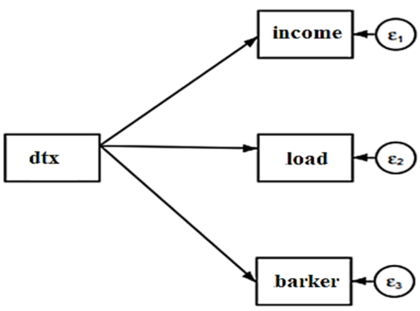

Figure 1: The impact of the use of digital technologies on the railway infrastructure.

In scientific research, we use n-index and linear regression models. To do this, we used the method of least squares to create process regression models:

$$F = \sum(Y - Y_x)^2 \to m \text{ or } F = \sum(Y - \beta_0 - \beta_1 x - \beta_2 x^2 - \cdots - \beta_k x^k)^2 \to min$$

if we take a special derivative from it, the following system of equations is formed. The level multifactor econometric model looks like this:

$$Y = \beta_0 * x_1^{\beta_1} * x_2^{\beta_2} * \ldots * x_n^{\beta_n} \quad (1)$$

Hence, Y is the resulting factor; $x_1, x_2, \ldots, x_n$ - influencing factors.

If we substitute the natural logarithm in the model presented in the (1), then we will have the following form:

$$\ln(y) = \ln(\beta_0) + \beta_1 \ln(x_1) + \beta_2 \ln(x_2) + \ldots + \beta_n \ln(x_n) \quad (2)$$

In the model ln(y) = y`, ln(ß0) = ß0', $\ln(x_1) = x_1', \ln(x_2) = x_2', \ldots, \ln(x_n) = x_n'$ if we make the designations, then we will have the following appearance:

$$y' = \beta_0' + \beta_1 x_1' + \beta_2 x_2' + \ldots + \beta_n x_n' \quad (3)$$

In the above model, the following system of normal equations is created to find unknown parameters (b_0,)`(b`_1,…,(b_ n)` )`

$$\begin{cases} n\grave{\beta}_0 + \grave{\beta}_1 \sum x'_1 + \grave{\beta}_2 \sum x'_2 + \cdots \grave{\beta}_n \sum x'_n = \sum y' \\ \grave{\beta}_0 \sum x'_1 + \grave{\beta}_1 \sum x'^2_1 + \grave{\beta}_2 \sum x'_1 x'_2 + \cdots \grave{\beta}_n \sum x'_1 x'_n = \sum x' \\ \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \\ \grave{\beta}_0 \sum x'_n + \grave{\beta}_1 \sum x'_n x'_1 + \grave{\beta}_2 \sum x'_n x'_2 + \cdots \grave{\beta}_n \sum x'^2_n = \sum x'_n y' \end{cases} \quad (4)$$

If this system of normal (4) is analytically solved by several methods of mathematics, then the values of unknown parameters (b_0,)`(b`_1,…, (b_ n)` )` are found.

The main goal in any econometric modeling is to determine the beta coefficient, for which the beta coefficient should be BLUE (best linear unbiased estimator), i.e. F-test, t-test, in addition to statistical significance, the model meets the requirements of Gauss Markov, especially the requirements of heteroscedasticity and autocorrelation [7]. The model residuals should not be correlated with the corresponding model values, nor should the lags of the residuals be systematically related. To get rid of heteroscedasticity in general, model data were normalized, that is, logarithmzed.

2.1 Econometric Modeling Approach to Railway Income under Digital Transformation

Econometric methods were used in the "STATA" program to analyze the changes of revenues and expenses in the field of railway transport using cloud-based computational approaches [16], and the relevant coefficients were calculated in Table 1 [8].

The analysis of the results obtained in Table 1 shows that the coefficient of determination in the regression model is $R^2 = 0.95$; F count 107.20; (when, F tab=2.17) is equal to. When we compared each coefficient using by given criterion, we found that the calculated values exceeded the values in the table.

Typically, the coefficient of determination ranges from [0;1]. The closer it is to 1, the stronger the relationship. In this context, a coefficient of determination of 0.95 indicates a strong connection between the economic indicators in the model. To compare models with varying numbers of factors, and ensure the number of factors doesn't influence the R^2 statistic, a corrected version of the coefficient of determination is generally employed:

$$R^2_{\text{plain}} = 1 - \frac{s^2}{s_{\hat{y}}^2}. \quad (5)$$

In this instance, the adjusted coefficient of determination is 0.94, and its proximity to R^2 suggests that the change in the number of influencing factors in the model is acceptable. To assess the statistical significance of the constructed multifactor econometric model and its alignment with the studied process, we use Fisher's F-criterion. The true value of the F-criterion is calculated using the following (6):

$$F_{\text{count}} = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m}, \quad (6)$$

here: R^2 - determination coefficient; n - The total count of data points; m - the total count of variables. The actual value of the F-criterion is F_ count =107.20. When the actual F-statistic value surpasses the critical value from the statistical table, it suggests that the multifactor econometric model effectively explains the variability in the dependent variable. This implies that the factors in the model have a substantial influence on the studied process, making the model statistically significant and appropriate for further analysis or forecasting.

2.2 Strengths and Weaknesses of Econometric Modeling of Railway Income

Advantages of the econometric modeling of income changes based on digital technologies:

1) To determine the tabular value of the F-criterion, we first calculate the degrees of freedom and the significance level. With a significance level and degrees of freedom of k1=2 and k2=20-2-1=17, the F-criterion value from the table is F table = 2.17;
2) F calculation>>F table satisfies the condition, which means that the calculated value of F_ count is statistically significant than the value in the table, and can be used in forecasting for future periods;
3) The student's t-test is used to check the reliability of multifactor econometric model parameters and correlation coefficients. In this case, their value is compared with the values of random errors [9], [15].

Disadvantages of the econometric modeling of income changes based on digital technologies:

1) By comparing the calculated and table values of Student's t-test, we accept or reject the H_0 hypothesis. For this, we find the tabular value of the t-criterion based on the conditions of the selected reliability probability and degree of freedom. Here are the number of observations and factors [10], [13], [14].

Table 1: Econometric modeling of income changes based on digital technologies.

Variable	Model coefficients	Standard errors	t-Student test	P-value
β_1	2.13	0.20	10.65	0,0000
C	-1.17	0.17	-6.88	0,0000
R^2-Coefficient of determination	0,95	F - Fisher's criterion	107.20	
Adjusted R^2-Coefficient of Determination	0,94	Prob (F - Fisher's criterion)	0,0000	
Breusch Pagan test	1.44 (0.23)	DW- Durbin-Watson criterion	2.02	
Akaike's information criterion	-13.97	Schwartz's information criterion	-14.08	

2) When the reliability probability and the degree of freedom are d. f.=20-2-1=17, the table value of the t-criterion is equal to t table=2.14. We assume that t_ count is equal to 10.65 for the beta coefficient, that the |t_ count |≫t _tab condition is satisfied, and that the coefficient is statistically significant.
3) There are several tests for heteroscedasticity, the main one being the Breusch-Pagan test. In this test, the probability level is less than 0.05, so the main hypothesis is accepted, and the alternative hypothesis is rejected. It is expressed that the model's residuals are not associated with the corresponding values [9].

3 ANALYSIS OF THE ECONOMETRIC MODELING

The Breusch is 1.44 in the Pagan test, and the probability is 0.23. The model has a homogeneity property, the beta coefficient is BLUE.

We use the Darbin -Watson (DW) criterion to check autocorrelation in the residuals of the resulting factor according to the model [9]

$$DW = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2} = \frac{\sum_{t=2}^{T} e_t^2 + \sum_{t=2}^{T} e_{t-1}^2 - 2\sum_{t=2}^{T} e_t e_{t-1}}{\sum_{t=1}^{T} e_t^2} = \quad (7)$$

$$= 2 - 2\frac{\sum_{t=2}^{T} e_t e_{t-1}}{\sum_{t=1}^{T} e_t^2} \approx 2(1-\rho_1),$$

This is the first-order correlation coefficient ρ_1. If there is no autocorrelation among the residuals, the resulting factor approaches zero in the case of positive autocorrelation and 4 in the case of negative autocorrelation.

$$\begin{cases} \rho_1 = 0 \rightarrow DW = 2; \\ \rho_1 = 1 \rightarrow DW = 0; \\ \rho_1 = -1 \rightarrow DW = 4. \end{cases}$$

The calculated DW is compared with the DW in the table. If there is no autocorrelation in the residuals of the resulting factor, then the value of the calculated DW criterion will be around 2. In this example, the value of the calculated DW criterion expresses the same dimension with 2.02. This suggests that there is no autocorrelation in the residuals of the resulting factor.

Therefore, we chose this following model of regression as adequate:

$$lnd = -1.17 + 2.13lndtx; \quad (8)$$

The coefficient of the model is BLUE and it is statistically significant according to F test, t test and Gauss Markov. Akaike and Schwartz criteria indicate that the model is optimal [11]. The beta coefficient is 2.13. That is, a 1% increase in expenses for the development of digital software in the railway system leads to a 2.13% increase in revenue from this sector.

3.1 Econometric Modeling of Rail Freight

To further test our hypothesis, we model the dependence on rail freight based on digital software development (Table 2).

The analysis of the results obtained in Table 2 shows that the coefficient of determination in the regression model is R^2 = 0.857; F count= 30.02; (when, F tab=2.17) is equal to. When we compared each coefficient according to the student criterion, it was found that the calculated values are greater than the table values. In this case, the coefficient of determination is equal to 0.857, which means that there is a sufficiently strong relationship between these economic indicators in the model.

Table 2: Econometric modeling of changes in railway freight on the basis of digital technologies.

Lnyuk	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Lndtx	.389	.071	5.48	.003	.206	.571	***
Constant	8.342	.568	14.70	0	6.883	9.802	***

Mean dependent var	11.452	SD dependent var	0.066
R-squared	0.857	Number of obs	7
F-test	30.032	Prob > F	0.003
Akaike crit. (AIC)	-28.882	Bayesian crit. (BIC)	-28.990
*** p<.01, ** p<.05, * p<.1			

The adjusted coefficient of determination equal to the value of 0.82 and its closeness to R^2 means that the change in the number of influencing factors of the model is accepted around the values.

The calculated value of the F-criterion is F count = 30.02. To determine the tabular value of the F-criterion, we first calculate the degrees of freedom and the significance level. Given the significance level and degrees of freedom k1=2 and k2=20-2-1=17, the table value of the F-criterion is F table = 2.17.

F calculation>>F table satisfies the condition, which means that the calculated value of F-criterion is statistically significant than the value in the table, and can be used in forecasting for future periods.

By comparing the calculated and tabular values of Student's t-test, we determine whether to accept or reject the null hypothesis H_0. To do this, we find the tabular value of the t-criterion based on the chosen level of confidence (α) and degree of freedom ($\text{d.f.} = n - m - 1$). Here, n- represents the number of observations, and m- denotes the number of factors.

When the probability of reliability and the degree of freedom are d. f.=20-2-1=17, the table value of the t-criterion is equal to t_ tab=2.14. We assume that t_ count is equal to 5.48 for the beta coefficient, that the |t_ count |≫t_ condition is satisfied, and the coefficient is statistically significant.

We found the model to be statistically significant by F test, t test, coefficient of determination. Now we check the Gaussian Markov conditions.

Breusch-Pagan / Cook-Weisberg test for heteroscedasticity. Ho: constant variance. Variables: fitted values of lnyuk chi2(1) = 0.63 Prob > chi2 = 0.4284.

When heteroskedasticity was tested using the Breusch Pagan test using STATA software, the chi2 (X square) value was 0.63, and the probability level was 0.42. In this case, we accept the main hypothesis and reject the alternative hypothesis. There is no heteroscedasticity in the model, the model residual is a free variable.

3.2 Econometric Modeling of Rail Passenger Traffic

In the analysis of the third part of our hypothesis, we model the relationship between the development of digital software in the system of passenger transportation by railway (Table 3).

The analysis of the results obtained in Table 3 shows that the coefficient of determination in the regression model is R^2 = 0.49; Faccount= 19.88; (when, F table=2.17) is equal to. When we compared each coefficient according to the student criterion, it was found that the calculated values are greater than the table values.

There are a number of tests for heteroscedasticity, the main one being the Breusch Pagan test, where the probability level is less than 0.05, the main hypothesis is accepted and the alternative hypothesis is rejected. Model residuals are represented as uncorrelated with fit values. Breusch is 1.65 in Pagan's test and the probability is 0.95. The model has homogeneity property, the beta coefficient is BLUE.

A 1% increase in spending on the development of digital software in the railway system affects a 2.53% increase in railway passenger traffic (Fig. 2).

Table 3: Econometric modeling of rail passenger changes based on digital technologies.

Lnyoovchi	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Lndtx	2.53	0.145	17.44	.001	2.473	8.413	***
Constant	29.807	9.162	3.25	.023	6.256	53.358	**

Mean dependent var	9.563	SD dependent var	0.566
R-squared	0.494	Number of obs	7
F-test	19.884	Prob > F	0.000
Akaike crit. (AIC)	10.057	Bayesian crit. (BIC)	9.949
*** *p<.01*, ** *p<.05*, * *p<.1*			

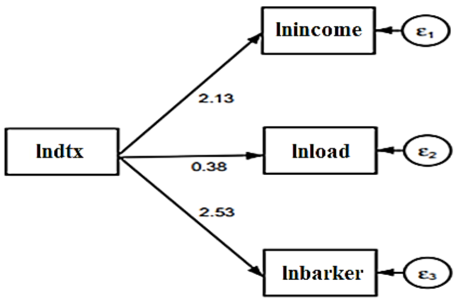

Figure 2. Results of the impact of the use of digital technologies on the railway infrastructure.

All three coefficients are BLUE and statistically significant under all conditions of F test, t test and Gauss Markov. Akaike and Schwartz criteria indicate that the models are optimal.

4 CONCLUSIONS

Modeling the process of development of service industries and its correlation with performance should be considered as a way to determine the direction of development of the service industry using these factors. This problem can be solved with the help of modern digital technologies and tools, which include models for predicting the relationship between the costs and revenues of passenger transportation. Thus, the main requirements for the railway transport development model include: introduction of organizational management mechanisms; using the efficiency of the transport sector as an objective function, taking into account the impact factors; expansion of means aimed at improving the efficiency of passenger transportation.

A 1% increase in digital software processing costs in the railway system will lead to a 2.13% increase in railway freight transportation, a 0.38% increase in railway passenger transportation, and a 2.51% increase in railway total transportation.

The high coupling type of the tested model indicates the effectiveness and performance of the developed model. The study showed that the information obtained helps to reduce ineffective subjective management decisions. In the process of forecasting, it is necessary to identify possible alternative methods for the implementation of tasks aimed at increasing the efficiency of management of railway transport organizations and justify their effectiveness.

This proposal will lead to the formation of a "logistic transport chain" that integrates all participants in the freight transportation process into a unified management system on the railways of the Republic of Uzbekistan, ensuring the optimization of all cargo flows. Additionally, it provides an opportunity to reduce the annual volume of operating costs related to wagon turnover by an average of 17-18%. Based on the e-model, it is possible to develop forecast values of the pessimistic, optimistic, and average scenario variants for the impact of infrastructure improvement costs on the volume of freight and passenger transportation by rail until 2027.

REFERENCES

[1] Transport Investment, Moving Britain Ahead. Presented to Parliament by the Secretary of State for Transport by Command of Her Majesty. [Online]. Available: www.gov.uk/government/uploads/system/uploads/attachment_data/file/624990/transport-investment-strategy-web.

[2] I. D. Dimitrov, "The impact of the digital economy on the development of the transport industry in Russia," Transport of the Russian Federation. Journal of Science, Practice, Economics, no. 6, pp. 73–77, 2017.

[3] R. Howe-Teo, "Singapore's Smart Mobility 2030: Big Data and Car-Lite Society." [Online]. Available: https://www.nscs.gov.sg/public/download.ashx?id=1005.

[4] Digital Australia: State of the Nation. The 2020 edition. [Online]. Available: https://digitalaustralia.ey.com.

[5] U.S. Department of Transportation, Strategic Plan for FY 2020–2024. Draft for Public Comment, Oct. 19, 2021.

[6] Resolution of the President of the Republic of Uzbekistan, "On measures for the widespread introduction of the digital economy and electronic government," No. PP-4699, Apr. 28, 2020.

[7] D. Asteriou and S. G. Hall, Applied Econometrics. A Modern Approach Using EViews and Microfit, revised ed., New York: Palgrave Macmillan, 2007, p. 397.

[8] D. Rasulev, Sh. Nurullaeva, N. Rumetova, and M. Muminova, Fundamentals of Econometrics, Textbook, Tashkent: Economics, 2019, 248 p.

[9] R. N. Williams, "Elementary guide to CRC algorithms of error detection," Aug. 19, 1993.

[10] S. M. Sultanova, "Improvement of financial flow management of a railway transport enterprise," Turkish Online Journal of Qualitative Inquiry (TOJQI), vol. 12, no. 6, pp. 7117–7121, 2021. [Online]. Available: https://www.tojqi.net/index.php/journal/article/view/2993/1994.

[11] N. Babaxanova and S. Sultanova, “Statistical model of operational costs indicators in the intelligent transport system,” in Proc. Int. Sci. Conf. Construction Mechanics, Hydraulics & Water Resources Engineering, Topic 5: Engineering Materials Science, Intelligent Transport Systems and Transport Logistics (CONMECHYDRO 2021). [Online]. Available: https://us02web.zoom.us/j/7514330617?pwd=TEoycnExTlVqYXV2T3RMRFZPVWljdz09.

[12] C. Dougherty, Introduction to Econometrics, 3rd or 4th ed., Oxford: Oxford University Press, 2011.

[13] J. F. Kurbanov, N. B. Yaronova, O. R. Achilov, and J. Xidirov, “The module development for remote control signals of an automatic locomotive signaling system,” Proc. SPIE – International Society for Optical Engineering, vol. 13065, Dushanbe, Tajikistan, Feb. 20, 2024, doi: 10.1117/12.3024860.

[14] A. I. Khalil and M. M. Khalil, “Image data compression based on two hybrids algorithms,” Turkish Journal of Computer and Mathematics Education (TURCOMAT), vol. 12, no. 9, pp. 1403–1415, 2021.

[15] A. Khomonenko and M. M. Khalil, “Quantum computing in controlling railroads,” E3S Web of Conferences, vol. 383, 2023.

[16] M. M. Khalil, A. D. Khomonenko, and M. D. Matushko, “Measuring the effect of monitoring on a cloud computing system by estimating the delay time of requests,” Journal of King Saud University – Computer and Information Sciences, 2021.