Article

Tools for Correlation and Regression Analyses in Estimating a Functional Relationship of Digitalization Factors

Svetlana Zemlyak *, Olga Gusarova and Galina Khromenkova

Financial University under the Government of the Russian Federation, Smolensk Branch,
Smolensk 214018, Russia; omgusarova@fa.ru (O.G.); smolensk@fa.ru (G.K.)

* Correspondence: svzemlyak@fa.ru

Abstract: Digitalization processes affect all levels and spheres of human activities, from personal communications to public events. The widespread implementation of digital technologies has an ambiguous effect on the personal, social, and economic paths of modern society's development. At the moment, there is no single approach to the estimation of digitalization's impact, particularly on the financial and economic properties of a company. The objective of this study was to create multiple models for the assessment of digitalization's impact on company performance. To accomplish this objective, the following steps were performed: conducting a literature survey on the experience with digital technologies' implementation; selecting the most appropriate mathematical tools for correlation and regression analyses; determining a functional relationship between the factors and consequences of digitalization in terms of companies' performance; identifying digitalization factors, risks, and their impact on the financial sustainability of companies; and creating a multiple regression model of the functional relationship between digitalization factors and the financial sustainability of companies. In the course of the study, a correlation analysis of the dependence of companies' financial sustainability on a number of digitalization factors has been conducted, and different ways of using company performance data as effective features and predictive factors are offered. The article includes sampling data on the parameterization and quality evaluation of multiple regression models. The validity of the multiple models suggested was tested with actual statistical data obtained from 16 Russian companies. The application of the multiple regression model was devised to estimate digitalization's impact on companies' performance, and their financial sustainability can be seen as the most important practical implication of this study.

Keywords: digital transformations; digitalization factors; multiple regression model; company financial sustainability

Citation: Zemlyak, S.; Gusarova, O.; Khromenkova, G. Tools for Correlation and Regression Analyses in Estimating a Functional Relationship of Digitalization Factors. *Mathematics* **2022**, *10*, 429. <https://doi.org/10.3390/math10030429>

Academic Editor: Manuel Alberto M. Ferreira

Received: 16 November 2021

Accepted: 30 December 2021

Published: 28 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the modern context of digital transformations and digitalization in all spheres of the economy and management, studying the impact of information technology (IT) on company performance has become a popular research avenue. Modern digital technologies make it possible for companies to attain a new level, providing information about services and products virtually to any international/national/regional market. The rapid development of digital financial tools has drastically changed the work of financial organizations, transforming every aspect of their business models. Literature reviews [1–8] have shown that regardless of social and economic development level, most countries consider digitalization as a tool to strengthen their economic position in the world, which contributes to improving their stability and the welfare of their populations.

The implementation of information and communication technologies (ICTs) has a crucial effect on the financial sustainability of companies. The high financial sustainability and performance of a company, first and foremost, reflect the priorities of its internal financial policy and a high resiliency toward environmental concerns.

Modern ICTs facilitate the processing of financial and other documents, which leads to increased labor efficiency in all departments.

The effectiveness of digitalization has been described in a large number of scientific articles. The range of issues discussed is constantly expanding, which testifies to the complexity of the phenomenon. However, a vast majority of investigators have pointed to the effect of business digitalization. Highlighting the advantages of digitalization, reference [9] called it the Holy Grail that helps to create “fundamentally new, disruptive business models”. Reference [10] examined the pros and cons of digitalization in leading European companies and came to the conclusion that digital transformation has a favorable effect on company performance, proven by the high take-up of digital technologies, higher sales, better access to customers, the greater flexibility of price setting, the greater ease of sharing knowledge, and more efficient production processes. The study conducted by [11] in the automotive industry proved the benefits of digital transformations, forecasting higher profits, better productivity, greater competitiveness, better services, and greater satisfaction. Small-to-medium-sized European enterprises were examined by [12], and it was demonstrated that business model innovation driven by social media and big data has a positive impact on business performance. The favorable opportunities of the digital world were emphasized by [13] in the present-day pandemic.

Many studies have been devoted to information and the way in which we structure and use it. Investigating the role of the Internet of Things (IoT) and big data in businesses’ digital transformation, reference [14] pointed out how disorganized knowledge in this sphere leads to inconsistent paths toward developing both research and practice. Having analyzed different aspects of understanding and applying data, reference [15] concluded that digitalization results in “information literacy” challenges and leads to new possibilities for providers of accounting information.

Some individual aspects of digitalization have been addressed in numerous works. The impact of digital transformation in marketing was examined in [16]. Reference [17] considered the crucial importance of choosing appropriate digital tactics and developing business strategies aimed at supporting managers. Reference [18], admitting the key role of digital technologies in maximizing profits, investigated the impact of social media as a marketing tool. Reference [19] investigated the way in which three social partners (governments, business associations, and trade unions) adapt to a new digitized context and their level of digitization. Digital twins (DTs) technology and its benefits in industrial applications were described in [20].

Although some investigators have concentrated on individual digitalization factors, other scientists have come up with different ideas for creating models to estimate and analyze digitalization’s impact. Reference [21] developed “a model for digital innovations-driven business model regeneration” that can be used to analyze digital innovations’ effects and update existing business models. Reference [22] looked at innovations in digital energy service business models and revealed two significant value gaps related to the ability of these models to increase social and environmental benefits.

Some works have provided a detailed overview of the literature devoted to digitalization in business. The study conducted in [23] revealed three thematic clusters: technological innovation, strategic management, and digital transformation. Reference [24] paid attention to an under-researched sphere of digitalization: self-employed workers and freelancers running their business online. Reference [25] examined a poorly investigated question about the impact of digital servitization on sustainability based on the experience of small and medium manufacturing companies in Italy.

Thus, the digital transformation of the economic environment carried out by a large-scale implementation of ICTs implies the transformation of all activities of modern companies to digital platforms. Such a transformation entails changing business-running standards, expanding horizons and creating new types of business (digital business), changing company structure, seizing opportunities that are provided by AI for business transformation and audience accumulation, broadening spheres, and conducting commu-

nication in a web [26–31]. However, there is no single approach that could be applied to estimate the impact of digitalization on company performance and the financial and economic features of companies. It can be claimed that the development of such approaches to estimate company performance in terms of digitalization's impact is seen as an urgent issue. The literature review has identified the two most important digitalization factors that impact company performance. Among these factors, the most common ones are labor productivity and digitalization costs [32–39]; hence, they must be used in creating models for estimating digitalization's impact on company performance.

2. Materials and Methods

The objective of the study is to create multiple models for the assessment of digitalization's impact on company performance.

The study material is a set of financial and economic indicators of 16 Russian companies obtained as of 2019.

The digital transformation of society should be considered using an interdisciplinary scientific approach that integrates methods of economic theory, philosophy, sociology, ICTs, and applied mathematics, and in particular, correlation and regression analyses.

To estimate digitalization's impact on a company's financial sustainability, the following multiple regression model can be used [40]:

$$y(t) = f(x_1, x_2, x_3 \dots x_m) + \varepsilon(t), \quad (1)$$

where $y(t)$ is an effective feature that describes company performance and a company's financial sustainability; f is a functional relationship between the effective feature and externalities; $x_1, x_2, x_3 \dots x_m$ is a set of externalities; and $\varepsilon(t)$ is a stochastic component that shows the impact of factors that are not explicitly included in the regression equation.

When determining the parameters of a multiple linear regression (MLR), an equation can be derived using the least square (LS) method. To do that, a system of standard equations must be solved:

$$\begin{cases} \Sigma y = na + b_1 \Sigma x_1 + b_2 \Sigma x_2 + \dots + b_m \Sigma x_m \\ \Sigma x_1 y = a \Sigma x_1 + b_1 \Sigma x_1^2 + b_2 \Sigma x_1 x_2 + \dots + b_m \Sigma x_1 x_m \\ \dots \\ \Sigma x_m y = a \Sigma x_m + b_1 \Sigma x_1 x_m + b_2 \Sigma x_2 x_m + \dots + b_m \Sigma x_m^2 \end{cases} \quad (2)$$

where y is a dependent variable that denotes any indicator of a company's financial and economic performance; x_i are the externalities that indicate digitalization's impact on a company's performance and sustainability; and a, b_i are MLR equation coefficients that describe changes in the average value of the effective feature, and take place with every unit change in the average value of the factorial feature.

When specifying a model, it is very important to compose a system of indicators for choosing appropriate predictive factors. Moreover, the coefficients of the pair correlation between the endogenous variable and externalities are to be calculated. The following correlation matrix is to be used:

$$R = \begin{pmatrix} r_{yy} & r_{yx_1} & \dots & r_{yx_m} \\ r_{yx_1} & r_{x_1x_1} & \dots & r_{x_1x_m} \\ r_{yx_2} & r_{x_2x_1} & \dots & r_{x_2x_m} \\ \dots & \dots & \dots & \dots \\ r_{yx_m} & r_{x_mx_1} & \dots & r_{x_mx_m} \end{pmatrix} \quad (3)$$

where r_{yxi} is a pair correlation coefficient that shows the impact of a certain predictor on the endogenous variable, and r_{xixj} is a coefficient of the predictor's pair correlation that describes the interdependence of factors.

To identify the predictor's interdependence, matrix R^* must be introduced into the analysis to exclude a multicollinearity:

$$R^* = \begin{pmatrix} r_{x_1x_1} & r_{x_1x_2} & \dots & r_{x_1x_m} \\ r_{x_2x_1} & r_{x_2x_2} & \dots & r_{x_2x_m} \\ \dots & \dots & \dots & \dots \\ r_{x_mx_1} & r_{x_mx_2} & \dots & r_{x_mx_m} \end{pmatrix} \quad (4)$$

The value of a pair correlation coefficient can be calculated by a formula, using a covariance of features:

$$r_{yx} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} = \frac{\overline{yx} - \bar{y} \times \bar{x}}{\sigma_x \sigma_y} \quad (5)$$

where

$$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} \quad (6)$$

is a mean square deviation of a certain factorial feature; and

$$\sigma_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n}} \quad (7)$$

is a mean square deviation of the effective feature,

$$\overline{yx} = \frac{\sum yx}{n}. \quad (8)$$

To determine the value of the coefficient of the pair correlation between the effective feature and the predictor, the following formula can be used:

$$r_{yx} = \sqrt{\frac{\sum (\tilde{y}_x - \bar{y})^2}{\sum (y - \bar{y})^2}} \quad (9)$$

where $\sum (\tilde{y}_x - \bar{y})^2$ is the sum of the squared deviations of the predictor-affected modelled values of the effective feature against the average y value, calculated using empirical data; and $\sum (y - \bar{y})^2$ is the aggregate sum of the squared deviations of y 's empirical values against its average value.

Traditionally, there are two directions in analyzing a matrix of a pair correlation. The first is estimating and analyzing the power of the predictor's impact on the effective feature, and the second is determining and analyzing the interdependence/closeness of the factorial features.

To identify a multicollinearity of predictors that can misrepresent the functional relationship of the variables under study and lead to creating invalid models, the Glauber method can be applied [41].

To check for a multicollinearity in the system of the chosen factorial features, χ^2 criteria can be used. It is calculated by the following formula:

$$\chi^2 = -\left(n - 1 - \frac{2m + 5}{6}\right) \ln(\det|R^*|) \quad (10)$$

the χ^2 criteria are calculated with regards to the degrees of freedom $k = \frac{m(m-1)}{2}$.

At a given reliability (probability) level of p calculation, a designated number of degrees of freedom $k = \frac{m(m-1)}{2}$ can be referred to a table (critical) value of χ_{kp}^2 criteria. If $\chi_p^2 \leq \chi_{kp}^2$, then the multicollinearity between the predictive factors is considered to be zero, and all of the factorial features can be included into the investigation, according to the researcher's judgment.

If $\chi_p^2 \geq \chi_{kp}^2$, then there is a multicollinearity between predictors at a given probability level of p calculation, and the further selection of factorial features to derive the regression equation is subject to additional limitations.

The least square method allows for the determination of the regression model's parameters. In similar cases [42], the identification of multiple regression parameters comes from calculating the minimum of the G function, which delineates the sum of the empirical data's deviation from the effective feature's modeled values. The G function is calculated as follows:

$$G = \sum_{i=1}^n f(y_i - \hat{y}_i) = \sum_{i=1}^n f(e_i) \quad (11)$$

where $f(\)$ is a functional relation that represents a mathematical form of measuring the scatter of empirical data regarding the modeled values. In most cases, the square of the scatter of the effective feature's empirical and modeled values can be used as a mathematical form.

With the regression model's parameters identified, the quality and reliability analyses for the devised regression model are to be performed. To do so, several quality characteristics can be considered.

To estimate the general quality of the regression equation, a coefficient of determination is commonly used. The value of the coefficient can be calculated by the following formula:

$$R^2 = 1 - \frac{\hat{\sigma}^2}{\hat{\sigma}_y^2} = 1 - \frac{ESS/n}{TSS/n} = 1 - \frac{ESS}{TSS} \quad (12)$$

where

$$ESS = \sum_{t=1}^n e_t^2 = \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (13)$$

is the value of residual dispersion that indicates the deviation of the empirical and modeled values of the effective feature; and

$$TSS = \sum_{t=1}^n (y_t - \bar{y})^2 = n\hat{\sigma}_y^2 \quad (14)$$

is the value of the aggregate dispersion that denotes the scatter of the empirical values against the average empirical value of the effective feature, which is measured as follows:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i. \quad (15)$$

The value of the determination coefficient varies from 0 to 1. The closer its value is to 1 (which means a lesser relationship between the residual and aggregate dispersion), the higher the general quality of the created regression equation will be.

Additionally, the quality of the created regression model can be examined using a coefficient of the multiple correlation, which describes the value of the joint impact of the factorial features included in the regression equation on the effective feature. To measure the value of the multiple correlation coefficient, the following formula can be used:

$$R = \sqrt{1 - \frac{ESS}{TSS}} = \sqrt{R^2}. \quad (16)$$

The closer the empirical values of the effective feature are to the regression line, the higher the value of the multiple correlation coefficient will be. Eventually, the reliability of the devised regression equation and its outcome will increase.

The second direction in the quality assessment of the devised regression equation is the estimation of the statistical significance of the regression equation, which can be carried out using the F-test.

There is an assumption that if a sustainable relationship exists between predictors and the effective feature, which is characterized by the devised regression equation, the factor dispersion of the effective feature affected by regression equation predictors must be much greater than the residual dispersion of the effective feature affected by other factors, which are not included in the regression equation.

To determine Fisher's criterion, the ratio of the factor dispersion to residual dispersion of the effective feature is calculated:

$$F = QSS/ESS \quad (17)$$

where

$$QSS = \sum (\hat{y}_x - \bar{y})^2 \quad (18)$$

is a factor dispersion.

Fisher's criterion can also be calculated by the following formula:

$$F = \frac{R^2}{1 - R^2} \times (n - m - 1) \quad (19)$$

where n is the length of the effective feature's time series and m is the number of factorial features included in the regression equation.

Comparing the calculated value of Fisher's criterion to the critical value that is computed based on the given significance level α (0.05 or 0.01) and the degree of freedom $v_1 = m - 1$, $v_2 = n - m$, the following conclusion is made:

If $F_{fact} > F_{table}$, then the devised regression equation is considered statistically significant; otherwise, it will be statistically insignificant.

The assessment of the quality of the devised regression equation can be performed by estimating the statistical significance of the regression coefficients; that, in turn, confirms the statistical significance of the factorial features included in the equation. This estimation can be performed with the help of Student's criterion:

$$t_{b_i} = \frac{b_i}{\mu_{b_i}} \quad (20)$$

where b_i are the regression coefficients of factorial features and μ_{b_i} is the standard deviation of the regression coefficient calculated by the following formula:

$$\mu_{b_i} = \sqrt{\frac{\sum (y - \hat{y}_x)^2 / (n - 2)}{\sum (x_i - \bar{x}_i)^2}}. \quad (21)$$

Since the residual dispersion can be determined by formula (13), the standard deviation of the regression coefficient is calculated through the residual dispersion:

$$\mu_{b_i} = \sqrt{\frac{ESS}{\sum (x_i - \bar{x}_i)^2}}. \quad (22)$$

To calculate the regression parameter a needed for determining Student's criterion, the following formula can be used:

$$t_a = \frac{a}{\mu_a} \quad (23)$$

where a is the absolute term of a regression equation that represents the value of the effective feature when the factorial features included in the regression equation do not affect the effective feature or correspond to zero.

The value of μ_a can be calculated in the following way:

$$\mu_a = \sqrt{ESS \times \frac{\sum x_i^2}{n \times \sum (x_i - \bar{x}_i)^2}}. \quad (24)$$

The computed values of Student's criterion are compared to the critical values of the criterion, and are determined with regards to freedom degrees $n-m$ and a given calculation probability for α (0.01; 0.05). If $t_{calculated} > t_{critical}$, then investigators agree to the statistical significance of the regression coefficient against a certain predictor; otherwise, the regression coefficient will not be statistically significant.

3. Results

After using correlation and regression analysis tools to devise the multiple models for the relationship of digitalization factors, the analysis of the digitalization factors and their impact on the effectiveness and sustainability of a business was carried out (Table 1).

Table 1. Effects of digitalization on company activities in terms of financial sustainability.

Activity	Effects	
	Positive	Negative
Economic	<ul style="list-style-type: none"> – Increased labor productivity of the financial and audit department, as well as the financial activity-controlling department; – Cost cuts caused by e-document management; – Automatization of the creation of projects for strategic development and tactic monitoring of the company's financial activity; 	<ul style="list-style-type: none"> – Layoffs in financial departments; – Necessity in additional staff training; – Expenses of a software purchase; – Cybercrime-related risks;
Production	<ul style="list-style-type: none"> – Automatization of engineering and manufacturing processes; – Automatization of inside and outside business processes; – Creating digital business models; – Creating new kinds of business; – Increased labor productivity in production; – New job formation; 	<ul style="list-style-type: none"> – Advanced professional training for employees working with automated systems; – Potential job losses;
Marketing	<ul style="list-style-type: none"> – Creation of a company's digital information space; – Entering new markets; – Increasing the number of customers; – Rapid distribution of information related to new goods and services, and events; – Creating new communication channels with customers/suppliers; 	<ul style="list-style-type: none"> – Competition toughening; – Increased urgency of personal data protection; – Cybercrime-related risks;
Financial	<ul style="list-style-type: none"> – Implementing and expanding the sphere of providing financial services and tools; 	<ul style="list-style-type: none"> – Threats to economic stability; – Cybercrime-related risks;
Data analytics	<ul style="list-style-type: none"> – Automatization and the facilitated creation of strategic plans and tactic programs related to company development; – Electronic database; – Big data analysis; – Centralization of company information resources; 	<ul style="list-style-type: none"> – Creating info-analytic departments; – Advanced professional training for employees; – Cybercrime-related risks;

Table 1. Cont.

Activity	Effects	
	Positive	Negative
Human resources	<ul style="list-style-type: none"> – Digitalization of hiring/position change operations; – E-document management; – High responsiveness to HR-related inquiries from both internal departments and external regulators; 	<ul style="list-style-type: none"> – Advanced professional training for HR employees; – Increased urgency of personal data protection; – Cybercrime-related risks;
Organizational	<ul style="list-style-type: none"> – Creating and developing new communication links on the web using social media sites and the Internet; – Improved management techniques; – Developing new methods and ways aimed at tackling the organizational tasks of the company; – An opportunity for a rapid reallocation of management tasks; – Reduced paperwork; – High responsiveness to force majeure circumstances in schedule adjustment; – High efficiency of controlling actions in the execution of orders; – Rapidly informing the staff of orders and errands to be executed; 	<ul style="list-style-type: none"> – Threat of information disclosure; – Necessity of updating hardware and software systems; creating supporting subsystems; – Cybercrime-related risks;
Logistic activities	<ul style="list-style-type: none"> – Automatization of supply/delivery processes; – Electronic database for suppliers/consumers; – Facilitated exchange of supply/delivery documents; – Immediate receiving of information. 	<ul style="list-style-type: none"> – Competition toughening; – Increased urgency of personal data protection; – Cybercrime-related risks.

Table 1 contains the effects of digital technologies on a company's financial sustainability.

As a result of the study conducted, the multiple regression model for the estimation of digitalization's impact on the financial sustainability of companies has been created:

$$Y(FU)_t = f((IRPT)_t, (ISZ)_t, (IASR)_t, (IABP)_t, (HBI)_t, (RKB)_t, (ZPK)_t, (HDI)_t), \quad (25)$$

where $Y(FU)_t$ is the effective feature that characterizes the effectiveness and sustainability of business activity, and f is a functional relationship between the effective feature and a set of factorial features.

The following parameters can be used as the effective and factorial features based on the investigator's judgment (see Table 2):

Derived values that describe the predictors' dynamic patterns with time, i.e., their indexes, can be used as predictors rather than absolute values of exogeneous features. An autoregressive model of the m -dimensional vector of exogeneous variables can be used to create a system of predictor values and to calculate the index of factorial features:

$$X_{it} = \alpha_i + \beta_{1i} X_{it-1} + \beta_{2i} X_{it-2} + \dots + \beta_{pi} X_{it-p} + \varepsilon_{it}, \quad (26)$$

where α_i and β_i are the parameters of the multiple regression model used to compute dynamics indexes for a number of exogeneous variables;

t is the index that shows the factorial features changing over time;

p is the value of the lag of the interrelated observations of the predictors' time series. It characterizes the structure of the autoregressive model and the extent of the relationships between the observations of the exogeneous variables' time series;

ε_{it} is called white noise errors. The variable describes the errors in measuring the numerical value of regression model parameters, as well as the violation of some prerequisites for

the LS method. Moreover, it shows the impact of slack (qualitative) variables on the predictor value.

Table 2. Parameters to be chosen for the regression model.

The Effective Feature	
Y(FU) _t	A set of financial sustainability indicators
	Cost effectiveness
	Liquidity
	Creditworthiness
	Amount and structure of equity
	Other parameters according to the investigator's judgment
Predictors	
IRPT _t	Index of labor efficiency increase
ISZ _t	Index of cost cutting related to implementing digital technologies into business activities
IASR _t	Index of the automatization of strategic planning and day-to-day management
IABP _t	Index of business process automatization
HBI _t	Expenses on ICTs (purchase, implementation, maintenance, automated information system security)
RKB _t	Cybercrime-related risks (a qualitative factor that is hard to measure)
ZPK _t	Expenses on the advanced professional training of employees working with ICTs
HDI _t	A number of qualitative parameters (competition toughening, entering new markets, increased number of customers, new level of organizational management)

To validate the created multiple models aimed at estimating digitalization factors' impact on a company's performance and sustainability, data from 16 companies of the Smolensk region were obtained.

Only companies from the Smolensk region were studied, since an expert survey in other regions involves significant expense.

The choice of the study objects (companies) was justified by the availability of the benchmark statistics (original statistical data) on the financial activity of the companies, as well as by the desire of these companies to take part in the expert survey. Some companies, due to the worsening of their performance indicators, refused to provide their statistics, referring to a commercial confidentiality policy to explain this refusal. Moreover, the pandemic has had a negative impact on company performance in most cases, and a number of firms have ceased to exist.

The following parameters of financial sustainability were taken into consideration: the coefficient of financial independence (KFN); the coefficient of the current liquidity (KTL); the coefficient of the absolute liquidity (KAL); and the coefficient of a short-term debt (KKZ). Table 3 contains the sampling data for a number of companies under study.

Table 3. Indicators of the efficiency and sustainability of companies in the Smolensk region as of 2019.

Company	KFN	KTL	KAL	KKZ
AO "NPP "Izmeritel""	0.37395	12.4765	11.1247	0.06540
FGUP "SPO "Analitpribor""	0.86739	7.05104	4.07013	0.10421
AO "Smolenskiy aviatzionny zavod"	0.47420	2.40083	0.75765	0.30016
AO "NII sovremennykh telekommunikatsionnykh tekhnologiy"	0.70132	5.40062	0.56761	0.15434

For validating the multiple models created for estimating the impact of digitalization factors on a company's financial sustainability, the following factorial features were chosen:

X1—Index of labor efficiency increase (IRPT);

X2—Expenses on ICTs (HBI 1);

X3—Expenses on implementing/installing/maintaining new software and ICTs (HBI 2);

X4—Expenses on the advanced professional training of employees working with ICTs (ZPK).

As previously mentioned, the choice of these factorial features is justified by the availability of the benchmark statistics (original statistical data) on the financial activity of the companies, as well as by the desire of these companies to take part in the expert survey.

For some factorial features included in the regression model for estimating digitalization's impact on a company's financial sustainability, it was impossible to obtain statistical data that were of a qualitative nature. These include data such as the index of cost cutting related to implementing digital technologies into business activities (ISZ), the index of the automatization of strategic planning and day-to-day management (IASR), the index of business process automatization (IABP), cybercrime-related risks (RKB), and a number of qualitative parameters (HDI). Statistical data on these factorial features within this study are difficult to obtain, as special (experimental) observation and day-to-day monitoring of these factors are required for every company involved in the study.

Table 4 presents the sampling data on the factorial features that have numeric expression and that were obtained from the expert interviews and the accounting documents of the companies. The numbers are shown in thousands of rubles.

Table 4. Data on the digitalization-related costs of companies in the Smolensk region as of 2019.

Company	X1 IRPT, %	X2 HBI1	X3 HBI2	X4 ZPK
AO "NPP "Izmeritel""	102.5	11,153	9513	3608
FGUP "SPO "Analitpribor""	105.0	10,159	5605	2977.7
AO "Smolenskiy aviatzionny zavod"	106.3	2335	1001	333.6
AO "NII sovremennykh telekommunikatziionnykh tekhnologiy"	104.5	1932	1118	254

To identify digitalization's impact on the financial sustainability of the above-mentioned companies, the tools of correlation and regression analyses, as well as statistical methods of modelling functional relationships for company performance indicators, were used.

The correlation dependence of the financial sustainability indicators of some companies and digitalization factors is shown in the matrix of the pair correlation of the effective and factorial features. Table 5 furnishes the sampling data on the correlation analysis.

Table 5. Matrix of the pair correlation of financial sustainability indicators and digitalization factors.

	KFN	KTL	KAL	KKZ	X1/IRPT	X2/HBI1	X3/HBI2	X4/ZPK
KFN	1							
KTL	0.2167	1						
KAL	0.1496	0.5530	1					
KKZ	−0.8458	−0.4730	−0.2988	1				
X1/IRPT	0.3441	−0.0311	−0.2756	−0.4529	1			
X2/HBI1	0.1315	0.2182	0.5345	−0.1455	0.1956	1		
X3/HBI2	0.0974	0.2451	0.6087	−0.1487	0.1031	0.9718	1	
X4/ZPK	0.1228	0.2257	0.5723	−0.1442	0.1235	0.9917	0.9832	1

4. Discussion

The results of the study, conducted with the use of the tools of correlation and regression analyses, reveal the impact of digitalization factors on the financial sustainability of companies; however, the intensity of the intercorrelation of the examined features varies from moderate to mild. This can be explained by the fact that digitalization factors do not play a pivotal role in improving a company's performance and financial sustainability. There are other factors, which are not related to digitalization, that have a greater effect on company performance. Nevertheless, in the context of digital globalization, companies cannot develop effectively unless they correspond to modern digital requirements in their activities [43–48].

The study results demonstrate that the coefficient of financial independence (KFN) is directly correlated with factorial features that depict the labor efficiency increase ($X1/IRPT$), expenses on software and ICTs ($X2/HBI1$, $X3/HBI2$), and expenses on the advanced professional training of employees ($X4/ZPK$). Thus, digitalization has a positive effect on the coefficient of the company's financial independence.

The data obtained point to the fact that digitalization factors affect financial sustainability, but their influence is ambiguous. For example, the index of labor efficiency increase ($X1/IRPT$) has a direct correlation with indicators of a company's financial sustainability, such as the coefficient of financial independence (KFN). This is due to the fact that an increase in labor efficiency leads to increased output, which positively affects the specified indicators of a company's financial sustainability.

However, the index of labor efficiency increase ($X1/IRPT$) has an inverse correlation with the indicators of a company's financial sustainability, such as the coefficients of the current and absolute liquidity (KTL and KAL), and the coefficient of a short-term debt (KKZ). These findings tell us that increases in labor efficiency lead to increased output and increased sales, which reduce receivables and payables, and decrease the coefficient of a short-term debt.

Digitalization costs ($X2/HBI1$, $X3/HBI2$, $X4/ZPK$) have a mild inverse correlation with some financial sustainability indicators, such as the coefficient of a short-term debt (KKZ), since expenses on the advanced professional training of employees decrease company profits and have a negative impact on the values of the specified parameters.

In most cases, the impact of digitalization factors ($X2/HBI1$, $X3/HBI2$, $X4/ZPK$) has a direct correlation with the indicators of a company's financial sustainability.

It is necessary to point out the strong correlation between predictors $X2/HBI1$, $X3/HBI2$, and $X4/ZPK$, which is quite obvious, since a significant part of the expenses on ICTs ($X2/HBI1$) causes considerable expenses in the implementation of digital technologies ($X3/HBI2$) and expenses on the advanced professional training of employees working with ICTs ($X4/ZPK$). According to the correlation and regression analysis, to ensure the high validity of the created models, it is possible to reject one of the highly correlated factors ($X2/HBI1$, $X3/HBI2$, $X4/ZPK$), or to unite these factors into one ($X2/HBI$) that embraces the influence of the other related factors and is equal to the sum of these factors.

The strength of correlation, i.e., the degree of the impact of digitalization factors on a company's financial sustainability, can be considered mild or moderate based on the study's findings. It can be concluded that digitalization factors are not determinants of the level of financial sustainability.

In addition to the correlation between digitalization factors and indicators of a company's financial sustainability, multiple regression models for the dependence of each indicator on digitalization factors have been built. Table 6 contains sampling data on the calculation of the multiple regression models for the dependence of a company's financial sustainability indicators on the above-mentioned digitalization factors.

Table 6. Results of creating multiple regression models for the dependence of a company's financial sustainability indicators.

Financial Sustainability Indicators	Equation of Multiple Regression Models
A coefficient of financial independence (KFN)	$\text{KFN} = -8.47598 + 0.08623 \times 1 - 1.91610 \times 10^{-4} X_2 - 1.09783 \times 10^{-4} X_3 + 8.92693 \times 10^{-4} X_4$ (27)
A coefficient of the current liquidity (KTL)	$\text{KTL} = 18.71124 - 0.12884X_1 + 3.18984 \times 10^{-8} X_2 + 1.636421 \times 10^{-3} X_3 - 2.47296 \times 10^{-3} X_4$ (28)
A coefficient of the absolute liquidity (KAL)	$\text{KAL} = 50.04526 - 0.47469X_1 - 4.94790 \times 10^{-4} X_2 + 1.70116 \times 10^{-3} X_3 - 3.11825 \times 10^{-4} X_4$ (29)
A coefficient of a short-term debt (KKZ)	$\text{KKZ} = 15.18814 - 0.14480X_1 + 3.96299 \times 10^{-4} X_2 - 6.14101 \times 10^{-5} X_3 - 1.14327 \times 10^{-4} X_4$ (30)

When determining the values of the indicators used in the models, the companies' financial statistics have been taken for a certain period, which is stipulated in the assignment (the study was conducted under government orders).

An insignificant value of the regression coefficients of the multiple models is justified by the fact that the majority of the financial sustainability indicators range from 0 to 1. The factorial features that characterize digitalization's impact on the company's financial sustainability are shown in relatively greater numbers. Moreover, as stated above, digitalization's impact on a company's financial sustainability is not pivotal in the numerical estimation of a company's financial sustainability indicators.

In the study, the quality of the devised regression equation was estimated to determine the reliability of further study results. Table 7 furnishes the results of the regression analysis of the functional dependence of the financial independence coefficient (27) on the examined factorial features.

Table 7. The major parameters of the regression statistics for the multiple models of the functional relationship of digitalization factors and the coefficient of the company's financial independence.

Model	R	R ²	R ² _A	μ	d
1	0.707	0.499	0.317	3.101	1.841

The values of the following parameters of the regression statistics are depicted in Table 7: multiple correlation coefficient (R); determination coefficient (R²); determination coefficient adjusted for degrees of freedom (R²_A); standard error (μ); and Durbin-Watson statistics (d). All of these parameters describe the quality of the multiple models for the estimation of digitalization factors' impact on the coefficient of a company's financial independence.

The value of the multiple correlation coefficient (R) equaling to 0.707 shows that there is a positive correlation between the performance indicator and the predictors under study. The low value of the determination coefficient (R²) equaling to 0.499 speaks to the insufficient quality of the created regression equation. The low value of the adjusted determination coefficient implies that substantial factors, which are more influential in terms of the effective feature, are not included in the regression equation.

Taking into consideration the high correlation between the factors X2/HBI1, X3/HBI2, and X4/ZPK, regression models with two factors have been created as well. These factors are X1/IRPT (index of the labor efficiency increase) and X2/HBI (expenses on ICT purchasing, implementing, and staff training). Factor X2/HBI is equal to the sum of the values of X2/HBI1, X3/HBI2, and X4/ZPK.

Sampling data on the results of creating a two-factor regression model of the dependency of a company's financial sustainability indicators on digitalization factors can be seen in Table 8.

Table 8. Results of creating two-factor regression models of the dependency of a company's financial sustainability indicators.

Financial Sustainability Indicators	Equation of Two-Factor Regression Models
A coefficient of financial independence (KFN)	$KFN = -6.56115 + 0.06714X_1 - 4.46438 \times 10^{-6} X_2$ (31)
A coefficient of the current liquidity (KTL)	$KTL = 24.33957 - 0.18395X_1 + 2.13746 \times 10^{-4} X_2$ (32)
A coefficient of the absolute liquidity (KAL)	$KAL = 62.33305 - 0.59600X_1 + 3.2940 \times 10^{-4} X_2$ (33)
A coefficient of a short-term debt (KKZ)	$KKZ = 10.59673 - 0.09921X_1 - 5.9165 \times 10^{-6} X_2$ (34)

The analysis of the quality of the two-factor regression models was carried out, and the sampling data on the absolute liquidity ratio can be seen in Table 9.

Table 9. The basic parameters of the regression statistics for the two-factor model of the functional relationship between digitalization factors and the absolute liquidity ratio.

Model	R	R ²	R ² _A	μ	d
2	0.678	0.460	0.263	2.962	1.748

The calculations show that the regression statistics' parameters of the two-factor regression do not differ much from the multiple models. On the whole, it can be unambiguously claimed that there is a correlation between the chosen predictors that reflect the digitalization processes. On the one hand, there is a correlation between a company's financial sustainability and performance; however, on the other hand, these factors are not determinants in the estimation of a company's financial sustainability.

As a result of this study, the following contributions to research into financial sustainability are made:

- The real-world experience of digitalization, in terms of its impact on company performance, has been examined;
- Mathematical tools of correlation and regression analyses have been investigated to identify the functional relationship of digitalization indicators;
- Digitalization factors that have an effect on a company's performance and financial sustainability have been systematized;
- The nature of the manifestations and consequences of digitalization's impact on a company's financial sustainability have been described;
- The multiple regression model for the estimation of digitalization's impact on financial sustainability has been created;
- The multiple models that were created were validated with statistics on some Russian companies;
- The matrix of the pair correlations of the system of the effective and factorial features that reflect digitalization's impact on financial sustainability has been built and described;
- Multiple regression models for the dependence of financial sustainability indicators on digitization factors have been created;
- The results of the correlation and regression analyses of the multiple models created have been described.

5. Conclusions

The results of this research contribute to the studying of digitalization's impact on company performance and financial sustainability. The selected digitalization factors can be referred to both on a macro and micro level when analyzing the experience of different countries. The tools of correlation and regression analysis have allowed us to identify a certain correlation between a company's financial sustainability indicators and digitalization factors. With the multiple models created, it has become possible to determine the invariable regression models of the dependence of different exogenous and endogenous factors.

This research has a number of restrictions and conditions.

Firstly, the sampling included regional medium-sized companies. With expanded sampling that takes into account large-sized companies, the results could be different.

Secondly, factors such as the index of labor efficiency increase and the expenses of ICT purchasing, implementation, and staff training have been taken as predictors that characterize digitalization's impact on a company's performance and financial sustainability. With expanded fields of research and additional statistical data, other indicators of digitalization's impact can be included.

Thirdly, an extended study time period, bigger sampling, and additional information on digitalization's effects at the macro and micro levels could provide a background for data validation and further research.

Despite the restrictions and conditions, the practical implications of the research assume the application of the devised approach to create multiple models for assessing digitalization's impact on a company's financial sustainability.

Author Contributions: Conceptualization, S.Z.; methodology, O.G.; software, O.G.; validation, S.Z., O.G. and G.K.; formal analysis, O.G.; investigation, S.Z.; resources, G.K.; data curation, O.G.; writing—original draft preparation, O.G.; writing—review and editing, S.Z.; visualization, G.K.; supervision, S.Z.; project administration, S.Z.; funding acquisition, S.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Government of the Russian Federation, grant number 44, and the APC was funded by the Government of the Russian Federation.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: <https://cloud.mail.ru/public/EsE4%2FFi5qbfKuA>.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Juliy, L.V.; Emchoo, L.V. Information systems and their role in the activities of modern companies. *Econ. Financ.* **2015**, *6*, 130–134.
2. Gebauer, H.; Arzt, A.; Kohtamäki, M.; Lamprecht, C.; Parida, V.; Witell, L.; Wortmann, F. How to convert digital offerings into revenue enhancement—Conceptualizing business model dynamics through explorative case studies. *Ind. Mark. Manag.* **2020**, *91*, 429–441. [\[CrossRef\]](#)
3. Ritter, T.; Pedersen, C.L. Digitization capability and the digitalization of business models in business-to-business firms: Past, present, and future. *Ind. Mark. Manag.* **2020**, *86*, 180–190. [\[CrossRef\]](#)
4. Bresciani, S.; Huarng, K.-H.; Malhotra, A.; Ferraris, A. Digital transformation as a springboard for product, process and business model innovation. *J. Bus. Res.* **2021**, *128*, 204–210. [\[CrossRef\]](#)
5. Gartner Hype Cycle for Digital Marketing 2020. Available online: <https://vc.ru/marketing/171411-gartner-hype-cycle-for-digital-marketing-2020-k-kakim-napravleniyam-i-tehnologiyam-stoit-prismotretsya-marketologam> (accessed on 11 November 2021).
6. Härting, R.-C.; Reichstein, C.; Schad, M. Potentials of Digital Business Models—Empirical investigation of data driven impacts in industry. *Procedia Comput. Sci.* **2018**, *126*, 1495–1506. [\[CrossRef\]](#)
7. Trinugrohoab, I.; Pamungkasab, P.; Wiwoho, J.; Damayantid, S.M.; Pramono, T. Adoption of digital technologies for micro and small business in Indonesia. *Financ. Res. Lett.* **2021**, *5*, 102–156.
8. Wood, S.; Faulconbridge, J.; Watson, I.; Teller, C. 'Business Code/Spaces' in digital service firms: The case of online multinational fashion retailing. *Geoforum* **2020**, *112*, 13–23. [\[CrossRef\]](#)

9. Hinterhuber, A.; Nilles, M. DIGITAL transformation, the HOLY GRAIL and the disruption of business models. *Bus. Horiz.* **2021**, *2*, 57–63.
10. Elding, C.; Morris, R. Digitalisation and its impact on the economy: Insights from a survey of large companies. *ECB Econ. Bull.* **2018**, *7*, 12.
11. Llopis-Albert, C.; Rubio, F.; Valero, F. Impact of digital transformation on the automotive industry. *Technol. Forecast. Soc. Chang.* **2021**, *162*, 120–134. [\[CrossRef\]](#)
12. Bouwman, H.; Nikou, S.; Reuver, M. Digitalization, business models, and SMEs: How do business model innovation practices improve performance of digitalizing SMEs? *Telecommun. Policy* **2019**, *43*, 101–118. [\[CrossRef\]](#)
13. Viswanathan, R.; Telukdarie, A. A systems dynamics approach to SME digitalization. *Procedia Comput. Sci.* **2021**, *180*, 816–824. [\[CrossRef\]](#)
14. Sestinoa, A.; Preteb, M.I.; Piperb, L.; Guido, G. Internet of Things and Big Data as enablers for business digitalization strategies. *Technovation* **2020**, *98*, 102–113. [\[CrossRef\]](#)
15. Bhimani, A.; Willcocks, L. Big Data and the transformation of accounting information. *Account. Bus. Res.* **2014**, *44*, 469–490. [\[CrossRef\]](#)
16. Melović, B.; Jocović, M.; Dabić, M.; Backović, T.; Dudic, B. The impact of digital transformation and digital marketing on the brand promotion, positioning and electronic business in Montenegro. *Technol. Soc.* **2020**, *63*, 101–114. [\[CrossRef\]](#)
17. Olson Eric, M.; Olson Kai, M.; Czaplewski, A.J.; Key, T. Business strategy and the management of digital marketing. *Bus. Horiz.* **2021**, *64*, 285–293. [\[CrossRef\]](#)
18. Dolega, L.; Rowe, F.; Branagan, E. Going digital? The impact of social media marketing on retail website traffic, orders and sales. *J. Retail. Consum. Serv.* **2021**, *60*, 142–154. [\[CrossRef\]](#)
19. Mas, J.M.; Andrés Gómez, A. Social partners in the digital ecosystem: Will business organizations, trade unions and government organizations survive the digital revolution? *Technol. Forecast. Soc. Chang.* **2021**, *162*, 120–134. [\[CrossRef\]](#)
20. Glatta, M.; Kölscha, P.; Siedlera, C.; Langlotza, P.; Ehmsena, S.; Auricha Jan, C. Edge-based Digital Twin to trace and ensure sustainability in crosscompany production networks. *Procedia* **2021**, *7*, 276–281.
21. Coskun-Setirek, A.; Tanrikulu, Z. Digital innovations-driven business model regeneration: A process model. *Technol. Soc.* **2021**, *64*, 161–174. [\[CrossRef\]](#)
22. Hiteva, R.; Foxon, T.J. Beware the value gap: Creating value for users and for the system through innovation in digital energy services business models. *Technol. Forecast. Soc. Chang.* **2021**, *166*, 120–145. [\[CrossRef\]](#)
23. Caputo, A.; Pizzi, S.; Pellegrini, M.; Dabić, M. Digitalization and business models: Where are we going? A science map of the field. *J. Bus. Res.* **2021**, *123*, 489–501. [\[CrossRef\]](#)
24. Reuschke, D.; Mason, C.; Syrett, S. Digital futures of small businesses and entrepreneurial opportunity. *Futures* **2021**, *128*, 362–374. [\[CrossRef\]](#)
25. Paiola, M.; Gebauer, H. Internet of things technologies, digital servitization and business model innovation in BtoB manufacturing firms. *Ind. Mark. Manag.* **2020**, *89*, 245–264. [\[CrossRef\]](#)
26. Maier, C.D.; Engberg, J. Harvard Business Review's reframing of digital communication: From professional expertise to practical guidance. *J. Pragmat.* **2021**, *176*, 186–197. [\[CrossRef\]](#)
27. Trevisan, A.H.; Zacharias, I.S.; Castro, C.G.; Mascarenhas, J. Circular economy actions in business ecosystems driven by digital technologies. *Procedia* **2021**, *100*, 325–330. [\[CrossRef\]](#)
28. Looy, A.V. A quantitative and qualitative study of the link between business process management and digital innovation. *Inf. Manag.* **2021**, *58*, 431–445.
29. Paiola, M.; Schiavone, F.; Khvatova, T.; Grandinetti, R. Prior knowledge, industry 4.0 and digital servitization. An inductive framework. *Technol. Forecast. Soc. Chang.* **2021**, *171*, 120–143. [\[CrossRef\]](#)
30. Montealegre, R.; Iyengar, K. Managing digital business platforms: A continued exercise in balancing renewal and refinement. *Bus. Horiz.* **2021**, *64*, 51–59. [\[CrossRef\]](#)
31. Paiola, M.; Schiavone, F.; Grandinetti, R.; Chen, J. Digital servitization and sustainability through networking: Some evidences from IoT-based business models. *J. Bus. Res.* **2021**, *132*, 507–516. [\[CrossRef\]](#)
32. Metlyakhin, A.I.; Nikitina, N.A.; Yarygina, L.V.; Orlova, E.O. Analysis of the Impact of Economy Digitalization on Labor Productivity in Russia. Digital Economy: Theory and Practice. Available online: <https://cyberleninka.ru/article/n/analysis-of-the-impact-of-economy-digitalization-on-labor-productivity-in-russia> (accessed on 12 November 2021).
33. Boikova, T.; Zeverte-Rivza, S.; Rivza, P.; Rivza, B. The Determinants and Effects of Competitiveness: The Role of Digitalization in the European Economies. *Sustainability* **2021**, *13*, 11689. [\[CrossRef\]](#)
34. Pasqualino, R.; Demartini, M.; Bagheri, F. Digital Transformation and Sustainable Oriented Innovation: A System Transition Model for Socio-Economic Scenario Analysis. *Sustainability* **2021**, *13*, 11564. [\[CrossRef\]](#)
35. Popov, N.A. Business process optimization in the digitalization era of production. *Strateg. Decis. Risk Manag.* **2019**, *10*, 28–35. [\[CrossRef\]](#)
36. The Effects of Digital-Technology Adoption on Productivity and Factor Demand: Firm-Level Evidence from Developing Countries. Available online: <https://openknowledge.worldbank.org/handle/10986/34251> (accessed on 12 November 2021).
37. Competing in the Digital Age: Policy Implications for the Russian Federation. Available online: <https://openknowledge.worldbank.org/handle/10986/30584> (accessed on 11 November 2021).

38. Digital Health Assessment Toolkit Guide. Available online: <https://openknowledge.worldbank.org/handle/10986/36547> (accessed on 13 November 2021).
39. The Impact of Digital Technologies on Routine Tasks: Do Labor Policies Matter? Available online: <https://openknowledge.worldbank.org/handle/10986/28364> (accessed on 14 November 2021).
40. Greene, W. *Econometric Analysis*; Pearson Education: Harlow, UK, 2012; Volume 34, pp. 41–47.
41. Wooldridge, J. *Introductory Econometrics: A Modern Approach*; South-Western, Cengage Learning: Belmont, CA, USA, 2013.
42. Verbeek, M. *A Guide to Modern Econometrics*, 2nd ed.; John Wiley & Sons, Ltd.: Chichester, UK, 2004; p. 446.
43. Gregori, P.; Holzmann, P. Digital sustainable entrepreneurship: A business model perspective on embedding digital technologies for social and environmental value creation. *J. Clean. Prod.* **2020**, *272*, 122–138. [[CrossRef](#)]
44. Jafari-Sadeghi, V.; Garcia-Perez, A.; Candelo, E.; Couturier, J. Exploring the impact of digital transformation on technology entrepreneurship and technological market expansion: The role of technology readiness, exploration and exploitation. *J. Bus. Res.* **2021**, *124*, 100–111. [[CrossRef](#)]
45. Chinoracky, R.; Kurotova, J.; Janoskova, P. Measuring the impact of digital technologies on transport industry—Macroeconomic perspective. *Transp. Res. Procedia* **2021**, *55*, 434–441. [[CrossRef](#)]
46. Kohtamäki, M.; Parida, V.; Oghazi, P.; Gebauer, H.; Baines, T. Digital servitization business models in ecosystems: A theory of the firm. *J. Bus. Res.* **2019**, *104*, 380–392. [[CrossRef](#)]
47. Hanafizadeh, P.; Kim, S. Digital Business: A new forum for discussion and debate on digital business model and digital transformation. *Digit. Bus.* **2020**, *1*, 572–586. [[CrossRef](#)]
48. Ardito, L.; Raby, S.; Albino, V.; Bertoldi, B. The duality of digital and environmental orientations in the context of SMEs: Implications for innovation performance. *J. Bus. Res.* **2021**, *23*, 44–56. [[CrossRef](#)]