ORIGINAL PAPER

PANEL DATA MODEL – MATHEMATICS ACHIEVEMENT OF COST RE-EDUCATION BASED ON THE IMPACT OF TOTAL COST VARIATION

CONSTANTIN AURELIAN IONESCU[12], MELINDA TIMEA FÜLÖP[3], DAN IOAN TOPOR[4], ALEXANDRA DELIA BUGNARIU[4], NICOLAE MĂGDAŞ[3]

Manuscript received: 18.01.2024; Accepted paper: 22.08.2024;
Published online: 30.09.2024.

Abstract. *By taking into account temporal and individual effects, panel data models enable a precise analysis of the effects of total cost variations on cost redistribution, which is crucial for cost structure optimization. With overall price increases, businesses are also seeing higher costs. However, one piece of data is a specific concern: a business's total cost. This research examined the total cost components to determine the factors affecting the total cost variation. We examined a listed business that has its activity structured in cost centers, presenting descriptive statistics of the independent and dependent variables of the study. A panel analysis was applied for the business to determine the factors that influence manufacturing costs and how these costs can be improved. The results show that personnel costs, costs allocated to the application of new software and hardware technologies, and quality management costs have significant statistical and economic effects. The outcomes indicate that the cost centers are crucial in organizing a business that aims to decrease expenses. Thus, reducing or eliminating costs linked to some factors with a noticeable arithmetical influence on the total expenses is possible. This examination is the initial effort to establish the impact of determining factors on total cost variation and associate this with businesses' opinions of the cost. The results provide an amplified appreciation of the bases and levels of the costs of sourcing.*

Keywords: *costs; quality management costs; efficiency; cost centers; cost accounting.*

1. INTRODUCTION

Econometric and mathematical analysis of panel data models plays a central role in determining cost redistribution based on total cost variation. Panel data models combine data from different time points and units, allowing both temporal and individual effects to be included in the analysis. Mathematically, such models often use extended linear regression approaches supplemented with fixed or random effects to reflect the heterogeneity of observation units. This allows for a nuanced analysis of the impact of changes in total costs on the distribution of costs between different units. In particular, by using such models, systematic influences can be isolated and accurate estimates of the effects of total cost

[1] Valahia University of Targoviste, Institute of Multidisciplinary Research for Science and Technology, 130004 Targoviste, Romania. E-mail: ionescu.aurelian@icstm.ro.
[2] Hyperion University of Bucharest, Faculty of Economic Sciences, 030615 Bucharest, Romania.
[3] Babeş-Bolyai University, Faculty of Economics and Business Administration, 400084 Cluj-Napoca, Romania. E-mail: melinda.fulop@econ.ubbcluj.ro; nicolae.magdas@econ.ubbcluj.ro.
[4] 1 Decembrie 1918 University of Alba Iulia, Faculty of Economic Science, 510009 Alba Iulia, Romania. E-mail: dan.topor@uab.ro; alexandra.bugnariu@uab.ro.

https://doi.org/10.46939/J.Sci.Arts-24.3-a16

variation on the cost structure can be created. This is crucial for the development of effective cost optimization and reallocation strategies in different economic contexts.

In an uncertain and ever-changing world, businesses are increasingly faced with saturation and competition in the marketplace. To remain competitive, managers aim to optimize a balance of future costs and revenues [1]. Managers must decide how to control production costs to maintain profitability; implement cost-saving measures, such as automation, negotiating with suppliers for lower raw material prices; find ways to reduce waste reduction, and allocate resources to minimize production costs [2]. The significance of comprehensive examination and reorganization of decision-making is cumulative in today's multifaceted commercial processes. Accountability bookkeeping was introduced to encounter the necessity for decentralization, truthful evidence, and analysis. Understanding the association between responsibility centers provides an organized flow of evidence that can improve decision-making and cost control [3].

Cost efficiency is a critical factor in determining the success of a business. It significantly impacts managerial performance by focusing on the ability to produce goods or services at the lowest possible cost while maintaining the desired quality. This objective can be achieved through strategies such as process improvement, cost reduction measures, and efficiency improvement through the use of technology, downsizing, consolidating operations, and outsourcing secondary functions [4,5]. Effective cost management can improve financial performance as lower costs lead to higher profits. Cost efficiency can also increase competitiveness and improve operational performance, as managers can focus on core business functions and allocate resources more efficiently [6].

A major benefit of cost efficiency is its direct impact on a business's bottom line. By reducing costs, organizations can increase their profit margins and become more competitive, leading to more market share and improved financial performance. In addition, cost efficiency helps organizations become more resilient to economic downturns and external pressures, such as increased competition or changes in market conditions. However, if managers become too focused on cost efficiency, they can neglect other important aspects of the business, such as innovation, product development, and employee engagement [7,8].

Cost and performance accounting, with its various design options, is the central tool of operational control. It processes information about costs and revenues based on the company's past, present, and future planning and control processes, including the relationships between the company and its environment [9]. According to this confirmation, this study of empirical efficiency aims to identify, understand, and analyze influencing factors and various production costs. This objective will yield guidance for decisions on the overall cost variance, help in policy making, and contribute to achieving organizational goals.

2. MATERIALS AND METHODS

One essential responsibility center in organization management is the cost center (CC). In this empirical study, we analyze the impact of some determining factors on the variation of the total cost (*Ctotal*). This study is based on data from a company listed on the stock exchange, and its activity is organized in ten centers of responsibility, as listed in Table 1.

We analyze the impact of the following determining factors: personnel expenses (*Cpersonal*), general expenses (*Cgeneral*), energy expenses (*Cenergy*), expenses allocated to the implementation of new software and hardware technologies to increase the quality and productivity of work (*Cit*), expenses for increasing quality management (*Cmanagement*) and transport expenses (*Ctransport*), as listed in Table 2.

Table 1. Cost centers that form the object of study.

CC maintenance Cost center 0 Cost center 1	Personnel expenses General costs and ongoing equipment maintenance costs Energy costs IT expenses Quality management costs Transportation expenses
CC production Cost center 2 Cost center 3 Cost center 4 Cost center 5 Cost center 6 Cost center 7 Cost center 8 Cost center 9	Personnel expenses General cost and additional disposal costs Energy costs IT expenses Quality management costs Transportation expenses

The organization of the data includes the systematization of the statistical data obtained by collecting data from the company's financial accounting documents structured in ten cost centers. Panel data were chosen as a form of organization, taking into account the purpose of the empirical study.

Table 2. Selected statistical variables.

Variable	**Description**
Ctotal	Total cost
Cpersonal	Personnel expenses
Cgeneral	General expenses
Cenergy	Energy costs
Cit	IT expenses
Cmanagement	Quality management costs
Ctransport	Transportation expenses

Panel data assume repeated measurements at different moments (or periods) on the same statistical entities, such as individuals, companies, regions, or states. In this case, the regression can highlight the variation in relation to statistical units and time. Specifically, in our case, the statistical units are cost centers, and time is measured in months. The parameter estimation methods used in this case are much more complex than those using cross-sectional data mirrored at a given period (year, month, quarter, etc.). The standard errors of the parameter estimators need to be adjusted because of the temporal correlation of the data. Panel data presuppose the use of much more complex models and estimation methods. The approach to statistical methodology using panel data requires us to distinguish between fixed- and random-effect models [10,11]. The categories of panel data are numerous, and analyzing them leads to various models. Generally, panel data are perceived at systematic period intervals. In our case, we have a balanced panel, meaning that all statistical units are observed in each period. Of note, this empirical study falls within the methodology specific to microeconometrics. Regardless of the assumptions, certain corrections are needed on the estimators based on the method of ordinary least squares (OLS), and to assess the increase in efficiency [12], we use the method of generalized least squares (GLS) [13].

The model with the specific effect of statistical entities for a dependent variable y_{it} has the following form:

$$y_{it} = \alpha_i + x_{it}^{'}\beta + \varepsilon_{it} \tag{1}$$

where: x_{it} is the regressor,
α_i is the specific effect, and

ε_{it} is the error of the i^{th} statistical entity at time t, where $i = \overline{1,N}$ and $t = \overline{1,T}$; N represents the number of statistical entities; and T is the number of time periods.

With respect to α_i, we have two types of models: fixed-effect (FE) and random-effect (RE). In the FE model, $\alpha_i$in (1) can be correlated with the regressors x_{it}. This allows for a limited form of endogeneity. The errors in (1) can also be written as follows:

$$u_{it} = \alpha_i + \varepsilon_{it} \quad (2)$$

where x_{it} can be correlated with the time-invariant component of the error α_i but uncorrelated with the time-varying component ε_{it}. We can obtain a consistent estimate of the β parameters for the time-varying regressors from (1) using special transformations applied to the equation to remove the α_i component.

In the RE model, we assume that α_i in (1) has random behavior. This restrictive condition implies a non-correlation of the regressors with α_i. Parameter estimation, in this case, involves the use of a feasible method applied to GLS, called feasible GLS (FGLS) [14]. Applying both estimation models has advantages and disadvantages. The largest disadvantage is that the estimates obtained are inconsistent if we use the RE model when effects are actually fixed. As a result, it is important to determine the type of effect specific to these panel data using special statistical tests.

Researchers use the combined (or group) econometric model (also called the average model of the statistical population, population-averaged model, or pooled model in the literature) when the regressors are exogenous [15]. In this case, we use the OLS estimator, provided we have control over the errors. If the errors are correlated over a period, for a given statistical entity, we use the within-group correlation estimator, and if a correlation between entities is possible, then we use the between-unit correlation estimators. Typically, the group model constitutes the starting model, after which the error behavior is analyzed. We apply this recommendation of the theoreticians during this empirical study, employing different estimators to estimate the parameters in the models presented above [16].

3. RESULTS AND DISCUSSION

The data panel contains statistical data for the variables in Table 4 relating to ten cost centers in the period of 2017–2021 and the first three months of 2022. More specifically, we have a database with 630 observations (10 centers × 5 years × 12 months + 3 months × 10 centers) for the seven variables. Table 3 lists the variables and their type, format, and description, which help explain how statistical processors work and help users find errors.

Table 3. Internal data format for the 630 observations.

Variable	Storage Type	Display Format	Variable Label
Ctotal	Float	%9.0g	Total cost
Cpersonal	Float	%9.0g	Personal expenses
Cgeneral	Float	%9.0g	General expenses
Cenergy	Float	%9.0g	Energy cost
Cit	Float	%9.0g	IT expenses
Cmanagement	Float	%9.0g	Quality management cost
Ctransport	Float	%9.0g	Transportation expenses

Table 4 contains descriptive statistics regarding the behavior of the model variables involved. The empirical study aims to analyze the impact of the independent variables

C*personal*, C*general*, C*energy*, C*it*, C*management*, and C*transport* on the dependent variable C*total*.

Table 4. Descriptive statistics.

Variable	No. of Observations	Mean value	Standard deviation	Minimum value	Maximum value
CC	630	5.5	2.874564	1	10
Month	630	32	18.19869	1	63
Ctotal	630	1,323,673	772,273.5	91,154	5,732,462
Cpersonal	630	24.32202	6.732336	10.01	43.62
Cgeneral	630	19.14043	7.634567	10.16	41.78
Cenergy	630	9.532571	4.362664	1.85	28.99
Cit	630	19.75997	8.688497	9.06	53.79
Cmanagement	630	2.777587	1.121272	0.63	6.88
Ctransport	630	6.707492	2.9575	3	16.71

Table 5 provides an overview comparing the variability between entities and that within entities for each variable. In other words, global variation, measured around the mean, can be divided into two components: within entities and between entities. The variables considered in the model show both internal and external variations. Therefore, we can use either the between or the within estimator.

Table 5. Internal and external variation.

Variable	Variability Type	Mean value	Standard deviation	Minimum value	Maximum value	No. of Observations
CC	Overall	5.5	2.874564	1	10	$N = 630$
	Between		3.02765	1	10	$n = 10$
	Within		0	5.5	10	$T = 63$
Month	Overall	32	18.19869	1	63	$N = 630$
	Between		0	32	32	$n = 10$
	Within		18.19869	1	63	$T = 63$
Ctotal	Overall	1323673	772,273.5	91,154	5,732,462	$N = 630$
	Between		432,983.4	568,348.5	2,047,564	$n = 10$
	Within		653,766.8	−21,423.79	5,664,724	$T = 63$
Cpersonal	Overall	24.32202	6.732336	10.01	43.62	$N = 630$
	Between		5.02945	18.59159	32.99746	$n = 10$
	Within		4.745769	9.433127	44.41408	$T = 63$
Cgeneral	Overall	19.14043	7.634567	10.16	41.78	$N = 630$
	Between		6.769154	12.44524	32.21778	$n = 10$
	Within		4.120892	4.922651	36.339	$T = 63$
Cenergy	Overall	9.532571	4.362664	1.85	28.99	$N = 630$
	Between		0.9065732	7.55381	10.7619	$n = 10$
	Within		4.276911	1.579397	28.3194	$T = 63$
Cit	Overall	19.75997	8.688497	9.06	53.79	$N = 630$
	Between		3.772698	15.1881	27.20302	$n = 10$
	Within		7.915787	4.975365	54.30346	$T = 63$
Cmanagement	Overall	2.777587	1.121272	0.63	6.88	$N = 630$
	Between		0.3222514	2.02381	3.113651	$n = 10$
	Within		1.078722	0.9640953	6.83219	$T = 63$
Ctransport	Overall	6.707492	2.9575	3	16.71	$N = 630$
	Between		2.707592	4.179048	11.55905	$n = 10$
	Within		1.462317	1.080667	12.41067	$T = 63$

3.1. MAPPING PANEL DATA RESULTS

An analysis based on tables should be supplemented with a series of graphical representations. Tables and graphs complement each other, allowing us to hypothesize the type of estimators used to assess parameters. The choice of an econometric model and estimator of its parameters depends largely on the behavior of panel data.

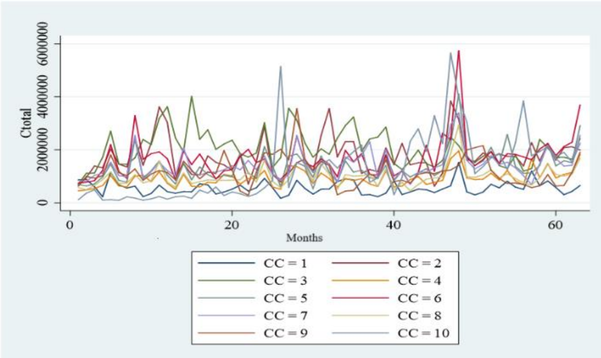

Figure 1. Evolution of *Ctotal*

Fig. 1 shows the slightly increasing evolution of the dependent variable *Ctotal* for all cost centers. We are interested in analyzing which of the six influencing factors have a significant impact on the variation of the total cost.

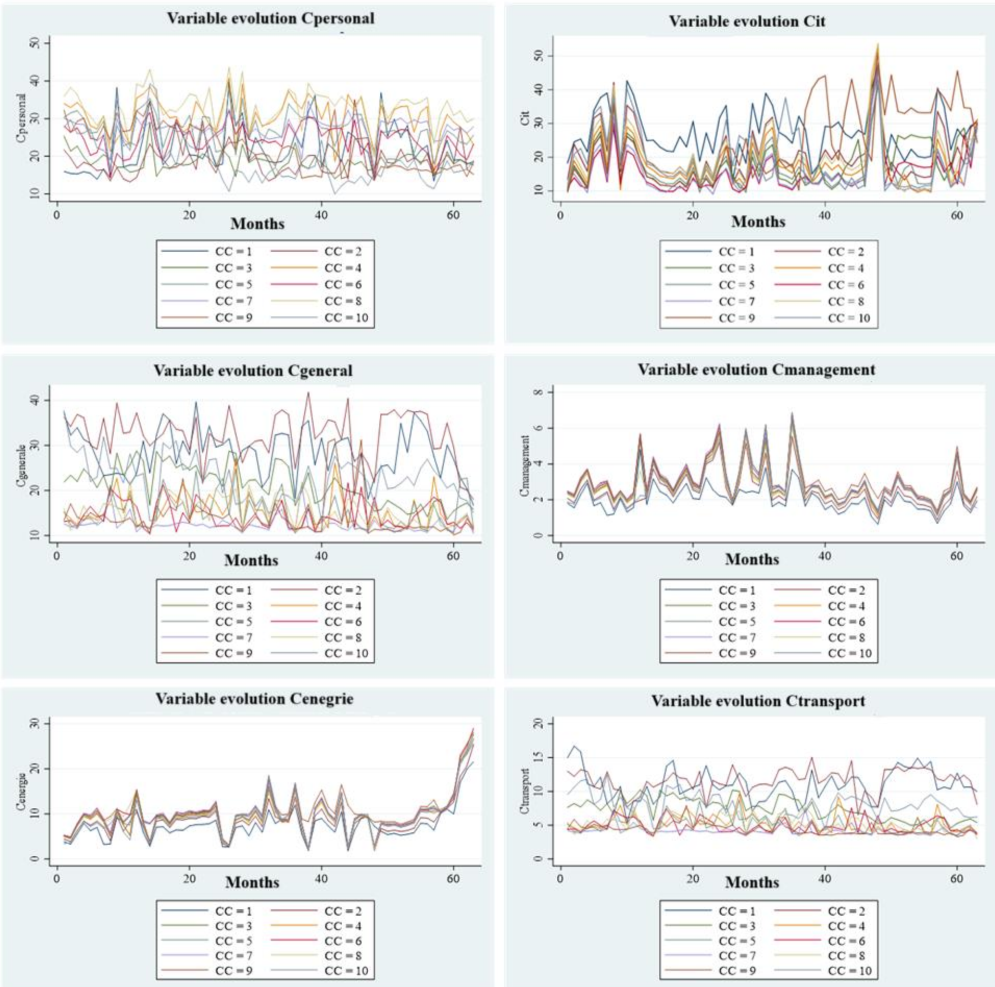

Figure 2. Evolution of independent variables.

The graphs on the evolution of independent variables in Fig. 2 are not sufficiently conclusive from this point of view. However, statistics enables us to analyze the interaction of variables more deeply to detect statistically significant but not so economically relevant influences. Statistical tests can highlight the significance of the influence of each factor or group of factors.

3.2. EFFECT OF DUMMY VARIABLES

Depending on the specifics of the panel statistics, we check for the existence of fixed/random properties caused by cost centers or caused by time. The essential variance between an FE model and an RE model involves studying the characteristics of dummy variables entered into the model. The analytical (functional) forms of the two models constitute adapted forms of (1) [17]:

$$\text{FE model: } y_{it}=(\alpha+\alpha_i)+x_{it}'\beta+\varepsilon_{it} \quad (3)$$

$$\text{RE model: } y_{it}=\alpha+x_{it}'\beta+(\alpha_i+\varepsilon_{it}) \quad (4)$$

The effect is referred to as α_i, and the FE or RE is specific to an individual or a time period and is not (and cannot be) included in the regression analysis. The errors are independent and identically distributed, i.e., $\varepsilon_{it} \sim IID(0, \sigma_\varepsilon^2)$. The FE model is used when the effects of time-varying variables are to be examined. It analyzes the relationship between the dependent variable and the independent variables within a given entity (such as a country, region, or company). Each of these units has individual characteristics that can influence the dependent variable. When using this model, we assume that certain aspects of the entity can influence both the dependent and independent variables, which requires a controlled consideration of this influence [18].

A central tenet of the FE model is the assumption that period-invariant appearances are exclusive to respectively separate characteristics and do not correlate with other characteristics. Since each unit is diverse, the error tenure and the constant that imprisons the separate appearances need not be connected. If the error relations are associated, the FE model is inappropriate, and switching to the RE model is necessary; otherwise, the conclusions may be erroneous [19].

This forms the basis of the Hausman test used in this study. The FE model analyzes individual changes in interrupts, supposing equal gradients and constant adjustment. Because the separate effect is period-invariant and preserved as a measure of the intercept, it may α_i and be correlated with other regressors without violating the exogeneity assumption. This model is usually estimated using OLS regression using dummy variables (LSDV). Different methods are used for estimating an FE model, e.g., LSDV and internal estimation methods, which do not use dummy variables. Both approaches lead to identical estimation results for the parameters of the regressors (non-dummy independent variables) [20,21].

Internal estimation applies a model that considers the individual means of the dependent and independent variables without resorting to dummy variables. The LSDV model is particularly popular because it is easy to use and interpret. However, this model proves problematic when analyzing panel data that include large numbers of individuals (or groups). If the time dimension T remains constant while the number of units $n \to \infty$, the regressor parameter estimates are consistent, the coefficients of the individual effects, $\alpha + \alpha_i$, , but it cannot be determined [22,23].

The data panel used in the empirical study meets all the criteria to be classified as a long panel. It includes ten cost centers analyzed over a 63-month period to examine the sensitivity of total costs to changes in costs regarding personnel, overhead, energy, IT, quality management, and transportation. The RE model assumes that individual heterogeneity (the individual effect) is uncorrelated with the regressors and then estimates the group- or period-specific error spread. For this reason, α_i represents an individual-specific random heterogeneity or a component of the compound error term. Therefore, the RE model is also called the error component model. In this model, the intercepts and slopes of the regressors are identical for each individual. Differences between individuals (or periods) exist in terms of their specific errors, not in their sections [17].

We examine the fixed (group) effects in more depth by introducing dummy variables into (5). The dummy variable *C1* is set to 1 for cost center 1 and to 0 for all other cost centers. Analogously, *C2* is set to 1 for cost center 2 and 0 for the remaining cost centers, and so on. Formula (5) defines the LDSV model. Nine dummy variables, *C1* to *C9*, are inserted into the OLS group model. One of the ten dummy variables in the model, in this case, *C10,* is dropped from the regression equation to prevent perfect multicollinearity. Both dummy variables and regressors can be integrated into an FE model. The coefficients γ_1–$\gamma_9$represent the regression coefficients for the dummy variables *C1* to *C9*.

$$Ctotal_i = \beta_0 + \beta_1 Cpersonal_i + \beta_2 Cgeneral_i + \beta_3 Cenergy_i + \beta_4 Cit_i + \beta_5 Cmanagement_i + \beta_6 Ctransport_i + \gamma_1 C1 + \gamma_2 C2 + \gamma_3 C3 + \gamma_4 C4 + \gamma_5 C5 + \gamma_6 C6 + \gamma_7 C7 + \gamma_8 C8 + \gamma_9 C9 + \varepsilon_i \quad (5)$$

Table 6 presents the estimates of the coefficients of the independent variables and the dummy variables from (5). The statistically and economically significant influence of three factors are noteworthy: personnel, IT, and quality management expenses. The corresponding values in the fifth column express the statistical significance of three factors (personnel, IT, and quality management expenses, (p-value < 0.05 except *Cit* (p-value < 0.064).

Table 6. OLS regression with dummy variables (LSDV model).

Cost Variable	Coef.	Std. Err.	*t*	*P*>|*t*|	95% Conf. Interval
Cpersonal	−61669.8	6348.324	−9.71	0.000	[−74,136.87, −49,202.74]
Cgeneral	−2850.118	9099.411	−0.31	0.754	[−20,719.86, 15,019.63]
Cenergy	−866.1339	5902.182	−0.15	0.883	[−12,457.05, 10,724.78]
Cit	7027.604	3792.921	1.85	0.064	[−421.0682, 14476.28]
Cmanagement	71611.35	22,631.53	3.16	0.002	[27,166.75, 116,056]
Ctransport	713.5072	25,141.51	0.03	0.977	[−48,660.28, 50,087.29]
CC2	966,172.1	110,919.8	8.71	0.000	[748,343.9, 1,184,000]
CC3	1,240,911	142,285.3	8.72	0.000	[961,486.1, 1,520,336]
CC4	832,688.8	148,773.3	5.60	0.000	[540,522.6, 1,124,855]
CC5	1,084,646	156,341.4	6.94	0.000	[777,617.2, 1,391,675]
CC6	1,217,893	165,576.4	7.36	0.000	[892,727.8, 1,543,057]
CC7	1,156,189	171,823.8	6.73	0.000	[818,755.2, 1,493,622]
CC8	1,053,533	146,189.7	7.21	0.000	[766,440.9, 1,340,626]
CC9	535,788.2	161,709.2	3.31	0.001	[218,218, 853,358.5]
CC10	549,154.9	132,708.1	4.14	0.000	[288,538.1, 809,771.6]
cons	1,680,160	407,400.5	4.12	0.000	[880,092.9, 2,480,228]
Number of obs. = 630			$F(15, 614) = 32.87$		
Prob > F = 0.0000			Adj R^2 = 0.4319		

The parameter estimates for the three regressors are different from those in the OLS group model. The coefficient of personnel expenses increased from −72,595.67 to −61,669.8, but the statistical significance remained almost unchanged (p-value < 0.05). IT and quality management expenditures became statistically significant (p-value < 0.05). The positive sign of the estimates of the two factors shows us that an increase in their values leads to an increase in total costs, which was expected. A more careful analysis of the sign of the estimates and their absolute values is carried out in the next step regarding the improvement of the econometric model.

We need to answer the following questions: How do we interpret the coefficients of the dummy variables, and what additional information does the new model bring? How do we know if there is a significant fixed group effect? The F-statistic test is helpful here. The null hypothesis (H_0) of this F-test assumes that all but one of the parameters of the dummy variables are zero. The alternative hypothesis (H_a) assumes that at least one parameter of a dummy variable is not zero [22].

The test pits the LSDV model (the robust model) against the OLS model (the more efficient model) and examines whether the introduction of dummy variables improves the panel data model. We observe the result of the F-te:

$$H_0 : \gamma_1 = \cdots = \gamma_9 = 0 \tag{6}$$

(1) CC2 = 0 (2) CC3 = 0 (3) CC4 = 0 (4) CC5 = 0 (5) CC6 = 0

(6) CC7 = 0 (7) CC8 = 0 (8) CC9 = 0 (9) CC10 = 0

$F(9, 614) = 20.41$ Prob > F = 0

If the null hypothesis is disallowed, we infer that there is an important FE or an important strengthening in model quality. The probability (prob > F = 0.0) confirms our hypothesis that at least one dummy variable has a significant impact on profit variation. As we observed, only three factors have a significant influence on the total cost, which means an improvement of the econometric model. The impact level is given by R^2. The absolute size and statistical significance of the coefficients of the cost centers show us the relevant impact on the variation of the dependent variable. All cost centers have a positive impact on the total cost variance. The magnitude of the impact is between RON 535,788 and 1,240,911. The amount of information provided by the dummy variables is significant for increasing the quality of the model.

Table 7 introduces the impact of the combined effect of the individual effects (in this case, cost centers) and the time effect (by introducing the variable *Month*). Applying the same method as in the case of short panels is difficult; using a linear or quadratic trend over time (in our case, the linear trend of *Month*) is more efficient. The impact of the variable *Month* is positive and statistically significant. Total expenses increase over time, from one month to the next, by approximately RON 8600. R^2 increases to 46%, which means a more significant impact on the variation of the dependent variable. Both effects (related to cost centers) and periods considered are significant.

When analyzing panel data, the difference between the RE model and the FE model is of great importance. If the model has a fixed effect, both the MOL and RE estimates are inconsistent, so the FE estimator is preferred. Assuming that the individual effects are random (the null hypothesis), both estimators are consistent. In the alternative hypothesis, however, the estimates differ. Hausman based his eponymous test on this idea, which is used to compare the two estimation approaches [23].

Table 7. OLS regression with dummy variables and the time variable.

Cost Variable	**Coef.**	**Std. Err.**	***t***	***P*>|*t*|**	**95% Conf. Interval**
Cpersonal	−52,844.75	6356.421	−8.31	0.000	[−65,327.76, −40,361.75]
Cgeneral	6004.124	8982.378	0.67	0.504	[−11,635.84, 23,644.09]
Cenergy	−10,936.61	5991.415	−1.83	0.068	[−22,702.8, 829.5767]
Cit	10,697.72	3743.56	2.86	0.004	[3345.964, 18,049.48]
Cmanagement	106,034.9	22,782.9	4.65	0.000	[61,292.94, 150,776.9]
Ctransport	10,327.29	24,525.65	0.42	0.674	[−37,837.2, 58,491.78]
CC2	975,307.2	107,976.4	9.03	0.000	[763,258.7, 1,187,356]
CC3	1,395,187	140,924.8	9.90	0.000	[1,118,433, 1,671,941]
CC4	957,873.1	146,345.7	6.55	0.000	[670,473.4, 1,245,273]
CC5	1,246,333	154,607.5	8.06	0.000	[942,708.6, 1,549,958]
CC6	1,415,586	164,587.9	8.60	0.000	[1,092,362, 1,738,811]
CC7	1,354,542	170,568.8	7.94	0.000	[1,019,572, 1,689,512]
CC8	1,150,702	143,238.9	8.03	0.000	[869,403.1, 1,432,000]
CC9	727,369.5	160,692.8	4.53	0.000	[411,794.4, 1,042,945]
CC10	677,407.8	130,976.6	5.17	0.000	[420,190.6, 934,625]
Month	8593.846	1451.314	5.92	0.000	[5743.696, 11,444]
cons	758,084.1	426,027.6	1.78	0.076	[−78,566.48, 1,594,735]
Number of obs. = 630			F(16, 613) = 34.72		
Prob > F = 0.0000			Adj R^2 = 0.4617		

Table 8 compares the estimated coefficients from the two models based on Hausman's test. The implementation of the test requires the prior estimation of the coefficients of the two models. Since Prob>chi2 = 0.0099, we reject the null hypothesis that we have a consistent RE model. Therefore, our model is an FE model.

Table 8. Hausman's test.

	b	*B*	*b*−*B*	sqrt(diag(*V_b*−*V_B*))
	FE	**RE**	**Difference**	**S.E.**
Cpersonal	−61,669.8	−64,347.02	2677.216	1883.988
Cgeneral	−2850.118	−4601.13	1751.012	1239.086
Cenergy	−866.1339	−1912.908	1046.774	607.6154
Cit	7027.604	3953.792	3073.813	1063.488
Cmanagement	71,611.35	71,082.77	528.5784	1393.493
Ctransport	713.5072	−15,258.65	15,972.16	6867.651
Test: H_0 = difference in coefficients not systematic				
chi2(6) = (*b*−*B*)'[(*V_b*−*V_B*)^(−1)](*b*−*B*) = 16.85				
Prob>chi2 = 0.0099				

Since T is greater than n, we need to find a model for the serial correlation of the errors. Since n is relatively small compared to T, we can relax the assumption of independence of u_{it} errors over i (for each i). We use a model in which the errors are correlated over i, an AR(1) autoregressive model regarding the autocorrelation of the errors over t, allowing for their heteroscedasticity. At the highest level of generality, the AR(1) model is as follows:

$$u_{it}= \rho_i u_{i,t-1}+ \varepsilon_{it} \tag{7}$$

where ε_{it} is an uncorrelated series but correlated over i with $Cor\ (\varepsilon_{it}, \varepsilon_{is}) = \sigma_{ts}$. The estimator used is efficient if the proposed model for error behavior is correctly specified. The estimator used in Table 9 is based on a flexible correlation model for errors between cost centers and with an AR(1) autocorrelation process of errors for each cost center.

Table 9. GLS regression with heteroskedasticity, cross-sectional correlation, and panel-specific autocorrelation AR(1).

Cost Variable	Coef.	Std. Err.	*z*	P>|z|	95% Conf. Interval
Cpersonal	−506,20.34	2931.996	−17.26	0.000	−56366.95 −44873.74
Cgeneral	4307.87	4830.293	0.89	0.372	−5159.33 13775.07
Cenergy	8205.791	6594.807	1.24	0.213	−4719.792 21131.38
Cit	−1269.969	2439.99	0.52	0.603	−6052.261 3512.323
Cmanagement	71143.12	24,008.59	2.96	0.003	24087.15 118199.1
Ctransport	1441.389	12,149.37	0.12	0.906	−22370.95 25253.72
cons	2,237,736	185,439.4	12.07	0.000	1874281 2601190
Estimated covariances = 55			Number of obs. = 630		
Estimated autocorrelations = 10			Number of groups = 10		
Estimated coefficients = 7			Time periods = 63		

The error correlation matrix for the ten cost centers forms the basis of the error independence test. Since *Pr* = 0.000 (Table 10), a correlation exists between the errors corresponding to the ten cost centers. Therefore, GLS regression with the specified characteristics was used.

Table 10. Error correlation matrix.

	e1	e2	e3	e4	e5	e6	e7	e8	e9	e10
e1	1.0000									
e2	0.0313	1.0000								
e3	0.1914	0.2227	1.0000							
e4	0.6477	0.3023	0.2752	1.0000						
e5	0.4139	0.4402	0.1715	0.7337	1.0000					
e6	0.2979	0.4047	0.1368	0.5845	0.9038	1.0000				
e7	0.3945	0.4514	0.2214	0.7824	0.9205	0.8423	1.0000			
e8	0.5582	0.3414	0.2997	0.8848	0.8380	0.7017	0.8402	1.0000		
e9	0.1661	−0.0048	0.3088	0.1136	0.0951	0.0226	0.1026	0.2138	1.0000	
e10	0.3876	0.2888	−0.1758	0.5045	0.4567	0.4231	0.4855	0.4769	−0.1348	1.0000
Breusch–Pagan LM test of independence: chi2(45) = 655.499									*Pr* = 0.0000	
Based on 63 complete observations over panel units										

The presence of heteroscedasticity in the model in Table 9 is marked by the test result from Table 11. The indicator Prob>chi2 = 0.0000 shows us that the model is heteroscedastic. Based on the two tests, we can rely on the estimates of the regression coefficients and further detail the significance of the impact factors on the total cost variation.

Table 11. Test for heteroscedasticity.

Modified Wald test for groupwise heteroskedasticity in cross-sectional time-series FGLS regression model
H_0: sigma(i)^2 = sigma^2 for all *i*
chi2 (10) = 848.25
Prob>chi2 = 0.0000

The presented econometric model represents the last theoretical form in which the analyzed behavioral specifics of the panel data were taken into account. Regression coefficients are estimated based on GLS regression accounting for heteroscedasticity, serial correlation, and error correlation between cost centers. Of the six impact factors analyzed, only personnel and quality management expenses are statistically significant, with *p*-values of less than 0.05. If the quality management expenses increase by 1%, then the total cost increases by approximately RON 71,143 on average. Similarly, we can say that with a 1% increase in personnel expenses, the total cost decreases by approximately RON 50,620. Transport and IT costs do not have a statistically significant impact on the total cost variation.

Similarly, general expenses (p-value = 0.372) and energy expenses (p-value = 0.213) do not have a statistically significant impact; however, the trend is expected (positive).

Comparing the results obtained in Table 9 with those in Tables 6 and 7 is crucial. First, the degree of impact differs significantly. If, in Tables 6 and 7, we have an impact degree of 40%, in the case of Table 9, the corresponding impact is 30%. The significant decrease in the degree of impact is due to the behavior of the model errors, which is only considered in the model in Table 9, which is more comprehensive than the one presented previously. By including the cost center–specific and time effects, the model in Tables 6 and 7 benefited from more meaningful information. The signs of the impact of the variation in personnel cost and quality management expenditure are the same. However, the behavior of the errors (their correlation and heteroscedasticity) reduced much of the impact of the other determining factors. One model for each cost center can be built:

$$y_{it} = x'_{it}\beta_i + u_{it} \tag{8}$$

The ten models, each corresponding to a cost center, allow us to estimate the impact of the six factors for each entity. Table 12 shows the parameter estimates. The necessary statistical inferences are easier to make, given that we are studying a long panel with few statistical entities and several periods.

Table 12. Unit impact of influencing factors.

CC	***Cpersonal***	***Cgeneral***	***Cenergy***	***Cit***	***Cmanagement***	***Ctransport***	***Month***	**cons**
1	−9958.02	17162.97	1722.25	15073.23	−46582.82	14919.44	1337.98	−2.35e+05
2	−15506.67	95040.44	60728.21	70532.06	154810.95	−1.02e+05	11152.51	−2.62e+06
3	−90969.63	32804.41	4219.95	17281.14	265758.47	121770.91	4652.93	860621.31
4	−29845.37	−45043.05	−5168.89	13410.87	73243.59	125374.95	8473.08	1.18e+06
5	−33631.69	−15260.21	−13844.57	45654.11	140511.20	91690.32	15517.24	503686.25
6	46329.84	−3115.90	21079.52	99573.87	206714.38	430054.41	10090.08	−4.26e+06
7	−16897.08	106457.22	−11545.91	47658.04	189392.84	−2.00e+05	12798.03	−1.83e+05
8	−25451.13	12416.19	−15314.78	24044.80	159997.06	−66850.48	8513.42	942077.50
9	−15133.06	−4057.13	−11355.83	6993.98	145265.45	−37174.09	−6330.95	1.44e+06
10	−78160.64	−30962.72	−53922.83	14572.08	−89654.51	79023.45	33574.31	2.29e+06

We compare the results obtained in Table 12 to those in Table 9. First, the impact of personnel expenses on the total cost variation is of the same sign (negative), except for cost center 6, which is oriented to production. Second, the signs of the impact of quality management expenses are the same except for production cost center 10 and maintenance cost center 1. Thirdly, the total cost registered an average monthly increase between RON 4652 and 33,574 except for production cost center 9, which has an average monthly decrease of RON 6330. The methods used to test the stationarity (unit-root test) of the variables, in the case of panel data, are based on the methods developed for a single time series and assume that $T \to \infty$ [25]. This is possible because we can extend the study to a sufficiently large number of months. Table 13 presents the results of the unit-root tests in a simplified form for each variable included in the study.

Table 13. Levin–Lin–Chu unit-root test for the seven variables.

Ho: Panels contain unit roots		Number of panels = 10	
Ha: Panels are stationary		Number of periods = 63	
Variable	**Trend**	**Statistic**	***p*-value**
Ctotal	no trend	−13.6091	0.0000
Cpersonal	no trend	−16.4279	0.0000
Cgeneral	no trend	−12.4271	0.0000
Cenergy	ft. demean	−8.8282	0.0000
Cit	no trend	−9.4931	0.0000
Cmanagement	no trend	−15.5623	0.0000
Ctransport	no trend	−8.9199	0.0000

The *p*-values of *<0.05* allow us to reject the null hypothesis (H_0) in favor of the alternative hypothesis (H_a), which means that the six variables are stationary. The test is valid for six variables that do not use the trend, less the *Cenergy* variable, where from each attached series for each cost center, the average value was subtracted to mitigate the impact of the dependencies between the series.

4. CONCLUSIONS

Cost variance refers to the changes that occur in the cost of producing a product or providing a service over a period. Cost management, on the other hand, is the process of identifying, measuring, analyzing, and controlling the costs associated with a business operation. This research analyzed the results of a listed company that has its activity organized in cost centers, presenting descriptive statistics of the independent variables of the study and the descriptive statistics of the dependent variable of the study. The purpose of cost management is to control and reduce production costs while maintaining or improving the quality of the product or service. This involves analyzing the costs of all aspects of production, including materials, labor, and overhead, using tools and techniques such as budgeting, cost-volume-profit analysis, and cost-benefit analysis. A key challenge in cost management is managing cost variance because of a variety of factors, such as changes in raw material prices, fluctuations in labor costs, and changes in market demand for the product or service in question. Meeting this challenge requires being proactive in tracking and analyzing cost trends and implementing strategies to mitigate the impact of cost variations. Internal factors are those under the organization's control, and effective cost management requires a deep understanding of these internal factors and of strategies that can be used to control and mitigate their impact.

The panel data from the company being studied were organized in a long format, and by mapping the results, we analyzed the evolution of the seven variables in 63 months for each cost center. The LSDV model outperforming the OLS group model shows that R^2 increased from 28% to 43%, which means a substantial increase in impact factors. Furthermore, three factors show a significant influence, statistically and economically: C*personal*, C*it*, and C*management*. The corresponding values express their statistical significance (*p*-value < 0.05 except C*it*, where p-value < 0.064). The parameter estimates for the three regressors are different from those in the OLS group model, but the statistical significance remained almost unchanged (*p*-value < 0.05). C*it* and C*management* became statistically significant (*p*-value < 0.05), and the positive sign of the estimates of the two factors shows us that an increase in their values leads to an increase in total costs, which was expected.

A more careful analysis of the sign of the estimates and their absolute values was carried out in the next step regarding the improvement of the econometric model by introducing the impact of the combined effect of the individual effects (the effects of the cost centers) and the time effect (by introducing the variable *Month*). The impact of the *Month* variable is positive and statistically significant. R^2 increases to 46%, which means a more significant impact on the variation of the dependent variable. Both effects considered (relating to cost centers and time periods) are significant. In the panel data analysis, we note the distinction between the RE model and the FE model (Hausman's test). Since Prob>chi2 = 0.0099, we reject the null hypothesis that we have a consistent RE model; therefore, our model is an FE model. The error correlation matrix for the ten cost centers forms the basis of the error independence test. *Pr* = 0.000 implies a correlation between the errors corresponding to the ten cost centers. Therefore, GLS regression with the specified characteristics was used. Since *T* is greater than *n*, we needed to find a model for the serial correlation of the errors.

The indicator Prob>chi2=0.0000 shows us that the model is heteroscedastic. Based on the two tests, we can rely on the estimates of the regression coefficients and further detail the significance of the impact factors on the total cost variation.

The presented econometric model represents the last theoretical form in which the analyzed behavioral specifics of the panel data were taken into account. Regression coefficients are estimated based on GLS regression accounting for heteroscedasticity, serial correlation, and error correlation between cost centers. Analyzing the unitary impact of the influencing factors, we observe that the impact of personnel expenses on the variation of the total cost is of the same sign (negative), with the exception of production cost center 6. In conclusion, the determinants that influence the variation of total costs can have a significant relationship with managerial decision-making. Managers must consider these factors in their decision-making process regarding managing costs, allocating resources, and improving efficiency.

REFERENCES

[1] Novianty, I., *South Asian Journal of Contemporary Business, Economics & Law*, **12**(1), 1, 2019.
[2] Grosu, M., *European Journal of Accounting, Finance & Business*, **16**(26), 117, 2021.
[3] Mahmud, I., Anitsal, I., Anitsal, M.M., *Global Journal of Accounting and Finance*, **2**(1), 84, 2018.
[4] Rieckhof, R., Bergmann, A., Guenther, E., *Journal of Cleaner Production*, **108**, 1262, 2015.
[5] Aničić, D., Aničić, J., Miletić, V. *Ekonomika*, **66**(1), 37, 2020.
[6] Buys, P., *Designing Cost Management Systems to Support Business Decision-Making*. Springer, Singapore, 2021.
[7] Tagashira, T., Minami, C., *Journal of Interactive Marketing*, **47**(1), 68, 2019.
[8] Zhu, C., Ji, Y., Du, X., Kong, F., Chi, Z., Zhao, Y., *Science of The Total Environment*, **852**, 158515, 2022.
[9] Gu, L., Zeng, D., Guo, S., Barnawi, A., Xiang, Y., *IEEE Transactions on Emerging Topics in Computing*, **5**(1), 108, 2015.
[10] O'Connell, V., *International Journal of Managerial Finance*, **3**(4), 372, 2007.
[11] Hassan, M.K., Sanchez, B., Yu, J.S., *The Quarterly Review of Economics and Finance*, **51**(1), 88, 2011.
[12] Kilmer, J.T., Rodríguez, R.L., *Journal of evolutionary biology*, **30**(1), 4, 2017.
[13] Bai, J., Choi, S.H., Liao, Y., *Empirical Economics*, **60**, 309, 2021.
[14] Miller, S., Startz, R., *Economics Letters*, **175**, 28, 2019.
[15] Griffith, E.J., Sukhbold, T., Weinberg, D.H., Johnson, J.A., Johnson, J.W., Vincenzo, F., *The Astrophysical Journal*, **921**(1), 73, 2021.
[16] Muhammad, S., Shamshuddeen, S., Baoku, I.G. *Fudma Journal of Sciences*, **5**(3), 93, 2021.
[17] Zulfikar, R. *Estimation model and selection method of panel data regression: an overview of common effect, fixed effect, and random effect model*. INA-Rxiv Papers https://doi.org/10.31227/osf.io/9qe2b, 2018.
[18] Mills, T.C. *Applied time series analysis: A practical guide to modeling and forecasting*. Academic Press, London, 2019.
[19] Pole, A., West, M., Harrison, J. *Applied Bayesian forecasting and time series analysis*, Chapman and Hall/CRC Press, London, 2018.
[20] Okeke, J.U., Okeke, E.N., *European Journal of Engineering and Technology Research*, **1**(6), 77, 2016.
[21] Hashimzade, N., Thornton, M.A. *Handbook of Research Methods and Applications in Empirical Microeconomics*, Edward Elgar Publishing, 2021.
[22] Baltagi, B.H., *Econometric analysis of panel data*, Vol. 4, Wiley, Chichester, 2008.
[23] Baltagi, B.H., Liu, L., *Econometric Reviews*, **35**(4), 638, 2016.