PREDICTING COST GROWTH ON MAJOR DEFENSE ACQUISITON PROGRAMS THROUGH FORECAST MODELING

Dr. Charles A. Moore Jr., DEng in Engineering Management George Washington University , M.S. in Engineering Management Pennsylvania State University, B.S. in Technical Management American Military University, Madison, Alabama 35758, 256-361-7877, thecharlesamoore@gmail.com

ABSTRACT

Department of Defense (DoD) overspending and unpredictability through its major defense acquisition programs (MDAP) challenge the U.S. budget. This review of data internal and external to the program establishes correlations and predicts whether MDAP will meet the following year's budget. A forecasting model to assess MDAP factors to predict execution in future budgets contribute to reducing DoD unpredictability on the top programs and priorities. This research incorporated multiple independent and considered multiple dependent variables to find the best forecasting model and produce the most actionable predictions for MDAP annual performance. The independent variables were internal and external, and the dependent variables were specific to annual cost overrun. The results show new methods to program annual prediction with the external variables and provide an actionable model for risk reduction when considering annual program predictive performance.

Keywords: Business Management, Program Management, Defense, Engineering Management

BODY

Introduction

The U.S. Department of Defense (DoD) is the largest government agency in the United States. When activated, the DoD acquires programs with authority and guidance under DoDD 5000.01, The Defense Acquisition System, and DoDI 5000.02, Operation of the Defense Acquisition System. The Milestone Decision Authority (MDA) addresses all major capability acquisition procedures under DoDI 5000.85 and oversees all major defense acquisition programs (MDAPs).

The United States (U.S.) Government Accountability Office (GAO) office has now conducted 21 annual assessments of the Department of Defense's (DoD) weapon system acquisitions. In the 2023 report, of the MDAPs reviewed over half are reporting schedule slips and continued cost overruns are being experienced on an annual basis (U.S. Government Accountability Office [GAO], 2022).

In 2023, MDAPs accounted for $98.8 billion of the $276 billion defense budget (Office of the Under Secretary of Defense (Comptroller)/Chief Financial Officer, 2022). The 2020 U.S. Government Accountability Office (GAO) report showed that MDAPs exceeded initial operating capabilities (IOC) schedules by over 30% and the allocated budget by $628B, or 54%. (U.S. Government Accountability Office [GAO], 2020). Poor performance and the potential loss of over $100 billion of DoD buying power due to inflation indicate the need to consider additional variables for budget allocation (Bacon, 2022).

Over the years, the budgeting process has unexpectedly changed due to world economics. DoD officials have sought to combat changing globalization models, international policies, and unpredicted inflation while focused on advancing technology to keep on pace with international adversaries. With current world events such as the war in Ukraine, Chinese technology and military advancement, inflation higher than 6%, and trends similar to the 1970s, the DoD should have predictable, planned budgets.

The DoD has begun addressing inflation's impact on defense contracts while attempting to control the impact of the COVID-19 pandemic, which has been annotated contractually as force majeure. The Office of the Under Secretary of Defense provided guidance on inflation and economic price adjustments for all DoD contracts. The guidance for contracts varies based on contract type; newly negotiated contracts may require additional clauses for contractor protection to reduce unnecessary costs or risk to the federal government (USD-A&S, 2022). Bacon (2022) predicted that the DoD could lose over $100 billion in buying power due to inflation and budgeting processes.

More than 35% of the DoD spending plan focuses on programs with a history of overspending by 54%. Thus, there is a need for additional insight, research, and collaboration. Although abundant research has resulted in shifting methodology and program execution, these shifts have produced additional barriers to entry and contributed to poor execution. Extant reviews have focused on program specifics; internal-facing variables; and major triggers, such as Nunn McCurdy breaches.

This study presents a model that includes the external variables related to program-specific execution to provide additional insight into the influence of cost and schedule performance. Annual performance compared to the presidential allocated budget as a binary 1 for exceeding budget and 0 for meeting budget was the dependent variable. This model is a means of leveraging internal program data and the historical performance of the external variables to examine whether the allocated budget will be exceeded. Senior leaders should use the model to better plan budget allocation to reduce the strain of unpredictable cost growth.

Scope of Research

This research commenced with reviewing and collecting the data for all MDAPs between 2008 and 2023. The scope of variable identification will focus on the MDAP presidential allocation of budget by each independent MDAP by the fiscal year allocation. For independent variables we will look at internal and external variables across these identified programs. Internal variables will include drivers of program performance previous identified by research or by DoD criteria.

The external MDAP variables fall into company and environmental categories. Company details include the prime and major contractors holding the contract from the acquisition command in private and public sectors. The contractors are both U.S. or foreign entities. Economic scope is limited to the United States economy since the budget is U.S. driven and is U.S. Defense. Economic variables depend on the United States' economic and political positions. This scope ensures we captured DoD standards, historical research, and further expand variable consideration across standard accepted variables through industry.

Limitations

Program data is readily available and released publicly through SARs. Programs must provide information and an annual report once they become MDAPs; however, data analysis requires adjusting several concerns (Easterling, 2020). Over the period, 130 different MDAPs were used for data collection and model building. The MDAPs were further reduced due to incomplete data or program cancelation before a phase of measurable progress over time, which, for this research, was having a requested, planned, and actual budget for 2 years. Prime contractors are identified through SAR reports, however, in some instances there are multiple prime contractors listed. Independent variables for the predictive model looks at prime contractor company performance and for this model we remove contracts with multiple prime contractors to eliminate the complexity of multiple entity performance evaluations. Limitations are present based on GAO reports and SAR data based on thresholds for the reports being reviewed. Simply because a program is listed as an MDAP does not mean all data is readily available due to those thresholds and that the 2021 GAO reports were not produced due to the lack of funding data in the 2022 request.

Additional Literature

Schedule and cost management are critical to expectations and appropriate planning. The U.S. Armed Services have been part of the federal spending plan since the government's establishment. The U.S. Armed Services historically accounted for 1% of the gross domestic product (GDP). However, globalization, warfare, and World War II resulted in escalating defense spending. Since the 1920s, the United States has had heights of more than 40% of GDP allocation to defense spending (U.S. Government Spending, n.d.). With an estimated investment greater than $1.8 trillion for current and future MDAPs, significant time and review are spent on the defense spending plan (U.S. GAO, 2020). U.S. defense spending is significantly higher than many other nations. In 2021, the United States had an expenditure of $801 billion, more than the next nine countries combined on defense.

In 2022, the GAO presented the 20th annual assessment of the DoD Weapon Acquisitions Report. Over 20 years, there has been significant cost and schedule growth across MDAPs. Annual GAO reports provide necessary data. The 2022 GAO report included 86 MDAPs across all services and presented the changes in MDAPs, future MDAPs, and middle-tier acquisition (MTA) programs over the past 5 years. The literature review of MDAP programs found several key themes: Cost overruns, Schedule delays, Requirements management, and Testing and evaluation.

Since 1981, researchers have examined processes and policies and made suggestions regarding the largest part of U.S. spending. Arnold reviewed the budgeting process and how the purpose of defense policy is to ensure appropriate decisions and execution, indicating that focusing on defense spending across resources, not just capabilities, is the only path to success (Korb, 1981).

Easterling (2020) conducted a variable definition for regression analysis to predict Nunn McCurdy breaches and found that technology advancement and elements aligned to the technology readiness level (TRL) significantly impacted multiple phases of the contract lifecycle. Easterling (2020) found that technology readiness in Milestone B was the most important determinant for research, development, test, and evaluation cost growth and that technology elements are critical at or above TRL6.

Betreau and Hofbauer conducted a general assessment with factors outside the traditional contract variables, such as the company priming the contract along with the actual competition and contract type, and found a potential correlation between these variables (Berteau et al., 2010).

Although enacted annually, the budget process may begin up to 2 years before budget enactment. The annual defense budget program has spending categories of projected, planned, and actual. The *projected budget* is the planned budget requested 2 years before execution; the *planned budget* is the budget awarded and allocated the year before execution; and the *actual budget* consists of the costs incurred during the year of execution. Figure below shows the PPBE process for a 5-year plan across 4 calendar year categories. (U.S. DoD, 2017)

For major capabilities, reviews occur at five key points, indicating the milestones and capability to move to the next funding phase for a DoD resource; Milestone A, Milestone B, Milestone C, Initial Operating Capability, Full Operational Capability. GAO and SAR reports also discreetly look at schedules and standardized milestones. The Milestone of a program drives the lifecycle and funding allocation phases for a program as outlined in the DoD Financial Management Regulation (FMR).

From a DoD perspective, the purpose of Project RAND and the Air Force is to better anticipate costs and schedule challenges to address the DoD's issue of routinely underestimating MDAP costs. Project RAND found three cost growth factors and one schedule slip factor. The project focused on cost growth by unit cost and variables such as program phase, schedule slips in phases, concurrency, program type, technical design concepts, and type of technology and years (RAND, 2017).

McCubbins (1991) described how a divided government impacts defense spending and budgets, contending that division and party misalignment is a disadvantageous position for spending and reform. Alt and Lowry (1994) also contended that a divided government has fiscal and schedule consequences.

Worger and Jalao write about strategies for the Department of Defense acquisition process and suggested using their simulation to increase fidelity in policy changes to support the contract (Worger et al., 2016). The gap in Worger and Jalao research is that the research was simply done from a statistical perspective and not from a predictive modeling perspective.

McNicol et al. (2015) found that the budgets changed and costs increased during Milestone B (McNicol et al., 2015). The consideration provided specific to how funding cycles were within the Department of Defense and Congress is a great start to looking at how programs are impacted by external variables.

Baker introduced the Technology–Organization–Environment (TOE) framework for innovation adoption and implementation. The TOE framework could provide guidance for researchers and theorists to address innovation comprehensively (Baker, 2011). Scholars have used the framework to support innovation appropriation, use, and development (Bryan & Zuva, 2021).

The importance of the TOE framework for this analysis was that historical regressions and models have not included organizational or environmental data. Each has shown the importance of technology and technology maturation in the management framework.

Methodology, Data Collection, and variable Identification

The purpose of the research was to develop a model for supporting government contract programs and increasing the fidelity of DoD budget allocation. The research was focused on answering the multiple questions with one ultimate goal:
Leverage all research and metrics to predict contractor annual cost overrun against the DoD planned budget.

The analysis involved more than predicting the impact of variable types to understand the data. With an opportunity to further divide data and most accurately predict: First Overrun, Not first overrun, and Any/all overruns.

The method involved collecting data on exhaustive independent variables and variations of those variables for data analysis and model testing.

The research included quantitative data from annual GAO and SAR reports. The President's Budget presents budget allocations and constraints; however, it does not account for changes within program variables, such as quantity changes through execution years. The SAR reports provided critical data on the independent variables used for the analysis. SAR and GAO reports offer the most logical data for MDAP analysis. Although the SAR reports provided baseline data on the independent variable and dependent variable, there is a need for further research. Therefore, this analysis went beyond SAR data, which superficially focus on programs, to address the external factors in the program data categorized as entity and economic data.

Entity data include annual mandated business financial reports for publicly traded U.S. organizations. The 10-K reports provide a standardized and strict guideline for consistent methodology across different organizations for data collection. Economic data are collected through U.S. economic reporting to the global forum. The boundary of U.S. economic and entity data is in place to reduce differing laws and regulations and due to the view that DoD budgets are U.S. budgets, so consistently reporting information United States basis is used. The data resulted in the following entity variables for trading and public holdings: price per share, operating income, gross profit, and annual revenue.

These variables, categorized as program, economic, and entity, provided a novel opportunity to predict DoD spending. Additionally, this model included the variables with the most significant impact on cost growth in the SAR reports.

In addition to the program data from the SAR and GAO reports, the data collection included analyzing annual reports for the variables reported and tracked by program and government officials. The annual reports included the variables Milestone B growth, PDR growth, CDR growth, Milestone C growth, IOT&E growth, FRP growth, IOC growth, cost variance reporting, schedule variance reporting, program protested, and Nunn McCurdy breaches.

The collected data included the economic variables contributing to program cost growth. Government reports provided the economic variables specific to the U.S. economy. This research focused on the U.S. economic variables impacting DoD budget approvals and allocation to narrow the number of economic variables. The economic variables included in the study were Senate party control, House of Representatives party control, White House party control, and annual inflation.

MDAP planning occurs over multiple years and should reflect how the data is analyzed and associated independent variable data are considered. So, for each variable identified, there was a need to account for the prior 2 years. This variable and data collection approach produced a data matrix with 129 MDAPs (rows) and 917 internal and external variables from 2010 to 2023. Data collection produced more than 100,000 data points for consideration.

The actual cost should align with the President's Budget awarded to the contractor for the year of execution. However, the issue is unpredicted growth in cost versus the planned cost in DoD budgets. Therefore, the actual program cost was a dependent variable in the model.

For this model, programs that exceeded the cost for the year were labeled 1, and those that did not exceed the planned budget were labeled 0. This binary approach aligned with previous research for MDAP programs and was a suitable means of predicting project or program management.

To expand knowledge on how the overrun can be predicted the dependent variable was further tagged as what type of overrun occurred and the model was categorized into three instances; first overrun, not first overrun, and any/all overruns. Doing this is to understand if predictability of the type of overrun was significant in any way for modeling purposes.

Internal Independent Variables

The GAO and SAR report reviews included identifying the internal program variables. The internal program variables were consistently reported and expected from government reports of ACAT1 programs. Additionally, the variables were historically used to predict program execution and actual program spending.

All variables are collected specifically against the individual MDAP. The study included 2010–2023 data on the variables, with 2010 reports providing significant historical data for a wide range of programs. The three categories of internal variables were planned, requested, and actual. The 2010–2023 timeframe provided a snapshot-in-time view compared to previous data regarding DoD and GAO report organization. The Comptroller reports provided data on the discrete financials of individual programs with 3-year snapshots that included the next year's request, the current year's planned budget, the presidential authorized budget, and the prior year's actual budget. The study included collecting program milestone reviews and the variables affecting schedules and milestones compared to the initial procurement plan. Because costs and schedules correlate, the study included organizing the cost and schedule variables to find whether the year of schedule growth impacted costs. The variable was also maintained to find whether the schedule growth of a critical milestone impacted future cost growth. Finally, the study included internal variables related to the reporting and identification of program execution within the program office and DoD reviews. MDAP programs require earned value techniques for execution and management. However, they may not always have the requirement within the DoD to report cost and schedule variance once within the government reporting. Therefore, there could be a disconnect between contractor requirements and internal government reporting requirements.

The 129 MDAPs and program variables resulted in 22 internal independent variables with more than 15,000 data points. This study had a data set larger than other methodologies and analyses conducted to predict and review MDAP performance and execution.

External Independent Variables

The external variables were grouped and identified for analysis of program cost growth and the variables historically not considered when discussing specific program execution. These variables correlated with those in the GAO and SAR reports, such as the prime contractor. In this analysis, the prime contractor independent variable included entity- and contractor-specific data. The analysis also focused on the U.S. economy and budget for programs.

The entity variables collected were baselined for broader information about the entity to understand different ways to look at an organization. The GAO and SAR reports provided the prime contractor information. A piece of data for inclusion as an independent variable was whether the entity was public or private. Analysis occurred to find and allocate all prime contractors' public or private status against the MDAP data instance. Categorization as either a U.S. or foreign entity occurred for all prime contractors. The remaining entity independent variables emerged from the financial data provided by publicly traded companies in annual 10-K reports per the SEC. The 10-K is a comprehensive annual report that contains much more data than just annual reports distributed by organizations. The data collected from the 10-K reports included annual revenue, price per share, operating income, and gross profit. Therefore, the analysis included the financial data in a binary versus continuous for comparison to the prior year.

There were MDAP data collected for any given year in three instances (requested, planned, and actual). The entity financial data also covered 3 years for each annual instance of the program data (a year prior and 2 years prior). There were binary independent variables created for growth as well. The data review and collection produced 15 entity independent variables, one string variable, and 14 standardized continuous variables.

The second external variables considered were economic variables, which showed any predictive correlation between the U.S. economy and MDAP spending. Economic data affected both sides of the MDAP equation. The economic data included MDAP budgets in the presidential budget and entity execution of the contractual performance through policy and government requirements.
The study had a variable position on inflation based on the Federal Bank, which suggests that an inflation rate of 2% over a long period could impact the personal consumption expenditure price index (Federal Reserve System, 2020)

Additional data collected pertained to government representation in the House, Senate, and White House. The data underwent analysis for a change in government leadership for each of the three government branches. Because MDAP data have a range of 3 years, these independent variables also received annotations for a year prior and 2 years prior, consistent with the entity variable annotation. The review of the economic data resulted in 16 independent variables.

Creating Validated Data Set

Creating a data set from the independent variables involved expanding the instances from the 129 MDAPs. This process involved aligning all the data to the year of execution and examining the data set to look at each MDAP program by year as an instance against its planned performance. This step increased the instances from 129 to 1,677.

After creating the instances the execution year was used to separate the data into annual execution, data cleaning occurred to look at missing data across planned cost and actual cost. The purpose of the data cleaning was to ensure the instances in the final data set reflected the annual program data with planned and actual costs from an execution year. Data cleaning showed that the dependent variable of cost growth actual plan had complete information and did not lack variable dependency. Data cleaning by removing any variables related to program performance for the year predicted reduced the 1,677 instances to 642 across 55 attributes and more than 37,000 data points. This was reduced one further by removing the programs that had multiple prime contractors listed. This reduction was for clarity of all the prime contractor variables used and looking at the prime contractor performance. The 642 unique instances ended at 615 unique instances and more than 33,000 data points and 55 independent variables against the one dependent variable. Fifty-five attributes across 615 instances aligned with the commonly accepted 10-1 heuristic rule. The rule suggests that every independent variable in the model should have at least 10 instances in the dataset (Harrell, 2001).

Modeling

The goal of the research questions was to develop a predictive model for cost growth to understand how the internal and external variables correlated with the dependent variables. To do this we will use Pearson Correlation Coefficient along with statistical p value. We will further look into the confusion matrices of the modeling to ensure that we are increasing the value of the model to most impact predicting cost overruns.

One of the benefits of Pearson's correlation for a binary variable is the simplicity and ease of interpretation. The coefficient is straightforward to compute, with a relatively intuitive interpretation. It is given by the formula:

$$\mathrm{r} = \frac{\Sigma\big((\mathrm{x} - \mu\mathrm{x})(\mathrm{y} - \mu\mathrm{y})\big)}{\sqrt{(\Sigma(\mathrm{x} - \mu\mathrm{x})^2)} * \sqrt{(\Sigma(\mathrm{y} - \mu\mathrm{y})^2)}}$$

where Σ represents the sum, μx and μy are the means of x and y respectively, and sqrt denotes the square root.

Statistical significance is important for correlation. This study focused on the correlation of variables at .1 alpha considering researchers may use an alpha level of 0.1, particularly in exploratory studies (Levin, 1999).

Modeling consideration occurred with a range of machine learning models to understand the most beneficial for the data set and also the consideration of the dependent variable being a binary variable. The algorithms considered and run for modeling were: Random forest, Logistic regression, Stacked ensemble base random forest with logistic regression, Stacked ensemble base logistic regression with random forest, K-Nearest Neighbor, Random Tree, and Decision Stump.

Models were run with all collected variables and reduced to attribute significance and gain better predictive results within the previously described bounds. This reduction occurred within WEKA. The reduction involved evaluating the worth of a subset of attributes by considering the individual predictive ability of each feature and the degree of redundancy between them. In addition, the reduction occurred with subsets of features highly correlated with the class and had low intercorrelation. The modeling and reduction of attributes included greedy stepwise evaluation criteria for searching backward for the best attribute combination. With increased interpretability, researchers and practitioners can more easily understand the underlying relationships between the predictors and response variables (Hastie et al., 2009).

Model evaluation can occur via multiple techniques. The techniques in this study were *k*-fold cross-validation and the 66/33 train split.

In the context of predictive modeling, 10-fold cross-validation can help in model selection by providing a robust estimate of the model's generalization ability. It can assist in identifying the optimal complexity of the model, thereby preventing overfitting or underfitting (Hastie et al., 2009). This study uses the 10-fold cross validation as one of the evaluation criteria to determine the best model.

The 66/33 split ratio is commonly used in predictive modeling, as it allows models to learn from a large amount of data, while still maintaining a substantial portion for testing. This ratio helps in avoiding overfitting, where the model performs well on the training data but poorly on unseen data. It also assists in ensuring the model's generalization ability, which is vital in predictive modeling (Kohavi, R., 1995). This study uses 66/33 split ration as one of the evaluation criteria to determine the best model.

This study further looks at performance metrics as an output of the confusion matrix to determine best model. Confusion matrix performance metrics include:

- Accuracy: The proportion of correctly classified instances out of the total instances.
- Precision: The proportion of true positive predictions out of all positive predictions.
- Recall (sensitivity): The proportion of true positive predictions out of all actual positive instances.

For clarity and consistency True Positives represent correctly predicted positive instances, True Negatives represent correctly predicted negative instances, False Positives represent incorrectly predicted positive instances, and False Negatives represent incorrectly predicted negative instances.

The confusion matrix for each model underwent consideration due to the importance of predicting the overrun and the goal of increasing the precision of *Yes–True Positive* while remaining attentive to overall *Yes* recall. It was necessary to avoid sacrificing precision due to recall when considering the impact of predicting a *Yes* on DoD reaction. A high false positive rate could result in unnecessary resources expended to prevent a cost overrun.

The goal was to develop a more accurate model for predicting cost overruns than the current planned allocation of the DoD and presidential budget when considering true positives would be mitigated based on the DoD response.

Reviewing the instances we went to review the total number of overruns that are within our dataset for consideration. Of the 615 instances 167 are overruns as compared to planned cost. The equation we are using to determine program performance will be:

$$\frac{Total\ in\ budget}{Total\ Instances}$$

This means currently 27% of programs overrun allocated and predicted planned costs across the Department of Defense Major Defense Acquisition Programs as considered by this data set. Our baseline for modeling should be able to predict these overruns and better predicted execution within the planned cost to reduce the overrun and unexpected budget necessary for defense programs.

With the consideration provided for type of overrun the same methodology was applied to perform better than the dataset itself. Meaning that if the model is used and it would predict an overrun than there would be an expectation that resources or additional considerations to that prediction would be allocated to further reduce the risk of overrun. With that methodology our models precision of overrun prediction must stay above 50% and the overall performance of the dataset would then be adjusted for within budget execution of those true positive predictions creating a lower amount of programs that overrun allocated budgets.

Data was split into three models to help understand how the variables interacted with the independent variable.

Model A included only internal program metrics for consideration against the cost growth dependent variable.

Model B only included the external program metrics for consideration against the cost growth dependent variable.

Model C included all variables and performed across numerous algorithms and validation methods and further divides for the first overrun, next overrun, or any cost growth-dependent variable.

Results

The data collected underwent multiple reductions and iterations for accuracy and to cover all factors for consideration. Through the modeling described this resulted in a data set for Model A and Model B had 642 unique instances and 55 unique independent variables. The model involved splitting the 55 unique independent variables for Model A and B. For Model C finalized from 642 to 615 unique instances.

The development of Model A included the entire internal independent variable data set. There were 24 internal variables used for Model A. Pearson‘s correlation and statistical correlation was used to identify the internal program metrics for better cost execution in Model A.

The study's results were consistent and expanded beyond with previous research. The quantities and changes in quantities across the MDAP programs affected cost adjustments. The results also showed the statistical significance of historical performance. Historical growth had the highest Pearson's correlation and lowest *p* values, while program protested had a negative correlation. The following inferences were made:

- Any historical growth (r = .108, p = .006) indicating a direct correlation to growth.
- 2 years prior qty growth (r = .089, p = .025) indicating a direct correlation to QTY growth.
- Year prior cost growth (r = .080, p = .043) indicating a direct correlation to the prior year of cost growth.
- Program protested (r = -.076, p = .053) indicating a negative correlation to the program being protested.
- Any historical cost growth (r = .066, p = .097) indicating a direct correlation to any historical cost growth.

Model B included the entire external independent variable data set. There were 32 external variables for consideration on Model B. Pearson's correlation and statistical correlation occurred to identify the external program metrics for cost execution.

These results showed the significance of external variables for cost growth of plan execution. MDAP cost growth significantly correlated with the external entity and economic variables. The White House change had the highest Pearson's correlation and the lowest *p* value, with year prior inflation above 2 slightly behind it. Further review showed negative correlations with multiple entity and economic variables.

- WH change (r = .203, p < .001) there was a positive correlation between a U.S. White House change in control.
- Year prior above 2 (r = .173, p < .001) there was a positive correlation between U.S. inflation above 2% the year prior
- Entity gross profit change year prior versus 2 years prior (r = .123, p = .002) there was a positive correlation between the growth of entity gross profit increasing the year prior versus 2 years prior
- House 2-year prior change (r = -.107, p = .007) there was a negative correlation between the U.S. House control change 2 years prior
- House change (r = .096, p = .015) there was a positive correlation between a change of control in the House of Representatives

- WH year prior change ($r = .096$, $p = .015$) there was a positive correlation between a change in party control at the U.S. White House the year prior
- WH ($r = -.077$, $p = .051$) there was a negative correlation with Democrat White House control. Said differently there was a direct correlation between Republican White House control
- Entity location ($r = -.077$, $p = .053$) there was a negative correlation, pushing further when international headquarters was part of the variable.
- Entity operating income change year prior versus 2 years prior ($r = .074$, $p = .060$) there was a positive correlation between the entity operating income increasing the year prior versus 2 years prior
- Entity price per share change year prior versus 2 years prior ($r = .067$, $p = .088$) there was a direct correlation between the growth of entity price per share the year prior versus 2 years prior
- Entity operating income change 2 years prior versus 3 years prior ($r = -.065$, $p = .099$) there was a negative correlation between entity operating income growth 2 years prior versus 3 years prior and program cost growth. Thus, the results showed a positive correlation between the loss of operating income and the annual cost growth of the program.

Model C included the entire independent variable data set. Multiple regression models for the external program metrics occurred to better predict cost execution.

Model C focused on cost growth and included 615 instances across all 55 independent variables. The runs conducted for the regression models used were:

- Random forest
- Logistic regression
- Stacked ensemble base random forest with logistic regression
- Stacked ensemble base logistic regression with random forest
- K- Nearest Neighbor
- RAND Tree
- Decision Stump

Additionally, the model underwent a *k*-fold and a test train split performed as 66/33. An additional measure was searching for the most significant variables for the predictive model by reducing the independent variables with a backward greedy stepwise regression and running through the above scenarios again.

The goal for the model was a precision above 50%. The false negative was any program that was predicted to not overrun but with an actual budget overrun. The result was improving the model's performance for budget allocation to account for false negatives over the entire number of instances for a performance increase. Further the chosen model will be the one with the highest performance to allocated budget.

Modeling for any cost growth ranged from 62.67% predictability to 76.56% model predictability.

The best model was a logistic regression model with independent variables reduced based on backward greedy stepwise regression and run with *66/33* test train split. After reducing the independent variables, the variables in the model were:

- Prime contractor
- Entity gross profit change year prior versus 2 years prior
- Requested QTY
- Inflation year prior above 2%
- White House change

The model showed the connection between internal and external variables in MDAPs. The model included external entity and economic variables. The final model output provided a 76.55% correctly classified model. The final model included 615 instances, when reduced by 66/33 it had 209 with 50 showing an overrun of actual and planned costs resulting in prior to model being run a performance of 76.07%.

- The model reduced overruns from 50 to 34.
- The model predicted 31 overruns, of which 16 were actual overruns, resulting in a 51.6% precision.
- Ultimately, the model could increase annual performance by 7.66% or a total performance of 83.73%.

Total Instances	209		Confusion Matrix	
Correctly Classified	160		TN	FP
Incorrectly Classified	49		FN	TP
Model Performance	76.55%			
			Confusion Matrix	
Positive Recall	32.00%		144	15
Positive Precision	51.60%		34	16
	Original	After Model		
Overruns	50	34		
Within Budget %	76.08%	83.73%		

Modeling for the next cost growth ranged from 50.28% predictability to 84.07% predictability.

The best model was a k-nearest neighbor with independent variables reduced via backward greedy stepwise regression and run with Kfold cross validation of 10. After reducing the independent variables, the variables maintained in the model were:

- Military department
- Entity gross profit change year prior versus 2 years prior
- Requested QTY
- Inflation year prior above 2%
- House two year prior change
- White House change

The model showed the connection between internal and external variables in the MDAP. The model included external entity and economic considered variables. The final model output provided a 72.17% correctly classified model. The final model included 521 instances, of which 149 had actual cost overruns compared to the planned costs which leads to an original performance of 71.40% to budget and after model use a performance of 84.07%.

- The model reduced overruns from 149 to 83.
- The model predicted 128 overruns, of which 66 were actual overruns, resulting in 51.60% precision.
- Ultimately, the model could increase annual performance by 12.67%.

Total Instances	521		Confusion Matrix	
Correctly Classified	376		TN	FP
Incorrectly Classified	145		FN	TP
Model Performance	72.17%			
			Confusion Matrix	
Positive Recall	44.30%		310	62
Positive Precision	51.60%		83	66
	Original	After Model		
Overruns	149	83		
Within Budget %	71.40%	84.07%		

Modeling the MDAP internal and external variables and exploring multiple algorithms provided more accurate predictions of MDAP performance with planned budgets. With an increase of performance greater than 12% and sufficient precision, the model could be a suitable means of predicting program overruns.

Discussion and Conclusion

The most revealing part of the praxis analysis was the ability to improve predictions of MDAP program spending by focusing on internal and external MDAP variables. Additionally, removing the first overrun in a program resulted in increased prediction of program execution falling within the allocated budget. Whether predicting any overrun or the next overrun, the best-performing models included the internal and external variables. This finding indicates that external program variables with economic and entity data could contribute to planning and program performance predictability when allocating DoD MDAP budgets.

The largest adjustment from the plan on this research was the adjustment from looking at the MDAP life as the instance to adjusting instances based on year of performance to create a larger data set affording more learning across annual execution. This adjusted view of the MDAP data allowed for historical performance to be contributed to current-year actual expectation to planned budget execution. This adjustment was a considerably different view than previous MDAP planning and research. The model could provide the opportunity to address cost overruns sooner when planning and distributing MDAP budgets.

Model A showed a correlation between internal variables and program measures when predicting cost overruns.

Model B showed the correlation between external variables and program measures when predicting cost overruns.

Model C increased the predictability of contracts aligning with the annual presidential budget for the DoD. Model C showed that the most significant independent variables for predicting overruns were a combination of internal and external variables, including economic and entity variables.

The models in this study provide better predictions of MDAP budgets. This study's model addressed the need to consider additional variables in budget allocation due to poor MDAP performance and the potential impact of inflation.

The research showed the statistical significance of numerous internal and external variables when predicting whether an MDAP will fall within the allocated budget.

Through modeling we found the most contributing variables were across all identified categories of both internal and external variables but additionally across the economic and entity external variables. Of the six variables ultimately used to model annual cost growth for MDAP programs 4 were external variables while only 2 were internal variables.

The ultimate goal was to answer: Can a forecasting model with external and internal variables more accurately predict adherence to the President's Budget FY plan for an MDAP?

Through combination of data and variables analyzed we see that the model creates a k-nearest neighbor predictive model that predicted greater than 12% better planned costs adherence for DoD MDAP programs whose variables are:

- Military Department
- Entity Gross Profit Change year prior vs two years prior
- Requested QTY
- Inflation year prior above 2.0%
- House two year prior change
- White House Change

The overall prediction of the model is 72.17% of the time accurate to historical execution. Further, the model predicts performance within budget 84.07% of the time vs the current state of instances which was 71.40%

The 2 internal variables of military department and requested quantity shows that the branch in control of the contract is a factor for predicting cost growth and could be further expanded on if individual studies were done on changes in branch control or of significant ideology shifts of technology or how that branch performs defense to be a part of the broader DoD defense structure. The internal variable requested quantity aligns to previous research that quantity impacts cost but adds that the original requested quantity could be more impactful to determining cost overrun.

The four external variables significantly add to the body of knowledge and how program cost is predicted showing that there are influences and factors far outside program controls impacting program growth. The external entity independent variable Profit Change year prior vs two years prior shows us that historical performance of the entity in addition with the other variables is most useful to predict program cost growth. This drives directly to business performance measures driving program resources and ability to perform contract requirements. This further shows that resourcing of programs and impacts of business measures is not felt by the DoD market the year the entity is impacted but rather a year or longer after the business results are published. This speaks to the impacts that program managers face when executing to customer requirements while working to meet business demands based on the broader entity or economic adjustments.

The external economic variables were spread across political scene and economic performance. Inflation the year prior above 2.0% shows that with unpredicted inflation the impacts to programs become more predictable. This could be across the business and entity impacts but most likely due to the broader impact the inflation has on all the DoD market and the ripple across the industrial supply base. The political independent external variables were both the House and White House driven. For the House two years prior change this speaks directly to the budgeting cycle of the Department of Defense but also speaks to program performance being planned out prior to budget being allocated and how changes in budgets during the cycle will impact the execution and adherence to budget. Finally, White House change being part of the final model shows the impact that white house control has on performance of programs. This could be for a multitude of reasons varying from instantly seen policy changes and methodology shifts in DoD priorities but also speaks to how business and programs react to the execution of programs during years when the white house control changes.

We further see that through looking across three years of historical data both internal to the program and external we use data across all three years to build the most predictive and best performing model. For the year of execution the variables that are used in the final model are military department and white house change. For variables that happen the year prior to execution we see Entity Gross Profit Change year prior vs two years prior and inflation the year prior being above 2.0%. This shows us the both economic and entity performance the year prior to program execution is impactful. Concluding with variables that were two years prior to program execution; requested quantity and House two years prior change in control. This shows us that the plan of the program executing quantities and the political shift in the House impact program performance in the future.

Potentially most notably we see that the inclusion of variables not just across internal and external variables but also across 3 years of execution, which is how DoD budgeting is conducted, all are included in the final model and speak directly to the performance of a program performance.

Conclusions and Significant Contributions

- Internal program data from the GAO, SAR, and Comptroller reports contributed to the model's significance (Military Department, Requested QTY). There is a need for stronger consideration of variances within these reports year over year.
- The ability to better predict the next overrun compared to any overrun suggests a program's first overrun is harder to predict and model. Increased awareness after a program has an overrun could be a way to improve annual budget adherence.
- MDAPs have annual reports, with special reports created annually based on thresholds. However, there are gaps in the yearly data due to changes in reporting and thresholds. After reaching an MDAP threshold, the MDAP reporting and reviews should continue to avoid gaps in the data.
- The Model C results regarding entity gross profit change year prior versus 2 years prior, inflation year prior above 2%, House two years prior change, and White House change suggest the importance of these external variables for current year program execution.
- The model in this study could predict whether the annual cost performance will exceed the allocated budget with a *yes* or *no*.
- Entity variable of entity gross profit change year prior versus 2 years contributes to predicting cost adherence to presidential budget MDAP allocation.
- Combining internal and external independent variables produced a model for predicting adherence to presidential budget MDAP allocation and annual performance.

Future Research

Despite this study's significance and contribution, many areas still require significant improvement. Future researchers could build on this study's internal and external variables to predict the performance of programs other than MDAPs. An example of this expansion could be looking at the addition of political power across the House, Senate, and White House and see if there is an impact when a larger majority is in control or when the White house and another branch is in control together. Such research could provide additional instances and data and find additional correlations. A larger sample size could also show the importance of different variables at different program levels. Finally, a suggestion for increased testing across different methodologies or expanding the algorithms tested could provide for new insights. Although, the inclusion of stacked algorithms showed promise ultimately K-nearest neighbor was the chosen model. With further expanding the specific tests and potentially even the nuances of each test there could be additional information gained or predictability and performance enhancements.

References

Alt, J. E., & Lowry, R. C. (1994). Divided government, fiscal institutions, and budget deficits: Evidence from the states. *The American Political Science Review*, *88*(4), 811–828. https://doi.org/10.2307/2082709

Armed Forces, 10 U.S.C. § 2430 (2016).

Bacon, P. (2022, March 28). *The $773 billion question: Inflation's impact on defense spending*. McKinsey & Company. https://www.mckinsey.com/industries/aerospace-and-defense/our-insights/the-773-billion-question-inflations-impact-on-defense-spending

Baker, J. (2011). The Technology–Organization–Environment framework. In Y. K. Dwivedi, M. R. Wade, & S. L. Schneberger (Eds.), *Information systems theory: Explaining and predicting our digital society* (Vol. 1, pp. 231–245). Springer. https://doi.org/10.1007/978-1-4419-6108-2_12

Berteau, D., Hofbauer, J., Sanders, G. F., & Ben-Ari, G. (2010, April 30). *Cost and time overruns in major defense acquisition programs*. Defense-Industrial Initiatives Group Center for Strategic & International Studies. https://apps.dtic.mil/sti/citations/ADA530059

Bryan, J. D., & Zuva, T. (2021). A review on TAM and TOE framework progression and how these models integrate. *Advances in Science, Technology and Engineering Systems Journal*, *6*(3), 137–145. https://doi.org/10.25046/aj060316

Easterling, C. A., III. (2020). A method for the prediction of Nunn-McCurdy breaches in Department of Defense acquisition programs (Publication No. 28088424) [Doctoral dissertation, The George Washington University]. ProQuest Dissertations and Theses Global.

Federal Reserve System. (2020, August 27). Why does the Federal Reserve aim for inflation of 2 percent over the longer run?. Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/faqs/economy_14400.htm

Harrell, F. E. (2001). *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*. Springer.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

Kohavi, R. (1995, August 20-25). *A study of cross-validation and bootstrap for accuracy estimation and model selection* [Conference presentation]. 14th International Joint Conference on Artificial Intelligence, Montreal, Quebec, Canada.

Korb, L. F. (1981). [Review of the books *Defense Politics: A Budgetary Perspective*, by A. Kanter, and *From Guns to Butter: Technology Organization and Reduced Military Spending in Western Europe*, by B. Udis]. *Armed Forces & Society*, *7*(2), 317–321.

Levin, J. (1999). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology*, *149*(12), 1079–1081. https://doi.org/10.1093/aje/kwk052

McCubbins Matthew D. 1991. “Party Governance and U.S. Budget Deficits: Divided Government and Fiscal Stalemate.” In *Politics and Economics in the Eighties*, edited by Alesina Alberto, Carliner Geoffrey. Chicago: University of Chicago Press.

McNicol, D. L., Burns, S. K., & Wu, L. (2015, May). Evidence on the effect of DoD acquisition policy and process and funding climate on cancellations of major defense acquisitions programs. Institute for Defense Analyses. https://apps.dtic.mil/sti/pdfs/AD1001295.pdf

Office of the Under Secretary of Defense Comptroller. (2022, April). *Defense budget overview*. U.S. Department of Defense. https://comptroller.defense.gov/Portals/45/Documents/defbudget/FY2023/FY2023_Budget_Request_Overview_Book.pdf

U.S. Department of Defense. (2017, August 29). *The planning, programming, budgeting, and execution (PPBE) process* (DoD Directive 7045.14). https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/704514p.pdf

U.S. Department of Defense. (2022a, July 28). *The Defense Acquisition System* (DoD Directive 5000.01). https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/500001p.pdf

U.S. Government Accountability Office. (2020, June). *Defense acquisitions annual assessment: Drive to deliver capabilities faster increases importance of program knowledge and consistent data for oversight* (GAO-20-439). https://www.gao.gov/assets/gao-20-439.pdf

U.S. Government Accountability Office. (2022, June). *Weapons systems annual assessment: Challenges to fielding capabilities faster persist*. https://www.gao.gov/assets/gao-22-105230.pdf

U.S. Government Spending. (n.d.) *History with charts*. https://www.usgovernmentspending.com/defense_spending_history

Worger, D., Jalao, E. R., Wirthlin, J. R., Colombi, J., & Wu, T. (2016). Intervention strategies for the department of defense acquisition process. *The Journal of Defense Modeling and Simulation*, *13*(2), 139–151. https://doi.org/10.1177/1548512914548615