JOURNAL OF PUBLIC PROCUREMENT, VOLUME 5, ISSUE 2, 187-209 2005

A REGRESSION APPROACH FOR ESTIMATING PROCUREMENT COST

Gary W. Moore and Edward D. White III*

ABSTRACT. Cost growth in Department of Defense (DoD) weapon systems continues to be a scrutinized area of concern. One way to minimize unexpected cost growth is to derive better and more realistic cost estimates. In this vein, cost estimators have many analytical tools to ply. Previous research has demonstrated the use of a two-step logistic and multiple regression methodology to aid in this endeavor. We investigate and expand this methodology to cost growth in procurement dollar accounts for the Engineering and Manufacturing Development phase of DoD acquisition. We develop and present two salient statistical models for cost estimators to at least consider if not use in mitigating cost growth for existing and future government acquisition programs.

INTRODUCTION

An ongoing problem for over three decades, cost growth during the acquisition of major weapon systems concerns not only those who work in the acquisition environment, but also the members of Congress and the general public. According to reports by the General Accounting Office (GAO), RAND, and the Institute for Defense Analysis, the average cost growth in major defense acquisition programs ranges anywhere from 20 – 50 % (Calcutt, 1993, p. i). This fiscal escalation in major acquisition programs adversely impacts the Defense Department, the defense industry, and the nation.

----

* *Gary W. Moore, Captain, US Air Force, is a cost analyst with the Air Force Cost Analysis Agency in Washington, DC. Edward D. White III, Ph.D., is an Associate Professor of Statistics, Department of Mathematics and Statistics, Air Force Institute of Technology, Wright-Patterson AFB, Ohio. His teaching and research interests are in design of experiments, linear and nonlinear regression, and statistical consulting.*

The Department of Defense (DoD) coined the phrase "realistic costing" for the current reform being undertaken in the defense acquisition community. "Under the new costing approach, the Pentagon will adopt program estimates developed by the Cost Analysis Improvement Group (CAIG[1]) in conjunction with a service estimate (Grossman, 2002, p. 2)." Realistic costing utilizes the CAIG's cost estimating expertise to provide higher quality estimates. DoD's dedication to realistic costing contributed significantly to the cancellation of the Navy Area missile defense program, sending a strong message to the acquisition community. If managers overrun their budget and breach the Nunn-McCurdy law, their program will be terminated (Grossman, 2002, p. 2).

For managers to understand and to contain cost growth, they must identify and control the root causes of cost growth. Program managers often resort to a process known as "buffering" in order to increase the accuracy of the baseline estimate and to limit the program's likelihood of incurring cost overruns. Buffering of an estimate entails assigning a cost estimate (dollar value) to each of the cost risks, e.g. additional engineering effort because of a new weapon system, avionics package, or stealth technology. The plan is to have sufficient funds available if/when the risk comes to fruition so that the program does not have to request additional funding. In the past, costs have been assigned to these risks; however, they have been shown to be sometimes underestimated. According to McCrillis (2003), who presented the conclusions of a ten-year study by OSD CAIG (Office of the Secretary of Defense Cost Analysis Improvement Group), procurement cost growth has occurred primarily because of optimistic learning curves.

An example of a current acquisition program struggling to keep costs under control is that of the F/A-22. The production quantity of the Air Force's newest air-to-air fighter has fluctuated considerably over the past decade in the attempt to maintain some modicum of control over increased cost growth. Because of rising program costs, the F/A-22 program has reduced the number of initial desired aircraft from 750 to 658 in 1991 to a more recently lowered number of 276 in 2002. In 1997 and 2001, the DoD conducted reviews of the F/A-22 Raptor program. During these reviews the Air Force attributed estimated production cost growth to increased labor, airframe, and engine costs. These factors totaled almost 70 percent of the overall cost growth (GAO, 2003).

Although some aspects of cost growth will always be hard to control or even to predict (e.g., the political arena), accurate estimates of costs to assign to risks remain a real challenge. This process necessitates that the manager accurately identify risks related to potential cost growth in program estimates and assign appropriate dollar values to these risks. While ultimately responsible for their programs, managers rely on the cost estimating community, which is made up of engineers, mathematicians, financial analysts, and acquisition personnel (procurement specialists), to assign accurate dollar values to specific risk factors and include these dollar amounts in the cost estimate.

Cost estimators determine and assign dollar amounts to risk factors by utilizing a vast assortment of tools. Cost estimators oftentimes use subjective means, such as expert opinions, for making these dollar assignments. When available, the estimator may utilize more objective methods, such as gathering historical data and comparing analogous systems. If possible, the analyst should group historical cost growth data into different categories and then analyze these categories to determine if different types of cost growth have different and distinct predictors. Statistical regression techniques prove useful for determining such relationships provided they are based upon sound framework and past research.

Once such historical encapsulation is that by Sipple, White and Greiner (2004) who conducted an exhaustive review of all cost growth studies performed during the past ten years. From this review, they gained valuable insight into the root causes of cost growth, which consists of funding increases in research and development, procurement, or operating and support. They also found extensive amounts of research devoted to establishing predictive relationships and determining predictor variables. For example, their consolidated review revealed that the average production cost growth is 19 percent, that the urgency of the program, difficulty of technology, and degree of testing affect cost growth, and that a relationship between cost growth and schedule growth in both the development and the production phases exist. That is, if there is cost growth in one aspect, there is a likelihood that the other one will also realize increased growth.

From this foundational framework, White, Sipple and Greiner (2004) assembled an extensive database with over 70 predictor variables from a collection of data gathered from the Selected Acquisition Reports

(SARs) on 115 major acquisition programs from 1990 to 2000. From this database, they constructed regression models aimed at predicting Engineering and Manufacturing Development (EMD[2]) cost growth directly related to engineering changes[3]. They found that combining logistic and multiple regression techniques most accurately represented the projected cost growth without violating the underlying regression assumptions.

Additionally, White, Sipple and Greiner (2004) suggested that the usefulness of this two-step approach might extrapolate beyond predicting engineering cost growth during the EMD phase of the acquisition life cycle. Their research hinted that the methodology might be extended to other cost areas. While they focused on engineering cost growth, we investigate the feasibility of this joint approach in modeling procurement cost growth during EMD. Much of the performed research that we annotate in this article mirrors the efforts carried forth in White, Sipple and Greiner, although we do allow for more complicated regression models via interactions of predictors, which we discuss in detail later.

METHODS

We use the SARs database as the sole source for cost variances and other information included in this analysis as in White, Sipple and Greiner (2004). This database contains historical, schedule, cost, budget, and performance information for major acquisition programs from all military services. Therefore, the programs listed in the SARs consistently represent programs with high-level government interest[4]. The SARs provide cost variance data in both base year and then year dollars. We use base year dollars and then covert to 2002 dollars for analysis, making comparison across programs more feasible. Although, SARs record cost variances in seven different categories, we focus exclusively on total procurement cost variance during the EMD phase, because most dramatic program changes occur during this phase. We do not focus specifically on one targeted cost category because most cost estimators are only concerned with overall cost growth. [Note: this remains an active research area, and we plan to address individual cost variance categories in other articles].

White, Sipple and Greiner (2004) constructed a database that contains SARs data from 1990 through the summer of 2000. Because this paper's effort follows theirs, we had access to their established database. We verified the existing information by checking past SARs and then updated the newest/most recent data by program in the SARs database to capture recent trends. Specifically, the latest SARs at our disposal at the time of this research were from December 2001. Thus, our data collection efforts begin with those SARs and work backwards through the summer of 2000. These reports are then incorporated into the previous database, resulting in a record 1990 through 2001.

We use the same predictor variables as White, Sipple and Greiner. A complete list of these predictors can be found in the Appendix along with a description of each. The continuous variable classification denotes a quantitative measurement variable, while a binary classification is reserved for the typical, dummy categorical variable. Similar to what White, Sipple and Greiner (2004) encountered, we learn that when we incorporate into our model the EMD maturity variables[5] that use Initial Operational Capability (IOC) or First Unit Equipped (FUE) computations, we face a scarcity of data points. Consequently, this limits the potential use of these as predictors, but not necessarily their potency in incorporating within predictive regression models. In fact, we demonstrate that one of these variables is highly predictive of procurement cost growth[6] in the EMD phase of a weapon system.

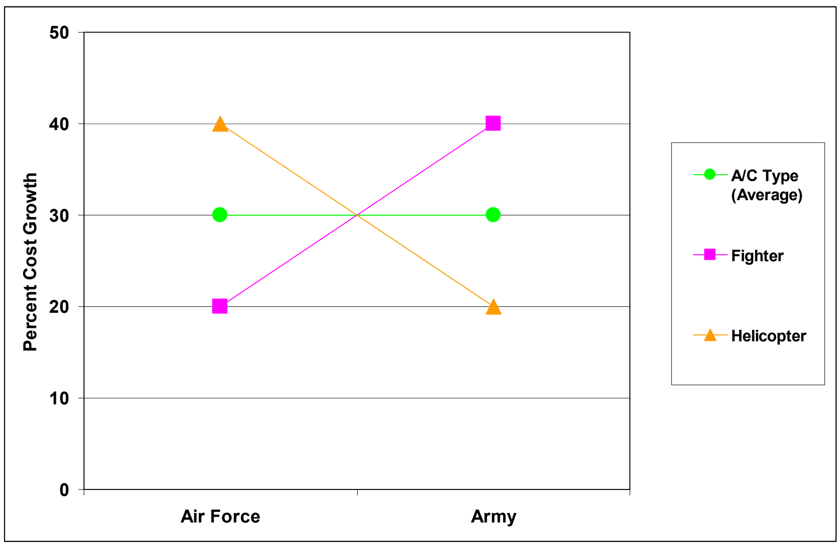
We concern ourselves with two different response variables – one that indicates if procurement cost growth occurs and another that expresses the degree to which procurement cost growth occurs. The first of the two, we express as a binary variable where the value '1' means that we estimate a program will have cost growth in procurement dollars, while the value '0' means that it will not. We call this variable *Procurement Cost Growth.* In order to construct the most useful model possible, we decide that the second response variable should be the percentage of procurement cost growth rather than just relative cost growth. This format applies equally well to programs with both large and small acquisition costs. We call this second response variable *Procurement Cost Growth %*, and we define it as the difference between the Current Estimate and Development Estimate, divided by the Development Estimate. [Note: the current estimate may or not may be the same as the initial contract award. The current estimate may have been revised/rebaselined since the program's inception.]

To determine if our data falls in line with what White, Sipple and Greiner (2004) discovered, we preliminarily investigated our response variable, which could consist of positive, negative, or even no cost growth. We soon discovered that a discrete point mass, representing approximately 20% of the data, lay at zero. A basic assumption of ordinary least squares (OLS) regression requires that the response variable be from a reasonably continuous distribution without such 'spikes' of point probability. As such, our findings mirrored what White, Sipple and Greiner realized. Hence, our methodology and subsequent analysis follow precisely what they established. Therefore, we duplicate those said procedures, which we now briefly review.

Before initiating any regression analysis, we randomly set aside approximately 20 percent of our data (25 out of a total of 122 programs) for model validation and to ensure that we construct a robust statistical model. Afterwards, we 'split' the remaining 80% data into two pieces. Our first cut involves coding the data into 1's and 0's. We code each program that incurs cost growth with a '1' and each program that has either no cost or negative cost growth with a '0.' Since an estimator would not realistically assess negative cost growth in an estimate, we do not consider negative cost growth in our model. As did White, Sipple and Greiner, we then utilize logistic regression to analyze this discrete distribution since this particular type of regression is appropriate to model binary outcomes, that is those usually coded '0' and '1' (Neter, Kutner, Nachtsheim, and Wasserman 1996, p. 567). Logistic regression has become the standard method for regression analysis of dichotomous data in many fields, especially in the health sciences (Hosmer & Lemeshow, 2000, p. vii). The reader is referred there for further background if needed.

Our second, and last cut so to speak of the data, involves grouping just those programs that have positive cost growth. The idea is to have available a model that will predict the percentage amount of cost growth given the logistic model's cue suggesting the likelihood of procurement cost growth. This resultant second pool of data is relatively continuous and hence OLS regression becomes the standard analysis to employ. Our multiple regression efforts focus not only on individual variables, but also include logical interactions between variables that may enhance their predictive relationships. In explaining interactions, we present the following scenario depicted in Figure 1. If interactions are not considered, then the center line shows the amount of cost growth (30%) associated with aircraft type across Air Force and Army. When we consider interactions, we find that the cost growth varies depending on both the aircraft type and the lead service involved (i.e. 40 percent for Air Force helicopters).

FIGURE 1
Statistical Interaction between Two Model Variables



RESULTS

Logistic Regression

The immense number of possible predictor variable combinations makes finding a true "best" logistic model an unattainable goal. So, we set out to produce the most predictive model possible within our resource constraints. Given the enormity of exploring all of the possible combinations, we narrow our predictor combinations to only those that show the most promise as we progress from simple to more complicated models. We begin by regressing all one-variable models and recording the results. From these findings, we select the ten "best" one-variable models and regress all possible two-variable models that stem from each of those models. Next, we select the nine models that appear most significant from the two-variable results and regress all possible three-variable models that stem from each of those models. We continue, keeping in mind not to violate the suggested 10:1 data point to variable ratio as admonished by Neter et al. (1996, p. 437).

After many analyses, we pick as our 'best' logistic regression model as shown in Table 1. This three-variable model's overall p-value is less than 0.0001, making us 99.99% confident that the presented statistical model is highly predictive. The listed R^2 (U) that JMP® uses is the mathematical difference of the negative log likelihood of the fitted model minus the negative log likelihood of the reduced model divided by the negative log likelihood of the reduced model. In other words, this R^2 (U) statistic "is the proportion of the total uncertainty that is attributed to the model fit (JMP® 5.0, 2002: Help)". As with ordinary least squares (OLS) a value of 0 indicates a weak model and that the explanatory variables have no predictive effect, while an R^2 (U) of 1 indicates a perfect fit.

As for the other model characteristics, we do not violate the 10:1 data point to variable ratio. The number of data points a model utilizes is particularly important for two reasons. First, the larger the sample size, the more of our population we capture in our sample. Second, the greater the number of data points, the more predictor variables we can add before the model becomes invalid statistically. According to Neter et al., a model should have at least six to ten data points for every predictor used (Neter et al., 1996, p. 437). Lastly, we consider the area under the receiver operating characteristic (ROC) curve. According to

TABLE 1
Logistic Regression Results

Model Characteristics		
Overall Model P-value	< 0.0001	
R^2(U)	0.831	
Data point/Variable ratio	11.67	
Area under ROC Curve	0.993	
Predictor Variables		
Variable	Estimate	P-value
Intercept	21.61	0.0349
X_1: Class – S	– 9.53	0.0689
X_2: Length of Prod in Funding Yrs	– 1.10	0.0390
X_3: FUE-based Maturity of EMD%	– 8.58	0.0594

JMP®, the ROC curve maps out the proportion of the true positives out of all actual positives versus the proportion of false positives out of actual negatives, both calculated across all possible calibrations of the model. We classify a true positive as a program incurring cost growth when the model predicts that cost growth will occur. Further, we define a false positive when the model predicts that cost growth will occur, but the program does not incur cost growth. The area under the ROC curve, then, gives an idea of the probability associated with ability of the model to accurately predict whether a program will have cost growth, based on results from the fitted values (Goodman, 1998: Appendix A).

To test the robustness of our logistic model, we use the 25 data points that we randomly selected from the original 122-point data set. Of these 25 data points, however, 21 data points have missing values for some of the variables, overwhelmingly because of *FUE-based Maturity %*. This leaves only 4 for validation, clearly too small a number. So, we pursue more extensive measures, namely by looking at all the viable data points in our database. Again, the FUE-based variable is a limiter, as mentioned earlier in this article. This limitation narrows our 'usable' validation pool to 39. The validation process entails saving the functionally predicted values ('1' or '0', cost growth or no cost growth) in JMP® for each of the validation data points and comparing those predicted values to the actual values. We find the model to be accurate for 37 out of the 39 useable data points (i.e., a 95% success rate), establishing that this model has a high degree of predictive ability.

In terms of the actual model's mathematical structure and of a form a user can directly use and incorporate in perhaps Microsoft Excel®, it takes the following equation:

Estimated probability of cost growth = $[\exp(-\boldsymbol{XB})] / [1+\exp(-\boldsymbol{XB})]$. (1)

Where:

XB in (1) would consequently come from the parameter estimates in Table 1 and would be represented as, $21.61 - 9.53\ X_1 - 1.10\ X_2 - 8.58\ X_3$, and

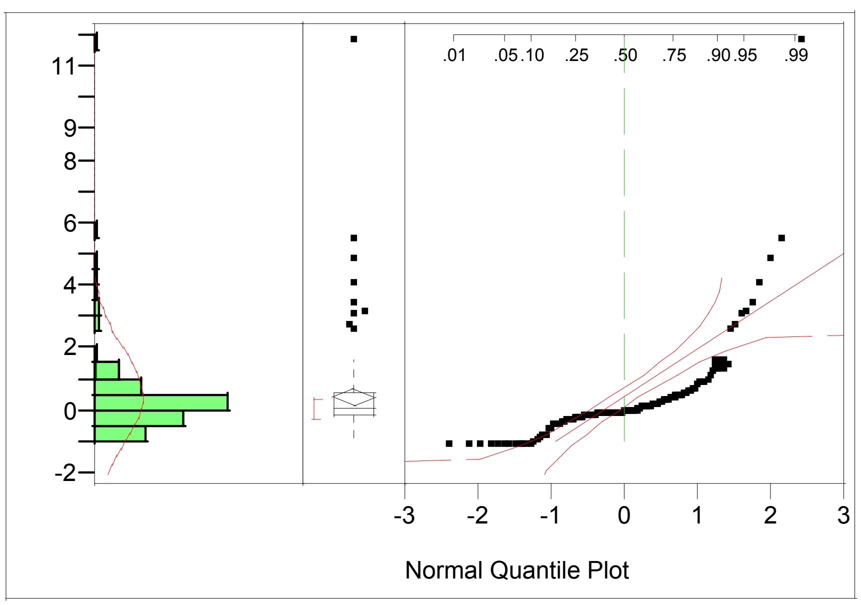
'exp' in (1) refers to the natural exponent function.

OLS Regression

Turning to multiple ordinary least squares (OLS) regression, we build this model for those occasions where a decision maker believes a program will have cost growth and wants to predict the amount of incurred cost growth. We begin model construction with our randomly selected 97 data points and exclude programs that have negative or no cost growth, leaving us with 55 data points. Focusing our efforts on only these points increases the model's prediction accuracy, because it prevents data points outside the range of interest from skewing the results. We utilize the same 78-predictor variables as in logistic regression and we consider all possible interactions between variables. For the response variable, *Y*, we use "procurement cost growth %," which measures the percent increase of procurement cost growth from the development estimate.

Because of what White, Sipple and Greiner (2004) discovered prior to their OLS model building effort, we perform a preliminary analysis of the response variable to ensure that it is continuous in nature. From the results (Figure 2), we determine that the *Y* variable exhibits a lognormal distribution, suggesting that it might suffer the same non-constancy of model residuals alluded to by White, Sipple and Greiner. We perform a

FIGURE 2
Histogram, Boxplot and Quantile Plot of Procurement Cost Growth (In %)



few test regressions and analyze the resulting residual plots (Figure 3). The plots fail to pass the visual inspection for constant variance as well as the Breusch-Pagan test (Neter et al., 1996, p. 112) at an alpha level of 0.05. Based on these findings, we transform the *Y* variable by taking the natural logarithm. This transformation successfully removes the heteroskedasticity (Figure 4) previously found and results in a distribution shape that is approximately normal (Figure 5). The distribution easily passes the Shapiro-Wilk Test for normality at an alpha level of 0.05 considering its p-value is 0.85.

We utilize the automated stepwise regression found in JMP® to aid us in narrowing the number of possible predictor variable combinations. Since we start with only 55 data points, we limit the number of predictors to six in order to prevent the predictor to data point ratio from going too far below ten to one (Neter et al., 1996, p. 437). Additionally, since we consider all variable interactions, we further constrain all models to contain at least three variables. We then analyze a multitude of

FIGURE 3
Plot of Residuals where Y = Procurement Cost Growth (In %)

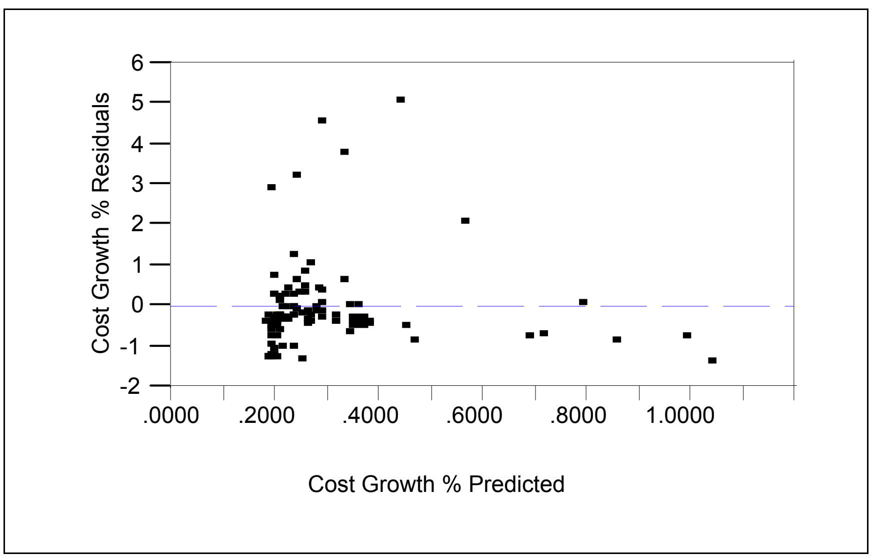


FIGURE 4
Plot of Residuals where Y = Natural Logarithm of Procurement Cost Growth %

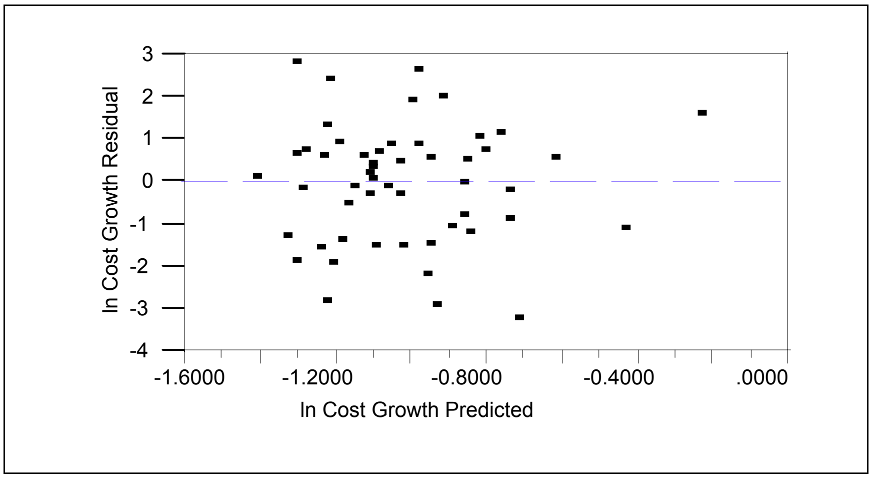
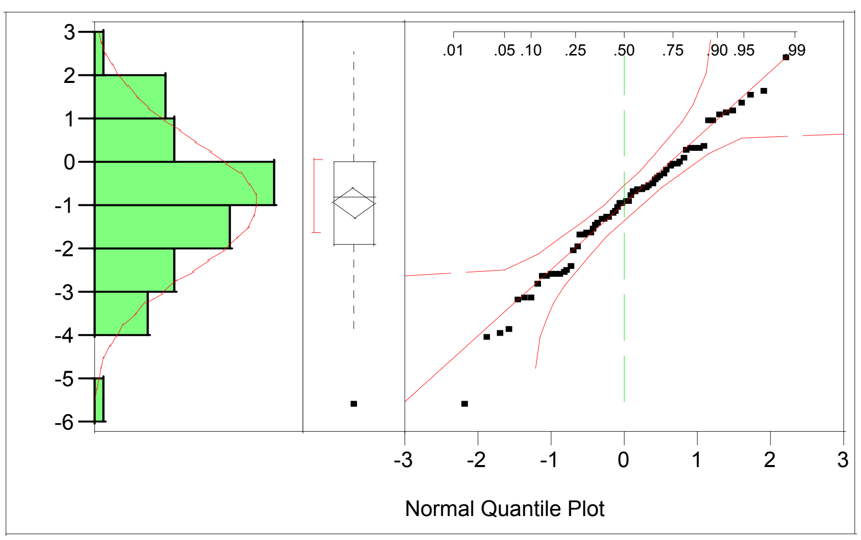


FIGURE 5
Histogram, Boxplot, and Quantile Plot of the Natural Logarithm of Procurement Cost Growth %



regression models for each number of predictors, just as we do for logistic regression, choosing the model that appears to provide the best prediction capability without violating any underlying statistical assumptions. Table 2 summarizes the details of our best OLS model. Although we dip below the desirable 10:1 data point/variable ratio, we still remain above the minimum requirement cut-off of 6:1.

In terms of the actual model's mathematical structure and in a form a user can readily incorporate in an estimating role, again perhaps in Microsoft Excel®, it takes the following equation:

$$\text{Estimated percent of cost growth} = \exp(\boldsymbol{XB}). \tag{2}$$

Where ***XB*** in (2) would consequently come from the parameter estimates in Table 2 and would be represented as, $-0.892 + 0.001\,X_1 - 0.098\,X_2 + 0.058\,X_3 - 0.569\,X_4 + 0.032\,X_5$.

We analyze the resultant model to ensure compliance with the underlying assumptions of constant variance, normality, and independence. We find the model's residuals meet the required assumptions of normality and constant variance at an alpha level of 0.05. Furthermore, having removed all dependent programs during our initial data scrubbing, we find no obvious serial correlation present. Consequently, we assume independence within the data set. As an additional precaution, we investigate all predictors for multicollinearity

TABLE 2
Multiple Regression Results

Model Characteristics		
Overall Model P-value	0.0011	
Adjusted R^2	0.595	
Data point/Variable ratio	7.33	
Predictor Variables		
Variable	Estimate	P-value
Intercept	– 0.892	0.0967
X_1: FUE-Based Length of EMD	0.001	0.7787
X_2: Service = Army	– 0.098	0.8030
X_3: FUE-Based Length of EMD * Service = Army	0.058	0.0002
X_4: Electronic	– 0.569	0.2262
X_5: FUE-Based Length of EMD * Electronic	0.032	0.0189

(i.e., linear redundancy) by ensuring all variance inflation factors as calculated by JMP® are less than ten (Neter et al., 1996, p. 387).

We choose the Adjusted R^2 to measure the model's predictive ability over the standard R^2 because of its conservative nature. The customary R^2 value is subject to artificial inflation from simply adding additional independent variables to the model. Adjusted R^2 penalizes the model builder for adding variables that do not significantly increase the model's predictive ability. Thus, by utilizing Adjusted R^2, we ensure that the variables within our model are significant, and not just used to 'pad' the model's apparent predictive ability. The p-values of the individual predictor variables are influenced by the interactions used in the models, and thus are not proper to address individually. However, all of the interaction predictors significantly add to the model at an alpha level of 0.05.

For validation, we use the same candidate data as for logistic regression. Only 17 of the original 25 validation data points have cost growth; the other 8 do not. The 17 represent approximately 25 percent of the programs within the data that contain cost growth. Therefore, we feel reasonably confident in the validation results. During model validation, we realize that only 4 of the 17 data points are usable because of missing data for some of the predictor variables, specifically FUE-based length of EMD. These results are not surprising as they mirror the results from logistic regression. Thus, we feel confident proceeding with the amalgamated validation process. That is, to further ensure the validity of the results, we perform validation on 100 percent of the data set just as was done with the logistic regression model.

We create an upper bound for validation as opposed to a prediction interval for practicality reasons. In the cost-estimating environment very few decision makers are concerned with having too much money. Consequently, our goal is to accurately predict the amount of cost growth while ensuring that the program is not underestimated. We consider an 80 percent upper prediction bound. For an 80 percent upper bound, we expect to see approximately 80 percent of the validation data points fall under the bound. From the results of our validation, we determine that for the validation data our model is 100 percent accurate at a prediction bound of 80 percent. As with the logistic regression, we are confident of our proposed model.

CONCLUSION

Defense spending has undergone great change in the last 20 years — large increases during the Reagan Administration of the 1980s, and record setting reductions under the Clinton Administration of the 1990s. The threat to the security of the United States, however, has not declined; it has merely changed form. This puts the defense acquisition community in the position of having to find ways to do more with less. For this reason, elected representatives, as well as higher ranking members of the Department of Defense pay close attention to the cost performance of major defense acquisition programs.

In this vein, cost estimators need viable tools to produce better cost estimates. A previous cost-growth research by White, Sipple and Greiner (2004) established a statistically valid methodology for predicting engineering cost growth in the research and development dollars for the EMD phase. In that article, they mention their two-step approach of multiple and logistic regression may crossover to other cost areas. We investigated the plausibility of that with respect to predicting total procurement cost growth during the EMD phase of development and discovered that this methodology works well. In that end, we supplemented the selected acquisition report database first constructed by White, Sipple and Greiner (2004). From this database, we then constructed, analyzed, and validated cost estimating relationships (CERs) for use by the cost estimating community.

Overall, our presented logistic and multiple regression models perform reasonably well in determining whether a program will have cost growth and if so how much expected cost growth (in percent) a program manager might observe. We find that the EMD maturity measure 'First Unit Equipped-based variable' is the proverbial 600-lb gorilla in the statistical models. We also learn that this information is infrequently recorded in the SARs database, perhaps because estimators are unaware of the importance of their contributions to CERs. However, when these variables are present, they appear to be significant predictors of procurement cost growth. As with any on-going research, we do not pretend that our models are the absolute best. However, they do display good predictive capability of procurement cost growth. This ability should allow the program manager to budget dwindling resources with greater confidence; thereby promoting greater credibility of the Department of Defense acquisition community to the American public.

NOTES

1. Created in 1972, the Office of the Secretary of Defense Cost Analysis Improvement Group (OSD CAIG) serves as an advisor to the Armed Services and Congress. Their role is to form an independent cost estimate/assessment for major acquisition programs in addition to that developed by the service program office governing the acquisition. The CAIG is comprised of OSD cost analysts and senior OSD officials and normally involves key members from the program that is under review.

2. In terms of the government acquisition timeline, the standard order is Engineering Manufacturing and Development (EMD), Production, and Operating and Support (O&S). EMD is normally the riskiest stage of the acquisition process. This is where the program goes through dramatic changes. The item being built goes from drawing/concept to an actual physical product (normally a prototype). Also, it is not uncommon for the initial design to undergo configuration changes either due to user changes or possibly just because of engineering constraints (e.g., cutting-edge technology).

3. Engineering changes are more commonly referred to as Engineering Change Orders (ECOs) and normally equate to approximately 5% of a program's total cost. These changes can result from a shift in the user's requirements (e.g., longer range, more speed, greater load capacity, change in the political arena (such as the fall of the Soviet Union for the F/A-22), etc.). More commonly, these changes result from a shortfall in the original system's design or from some advancement in technology. Therefore, changes (ECOs) are required to fulfill the system's operational requirements.

4. These are the largest DoD acquisition programs, and are also known as Major Defense Acquisition Programs (MDAP). To achieve this level of designation, a program must exceed $365 million in Research and Development funding or exceed $2.190 billion in Procurement funding. The Air Force's F/A-22 Raptor and the Marine's V-22 Osprey are two examples of a MDAP.

5. EMD maturity variables are utilized to determine how far a program has progressed into the development stage and how quickly it achieved certain milestones. Initial Operational Capability (IOC) is the point at which the system in development first achieves

operational capability. First Unit Equipped (FUE) is when the first operational unit is provided to the customer. To calculate a system's maturity, one would determine how far into development a program was when it reached either the IOC or FUE date and then divide that by the total length of EMD.

6. Procurement cost growth refers to any growth (schedule slips, estimates changes, engineering changes, etc.) that has affected procurement dollars. This is a categorization of cost growth based on the "color" of money that experiences cost growth. The government uses this saying to designate specific areas where certain portions of funds are designated for specific expenditures. For example, research and development dollars cannot be used to offset procurement expenditures and vice versa. They are deemed to have different "colors" of money.

REFERENCES

Calcutt, H. M. (1993). Cost Growth in DoD Major Programs: A Historical Perspective. Executive Research Project. The Industrial College of the Armed Forces, National Defense University. Fort McNair, Washington DC.

General Accounting Office (GAO) (2003). *TACTICAL AIRCRAFT: DoD Needs to Better Inform Congress about Implications of Continuing F/A-22 Cost Growth*, GAO-03-280.

Goodman, C. S. (1998). "TA101: Introduction to Health Care Technology Assessment." A Presentation Prepared for Health Care Professionals by The Lewin Group. Bethesda, MD: US National Library of Medicine. [Online]. Available at http://www.nlm.nih.gov/nichsr/ta101/ta101_c1.htm.

Grossman, E. (2002, February 7). "Defense Budgets Include $7.4 Billion Boost For Realistic Costing." [Online]. Available at https://www.asc.wpafb.af.mil/news/earlybird.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). New York: Wiley & Sons.

Jarvaise, J. M., Drezner, J. A., & Norton, D. (1996). *The Defense System Cost Performance Database: Cost Growth Analysis Using Selected Acquisition Reports*. Santa Monica, CA: RAND (MR-625-OSD).

JMP® Academic Version 5.0. (2002). Computer software, CD-ROM. Cary, NC: SAS Institute Inc.

McCrillis, J. (2003). "Cost Growth of Major Defense Programs." Briefing at the 36th Annual DoD Cost Analysis Symposium. Williamsburg, VA, 30 January.

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied Linear Statistical Models*, Fourth Edition. Boston, MA: McGraw-Hill.

Sipple, V. P., White, E. D., & Greiner, M. A. (2004). "Surveying Cost Growth." *Defense Acquisition Review Journal*, *11* (1): 78–91.

White, E. D., Sipple, V. P., & Greiner, M. A. (2004). "Using Logistic and Multiple Regression to Estimate Engineering Cost Risk." *Journal of Cost Analysis and Management*, Summer: 67–79.

APPENDIX

Program Size Variables

Total Cost Calendar Year (CY) $M 2002 – continuous variable which indicates the total cost of the program in CY $M 2002.

Total Quantity – continuous variable which indicates the total quantity of the program at the time of the SAR date; if no quantity is specified, we assume a quantity of one (or another appropriate number) unless the program was terminated.

Program Acquisition Unit Cost – continuous variable that equals the quotient of the total cost and total quantity variables above.

Quantity during PE – continuous variable that indicates the quantity that was estimated in the Planning Estimate (PE).

Quantity planned for R&D – continuous variable which indicates the quantity in the baseline estimate.

Physical Type of Program

Domain of Operation Variables:

Air – binary variable: 1 for yes and 0 for no; includes programs that primarily operate in the air; includes air-launched tactical missiles and strategic ground-launched or ship-launched missiles.

Land – binary variable: 1 for yes and 0 for no; includes tactical ground-launched missiles; does not include strategic ground-launched missiles.

Space – binary variable: 1 for yes and 0 for no; includes satellite programs and launch vehicle programs.

Sea – binary variable: 1 for yes and 0 for no; includes ships and ship-borne systems other than aircraft and strategic missiles.

Function Variables:

Electronic – binary variable: 1 for yes and 0 for no; includes all computer programs, communication programs, electronic warfare programs that do not fit into the other categories.

Helo – binary variable: 1 for yes and 0 for no; helicopters; includes V-22 Osprey.

Missile – binary variable: 1 for yes and 0 for no; includes all missiles.

Aircraft – binary variable: 1 for yes and 0 for no; does not include helicopters.

Munition – binary variable: 1 for yes and 0 for no.

Land Vehicle – binary variable: 1 for yes and 0 for no.

Ship – binary variable: 1 for yes and 0 for no; includes all watercraft.

Other – binary variable: 1 for yes and 0 for no; any program that does not fit into one of the other function variables.

Management Characteristics

Military Service Management:

Services > 1 – binary variable: 1 for yes and 0 for no; number of services involved at the date of the Selected Acquisition Report (SAR).

Services > 2 – binary variable: 1 for yes and 0 for no; number of services involved at the date of the SAR.

Services > 3 – binary variable: 1 for yes and 0 for no; number of services involved at the date of the SAR.

Service = Navy Only – binary variable: 1 for yes and 0 for no.

Service = Joint – binary variable: 1 for yes and 0 for no.

Service = Army Only – binary variable: 1 for yes and 0 for no.

Service = Air Force Only – binary variable: 1 for yes and 0 for no.

Lead Service = Army – binary variable: 1 for yes and 0 for no.

Lead Service = Navy – binary variable: 1 for yes and 0 for no.

Lead Service = DoD – binary variable: 1 for yes and 0 for no.

Lead Service = Air Force – binary variable: 1 for yes and 0 for no.

Air Force Involvement – binary variable: 1 for yes and 0 for no.

Navy Involvement – binary variable: 1 for yes and 0 for no.

Marine Corps Involvement – binary variable: 1 for yes and 0 for no.

Army Involvement – binary variable: 1 for yes and 0 for no.

Contractor Characteristics:

Lockheed-Martin – binary variable: 1 for yes and 0 for no.

Northrop Grumman – binary variable: 1 for yes and 0 for no.

Boeing – binary variable: 1 for yes and 0 for no.

Raytheon – binary variable: 1 for yes and 0 for no.

Litton – binary variable: 1 for yes and 0 for no.

General Dynamics – binary variable: 1 for yes and 0 for no.

No Major Defense Contractor – binary variable: 1 for yes and 0 for no; a program that does not use one of the contractors mentioned immediately above = 1.

More than 1 Major Defense Contractor – binary variable: 1 for yes and 0 for no; a program that includes more than one of the contractors listed above = 1.

Fixed-Price EMD Contract – binary variable: 1 for yes and 0 for no.

Schedule Characteristics

Research and Development, Test, and Evaluation (RDT&E) and Procurement Maturity Measures:

Maturity (Funding Yrs complete) – continuous variable which indicates the total number of years completed for which the program had RDT&E or procurement funding budgeted.

Funding Year (Yr) Total Program Length – continuous variable which indicates the total number of years for which the program has either RDT&E funding or procurement funding budgeted.

Funding Years (Yrs) of R&D Completed – continuous variable which indicates the number of years completed for which the program had RDT&E funding budgeted.

Funding Yrs of Prod Completed – continuous variable which indicates the number of years completed for which the program had procurement funding budgeted.

Length of Prod in Funding Yrs – continuous variable which indicates the number of years for which the program has procurement funding budgeted.

Length of R&D in Funding Yrs – continuous variable which indicates the number of years for which the program has RDT&E funding budgeted.

R&D Funding Yr Maturity % – continuous variable which equals *Funding Yrs of R&D Completed* divided by *Length of R&D in Funding Yrs.*

Proc Funding Yr Maturity % – continuous variable which equals *Funding Yrs of R&D Completed* divided by *Length of Prod in Funding Yrs.*

Total Funding Yr Maturity % – continuous variable which equals *Maturity (Funding Yrs complete)* divided by *Funding YR Total Program Length.*

Engineering and Manufacturing Development (EMD) Maturity Measures:

Maturity from Milestone (MS) II in months (mos) – continuous variable calculated by subtracting the earliest MS II date indicated from the date of the SAR.

Actual Length of EMD (MS III-MS II in mos) – continuous variable calculated by subtracting the earliest MS II date from the latest MS III date indicated.

MS III-based Maturity of EMD % – continuous variable calculated by dividing *Maturity from MS II in mos* by *Actual Length of EMD (MS III-MS II in mos).*

Actual Length of EMD using IOC-MS II in mos – continuous variable calculated by subtracting the earliest MS (Milestone) II date from the IOC (Initial Operational Capability) date.

IOC-based Maturity of EMD % – continuous variable calculated by dividing *Maturity from MS II in mos* by *Actual Length of EMD using IOC-MS II in mos.*

Actual Length of EMD using FUE-MS II in mos – continuous variable calculated by subtracting the earliest MS II date from the FUE (First Unit Equipped) date.

FUE-based Maturity of EMD % – continuous variable calculated by dividing *Maturity from MS II in mos* by *Actual Length of EMD using FUE-MS II in mos.*

Concurrency Indicators:

MS III Complete – binary variable: 1 for yes and 0 for no.

Proc Started based on Funding Yrs – binary variable: 1 for yes and 0 for no; if procurement funding is budgeted in the year of the SAR or before, then = 1.

Proc Funding before MS III – binary variable: 1 for yes and 0 for no.

Concurrency Measure Interval – continuous variable which measures the amount of testing still occurring during the production phase in months; actual IOT&E completion minus MS IIIA (Jarvaise, 1996: 26).

Concurrency Measure % – continuous variable which measures the percent of testing still occurring during the production phase; (MS IIIA minus actual IOT&E completion) divided by (actual minus planned IOT&E dates) (Jarvaise, 1996: 26).

Other Characteristics

 # Product Variants in this SAR – continuous variable which indicates the number of versions included in the EMD effort that the current SAR addresses.

Class – S – binary variable: 1 for yes and 0 for no; security classification Secret.

Class – C – binary variable: 1 for yes and 0 for no; security classification Confidential.

Class – U – binary variable: 1 for yes and 0 for no; security classification Unclassified.

Class at Least S – binary variable: 1 for yes and 0 for no; security classification is Secret or higher.

Risk Mitigation – binary variable: 1 for yes and 0 for no; indicates whether there was a version previous to SAR or significant pre-EMD activities.

Versions Previous to SAR – binary variable: 1 for yes and 0 for no; indicates whether there was a significant, relevant effort prior to the DE; a pre-EMD prototype or a previous version of the system would apply.

Modification – binary variable: 1 for yes and 0 for no; indicates whether the program is a modification of a previous program.

Prototype – binary variable: 1 for yes and 0 for no; indicates whether the program had a prototyping effort.

Dem/Val Prototype – binary variable: 1 for yes and 0 for no; indicates whether the prototyping effort occurred in the PDRR phase.

EMD Prototype – binary variable: 1 for yes and 0 for no; indicates whether the prototyping effort occurred in the EMD phase.

Did it have a PE – binary variable: 1 for yes and 0 for no; indicates whether the program had a Planning Estimate.

Significant pre-EMD activity immediately prior to current version – binary variable: 1 for yes and 0 for no; indicates whether the program had activities in the schedule at least six months prior to MSII decision.

Did it have a MS I – binary variable: 1 for yes and 0 for no.

Terminated – binary variable: 1 for yes and 0 for no; indicates if the program was terminated.