“Modelling budgetary decision utility and institutional risk: Evidence from the Portuguese Navy”

AUTHORS	Ricardo de Moraes e Soares
ARTICLE INFO	Ricardo de Moraes e Soares (2026). Modelling budgetary decision utility and institutional risk: Evidence from the Portuguese Navy. *Public and Municipal Finance*, *15*(4), 14–31. doi:10.21511/pmf.15(4).2026.02
DOI	http://dx.doi.org/10.21511/pmf.15(4).2026.02
RELEASED ON	Thursday, 24 September 2026
RECEIVED ON	Wednesday, 10 June 2026
ACCEPTED ON	Wednesday, 02 September 2026

JOURNAL	"Public and Municipal Finance"
ISSN PRINT	2222-1867
ISSN ONLINE	2222-1875
PUBLISHER	LLC “Consulting Publishing Company “Business Perspectives”
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
50	6	4

BUSINESS PERSPECTIVES

LLC "CPC "Business Perspectives"
Hryhorii Skovoroda lane, 10,
Sumy, 40022, Ukraine
www.businessperspectives.org

Ricardo de Moraes e Soares (Portugal)

MODELLING BUDGETARY DECISION UTILITY AND INSTITUTIONAL RISK: EVIDENCE FROM THE PORTUGUESE NAVY

Abstract

Budgetary decision-making in the defense sector involves making trade-offs between interdependent capabilities, subject to financial, operational and institutional constraints. The analysis of aggregate expenditure is insufficient to evaluate these decisions, as the impact of the allocated resources depends on their composition and the interplay between categories of expenditure. The study analyzes the budgetary headings of the Portuguese Navy between 2023 and 2025 and assesses their compatibility with an institutional utility function. To this end, possible adjusted structural model has been proposed, which includes the complementarity between maintenance and investment, the adequacy of personnel expenditure, potential performance and institutional stability. The robustness of the optimization was assessed through 200,000 Monte Carlo simulations, with simultaneous variation of parameters, thresholds and weights. Total expenditure increased by 20.66 per cent over the period, but the highest utility was achieved in 2024 (0.691554), followed by 2025 (0.683443) and 2023 (0.659761). The 2024 financial year ranked first in 91.64% of the simulations and outperformed 2023 in all configurations analyzed. However, alternative normalization of the maintenance-investment component altered the ranking to 2023 > 2024 > 2025, thereby distinguishing the model's parametric robustness from its structural robustness. The model enables to explain budgetary commitments and test the stability of decisions, but it does not constitute a causal measure of operational readiness, efficiency or institutional risk. The model's relevance lies in the transparent support it provides for comparing the various allocation alternatives.

Keywords naval administration, efficiency, modelling, budgetary optimization, operational readiness, risk, utility

JEL Classification C61, D81, H61, H56

Type of the article: Research Article

Received on: 10th of June, 2026
Accepted on: 2nd of September, 2026
Published on: 24th of September, 2026

Ricardo de Moraes e Soares, Ph.D. in Public Administration, Professor, Department of Humanities and Management, Naval School, Military University Institute, Portugal.

Conflict of interest statement:
Author(s) reported no conflict of interest

INTRODUCTION

The allocation of financial resources within military organizations constitutes a public decision-making problem characterized by budgetary constraints, multi-annual commitments and operational requirements with different time horizons. The budgetary dimension does not, in itself, define institutional capacity, as the effects of expenditure depend on the functional composition of the links between budget headings and the organizational capacity to transform resources into usable capabilities.

Within naval structures, this relationship is particularly demanding. The acquisition and modernization of assets increase capacity but also generate additional requirements for maintenance, technical support, infrastructure and training. Conversely, an excessive focus on maintenance may limit technological renewal. Expenditure on personnel introduces further tension between institutional continuity and budgetary rigidity.

The scientific problem concerns the integrated representation of these relationships without confusing formal assumptions with observed

http://dx.doi.org/10.21511/pmf.15(4).2026.02

evidence. The calibrated utility function allows for the comparison of configurations and the explanation of trade-offs, but the results depend on the selected functional forms, weights, benchmarks and thresholds. Research must therefore assess the internal consistency of the models and the stability of the conclusions in the face of parametric changes by distinguishing between normative analysis, empirical application and institutional validation.

1. LITERATURE REVIEW AND HYPOTHESES

Financial decision-making in the defense sector is not merely a matter of accounting for the allocation of a budget. The budget reflects choices between competing capabilities, subject to financial, technological, legal and operational constraints, the effects of which materialize over different time horizons (Schick, 1966). Personnel expenditure ensures continuity, expertise and availability but offers little flexibility. Maintenance aims to keep existing assets in operational condition, minimizing breakdowns, downtime and loss of capability. Investment funds the acquisition of new resources, the replacement of obsolete assets and technological modernization. Day-to-day operations support the allocation of capacities through services, consumables, energy, communications, transport and other resources necessary for the activity. These categories perform different but interdependent functions: investment generates potential capacity, maintenance preserves it, and day-to-day operations enable operational mobilization.

Accounting information is essential for identifying the allocation of resources, monitoring implementation and underpinning accountability (Lopes & Rosado, 2018). However, financial figures describe resources that have been authorized, committed or allocated; on their own, they do not indicate how these have been transformed into institutional outcomes. A higher budget allocation may coexist with contractual delays, low implementation rates, inadequate maintenance or the unavailability of resources. Consequently, there is no automatic equivalence between expenditure, capacity and performance (Pollitt & Bouckaert, 2017). This distinction is particularly important in public sector organizations subject to incrementalism (Wildavsky, 1984), regulatory rigidity, multi-annual commitments, and dependence on past decisions, rules and institutional routines (Jones & Baumgartner, 2005; March, 1994; Scott, 2014).

The utility theory provides a systematic way of ranking alternative configurations when the criteria and preferences are sufficiently specified (von Neumann & Morgenstern, 1944; Savage, 1972). However, utility function does not measure a specific physical property of an organization. The result reflects a ranking conditioned by the outcomes included, the substitution rates allowed and the parameters adopted. In multi-criteria problems, the ranking depends on how the various objectives are aggregated and balanced (Belton & Stewart, 2002; Keeney & Raiffa, 1993). When the weights do not stem from revealed choices, institutional decisions or formal procedures, the model takes on a normative nature, and the alternatives are assessed according to a value-based logic, which may not directly correspond to the organization’s actual preferences.

The distinction between formal rationality and effective organizational decision-making constitutes a limitation. Organizations do not have complete information, unlimited computational capacity or the freedom to fully reconsider the choices made in each period (Simon, 1957). Choices are constrained by information costs (Conlisk, 1996), internal routines and coalitions (Cyert & March, 1963), administrative rules and feasible alternatives (Lindblom, 1959). The fact that a given budget line item is close to the maximum of a calibrated function does not allow the decision to be characterized as rational or efficient. The result may stem from the mathematical specification itself, previously assumed obligations and/or constraints that the model does not represent (Battaglio et al., 2019).

Behavioral economics reinforces this reservation by pointing out that decisions, in a context of uncertainty, depend, in particular, on reference points and the asymmetric evaluation of gains and losses (Kahneman & Tversky, 1979).

In complex contexts, heuristics can give rise to biases, but they can also serve as adaptive responses to a lack of time and information (Gigerenzer &

http://dx.doi.org/10.21511/pmf.15(4).2026.02

Gaissmaier, 2011). In the military domain, these conditions are exacerbated by strategic uncertainty, the irreversibility of certain decisions and the severity of their consequences (Knighton, 2004; Liwång, 2017; Roth, 2004). Aggregate budgetary data do not allow for the identification of loss aversion, status quo bias or any specific cognitive mechanism. Identification requires information on possible alternatives, perceived probabilities and the choices actually made. Behavioral literature should therefore delimit the scope of the normative model and not attribute to decision-makers' behaviors that have not been observed.

The economics literature is unanimous in its criticism of the use of aggregate expenditure as a measure for analyzing produced capability (Hitch & McKean, 1960; Hartley, 2012; Smith, 2009). The analysis must take into account opportunity costs, the functional composition of resources, the timing of expenditure, and the potential to convert appropriations into effective capability. An increase in the budget may reflect greater strategic ambition, inflation in asset costs, technological renewal or simply the accumulation of commitments. Without disaggregation, these situations are indistinguishable. The assessment must disaggregate allocation, commitment, payment, physical delivery, entry into service and operational outcome and, in this regard, avoid treating the stages of the process as equivalent. The need for specification is consistent with the evidence that conclusions regarding the effects of military expenditure are sensitive to the theoretical and empirical framework adopted (Dunne et al., 2005).

The composition of the budget is also influenced by the institutional framework and political priorities (Bove & Nisticò, 2014). The efficiency of expenditure depends not only on the amount but also on its purpose and the conditions under which it is implemented, given that expenditure on staff, equipment, infrastructure, operations and maintenance contributes in different ways to the production of effective capabilities (Becker et al., 2025; Dimitriou et al., 2025). The structure with high levels of investment may exhibit poor material performance when processes are complex, contractual timelines span several years, or administrative capacity is limited (Markowski et al., 2010). High level of financial execution does not necessarily attest to the quality of expenditure, as payments may be made before the operational effects are realized.

The relationship between investment and maintenance is central to naval organizations. The acquisition of new assets does not eliminate the need for sustainment; on the contrary, it tends to generate new requirements for technical support, spare parts, infrastructure, training and up-to-date maintenance. Increasing technological complexity can drive up both acquisition and life-cycle costs (Asiedu & Gu, 1998; Kirkpatrick, 1995; Pugh, 2007). Consequently, investment and maintenance should not be regarded as substitutes, given that procurement, industrial support and technical sustainment form interdependent dimensions of capability (Hall et al., 1998). However, complementarity cannot be demonstrated merely by adopting a function that integrates them a priori. CES function with elasticity less than unity necessarily penalizes unbalanced combinations (Arrow et al., 1961). The result describes the logic of the specification, not an estimated relationship between expenditure and operational availability.

The distinction between mathematical properties and empirical results also applies to the logistic transformation used to present potential implementation capacity. Logistic function incorporates limited growth and a change in the marginal rate of increase along the curve. The curve is useful for exploring possible saturation scenarios but does not measure actual financial implementation. However, it requires data on adjusted appropriations, commitments, payments and physical implementation. The availability of financial resources does not guarantee their immediate conversion into results, as implementation depends on the coordination between planning and resource allocation decisions (Webb et al., 2010), program-based management (Tagarev, 2006), administrative capacity (Chen, 2021), process sequencing and coordination between units (Pufahl et al., 2025).

Risk analysis requires the same degree of prudence. Strictly speaking, risk involves recognizing possible events, their probabilities and/or degrees of uncertainty, and their associated effects (Kaplan & Garrick, 1981; Aven, 2016). Exceeding a budgetary threshold does not, in itself, indicate

http://dx.doi.org/10.21511/pmf.15(4).2026.02

the presence of institutional risk. The result merely indicates potential pressure on the organization, provided that the threshold is properly justified and the indicator is presented as an analytical construct rather than as a measure of observed risk. Validation requires information on delays, unavailability, non-compliance, cost overruns, non-conformities or aggregate maintenance – factors that are particularly relevant to the execution of defense contracts (Oudot, 2010). Without this information, the model merely confirms what has already been incorporated into the parameterization, rather than demonstrating the existence of actual institutional risk.

The aggregation of components also raises three main issues: scale, weighting, and compensation of values. Standardization mitigates the problem but does not resolve the choice of weights. Weighted summation allows a high score in one dimension to compensate for a low score in another, a property inherent in various multi-criteria decision-making models (Goodwin & Wright, 2009; Marler & Arora, 2010). This property may be inappropriate when there are non-substitutable minimum requirements, such as critical levels of maintenance or availability. The ranking produced using a composite index must therefore be compared with alternative specifications, veto constraints or dominance analyses, given the influence that normalization, weighting and aggregation decisions exert on the final results (Greco et al., 2019; Nardo et al., 2008; Roy, 1996).

It is in this context that sensitivity analysis is not an optional extra but a fundamental element of the research. Simultaneously varying elasticities, benchmarks, thresholds, penalties and weights enables to identify which assumptions constrain the classification (Saisana et al., 2005; Saltelli et al., 2008; Sobol, 2001). The given configuration can only be considered robust within the model if it maintains the same ranking in the face of a wide variation in parameters. If the ranking changes frequently, the conclusion is not that a particular alternative is superior, but rather that the decision depends on the preferences and thresholds adopted (Cherchye et al., 2008).

The theory therefore supports the adoption of a normative quantitative framework, provided it is made explicit that the model reflects analytical assumptions and not a direct empirical measure of institutional performance. With just three observations, the model allows for the organization of trade-offs, the comparison of configurations and the analysis of the sensitivity of the results, but it does not allow for the estimation of causal relationships nor for the direct assessment of institutional readiness, efficiency or risk. The scientific problem centers, above all, on analyzing how different specifications alter the classification of budgetary compositions and distinguish between results that remain stable and those that depend on adjustments.

The aim of this study is to develop and apply a quantitative decision-support framework for comparing the budgetary compositions of the Portuguese Navy between 2023 and 2025, assessing how the interaction between maintenance, investment, personnel expenditure, potential execution and institutional stability affects institutional utility.

In light of the theoretical framework and the formal structure of the model, five hypotheses were formulated:

H1: *Lower relative proportion of maintenance reduces the balance between maintenance and investment when the elasticity of substitution is less than one.*

H2: *The greater the deviation of staff expenditure from the reference value, the lower the score in the model, with the reduction being proportional to the square of the observed deviation.*

H3: *Higher proportion of investment increases the potential execution index, but the increase is progressively smaller as the logistic function approaches the upper limit.*

H4: *When the proportion of expenditure on staff or investment exceeds the defined thresholds, the modelled institutional stability index decreases.*

H5: *The classification of budgetary configurations can only be considered robust if it remains stable when the model's parameters, weights and thresholds are altered.*

The first four state the conditional implications of the functional forms adopted, whilst the fifth establishes the methodological criterion for assessing the stability of the results. The analysis therefore focuses on the model's econometric consistency, the transparency of the assumptions and the sensitivity of the conclusions to parametric uncertainty.

2. METHOD

The study develops a fitted structural model designed to represent the institutional assessment of the Portuguese Navy's various budgetary compositions. The unit of analysis is the budget year, and the empirical focus covers the distribution of total expenditure (B_t), across maintenance (M_t), investment (I_t), personnel (P_t), goods and services (F_t) and other expenditure (O_t), for the period from 2023 to 2025. The data are available on the institution's official website and have been processed at the level of economic headings. For each year t the following vector is defined:

$$x_t = (m_t, i_t, p_t, f_t, o_t), \; \sum_j x_{jt} = 1. \tag{1}$$

For

$$m_t = \frac{M_t}{B_t}, \quad i_t = \frac{I_t}{B_t}, \quad p_t = \frac{P_t}{B_t}, \\ f_t = \frac{F_t}{B_t}, \quad O_t = 1 - (m_t, i_t, p_t, f_t). \tag{2}$$

Financial soundness is verified by

$$B_t = M_t + I_t + P_t + F_t + O_t, \quad o_t = \frac{O_t}{B_t}. \tag{3}$$

The balance between maintenance and investment is given by

$$C_t = \frac{\left(0.5m_t^{-1} + 0.5i_t^{-1}\right)^{-1}}{0.5m_t + 0.5i_t}. \tag{4}$$

The index approaches 1 when the two proportions are similar and decreases when investment grows without a proportional increase in maintenance. The labor element is

$$P_t^* = \max\left\{0.1 - 2\left(p_t - 0.50\right)^2\right\}. \tag{5}$$

Measuring only the proximity to the 50 per cent benchmark. The potential execution index is calculated as

$$E_t = \frac{1}{1 + \exp\left[-\left(-0.5 + 4i_t - 1.5f_t\right)\right]}. \tag{6}$$

Budgetary pressure results from

$$Q_t = 10\max\left(0, p_t - 0.50\right)^2 \\ +120\max\left(0, i_t - 0.33\right)^2. \tag{7}$$

This is converted into the stability index $S_t = \exp(-Q_t)$. The relationship between budget composition and institutional performance is given by

$$U_t = w_C C_t + w_p P_t + w_E E_t + w_S S_t \\ = 0.25C_t + 0.25P_t + 0.25E_t + 0.25S_t. \tag{8}$$

The procedure is repeated separately for the three financial years, and the values of U_t are ranked.

Robustness is assessed using 200,000 Monte Carlo simulations. In each simulation, the following parameters are drawn simultaneously, $\alpha_m \in [0{,}40; 0{,}60]$, $\rho \in [-2; -0{,}5]$, $p^* \in [0{,}48; 0{,}52]$, $k \in [1; 4]$, $\gamma_0 \in [-0{,}75; -0{,}25]$, $\gamma_1 \in [3; 5]$, $\gamma_2 \in [1; 2]$, $P^+ \in [0{,}48; 0{,}52]$, $i^+ \in [0{,}31; 0{,}35]$, $\eta_p \in [5; 15]$ and $\eta_i \in [80; 160]$. The four weights are generated from a Dirichlet distribution (4, 4, 4, 4). In each simulation, the same set of parameters is applied to the three years, enabling the respective ranking to be determined. The analysis shows the position occupied by each financial year, the Spearman's correlations between the randomly selected parameters, and the differences in utility between the years.

The model enables a systematic and reproducible comparison of the budgetary configurations from 2023 to 2025 and assesses the extent to which the ranking remains stable in the face of changes to the model's specification, parameters, thresholds and weights.

http://dx.doi.org/10.21511/pmf.15(4).2026.02

3. RESULTS

The results characterize the relative position of the budgetary compositions observed in 2023–2025 in relation to an adjusted institutional utility function. The score does not directly measure operational readiness, availability of resources, physical implementation and/or fulfilment of missions. It measures the compatibility of each expenditure structure with the functional relationships, thresholds and weightings incorporated into the model. This distinction is relevant, as the ranking obtained is a joint property of the budgetary data and the analytical specification, not a causal assessment of the Navy's performance.

The budgeted amount under consideration rose from 641.65 million euros in 2023 to 723.52 million in 2024 and 774.23 million in 2025. The nominal increase amounts to 12.76 per cent between 2023 and 2024 and 7.01 per cent between 2024 and 2025, representing a cumulative increase of 20.66 per cent over the period (Table 1). However, the utility function does not depend on the absolute amount but on how the resources have been allocated across the different categories of expenditure. The increase in total expenditure only alters when it changes the relative composition of the budget.

The transition from 2023 to 2024 saw a notable change in the budgetary structure. Investment rose by 68.62 per cent, from 127.45 to 214.90 million euros, which increased the budget by 9.84 percentage points. Staff costs remained stable in nominal terms, rising by just 0.53 per cent, meaning that their share of the budget fell from 53.66 per cent to 47.84 per cent (Table 1). Expenditure on goods and services fell by 7.56 per cent, whilst maintenance expenditure rose by 3.05 per cent – a figure below the increase in the total budget – representing a decrease of 0.87 percentage points. The result is therefore a relative reallocation towards investment without a nominal reduction in maintenance expenditure.

In 2025, investment continued to rise, reaching a total of 244.71 million euros, but maintenance fell to 39.42 per cent, amounting to 40.56 million. Between 2023 and 2025, investment grew by 92.01 per cent, whilst maintenance fell to 37.58 per cent. Consequently, maintenance gradually lost ground relative to investment. In 2023, for every euro invested, 0.51 euros were allocated to maintenance. In 2024, this figure fell to 0.31 euros and, in 2025, to 0.17 euros. The 2025 budget expansion does not, therefore, represent a linear continuation of the 2024 pattern; rather, it reflects an intensified focus on investment, accompanied by a sharp reduction in the component allocated to asset maintenance.

The change is reflected in the four components of the function. The estimated utility for 2024 exceeded that of 2023 by 0.031792 points. The difference stems, in particular, from the improvement in the implementation efficiency component, which increased by 0.026284 points and accounts for 82.67% of the total variation. The difference is mainly due to the removal of the stability penalty, which adds 0.00328 points, or 10.477%. In 2023, the penalty was applied because the proportion of staff costs exceeded the 50% threshold; in 2024, the staff costs component fell to 47.84%, rendering the penalty unnecessary. The components relating to the balance between maintenance and investment and the adequacy of staff expenditure have a significantly reduced effect, contributing 0.001744 and 0.000437 points, respectively. The higher figure for 2024 results from the potential implementation logistics function. The increase in the investment component and the reduction in goods and services underpin the change in the logistics function from 0.083617 to 0.515207, causing Et to rise from 0.520892 to 0.626026 (Table 2).

The stability index also deviates from the 2023 figure. The personnel component, at 53.66 per cent, exceeds the baseline threshold of 50 per cent and yields a Q_t = 0.013403 and S_t = 0.986687. In 2024 and 2025, the proportions of the personnel and in-

Table 1. Changes in the budget breakdown

Year	Total expenditure (millions of €)	Recurrent expenditure	Investments	Personnel costs	Goods and services	Other expenditure
2023	641.65	10.13%	19.86%	53.66%	14.06%	2.29%
2024	723.52	9.25%	29.70%	47.84%	11.53%	1.68%
2025	774.23	5.24%	31.61%	49.71%	11.46%	1.98%

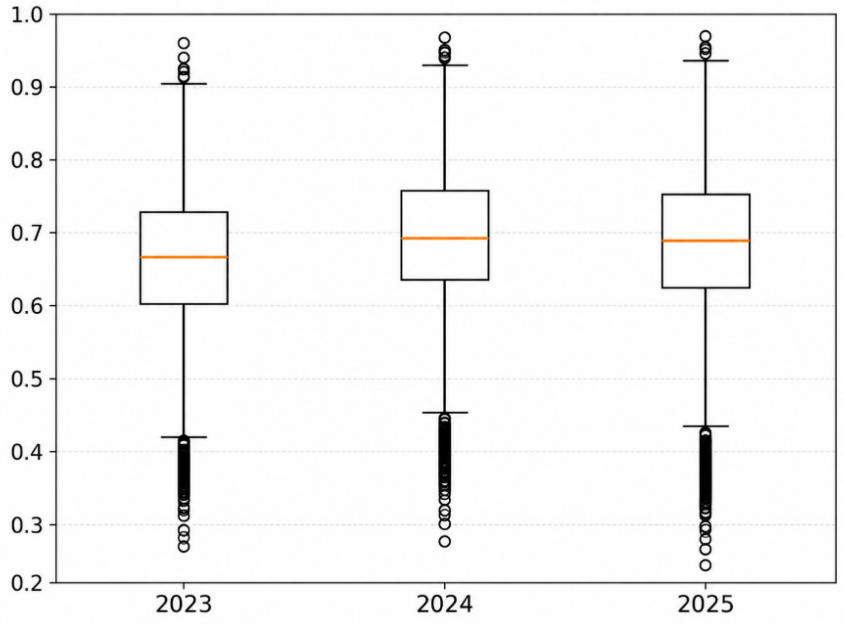

Figure 1. Monte Carlo utility distributions by budget year

vestment components do not simultaneously exceed the relevant thresholds of the base configuration; therefore, $S_t = 1$. The penalty is applied discontinuously; that is, differences below the threshold incur no cost, whilst excesses are penalized quadratically. In this case, stability distinguishes 2023 from the other financial years but does not allow for a distinction between 2024 and 2025.

The comparison between 2024 and 2025 reveals the model's trade-off. In 2025, potential performance increases by 0.017882, which adds 0.004470 points to the weighted utility. The component improves marginally and contributes 0.000229 points. These gains are offset by the reduction in C_t, whose contribution, which decreases by 0.012810 points. Total utility falls by 0.008110. The 2025 configuration achieves the highest score for potential performance, but the concentration of the budget on the investment heading, without proportional maintenance, reduces the combined assessment of the two headings. The model selects 2024, as this year combines a level of investment substantially higher than that of 2023 with maintenance that is still sufficient to prevent the deterioration observed in 2025 (Figure 1).

The analysis also highlights an asymmetry between the formal weights and the components' actual discriminatory power. P_t^* varies only between 0.997319 and 0.999984, whilst S_t varies between 0.986687 and 1. Although the two components together account for 50 per cent of the nominal weight, they generate reduced cross-sectional variation. The result is a function whose ranking is determined mainly by C_t and E_t. The staff component is practically saturated around its maximum, whilst stability equals one in two of the three exercises. Equal weights do not, therefore, imply equal empirical influence.

The comparison between 2024 and 2023 is the most stable. The difference U2024 – U2023 is posi-

Table 2. Utility components in the baseline configuration

Year	C_t	P_t^*	E_t	S_t	U_t	Position
2023	0.134146	0.997319	0.520892	0.986687	0.659761	3ª
2024	0.141122	0.999067	0.626026	1.000000	0.691554	1ª
2025	0.089882	0.999984	0.643908	1.000000	0.683443	2ª

Note: Ct denotes the maintenance–investment complementarity component; Pt* denotes personnel adequacy; Et denotes potential execution; St denotes institutional stability; and Ut denotes aggregate institutional utility. Position refers to the ordinal ranking of each budget year according to Ut.

http://dx.doi.org/10.21511/pmf.15(4).2026.02

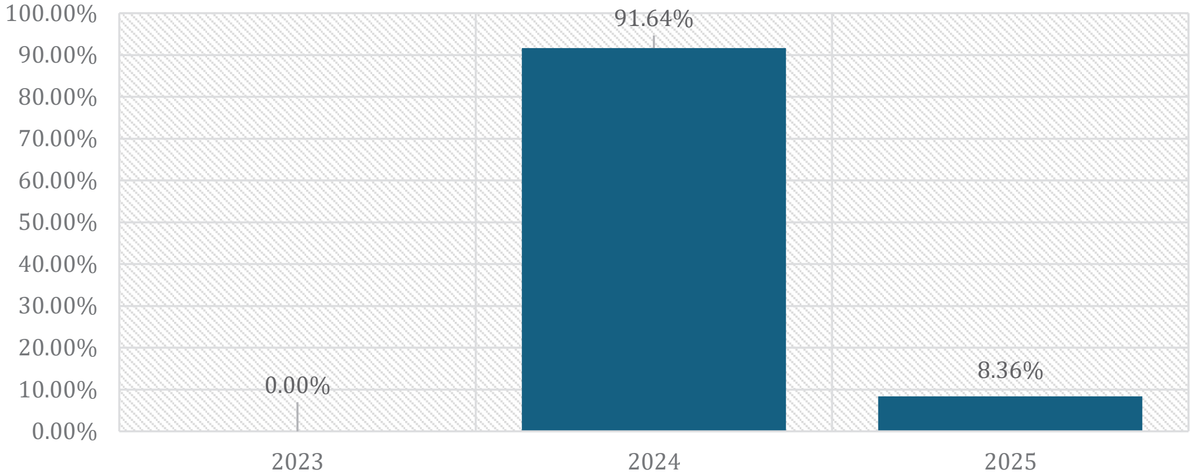

Figure 2. Frequency of first place across Monte Carlo simulations

tive in all simulations, with a mean of 0.031403, a median of 0.030331 and a minimum of 0.001816. Even under the combination of parameters most unfavorable to 2024, utility remains higher than that of 2023. This inference is valid only for the values tested and should not be generalized to other model specifications (Figure 2).

The difference U2025 – U2023 is positive in 94.4035% of the simulations. On average, the utility in 2025 was 0.023166 points higher than that in 2023, whilst the 5th percentile is –0.008236. In around 5 per cent of the simulations, the 2025 period shows a utility equal to or lower than that of 2023, which demonstrates that the advantage is not stable across all parameter combinations. The comparison between 2024 and 2025 is more uncertain; that is, U2025 – U2024 shows a mean of -0.008236 and a median of –0.007749, making it positive in 8.3575% of the simulations. The 95th percentile is just 0.001275, which confirms that, even under parameters relatively favorable to 2025, the advantage tends to be small (Figure 3).

The sensitivity test indicates the possible cause of the inversion. The Spearman's correlation between w_C and U2025 – U2024 is –0.945865, whilst the correlation with w_E is 0.554671. When 2024 is in the first position, the average weight of C_t is 0.263067 and that of E_t is 0.237148. Conversely, in 2025, the values change to 0.102046 and 0.393427, respectively. In standard deviation units, the conditional difference is –1.53 for w_C and 1.49 for w_E. The shift in leadership from 2024 to 2025 is not explained by the curve parameters or the thresholds; it results from the shift in normative importance

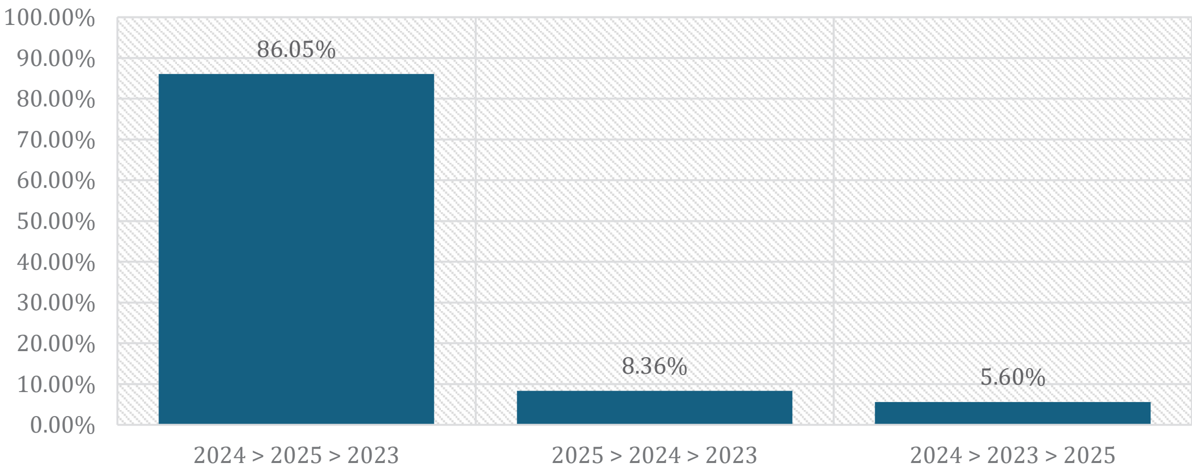

Figure 3. Frequency of complete ranking orders

between complementarity and potential implementation. The parameter γ_1, which governs the effect of investment on the logistic function, shows a correlation of just 0.063063 with the difference between 2025 and 2024.

The evidence alters the interpretation of robustness. The relevant uncertainty is neither essentially statistical nor technological; it is an uncertainty of institutional preference. The 2024 ranking is stable when substantial weight is given to the maintenance-investment balance. The 2025 scenario serves as a reference only when an evaluation function is accepted that substantially reduces the weighting and prioritizes potential performance. The simulated frequencies do not reveal objective probabilities of superiority. The table shows the percentage of combinations of the tested parameters in which each ranking of the three years occurs. These depend on the uniform intervals, the Dirichlet distribution and its concentration. In the Dirichlet distribution (4, 4, 4, 4), each weight has a mean of 0.25, but the empirical range, between the 5th and 95th percentiles, is approximately [0.097; 0.440], which allowed for appreciable variations in the structure of preferences.

There is also the issue of specification with substantive effect. With regard to C_t, this is calculated as an unnormalized CES average. The term does not merely measure balance; it also captures the degree of complementarity with the combined level of maintenance and investment proportions. Although 2023 shows a more balanced relationship between the two categories, C2024 > C2023 because the sum of the proportions is higher in 2024. The term 'equilibrium' is therefore incomplete. The term should be interpreted as joint capacity adjusted for complementarity.

A counterfactual test confirms the importance of this choice. By normalizing the harmonic mean by the arithmetic mean, so as to isolate relative equilibrium exclusively, the following approximate values were obtained: 0.8946, 0.7245 and 0.4879 for 2023, 2024 and 2025, respectively. Whilst retaining the remaining components and base weights, the ranking changed to 2023 > 2024 > 2025 (Figure 4). Monte Carlo simulations suggest that the 2024 ranking is stable in the face of changes to the parameters. However, the alternative test reveals that this position does not hold when the indicator's assumptions are altered. The stability observed in 2024 should be regarded as internal robustness within the adopted specification, rather than as robustness to plausible alternative specifications.

Table 3. Results of applying the Monte Carlo model with 200,000 simulations

Results	Frequency	Probability
2024 in first place	183,285	91.6425%
2025 in first place	16,715	8.3575%
2023 in first place	0	0.0000%
(2024 > 2025 > 2023)	172,092	86.0460%
(2024 > 2023 > 2025)	11,193	5.5965%
(2025 > 2024 > 2023)	16,715	8.3575%

Table 3 assesses the stability of the ranking of the three financial years in the face of parameter uncertainty. In 183,285 of the 200,000 simulations, there is a 91.6425 per cent probability that 2024 will be ranked first (Figure 5). This result means that, in more than 9 out of every 10 combinations

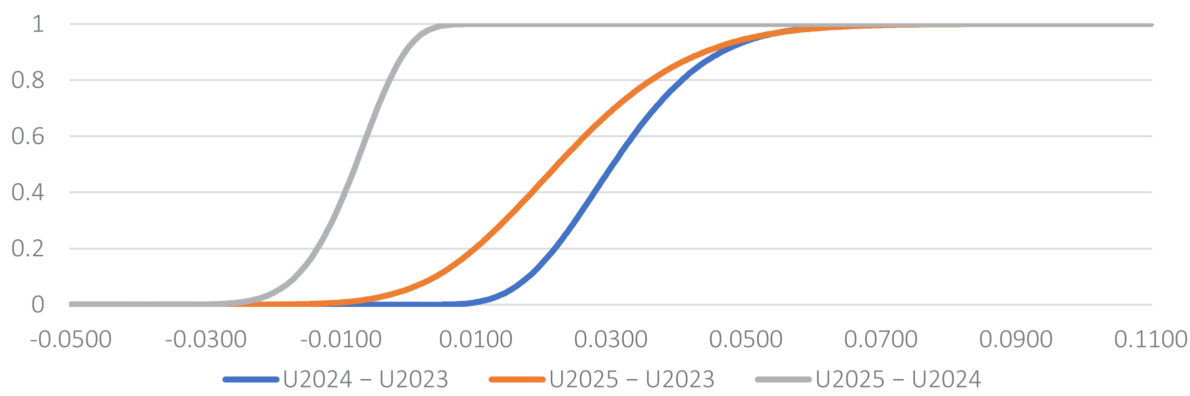

Figure 4. Empirical cumulative distributions of pairwise utility differences

http://dx.doi.org/10.21511/pmf.15(4).2026.02

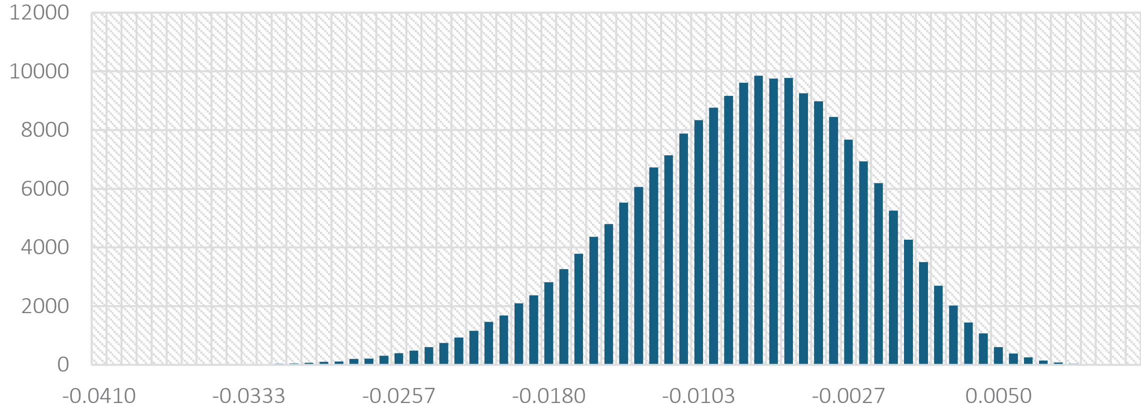

Figure 5. Distribution of U2025 – U2024

of parameters tested, the year 2024 maintained the highest score amongst the three financial years. The ranking does not, therefore, depend solely on the values defined in the base configuration, although it remains subject to the ranges and assumptions adopted in the simulation. However, this lead is not absolute: in 16,715 simulations, or 8.3575 per cent, the 2025 period outperforms 2024. The difference between the two financial years is therefore the main source of instability in the ranking.

The 2024 > 2025 > 2023 configuration is clearly predominant and accounts for a total of 86.046 per cent of the simulations. The results indicate that, across all combinations of parameters and weights, 2024 exhibits the budget composition most compatible with the utility function, whilst 2025 occupies an intermediate position and 2023 remains in last place. The ranking 2025 > 2024 > 2023, observed in 8.3575 per cent of the simulations, shows, however, that 2025 is preferable when the adjustment places greater emphasis on the components in which the exercise performs best – particularly potential execution – and less emphasis on the balance between maintenance and investment.

The 2024 > 2023 > 2025 configuration occurs in 5.5965 per cent of the simulations. This result is particularly relevant, as it shows that the predominance of 2025 over 2023 is less stable than that observed for 2024. Consequently, 2023 does not take first place but outperforms 2025 in a limited proportion of the configurations. Conversely, 2024 always remains above 2023, given that none of the observed rankings place 2023 in first place (Figure 6).

The scale of the simulation substantially reduces the numerical error in the estimated frequencies. For the probability of 2024 taking the lead, the Monte Carlo standard error is approximately 0.062 percentage points. The proportions do not reveal objective probabilities of institutional leadership.

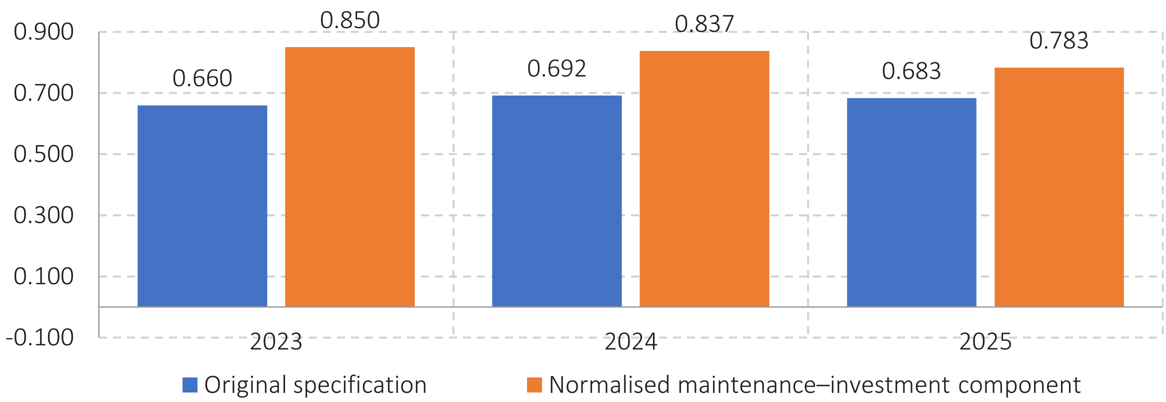

Figure 6. Sensitivity of utility levels to the specification of the maintenance-investment component

They are frequencies conditioned by the functional specification, the parametric ranges and the distributions adopted. The table thus underpins the robustness of the model. The 2024 period broadly leads the adopted model, but the comparison with 2025 remains sensitive to the structure of preferences incorporated into the function.

In summary, in the model, the 2024 period reflects the composition most consistent with the combination of investment, maintenance, staffing, potential performance and stability of the function. The lead over 2023 is substantial and persists across the entire simulated parameter range. For 2025, the lead is smaller and depends on the weighting given to the combination of maintenance and investment. The 2025 period does not show a shortfall in resources; it presents a more polarized structure, in which an increase in investment occurs simultaneously with a decrease in maintenance. The institutional interpretation must, however, remain subject to the absence of physical performance indicators, the accounting classification of the budget headings and the adjusted – rather than estimated – nature of the parameters.

4. DISCUSSION

The key findings do not lie in the fact that 2024 recorded the highest profit but in the reason why the increase in the budget in 2025 did not result in a further improvement in the rating. Between 2023 and 2025, total expenditure rose by 20.66 per cent. However, utility peaked in 2024 and declined in the following year. This discrepancy demonstrates the inadequacy of aggregate expenditure as an indicator of capability, as noted by Hitch and McKean (1960), Smith (2009) and Hartley (2012). The increase in the budget boosts the availability of resources but does not determine the coherence of the budget's composition, its actual implementation or the effective capability produced. The result is also consistent with Dunne et al. (2005), in demonstrating that the assessment of military expenditure depends on the theoretical framework adopted. Budgetary growth can be viewed as a strengthening of capability, an accumulation of commitments and/or a functional imbalance, depending on the variables and relationships considered.

The 2024 figure reflects, above all, a combination that the model assesses as intermediate between the 2023 and 2025 structures. Investment increased significantly compared with 2023, but maintenance expenditure has not yet seen the reduction observed in the 2025 financial year. In 2025, investment reached its highest level, whilst maintenance fell to 5.24 per cent of total expenditure. This result is consistent with Kirkpatrick (1995), Pugh (2007) and Hall et al. (1998), who reject the interpretation of procurement and sustainment as independent decisions. The adoption of new equipment increases potential capacity but also generates subsequent requirements for maintenance, spare parts, infrastructure, technical support and training. Thus, the 2025 structure may reflect short-term modernization but also the possible accumulation of sustainment requirements that only materialize in later periods.

This interpretation, however, calls for caution. The model does not take into account operational availability, breakdowns, operating hours, maintenance delays and/or life-cycle costs. The reduction in the maintenance budget does not prove that maintenance has been materially insufficient. It may reflect changes in accounting classification, deferred payments, multi-year contracts or a concentration of acquisitions over a specific period. In fact, the convergence with the literature on complementarity is conceptual, not causal. The results indicate a budgetary structure consistent with a possible tension between procurement and sustainment but do not prove any operational effects.

With regard to the propositions, Proposition 1 must be assessed in light of the limitations. The reduction in C_t in 2025 is consistent with the assertion that a lower relative share of maintenance reduces the joint valuation of maintenance and investment when the elasticity of substitution is less than 1. However, the result does not constitute empirical confirmation of complementarity. As argued by Arrow et al. (1961), CES function with a chosen parameter already incorporates the difficulty of substitution between factors. The reduction in C_t is thus a logical consequence of the specification. The analysis demonstrates how the model responds to disequilibrium but does not estimate the actual elasticity, nor does it examine whether maintenance and investment jointly lead to greater operational readiness.

http://dx.doi.org/10.21511/pmf.15(4).2026.02

The issue is all the more important because the aggregator used does not simply measure the balance. The non-standardized formulation combines the ratio between the two components with their combined weight in the budget. For this reason, C2024 surpasses C2023, even though 2023 exhibits a less asymmetrical ratio between maintenance and investment. When the indicator is standardized to isolate, exclusively, the relative balance, the ranking changes to 2023 > 2024 > 2025. This reversal is not merely a technical detail; it demonstrates that the position of 2024 depends on the substantive interpretation attributed to C_t. If the measure is intended to reflect the joint capacity adjusted for complementarity, the original formulation of the model is defensible. If it is intended to measure only the equilibrium, the normalization is conceptually more consistent. The discussion of Proposition 1 must therefore acknowledge that the conclusion depends on the definition of the construct itself.

The trend in E_t shows a similar pattern. The index rises from 0.520892 in 2023 to 0.626026 in 2024 and 0.643908 in 2025. This progression is consistent with Proposition 3 and with the saturation property of the logistic function. The rate of increase slows as the value approaches the upper limit. The pattern does not show diminishing marginal returns from the observed performance. The increase in the investment ratio is considerably greater between 2023 and 2024 than the change between 2024 and 2025; therefore, the reduction in the increment of E_t results both from the curvature of the function and from the smaller variation in the data.

A comparison with Webb et al. (2010), Tagarev (2006), Chen (2021) and Pufahl et al. (2025) reinforces this reservation. For the authors, implementation depends on planning, scheduling, administrative capacity, coordination and the sequence of processes. The model does not take into account commitments, payments, implementation rates, deliveries or physical realization. E_t should be interpreted as the potential capacity associated with the budgetary composition and not as actual implementation. The '2025 lead' simply means that the organization in question scores higher in terms of the logistical transformation adopted. It does not mean that the 2025 period has converted resources into results more effectively.

The personnel component has a lower explanatory power. P_t^* remains very close to 1 in all financial years, so its influence on the ranking is limited. The nominal stability of expenditure is consistent with Wildavsky (1984), March (1994) and Scott (2014), as they highlight incrementalism, path dependence and institutional rigidity. Expenditure stems from career structures, employment contracts, permanent requirements and accumulated decisions, which cannot be adjusted as quickly as investment. However, proximity to the 50 per cent benchmark does not attest to structural adequacy. Proposition 2 is satisfied, as the function imposes a quadratic penalty on deviation; however, the benchmark is neither estimated nor validated by the operational results. In fact, the model measures the distance to a given normative benchmark but does not identify the efficient proportion of staff expenditure.

As for Proposition 4, the same limitations can be identified. In 2023, the proportion of expenditure on staff exceeds the threshold and reduces S_t in 2024 and 2025. This pattern confirms the mathematical operation of the function but does not validate the concept of institutional risk. Kaplan and Garrick (1981) and Aven (2016) associate risk with possible events, uncertainty and consequences. Oudot (2010) refers to delays, non-compliance, cost overruns and specific contractual problems. None of these dimensions are observed. The index merely reflects analytical exposure to a pressure defined by the model, rather than the actual risk that may be incurred. The term 'institutional stability' is therefore more appropriate than 'risk', provided it corresponds to a normative construct.

The ranking also highlights the importance of trade-offs between components. The improvement in E_t in 2025 partially offsets the reduction in C_t, but does not cancel it out. This property reflects the behavior of the additive models analyzed by Keeney and Raiffa (1993), Goodwin and Wright (2009) and also Marler and Arora (2010). The issue is a substantive one: insufficient maintenance can be mathematically offset by improved potential performance, even though, operationally, there is a minimum level of maintenance that cannot be substituted. If maintenance is a necessary condition for availability, a fully compensatory rule is debatable. The recommendations of Nardo et al.

(2008) and Greco et al. (2019) justify comparison with minimum thresholds, veto rules or partially non-compensatory models.

Sensitivity analysis allows for a more rigorous assessment of proposition 5. The lead in 2024, observed in 91.6425 per cent of the simulations, does not depend solely on the adjustments. The lead over 2023 is particularly stable, as it holds true across all the parametric combinations analyzed. However, the comparison with 2025 is clearly conditional. In 8.3575 per cent of the simulations, the 2025 period takes first place. The reversal occurs when the weight assigned to C_t decreases and the weight of E_t increases. The choice between 2024 and 2025 is not determined by an unequivocal difference in performance, but by the institutional priority given to joint sustainability or potential implementation.

The results are supported by Saisana et al. (2005), Saltelli et al. (2008) and Cherchye et al. (2008). Sensitivity analysis serves not only to confirm the classification, but also to identify the possible assumptions underlying it. Proposition 5 concerns the validity in terms of parametric stability for 2024, particularly in relation to 2023, but not in terms of structural robustness. A change in the functional structure C_t alters the ranking; this situation demonstrates that the uncertainty in the specification is greater than the uncertainty in the parameters. The 200,000 simulations provide a high degree of computational precision; however, they do not validate the model beyond the adopted architecture.

The disparity brings the results into line with those of Simon (1957), Cyert and March (1963) and Lindblom (1959). The function generates a rational ranking under certain explicit criteria but does not represent the actual decision-making process. Decision-makers do not freely choose the proportions for each period and are faced with commitments from previous periods, legal constraints, incomplete information and, furthermore, limited alternatives. A lead in 2024 does not mean that the same composition is viable in 2023 or 2025. Furthermore, it does not allow for the attribution of Behaviors such as loss aversion, status quo bias and/or the adoption of heuristics, as defined by Kahneman and Tversky (1979) and Gigerenzer and Gaissmaier (2011), since alternatives, expectations and/or individual choices were not examined.

The contribution of the results is therefore analytical. The model does not identify the optimal budget, nor does it directly measure readiness, efficiency and/or risk. It merely confirms that 2024 is the preferred option when the decision simultaneously prioritizes investment, sustainability and the absence of structural penalties. The 2025 period is preferable when potential implementation is given greater importance, and 2023 becomes more relevant when the relative balance between maintenance and investment is considered in isolation from the overall picture.

CONCLUSION

The study devised a possible quantitative framework to compare the budgetary compositions of the Portuguese Navy between 2023 and 2025, taking into account the interplay between maintenance and investment, the appropriateness of personnel expenditure, potential expenditure, and institutional stability. The results show that an increase in the budget does not, in itself, lead to an improvement in institutional utility. Although 2025 has the largest budget allocation and the highest potential execution value, 2024 achieves the highest aggregate utility, as it combines increased investment with a relatively smaller reduction in maintenance.

The sensitivity analysis confirms the parametric stability of the ranking: 2024 occupies the top position in 91.64 per cent of the 200,000 simulations and outperforms 2023 in all configurations analyzed. The comparison between 2024 and 2025 is, however, sensitive to the importance attributed to the complementarity between maintenance and investment and to potential expenditure. The results indicate that the ranking depends not only on budgetary data but also on the choices regarding specification, weighting and aggregation.

http://dx.doi.org/10.21511/pmf.15(4).2026.02

The analytical propositions are consistent with the formal behavior of the model but do not reflect causal relationships. The classification remains limited by the weightings, thresholds and functional forms adopted. The model is therefore a decision-support tool rather than a direct measure of institutional readiness, efficiency or risk. Future research should incorporate operational indicators, physical implementation, life-cycle costs and non-compensatory specifications.

AUTHOR CONTRIBUTIONS

Conceptualization: Ricardo de Moraes e Soares.
Formal analysis: Ricardo de Moraes e Soares.
Investigation: Ricardo de Moraes e Soares.
Methodology: Ricardo de Moraes e Soares.
Project administration: Ricardo de Moraes e Soares.
Resources: Ricardo de Moraes e Soares.
Software: Ricardo de Moraes e Soares.
Supervision: Ricardo de Moraes e Soares.
Validation: Ricardo de Moraes e Soares.
Visualization: Ricardo de Moraes e Soares.
Writing – original draft: Ricardo de Moraes e Soares.
Writing – review & editing: Ricardo de Moraes e Soares.

ACKNOWLEDGMENT

This article is financed by CINAV–Navy Research Centre [CINAV–Centro de Investigação Naval].

REFERENCES

1. Arrow, K., Chenery, H., Minhas, B., & Solow, R. (1961). Capital-labor substitution and economic efficiency. *The Review of Economics and Statistics, 43*(3), 225-250. https://doi.org/10.2307/1927286
2. Asiedu, Y., & Gu, P. (1998). Product life-cycle cost analysis: State of the art review. *International Journal of Production Research, 36*(4), 883-908. https://doi.org/10.1080/002075498193444
3. Aven, T. (2016). Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research, 253*(1), 1-13. https://doi.org/10.1016/j.ejor.2015.12.023
4. Battaglio, R., Belardinelli, P., Bellé, N., & Cantarelli, P. (2019). Behavioral public administration ad fontes: A synthesis of research on bounded rationality, cognitive biases, and nudging in public organizations. *Public Administration Review, 79*(3), 304-320. https://doi.org/10.1111/puar.12994
5. Becker, J., Benson, S., Dunne, J., & Malesky, E. (2025). Disaggregated defense spending: Introduction to data. *Journal of Peace Research, 62*(3), 772-788. https://doi.org/10.1177/00223433231215785
6. Belton, V., & Stewart, T. (2002). *Multiple criteria decision analysis: An integrated approach*. Kluwer Academic Publishers. https://doi.org/10.1007/978-1-4615-1495-4
7. Bove, V., & Nisticò, R. (2014). Military in politics and budgetary allocations. *Journal of Comparative Economics, 42*(4), 1065-1078. https://doi.org/10.1016/j.jce.2014.02.002
8. Chen, L. (2021). The impact of performance budgeting on defense resource allocation. *International Journal of Economic Behavior and Organization, 9*(4), 151-156. https://doi.org/10.11648/j.ijebo.20210904.15
9. Cherchye, L., Moesen, W., Rogge, N., Van Puyenbroeck, T., Saisana, M., Saltelli, A., Liska, R., & Tarantola, S. (2008). Creating composite indicators with DEA and robustness analysis: The case of the Technology Achievement Index. *Journal of the Operational Research Society, 59*(2), 239-251. https://doi.org/10.1057/palgrave.jors.2602445
10. Conlisk, J. (1996). Why bounded rationality? *Journal of Economic Literature, 34*(2), 669-700. Retrieved from https://www.jstor.org/stable/2729218
11. Cyert, R., & March, J. (1963). *A behavioral theory of the firm*. Prentice-Hall. Retrieved from https://archive.org/details/behavioraltheory0000cyer
12. Dimitriou, D., Goulas, E., & Kallandranis, C. (2025). Spend on what? Insights on military spending efficiency. *European Journal of Political Economy, 88,* Article

102696. https://doi.org/10.1016/j.ejpoleco.2025.102696

13. Dunne, J., Smith, R., & Willenbockel, D. (2005). Models of military expenditure and growth: A critical review. *Defence and Peace Economics, 16*(6), 449-461. https://doi.org/10.1080/10242690500167791
14. Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62,* 451-482. https://doi.org/10.1146/annurev-psych-120709-145346
15. Goodwin, P., & Wright, G. (2009). *Decision analysis for management judgment*. Wiley. Retrieved from https://archive.org/details/decisionanalysis0004good
16. Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research, 141*(1), 61-94. https://doi.org/10.1007/s11205-017-1832-9
17. Hall, P., Markowski, S., & Thomson, D. (1998). Defence procurement and domestic industry: The Australian experience. *Defence and Peace Economics, 9*(1-2), 137-165. https://doi.org/10.1080/10430719808404898
18. Hartley, K. (2012). *The economics of defence policy: A new perspective*. Routledge. https://doi.org/10.4324/9780203838778
19. Hitch, C., & McKean, R. (1960). *The economics of defense in the nuclear age*. Harvard University Press. Retrieved from https://archive.org/details/economicsofdefen0000unse_m9h6/page/n7/mode/2up
20. Jones, B., & Baumgartner, F. (2005). *The politics of attention: How government prioritizes problems*. University of Chicago Press. Retrieved from https://archive.org/details/politicsofattent0000jone
21. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-291. https://doi.org/10.2307/1914185
22. Kaplan, S., & Garrick, B. (1981). On the quantitative definition of risk. *Risk Analysis, 1*(1), 11-27. https://doi.org/10.1111/j.1539-6924.1981.tb01350.x
23. Keeney, R., & Raiffa, H. (1993). *Decisions with multiple objectives: Preferences and value trade-offs* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139174084
24. Kirkpatrick, D. (1995). The rising unit cost of defence equipment - The reasons and the results. *Defence and Peace Economics, 6*(4), 263-288. https://doi.org/10.1080/10430719508404831
25. Knighton, R. (2004). The psychology of risk and its role in military decision-making. *Defence Studies, 4*(3), 309-334. https://doi.org/10.1080/1470243042000344786
26. Lindblom, C. (1959). The science of "muddling through." *Public Administration Review, 19*(2), 79-88. https://doi.org/10.2307/973677
27. Liwång, H. (2017). Risk communication within military decision-making: Pedagogic considerations. *Defense & Security Analysis, 33*(1), 30-44. https://doi.org/10.1080/14751798.2016.1269389
28. Lopes, H., & Rosado, D. (2018). A importância da contabilidade e dos sistemas de informação de apoio à decisão na gestão organizacional do Exército Português [The importance of accounting and decision-support information systems in the organizational management of the Portuguese Army]. *Proelium, 8*(1), 91-119. (In Portuguese). Retrieved from https://ru.scribd.com/document/603359703/Proelium-VIII-1-2018-91-119
29. March, J. (1994). *A primer on decision making: How decisions happen*. Free Press. Retrieved from https://archive.org/details/primerondecision00marc
30. Markowski, S., Hall, P., & Wylie, R. (Eds.). (2010). *Defence procurement and industry policy: A small country perspective*. Routledge. https://doi.org/10.4324/9780203013694
31. Marler, R., & Arora, J. (2010). The weighted sum method for multi-objective optimization: New insights. *Structural and Multidisciplinary Optimization, 41*(6), 853-862. https://doi.org/10.1007/s00158-009-0460-7
32. Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffmann, A., & Giovannini, E. (2008). *Handbook on constructing composite indicators: Methodology and user guide*. OECD Publishing. https://doi.org/10.1787/9789264043466-en
33. Oudot, J. (2010). Performance and risks in the defense procurement sector. *Journal of Public Policy, 30*(2), 201-218. https://doi.org/10.1017/S0143814X10000073
34. Pollitt, C., & Bouckaert, G. (2017). *Public management reform: A comparative analysis - Into the age of austerity* (4th ed.). Oxford University Press.
35. Pufahl, L., Stiehle, F., Ihde, S., Weske, M., & Weber, I. (2025). Resource allocation in business process executions: A systematic literature study. *Information Systems, 132,* Article 102541. https://doi.org/10.1016/j.is.2025.102541
36. Pugh, P. (2007). Retrospect and prospect: Trends in cost and their implications for UK aerospace. *Defence and Peace Economics, 18*(1), 25-37. https://doi.org/10.1080/10242690600900505
37. Roth, R. (2004). The rational analytical approach to decision-making: An adequate strategy for military commanders? *Connections: The Quarterly Journal, 3*(2), 71-92. https://doi.org/10.11610/Connections.03.2.08
38. Roy, B. (1996). *Multicriteria methodology for decision aiding*. Springer. https://doi.org/10.1007/978-1-4757-2500-1
39. Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 168*(2), 307-323. https://doi.org/10.1111/j.1467-985X.2005.00350.x
40. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Taran-

http://dx.doi.org/10.21511/pmf.15(4).2026.02

tola, S. (2008). *Global sensitivity analysis: The primer*. Wiley. https://doi.org/10.1002/9780470725184

41. Savage, L. (1972). *The foundations of statistics* (2nd rev. ed.). Dover Publications. Retrieved from https://archive.org/details/foundationsofsta00leon

42. Schick, A. (1966). The road to PPB: The stages of budget reform. *Public Administration Review, 26*(4), 243-258. https://doi.org/10.2307/973296

43. Scott, W. (2014). *Institutions and organizations: Ideas, interests, and identities* (4th ed.). SAGE Publications. Retrieved from https://uk.sagepub.com/en-gb/eur/institutions-and-organizations/book237665

44. Simon, H. (1957). *Administrative behavior: A study of decision-making processes in administrative organization* (3rd ed.). Free Press. Retrieved from https://archive.org/details/administrativebe00simo

45. Smith, R. (2009). *Military economics: The interaction of power and money*. Palgrave Macmillan. https://doi.org/10.1057/9780230244672

46. Sobol, I. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation, 55*(1-3), 271-280. https://doi.org/10.1016/S0378-4754(00)00270-6

47. Tagarev, T. (2006). Introduction to program-based defense resource management. *Connections: The Quarterly Journal, 5*(1), 55-69. https://doi.org/10.11610/Connections.05.1.05

48. von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press. Retrieved from https://archive.org/details/in.ernet.dli.2015.215284

49. Webb, N., Richter, A., & Bonsper, D. (2010). Linking defense planning and resource decisions: A return to systems thinking. *Defense & Security Analysis, 26*(4), 387-400. https://doi.org/10.1080/14751798.2010.534647

50. Wildavsky, A. (1984). *The politics of the budgetary process* (4th ed.). Little, Brown. Retrieved from https://archive.org/details/politicsofbudget0004wild

APPENDIX A

Table A1. Integrated evidence on institutional utility, ranking stability and model sensitivity

Analytical dimension	Quantitative evidence	Sensitivity and robustness evidence	Scientific interpretation
Simulation design and reproducibility	200,000 Monte Carlo simulations; pseudo-random seed: 20260726. The same parameter vector was applied to 2023, 2024 and 2025 in each simulation. Parameters were drawn from: $a_m \sim U[0.40, 0.60]$; $p \sim U[-2.00, -0.50]$; $p^* \sim U[0.48, 0.52]$; $k \sim U[1, 4]$; $\gamma_0 \sim U[-0.75, -0.25]$; $\gamma_1 \sim U[3,5]$; $\gamma_2 \sim U[1,2]$; $p^+ \sim U[0.48, 0.52]$; $i^+ \sim U[0.31, 0.35]$; $\eta_p \sim U[5,15]$ and $\eta_i \sim U[80,160]$. Component weights followed $(w_C, w_P, w_E, w_S) \sim$ Dirichlet (4, 4, 4, 4). $\sim U[5,15]$	All parameters, thresholds, penalty coefficients and weights varied simultaneously. Pairwise comparisons were therefore conditional on an identical parameter draw for the three budget years	The simulation evaluates global parametric uncertainty within the adopted model. The paired design prevents differences between years from being generated by year-specific parameter draws. The resulting frequencies remain conditional on the selected functional forms, parameter intervals and distributions
Distribution of annual utility	2023: mean = 0.659017; SD = 0.088060; P5 = 0.508821; median = 0.662146; P95 = 0.798861. 2024: mean = 0.690419; SD = 0.086966; P5 = 0.539726; median = 0.694962; P95 = 0.825638. 2025: mean = 0.682183; SD = 0.091700; P5 = 0.522179; median = 0.687608; P95 = 0.823553	The distributions overlap, but the distribution for 2024 is shifted towards higher values. The 2025 distribution has the largest standard deviation and the highest observed maximum, 0.971673, whereas its minimum, 0.229137, is the lowest of the three years	The higher central utility of 2024 does not imply first-order stochastic dominance. The wider distribution for 2025 indicates greater sensitivity to the parameter and weight configurations: it can generate the highest utility under favourable preferences, but also the lowest outcomes
Comparison between 2024 and 2023	$U_{2024}-U_{2023}$: mean = 0.031403; SD = 0.011191; minimum = 0.001816; P5 = 0.014974; median = 0.030331; P95 = 0.051450; maximum = 0.098680. The difference was positive in 100,0000% of the simulations	Principal Spearman correlations: w_E = 0.901455; w_P = −0.346597; w_C = −0.311945; γ_1 = 0.241940; w_S = 0.194246	The superiority of 2024 over 2023 is the most stable result within the simulated parameter domain. Its magnitude increases mainly when greater importance is assigned to potential execution. Nevertheless, the conclusion is restricted to the adopted model and the tested parameter intervals
Comparison between 2025 and 2023	$U_{2025} - U_{2023}$: mean = 0.023166; SD = 0.015615; minimum = −0.036813; P5 = −0.000704; median = 0.022148; P95 = 0.050522; maximum = 0.109361. The difference was positive in 94,4035% of the simulations	Principal Spearman correlations: w_E = 0.871993; w_C = −0.605012; γ_1 = 0.197519; w_P = −0.162283; P^+ = −0.124854	The advantage of 2025 over 2023 is predominant but not uniform. The negative lower tail shows that 2023 becomes preferable under combinations that attach greater importance to maintenance–investment complementarity and less importance to potential execution
Comparison between 2025 and 2024	$U_{2025}-U_{2024}$: mean = −0.008236; SD = 0.006304; minimum = −0.041401; P5 = −0.019400; median = −0.007749; P95 = 0.001275; maximum = 0.012304. The difference was positive in 8,3575% of the simulations	Principal Spearman correlations: w_C = −0.945865; w_E = 0.554671; w_P = 0.193854; w_S = 0.150458; γ_1 = 0.063063	The comparison between 2024 and 2025 constitutes the principal source of ranking instability. The near-unit negative association with w_C shows that the result is governed primarily by the relative importance assigned to maintenance–investment complementarity, rather than by the parameters of the logistic function
Ordinal stability of the budget years	First position: 2023: 0 simulations; 2024: 183,285 simulations, or 91.6425%; 2025: 16,715 simulations, or 8.3575%. Expected rank: 2023: 2.9440; 2024: 1.0836; 2025: 1.9724 Complete rankings: 2024 > 2025 > 2023: 172,092, or 86.0460%; 2025 > 2024 > 2023: 16,715, or 8.3575%; 2024 > 2023 > 2025: 11,193, or 5.5965%. No other ranking was observed	The Monte Carlo standard error associated with the 91.6425% leadership frequency of 2024 is approximately 0.062 percentage points. Increasing the number of simulations would therefore have a negligible effect on the reported frequency, conditional on the same model and distributions	The ordinal evidence distinguishes computational precision from substantive validity. The ranking frequencies are estimated with high numerical precision, but they do not represent objective probabilities that one budget year is institutionally superior

http://dx.doi.org/10.21511/pmf.15(4).2026.02

Table A1 (cont.). Integrated evidence on institutional utility, ranking stability and model sensitivity

Analytical dimension	Quantitative evidence	Sensitivity and robustness evidence	Scientific interpretation
Preference structure associated with the leadership of 2024 and 2025	Unconditional means: w_C = 0.249609 and w_E = 0.250209 When 2024 ranked first: $E(w_C)$ = 0.263067 and $E(w_E)$ = 0.237148 When 2025 ranked first: $E(w_C)$ = 0.102046 and $E(w_E)$ = 0.393427 The conditional differences correspond to approximately –1.53 standard deviations for w_C and +1.49 standard deviations for w_E	The leadership of 2025 is associated with a reduction of 0.161021 in the mean weight assigned to C_t and an increase of 0.156279 in the mean weight assigned to E_t, relative to the configurations in which 2024 leads	The reversal between 2024 and 2025 reflects uncertainty regarding institutional preferences. The 2024 composition is favored when joint investment and sustainment receive material importance; 2025 is favored when potential execution dominates the evaluation
Effective discriminatory capacity of the components	Cross-year ranges: C_t = 0.051240; P_t^* = 0.002664; E_t = 0.123016; S_t = 0.013313. With equal weights of 0.25, the corresponding weighted ranges are: C_t = 0.012810; P_t^* = 0.000666; E_t = 0.030754; S_t = 0.003328	Although each component receives the same nominal weight, E_t produces approximately 46 times the weighted variation of P_t^*, while C_t produces approximately 19 times that variation	Equal nominal weights do not imply equal empirical influence. The ranking is driven predominantly by potential execution and the maintenance–investment component. Personnel adequacy is almost saturated around its maximum, while institutional stability discriminates only between 2023 and the other years.
Structural robustness of the maintenance–investment specification	Original specification: C_{2023} = 0.134146, C_{2024} = 0.141122, C_{2025} = 0.089882; utilities = 0.659761, 0.691554 and 0.683443; ranking 2024 > 2025 > 2023 Normalized specification: C_{2023} = 0.894604, C_{2024} = 0.724493, C_{2025} = 0.487873; utilities = 0.849876, 0.837396 and 0.782941; ranking 2023 > 2024 > 2025	The original CES component combines the joint budgetary magnitude of maintenance and investment with their relative balance. Normalization by the corresponding arithmetic mean isolates relative balance and reverses the complete ranking	The leadership of 2024 is parametrically robust within the original specification but not structurally robust to an alternative and substantively plausible definition of C_t. The principal source of model uncertainty is therefore the definition of the construct, rather than Monte Carlo error or marginal parameter variation

Note: Ct denotes the joint maintenance–investment component adjusted for complementarity; Pt* denotes personnel adequacy; Et denotes potential execution; St denotes modelled institutional stability; and Ut denotes aggregate institutional utility. SD denotes standard deviation and P5 and P95 denote the 5th and 95th percentiles. The positive-difference frequency is the proportion of parameter combinations in which the first year identified in the comparison obtained a higher utility than the second. Expected rank is the arithmetic mean of the ordinal position across the 200,000 simulations. Spearman coefficients measure monotonic associations within the simulated sample and must not be interpreted as causal effects. Simulated frequencies are conditional on the model specification, parameter intervals and probability distributions adopted.