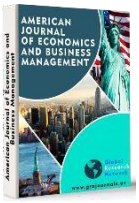

American Journal of Economics and Business Management

Vol. 9 Issue 4 | **pp.** 570-577 | **ISSN:** 2576-5973
Available online @ https://www.globalresearchnetwork.us/index.php/ajebm

Article

Application of Econometric Methods in Processing Statistical Data Based on Cloud Technologies

Ablazov Lazizbek Abdiqosimovich

1. Senior Lecturer, Department of Digital Economy, Tashkent State University of Economics
* Correspondence: lazizbekablazov@gmail.com

Abstract: This paper develops a multifactor econometric model to assess the determinants of costs associated with processing statistical data using cloud technologies. The empirical analysis is based on semiannual data for 2010-2025 obtained from the Statistics Agency of the Republic of Uzbekistan. The dependent variable is the cost of processing statistical data in cloud environments, while the explanatory variables include data volume, internet connection speed, level of automation, and the number of cloud users. The results substantiate the relationships among variables, confirm the statistical significance of the model, and demonstrate its forecasting capability. In addition, projections for 2026–2030 are produced, enabling an evaluation of how the adoption of cloud technologies influences both the scale and the dynamics of data-processing costs.

Keywords: cloud technologies, statistical data, econometric model, multiple regression, ordinary least squares (OLS), correlation analysis, multicollinearity, forecasting, digital economy, data volume, internet speed, level of automation, number of users.

1. Introduction

The development and construction of a multifactor econometric model aimed at improving the methodology of processing statistical data using cloud technologies represents a relevant scientific and practical task. Addressing this problem contributes to enhancing the efficiency of statistical analysis, expanding analytical capabilities, and forming a modern data processing infrastructure that meets the requirements of the digital economy.

The application of econometric models makes it possible not only to assess the current state of the sector, but also to identify its problematic aspects, as well as to determine priority directions of state policy regarding the use of modern technologies in data processing [1].

Based on the objectives of the dissertation research, the construction of multifactor econometric models for the costs of data processing in cloud technologies in Uzbekistan is aimed at evaluating the key factors influencing the development of this field.

Literature Review

Foreign literature extensively covers the issues of digital platforms and ecosystems. In particular, studies conducted by Annabelle Gawer substantiate that digital platforms have become the dominant institutional form of the modern economy. The author demonstrates that data flows through platforms and their processing mechanisms play a central role in the creation of economic value. At the same time, these studies do not sufficiently address the econometric evaluation of the costs of processing statistical data at the micro level (i.e., at the level of enterprises or public systems) [2].

The Russian scholar M.N. Tolchinskaya, in her research, systematically examines the impact of digital transformation on the efficiency of the economy and business. The author emphasizes that digital technologies, including Big Data and cloud services, contribute to the

Citation: Abdiqosimovich, A. L. Application of Econometric Methods in Processing Statistical Data Based on Cloud Technologies. American Journal of Economics and Business Management 2026, 9(4), 570-577

Received: 10th Jan 2026
Revised: 21th Feb 2026
Accepted: 08th Mar 2026
Published: 15th Apr 2026

Copyright: © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

formation of new models of enterprise activity. However, this approach is primarily theoretical in nature, and the aspects of applied econometric modeling and forecasting are not sufficiently developed [3].

A deeper analysis of domestic scientific literature indicates that the studies conducted by Sh. Shodmonov on the evaluation of the effectiveness of the digital economy and electronic services are of particular importance. In his works, the author provides a comprehensive analysis of the institutional foundations of the digital economy, its integration into economic systems, and the mechanisms for the development of electronic services [4].

Among domestic researchers, the studies conducted by O. Astanaqulov emphasize the significant role of information technologies in enhancing investment attractiveness in the context of the digital economy. The author substantiates that the development of digital infrastructure is a key factor in improving economic efficiency [5].

2. Materials and Methods

This study applies an econometric approach to identify and evaluate the factors influencing the costs of processing statistical data using cloud technologies. The research methodology consists of several sequential stages, including descriptive statistics, correlation analysis, regression modeling, and forecasting procedures [6].

3. Results and Discussion

The database for the multifactor econometric models consists of semiannual data for the period 2010-2025 obtained from the Statistics Agency of the Republic of Uzbekistan. In the model, the dependent variable is the cost of processing statistical data using cloud technologies (Y, in UZS), while the explanatory variables include data volume (X_1, in MB), internet connection speed (X_2, in Mbit/s), level of automation (X_3, in %), and the number of users of cloud technologies (X_4) [7].

The ordinary least squares (OLS) method was employed to construct econometric models and assess the impact of the selected factors on data processing costs. As an initial step, descriptive statistical analysis was conducted for all variables. This analysis provides a general understanding of the level, variability, and distributional characteristics of the indicators describing cloud-based data processing costs and serves as a foundation for subsequent econometric modeling.

The results of the descriptive statistical analysis for the selected variables are presented in Table 1 [8].

Table 1. Descriptive statistics of factors influencing the cost of processing statistical data using cloud technologies.

	Y	X1	X2	X3	X4
Mean	6128880	16166.94	2.341129	1.212903	2.096774
Median	6176652	14213.00	2.200000	1.200000	2.000000
Maximum	22318722	49255.00	8.400000	4.000000	7.000000
Minimum	17785798	21158.00	4.375000	2.500000	3.000000
Std. Dev.	9342034.	18829.33	3.283440	1.466684	2.165715
Skewness	-0.139230	0.150243	0.076690	-0.380550	0.073789
Kurtosis	2.985733	2.225396	2.390566	2.882521	3.689660
Jarque-Bera	0.100419	0.891642	0.510125	0.766056	0.642488
Probability	0.951030	0.640298	0.774868	0.681794	0.725246
Sum	1.90E+08	501175.0	72.57500	37.60000	65.00000
Sum Sq. Dev.	2.62E+15	1.06E+10	323.4294	64.53484	140.7097
Observations	31	31	31	31	31

From the table, it is possible to observe the mean, median, maximum, and minimum values of each variable in the sample under study.

The Jarque-Bera statistic is used to test whether the variables follow a normal distribution. In addition, it provides probability values for each variable. If the probability value exceeds 0.05, the variable is considered to follow a normal distribution and can be included in the multifactor econometric model. Since the calculated probabilities for all variables are greater than 0.05, their normal distribution is confirmed.

Before deciding whether to include the selected variables in the multifactor econometric model, it is necessary to examine the existence of relationships among them. For this purpose, both partial and pairwise correlation coefficients are calculated.

The results of the correlation analysis are presented in Table 2. which shows the correlation matrix of the variables.[9]

Table 2. Correlation matrix of the variables.

Probability	Y	X1	X2	X3	X4
Y	1.000000				
X1	0.823189	1.000000			
	7.807791	-----			
	0.0000	-----			
X2	0.706187	0.506654	1.000000		
	5.371176	3.164666	-----		
	0.0000	0.0036	-----		
X3	0.677745	-0.293890	0.303870	1.000000	
	3.819942	-1.655767	1.717611	-----	
	0.0016	0.1086	0.0965	-----	
X4	0.691419	0.286373	0.339388	0.080398	1.000000
	4.104904	1.609579	1.942983	0.434360	-----
	0.0003	0.1183	0.0618	0.6672	-----

The correlation between the cost of processing statistical data using cloud technologies (Y) and the volume of data to be processed (X_1) is 0.8232, indicating a strong positive relationship between these two variables. The correlation between Y and internet connection speed (X_2) is 0.7062, which also reflects a strong positive relationship. The relationship between Y and the level of automation (X_3) is measured at 0.6777, suggesting a moderately strong positive correlation. Similarly, the correlation between Y and the number of cloud users (X_4) is 0.6914, indicating a moderately strong positive relationship between these variables.

Another method for testing the presence of multicollinearity among the explanatory variables is the calculation of the Variance Inflation Factor (*VIF*). The *VIF* values computed for each variable are presented in Table 3 [10].

Table 3. Measurement of multicollinearity among explanatory variables.

Variable	CoefficientVariance	CenteredVIF
X1	2522.029	1.959688
X2	8.39E+10	1.981723
X3	3.34E+11	1.573891
X4	1.13E+11	1.157602
C	1.60E+12	NA

If multicollinearity exists among the explanatory variables, the condition *Centered VIF >10* would be observed. However, as shown in Table 3.6, all calculated *VIF* coefficients for the explanatory variables affecting the cost of data processing using cloud technologies (Y) are less than 10. This indicates, similar to the correlation analysis, that there is no multicollinearity among the selected variables.

In addition, to test the stationarity of the dependent variable included in the multifactor econometric model, that is, to determine the absence of a unit root, the

Augmented Dickey-Fuller (*ADF*) test was conducted. The results of this test are presented in Table 4 [11].

Table 4. Augmented Dickey–Fuller test for stationarity.

Null Hypothesis: LNY has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=1

			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-4.015743	0.0027
Testcriticalvalues:	1% level		-3.724070	
	5% level		-2.986225	
	10% level		-2.632604	
*MacKinnon (1996) one-sided p-values.				

According to the results of the Augmented Dickey–Fuller (*ADF*) test, the null hypothesis of a unit root in the dependent variable (Y) is rejected at the 1%, 5%, and 10% significance levels. The *ADF* test statistic (−4.015743) is lower than the corresponding critical values, and the p-value is 0.0027. Therefore, the time series of the dependent variable (Y) is confirmed to be stationary.

Thus, taking into account the absence of multicollinearity among the explanatory variables and the absence of heteroskedasticity in the dependent variable, a multifactor econometric model is constructed to analyze the cost of processing statistical data using cloud technologies (Y) and its influencing factors (X_i). The general form of the model is as follows:

$$y = a_0 + a_1 x_1 + a_2 x_2 + ... + a_n x_n + \varepsilon, \quad (1)$$

where y – is the dependent variable, x_i – are the explanatory variables and ε – is the random error term.[12]

The unknown parameters of the multifactor econometric model (1) were estimated using the Ordinary Least Squares (*OLS*) method. The results of the estimation are presented in Table 5.

Table 5. Estimated parameters of the multifactor econometric model.

Dependent Variable: LNY
Method: Least Squares
Date: 01/21/26 Time: 17:19
Sample: 2010S1 2025S1
Includedobservations: 31

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	356.4008	50.21981	7.096818	0.0000***
X2	-580828.0	289606.9	-2.005574	0.0554**
X3	1332311.0	577786.4	2.305889	0.0293***
X4	947283.8	335579.0	2.822834	0.0090***
C	-4595027.0	1263995.0	-3.635322	0.0012***
R-squared	0.864067	Meandependentvar		6128880.
Adjusted R-squared	0.843155	S.D. dependentvar		9342034.
S.E. ofregression	3699791.	Akaikeinfocriterion		33.23214
Sumsquaredresid	3.56E+14	Schwarzcriterion		33.46343
Loglikelihood	-510.0982	Hannan-Quinncriter.		33.30753
F-statistic	41.31783	Durbin-Watsonstat		2.088836
Prob(F-statistic)	0.000000			

Note: *** - significant at the 0.05% level, ** - significant at the 0.1% level.

Using the results presented in Table 3.9, the estimated multifactor econometric model for the cost of processing statistical data using cloud technologies (Y) can be expressed in analytical form as follows:

$$\hat{Y} = -4595027{,}0 + 356{,}4008X_1 + 580828{,}0X_2 + 1332311{,}0X_3 + 947283{,}8X_4 \quad (2)$$

The estimated multifactor econometric model (2) shows that if the volume of data to be processed (X_1) increases by 1 MB on average, the cost of processing statistical data using cloud technologies (Y) increases by 356.4008 UZS on average. If internet connection speed (X_2) increases by 1 Mbit/s on average, the cost of processing statistical data using cloud technologies (Y) decreases by 580,828 UZS on average. If the level of automation (X_3) increases by 1 percent on average, the cost of processing statistical data using cloud technologies (Y) rises by 1,332,311 UZS on average. Likewise, an average increase of one unit in the number of cloud technology users (X_4) leads to an average increase of 947,283.8 UZS in the cost of processing statistical data using cloud technologies (Y).[13]

To evaluate the quality of the multifactor econometric model (2) constructed for the cost of processing statistical data using cloud technologies (Y), the coefficient of determination is used. The coefficient of determination shows what proportion of the variation in the dependent variable is explained by the factors included in the model. The calculated coefficient of determination (R^2, see Table 3.9) is 0.8641. This means that 86.41% of the variation in the cost of processing statistical data using cloud technologies (Y) is explained by the variables included in the estimated multifactor econometric model (2), while the remaining 13.59% is attributable to factors not included in the model.

To test the statistical significance of the multifactor econometric model (2) constructed for the cost of processing statistical data using cloud technologies (Y), Fisher's F-test is applied. For this purpose, the corresponding degrees of freedom and significance level are used to determine the critical tabulated value of the F-statistic. Based on these values, the critical value is $F_{table} = 2.69$. Since the calculated F-statistic is $F_{calc} = 41.3178$ and the condition $F_{calc} > F_{table}$ is satisfied, the multifactor econometric model (2) can be considered statistically significant. Therefore, it may be used for forecasting the future costs of processing statistical data using cloud technologies (Y).

The reliability of the estimated parameters in the multifactor econometric model (2) is then tested using Student's t-test. The critical value of the t-statistic is determined according to the selected confidence level and degrees of freedom, where n denotes the number of observations and k denotes the number of explanatory variables. At the chosen confidence level, the tabulated value of the t-statistic is $t_{table} = 2.059$.

As shown in Table 3.9, the calculated t-statistics for all explanatory variables included in the multifactor econometric model exceed the tabulated value at the relevant significance levels. This confirms that all variables are statistically reliable and can therefore be retained in the multifactor econometric model.

To test for autocorrelation in the residuals of the multifactor econometric model (2) for the cost of processing statistical data using cloud technologies (Y), the Durbin-Watson (DW) test is used. The calculated DW statistic is compared with the lower and upper critical bounds, DW_L and DW_U. If $DW_{calc} < DW_L$, residual autocorrelation is present; if $DW_{calc} > DW_U$, residual autocorrelation is absent. In this study, the lower critical bound is $DW_L = 1.16$ and the upper critical bound is $DW_U = 1.74$, while the calculated value is $DW_{calc} = 2.089$. Since $DW_{calc} > DW_U$, it can be concluded that there is no autocorrelation in the residuals of the dependent variable, namely the cost of processing statistical data using cloud technologies (Y) [14].

Based on these findings, the multifactor econometric model can be used to generate forecast estimates for future costs of processing statistical data using cloud technologies. For this purpose, trend models are first developed for each explanatory variable. A trend model represents a time-dependent function of the explanatory variable. The analysis of the trend models shows that all estimated coefficients are statistically significant and that their parameters are reliable. Therefore, the predicted values obtained from the trend models are substituted into the multifactor econometric model, allowing first the forecast

of the explanatory variables and then the forecast of the dependent variable. As a result, forecast values of the costs of processing statistical data using cloud technologies are obtained for the forecast period, as presented in Table 6.

Table 6. Forecast values of the costs of processing statistical data using cloud technologies and the factors affecting them.

Years	Cost of processing statistical data using cloud technologies (UZS), Y	Volume of data to be processed, MB, X1	Internet connection speed, Mbit/sek., X2	Level of automation, %, X3	Number of cloud technology users, units, X4
2010	83479036	237529	8,28	25,8	16
2011	93496522	265055	13,53	29,9	22
2012	104716104	295771	19,95	34,2	29
2013	117282037	330046	27,23	38,7	35
2014	131355880	368293	35,20	43,3	41
2015	147118588	410972	43,83	50,0	47
2016	164772818	458597	53,03	52,8	53
2017	184545556	511741	62,73	57,6	59
2018	206691023	571043	72,90	62,5	67
2019	231493944	637218	77,50	67,4	76
2020	259273218	711061	94,50	72,4	87
2021	290386005	793461	105,90	77,5	101
2022	325232326	885411	117,68	82,5	113
2023	364260203	988016	129,83	87,6	125
2024	407971429	1102511	142,33	92,7	134
2025	456927999	1230275	155,15	93,8	141
2026*	**426844570**	**1154539**	**156,38**	**94,7**	**144**
2027*	**450818914**	**1218330**	**166,25**	**95,6**	**152**
2028*	**474793258**	**1282122**	**176,12**	**96,4**	**161**
2029*	**498767602**	**1345913**	**185,99**	**98,2**	**169**
2030*	**522741947**	**1409704**	**195,87**	**99,4**	**178**

* Note: yillarda * indicates the forecast period.

Table 3.10 above presents the dynamics of the cost of processing statistical data using cloud technologies (Y) and its key influencing factors (X_1–X_4) for the period 2010–2025, as well as forecast values for 2026–2030. The table shows that during the historical period (2010–2025), all indicators demonstrate a steady upward trend, which generally continues in the forecast period. However, certain differences in growth rates and trajectories can be observed: in particular, a short-term decline in Y and X_1 is evident in 2026, followed by a return to growth in 2027–2030. This indicates the presence of short-term adjustment effects or base effects within the forecasting mechanism (trend models and multifactor regression model).

1) Dynamics of the dependent variable - costs (Y)

Over the period 2010-2025, the cost of processing statistical data using cloud technologies increased significantly, rising from 83,479,036 UZS in 2010 to 456,927,999 UZS in 2025, representing an approximate 5.5-fold increase. This trend can be explained by the expansion of digital services, increasing demand for computational infrastructure, and the growing need for data storage and processing.

In the forecast period (2026-2030), Y is expected to continue increasing overall, rising from 426,844,570 UZS in 2026 to 522,741,947 UZS in 2030. However, a temporary decline is observed in 2026 compared to 2025. This can be interpreted in two ways: (*i*) improved

cost efficiency due to optimization, automation, or infrastructure enhancements; and (*ii*) short-term fluctuations arising from the interaction between trend projections and the econometric model. From 2027 onward, costs return to a stable growth trajectory, increasing annually by approximately 23-24 billion UZS.

2) Changes in data volume (X_1)

The volume of data to be processed increased from 237,529 MB in 2010 to 1,230,275 MB in 2025, representing more than a fivefold growth. This reflects the expansion of digital services, electronic reporting, online platforms, and data generated from IoT and other digital sources.

In the forecast period, X_1 declines slightly in 2026 before increasing again to 1,409,704 MB by 2030. This temporary decrease corresponds with the decline observed in Y, as X_1 has a positive coefficient in the model and directly influences cost levels.

3) Growth of internet speed (X_2) and its impact on costs

Internet connection speed is one of the fastest-growing factors, increasing from 8.28 Mbit/s in 2010 to 155.15 Mbit/s in 2025, with a projected value of 195.87 Mbit/s by 2030. This growth enhances access to cloud infrastructure, reduces latency, and improves service continuity.

The negative coefficient of X_2 in the model indicates that higher internet speeds lead to lower data processing costs. This reflects increased operational efficiency, as faster data transmission reduces processing time and associated costs. Thus, X_2 acts as a compensatory factor, mitigating the cost increases driven by other variables such as X_1 and X_4.

4) Level of automation (X_3) and the saturation effect

The level of automation increased from 25.8% in 2010 to 93.8% in 2025 and is expected to reach 99.4% by 2030. This reflects a saturation effect, where growth naturally slows as automation approaches its maximum level.

Although automation improves efficiency, the positive coefficient of X_3 suggests that higher automation levels are associated with increased costs, particularly due to initial investments in IT infrastructure, software, integration, workforce training, and cybersecurity.

The forecast results for 2026-2030 confirm the continued expansion of cloud-based data processing. The growth of data volume (X_1), internet speed (X_2), automation (X_3), and the number of users (X_4) contributes to a long-term increase in the dependent variable (Y). The temporary decline observed in 2026 reflects short-term adjustment dynamics, while from 2027 onward a stable upward trend is re-established. Therefore, by 2030, the costs of processing statistical data using cloud technologies are expected to increase, emphasizing the importance of efficiency-oriented management decisions, resource optimization, and the implementation of advanced automation technologies [15].

4. Conclusion

A comprehensive analysis of the key factors influencing the cost of processing statistical data using cloud technologies was conducted, and a multifactor econometric model was developed. The results indicate that data volume, level of automation, and the number of users act as cost-increasing factors, whereas the growth of internet speed serves as a significant cost-reducing factor.

The forecast results for the period 2026-2030 demonstrate that the overall upward trend in both data processing volume and associated costs in cloud technologies is expected to persist. At the same time, the presence of short-term fluctuations was identified, which can be explained by the dynamics of the influencing factors and the characteristics of the model. The findings provide an important methodological and practical basis for optimizing data processing processes in cloud environments, ensuring efficient resource utilization, and supporting evidence-based managerial decision-making.

REFERENCES

[1] President of the Republic of Uzbekistan, "On the approval of the 'Digital Uzbekistan – 2030' strategy," Decree No. PF-6079, Oct. 5, 2020.

[2] President of the Republic of Uzbekistan, "On measures for the development of the digital economy," Resolution No. PQ-4699, Apr. 28, 2020.

[3] M. Iansiti and K. R. Lakhani, Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World. Boston, MA, USA: Harvard Business Review Press, 2020.

[4] A. Gawer, "Digital platforms and ecosystems: Remarks on the dominant organizational forms of the digital age," Innovation: Organization & Management, vol. 23, no. 1, pp. 110–124, 2021.

[5] M. N. Tolchinskaya, Digital Transformation of Economy and Business. Moscow, Russia: Yurayt, 2019.

[6] Sh. Shodmonov, Digital Economy: Theory and Practice. Tashkent, Uzbekistan: Iqtisodiyot Publishing House, 2020.

[7] O. Astanaqulov, "Increasing investment attractiveness in the digital economy," Journal of Economics and Innovations, vol. 4, no. 2, pp. 45–52, 2022.

[8] L. Ablazov, R. Khurramov, and F. Xoliyarov, "Current state of cloud computing services: advantages and disadvantages," Digital Economy and Information Technologies, vol. 4, no. 1, pp. 65–71, 2024. [Online]. Available: https://dgeconomy.tsue.uz/index.php/dgeco/article/view/138

[9] M. N. Tolchinskaya, Digital Transformation of Economy and Business. Moscow, Russia: Yurayt, 2019.

[10] G. Kh. Batova, "Analysis of the impact of digital technologies on enterprise efficiency," Economics and Management, no. 5, pp. 78–85, 2020.

[11] Yu. V. Belousov, "Big data and its role in the development of the digital economy," Bulletin of Economic Research, no. 3, pp. 33–40, 2021.

[12] A. Gawer, "Digital platforms and ecosystems: Remarks on the dominant organizational forms of the digital age," Innovation: Organization & Management, vol. 23, no. 1, pp. 110–124, 2021.

[13] M. Iansiti and K. R. Lakhani, Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World. Boston, MA, USA: Harvard Business Review Press, 2020.

[14] OECD, The Digital Transformation of SMEs. Paris, France: OECD Publishing, 2020.

[15] State Committee of the Republic of Uzbekistan on Statistics, "Digital economy statistics." [Online]. Available: https://stat.uz/uz/rasmiy-statistika/raqamli-iqtisodiyot