




Article

Application of Logistic Regression to Analyze The Economic Efficiency of Vehicle Operation in Terms of the Financial Security of Enterprises

Malgorzata Grzelak ^{1,*}, Paulina Owczarek ¹, Ramona-Monica Stoica ², Daniela Voicu ² and Radu Vilău ²¹ Faculty of Security, Logistics and Management, Military University of Technology, 00-908 Warsaw, Poland; paulina.owczarek@wat.edu.pl² Faculty of Aircraft and Military Vehicles, Military Technical Academy “Ferdinand I”, 050141 Bucharest, Romania; ramona.stoica@mta.ro (R.-M.S.); daniela.voicu@mta.ro (D.V.); radu.vilau@mta.ro (R.V.)

* Correspondence: malgorzata.grzelak@wat.edu.pl

Abstract: *Background:* A measurable feature of the efficiency of vehicle use in transportation companies is the revenue from transport orders, which has a significant impact on their profitability. Therefore, it is important to skillfully analyze the parameters related to the operation of vehicles and their impact on the bottom line. Transportation companies, when managing their operations, take steps to reduce operating costs. The above makes a large number of studies available in the literature on the analysis of vehicle damage or wear of system components, as well as ways to predict them. However, there is a lack of studies treating the impact of the parameters of specific orders on economic efficiency, which is a research niche undertaken in the following study. *Methods:* The purpose of this article was to analyze the economic efficiency of vehicle operation in terms of the financial security of enterprises. The main research problem was formulated in the form of the question of how the various parameters of a transport order affect its profitability. During our study, critical analysis of the literature, mathematical modeling and inference were used. A detailed analysis of transport orders executed by SMEs (small and medium-sized enterprises), which are characterized by a fleet of light commercial vehicles with a capacity of up to 3.5 t, was carried out in the FMCG (Fast-Moving Consumer Good) industry in Poland in 2021–2022. Due to the binary variable form, a logistic regression model was elaborated. The estimated parameters of the model and the calculated odds ratios made it possible to assess the influence of the selected factors on the profitability of orders. *Results:* Among other things, it was shown that in the case of daily vehicle mileage, the odds quotient indicates that with each additional kilometer driven, the probability of profitability of an order increases by 1%. Taking into account the speed of travel, it is estimated that with an increase in its value by 1 km/h, the probability of profitability of an order decreases by 3%. On the other hand, an increase in cargo weight by 1 kg makes the probability of a profitable order increase by 9%. *Conclusion:* Through this study, the limited availability of low-cost analytical tools that can be applied during transportation fleet management in SME companies was confirmed, as was the use of simple and non-expansive mathematical models. At the same time, they are not “black boxes” and therefore enable drawing and implementing model conclusions into operations. The results obtained can help shape the overall strategy of companies in the area of vehicle operation and can support the decision-making process related to the management of subsequent orders, indicating those that will bring the highest profit. The above is very important for SME companies, which often operate on the verge of profitability.

Keywords: logistic regression; economic efficiency; delivery vehicles; profitability

Citation: Grzelak, M.; Owczarek, P.; Stoica, R.-M.; Voicu, D.; Vilău, R. Application of Logistic Regression to Analyze The Economic Efficiency of Vehicle Operation in Terms of the Financial Security of Enterprises. *Logistics* **2024**, *8*, 46. <https://doi.org/10.3390/logistics8020046>

Academic Editor: Robert Handfield

Received: 7 March 2024

Revised: 25 April 2024

Accepted: 29 April 2024

Published: 1 May 2024



Copyright: © 2024 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

Dynamic changes in the business market and the constant increase in customer expectations force transportation enterprises to constantly seek methods and tools to optimize processes and improve the efficiency of the entire transport system, which, operating in a changing environment, is susceptible to negative effects [1]. The consequences may include the disruption of transportation tasks or a complete inability to undertake transportation activities.

The diverse set of qualitative and quantitative data describing the execution of a single transport order and the constantly changing conditions of vehicle use mean that taking rational and planned actions in transportation enterprises in terms of intensity of use, maintenance activities, control or monitoring is often driven by intuitive and subjective feelings, which does not strengthen the position of the enterprise's vehicle operation strategy. The large number of variables describing the process of vehicle use, especially in micro- and small enterprises, means that, in practice, the coefficients characterizing the operation of vehicles and their impact on the profitability of the tasks carried out are not analyzed [2]. Lack of awareness of the cause-and-effect relationship of the process of using and operating vehicles results in users often making decisions that are inappropriate and irrational from the point of view of the business outcome. This significantly affects the quality of the vehicle operating system, generates costs associated with its use and operation and may cause the risk of incurring financial losses when the cost of executing the order exceeds its value.

SME enterprises are characterized by the small number of vehicles they own as part of their fleet. Due to their limited budget, which is mainly due to the fact that they carry out less profitable transport orders, they are unable to implement advanced transportation management systems (e.g., TMS class) to optimize their operations. This causes them to rely solely on their experience when deciding whether to undertake an order and its pricing. The above causes them to often execute orders at a price below the break-even point. This causes the economic efficiency of such entities to decrease significantly. Therefore, it seems important to build simple and accessible tools to support decision-makers in the decision-making process for the execution of transport orders. The above formed the basis for further analysis [3].

The purpose of this article was to analyze the economic efficiency of vehicle operation in terms of enterprises' financial security. For the purpose of the study conducted, the research problem was formulated as follows: How do the various parameters of a transport order affect its profitability? The following research methods were used: critical analysis of the literature, mathematical modeling using a logistic regression model and inference (synthesis and deduction). In this article, an analysis was made of the execution of transport orders in an SME enterprise (which is characterized by a fleet of light commercial vehicles with a capacity of up to 3.5 t) providing services in Poland in the FMCG industry. Logistic regression was used in this article. The accumulated observations and preliminary analyses carried out made it possible to isolate the main factors affecting the profitability of a transport order and enabled the selection of the model's predictors. These included qualitative variables, characterizing lead time, and quantitative variables, describing operating parameters, i.e., daily mileage, driving speed or fuel consumption. The parameters estimated by the regression model made it possible to analyze and evaluate the relationships that occurred. The quality of the proposed model was verified by analyzing the significance of the parameters and testing the goodness of fit using the Receiver Operating Characteristic (ROC) Curve and the Area Under ROC (AUC) Curve.

2. Literature Analysis

Efficiency is a concept widely described, although it is often not clearly defined in the literature. With regard to the profitability of implemented activities, the literature most often uses the concept of economic efficiency. Efficiency, not as a technical but as an economic category, determines the outcome of the evaluation of the enterprise's activities

as well as its individual areas, i.e., the process of vehicle use, service and diagnosis. One of the most popular definitions of economic efficiency indicates that it is an activity that is waste-free and aimed at achieving the best result within the available resources and technology [4,5]. In general terms, it is defined as the ability to use one's resources in such a way as to achieve a given goal in the most efficient and least wasteful manner. The essence of measuring such efficiency is to compare the effect achieved with the inputs consumed [6]. Samuelson and Nordhaus defined efficiency as the most effective application of a society's available resources in the process of satisfying needs [7]. According to organizational theory, efficiency is considered the broadest in meaning among such concepts as performance, productivity, economy, rationality or effectiveness [8]. Efficiency can apply to both individual entities and the economy as a whole [9]. Moreover, it can be considered in the short and long term [10]. An enterprise is efficient if, making full use of its resources, it produces goods that are in demand in the market without unnecessary stockpiling [11]. In an attempt to capture the essence of efficiency, Pyszka points out the ambiguous interpretation of this concept, which is undergoing constant modification with the development of management theory [12].

When analyzing considerations of efficiency, it is necessary to adopt such a concept of it that most aptly characterizes the work of delivery vehicles [13]. Therefore, the following section of this paper will analyze the efficiency of use, which will be considered from two aspects, i.e., functional, related to efficiency and compliance with customer requirements, and economic, related to the profitability of orders. This approach was adopted because, in many publications, the authors confirm that the efficiency of transportation enterprises is a universal concept determined by the rational management of vehicles [14,15]. However, the transportation tasks carried out in transportation enterprises can be efficient from one point of view and inefficient from another [16]. Thus, the main objectives of rational vehicle management in transportation enterprises include minimizing expenditures on the operation of vehicles, thus increasing the enterprise's profit from the execution of orders [17].

For transportation enterprises, the measurable feature of the efficiency of the use of vehicles is always the revenue from the execution of transport orders and the resulting costs. The optimal management of the vehicle use process should always be aimed at maximizing revenue and minimizing costs [18].

Effective management of the system of operation of delivery vehicles expresses the entire ability to meet the specific needs of users and is focused on making rational decisions aimed at maximizing the degree of achievement of the objectives, including the bottom line [19]. Shaping and evaluating the effects of the vehicle use process in transportation enterprises is closely related to the need to optimally allocate tasks to vehicles along with maintaining their performance characteristics at an appropriate level under specific operating conditions [20]. These features, which meet the requirements of representative ones, should be controlled on an ongoing basis in order to maintain the fitness of the vehicle. With the help of the selected operating ratios, the fleet manager can verify the intensity of vehicle use in real time, analyze the profitability of the actions taken and plan the process of vehicle use. An analysis of the literature showed that there are a number of studies on the economic efficiency of vehicles, but they are conducted for heavy-duty vehicles whose payload exceeds 9.5 t [21]. However, there is a lack of studies dealing with fleets of light commercial vehicles with a cargo capacity of up to 3.5 t. In addition, due to costs, micro- and small enterprises rarely implement modern information systems to support decision-making in vehicle operation. The lack of access to specialized tools causes decision-makers to rely solely on their experience when valuing an order and undertaking its execution [22]. This causes them to rely solely on their experience when deciding to undertake an order and its valuation. They often undertake orders that are completely unprofitable for them (e.g., with a cargo weight above the permitted capacity, taking the risk of receiving a fine or damaging the vehicle) or price the order at or below the profitability limit. A characteristic feature of such entities is that they perform transports over short distances (within up

to 200 km of their headquarters) and for relatively small amounts. As a result of their decisions, these entities do not make a high profit and cannot invest in further development, including both IT systems and the purchase of a new fleet of vehicles. That is why it is important to provide simple decision-support tools, the implementation of which does not require large financial outlays. Such a tool can be the logistic regression model presented in this article for assessing the economic efficiency of vehicle operation in transportation enterprises. This research was conducted at an SME transportation enterprise that provides transportation services to a trade company operating in the FMCG (Fast-Moving Consumer Good) sector in Poland. The main characteristic of this industry, both in Poland and around the world, is the relatively low price of the products offered to customers, as well as the fact that they are sold on a mass scale and with very short delivery times [23]. The above makes it necessary for the transport processes implemented for this industry to be characterized by relatively low costs and high reliability. The performance of transportation services for this type of industry is characterized by a special approach due to the type of cargo being carried. The specific nature of the products transported, i.e., eggs, fresh vacuum-packed meat and cured meats, requires transportation enterprises to handle orders particularly quickly and efficiently, with the goal of ensuring the freshness of the products. Appropriate transportation methods and conditions, cargo security, flexibility and often taking on orders immediately are all challenges that a carrier providing food transportation services must address. From the point of view of trade companies, the most important criterion for the selection of a transportation enterprise is the price of the service provided, but given the nature of the product, the complex logistics chain and the very short delivery times, trading companies rely on the flexibility of the carrier and are willing to pay a higher price to pick up the cargo as quickly as possible and deliver it to the recipient. These types of orders are an opportunity for carriers to earn much more revenue, so it is important to be able to analyze and evaluate the factors that have a key impact on the profitability of the order being executed and the degree of this impact.

3. Methodology—Logistic Regression Model

The analysis and evaluation of the profitability of transport orders were made using a logistic regression model. It provides a mathematical notation of the effect of several variables x_1, x_2, \dots, x_n on the dichotomous variable Y , which takes the value of 1 when the order is profitable and 0 when the enterprise suffers financial losses. The mathematical function in this kind of model is the logistic function, which takes the form presented in (1) below [24]:

$$f(x) = \frac{e^x}{1 + e^x} = \frac{1}{1 + e^{-x}} \quad (1)$$

where

e —Euler number;

x —value of the explanatory variable X .

The logistic regression model is described by the following equation:

$$P(Y = 1 | x_1, x_2, \dots, x_k) = \frac{e^{\beta_0 + \sum_{i=1}^k \beta_i \cdot x_i}}{1 + e^{\beta_0 + \sum_{i=1}^k \beta_i \cdot x_i}} \quad (2)$$

where

$\beta_i \ i = 0, \dots, k$ —logistic regression coefficients;

x_1, x_2, \dots, x_k —independent variables, which can be both measurable and qualitative.

It is based on a way of expressing probability through odds, i.e., the ratio of the probability of an event occurring to the probability that it will not happen. For a given event A , the above definition takes the following form:

$$(\text{Odds})S(A) = \frac{P(A)}{P(\text{no}A)} = \frac{P(A)}{1 - P(A)} \quad (3)$$

The measure of the relationship between the variables in the above model is the odds ratio (OR), which, when comparing groups A and B, should be understood as the ratio of the chance of occurrence of A to the chance of occurrence of B. It is most often written in the following form:

$$OR_{A \times B} = \frac{S(A)}{S(B)} = \frac{P(A)}{1 - P(A)} : \frac{P(B)}{1 - P(B)} \quad (4)$$

Regression coefficients β_i , where $i = 0, \dots, k$, are estimated based on the maximum reliability method (MRM), which maximizes the reliability function or its square. The reliability function takes the following forms [25]:

$$L(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_k) = \prod_{i=1}^n p(x_i; \theta_1, \theta_2, \dots, \theta_k) \text{ for discrete variables} \quad (5)$$

$$L(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_k) = \prod_{i=1}^n f(x_i; \theta_1, \theta_2, \dots, \theta_k) \text{ for continuous variables} \quad (6)$$

where

$\theta_1, \theta_2, \dots, \theta_k$ —unknown distribution parameters;

x_1, x_2, \dots, x_n —observed values of variable X in a random sample;

$p(x)$ —distribution defined by a probability function;

$f(x)$ —distribution defined by the density function;

L —reliability function.

Taking into consideration that all observations are independent of each other, the probability of observing the entire data set is equal to the likelihood function of the individual samples.

The use of a logistic regression model is only possible after examining and meeting the following requirements:

- The dependent variable is binary;
- There is no correlation between the independent variables;
- There is a linear relationship between the probability logit and the independent variables;
- The model includes only those independent variables that have a significant impact on the dependent variable.

The assessment of the correctness of the built model and its goodness of fit is possible by doing the following:

- Using statistics used in evaluating diagnostic tests, i.e., accuracy (ACC), sensitivity (SE) and specificity (SP);
- Performing the Hosmer and Lemeshow test;
- Plotting the ROC curve and analyzing the area under the AUC curve.

To evaluate the research conducted and the results obtained, a value and the corresponding cut-off point should be determined. It is the value of the research variable for which it is possible to best divide the set into two groups (the phenomenon occurs and does not occur). This point is a fixed value of π_0 from the interval $[0,1]$, for which the following is assumed:

$$\hat{P}(Y = 1|x) > \pi_0 \quad (7)$$

To assess model goodness of fit, two quantities are most often used simultaneously: sensitivity and specificity (defined as difference $1 - SP$). For evaluation, the obtained pairs of $1 - SP$ and SE values are marked on a plane, where the horizontal axis represents $1 - SP$ values, while the vertical axis represents SE values. The combination of the obtained points forms the so-called ROC curve, which allows an overall evaluation of the predictive quality of the model. It shows all possible cut-off points and the statistics associated with them. In addition, it does not depend on the scale adopted and allows easy reading of sensitivity and specificity values.

If there is a continuous independent variable in a logistic regression model, a number of different cut-off points can be tested to distinguish occurrences from non-occurrences, depending on the predicted probability. It is important to find a point that has high values of both sensitivity and specificity. This is the optimal cut-off point, which is determined using the Youden index (J). It takes the following form:

$$J = \text{sensitivity} + \text{specificity} - 1 \quad (8)$$

The optimum cut-off point corresponds to the case where the J value reaches its maximum.

An analysis of the literature has shown that there are many scientific publications using logistic regression models to analyze diversified problems and variables. These models have also been applied to assess the profitability of companies. Zizi et. al. [26] used it to assess the determinants and predictors affecting SMEs' financial failure, considering only the macroeconomic indicators that characterize these companies. Budak and Sarvari [27] analyzed profit margin prediction in sustainable road-freight transport. Their study, however, considered a company without its own fleet of vehicles, which, when valuing orders, operates only by assessing its own margin, which is added up to the value of the order given by the truck driver. Conrad IV [28], on the other hand, carried out an analysis of the profitability of individual timber deliveries for truck owners doing business in the USA, indicating that it is often unprofitable (86%) and explaining the reasons for this. The above are only examples of the application of the regression model to similar cases; however, no studies dedicated to the analysis of the profitability of transport orders, especially for companies with a fleet of delivery vehicles, were found.

4. Results and Discussion

This article analyzes the bottom-line figures for transport orders registered in 2021–2022, which were carried out by transportation enterprises of the SME sector in the FMCG industry in Poland. Due to the knowledge of the Polish transport market and the fact that, in 2021–2022, Polish transport companies were the industry leaders among all EU countries, it was decided to choose this country as an appropriate example of process implementation [29]. Additionally, our study covered the years 2021–2022, i.e., the post-pandemic period, in which a significant increase in the number of transport operations was recorded. It focuses on the profitability of their execution and the analysis of the impact of vehicle performance on enterprises' bottom line. An analysis of operating costs showed that for a transport order to be profitable, its value should be min. PLN 500. Therefore, an order profitability index with a binomial distribution was indicated as the dependent variable, where 1 meant that the order was profitable and 0 meant that its cost exceeded revenue, generating a loss. Therefore, a logistic regression model was used. The data set size was 13,000 observations. They were divided into two subsamples, test and training, in a ratio of 80–20.

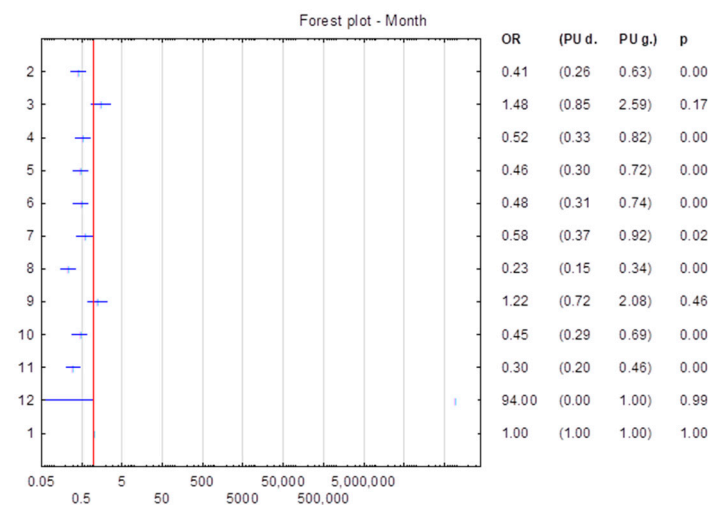
Building the model began by selecting the factors that could affect the explanatory variable. The ones selected were those related to the operating conditions of the vehicles, i.e., daily mileage [km], operating time [min], driving speed [km/h], cargo weight [kg], fuel consumption [l] and the seasonality of the deliveries made, i.e., day of the week and month of the order. The first part of this study analyzed and evaluated the impact of the selected parameters on the dependent variable. For this purpose, the LR test, which allows testing whether a univariate model containing a given variable is significantly better than a model containing only an absolute term, was used. The results are presented in Table 1.

Table 1. LR test results.

Variable	<i>p</i> -Value
daily mileage	0.00
operating time	0.00
driving speed	0.01
fuel consumption	0.00
cargo weight	0.00
month	0.00
day of the week	0.09

The calculated *p*-value for the variable *day of the week* was 0.09, which means that the null hypothesis of no significant effect of this variable on the analyzed phenomenon could not be rejected. Therefore, the variable was also eliminated and not considered further in our study. For the remaining variables, the calculated *p*-value of the test conducted indicated that the alternative hypothesis, stating that their effect on *profitability* is significant, should be accepted.

The next step examined the strength of the effect of the qualitative variables on the phenomenon under study and their statistical significance. For this purpose, a forest plot, which shows the odds ratios calculated from the regression parameter evaluations, was used (Figure 1).

**Figure 1.** Forest plot for *month* variable (red line means the same effect like January).

The plot above was made for the odds ratios, which refer to the regression parameter ratings for the *month* variable. It should be noted that for the *month* variable, a significant difference of 1 was not found only for the months of March, September and December; therefore, this variable was included in the further study.

The use of a logistic regression model is only possible if the assumption of linearity of the explanatory variables is met. For the selected quantitative variables, a graphical assessment of the scatterplot was performed along with the analysis of the corresponding logarithm of odds values, and a linearity test was performed [30]. The calculated *p*-value indicated that there was no basis to reject the null hypothesis of the linear effect of the studied variables on the analyzed phenomenon (Table 2).

Table 2. LR linearity test results.

Variable	<i>p</i> -Value
daily mileage	0.25
operating time	0.18
driving speed	0.21
fuel consumption	0.65
cargo weight	0.45

In the next step, the collinearity of the explanatory variables selected to build the model was checked. For this purpose, the correlation between the quantitative variables was tested (Table 3). The significance of the correlation is marked in red.

Table 3. Correlation coefficients between quantitative variables.

Variable	Daily Mileage	Operating Time	Driving Speed	Fuel Consumption	Cargo Weight
daily mileage	1.00	0.85	0.25	0.08	0.05
operating time	0.85	1.00	−0.13	0.06	0.03
driving speed	0.25	−0.13	1.00	0.03	0.04
fuel consumption	0.08	0.06	0.03	1.00	0.93
cargo weight	0.05	0.03	0.04	0.93	1.00

The analysis performed made it possible to isolate two clusters of correlated variables, i.e., *fuel consumption* and *cargo weight*, and *daily mileage* and *operating time*. Accordingly, it was decided to eliminate the variables *fuel consumption* and *operating time* from further analysis.

The next step was to build a multivariate model. For the qualitative variables, parameterizations with an assumed baseline were adopted. Model building was carried out using the backward-step method, which allowed the elimination of irrelevant predictors from the model. The 10-fold cross-validation method, which is an intermediate method between strategies based on splitting the data set into a training and a testing sample and the LOO (leave-one-out) strategy, was indicated as a way to validate the model. It involves dividing the sample into 10 subsets, each of which acts as a subset of the observations used to calculate the prediction errors. The global error score is the average of the values from each subset.

The estimated values of the model parameters are presented in Table 4. On the other hand, the results of the LR test for the significance of the model parameters are shown in Table 5.

All the estimated model parameters proved to be statistically significant. In the case of the *month* variable for March, April, September and December, the calculated value of the Wald statistic and the corresponding *p-value* were greater than the adopted significance level of 0.05, so they will be eliminated from the model. The equation of the logistic regression model takes the following form:

$$P(\text{profitability} = 1|X) = \frac{e^a}{1 + e^a} \quad (9)$$

$$\text{where } a = -6.248 - 0.597 \cdot \text{February} - 1.028 \cdot \text{May} - 0.807 \cdot \text{June} - 0.609 \cdot \text{July} - 1.584 \cdot \text{August} - 1.032 \cdot \text{October} - 1.471 \cdot \text{November} + 0.009 \cdot \text{mileage} - 0.031 \cdot \text{driving speed} + 0.588 \cdot \text{cargo weight} \quad (10)$$

Table 4. Estimated values of parameters in the logistics regression model.

Variable	Modeled Probability: Profitability Distribution: Binomial; Binding Function: LOGIT				
	Effect Level	Parameter	Estimation Standard Error	Wald's Test Statistics	<i>p</i>
intercept		−6.248	0.769	65.997	0.000
month	2	−0.597	0.231	6.712	0.01
month	3	0.284	0.29	0.962	0.327
month	4	−0.329	0.237	1.917	0.166
month	5	−1.028	0.231	19.829	0.000
month	6	−0.807	0.231	12.227	0.000
month	7	−0.609	0.241	6.372	0.012
month	8	−1.584	0.214	54.545	0.000
month	9	−0.154	0.278	0.308	0.579
month	10	−1.032	0.231	19.948	0.000
month	11	−1.471	0.222	43.966	0.000
month	12	20.685	5174.995	0.000	0.997
daily mileage		0.009	0.000	440.264	0.000
driving speed		−0.031	0.007	17.341	0.000
cargo weight		0.588	0.049	145.664	0.000

Table 5. Results of LR test for significance of model parameters.

Variable	<i>p</i>
month	0.00
mileage	0.00
driving speed	0.00
cargo weight	0.00

The equation presented above can take the equivalent form of the odds presented in (11) below:

$$P \frac{P(\text{profitability} = 1|X)}{1 - P(\text{profitability} = 1|X)} = e^a \quad (11)$$

$$\text{where } a = 0.55 \cdot \text{February} + 0.36 \cdot \text{May} + 0.45 \cdot \text{June} + 0.54 \cdot \text{July} + 0.21 \cdot \text{August} + 0.36 \cdot \text{October} + 0.23 \cdot \text{November} + 1.01 \cdot \text{mileage} - 0.97 \cdot \text{driving speed} + 1.09 \cdot \text{cargo weight} \quad (12)$$

Finally, the goodness of fit of the model was tested based on a comparison of predicted and observed values. For this purpose, the Hosmer–Lemeshow (HS) test was used. The HS test is based on comparing the incidence count of the modeled class with its actual realization. The null hypothesis is that the observed and expected counts are not different. The calculated *p*-value of the HS test was 0.42 and did not warrant the rejection of the null hypothesis, indicating that the model is a good fit for the data.

Next, calculations were made of the odds ratios that the order would be profitable when the various parameters associated with the operation of the vehicle were met. The

sign next to the calculated parameter values indicates whether the analyzed factor is a stimulant or a destimulant of the odds of the phenomenon under study.

The calculated odds ratio values (Table 6) indicated that each of the model parameters significantly affects the profitability of a transport order. In terms of daily vehicle mileage, the odds ratio indicated that with each additional kilometer driven, the probability of an order being profitable increases by 1%. When considering driving speed, it was estimated that as its value increases by 1 km/h, the probability of order profitability decreases by 3%. In contrast, an increase of 1 kg of cargo weight makes the odds of profitability of an order increase by 9%. This is because the increase in freight rates is a result of the increase in the weight of goods carried. With regard to the month of order execution, it is inferred from the results that, with reference to January, the probability of order profitability is 0.55 times lower in February, 0.36 times lower in May, 0.45 times lower in June, 0.54 times lower in July, 0.21 times lower in August, 0.36 times lower in October and 0.23 times lower in November.

Table 6. Estimated values of odds ratios.

Variable	Modeled Probability: Profitability Distribution: Binomial; Binding Function: LOGIT				
	Effect Level	Odds Ratio	95.00% OR	−95.00% OR	<i>p</i>
intercept		0.00	0.00	0.01	0.00
month	2	0.55	0.35	0.86	0.00
month	5	0.36	0.23	0.56	0.00
month	6	0.45	0.28	0.70	0.00
month	7	0.54	0.34	0.87	0.00
month	8	0.21	0.13	0.31	0.00
month	10	0.36	0.23	0.56	0.00
month	11	0.23	0.15	0.35	0.00
daily mileage		1.01	1.01	1.01	0.00
driving speed		0.97	0.96	0.98	0.00
cargo weight		1.09	1.64	1.98	0.00

The developed model can be used to predict the profitability of a transport order depending on the parameters of its execution, so an evaluation of the quality of the prediction was carried out. For this purpose, an ROC curve was plotted, with the AUC, or area under the ROC curve, being the most important parameter for evaluation. It takes values from 0 to 1. The larger the area under the curve, the greater the discrimination of the model, for the evaluation of which Kleinbaum and Klein proposed the following classification (Table 7) [31,32].

Table 7. AUC classification.

AUC Value	Score
0.9 < AUC < 1.0	Excellent discrimination
0.8 < AUC < 0.9	Good discrimination
0.7 < AUC < 0.8	Sufficient discrimination
0.6 < AUC < 0.7	Weak discrimination
0.5 < AUC < 0.6	Insufficient discrimination

As cross-validation was used to validate the model, ROC curves were plotted separately for the training and testing sample sets and AUC areas were calculated (Figures 2 and 3). In addition, a cut-off point and a classification matrix of actual and model-indicated observations were determined (Tables 8 and 9).

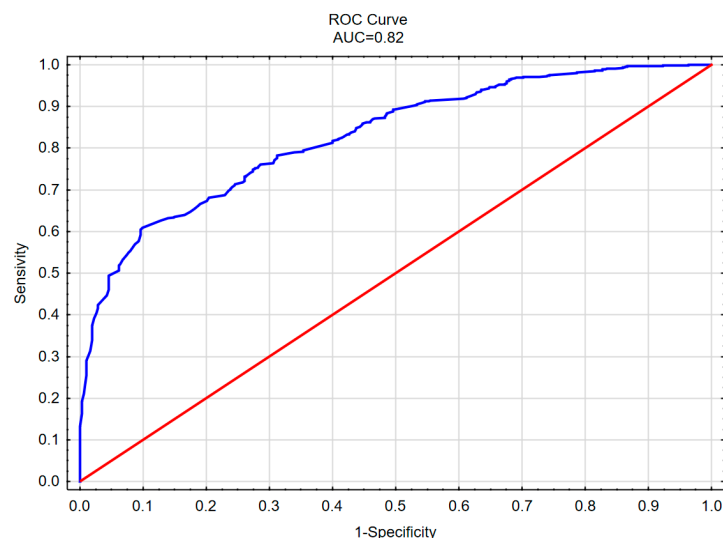


Figure 2. ROC curve for the training set (red line means random classifier).

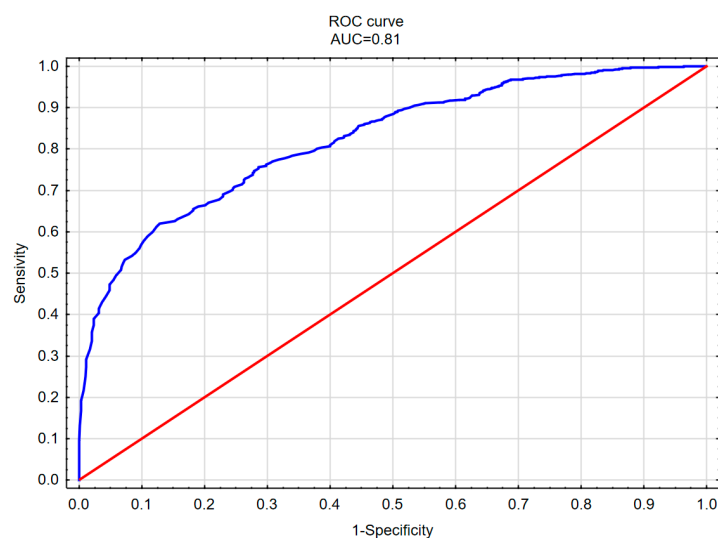


Figure 3. ROC curve for the testing set (red line means random classifier).

Table 8. Classification matrix—factual status and model indication.

	Anticipated: 1	Anticipated: 0	Percentage of Correct
Observed:1	7693	4733	62%
Observed:0	87	588	87%

Table 9. Cut-off point of the model studied.

Observation No.	Cut-Off Point	True Positive	True Negative	False Positive	False Negative	SE	1-SP
34	0.49	7693	588	87	4733	0.619	0.129

The calculated AUC values for each set were 0.82 for the training set and 0.81 for the testing set, which allows us to conclude that they are characterized by sufficient discrimination. This demonstrates both sufficient adaptation to the data and sufficient quality of the model when new data come in. In addition, the calculated AUC values were close to each other (the difference between them did not exceed 0.05), which allows the model to be considered correct.

For the proposed cut-off point, the sensitivity was 0.619 and the specificity was 0.129. There were 8281 correctly classified cases (7693 positives and 588 true negatives) and 4820 misclassified cases (87 false positives and 4733 false negatives). This allows the built model to be considered correct.

This literature analysis has shown that there are a number of studies on the economic efficiency of vehicles, but they are conducted for trucks weighing more than 9.5 tons. However, there is a lack of research on fleets of light commercial vehicles with a load capacity of up to 3.5 t. However, there are several studies that have examined the profitability of transport orders. Research available in the literature indicates, similarly to the study above, that profitability is influenced by the length of the transport route, increasing the probability as the transport distance increases [28]. It was also indicated that increasing the transported load increases the probability of making a profit, which is also consistent with the results obtained (an increase in fuel consumption increases the probability of profitability). Similar results were also obtained using other mathematical models and ratio analysis methods for vehicles with a load capacity of over 60 t [33] and for orders acquired from the transport exchange market [34]. It should be noted that despite the differences in the load capacity of trucks, similar results were obtained.

5. Conclusions

The study presented in this article was conducted on the basis of data on the execution of transport orders by SME enterprises in the FMCG industry. It was prompted by the peculiarities of the industry's operation and their impact on the profitability of orders.

The purpose of this article was to develop a model for evaluating the profitability of transport orders based on selected parameters related to vehicle operation. It was made using the value of the order and the costs it generates, thus indicating how the selected parameters affect the risk of loss for transportation enterprises. Based on the results obtained, it can be concluded that each of the selected parameters related to operation significantly affects the profitability of a transport order. In terms of daily vehicle mileage, the odds ratio indicated that with each additional kilometer driven, the probability of an order being profitable increases by 1%. When considering driving speed, it was estimated that as its value increases by 1 km/h, the probability of order profitability decreases by 3%. In contrast, an increase of 1 kg of cargo weight makes the odds of profitability of an order increase by 9%. This is because the increase in freight rates is a result of the increase in the weight of goods carried. With regard to the month of order execution, it is inferred from the results that, with reference to January, the probability of order profitability is 0.55 times lower in February, 0.36 times lower in May, 0.45 times lower in June, 0.54 times lower in July, 0.21 times lower in August, 0.36 times lower in October and 0.23 times lower in November.

SMEs do not have large budgets, so they often cannot invest in advanced software to support the management of transportation processes. In addition, the nature of the FMCG industry and the low price of the products sold force costs to be reduced to a minimum. Therefore, the developed tool, due to its simple form and ability to easily draw conclusions, can be an ideal tool to support decision-makers at the stage of accepting and pricing transport orders.

Author Contributions: Conceptualization, M.G. and P.O.; methodology, M.G., P.O., D.V., R.-M.S. and R.V.; software, M.G. and P.O.; validation, D.V., R.-M.S. and R.V.; formal analysis, M.G., P.O., D.V., R.-M.S. and R.V.; investigation, M.G. and P.O.; data curation, M.G. and P.O.; writing—original draft preparation, M.G., R.-M.S. and D.V.; writing—review and editing, M.G., R.-M.S. and D.V.;

visualization, R.-M.S. and D.V.; supervision, M.G.; project administration, P.O.; funding acquisition, M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to trade secret.

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

References

1. Młynarski, S.; Pilch, R.; Smolnik, M.; Szybka, J. Methodology of network systems reliability assessment on the example of urban transport. *Eksplloat. i Niezawodn.—Maint. Reliab.* **2018**, *20*, 278–283. [CrossRef]
2. Świdorski, A.; Jóźwiak, A.; Jachimowski, R. Operational quality measures of vehicles applied for the transport services evaluation using artificial neural networks. *Eksplloatacja i Niezawodnosc—Maint. Reliab.* **2018**, *20*, 292–299. [CrossRef]
3. Owczarek, P.; Brzeziński, M.; Zelkowski, J. Evaluation of light commercial vehicles operation process in a transport company using the regression modelling method. *Eksplloatacja i Niezawodnosc—Maint. Reliab.* **2022**, *24*, 13. [CrossRef]
4. Zabielska, A.; Jacyna, M.; Lasota, M.; Nehring, K. Evaluation of the efficiency of the delivery process in the technical object of transport infrastructure with the application of a simulation model. *Eksplloatacja i Niezawodnosc—Maint. Reliab.* **2023**, *25*, 1. [CrossRef]
5. Gołda, P.; Zawisza, T.; Izdebski, M. Evaluation of efficiency and reliability of airport processes using simulation tools. *Eksplloatacja i Niezawodnosc—Maint. Reliab.* **2021**, *23*, 659–669. [CrossRef]
6. Wasiak, M.; Jacyna-Gołda, I.; Markowska, K.; Jachimowski, R.; Kłodawski, M.; Izdebski, M. The use of a supply chain configuration model to assess the reliability of logistics processes. *Eksplloatacja i Niezawodnosc—Maint. Reliab.* **2019**, *21*, 367–374. [CrossRef]
7. Jóźwiak, A.; Owczarek, P.; Prochowski, L.; Świdorski, A. Analysis of the impact of the use time of n1 motor vehicles on the economic efficiency of their maintenance. *Eksplloatacja i Niezawodnosc—Maint. Reliab.* **2020**, *22*, 121–129. [CrossRef]
8. Skerlic, S.; Sololovskij, E.; Erculj, V. Maintenance of heavy trucks: An international study on truck drivers. *Eksplloatacja i Niezawodnosc—Maint. Reliab.* **2020**, *3*, 493–500. [CrossRef]
9. Loska, A.; Paszkowski, W. Geometric approach to machine exploitation efficiency: Modelling and assessment. *Eksplloatacja i Niezawodnosc—Maint. Reliab.* **2022**, *24*, 114–122. [CrossRef]
10. Wasiak, M.; Zdanowicz, P.; Nivette, M. Research on the effectiveness of alternative propulsion sources in high-tonnage cargo transport. *Arch. Transp.* **2021**, *60*, 259–273. [CrossRef]
11. Kwasiborska, A.; Gładyś, S.; Banaszek, K.; Manewrowski, J.; Roś, M. Analysis of aircraft equipment requirements to conduct RNP APCH approach operation. *Syst. Logistyczne Wojsk* **2021**, *55*, 101–120. [CrossRef]
12. Czerepicki, A. Study on effectiveness of using column-oriented databases in the processing of measurement characteristics of an electric vehicle. *Arch. Transp.* **2019**, *51*, 77–84. [CrossRef]
13. Rykała, Ł.; Typiak, A.; Typiak, R.; Rykała, M. Application of Smoothing Spline in Determining the Unmanned Ground Vehicles Route Based on Ultra-Wideband Distance Measurements. *Sensors* **2022**, *22*, 8334. [CrossRef] [PubMed]
14. Szkutnik-Rogoż, J.; Ziółkowski, J.; Małachowski, J.; Oszczypała, M. Mathematical Programming and Solution Approaches for Transportation Optimisation in Supply Network. *Energies* **2021**, *14*, 7010. [CrossRef]
15. Karwat, B.; Rubacha, P.; Stańczyk, E. Optimization of a screw conveyor's exploitation parameters. *Eksplloatacja i Niezawodnosc—Maint. Reliab.* **2021**, *23*, 285–293. [CrossRef]
16. Kozłowski, E.; Borucka, A.; Oleszczuk, P.; Jałowicz, T. Evaluation of the maintenance system readiness using the semi-Markov model taking into account hidden factors. *Eksplloatacja i Niezawodnosc—Maint. Reliab.* **2023**, *25*. [CrossRef]
17. Roman, M.; Waśkiewicz, M. Optimizing Municipal Waste Collection: A Case Study of a City in Poland. *Mil. Logist. Syst.* **2023**, *58*, 197–218. [CrossRef]
18. Gołębiowski, P.; Kukulski, J. Preliminary study of shaping the railway track geometry in terms of their maintenance costs and capacity. *Arch. Transp.* **2020**, *53*, 115–128. [CrossRef]
19. Gądek-Hawlena, T.; Heliosz, P. The impact of supply chain disruptions on the opportunity costs of road transport companies in Poland. A case study. *Mil. Logist. Syst.* **2023**, *58*, 181–196. [CrossRef]
20. Voicu, D.; Stoica, R.M.; Vilau, R.; Marinescu, M.; Digulescu, A.; Despina-Stoian, C.; Popescu, F. Frequency Analysis of Vibrations in Terms of Human Exposure While Driving Military Armoured Personnel Carriers and Logistic Transportation Vehicles. *Electronics* **2023**, *12*, 3152. [CrossRef]
21. Prochowski, L. Evaluation of the process of mileage growth during the operation of motor trucks, in several categories of engine cubic capacity. *Eksplloatacja i Niezawodnosc—Maint. Reliab.* **2018**, *20*, 359–370. [CrossRef]
22. Thrassou, A.; Uzunboyu, N.; Vrontis, D.; Christofi, M. Digitalization of SMEs: A review of opportunities and challenges. In *The Changing Role of SMEs in Global Business: Volume II: Contextual Evolution Across Markets, Disciplines and Sectors*; Springer: Berlin, Germany, 2020; pp. 179–200.
23. What Is the FMCG Industry? What Are the Earnings Like There? Available online: <https://businessinsider.com.pl/poradnik-finansowy/oszczedzanie/branza-fmcg-co-to-jest-i-jakie-produkty-obejmuje/hgpmc68> (accessed on 12 April 2024).

24. Stanisław, A. *Logistic Regression Models: Applications in Medicine, Life and Social Sciences*; Stasoft Editorial Office: Krakow, Poland, 2016.
25. Borucka, A. Logistic regression in modeling and assessment of transport services. *Open Eng.* **2020**, *10*, 26–34. [[CrossRef](#)]
26. Zizi, Y.; Oudgou, M.; El Moudden, A. Determinants and Predictors of SMEs' Financial Failure: A Logistic Regression Approach. *Risks* **2020**, *8*, 107. [[CrossRef](#)]
27. Budak, A.; Sarvari, P.A. Profit margin prediction in sustainable road freight transportation using machine learning. *J. Clean. Prod.* **2021**, *314*, 127990. [[CrossRef](#)]
28. Conrad, J.L., IV. Evaluating profitability of individual timber deliveries in the US South. *Forests* **2021**, *12*, 437. [[CrossRef](#)]
29. Road Freight Transport by Journey Characteristics. Available online: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Road_freight_transport_by_journey_characteristics#Road_freight_transport_performance_by_type_of_operation_28in_tonne-kilometres.29 (accessed on 12 April 2024).
30. Kozłowski, E.; Borucka, A.; Szymczak, T.; Gil, L. Predicting the Fatigue Life of a Ball Joint. *Transp. Telecommun. J.* **2021**, *22*, 453–460. [[CrossRef](#)]
31. Flis, M.; Czyżowski, P.; Beeger, S.; Rataj, B.; Karpiński, M. Logistic Regression Model for Determination of the Age of Brown Hare (*Lepus europaeus* Pall.) Based on Body Weight. *Animals* **2022**, *12*, 529. [[CrossRef](#)]
32. Joshi, R.D.; Dhakal, C.K. Predicting Type 2 Diabetes Using Logistic Regression and Machine Learning Approaches. *Int. J. Environ. Res. Public Health* **2021**, *18*, 7346. [[CrossRef](#)]
33. Väättäinen, K.; Laitila, J.; Anttila, P.; Kilpeläinen, A.; Asikainen, A. The influence of gross vehicle weight (GVW) and transport distance on timber trucking performance indicators—Discrete event simulation case study in Central Finland. *Int. J. For. Eng.* **2020**, *31*, 156–170. [[CrossRef](#)]
34. Woźniak, W.; Stryjski, R.; Mielniczuk, J.; Wojnarowski, T. The Concept of the Profitability for the Transport Orders Acquired from the Transport Exchange Market. In Proceedings of the 27th IBIMA Conference, Milan, Italy, 4–5 May 2016.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.