Scholar Publishing Group
Socio-Economic Statistics Research
https://doi.org/10.38007/SESR.2026.070204
ISSN 2790-2722 Vol. 7, Issue 2: 27-37

Research on Corporate Expense Allocation and Budget Execution Variance Analysis Based on Multi-Dimensional Operational Indicators

Siyao Li

Finance Department, OMD USA, New York, 10001, United States

Keywords: Cost Allocation; Budget Execution; Variance Analysis; Operating Indicators; Entropy Weight Method; Rolling Budget

Abstract: In the current era of digitalisation and changes in the business model, the single-revenue-ratio method can no longer accurately reflect the distribution of shared expenses, and fixed budgets are also likely to be out of date due to alterations in business scale, staffing structure and market strategy. This paper builds an all-encompassing model of "expense allocation - rolling budget - deviation diagnosis": First, entropy-weighted allocation factors are derived from revenue, gross profit, direct operating expenses and growth intensity; second, rolling budget benchmarks for expense categories are set according to driving factors such as service revenue, functional staff size and advertising investment; finally, anomalies are identified based on a deviation rate and importance threshold. Acquire publicly available annual report data for Microsoft from 2023 to 2025 and validate the model. Based on the above analysis, the multi-dimensional model has shifted general administrative expenses towards business units with higher growth rates and gross profit contributions; the average absolute budget deviation for the four expense categories in 2025 is 3.62%, a reduction of 56.69% from the 8.36% benchmark of single-revenue-growth areas. Therefore, this study suggests that the allocation of expenses should be changed from "scale average" to "equal emphasis on resource consumption and benefits", and the evaluation of budget execution should distinguish between efficiency improvements, delayed investment, and actual overspending to enhance the interpretability of responsibility centre assessment and the quality of resource allocation.

1 Introduction

The goal of the company's budget is to set an upper limit on total spending and to ensure that the money will be used in accordance with the strategy and goals for that period. As the scale of platform-based operations and cloud services and digital product revenues increases, expenses for shared technology platforms, data centres, brand marketing, human resources and compliance management are difficult to directly attribute to any single product or department. Although the previous mode of cost allocation was based on a single factor, such as revenue, headcount or working hours, it had defects; for instance, high-revenue departments bore a disproportionately

large share of the costs, growth businesses were overlooked, and support departments passively absorbed costs, thus failing to reflect the actual profit of products and departmental performance and investment priorities.

Budget execution deviations are also multi-causal. Actual costs exceeding the budget may be due to increased sales, changes in the structure of services, expansion of infrastructure or price adjustments; actual costs falling short of the budget may be due to efficiency improvements, project delays, staff vacancies or reduced necessary investments. Therefore, only by combining financial results with the volume of operating activities, resource consumption and the degree of benefit can we determine whether the deviation is due to controllable waste, reasonable flexibility or strategic investment. A broad focus of the senior management on the budget function will change the role of the budget in planning, coordination, resource allocation and performance assessment significantly [1]. Therefore, the scope of budget analysis should no longer be limited to a simple comparison of "budget versus actual", but a closed-loop system of operation-driven dynamic updates and accountability should be established.

The two problems in this paper are the unfair division of joint living expenses and the lack of reasons for budget overruns. It puts forward an all-encompassing, multiple-level operational indicator-driven analysis system. Selects publicly available operating data from Microsoft for the period 2023-2025 as the case study; its annual reports contain indicators such as expense categories, business segments, personnel structure and service revenue, etc., which can be used to test the model without access to confidential internal budgets. The "budget" in this paper refers to an auditable rolling baseline constructed from open data, and it is not an inference based on the company's own budget figures. The research provides the following contributions: First, it combines scale, profitability, input and growth indicators using entropy weighting to enhance the causal consistency of shared expense allocation; second, it sets differentiated driver combinations for different expense categories to avoid budget mismatch caused by uniform growth rates; and third, it demonstrates through a real-world case that even large favourable deviations require operational explanations.

2. Analysis of the Current Situation of the Research Topic

2.1 Budget Management is Changing from Fixed Control to Strategic Synergy.

Recently, many scholars have proposed that budget supervision and long-term plans should work together. Small and medium-sized enterprises in Australia have found through continuous budget monitoring that there are new development prospects for their company and better utilisation of funds [2]. The above studies show that budgets are not only used as an ex-post control device but also serve the functions of information integration and expectation management. For companies that are changing their business model, a single-year fixed budget will not be practical for a long time after implementation.

Longitudinal studies of manufacturing enterprises have further shown that the enterprises have not generally abandoned budgets but have improved the success of budgets by reforming budget types, shortening the update cycle, and enhancing operational participation [3]. Therefore, the goal of budget optimisation is not to disregard the old budget, but rather to separate the functions of fixed targets, flexible forecasts and performance evaluation; fixed targets are for strategic commitment, rolling forecasts are for resource adjustment, and deviation analysis is for identifying responsibilities.

Budget control also has some behaviour. According to a large Swedish organisation, when managers have psychological safety, information openness and a 'safe space' for error, the budget

control will likely be viewed as 'supportive' rather than 'punitive' [4]. Therefore, in explaining deviations, responsibility centres should be allowed to present factors such as business volume, price, structure and timing, and not all adverse differences should be attributed to management failure.

2.2 The cost system is changing from a single-basis allocation to a multi-driven and refined type.

Theoretically, the division of costs should be based on cost drivers and benefits. Cognitive patterns of management accountants and the culture of organisational innovation will affect the complexity, functions and continuous improvement of the cost system [5]. In the digital business environment, the cost objects have expanded from products to customers, channels, projects, platforms and data resources, and a single indicator is no longer sufficient to cover the formation mechanism of all types of costs.

Recently, due to changes in the economic environment, business scale and scope, etc., more complex cost systems have been introduced by enterprises [6]. Not all complex systems are better. If there are too many indicators, the weights lack a basis, and they cannot be explained, the model will likely increase public opposition. Therefore, the cost allocation model should meet the requirements of "a few key drivers, verifiable data, stable calculation rules and allowance for periodic calibration".

Research on management control systems has also shown that financial control and non-financial control have different effects on corporate behaviour and capital structure, and the financial literacy of managers will affect the utilisation efficiency of control information [7]. Therefore, the foundation for cost allocation and budget assessment should be a combination of financial and operating indicators; that is to say, financial indicators show the results and profitability, and operating indicators explain resource consumption and business activities.

2.3 Digital Forecasting Improves Efficiency but Demands Interpretable Constraints

Artificial intelligence is changing the way management accounting uses data and forecasts for decision-making by modifying the automation of budgeting, pricing and performance analysis, etc. [8] Budgeting by the company also covers how to use its funds, what the company will undertake and why these activities are selected. Only aiming for forecast accuracy may result in models that use uncontrollable or lagged variables and are thus difficult to implement in practice.

Machine learning can be employed in business forecasting to reduce the workload of rolling forecasts manually and provide a relatively objective benchmark for sales, inventory and resource demand [9]. However, if there has been a sudden change in the business structure or a merger and acquisition recently, the complex algorithm may not be better than the transparent driving model over a short period. Therefore, this paper will use an explainable weighted driving method and avoid a black-box forecast that is difficult to audit.

Research on explainable AI in finance shows that the main decisions should be accompanied by disclosures of the contributing factors, model rationale and results' sensitivity [10]. The same applies to the budget model for corporate expenses: it should be possible to answer whether the questions "which operating indicators led to the budget change", "why were expenses allocated to this responsibility centre", and "whether the deviation is controllable" can be addressed [11].

3. Problem Statement and Research Design

3.1 Management problems that need to be solved

First, the allocation basis is not in line with the causes of expenses. The revenue ratio method links payment capacity to resource consumption, and the headcount ratio method neglects business expansion and platform benefits, thereby causing an organised overestimation or underestimation of the profit in responsibility centres. Secondly, the budget target does not align with the changed circumstances [12]. Adjusting all expenses uniformly with the growth in revenue will overestimate back-office management and sales expenses and underestimate cloud service delivery costs. Third, the direction of deviation is mistaken for the level of performance. Falling short of the budget does not necessarily mean poor performance, and going over budget is not necessarily a waste. Fourth, there is no management of the indicators. If the source, update time and responsible party are unknown, the multidimensional model will generate a new risk of "accuracy but unauditable" [13].

3.2 Data Sources and Indicator System

The 2023, 2024, and 2025 annual reports of Microsoft are the data in this paper. Expense categories are operating costs, research and development expenses, sales and marketing expenses, and general and administrative expenses; operating indicators include total revenue, service and other revenue, functional staff size, advertising expenditure, segment revenue, gross profit, direct operating expenses and their growth rate. General and administrative expenses in 2025 will be treated as a shared expense pool for multi-dimensional allocation; the actual expenses in 2024 will serve as the starting point for the rolling budget, and changes in operating indicators in 2025 will be used to derive the budget baseline.

Table 1. Multidimensional Operating Indicator System and Management Logic

Dimension	Indicator	Symbol	Management rationale	Model use
Scale	Revenue	REV	overall benefit and business volume	Allocation and budget
Benefit capacity	Gross margin	GM	Capacity to absorb shared resources	Allocation
Resource intensity	Direct operating expense	DOE	Observable operating input	Allocation
Dynamics	Revenue growth	GROW	Incremental management complexity	Allocation
Digital activity	Service and other revenue	SERV	Cloud and service delivery intensity	Budget
Human resources	Functional headcount	HC	Payroll and support workload	Budget
Market activity	Advertising spend	ADV	Campaign and acquisition intensity	Budget

Table 1 Analysis: The four kinds of cost drivers in the indicator system are scale, benefit capacity, resource input and business dynamics. Revenue and gross profit are used to show the

extent of benefit from shared resources, direct operating expenses and personnel size reflect resource consumption, and growth rate and service revenue capture the incremental complexity introduced by new business development. Use "allocated indicators" and "budget indicators" separately to prevent the same indicator from having too much explanatory weight.

Table 2 Public Financial and Operating Indicator Data

Metric	2023	2024	2025	Unit
Revenue	211,915	245,122	281,724	USD million
Service and other revenue	147,216	180,349	217,778	USD million
Cost of revenue	65,863	74,114	87,831	USD million
R&D expense	27,195	29,510	32,488	USD million
Sales & marketing expense	22,759	24,456	25,654	USD million
General & administrative expenses	7,575	7,609	7,223	USD million
Total employees	221,000	228,000	228,000	Persons
R&D headcount	72,000	81,000	80,000	Persons
Sales & Marketing Headcount	45,000	45,000	44,000	Persons
G&A headcount	15,000	16,000	15,000	Persons
Advertising spend	904	1,700	2,100	USD million

Data source: Microsoft Corporation Annual Reports 2023-2025; calculated by the authors.

As shown in Table 2, revenue and service revenue continued to increase from 2023 to 2025, and the total number of employees remained the same in 2024 and 2025; however, the number of R&D, sales and management positions slightly decreased. Thus, the increase in expenses cannot be attributed solely to an increase in personnel or revenue. Advertising expenses rose by 23.53 per cent, but sales and marketing expenses increased only 4.90 per cent; therefore, the reasons for the results must be that the expense structure, channel efficiency and M&A integration all performed well.

4. Multidimensional Indicator Cost Allocation and Budget Deviation Model

4.1 Indicator Normalisation and Objective Weighting

Normalise the original value of business unit j for indicator i to a proportion so that indicators with different dimensions can be compared in the same system:

$$p_{ij} = \frac{x_{ij}}{\sum_{j=1}^{n} x_{ij}} \tag{1}$$

x_ij is the original value of the i-th indicator in business unit j, p_ij is the normalised proportion of that indicator, and n is the number of business units. As all the indicators in this paper are positively correlated, a larger value can be seen as a greater benefit of shared resources or higher management complexity; therefore, an inverse relationship does not need to be taken into account.

To reduce the bias of human-assigned weights, entropy weighting can be used to set a weight for each indicator based on how different the business units are:

$$e_i=-\frac{1}{\ln n}\sum_{j=1}^{n} p_{ij}\ln p_{ij},\quad w_i=\frac{1-e_i}{\sum_{i=1}^{m}(1-e_i)} \tag{2}$$

The formula is as follows: e_i is the indicator entropy value, 1-e_i is the difference coefficient, w_i is the final weight, and m is the number of indicators. If the distribution of the indicators is uneven, they can differentiate the business units better and be assigned higher weights. The weights of revenue, gross profit, direct operating expenses and growth intensity in the case study were 18.20%, 37.19%, 11.63% and 32.98%, respectively.

4.2 Sharing of Shared Costs

Obtain the total driving scores for each business unit, and then allocate the shared cost pool based on the overall score as follows:

$$A_j=C^S\sum_{i=1}^{m} w_i p_{ij} \tag{3}$$

C^S is the total amount of shared expenses to be distributed according to the formula, and A_j is the expense of business unit j. Given both the current benefit capacity and the demand for future growth, the model can also be applied to expenses that are difficult to trace directly, such as legal and financial fees, personnel costs, information-technology expenses, shared facility fees, and enterprise-level customer support costs. For expenses that can be directly attributed, they should still be included directly first, and the allocation pool should not be expanded in the pursuit of model uniformity.

4.3 Multi-drive Rolling Budget Baseline

The Drivers of Different Expense Categories Differ. The actual expenses in 2024 are taken as the starting point, and a rolling budget is built based on the relative changes in operating indicators in 2025:

$$B_{k,t}=A_{k,t-1}\sum_{r=1}^{q}\beta_{kr}\frac{D_{r,t}}{D_{r,t-1}},\quad \sum_{r=1}^{q}\beta_{kr}=1 \tag{4}$$

B_k,t is the budget baseline for expense category k in period t, A_k,t-1 is the actual expense in the previous period, D_r,t is the r-th operating driver, and β_kr is the association weight between the expense category and the driver in the formula. Operating expenses are mainly composed of revenue, service revenue and operating personnel; research and development expenses are driven by service revenue, R&D personnel and revenue; sales and marketing expenses are driven by revenue, sales staff and advertising investment; and general and administrative expenses are driven by the total number of employees, management staff and revenue. The total weight is 1, and thus the budget factor can be regarded as a weighted combination of the driving rates of change.

4.4 Deviation Rate and Importance Judgment

Budget execution deviation rate =

$$V_{k,t}=\frac{A_{k,t}-B_{k,t}}{B_{k,t}}\times 100\% \tag{5}$$

V_{k,t} in the formula is the cost deviation rate. A positive value indicates that the actual cost has exceeded the benchmark; otherwise, it is a negative number and thus the actual cost is lower than the benchmark. A normal threshold of 3% and an importance threshold of 5% have been set in this paper: an absolute deviation of no more than 3% is considered a normal fluctuation; a range of

3%-5% is considered to have attracted attention; and deviations exceeding 5% require an explanation. The threshold is only for filtering; the final selection will still be based on the scale of business, price, project time and strategic value.

5. Empirical Results and Solutions

5.1 Characteristics of the Actual Changes in Cost Structure

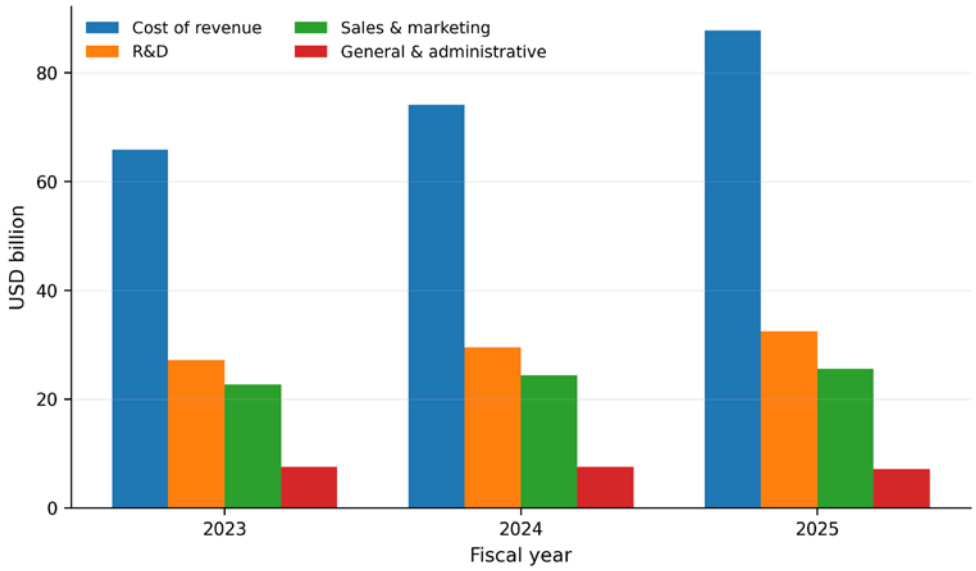

Figure 1. Changes in the Main Expense Items from 2023 to 2025

Data source: Microsoft Corporation Annual Reports 2023-2025; calculated by the authors.

Figure 1 Analysis: Operating expenses increased from US$65.863 billion in 2023 to US$87.831 billion in 2025; this rise was relatively more substantial than that of general and administrative expenses, R&D expenses continued to rise, and the growth speed of sales and marketing expenses gradually declined. According to the structure of expenses, the increase in cloud services and artificial intelligence infrastructure has led to higher delivery and R&D expenses; at the same time, back-office management expenses have fallen as revenue has increased significantly, and economies of scale and organisational reduction have both occurred.

5.2 Multidimensional Allocation Results

Using the $7.223 billion in general and administrative expenses for 2025 as the shared expense pool, the model allocates 42.23% ($305 million), 40.26% ($291 million) and 17.51% ($127 million) to Productivity and Business Processes, Intelligent Cloud, and More Personal Computing, respectively. Compared with the simple revenue-sharing mode, Intelligent Cloud has increased its share by $18 million and More Personal Computing has reduced its share by $14 million. The increase is due to a high-growth cloud business and a high-gross-margin productivity business; therefore, the multi-dimensional model does not show that revenue size is the sole determinant of the allocation results.

Therefore, the division of joint-use expenses should consider not only the scale of each department but also the speed at which they are developing new businesses and requiring shared platforms and management tools. However, if too much weight is given to the growth indicators, it

may place an undue financial burden on the new-business department. Therefore, the company will set weight caps, continuity constraints and annual calibration mechanisms, and perform manual reviews of substantial organisational changes, mergers and acquisitions or alterations in accounting standards.

5.3 Budget Execution Deviation Identification

Table 3.2025 Multi-driven Budget Baseline and Implementation Deviation (USD million)

Expense category	Driver-based budget	Actual	Amount variance	Variance rate	Assessment
Cost of revenue	85,627.0	87,831.0	+2,204.0	+2.57%	Unfavorable / Normal
R&D	32,781.2	32,488.0	-293.2	-0.89%	Favorable / Normal
Sales & Marketing	27,060.0	25,654.0	-1,406.0	-5.20%	Favorable / Material
General & administrative	7,669.8	7,223.0	-446.8	-5.83%	Favorable / Material

Table 3 Analysis: Operating costs are 2.57% higher than the model benchmark and within the normal fluctuation range; R&D expenses are generally in line with the benchmark. Sales and marketing expenses and general administrative expenses are 5.20% and 5.83% lower than the benchmark, respectively; these are relatively large favourable deviations. The above result is not to be considered a saving. Channel efficiency, staff reduction, project delays and functional service level need to be further studied to determine whether the required reduction in inputs will harm the short-term budget objectives.

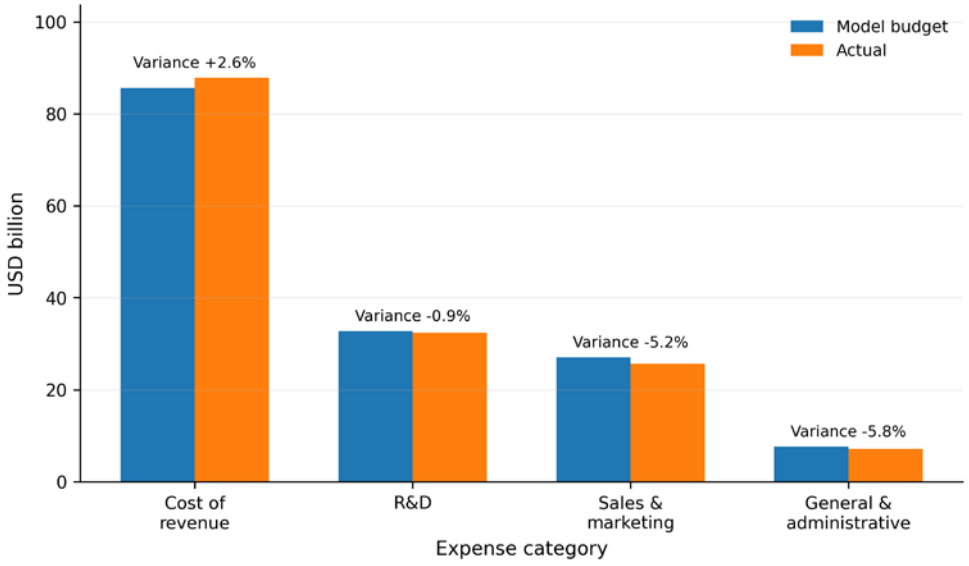

Figure 2. Comparison of Model Budget and Actual Costs in 2025

Data sources: Microsoft Corporation's 2023-2025 Annual Reports; calculated by the authors.

Figure 2 Analysis: The total amounts of the model budget and actual expenses are relatively close, but there are considerable structural differences among the categories. The average absolute deviation of the four categories of expenses is 3.62%, which is significantly lower than the 8.36% obtained by the uniform revenue growth method; that is, an error reduction of 56.69% has been achieved. Among them, the negative deviation in sales and administrative expenses is the largest; thus, the multi-dimensional model has indicated a structural change in the budget execution that needs to be corrected by management.

5.4 Management Strategies and Implementation Path

Set up a hierarchical rule for the cost pool first. Direct costs should be directly associated with contracts, work orders, projects or asset tags; traceable shared costs should be allocated according to the volume of activity; only enterprise-level costs that are difficult to trace should be allocated via multi-dimensional comprehensive allocation. The order of allocation should be "direct attribution first, causal factors second, and benefit capacity supplementation", and all costs should not be included in a uniform model.

Second, build a dictionary of operating indicators. All indicators need to have a clear business purpose, a source of data, an update time, an accountable person, and exception-handling rules. Personnel indicators should distinguish between working staff, average staff and full-time equivalent; revenue indicators should separate signed contracts, confirmed payments and received payments; advertising expenditure should be divided into brand-building and performance-driven advertising, etc. Any modifications to the definition of a metric should be accompanied by version information and recalculation records.

Third, set up a three-tier budget system of fixed goals, rolling forecasts and flexible evaluations. Set the general allocation of funds and purposes for each year; adjust requirements for funds and other resources several times a year according to changes in business circumstances; evaluate performance based on factors that can be controlled. Infrastructure, R&D and market-acquisition costs can be added to a special protected zone to prevent the responsibility centre from delaying the necessary investment due to a short-term budget.

Fourth, build a deviation decomposition tree. All major deviations should be broken down into at least five categories of reasons: volume, price, structure, efficiency and timing, and within these, controllable and uncontrollable factors need to be identified. Positive deviations need to be identified as either the result of growth and strategic investment, and negative deviations need to be determined to ascertain whether they are caused by efficiency improvements, lower procurement prices, staff vacancies or project delays. Only after completing the cause code is the deviation data used for weight calibration in the next cycle.

Fifthly, enhance model interpretability and governance. The finance department is responsible for the model framework and definitions; the business department is responsible for driving data and explanations of the reasons; the data department is responsible for interfaces, permissions and quality monitoring; and the audit or internal control department regularly reviews weights, thresholds and adjustments for anomalies. Show the details of the contributions and sensitivities of indicators in the model results, not only the final allocated amounts.

Sixth, use a staged release model. Select shared services or marketing expenses with a clear causal relationship between costs and high data quality during the pilot period, and after stabilisation, extend this to technology platforms, data centres and enterprise functions. The deployment assessment should also consider budget variations, the number of disputes, modification counts, settlement durations and business acceptance to prevent increased organisational costs due to higher model accuracy.

6. Summary

The problems of simple division of shared expenses and lack of explanations for budget execution deviations are addressed by establishing a multi-dimensional operational indicator-driven framework for expense allocation and rolling budget analysis in this paper. Based on the above empirical results, the entropy weight model can include gross profit, direct operating inputs and growth intensity in addition to the scale of revenue, thus making shared expense allocation more aligned with the actual benefits and management complexity. A separate budget-driven analysis can be conducted to find the source of abnormal fluctuations in operating expenses, R&D and sales-and-administration expenses, etc. The multi-dimensional model has reduced the average absolute deviation of the categories by 56.69% compared with the uniform revenue growth benchmark, and this transparent and auditable driving model has shown good practical results in a short sample.

Management implications of this study are to view cost allocation and budget deviation as interconnected problems rather than independent ones. Allocation sets the cost limits of responsibility centres; Budgets specify the resource demands within those centres; and deviation analysis determines if the allocation factors and budget weights are reasonable. Establish a closed-loop system of "indicator governance—cost collection—driving allocation—rolling budget—root cause analysis—parameter calibration" in the enterprise, and change deviation assessment from simple rewards and punishments to resource allocation and business learning.

This paper still has deficiencies. The case study only uses the public annual data, and monthly business volume, procurement prices, internal project progress and actual budget approval figures cannot be obtained. Thus, the model will only confirm the structural explanatory power of the enterprise and not rebuild internal budgets. Future research may introduce time-driven costing, causal inference and interpretable machine learning on enterprise-level detailed data, and compare the implementation effects under different weight stability, anomaly threshold and organisational context.

References

[1] *Kihn, L.-A., & Ström, E. (2025). Top management characteristics and comprehensive focus on budgeting. Journal of Applied Accounting Research, 26(6), 223–243. https://doi.org/10.1108/JAAR-02-2023-0036.*

[2] *Rice, J., Martin, N., Raziq, MM, Memon, MA, & Fieger, P. (2024). Strategic planning, budget monitoring and growth optimism: Evidence from Australian SMEs. European Business Review, 36(6), 962–980. https://doi.org/10.1108/EBR-09-2023-0285.*

[3] *Kihn, L.-A. (2023). Budgeting changes and success: Abandonment, reform or stability in Finnish manufacturing firms? Journal of Accounting & Organizational Change, 19(6), 91–111. https://doi.org/10.1108/JAOC-12-2022-0187.*

[4] *Johansson-Berg, T., & Wennblom, G. (2023). If managers feel safe, budget control becomes enabling: Evidence from a large local government organization in Sweden. Journal of Public Budgeting, Accounting & Financial Management, 35(6), 154–179. https://doi.org/10.1108/JPBAFM-01-2023-0001.*

[5] *Pavlatos, O., & Ioakimidis, M. (2024). Management accountants with a growth mindset and changes in the design of costing systems: The role of organizational culture. Accounting & Finance, 64(2), 2085–2110. https://doi.org/10.1111/acfi.13209.*

[6] *Diavastis, IE (2025). What drives cost system sophistication? Empirical evidence from the Greek hotel industry. Journal of Risk and Financial Management, 18(7), 401. https://doi.org/10.3390/jrfm18070401.*

[7] Graña-Alvarez, R., Gomez-Conde, J., Lopez-Valeiras, E., & González-Loureiro, M. (2024). Management control systems, business financial literacy and financial leverage in business-incubated start-ups. The British Accounting Review, 56(6), 101427. https://doi.org/10.1016/j.bar.2024.101427.

[8] Liu, B. (2026). Research On a Framework for Generating and Maintaining Automated Drawline Syntax Trees for Complex Programming Languages. Procedia Computer Science, 279, 420-428.

[9] Wu, L. (2026). Construction and Evolutionary Analysis of a Game Model for Supply Chain Finance Funding Based on Blockchain Technology. Procedia Computer Science, 282, 2004-2012.

[10] Ding, J. (2025, December). Research on Logistics Cost Control and Optimization in Automotive Manufacturing Supply Chain Based on DMAIC Model. In 2025 IEEE 1st International Conference on Recent Trends in Computing and Smart Mobility (RCSM) (pp. 1-7). IEEE.

[11] Sun, J. (2026). Automated Feature Engineering and Screening System for Large-Scale Factor Libraries. Procedia Computer Science, 281, 1282-1290.

[12] Hou, Y. (2026). Collaborative Regulation for Stable Data Center Operation under Energy Efficiency Constraints. Procedia Computer Science, 281, 612-620.

[13] Zhang, Z. (2026). Research on Performance Optimization Methods for Resource-Aware Model Services in AI Systems. Procedia Computer Science, 281, 1310-1317.