COMPARISON OF DEVELOPMENT TEST AND EVALUATION AND OVERALL PROGRAM ESTIMATE AT COMPLETION

THESIS

William R. Rosado, Capt, USAF

AFIT/GCA/ENC/11-02

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GCA/ENC/11-02

COMPARISON OF DEVELOPMENT TEST AND EVALUATION AND OVERALL PROGRAM ESTIMATE AT COMPLETION

THESIS

Presented to the Faculty

Department of Mathematics and Statistics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Cost Analysis

William R. Rosado, BS

Captain, USAF

March 2011

AFIT/GCA/ENC/11-02

COMPARISON OF DEVELOPMENT TEST AND EVALUATION AND OVERALL

PROGRAM ESTIMATE AT COMPLETION

William R. Rosado, BS
Captain, USAF

Approved:

Dr. Edward D. White (Chairman)	Date
Lt Col Eric J. Unger (Member)	Date
Steven P. Tracy (Member)	Date

AFIT/GCA/ENC/11-02

Abstract

Historically, cost growth regression models analyze aggregate, program-level information. Initiatives by the Office of Secretary of Defense, Cost Assessment and Program Evaluation (OSD CAPE) require direct, centralized reporting of the complete Work Breakdown Structure (WBS) Earned Value (EV) data. Centralized reporting allows access to unfiltered, unaltered, EV data for multiple programs. Using regression, we evaluate if WBS element Development Test and Evaluation (DT&E) EV data is related to program estimate at completion (EAC). Identifying a relationship provides evidence validating pertinence and reliability of low level EV data. Additionally, a relationship between a specific WBS element and program EAC establishes a basis for improved estimate development, and prediction capability. Our results show a strong relationship between DT&E and program EAC.

Although limited by sample size and assumptions regarding DT&E commonality, our findings lead us to believe that there is potential for improved prediction models using low level WBS EV data.

AFIT/GCA/ENC/11-02

I dedicate this to my Parents, Maureen and William, my brother Matthew, and my friends and peers in the AFIT GCA/GFA cohort.

Acknowledgements

I would like to thank my family for their continuing support in all aspects of my life, including the challenges associated with the thesis process.

I would like to thank Steven Tracy for his involvement and assistance as I interpreted the research results. I especially want to thank LtCol Eric Unger and Dr. Edward White who were both instrumental to my success in this endeavor, not only as members of my committee, but as my guides through the entire Cost Analysis Master's Program.

Finally, I would like to thank Lt. Charles "Grant" Keaton who was a tremendous asset throughout the thesis process.

Table of Contents

Page

Page

List of Figures

List of Tables

COMPARISON OF DEVELOPMENT TEST AND EVALUATION AND OVERALL PROGRAM ESTIMATE AT COMPLETION

I: Introduction

General Issue

Despite numerous efforts and various studies the Department of Defense (DoD) acquisition community continues to struggle with cost estimating and the accurate forecast of program's Estimate at Completion (EAC). In an increasingly harsh and demanding financial climate, coupled with continuing military operations, inaccurate estimates draw attention from stakeholders at every level. Most recently, the DoD implemented the Weapon Systems Acquisition Reform Act of 2009 to address the need for improved cost estimates. Simultaneously, academics continue to develop, test and analyze prediction models while acquisition and cost estimating professionals in the field strive to refine their estimation techniques.

The number of individual formulas and processes for developing an EAC are numerous but, as experts in the field of EAC research have found, they can be summarized into three general categories; index based, regression (linear and non-linear), and other (Christensen, 1995). The most abundant method in use and variety is the index based approach, followed by regression methods and finally other techniques.

Regardless of how an index is calculated, we can generally describe it as, "A measure of the level of performance attained in completing the work on the contract up to current time (Nystrom, 1995)." The index approach develops the EAC by adding the actual costs incurred to date, or the Actual Cost of Work Performed (ACWP) plus an adjusted value for the work remaining. The simplest index versions used are the Cost Performance Index (CPI), Schedule Performance Index (SPI), or a combination of both called the Schedule Cost Index (SCI). Calculation of these is done using the Budget Cost of Work Performed (BCWP), Budgeted Cost of Work Scheduled (BCWS), and previously mentioned ACWP.

The regression based approach attempts to use multiple or logistic regression techniques; multiple regression focuses on the magnitude of cost growth while logistic regression is aimed at identifying the existence of cost growth. The most common regression approach seen in academic studies focuses on modeling the cost growth profile of a program. This cost growth profile is also known as the S-Curve, curvilinear cost profile or the growth curve. The majority of the cost growth profile work utilizes the Budgeted Cost of Work Performed or percent complete (calculated as BCWP divided by BAC) as independent variables and the ACWP as the dependent variable, although other variations have been investigated. The equations follow a linear or non-linear form such as exponential or quadratic:

$$ACWP = A \times BCWP + B$$

$$ACWP = A \times BCWP^{B}$$

$$Percent\ Complete = A + (B \times \%Time) + (C \times \%Time^{2})$$

Where A and B are coefficient estimates of a given model or curve and BCWP and % Time are the independent, predictor variables.

Review of the major academic and field studies shows that most regression model efforts are plagued by small sample sizes (sample size ranging from one to fifty seven programs). These attempts are also modeled on program level data that "rolls up" the detailed, low-level, Work Breakdown Structure (WBS) elements into higher level aggregate values. Most recently, using a sample set of 114 programs Kristine Thickstun attempted to build a multiple regression model to predict if a program would experience an Over Target Baseline (OTB) adjustment. Her work expanded on Elizabeth Trahan's analysis, which was based on the Defense Acquisition Executive Summary (DAES) and the Systems Acquisition Review (SAR) (Trahan, 2009; Thickstun, 2010). In both studies the data analyzed was top level and aggregate in nature.

As part of improvement initiatives reporting procedures for program earned value data were changed (Augustus, 2011). As seen in Figure 1, the changes required that program CPRs flow directly to Defense Cost and Resource Center (DCARC). Previously the flow of this information routed through the Program Management Office (PMO).

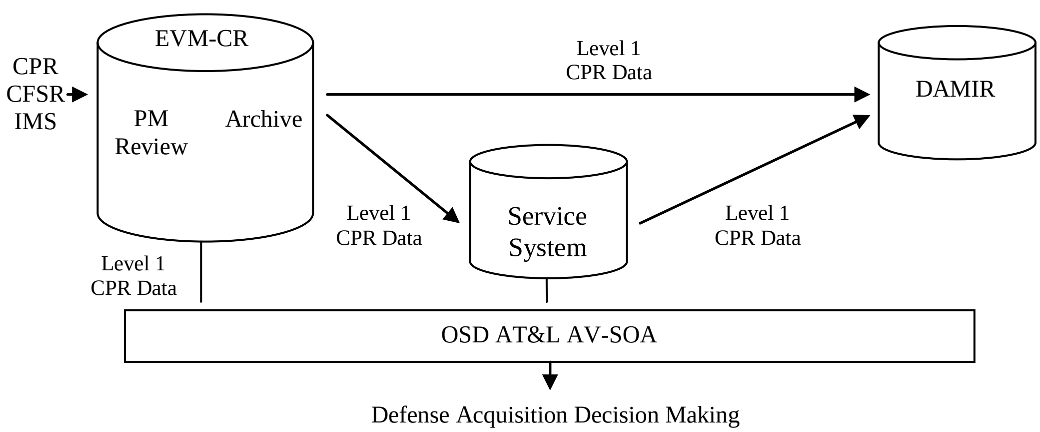


Figure 1: DCARC EVM Data Flow

As summarized on the DCARC portal page, "DoD established a single centralized repository where data can be carefully controlled and easily accessed. DoD identified

Earned Value Management (EVM) System products as the first series of data to be included in the repository." (DCARC Portal) This centralized repository of unfiltered, unaltered WBS EVM data provides opportunities for research and analysis at levels of detail previously unavailable.

DACIMS is simply an online portal used to access the monthly, detailed EVM data that resides within DCARC. Using the detailed work breakdown structure earned value information now available we hope to develop prediction models, using regression based techniques, which better define the cost behavior of DoD programs.

Purpose of This Study

Dr. Dave Christensen stated, "The purpose of variance reporting is not to find fault but to identify and correct problems before they worsen." (Christensen, 1995) A regression model does not necessarily show causality; therefore any relationship we find between independent variables in our study and EAC growth act as predictors. Further analysis and research is required to define the specific causality between any significant predictor variables and actual cost growth. Additionally, identification of problem areas within a program does not necessarily mean that we can fix the problems.

For those reasons our goal is to develop a model that can provide early warning that cost growth may occur. Additionally we hope to identify if there are specific program elements, as defined by the WBS, which contribute most to cost growth. This early prediction does not provide a solution but it does provide vital information necessary for successful management of the program. The earlier we are aware of potential issues, and the more knowledge we have regarding the source, the sooner we can implement strategies and processes to mitigate negative effects. Ultimately, improved models support successful management of acquisition programs.

Research Question

Previous work validated that growth curve equations such as Gompertz, Rayleigh, and Weibull are descriptive of growth patterns seen in various fields of study (Karsch, 1974; Watkins, 1982; Winsor, 1932). Subsequently, the relationship of growth curves to DoD program budget outlay and expenditure patterns was tested and validated (Karsch, 1974; Unger, 2001; Trahan, 2009). In these studies, a regression model incorporating the characteristics of growth curves was used to develop a model to predict EAC or the presence of cost and schedule growth. Other studies attempted to build prediction models using multiple regression; these attempts used characteristics of the program such as weapon system type, phase, and EVM data (Sipple, 2002; Thickstun, 2010).

The various works cover a wide spectrum of potential approaches to building a regression model for the purposes of estimating cost growth. However, a common characteristic of the works cited is that they utilize top level EVM data from programs. Our first research question focuses on the relationship between the lower level EVM data and the aggregate data. We believe it is important to validate that the accounting and reporting of lower level EVM data relates to the aggregate and follows common trends such as S-shaped cumulative expenditure patterns. Expanding the field of research beyond this limiting characteristic is the main basis for our research in the complete WBS structure of acquisition programs.

EVM data trends, such as s-shaped expenditure curves, are noticeable at aggregate levels. But we found no work testing if these trends are consistent down to the lower WBS levels. We expected these trends to be exhibited in the lower WBS structure, but we also expected degradation in the trend the deeper in the WBS structure we look. The potential for wide swings in EVM metrics are more likely at the lower levels where the EVM data represents specific elements of a program. Therefore, our first question is: *are the EVM characteristics and patterns of the overall program consistent in the lower levels of the WBS structure?*

We then wanted to analyze the relationship of specific WBS elements and their characteristics, such as EVM metrics, against the EAC growth behavior of the entire program. Doing so could potentially reveal cost drivers for programs given the program's weapon system type, service, phase, and so on. Our second question is: *can we show a statistical relationship between low-level WBS elements and overall program EAC growth?*

Our third question builds upon the previous two finds, we ask: *using the lower level WBS information now available through DCARC can we build a statistically robust model to predict EAC growth?*

Summary

We seek to expand upon previous work in this field by utilizing the complete WBS EVM data available to us. Better models and estimating techniques support proper management of DoD acquisition programs; ultimately, this translates into better system capabilities, fielded sooner, for use by the war fighter.

II: Literature Review

Introduction

The majority of EAC estimation research focuses primarily on index-based methods. Similarly, the most commonly employed methods in the field are index-based; this is most likely due to the simplicity of applying an index to develop an EAC (Trahan, 2009). Regression methods, as an alternative, require a more complicated process to develop but, as research has shown, could generate better estimates early in a programs life (Christensen, 1995; Tracy, 2005). Experts in this field generally agree that the usefulness of a robust parametric model will be highly useful early in a program's life. However, as shown in Figure 2 the effectiveness of parametric modeling declines as index-based EAC estimates become increasingly accurate (Holeman, 1975).

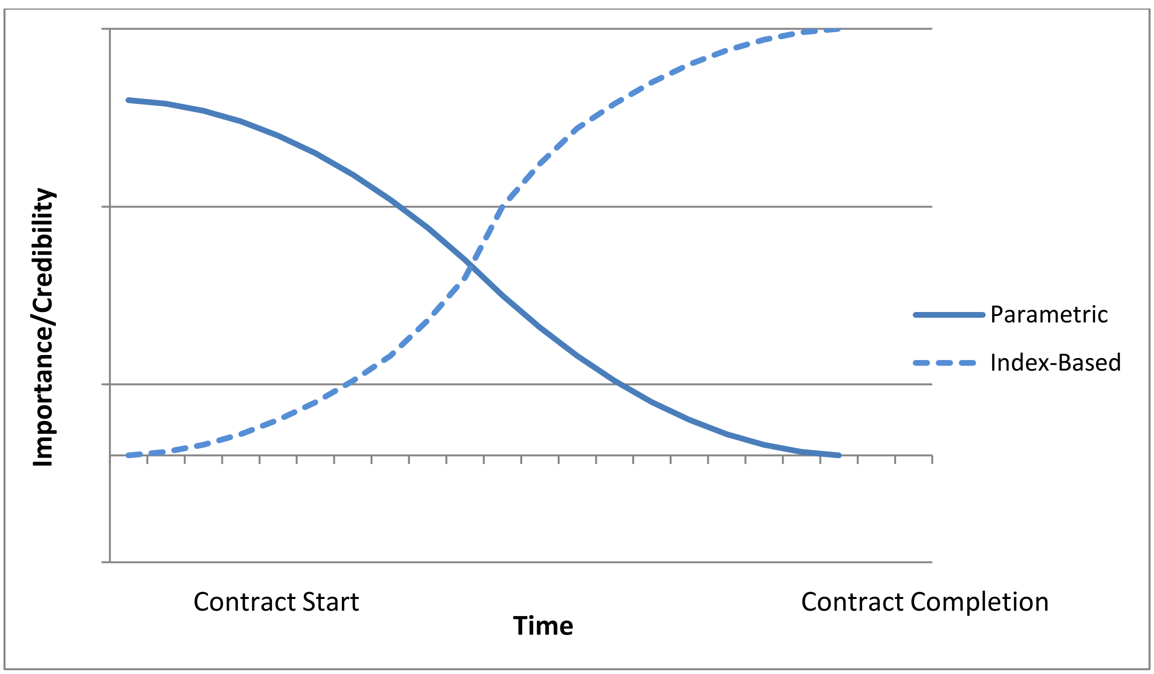


Figure 2: Tradeoff between Parametric and Index-Based Estimates, Holeman 1975

The Office of the Undersecretary of Defense for Acquisition reviewed over 500 completed contracts from the Defense Acquisition Executive Summary (DAES) database and found (Christensen, 1995):

> *Given that a contract is more than fifteen percent complete, the overrun at completion will not be less than the overrun incurred to date; and the percent overrun at completion will be greater than the percent overrun incurred to date.*

Knowing that early, accurate estimates are necessary to mitigate the risk of cost growth, numerous researchers have turned to regression in attempt to build a model with predictive capability early in the life of a program. Prior research has validated the growth curve characteristics of acquisition programs and also investigated and identified relationships between cost growth and program characteristics. This prior research provides a vector for our work by identifying potential predictor variables when developing a multiple regression model.

Regression Modeling Background

The most common characteristic of program's budget and expenditure patterns is the "S-Shape" curve (Weida, 1977). Aside from a few studies using time-series analysis, smoothing techniques or a combination of both (Olsen, 1976; Chacko, 1981), or multiple regression attempts (Sipple, 2002; Thickstun, 2010), the majority of regression techniques focused on the growth models.

Growth Models

The S-Shape curve can be accurately modeled and applied to program budget data using a variety of different growth equations (Karsch, 1974; Nystrom, 1995; Unger, 2001; Trahan, 2009). Previous regression work primarily focuses on ACWP, BCWP, CPI, and Time as the independent and dependant variables and follow one of given formats shown in Table 1 (McKinney, 1991):

Table 1: Regression Curve Formulas

$Y = a + bX$	**Linear Curve**
$Y = aX^{b}$	**Power Curve**
$Y = ae^{b(X)}$	**Exponential Curve**
$LnY = a + b \ Ln \ X$	**Log Curve**
$Y = a + bX + b_2X^{2}$	**Quadratic Curve**

Karsch developed a nonlinear model using least squares regression. His model assumes that identification of a reasonable trend relationship of ongoing activity sets the pace for future activity (Karsch, 1974). He developed a power curve to describe the relationship between ACWP and BCWP. Using a log transformation he calculated the coefficients of the model and evaluated their predictive capability against other program data. He found that between various samples the range of coefficient estimates was "narrow"; from 0.97 – 1.18 for the exponent parameter, most of the cases between 1.00 and 1.10 (Karsch, 1974). His work further showed that growth characteristics inherent in program expenditures are common across all types of programs.

Additional research on Karsch's model found it to be highly sensitive to various characteristics including the phase or stage of the contract (Busse, 1977; Heydinger, 1977; Land and Preston, 1980). Work done by Watkins and others varied the growth curve type and application to the data set to develop regression models (Watkins, 1982).

Watkins considered the impact of level of effort, represented by manpower buildup, on program costs. He developed a technique to apply an adaptive Rayleigh-Norden model to describe the relationship between manpower and the ACWP over time. An example of the Rayleigh cumulative distribution is shown in Figure 3.

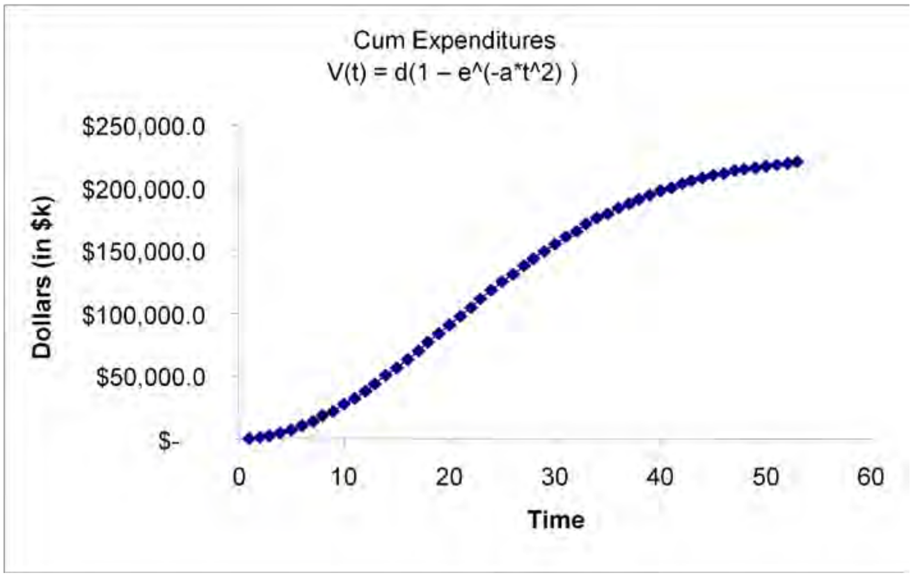


Figure 3: Rayleigh Cumulative Distribution, Lee 2002

Later, using the Weibull model and budget data, Unger was able to develop a robust model to predict the existence of cost growth. Weida recognized the relationship between growth models and the expenditure patterns in programs but did not try to apply a previously existing growth model type. Instead he developed an S-curve equation specific to the program data he had, unconstrained by a specific model specification (Weida, 1977).

Weida felt that the comparative and predictive capability of an S-shaped curve provided the rational for its use. As analysts, unless we wish to duplicate the effort inherent in the original "contract-letting" process, we must accept the proposed budget as the best estimate of total cost. Once we accept the proposed budget Weida believed that three analysis approaches become available to the analyst. First, using regression techniques a General S-shaped curve must be developed, preferably by weapon system (aircraft, avionics, etc.) using a large sample pool of programs. Weida postulated that while there are often changes in programs as they progress, development of a General curve using data from similar weapon systems final costs should incorporate these changes. Therefore, final cost figures generated based off the General curve would include a similar number of program changes, even if the changes are not visualized early in the program life (Weida, 1977).

Analysis approach number one was comparison of the General expenditure curve, based on actual program data, to the proposed expenditure pattern of the given program. If the S-shaped curve developed from the proposed expenditure pattern of a program was statistically different (outside one standard deviation confidence interval) compared to the Generalized curve then the contractor should explain why their program is unique. The second approach is a validation of the specific rationality of the program. Weida found a strong relationship between the cumulative completion of project milestones and budget expenditure pattern. Testing the relationship between the new programs proposed cumulative completion and budget expenditure patterns would test this specific rationality. Finally, Weida suggested that the S-shaped curve could be used as a forecasting tool for the EAC.

Weida felt that the critical point in a programs' life, when the majority of the inherent uncertainty has dispelled, is the inflection point seen in the budget expenditure pattern. Citing work done by Drake in 1970 Weida suggested that the uncertainty due to unknown unknowns followed and exponentially decreasing pattern from the start of the program to completion. Alternatively, uncertainty due to known unknowns is much lower during the program life but does not decrease at such a rapid pace as the program goes along. Therefore, a combination of the two uncertainties generates a "kinked curve". This curve has high uncertainty early in the program life, decreasing exponentially as time goes on, and kinks at the inflection point where remaining uncertainty plateaus and slowly decreases until a sharp drop to zero at 100% completion.

Multiple and Logistic Regression Models

Other analysis strayed away from the growth curve models and attempted to build multiple or logistic regression models using characteristic program data. We will address some of the recent significant attempts with respect to DoD acquisition programs. In 2002 Sipple attempted to develop both logistic and regression models; in his two-step procedure he sought to predict the occurrence of cost growth using a logistic model and, if possible, model the total increase using a multiple regression model. Using an exhaustive set of predictor variables he focused his efforts in predicting cost growth in the research and development dollars for the Engineering Manufacturing Development phase of the acquisition (Sipple, 2002).

Sipple grouped the predictor variables into five broad categories: program size, physical type of program, management characteristics, schedule characteristics, and other characteristics. The broad categories contain at most two subcategories, the total set of predictor variables allowed Sipple to build increasingly specific models for the programs in his sample set.

In the physical type of the program group the predictor variables fall under the physical domain the system operates in (air, land, space, sea), or by the functional characteristics of the system (electronic, helo, missile, aircraft, munition, land vehicle, ship, other). A similar approach for management characteristics was used to develop predictor variables. Sipple created a predictor variable not only to represent which service was involved with the program, but also predictors to explain more complex relationships such as multiple service involvement and identification of the lead service. He used the same type of dummy variables to define contractor involvement and included variables to account for complex nuances such as no major defense contractor involvement, more than one major defense contractor involvement, and type of contract for the development.

Within the schedule characteristics group Sipple considered various measure of maturity for the sample set. From simple proxies for maturity such as “funding years complete” and “years research and development complete” to “total funding year maturity %”, a calculation based on funding years complete divided by total program length. Additionally, he developed predictor variables to test for the impact of testing concurrency in a program. The “concurrency measure %” was generated as the percent of testing still occurring during production divided by actual minus planned test and evaluation dates.

Finally, Sipple included variables describing other aspects of the sample set characteristics to explore their predictive capability. The other characteristics group included variables defining the security classification of the system in question, number of variants, and identified if any risk mitigation activities had taken place.

Sipple found that a seven variable logistic model was able to predict the existence of cost growth in approximately 70 percent of the validation set. A three variable model was created to predict the actual amount of cost growth, the model considered maturity from Milestone II, the lack of a major defense contractor, and the program acquisition unit cost as the predictive variables. The model had an adjusted R^2 of .42 and passed the tests for constant variance and normality of residuals.

In 2005 Tracy explored multiple regression models using an expanded list of potential predictor variables. Tracy grouped his set of predictor variables by categorical, performance data, and other. The overall preliminary set contained some similar variables as Sipple's work, dropped others, and included unique predictors as well. The number, amount, and magnitude of OTB changes, and consideration of the Contract Budget Baseline in relationship to the BCWP and management reserve were among the unique variables considered. Tracy developed five models which, based on the literature review of previous work, showed commonality in predictor variables. Validation of his models showed improved performance over the comparison methods, cumulative index based models, with generally better measures. Given the goal of developing a model to predict cost growth using a 'snapshot' of cross-sectional data, Tracy found the results outperformed expectations (Tracy, 2005).

Elizabeth Trahan used a Gompertz growth model to develop the EAC; her model was successful for Over Target Baseline (OTB) or approaching OTB contracts. Looking to build upon that research Kristine Thickstun attempted to build a multiple regression model that could be used to accurately predict whether or not a contract would go OTB. Included in her models are various program characteristic variables such as service type, military handbook weapon type, percent complete, and percent change in production quantity. The models also included EVM categories such as BCWS, BCWP, EAC in a given base year, and the Schedule Cost Index. Using these variables Thickstun tried to predict if a program would experience an OTB. Although the models failed the validation stage, the analysis process did provide further basis for potential predictor variables in a multiple regression model.

Summary

There is no shortage of regression based models developed for the prediction of EAC. However, these models all have some common characteristics. The main commonality we are concerned with is that these models are based on summary level data. The development of these models provides credence for regression based modeling for the purposes of estimating cost growth. We wish to expand the knowledge in this field by evaluating the performance of a model using data from the complete WBS.

The majority of prior work incorporates some type of pre-existing growth curve to model the data in which the models use similar independent variables (ACWP, time, CPI). The studies that break away from the growth curve modeling in attempt to develop logistic and multiple regression models give a significant preliminary set of potential predictor variables. Additionally, the results of prior work show that a multiple regression model, used to analyze a cross section of acquisition data in time, does have predictive capability.

Successful development of regression models, identification of a large set of potential predictor variables and access to the complete WBS cost data provides the context for research and model development.

III: Data Collection and Methodology

Chapter Overview

This chapter establishes the methodology and approach used to answer our research questions. First, we explain the characteristics of our data source, the DCARC EVM central repository, and our sample set, DCARC history files; this explanation provides context to understanding our analysis approach, results, and conclusion. Additionally, we will further detail anomalies in the data set which impacted our ability to use some of our sample. Next, using our literature review as a vector, we attempt to gather cost predictors from the data set to develop a model in a way that has not previously been attempted. Finally, we will describe the results of our exploratory data analysis and explain the regression techniques we used to analyze our sample set.

Defense Cost and Resource Center

DCARC History and Intent

The Defense Cost and Resource Center, which is a part of the Office of the Secretary of Defense, Cost Assessment and Program Evaluation (OSD CAPE), is a centralized repository for DoD acquisition program data. DCARC, formally known as Contractor Cost Data Report (CCDR) Project Office (CCDR-PO), was established in 1998 to support adjustments in the CCDR process. According to Mike Augustus, the OSD Acting Director of DCARC, the original intent of DCARC was to collect acquisition program Contract Cost Data Reports (CCDR) in accordance with the objective of making Major Defense Acquisition Program (MDAP) cost and software resource data available to authorized Government analysts (Augustus, 2011).

Earned Value Management Application

The Under Secretary of Defense for Acquisition, Technology and Logistics (USD AT&L), acknowledging their responsibility for ACAT IC and ID programs, realized a need for a centralized repository of Contract Performance Reports. This responsibility demands "situational awareness of all programs within their cognizance", and therefore transparency of data is paramount. Although collection of cost performance reports occurred, the reports were first filtered through the Program Management Office (PMO). According to Mr. Augustus, many Program Offices had a unique submission file format or method of presentation. More importantly, some Program Offices chose what information to pass along or manipulated data to protect their interests. AT&L was concerned about the fidelity of the data due to the PMO filter; it lacked transparency and not all DoD stakeholders were reviewing the same data. As summarized on the DCARC Portal EVM Application site:

> *EVM products are the first series of data to be included in the centralized reporting. All DoD contractors for ACAT IC and ID program contracts will forward their CPRs, CFSRs, and IMS. The one new distribution point replaces all the previous multiple distribution points previously required.*

The directive for a new collection point was designed to take the CPR direct from contractors and manifested into the Earned Value Management application on the DCARC portal.

Mr. Augustus stated that in the interest of collecting the data as soon as possible DCARC accepted submissions in any format or state, essentially to "see what's out there." Subsequently one finds a massive amount of information in the DCARC portal albeit impaired by inconsistent document types, unusable formats, missing or incomplete submissions, and incorrectly filed submissions. Appendix A shows a recent status report of programs in DCARC, this report gives an idea of the somewhat sporadic reporting of CPR files common in DCARC at this time.

DCARC Portal

Documents available in DCARC include, but are not limited to, the Cost and Software Data Report (CSDR), DD Form 1921 (Contractor Business Data Report), Contract Performance Reports (CPR), Contract Funds Status Report (CFSR), and Integrated Master Schedules (IMS). This information is collected and organized in different applications with the intent that individual analysts can utilize it to support the cost estimate process and ultimately make DoD cost estimates more robust.

Analysts with appropriate access can submit and review these reports by accessing one or more of the following applications, as shown in Table 2, within the DCAR Portal. These applications provide access to vast libraries of programmatic data.

Table 2: DCARC Applications

cPet Web: Cost and Software Data Report Planning and Execution Tool
CSDR-SR: Submit & Review of 1921, 1921-1, 1921-2, 2630-1, 2630-2, 2630-3, Contract Cost Data Reporting, Software Resources Data Reporting, & Contract Work Breakdown Structure
1921-3 & FPR: Submit & Review of 1921-3 & Forward Pricing Rate
DACIMS: Cost and Software Data Reporting & Forward Pricing Rate Library
EVM: Submit & Review of Cost Performance Report, Contract Funds Status Review, & Integrated Master Schedule

The EVM Central Repository

The EVM Central Repository provides and supports the centralized reporting, collection, and distribution for Key Acquisition EVM Data, such as Contract Performance Reports (CPRs), Contract Funds Status Report (CFSR), and the Integrated Master Schedule (IMS) for ACAT 1C & 1D (MDAP) as well as ACAT 1A (MAIS) programs.

Authorized users can download the information in various file formats including Adobe Reader (.pdf), Power Point (.ppt), Excel (.xls), Extensible Markup Language (.xml), or Deltek wInsight (.wsa). The different formats are used depending on the product being submitted. For example, many contractors submit cost performance reports in Adobe Reader format so that the respective DoD analyst can easily open and review the information. Contractors are also required to submit the CPR Format 1 EVM data in a file type which allows for easy manipulation and analysis. Typically the monthly files are in Extensible Markup Language but may also be Excel format. The standard is that contractors must provide, per month, the "readable" version of the CPR (.pdf format) and the "analyzable" version (.xml or .xls); but this is not always the case. As we identified through our discussion with Mr. Augustus, and other analysts responsible for programs in our sample set, files may be missing for numerous reasons. Mr. Maringas, who currently works on the Mission Planning System programs for ESC and has nearly five decades of experience in the acquisition field, provided an extensive list of potential reasons for missing files. Files are submitted, rejected, and subsequently not resubmitted. Backing up historical files for a program may require multiple years of data to backup. The backup crashes and subsequently is never completed. Backups are also affected by different versions of the wInsight software. File strings are broken when the work breakdown structure changes. Additionally, if a restore or backup does not work individual .xml CPR files may be imported under and labeled under a different naming scheme, making them hard to find (Maringas, 2011).

Based on the level of authorization and purpose for access a user is granted greater or lesser control within the portal, this includes the ability to download and open the Extensible Markup Language files containing the EVM data. Although the readable CPRs contain all the EVM data they do not make for easy analysis.

For example, one CPR for the C-17 Avionics Modernization Program contains close to 200 work breakdown structure elements. If, as an example, we wanted to consider only two years of the program life then we must manually consolidate 4,800 data points from twenty four individual reports.

We can expedite the process using text recognition software; however the variability in WBS structure and naming convention between programs makes it nearly impossible and would require an impractical amount of time to adjust the recognition software to each program. The task becomes especially daunting when we consider collecting a large enough sample size to build a robust regression model. Using our C-17 example as a standard, 4800 data points, expanding to a sample of data from only twenty programs requires the consolidation and organization of 96,000 data points.

Other options for consolidation of the CPR data are the Extensible Markup Language documents and Excel documents. We found that we could open and download the EVM Excel files without any problems. The data was not in the ideal arrangement within the Excel file, by that we mean one row reporting EVM data for a single line item and each column identifying the associated EVM data; however we felt we could manipulate the information quickly into the desired matrix format for analysis.

By matrix format we mean the data in adjacent rows and columns. Additionally, we wanted each row to contain a set of records unique to a specific WBS element, relative to the month the data was recorded.

However, using the Excel files was not feasible because most contractors did not submit their reports in this format and even if they did it was not consistent. Some months were done in Excel, others in the Extensible Markup Language. Looking to the Extensible Markup Language files as our source for the data we encountered complications with the server and software access. Additionally there were gaps in reporting of the CPRs, especially at the beginning of the programs life.

In order to solve the problem with data collection we needed an automated method and submissions that were consistent and complete. A contact was made at DCARC who provided a web based tool which can has the ability to parse Extensible Markup files or wInsight files. The history files, which contain CPR data over a programs life rather than just one month, appeared to be our solution.

DCARC History Files

Within the DCARC portal, searching by contract, we had access to all received submissions including Contract Performance Reports Format 1 through Format 5, Integrated Master Schedule, and Contract Funds Status Report. For the purposes of our research we focused on EVM information found in Format 1 of the monthly CPR. However, the monthly CPR file is not the only source for this EVM data, the program history files contain consolidated EVM data over multiple months of a program. Coordinating with DCARC administrators, we developed a query to search the vast library of DCARC files specifically for history file submissions. The query identified 813 potential submission events as relevant history files.

We used the associated Submission ID and File ID to find the history file, download it, and using the provided parsing tool export the file for analysis.

Of 813 potential history files we were unable to locate 28 files. An additional 150 potential files were unusable due to incorrect history file format or because the associated files were not actually history files. For example, some files appeared to be historical data but the file was .xml versus .wsa format and thus we could not parse it into a usable format. In the other cases the file found in the submission event was some type of report such as a single CPR, IMS, or other but not a complete history file.

As we continued our research we found that there were also a substantial number of duplicates per program. The history file query which produced 813 results not only found the most recent history file for each program but all previous history files, which are generally updated and loaded annually. Therefore, the sample of program history files dropped dramatically from what we previously thought was potentially six hundred and thirty six to just over two hundred individual contracts.

Once we were able to open the history files for the programs we also identified anomalies in the data itself. Missing data or completely blank cells that should be filled, shifted decimals randomly adjusting months worth of cost information from the millions to billions then back again, and history files that export into Excel with shifted columns which overlap portions of essential cost data all plagued our sample set. Some of the issues were addressable or a nonfactor; if the missing information is determinable from elsewhere we replaced it or in other data sets we validated the correct decimal place and changed values to be consistent. In other cases we could not "fix" the history file output, specifically when cost data was blank or shifted cells bumped required EVM columns out of the file during export to Excel.

Data Screening Criteria

Our initial intent was to focus on a narrow subset of the acquisition field, gather the data, design a model, and test for significance. As far as we know this is the first research attempt at building a regression model using earned value data below the aggregate level; that being said we wished to preemptively eliminate between program variability. Between program variability being the inherent differences between acquisition programs by service, weapon system type, contract phase, and contract type (production versus research and development).

We decided to base our sample set on research and development contracts within the military handbook classification of Aircraft, and if at all possible, narrow the scope further to a single service, the Air Force. What we soon found was that our population of data did not support such a narrow focus. In fact, filtering our sample for any of the above criteria substantially reduced the possible sample size. We finally decided upon a single criterion; we wanted to analyze Research, Development, Test and Evaluation (RDT&E) contracts only. The reason we exclude procurement or production contracts from our sample set is that certain predictor variables may work in contrary ways when considering RDT&E versus procurement contracts (Sipple, 2002). A clear example is changes to production quantity. Such an adjustment would directly affect the cost of a production contract but does not specifically drive the cost for an RDT&E contract. Due to the limitations of our population any additional criteria beyond RDT&E left us with, at most, fourteen contracts across three services (our sample contains 14 Electronic/Automated Software contracts; 7 Air Force, 4 Army, 3 Navy).

When working with acquisition program data for analysis purposes it is ideal to have 100% of the data. Not meeting this criteria, or some acceptable threshold such as 80%, is reason for removal from the sample set. We must clarify what is meant by having 100% of the data, there are two commonly accepted perspectives. One way to define percent complete is as function of the Budgeted Cost of Work Performed divided by the Estimate at Completion. Another way is to simply consider how far along the contract is relative to the stated contract start and finish date. In most cases our sample set is not complete in either sense. In fact, some of the sample programs in our database are only a small fraction of the total program. In normal circumstances this lack of complete programmatic data would be grounds for removal from the sample; if we followed this criterion, even at the relaxed standard of 80% complete or more, our sample drops to five programs. The previously discussed limitation of DCARC and desire to have a sample larger than five programs reduces our flexibility to exclude programs on these criteria. Later discussion of the five complete programs in our sample set references those programs with data coverage starting before 10% and extending beyond 80% of total program life.

In prior research we found other criteria for elimination including unidentifiable definitive date that work started, data missing for period earlier than 10 percent mark of the contract life, and contracts that went OTB. Prior research required that a contract include a start date and that the start date was prior to initial cost report submittal. This was not an issue for most of our sample set since the majority of history file reports available were truncated at the beginning. For the small number of programs in our sample that have history file data back to the beginning of the program we were able to pull the start date from the Format 3 and validate they matched. Two programs fell into the category of cost reports initiating prior to the official start we found on the Format 3, B2 EHF and H1 BOA. Nystrom noted that all regression-based models require a known, correct, start date to calculate EAC (Nystrom, 1995). However, for the two identified programs in our sample with discrepancies in the given start date we did not feel it should be a basis for elimination since our response variable was % EAC Growth; as calculated from the beginning of the history file data set to the end. The program start and end dates provided on the Format 3 served mostly as markers to compare our history file coverage against the actual program life, not for model building purposes.

We use similar reasoning to ignore the requirement that the program we are analyzing has complete data prior to the 10 percent mark of the program life. First, our sample set simply would not support this requirement. Only six programs in our final sample set had the complete set of data for the period earlier than 10 percent completion point. Second, because we defined our response variable as the % EAC Growth from the beginning of a program data set to the end the point at which the sample data actually began was arbitrary. If nothing else, we can develop categorical variables in the regression model which would group programs by where their data set lie on a normalized timeline of total program life.

Programs that have gone OTB are also important to screen for, not necessarily as criteria for removal but at a minimum as criteria for appropriate adjustments in our analysis. Analyzing a program that experienced an OTB poses problems for various reasons. If we wanted to develop growth models using time series EVM data and metrics such as ACWP, BCWP, BCWS, CPI, SPI, or others the OTB resets these values and impacts analysis. Although we do not intend on building a growth model, knowing that a program was OTB could be a significant predictor in our final model. In the DCARC EVM Application analysts can access a top level chronological overview of CPR data by program under the reports and metrics section. This overview contains a section for variance adjustments. In our review we found no variance adjustments recorded for any programs in our sample set. Knowing that information in DCARC is sometimes incomplete we wanted to further investigate our data set for signs of OTB. We reviewed the CPR data by month for times when the BCWS and BCWP matched the ACWP. We expect to see these trends early in a program's life but as the program progresses and risks manifest into tangible impacts on schedule and cost we expect the BCWS, BCWP, and ACWP to be different. We also looked at the CPI for each program by month in attempt to identify a sudden adjustment back to the baseline of CPI = 1 after significant variance. Within the limits of our historical data coverage, which will be discussed in the next section, we found no evidence of either trend; suggesting there were no OTB or other significant variance adjustments.

Although we did not find evidence of readjustments in our data set we did identify an anomaly with two programs in our sample. The program data sets in question begin on 4/30/2006, 21 months into the program life and 8/29/2004, 94 months into the program life respectively. The BCWS, BCWP, and ACWP match for these two programs during these months at the beginning of the data set. This raised concern since these exactly matching values occurred later in each program's life. We felt it was important to understand why these values would match at this point in a program's life, but we did not have prior CPR data in the history file. We thought that we could further investigate by reviewing the previous month CPR in DCARC but we were unable to do this. The file submissions in DCARC only go as far back as 2007 and 2006 for the programs, thus preventing us from seeing what the cost data looked like previous to the first CPR in our history file. For the context of our analysis we did not need to know why the values were the same at the beginning of the history file data since we developed our model based on the change from the first set of data in the history file to the last.

Our final sample contained thirty four contracts covering all services and nearly every military handbook type as seen in Table 3. Refer to Appendix B for a complete list of programs used in our analysis.

Table 3: Sample Set Service and Military Handbook Type

Sample Set Characteristics					
	Air Force	Army	DoD	Navy	Total by Military Handbook Type
Aircraft	2	1	1	5	9
Electronic/Automated Software	7	4		3	14
Missile		1		2	3
Ship				1	1
Space	3				3
Surface		1		1	2
System of Systems		1			1
UAV				1	1
Total by Service	12	8	1	13	Total Sample = 34

DCARC History Files Sample Data Coverage

Three factors affect the amount of data contained in each history file. First, of our available sample many programs have contract completion dates, as identified on the CPR Format 3, well in the future. The infancy of the EVM central repository plays a role in this factor; the consolidation of current cost data takes precedence over legacy information in this initiative and therefore we do not see information for previously completed programs loaded into the EVM application.

Second, the history files are not submitted to DCARC on a monthly basis, according to DoD, DCARC administrators the history files are generally updated and submitted annually. This delayed reporting impacts our ability to analyze the programs up to current date but is understandable. In fact, the delay is relatively insignificant because an analyst can still pull the individual, monthly CPR, or view the dashboard to see the most current program EAC and other information. Additionally, this delay in history file submission becomes less important for programs which have been in existence for an extended period of time. The important factor to note is that we have varying degrees of truncation at the end of our data set depending on how recent the available historical file is.

The third factor, one that seriously impacts the sample dataset, is a truncation in history file data coverage at the beginning of the program (and sometimes at the end, in excess of what we would expect to be missing due to annual history file reporting). As previously discussed there are a number of potential factors that drive this truncation including issues with software, loading legacy data, changes to the contract WBS and filing of program CPR files under different naming conventions.

As seen in Figure 4 our sample set history file coverage varied widely. This was a main driver in our inability to use complete programs as a data screening criteria. This characteristic of the data set also drove us to define our response variable, *% EAC Growth*, in a very specific manner.

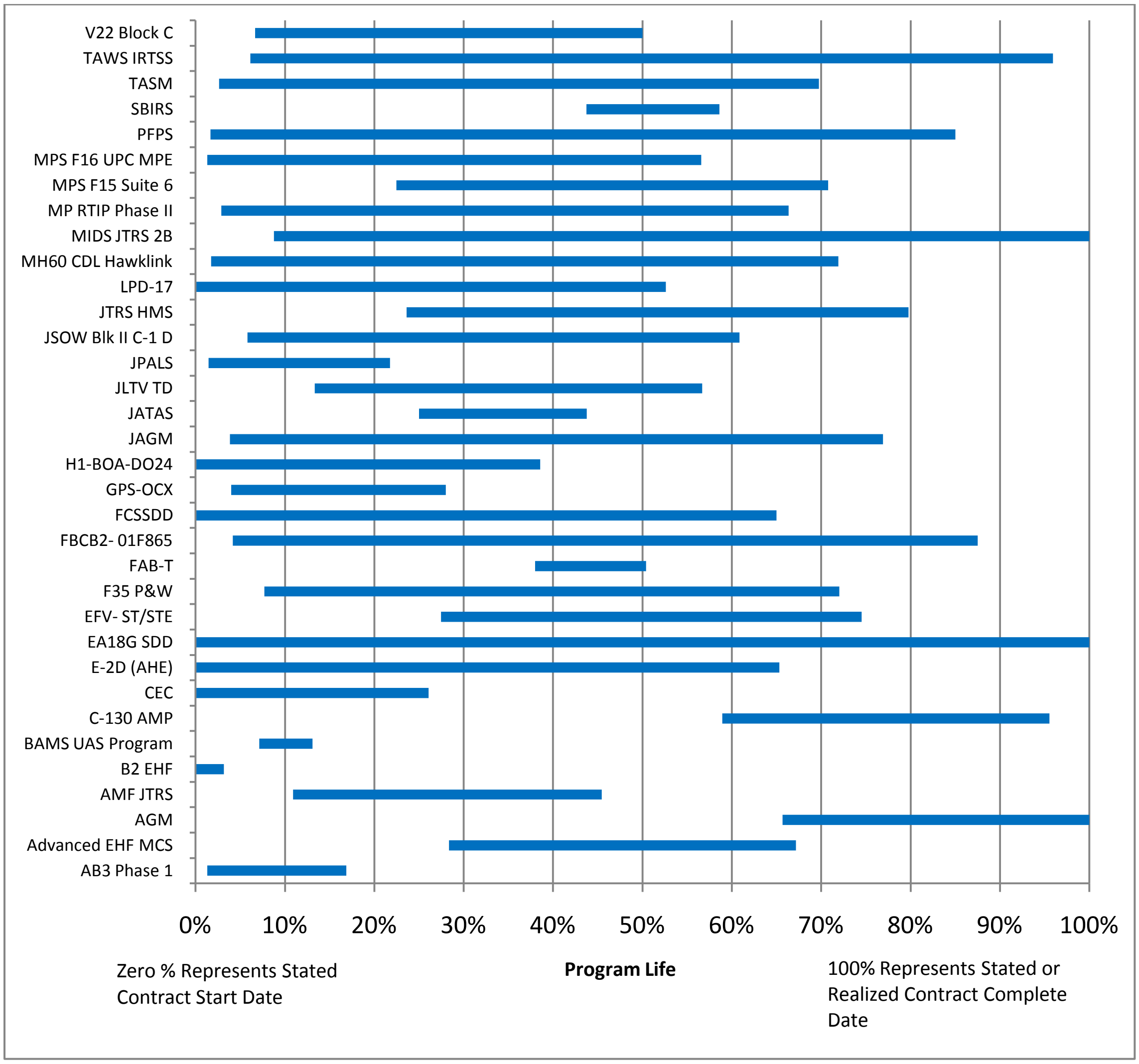


Figure 4: History File Data Coverage as a Percent of Total Program Life

Normalization of Data

After defining our sample set we normalized the data to a standard base year. Cost Performance Report data is submitted in Then Year dollars and therefore any analysis conducted without normalization would be skewed by inflation. Using the 2010 Weighted Indices, as seen in Appendix C, we adjusted the EVM data for each program to be in Base Year 2010 dollars.

The Response Variable: Percent EAC Growth

Previous regression work done with EVM data had the benefit of complete programs to review. Our sample set did not contain this characteristic which meant we had to craft our response variable carefully to handle this contingency. That is not to say that we were not able to collect either the final EAC (considering a program 80% or better as "complete"), or the most recent EAC. We were able to gather the most recent EAC from DCARC. However, we wished to avoid extrapolating our model beyond the constraints of the given history file data set. For this reason we decided to focus on the percent change in EAC, by program, from the first CPR in the respective program's history file to the last CPR. We calculated this percent change as the difference between the latest EAC and first EAC from the history file, divided by the first EAC. We named our response variable "*EAC % Growth*".

The added benefit of making our response variable a percentage is that it automatically normalizes for program magnitude. Because our sample set has budgets ranging from millions to billions of dollars the automatic normalization provided by analyzing percent growth in response and predictor variables greatly simplifies the process.

Defining the EAC

As we evaluated the sample set CPR Format 1 we noticed that there were inconsistencies in reporting the Best, Worst, and Most Likely (BWM) estimate at completion (as seen in block 6 subpart a, b, c). Some programs have logical values in place; a minimum below and a maximum above the most likely estimate. In other cases all three estimates were the same value. In some CPRs we saw all three blocks were blank. In addition to inconsistent reporting of the EAC we were also faced with the problem of varying methods. As we questioned DoD analysts responsible for different programs they quickly informed us that the methodology for developing the EACs on the Format 1 varied from program to program and even varied within the same program over time. A final layer of complexity surrounded the EAC on the Format 1, administration and overhead costs. The estimates include the administration and overhead costs applied to the contract after summation of the WBS elements. These additional costs include a Project Risk Adjustment (PRA), Administrative Costs, Undistributed Budget, and Management Reserve. This added another set of variables wrapped up in the EAC that were not consistent from program to program.

To prevent the need for normalization across all of our programs we decided to define our EAC in a way that is different than previously attempted. Using the CPR data in each program's history file we found the EAC as the summation of the lower level WBS elements. These EACs are prior to addition of a PRA or any administrative costs and prior to program specific adjustments related to the best, worst, or most likely values. We also felt this approach was better for our research purposes since it should potentially remove within program variability. Meaning the variability in the program EAC over time not related to the cost or schedule performance of the WBS elements. If the methods of developing the EAC for a program could and do potentially change, or if the management strategy or approach changes, and is therefore reflected in the administrative costs then the aggregate EAC suffers from this variability. Our intent was to model EAC growth using the lower level WBS information now available to us, it seemed logical to define our EAC in such a manner as well.

In order to validate that the values used for our EAC were consistent we looked for two things. First, the WBS breaks down into increasingly detailed levels of reporting, yet should remain consistent when we consider the reported values (EAC, BAC, ACWP, BCWP, BCWS) at any level when considering all elements of a given level. Therefore summation of all Level 4 elements under a given Level 3 should exactly match the values recorded for the individual Level 3 element. This pattern should continue up through the structure until we consider all Level 2 elements which, in summation, should equal the entire program value (excluding PRA and other adjustments). Validation of this consistency is simple at higher levels. However, as we look at the deepest levels of the WBS inconsistencies arise; summing all Level 5 elements, for example, does not match the sum of all Level 2 elements.

One clear reason for this inconsistency is that not every sublevel, below the overall program, has a subsequent sublevel. We can consider a simple example; our program has three Level 2 elements, two of the Level 2 elements contain two Level 3 elements, the third Level 2 has no sublevels. Summation of the Level 3 elements leaves out the value associated with the third Level 2 element and is inconsistent with the overall program value. Therefore, summation at the lowest WBS structure of a program could potentially not account for a significant portion of the program. Due to this we chose to calculate our monthly EAC values as the sum of all Level 2 (highest level) elements found in the DCARC history file.

Secondly, we compared the WBS elements and their structure present in the DCARC history file to the WBS elements on the Format 1 of the program to ensure consistency. In this way we made sure that the history file values and structure matched what the contractor was reporting on the Format 1.

Predictor Variables

We chose predictor variables based on prior research in the field of regression and developed new variables which use information now available to us. The main predictor variables are related to EVM values such as ACWP, BCWP, BCWS, and EAC associated with lower level WBS elements within the program. We included cost, schedule, and performance metrics of the lower level WBS elements in question as well. We also wanted to investigate the use of variables that may stand as proxies for program complexity such as WBS size, or total number of WBS elements, and WBS depth. Although we were unsure as to the predictive nature of these elements we decided to include them for exploratory purposes, especially since this type of predictor variable has not been analyzed in previous work we reviewed. Finally, we collected information on the number and magnitude of EAC adjustments across all Level 3 WBS elements, regardless of what element the adjustment occurred in. We wanted to test if the just the presence of activity, measured in number and magnitude of EAC changes, in the Level 3 WBS structure of a program was predictive of EAC growth.

Development Test and Evaluation

In 2002 Sipple identified a relationship between program cost growth and cost growth in the Engineering Manufacturing and Development phase of an acquisition program. His work established a baseline for further analysis by showing a relationship between overall cost growth and a specific area of expenditures within the acquisition program. Based on his success in modeling the EAC for development programs we felt that it was appropriate to consider the research and development phase of an acquisition program. Where our analysis differs from his is, instead of analyzing top level expenditures we wished to explore relationships of program cost growth and the lower level WBS elements. The calculation of our predictor variable was done in the same way as our response variable. We took the latest EAC for the specific WBS element and subtracted the first EAC for that element. Then we divided the difference by first EAC to show percent change. We decided to focus on the Development Test and Evaluation WBS element, or "*Dev EAC Growth*" as we named it, for three reasons.

First, we felt that the element is deep enough in the WBS structure to be considered significantly different from the overall, top-level cost data. Second, the Development Test and Evaluation element is a suggested element in all WBS structures, as provided by the DoD Handbook 881, except for the Space system structure. Our approach for dealing with the fact that the suggested space WBS structure does not contain Development Test and Evaluation is discussed in detail later on. Third, of the programs in our sample set the Development Test and Evaluation element was identifiable in the majority of our sample set history files and contained expenditure data. We initially intended on analyzing Operational Test and Evaluation as well for comparison purposes. However, we found that the majority of programs in our sample contained no expenditure data for this element and therefore excluded it. The DoD Military Handbook 881 suggested WBS structure for each weapon system type in our sample set is found in Appendix C.

Multiple Regression

We used the Ordinary Least Squares (OLS) multiple regression methodology to develop, analyze, and assess our models. The F-test, which evaluates the overall model, and T-test, which evaluates the individual parameters, were compared against at a significance level of .05 (alpha). Because we sought to maximize the predictive capability of our model yet we wanted to avoid "over-fitting" we consider the Adjusted R^2 value for any models built with more than one parameter. The Adjusted R^2 accounts for the artificial inflation of R^2 as more predictor variables are added to the model.

To analyze for normality we plotted the residuals, fit a normal distribution against them and used the Shapiro-Wilkes test. We evaluated the model's residual by predicted values plot to identify trends that are indicative of non-constant variance. Finally, we used the Cook's Distance plot to determine if there were overly influential data points in our models. Our data set is a collection of data over time; however our response variable and predictor variables are defined in a manner that makes them unrelated to time. For this reason we did not analyze our models for independence.

Analysis of our regression results led us to utilize a logarithmically transformed model as well. We applied a logarithmic transformation to our response variable, *EAC % Growth*, and the main predictor variable, *Dev EAC Growth*. The F-test and T-test were used again to evaluate the significance of the overall model and parameter. In addition to the Shapiro-Wilkes test, and Cook's Distance overlay plot we used the Breusch-Pagan test to evaluate our logarithmic model for constant variance of residuals.

IV: Analysis and Results

Chapter Overview

We intended to develop a regression model using the complete WBS EVM data to develop predictor variables. We wanted to specifically focus on a lower level WBS element to see if we could identify specific cost drivers that correspond to high EAC growth. Given the wide variety within our dataset we also hoped to identify significant differences in the statistical relationship of our lower level WBS element and overall program *% EAC Growth* based on service, military handbook weapon type or some other factor.

Our intent shifted due to the constraints of the dataset from a prediction model to a model that tested the relationship between overall program EAC growth and a lower level WBS element EAC growth. Analysis of the regression model diagnostics led us to investigate a non-linear relationship between *% EAC Growth* and *Dev EAC Growth*. The results of of a logarithmically transformed model show a significant relationship between the natural log of the change in reported DT&E EAC and the change in total reported EAC.

Preliminary Data Analysis

Upon initial review of our response variable distribution the predominant characteristic is a highly skewed right tail. We can attribute this to a number of potential causes. First, our sample set is small. Second, although we attempted to take a random sample of acquisition programs our sample pool is limited. Only more recent programs are currently loaded in DCARC, which may have some underlying affect on the data itself. Additionally we have truncated data and we considered the possibility that the limited window of information had an effect on our distribution.

Figure 5 shows the histogram of the response variable for our sample set. For clarification, the horizontal axis is formatted as a number not a percent. Therefore, the cohort between zero and one is all the programs that experienced between 0% and 100% EAC growth. All others outside this bin, to the right, experienced EAC growth in excess of 100%.

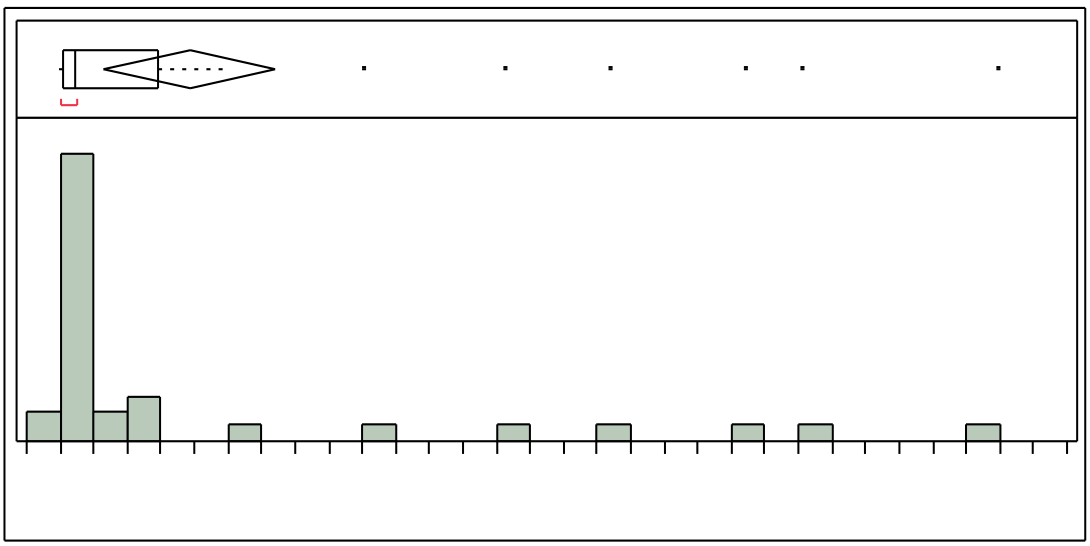


Figure 5: Histogram of Response Variable, EAC % Growth, November 2010

Of our sample, 12 programs or 35% of the total experienced EAC growth in excess of 100%. We see a highly skewed distribution to the right and seven significant outliers. The outliers seen here experienced 5 to nearly 28 fold increase in EAC *within* the window of the given programs data set.

Our first concern was that the "window" of data, specific to each program, was the cause for the highly skewed distribution of the response variable. Specifically we wondered if the outliers were those programs that had more complete history files, and therefore we actually captured more of the EAC growth. Of the seven programs identified, six have windows of data that cover greater than 50% of the stated program life. Of those six, one program had more than 65% of the program life represented, and three others had greater than 80%. One other program had approximately 5% of total life represented by the dataset, which was surprising. However, those programs with history file windows that reached back into the early stages of the program life, between 0% and 20% complete, all showed characteristics of drastic jumps in EAC. Additionally, of the five programs with data coverage from before 10% program life past the 80% mark of program life, two had EAC growth less than 100%. Of the entire sample more than a third of the programs with history file coverage greater than 50% of the entire program had growth less than 100%. These findings led us to believe that the distribution of our response variable was not related to the truncated history files. The program characteristics are summarized in Table 4.

Table 4: Percent EAC Growth Extreme Outliers, November 2010

Extreme Outlier Characteristics			
	Air Force	**Army**	**Total by Military Handbook Type**
Aircraft	1		1
Electronic/Automated Software	4	1	5
System of Systems		1	1
Total by Service	5	2	**Total Sample = 34**

The overwhelming characteristics are Electronic/Automated Software programs and Air Force programs. We already planned to test the impact of different weapon system types on the model using dummy variables, but this initial finding suggested that

Electronic/Automated Software programs and Air Force programs would prove to be predictive elements in our models.

We wanted to further analyze the cohort of less extreme EAC growth programs to determine if they exhibited behavior that is consistent with a normal distribution. We temporarily removed the seven extreme outliers to see what the grouping of response variables looked like.

The result is similar to the initial histogram of our entire sample set, only differing in magnitude of difference between the new cohort and outliers. As shown in Figure 6, similar to the histogram of our total sample we see a grouping with a skew to the right and outliers.

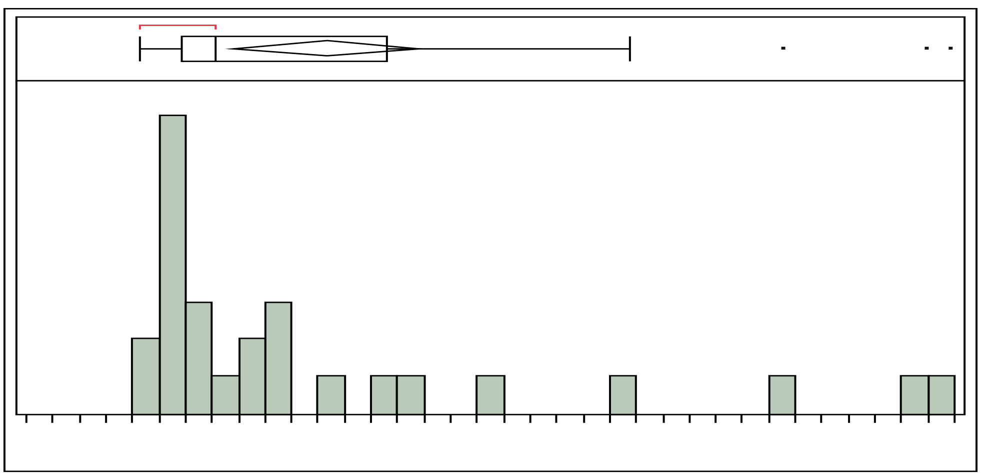


Figure 6: Histogram of Programs with <500% Growth, November 2010

Of our total sample 75% of the program history files cut off before the 80% complete mark. Of that 75% another 15 programs, or nearly 60%, do not have history file data beyond the 60% complete mark. For those reasons we wanted to observe if the growth patterns exhibited by the history data remained consistent to current date. Accessing the latest CPR submission event in the DCARC portal we are able to see a dashboard summary of the program, including the most recent EAC. We took the most recent EAC for each program and calculated the percent EAC growth from the first CPR in our history file dataset to the most recent value. The new percent growth calculations showed adjustments in percent growth for all the programs but had similarities to the percent growth associated with the DCARC dataset.

Adjustments in overall growth did occur and the trends were similar to what we saw in from the window of data. Of the total sample, six programs that did not have EAC greater than 100%, as calculated from the history file data, jumped to over 100% growth when the latest EAC was considered. Those programs which showed excessive EAC growth, identified as the seven outliers in our first histogram, either showed continued growth or stabilized at growth in excess of 500%. We see the distribution considering the current program EAC snapshot in Figure 7. The two distributions show similar trends including skew toward higher percent EAC growth and consistency in which programs are outliers.

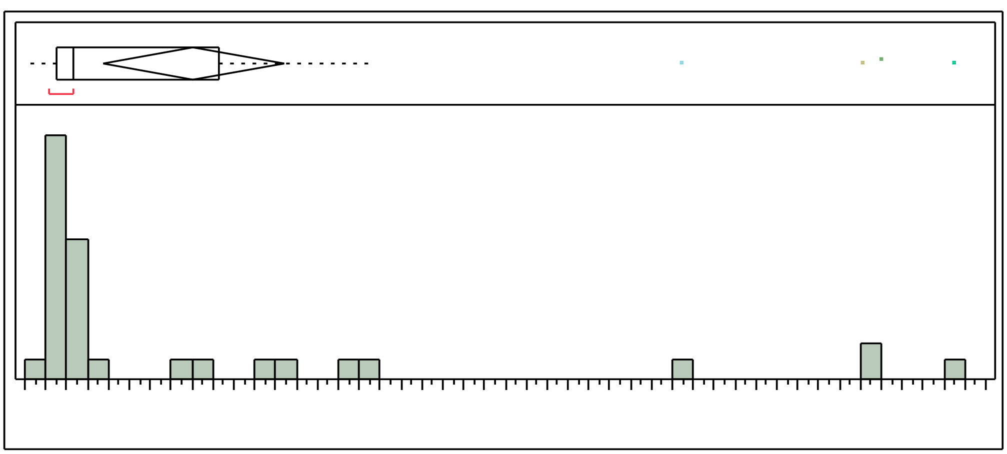


Figure 7: Percent Growth in EAC by Program using Current EAC Snapshot from DCARC, December 2010

The fact that the data show consistency in which programs are outliers and the overall skew of the distribution increases our confidence that the percent EAC growth as found in the historical file data is an accurate representation of the program behavior.

Preliminary analysis of the predictor variable, Percent EAC Growth of Development Test and Evaluation, shows that it exhibits similar behavior. Figure 8 shows the four excessive outliers in DT&E growth are four of the seven programs identified as excessive outliers in the response variable category. Of the remaining three of the seven excessive growth programs, two programs had incompatible WBS structures and we were not able to pull the DT&E information. However, the other program showed DT&E growth that fell on the right, skewed, side of the distribution at 327%. Of the total sample 11 programs showed DT&E growth greater than 100%, very similar to the 12 programs whose historical file overall EAC growth showed greater than 100%.

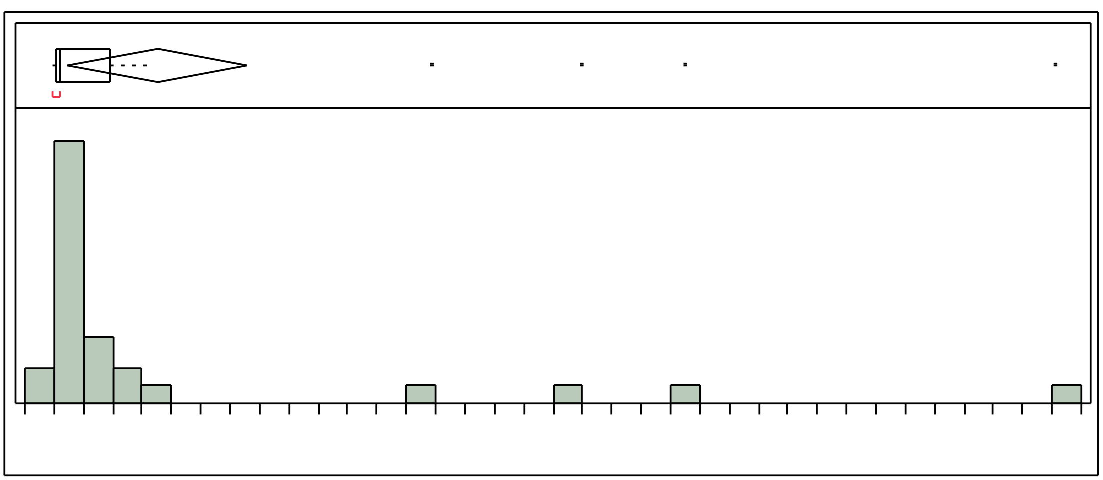


Figure 8: Percent Growth in Development Test and Evaluation by Program, November 2010

Analysis of Proxy Variables

We considered using variables that may be proxies for program complexity as predictors. This is based on the expectation that the more complicated a program is the more cost growth it will experience, which would reflect in the EAC. We tested if the total number of individual WBS elements, the depth of the WBS structure (lowest level), or count of Level 3 elements might relate to program EAC growth. Additionally, we wanted to test if activity of low level elements, measured as the count and magnitude of changes in the EACs, was predictive of overall program EAC growth. Multivariate analysis at a significance level of .05 (alpha) showed no significant relationship

Characteristics of Complete Programs in Sample Set

Our sample contained five programs with complete data sets. We wanted to briefly review the behavior of these programs to see if they followed the traits we assume are inherent in acquisition programs. We normalized all five programs and plotted their cumulative expenditures together. The plots, shown in Figure 9, do not have perfect s-shaped curves but generally the form of these expenditure patterns matches our expectations regarding the behavior or acquisition programs.

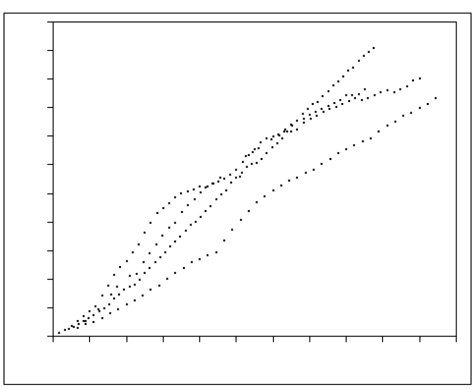


Figure 9: Actual Cost of Work Performed by Time for Complete Programs

Percent Change in Program WBS Level 3 EAC

Review of the EVM data for each program showed that at subsequently lower WBS levels the amount of recorded data diminished rapidly. The WBS structure is established and reported from month to month but there is no BCWP, ACWP, BWCS, EAC, or any other relevant values recorded. Our intent was to use the previously unavailable lower level WBS information to build a model yet we must balance this with the fact that we need data points to conduct analysis. Our review found that all of the history files in our sample with Level 4 elements and beyond had large gaps of data not recorded. It appears that the WBS structure is established but not all of it is used, especially as we look at the lowest levels. Additionally, similarities between programs at the fourth level became harder to identify as the unique structure of each sample become more and more apparent.

Using the DoD Military Handbook 881, Work Breakdown Structures for Defense Materiel Items, we analyzed if the program WBS aligned with the provided format. Appendix A contains all suggested WBS formats for the weapon system types present in our sample set. Another driving factor in us choosing to focus our analysis on the third level of the WBS structure was the following DoD 881 guidance:

> *WBS elements which are common (i.e. Integration, assembly, test and checkout; systems engineering/program management; system test and evaluation; training; and data) should be applied to the appropriate levels within the WBS for which they support. For example, if systems engineering is required to support a Level 3 WBS element, the systems engineering WBS element would appear at Level 4 of the WBS under the Level 3 element it supports*

Therefore, we expect and do find highly variable WBS structures by program below Level 3. For these reasons we decided to focus on the third WBS level in each program. We felt this was deep enough to significantly differentiate this analysis from previous aggregate level attempts yet the third level is populated with enough data points to support analysis.

Analysis of Common Level 3 Elements

In discussions with DoD DCARC analysts we were presented with the idea that if we focus on certain elements of the WBS architecture we may be able to find a predictor for EAC growth. Identification of such an element that corresponds with overall EAC growth would be a vector for future management and analysis attention. Areas commonly focused on include the research, test, and development portions of a contract since these areas inherently represent the risk associated with developing a new system. Using the WBS structure provided by DoD Handbook 881 we focused on the Level 2 element System Test and Evaluation; specifically, we chose to use the sub element Development Test and Evaluation. This was also a prime candidate for evaluation since the majority of our sample set WBS structure contained DEVELOPMENT Test and Evaluation and the CPR reported values for EAC, ACWP, BCWP, and BCWS.

Initial review of our sample set work breakdown structures showed little to no commonality. This lack of commonality held true even within the same service, military handbook system type, and contractor. Understandably the differences could be related to varying naming convention between service, contractor, and even different divisions of the same contractor. In addition to different nomenclature used for the respective elements we also found that the WBS structure, at the same level across history files, was highly variable in comparison to each other and the suggested format.

This difference prevented us from analyzing a single lower level variable across all programs, essentially putting a halt to our original research intent. In order to adjust for this roadblock we had to make a few important assumptions. First, some of the program history files have WBS elements that perfectly or near perfectly match the suggested format at Level 3; both in naming convention, structure, and number of elements. For example, the Apache Block III program history file contained a WBS structure that matched the suggested structure one for one, plus three additional elements: system engineering, program management, and Integrated Logistic Support management. While not a part of the suggested Level 3 structure the presence of the additional elements makes sense, they fall under the Level 2 element System Engineering/Program Management, and had no impact on our ability to identify the proper element for analysis. Because of this WBS structure the DT&E costs for AB3 and a few other programs was very easy to identify.

Other programs had close matches, whether differences were in naming convention or number and type of elements present in the history file Level 3 structure. In these instances we sorted by the Level 2 element, System Test and Evaluation, and subsequently identified if the naming of the Level 3 sub elements were close in nomenclature. For example, the Space Based Infrared Radar program has a Level 3 element named “Sys DT&E”; although not exactly titled “Development Test and Evaluation” we assume this is only a difference in naming. In other programs the Level 3 elements were not a match for the suggested guidance; further research found that in some program history files the elements matching the suggested Level 3 structure were recorded at a lower level. We chose to use these elements based on the assumption that the history files can be different depending on the report format and structure used by the contractor.

For example, the C-130 AMP history file contains one element at Level 1, "C-130 Summary WBS". This matches our common concept of the WBS structure which considers the overall program to be Level 1. As previously mentioned in our discussion on defining the EAC, we compared the history files to the CPR Format 1. This comparison revealed that most of the history files data actually began at the second WBS level, in other words the history file coded the elements Level 1 but in fact they represented Level 2 in the WBS. In the case of the C-130 history file, the coding actually contained an overall, system level WBS element in Level 1. Therefore, we were confident that the Development Test and Evaluation as identified in a lower level in the history file was actually representative of the suggested Level 3 WBS Development Test and Evaluation. A similar shift in the reporting structure to a lower level in the history file was seen in the H1 BOA program.

Finally we found that some program history files had a WBS structure that was completely incompatible with the suggested guidance and therefore we had to exclude their Level 3 Development Test and Evaluation % EAC Growth as a predictor variable. These programs include the B2 Extremely High Frequency SatCom Capability, F35 Pratt & Whitney Engine Development, Force XXI Battle Command Brigade and Below Program, Terrain Awareness and Warning System/Infrared Target Scene Simulation Program, and the LPD 17 Amphibious Transport Dock Ship. Three other programs had a WBS structure that matched the suggested format but had no recorded expenditures or associated EAC for DT&E.

Preliminary Model Results

For the preliminary model we wanted to analyze as many programs as possible in our sample set. Our intent was to garner a general idea of the relationship between overall EAC growth and EAC growth of Test and Evaluation. Then, in subsequent models, narrow the scope and focus to only those programs with a clear DT&E element. Analysis of the response variable, *EAC % Growth*, by the predictor, *Dev EAC Growth*, shows a strong relationship with an R^2 of .78, as seen in Figure 10. Full statistical results of the model can be found in Appendix F.

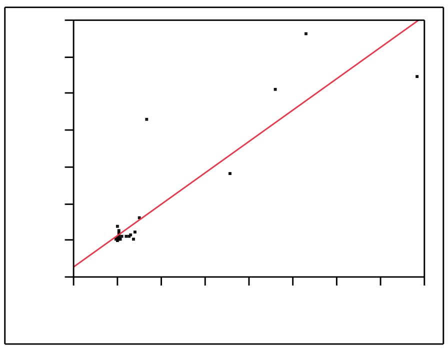


Figure 10: Percent EAC Growth of Programs by Percent EAC Growth in DT&E

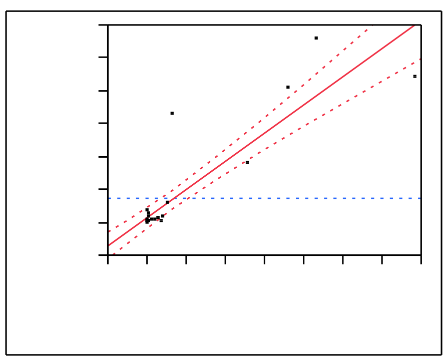


Figure 11: Preliminary Model Leverage Plot

Review of our leverage plot, Figure 11, identified potential influential data points and cohorts. Understandably, this preliminary model is subject to the most scrutiny. As previously mentioned we assume that Development Test and Evaluation in one program represents the same effort in another program in this model. Additionally, this model contains the most subjectivity based on our judgments of what expenditure data actually does represent Development Test and Evaluation in each program; while it was very clear in some programs in others it required interpretation using the suggested WBS structure, Contract Work Breakdown Structure Dictionary, and analysis of the history file structure. The CWBS Dictionary DT&E definitions for our sample are included in Appendix E.

Preliminary Model Diagnostics

Figure 12 shows a histogram of the preliminary model residuals. Using the Shapiro-Wilkes test we evaluated if the residuals followed a normal distribution. The Shapiro-Wilkes test returned a p-value of <.001, thus we must reject the null hypothesis that the residuals are normal.

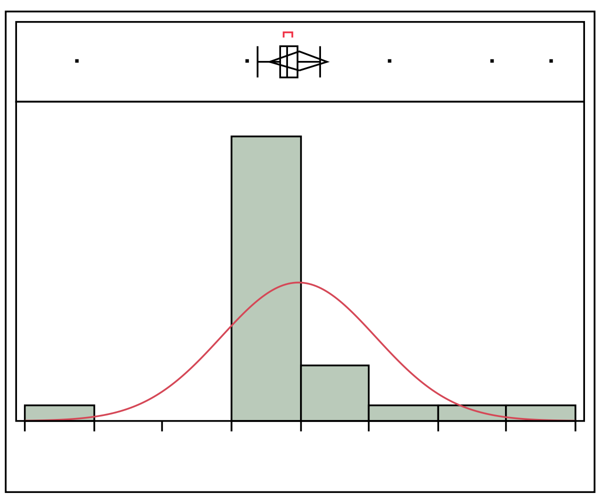


Figure 12: Histogram of Residuals from Preliminary Model

Figure 13 shows the residual by predicted plot from our preliminary model. We see a cone shaped pattern in the plot, suggesting that our residuals exhibit non-constant variance.

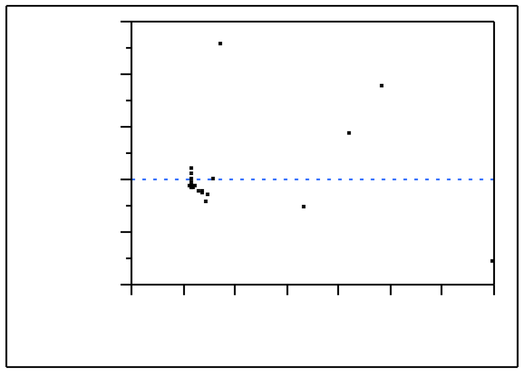


Figure 13: Preliminary Model Residual by Predicted Plot

Analysis of the Cook’s D overlay plot shows two potential programs that have overt impact on the results of the model: Mission Planning System F-15 Suite 6 and the Tanker Airlift Special Mission programs. We the results of the Cook’s Distance overlay plot in Figure 14.

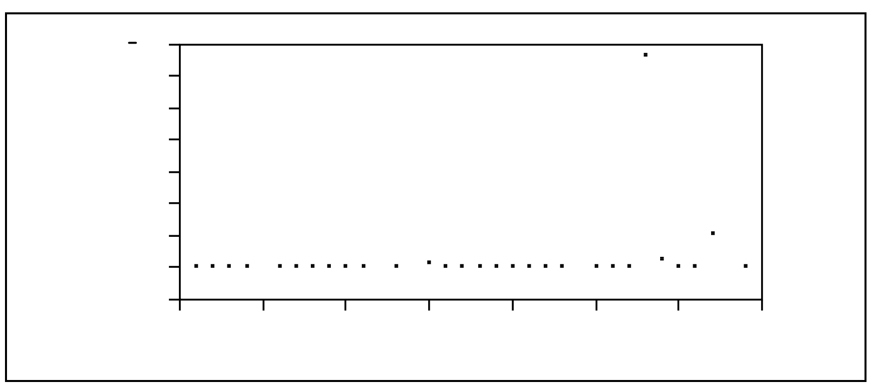


Figure 14: Cook's Distance Overlay Plot from Preliminary Model

To determine the true impact on the model we removed each program separately and evaluated how the model changed in each case. Removal of the Tanker Airlift

Special Mission program, which had a Cook's Distance value of 1.06, did little to change the model. The Adjusted R^2 shifted from .78 to .76 and the predictor variable *Dev EAC Growth* still had a significant p-value of <.0001, leading us to believe that the Tanker Airlift Special Mission program is not overly influential.

The other influential sample, Mission Planning System F-15 Suite 6, showed considerably higher potential for adverse influence on the model with a Cook's Distance value of 6.67. Removal of this program alone improved the overall predictive capability of the model, from Adjusted R^2 of .76 to .82, but did nothing to solve our issues of constant variance or normality. Additionally, we cannot arbitrarily remove an observation simply because it improves the predictive capability of our model.

Secondary Models

Preliminary data analysis, specifically our histogram of the response variable *EAC % Growth*, showed that we may have issues with normality and constant variance. Diagnostics of a preliminary model validated this assumption. In the secondary iteration of model development we attempt to distinguish between the cohorts in our data set and come closer to meeting our OLS assumptions using categorical variables. Additionally, we look at the impact of narrowing our scope to just those programs with specified Development Test and Evaluation WBS elements to determine how much our subjective interpretation of the WBS structure played a role in the Preliminary model.

First we focused our attention on the two programs which, based off of Cook's Distance plot, showed the most potential for influence over the model. We conducted further investigation of these two influential programs to determine why their results stand out in comparison to the other programs. We returned to our assumptions regarding the WBS and asked if the cost data we recorded and analyzed for these two programs was actually representative of the same effort in other programs. The WBS structure of the MPS F15 Suite 6 history file contains a System Test and Evaluation element with a Development Test and Evaluation sub element.

Both the overall program EAC and the DT&E EAC show similar trends in growth, suggesting this particular data point is a valid part of our sample. To test the sensitivity of our model we also calculated the adjustment in EAC of this element taking into consideration a later estimate in month 42 of the program that is closer in overall magnitude to the final DT&E EAC. The percent increase in EAC for Development Test and Evaluation using the later estimate from month 42 showed a significantly smaller increase, from over 34 fold to just 46%. Making no other adjustments to the Preliminary model aside from changing the value of the DT&E EAC percent increase from 34 times the original to just 46% increase has a drastic affect on the model. The Adjusted R^2 shows no predictive capability and the p-value for our predictor variable *Dev EAC Growth* is not significant. Further review of the program behavior showed that the overall EAC and EAC values for other elements exhibited a pattern of low EAC growth which spiked later in the program. For this reason we believed the initial DT&E EAC growth value of nearly 34 times was appropriate for analysis. Our research found no reason to change the values for the observation or remove it from our sample set.

The second influential program, Tanker Airlift Special Mission, was excluded from the secondary model because we could not identify which element under System Test and Evaluation represented their Development Test and Evaluation efforts. This same approach was used to further narrow entire sample set to only programs that have a specific WBS element Development Test and Evaluation. We wanted to see if the results drastically changed when we exclude those programs with no specific DT&E element.

The Secondary model has an Adjusted R^2 of .99 when we include MPS F15 Suite 6, but we know this program has an overt influence and therefore we removed it to see how the model would change. The new model, with a sample size of 17 programs, has an Adjusted R^2 of .83 suggesting a strong relationship between the response and predictor variables. Full statistical results of the model can be found in Appendix G. The residuals shifted further from a normal distribution; however this was not completely unexpected since decreased sample sizes affect normality. Additionally the residual by predicted plot shows an even clearer trend of non-constant variance as seen in Figure 15.

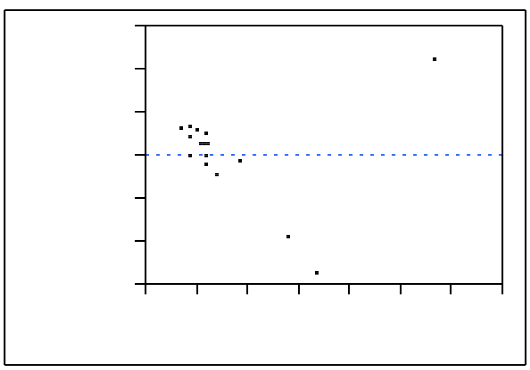


Figure 15: Secondary Model Residual by Predicted Plot

We attempted to use categorical variables to further improve both the Preliminary model, including all 29 programs, and at the Secondary model using our set of 17 programs. Admittedly, using more than one predictor for a sample set of 17 programs fails to meet the 10 observations to every 1 parameter rule. However, we wanted to identify if any variables improved normality and non-constant variance of our residuals.

In our preliminary analysis we saw that the majority of programs with excessive EAC growth were Electronic/Automated Software systems. We also saw that categorical grouping of programs by military weapon system type Aircraft showed significant differences in the mean using one-way Analysis of Variance. Using dummy variables, we grouped our sample set by military handbook type and service to test the model sensitivity when these categorical variables were included. Our results found that accounting for Electronic/Automated Software programs or Aircraft programs was not significant; both parameters failed to reject the null hypothesis at .05 (alpha).

Logarithmic Transformed Model

We identified an issue with heteroskedasticity in our preliminary and secondary models. To address the non-constant variance we took the logarithmically transformed values of our response and predictor variables and Figure 16 shows the transformed model, the model has a R^2 value of .56, with F-test and T-test both significant at .05 (alpha).

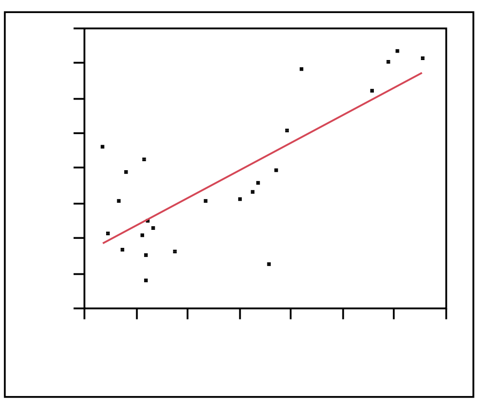


Figure 16: Logarithmic Transformed Model

We open our model up to the entire available sample of 29 program, however because we cannot calculate the log of a negative number or zero our sample size decreased. Two programs have negative overall EAC growth which prevented transformation of their response variable. Of those two one also showed negative EAC growth in the DT&E element, and the other showed zero growth. Two other programs showed a zero growth in DT&E. Of those two programs one had an overall EAC growth of 1.45%, very low in comparison to the total sample mean and median growth of 386% and 456% respectively. The other program had program growth of 298% but, due to an incompatible WBS structure, we were unable to identify neither the specific DT&E data nor summary level System Test and Evaluation data. Regardless, the trend shown in the negative growth of overall program EAC and DT&E mean provides further confidence that the relationship is strong even though we cannot include these programs in our log transformed model.

Transformation of the model data into log space helped improve the non-constant variance issues we saw in our residual and solved the problem of normality. Even for the programs we could not analyze, either due to incompatible WBS structures, or because the response or predictor variable showed negative growth; we were able to verify that the relationship seemed to hold. By that we mean, those programs with negative overall growth had either negative or zero DT&E growth. The number of programs with DT&E growth greater than overall program growth was 13. Of the programs with excessive EAC growth, we were able to model 5 of them in log space. Of those five programs three of them had DT&E growth in excess of overall growth. All but one had DT&E growth well above the mean of 351%, the one program that did not have growth above the mean was just below it at 327%. We evaluated this program for unique characteristics but the weapon system type, Electronic/Automated Software, contractor, service and other characteristics were not unique in comparison to the other excessive growth programs.

Figure 17 shows the studentized residual plot, and Figure 18 shows the residual by predicted plot. The model passed Shapiro-Wilkes test with a value of .88, failing to reject the null hypothesis that the sample is from a normal population. The predicted by residual plot shows a large improvement from the linear models. We see do not see the cone trend that was predominant in the earlier models.

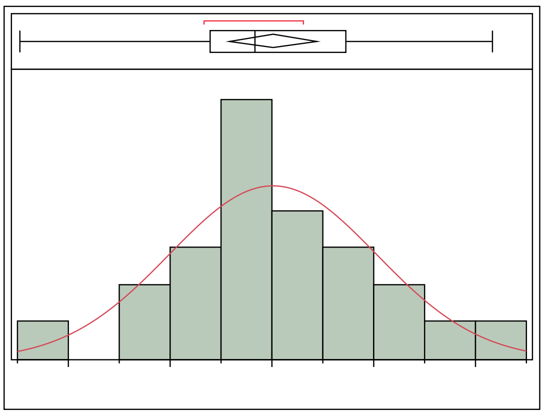


Figure 17: Studentized Residuals from Logarithmic Transformed Model

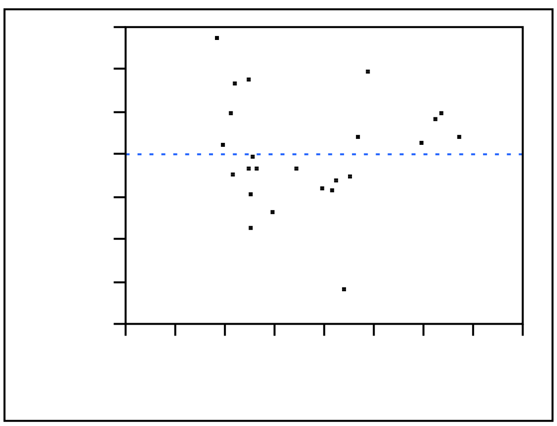


Figure 18: Residual by Predicted Plot of Logarithmic Transformed Model

The Cook's Distance overlay plot reveals that we have no overly influential data points in our sample. As seen in Figure 19 none of the sample points have a Cook's Distance value greater than one.

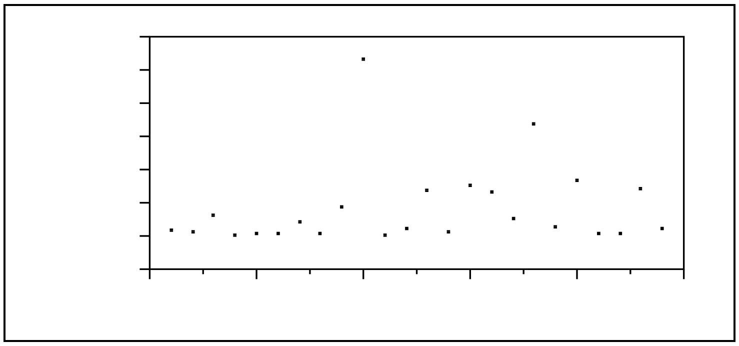


Figure 19: Cook's Distance Overlay Plot of Logarithmic Transformed Model

Finally, we evaluated our log transformed model using the Breusch-Pagan test. The results return a value of .51, thus confirming our residuals show constant variance.

Summary

Despite a small sample set and various data problems we were able to show a relationship between the overall program EAC growth and a lower level WBS element EAC growth. However, both the preliminary and secondary models failed to meet the expectation of normality and constant variance in the residuals.

Transformation of our model into log space revealed a non-linear relationship, improved the non-constant variance results and solved our issues regarding normality. We summarize that, in RDT&E contracts, the Level 3 WBS element Development Test and Evaluation is a significant driver for overall program EAC growth. This assessment is not limited to military handbook type, contractor, ACAT category, DAES group, or service.

V: Conclusions

Limitations

Our analysis hinges on the very significant assumption that the WBS structure of one program, and the cost data within that structure, can be compared to other programs. This is a tricky assumption to make between programs of the same weapon system type, service, phase, and, in some cases, contractor. This assumption incorporates increasingly more "unknown unknown" elements as the characteristics of our sample set expand beyond one service, weapon system type, and so on. Essentially, we are attempting to make a very specific diagnosis of what element is related to EAC growth yet from one program to another we are not sure the elements really are the same.

Given that we were not able to limit our sample to a set of programs with consistent characteristics (same weapon system type, service, and so on) we focused on using the suggested WBS structure as guidance. Admittedly, even after using the WBS structure as guidance we cannot say for sure that the work effort represented in the Development Test and Evaluation element of one program matches the effort of another. We can use the Contract Work Breakdown Structure (CWBS) as added guidance in the process of normalization; regardless the process is convoluted at best without close guidance from a subject matter expert to guide an external analyst. Our analysis was also limited and framed by the characteristics of the history file data set including truncation of the data at the beginning and end of the program.

Impact to the Acquisition Community

The centralized collection, reporting, and maintenance of EVM data is without question an additional burden on the acquisition community. The results of this analysis suggest that it is meaningful to collect this information and analysis conducted on it can be reliable. Therefore the role of the EVM centralized repository is justified and the initial groundwork for future research in the lower level WBS structure established.

We feel that the real value will manifest as subsequent research is done, using larger sample sizes, with more cohesive programmatic characteristics, and complete data. First, the comparative analysis of a WBS element and the overall EAC growth will provide historical context for the potential probability distribution surrounding a given WBS element. This has implications for developing EACs for new programs and potential uses in Monte Carlo simulation. Program managers and funding organizations seek “point” EAC estimates for their application in budget planning (Book, 2000). The trouble with rolling up the point estimate is that all WBS elements contain uncertainty and the simple summation of most likely estimates ignores that fact. Attempts to build input based simulation to derive an EAC regularly use subject matter expert opinion to provide a most likely estimate or a worst, best and most likely triangular distribution estimate. This methodology works better than summing the point estimate but can certainly be improved upon. If we have probability distributions for WBS elements developed and validated by analyzing the relationship between the elements and overall program behavior we can provide an increasingly accurate distribution of potential program costs.

Second, we feel that future analysis will identify WBS elements that are cost drivers specific to program characteristics. Our sample set was not specific to a weapon system type, service, and so on, but as the DCARC database grows increasingly robust sample sets should become available.

The combination of improved probability distributions for the purposes of Monte Carlo simulation and knowledge about specific WBS element cost drivers based on the program characteristics expand the cost estimator's toolset significantly. Improved EACs based on actual distributions provide a basis for rational decision making at the Program Manager level and above. Alternatively, knowledge of cost drivers provides program stakeholders a vector to focus their attention and program risk mitigation efforts.

Conclusion

Acknowledging that a major assumption regarding work effort represented by the DT&E element underpins our research; we conclude that analysis of the deeper WBS elements can be informative and beneficial in the field of EAC estimation. Despite high variability in our program sample set characteristics and a highly skewed distribution of our sample set response variable, percent EAC growth, we identified a statistically significant relationship between the overall program growth and Development Test and Evaluation WBS element.

We believe that future research potential can follow numerous paths. First, using subject matter expert input from different programs in a sample, attempt to organize, gather, and analyze lower level cost data related to a specific definition of work effort. This analysis could reveal how varied the grouping of work effort is between contractors and programs, even in the same WBS element. This analysis could also provide a robust distribution for a given element which could be applied to Monte Carlo EAC development. Second, because the EVM data is collected over time, there is potential for time series analysis between the lower WBS level elements and overall program. Using previously developed methods or creating a new method the relationship and behavior of the overall program can be modeled based on the lower level WBS elements.

Appendix A: DCARC EVM File Submission Example Status Report

Program	*Task*	*EDI Applied on CDRL*	CPR on Time	CPR Compliance	CFSR	IMS	CPR on Time	CPR Compliance	CFSR	IMS	CPR on Time	CPR Compliance	CFSR	IMS	*History File*
AMF JTRS – Joint Tactical Recone	Task 1														6/25/2010
BCS-F	BCS 3D Increment 3.2														1/26/2011
BCS-F	BCS SP3 Increment 3.1														1/26/2011
CEC – Cooperative Engagemen…	Sys Integrator/ Des Agent														6/14/2010
FAB-T – Family of Beyond Line	Task 1														10/26/2010
GCSS ARMY - Global Combat	T.O. 0001														11/24/2010

Legend			
CPR, CFSR, IMS SUBMISSION		CPR COMPLIANCE	
	Submitted on time		***Processes without errors***
	Submitted late		***Processes with minor errors***
	Submission in submitting status		***Multiple EDI files in 1 submission***
	Rejected-not resubmitted		***Multiple EDI files in 1 period***
	No submission recieved		***No EDI file recieved***
	Not required/Event not defined		***Not required/Event not defined***
	No data		***No data***

Appendix B: List of Programs in Sample Set

Program Name	Acronym	Military Handbook	Service
Apache Block III System Design and Development	AB3	Aircraft	ARMY
B-2 Extremely High Frequency SatCom Capability	B2 EHF	Aircraft	AIR FORCE
C-130 Avionices Modernization Program	C-130 AMP	Aircraft	AIR FORCE
E-2D Advanced Hawkeye	E-2D	Aircraft	NAVY
EA-18G Electronic Attack Variant	EA-18G	Aircraft	NAVY
F-35 Pratt & Whitney Engine Development	F-35 P&W	Aircraft	DoD
H1 Upgrades	H1-BOA	Aircraft	NAVY
MH-60R Multi-Mission Helicopter Upgrade	MH60	Aircraft	NAVY
V22 Block C ECS and Weather Radar	V22	Aircraft	NAVY
Cooperative Engagement Capability	CEC	Electronic/Automated Software	NAVY
Family of Beyond Line-of-Sight Terminals	FAB-T	Electronic/Automated Software	AIR FORCE
Force XXI Battle Command Brigade and Below	FBCB2	Electronic/Automated Software	ARMY
Joint and Allied Threat Awareness System	JATAS	Electronic/Automated Software	NAVY
Joint Precision Approach and Landing System	JPALS	Electronic/Automated Software	NAVY
Joint Tactical Radio System Airborne & Maritime/Fixed Station	AMF JTRS	Electronic/Automated Software	ARMY
Joint Tactical Radio System Handheld	JTRS HMS	Electronic/Automated Software	ARMY
Mission Planning System - F-15 - Suite 6	MPS F-15	Electronic/Automated Software	AIR FORCE
Mission Planning System - F-16 -UPC MPE	MPS F-16 UPC	Electronic/Automated Software	AIR FORCE
Multi-Functional Information Distribution System	MIDS/JTRS	Electronic/Automated Software	ARMY
Multi-Platform Radar Technology Insertion Program	MP RTIP	Electronic/Automated Software	AIR FORCE
Portable Flight Planning Software	PFPS	Electronic/Automated Software	AIR FORCE
Tanker Airlift Special Mission	TASM	Electronic/Automated Software	AIR FORCE
Threat Awareness and Warning System / Infrared Target Scene Simulation	TAWS IRTSS	Electronic/Automated Software	AIR FORCE
Advanced Anti-Radiation Guided Missile	AGM	Missile	NAVY
Joint Air to Ground Missile	JAGM	Missile	ARMY
Joint Stand-Off Weapon	JSOW	Missile	NAVY
LPD-17 Amphibious Transport Dock Ship	LPD-17	Ship	NAVY
Advanced Extremely High Frequency Mission Control System	EHF MCS	Space	AIR FORCE
GPS Next Generation Control Segment	GPS OCX	Space	AIR FORCE
Space-Based Infrared System	SBIRS	Space	AIR FORCE
Expeditionary Fighting Vehicle	EFV	Surface	NAVY
Joint Lightweight Tactical Vehicle	JLTV TD	Surface	ARMY
Future Combat Systems	FCS	System of Systems	ARMY
Broad Area Maritime Surveillance Unmanned Aircraft System	BAMS UAS	UAV	NAVY

Appendix C: Weighted Inflation Indices

USAF Weighted Inflation Indices	
Based on OSD Raw Inflation Rates	
Base Year (FY) 2010	
OPR: SAF / FMCEE	
Date of OSD Inflation Rates:	11-Dec-09
Date of SAF/FMCEE Issue:	8-Jan-10

Fiscal Year	Research, Develop., Testing, Evaluation (3600)
2000	0.836
2001	0.848
2002	0.857
2003	0.868
2004	0.890
2005	0.913
2006	0.940
2007	0.965
2008	0.984
2009	0.997
2010	1.008
2011	1.021
2012	1.038
2013	1.055
2014	1.073
2015	1.091

Appendix D: DoD Handbook 881 WBS Structures

Aircraft System

Level 1	Level 2	Level 3
Aircraft System		
	Air Vehicle (AV)	
		Airframe
		Propulsion
		AV Applications Software
		AV System Software
		Communications/Identification
		Navigation/Guidance
		Central Computer
		Fire Control
		Data Display and Controls
		Survivability
		Reconnaissance
		Automatic Flight Control
		Central Integrated Checkout
		Antisubmarine Warfare
		Armament
		Weapons Delivery
		Auxiliary Equipment
		Crew Station
	Sys Engineering/Program Management	
	System Test and Evaluation	
		Development Test and Evaluation
		Operational Test and Evaluation
		Mock-ups/System Integration Labs (SILs)
		Test and Evaluation Support
		Test Facilities
	Training	
		Equipment
		Services
		Facilities
	Data	
		Technical Publications
		Engineering Data
		Management Data
		Support Data
		Data Depository
	Peculiar Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Common Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Operational/Site Activation	
		System Assembly, Installation and Checkout on Site
		Contractor Technical Support
		Site Construction
		Site/Ship/Vehicle Conversion
	Industrial Facilities	
		Construction/Conversion/Expansion
		Equipment Acquisition or Modernization
		Maintenance (Industrial Facilities)
	Initial Spares and Repair Parts	

Electronic/Automated Software System

Level 1	Level 2	Level 3
Electronic/Automated Software System		
	Prime Mission Product (PMP)	
		Subsystem 1...n (Specify Names)
		PMP Applications Software
		PMP System Software
		Integration, Assembly, Test and Checkout
	Platform Integration	
	Systems Engineering/Program Management	
	System Test and Evaluation	
		Development Test and Evaluation
		Operational Test and Evaluation
		Mock-ups/System Integration Labs (SILs)
		Test and Evaluation Support
		Test Facilities
	Training	
		Equipment
		Services
		Facilities
	Data	
		Technical Publications
		Engineering Data
		Management Data
		Support Data
		Data Depository
	Peculiar Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Common Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Operational/Site Activation	
		System Assembly, Installation and Checkout on Site
		Contractor Technical Support
		Site Construction
		Site/Ship/Vehicle Conversion
	Industrial Facilities	
		Construction/Conversion/Expansion
		Equipment Acquisition or Modernization
		Maintenance (Industrial Facilities)
	Initial Spares and Repair Parts	

Missile System

Level 1	Level 2	Level 3
Missile System	Air Vehicle	
		Propulsion (Stages I...n,]
		Payload
		Airframe
		Reentry System
		Post Boost System
		Guidance and Control
		Ordnance Initiation Set
		Airborne Test Equipment
		Airborne Training Equipment
		Auxiliary Equipment
		Integration, Assembly, Test and Checkout
	Command and Launch	
		Surveillance, Identification and Tracking Sensors
		Launch and Guidance Control
		Communications
		Command and Launch Applications Software
		Command and Launch System Software
		Launcher Equipment
		Auxiliary Equipment
		Booster Adapter
	Systems Engineering/Program Management	
	System Test and Evaluation	
		Development Test and Evaluation
		Operational Test and Evaluation
		Mock ups/System Integration Labs (SILs)
		Test and Evaluation Support
		Test Facilities
	Training	
		Equipment
		Services
		Facilities
	Data	
		Technical Publications
		Engineering Data
		Management Data
		Support Data
		Data Depository
	Peculiar Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Common Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Operational/Site Activation	
		System Assembly, Installation and Checkout on Site
		Contractor Technical Support
		Site Construction
		Site/Ship/Vehicle Conversion
	Industrial Facilities	
		Construction/Conversion/Expansion
		Equipment Acquisition or Modernization
		Maintenance (Industrial Facilities)
	Initial Spares and Repair Parts	

Sea System (Ship)

Level 1	Level 2	Level 3
Sea System		
	Ship	
		Hull Structure
		Propulsion Plant
		Electric Plant
		Command, Communication and Surveillance
		Auxiliary Systems
		Outfit and Furnishings
		Armament
		Total Ship Integration/Engineering
		Ship Assembly and Support Services
	Systems Engineering/Program Management	
	System Test and Evaluation	
		Development Test and Evaluation
		Operational Test and Evaluation
		Mock-ups/System Integration Labs (SILs)
		Test and Evaluation Support
		Test Facilities
	Training	
		Equipment
		Services
		Facilities
	Data	
		Technical Publications
		Engineering Data
		Management Data
		Support Data
		Data Depository
	Peculiar Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Common Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Operational/Site Activation	
		System Assembly, Installation and Checkout on Site
		Contractor Technical Support
		Site Construction
		Site/Ship/Vehicle Conversion
	Industrial Facilities	
		Construction/Conversion/Expansion
		Equipment Acquisition or Modernization
		Maintenance (Industrial Facilities)
	Initial Spares and Repair Parts	

Space System

Level 1	Level 2	Level 3	Level 4
Space System			
	SEIT/PM and Common Elements		
	Space Vehicle (1...n as required)		
		SEIT/PM and Common Elements	
		Spacecraft Bus	
			SEIT/PM and Common Elements
			Structures and Mechanisms
			Thermal Control Subsystem
			Electrical Power Subsytem
			Attitude Control Subsytem
			Propulsion Subsystem
			Telemetry, Tracking, and Command
			Spacecraft Bus Flight Software
		Communication / Payload	
			SEIT/PM and Common Elements
			Communication (1&ldots;n as required)
			Payload (1&ldots;n as required)
			Communication / Payload Flight Software (Flight 1&ldots;n as required)
		Booster Adapter	
		Space Vehicle Storage	
		Launch Systems Integration	
		Launch Operations & Mission Support	
	Ground (1...n as required)		
		SEIT/PM and Common Elements	
		Ground Terminal Subsystem	
		Command and Control Subsytem	
		Mission Management Subsystem	
		Data Archive/Storage Subsystem	
		Mission Data Processing Subsystem	
		Mission Data Analysis and Dissemination Subsystem	
		Mission Infrastructure Subsystem	
		Collection Management Subsystem	
	Launch Vehicle		

Surface Vehicle System

Level 1	Level 2	Level 3
Surface Vehicle System		
	Primary Vehicle	
		Hull/Frame
		Suspension/Steering
		Power Package/Drive Train
		Auxiliary Automotive
		Turret Assembly
		Fire Control
		Armament
		Body/Cab
		Automatic Loading
		Automatic/Remote Piloting
		Nuclear, Biological, Chemical
		Special Equipment
		Navigation
		Communications
		Primary Vehicle Application Software
		Primary Vehicle System Software
		Vetronics
		Integration, Assembly, Test and Checkout
	Secondary Vehicle	(Same as Primary Vehicle)
	Systems Engineering/ Program Management	
	System Test and Evaluation	
		Development Test and Evaluation
		Operational Test and Evaluation
		Mock-ups/System Integration Lab (SILs)
		Test and Evaluation Support
		Test Facilities
	Training	
		Equipment
		Services
		Facilities
	Data	
		Technical Publications
		Engineering Data
		Management Data
		Support Data
		Data Depository
	Peculiar Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Common Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Operational/Site Activation	
		System Assembly, Installation and Checkout on Site
		Contractor Technical Support
		Site Construction
		Site/Ship/Vehicle Conversion
	Industrial Facilities	
		Construction/Conversion/Expansion
		Equipment Acquisition or Modernization
		Maintenance (Industrial Facilities)
	Initial Spares and Repair Parts	

System of Systems

Level 1	Level 2	Level 3	Level 4
System of Systems			
	Aircraft System		
		Air Vehicle (AV)	
			Appendix A- Level 3
		Sys Engineering/Program Management	
		System Test and Evaluation	
			Development Test and Evaluation
			Operational Test and Evaluation
			Mock-ups
			Test and Evaluation Support
			Test Facilities
		Training	
		Data	
	Electronic/Automated Software System		
		Prime Mission Product (PMP)	
			Appendix B- Level 3
		Platform Integration	
		Systems Engineering/Program Management	
		System Test and Evaluation	
			Development Test and Evaluation
			Operational Test and Evaluation
			Mock-ups/System Integration Labs (SILs)
			Test and Evaluation Support
			Test Facilities
		Training	
			Equipment
			Services
			Facilities
		Data	

System of Systems Continued

Level 1	Level 2	Level 3	Level 4
	UAV System		
		Air Vehicle	
			Appendix H- Level 3
		Payload (1&ldots;.n)	
			Appendix H- Level 3
		Ground Segment	
			Appendix H- Level 3
		UAV System Integration, Assembly, Test and Checkout	
		Sys Engineering/Program Management	
		System Test and Evaluation	
	System of Sys Engineering/Program Management		
	System of Sys Test and Evaluation		
		Development Test and Evaluation	
		Operational Test and Evaluation	
		Mock-ups/System Integration Labs (SILs)	
		Test and Evaluation Support	
		Test Facilities	
	Initial Spares/Repair Parts		

UAV System

Level 1	Level 2	Level 3
UAV System		
	Air Vehicle	
		Airframe
		Propulsion
		Communications/Identification
		Navigation/Guidance
		Central Computer
		Auxiliary Equipment
		Air Vehicle Application Software
		Air Vehicle System Software
		Integration, Assembly, Test and Checkout
	Payload (1&.n)	
		Survivability
		Reconnaissance
		Electronic Warfare
		Armament
		Weapons Delivery
		Payload Application Software
		Payload System Software
		Integration, Assembly, Test and Checkout
	Ground Segment	
		Ground Control Systems
		Launch and Recovery Equipment
		Transport Vehicles
		Ground Segment Application Software
		Ground Segment System Software
		Integration, Assembly, Test and Checkout
	Integration, Assembly, Test and Checkout	
	Sys Engineering/Program Management	
	System Test and Evaluation	
		Development Test and Evaluation
		Operational Test and Evaluation
		Mock-ups/System Integration Labs (SILs)
		Test and Evaluation Support
		Test Facilities
	Training	
		Equipment
		Services
		Facilities

UAV System Continued

Level 1	Level 2	Level 3
	Data	
		Technical Publications
		Engineering Data
		Management Data
		Support Data
		Data Depository
	Peculiar Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Common Support Equipment	
		Test and Measurement Equipment
		Support and Handling Equipment
	Operational/Site Activation	
		System Assembly, Installation and Checkout on Site
		Contractor Technical Support
		Site Construction
		Site/Ship/Vehicle Conversion
	Industrial Facilities	
		Construction/Conversion/Expansion
		Equipment Acquisition or Modernization
		Maintenance (Industrial Facilities)
	Initial Spares and Repair Parts	

Appendix E: Contract Work Breakdown Structure DT&E Definitions

AB3
This element includes test and evaluation conducted to: a. Demonstrate that the engineering design and development process is complete. b. Demonstrate that the design risks have been minimized c. Demonstrate that the system will meet specifications d. Estimate the system's military utility when introduced e. Determine whether the engineering design is supportable (practical, maintainable, safe, etc.) for operational use f. Provide test data with which to examine and evaluate trade-offs against specification requirements, life cycle cost, and schedule g. Perform the logistics testing efforts to evaluate the achievement of supportability goals, the adequacy of the support package for the system, (e.g., deliverable maintenance tools, test equipment, technical publications, maintenance instructions, and personnel skills and training requirements, etc.).

AGM
This summary element refers to the T&E conducted to demonstrate or determine: (1) the engineering design and development process is complete, (2) the system will meet specifications and (3) that the engineering design is supportable for operational use. Specific DT&E tasks include, but are not limited to, the following tests: system, reliability, maintainability, wind tunnel, ARH/MMW seeker, weapon fuzing, guidance and control, environmental, launch platform hardware and software integration tests, software verification/validation, shipboard compatibility, safety, Electromagnetic Interference (EMI)/Electromagnetic Capability (EMC), insensitive munitions and captive/free flights.

AMF JTRS
CWBS Dictionary not provided on DACIMS.

BAMS UAS
This effort is planned, conducted and monitored by the developing agency of the DoD component. It includes test and evaluation conducted to: demonstrate that the engineering design and development process is complete, demonstrate that the design risks have been minimized, demonstrate that the system will meet specifications, estimate the system's military utility when introduced, determine whether the engineering design is supportable (practical, maintainable, safe, etc.) for operational use, provide test data with which to examine and evaluate trade-offs against specification requirements, life cycle cost, and schedule, perform the logistics testing efforts to evaluate the achievement of supportability goals, the adequacy of the support package for the system, (e.g., deliverable maintenance tools, test equipment, technical publications, maintenance instructions, and personnel skills and training requirements, etc.).

C130 AMP
Conduct DT&E activities to demonstrate that the C-130 AMP/CAAP modifications meet the requirements of the system specification.

CEC Design Agent
N/A, no expenditure data recorded.

E2 Advanced EHF
Demonstrate that the engineering design and development process is complete, that the design risks have been minimized, that the system will meet specifications.

EA18 G

CWBS Dictionary not provided on DACIMS.

EFV STE

N/A, no expenditure data recorded.

FAB-T

CWBS Dictionary not provided on DACIMS.

JAGM

This WBS element includes test and evaluation conducted to: (a) Demonstrate that the engineering design and development process is complete. (b) Demonstrate that the design risks have been minimized (c) Demonstrate that the system will meet specifications (d) Estimate the system's military utility when introduced (e) Determine whether the engineering design is supportable (practical, maintainable, safe, etc.) for operational use (f) Provide test data with which to examine and evaluate trade-offs against specification requirements, life cycle cost, and schedule (g) Perform the logistics testing efforts to evaluate the achievement of supportability goals, the adequacy of the support package for the system, (e.g., deliverable maintenance tools, test equipment, technical publications, maintenance instructions, and personnel skills and training requirements, etc.). This element specifically includes non-operational, full-scale, and scale model vehicles for wind tunnel, safe separation, and other safety of flight demonstrations.

JATAS

N/A, no expenditure data recorded.

JLTV TD

This effort is planned, conducted and monitored by the developing agency of the DoD component for the JLTV FoV. It includes test and evaluation conducted to: a. Demonstrate that the engineering design and development process is complete. b. Demonstrate that the design risks have been minimized c. Demonstrate that the system will meet specificationsd. Estimate the system's military utility when introduced e. Determine whether the engineering design is supportable (practical, maintainable, safe, etc.) for operational use f. Provide test data with which to examine and evaluate trade-offs against specification requirements, life cycle cost, and scheduleg. Perform the logistics testing efforts to evaluate the achievement of supportability goals, the adequacy of the support package for the system, (e.g., deliverable maintenance tools, test equipment, technical publications, maintenance instructions, and personnel skills and training requirements, etc.). Efforts specifically included in the this element for LM's TD contract are: first article and/or EMI testing of automotive components and subsystems; all aspects of the Final Analysis Report; Prototype Vehicles Inspection and Test Plan and PVIR; Certification Documentation; Vehicle Integration; Integrated Subsystem Testing (supplier); Shakedown Testing of each vehicle; LM field support to testing (FSRs); shipment of vehicles to test sites; live fire test support.

JPALS

Systems level test and integration activities conducted at contractor facilities. This element includes all subsystem level and system level testing conducted at the Contractor System Integration laboratory, as well as subsystem and system level integration activities.

MPS F15 Suite 6

This element includes establishment of the MPE test program and vertical integration of MPE hardware and software over the DO period of performance. It includes the following activities for all new software development in versions 1.3, 2.0, and 2.1: MPE testing using scenario-based MPE level tests and delivery of MPE test documents (Software Test Plan, Description, and Report). It also includes Contractor on-site monitoring of the Government DT of MPE/UPC versions 1.3, 2.0, and 2.1 and monitoring of defects into the IKC.

SBIRS

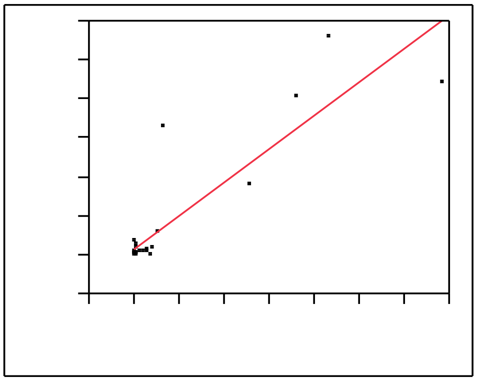
This included the effort to: (1) develop test requirements for system; (2) perform tests and compile test results to verify system requirements and confirm system performance capabilities; (3) perform early on-orbit test (EOT) of spacecraft and payload; (4) conduct ground and on-orbit system tests with Space and Ground Segment elements and interfacing elements external to the SBIRS system; (5) provide test equipment and interface connectivity specifically required for system tests; (6) develop test processes, plans, procedures and reports to accomplish the above activities, including the update and maintenance of the Integrated Test & Evaluation Plan (ITEP), the planning and coordination with AFOTEC on combined OT&E and DT&E test activities, and the activities associated with the readiness certification to enter IOT&E; and (7) perform risk assessment and risk mitigation measures addressing System I&T and Increment 3 ITW/AA certification.

V22 Block C

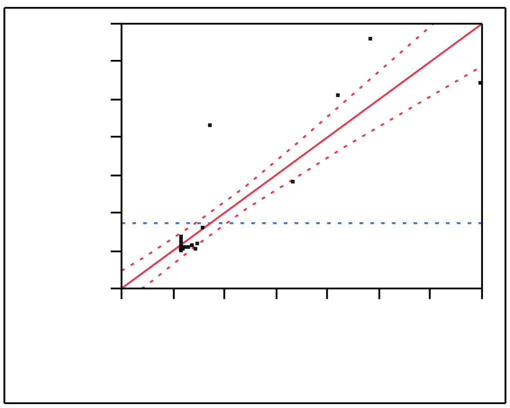
The Development Test element includes that contractor-conducted Test & Evaluation (T&E) held to (a) demonstrate that the engineering design and development process is reasonably complete, (b) ensure that all significant design problems have been identified and that solutions to these problems are in hand, (c) demonstrate that the system will meet specifications, (d) estimate the system's military utility when introduced, and (e) provide test data with which to examine and evaluate tradeoffs against specification requirements, Life Cycle Cost (LCC), and schedule.

Appendix F: Statistical Results of Preliminary Model

EAC % Growth by Development Test and Evaluation EAC % Growth



Actual by Predicted Plot



Summary of Fit

RSquare	0.78851
RSquare Adj	0.780677
Root Mean Square Error	3.606301
Mean of Response	3.696145
Observations (or Sum Wgts)	29

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	1309.1956	1309.20	100.6655
Error	27	351.1460	13.01	**Prob > F**
C. Total	28	1660.3416		<.0001*

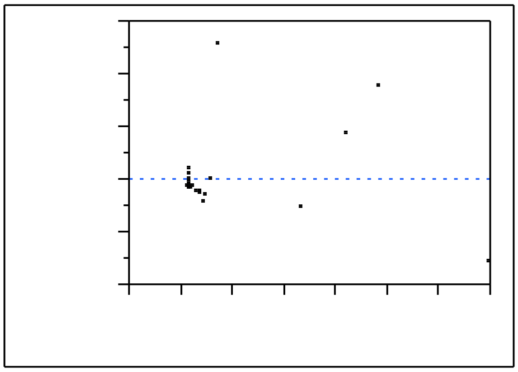
Lack Of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	25	351.13132	14.0453	1912.115
Pure Error	2	0.01469	0.0073	**Prob > F**
Total Error	27	351.14601		0.0005*
				Max RSq
				1.0000

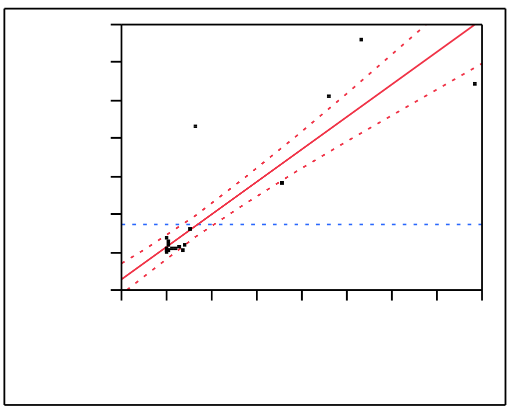
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.6887952	0.733693	0.94	0.3562
Dev EAC Growth	0.8556465	0.085281	10.03	<.0001*

Residual by Predicted Plot

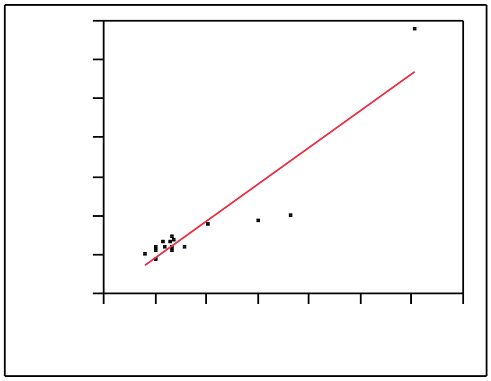


Dev EAC Growth
Leverage Plot

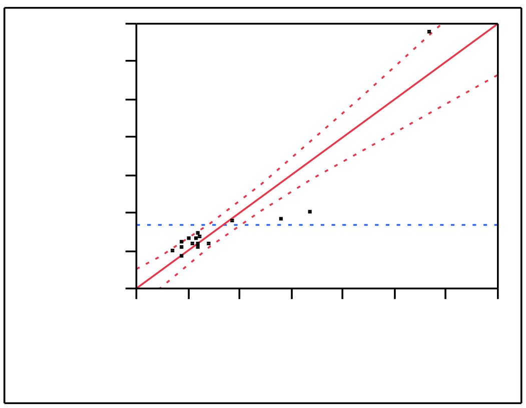


Appendix G: Statistical Results of Secondary Model

EAC % Growth by Development Test and Evaluation EAC % Growth



Actual by Predicted Plot



Summary of Fit

RSquare	0.847262
RSquare Adj	0.836352
Root Mean Square Error	0.283805
Mean of Response	0.32865
Observations (or Sum Wgts)	16

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	6.2551554	6.25516	77.6601
Error	14	1.1276335	0.08055	**Prob > F**
C. Total	15	7.3827889		<.0001*

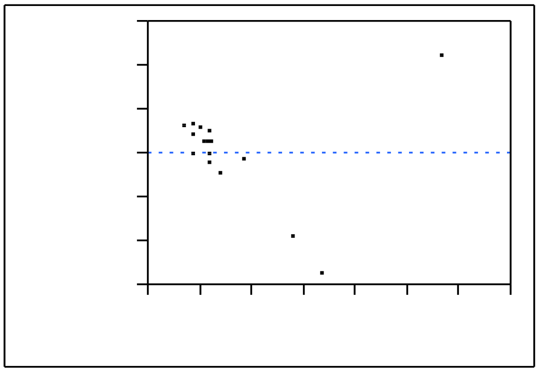
Lack Of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	12	1.1129427	0.092745	12.6263
Pure Error	2	0.0146908	0.007345	**Prob > F**
Total Error	14	1.1276335		0.0757
				Max RSq
				0.9980

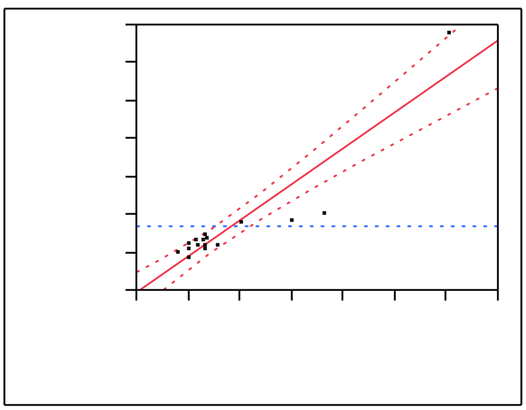
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0.061415	0.083626	-0.73	0.4748
Dev EAC Growth	0.949669	0.107764	8.81	<.0001*

Residual by Predicted Plot

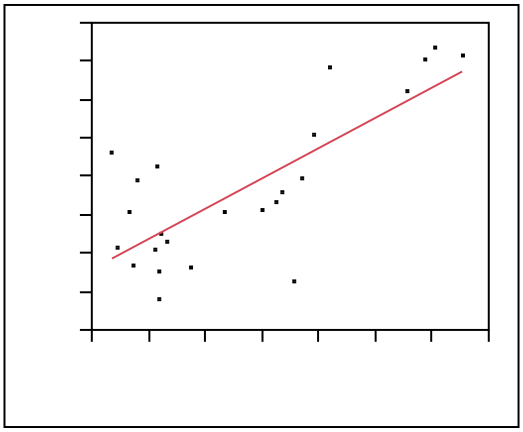


Dev EAC Growth
Leverage Plot

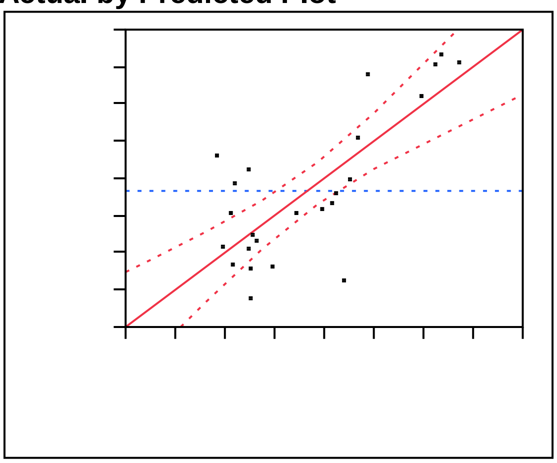


Appendix H: Statistical Results of Logarithmic Transformed Model

Log (EAC % Growth) by Log (Development Test and Evaluation EAC % Growth)



Actual by Predicted Plot



Summary of Fit

RSquare	0.585275
RSquare Adj	0.566424
Root Mean Square Error	1.323036
Mean of Response	-0.33868
Observations (or Sum Wgts)	24

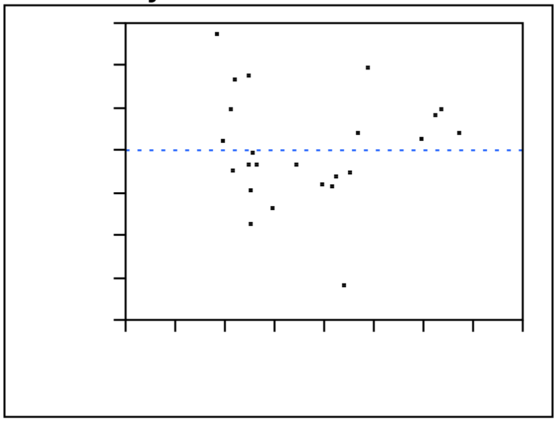
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	54.345761	54.3458	31.0472
Error	22	38.509339	1.7504	**Prob > F**
C. Total	23	92.855101		<.0001*

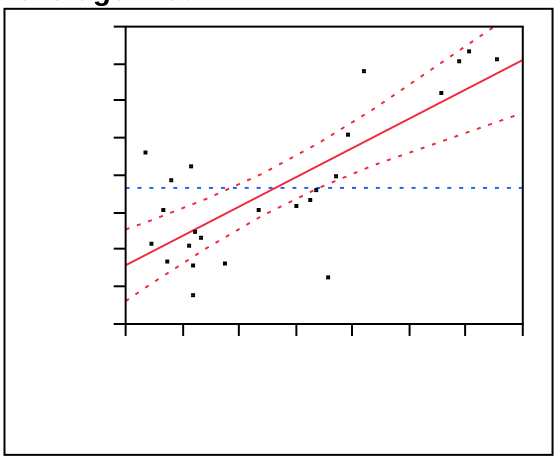
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0.054726	0.27483	-0.20	0.8440
Log(Dev EAC Growth)	0.7890038	0.141602	5.57	<.0001*

Residual by Predicted Plot



Log(Dev EAC Growth)
Leverage Plot



Works Cited

Augustus, Mike W. Office of the Secretary of Defense Acting Director of DCARC. Telephone interview. 7 February, 2011.

Book, Stephen A. *Do Not Sum Earned-Value-Based WBS-Element Estimates-at-Completion.* Society of Cost Estimating and Analysis. Los Angeles CA, June 2000.

Coleman, Richard et al. *Rayleigh Curves – A Tutorial.* Society of Cost Estimating and Analysis. Manhattan Beach CA, March 2004.

Defense Cost and Resource Center Portal (DCARC Portal). Office of the Secretary of Defense, Cost Assessment and Program Evaluation (OSD CAPE). n pag. http://dcarc.pae.osd.mil/EVM/Index.aspx

Department of Defense Handbook: Work Breakdown Structure. Appendix H, MIL-HDBK-881, 30 July 2005.

Holeman, J. B. *A Product Improved Method for Developing a Program Management Office Estimated Cost at Completion.* Defense Systems Management School, Fort Belvoir VA, January 1975.

Karsch, Arthur O. *A Cost Performance Forecasting Concept and Model.* Cost Research Report No. 117. Aeronautical Systems Division, Wright-Patterson AFB OH, November 1974.

Lee, David. *Norden-Raleigh Analysis: A Useful Tool for EVM in Development Projects.* Logistics Management Institute, The Measurable News. March 2002.

Maringas, Arthur. "WSA History Files." Electronic Message. 041800Z, 8 February 2011.

McKinney, John W. *Estimate At Completion Research – A review and Evaluation.* MS Thesis, Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1991.

Nystrom, Todd D. *A Comparative Study of Linear and Nonlinear Estimate At Completion Methods.* MS Thesis. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1995.

Scheinin, Warren. *Start Early and Often: the Need for Persistent Risk Management in the Early Acquisition Phases.*

Sipple, Vincent P. *Estimating Engineering Cost Risk using Logistic and Multiple Regression.* MS Thesis. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2002.

Thickstun, Kristine E. *Predicting Over Target Baseline (OTB) Acquisition Contracts.* MS. Thesis. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2010.

Tracy, Steven P. *Estimate at Completion: A Regression Approach to Earned Value*. MS Thesis. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2005.

Trahan, Elizabeth N. *An Evaluation of Growth Models as Predictive Tools for Estimates At Completion (EAC).* MS Thesis. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2009.

Unger, Eric J. *Relating Initial Budget to Program Growth with Rayleigh and Weibull Models.* MS Thesis. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2001.

Watkins, Harry. *An Application of Rayleigh Curve Theory to Contact Cost Estimation and Control.* MS Thesis. Naval Postgraduate School, Monterey CA, March 1982.

Weida, William J. A General Technique for R&D Cost Forecasting, USAFA-TR-77-12. US Air Force Academy Colorado, September 1977.

Winsor, Charles P. "The Gompertz Curve as a Growth Curve," *Proceedings of the National Academy of Sciences.* 18:1, 15 January 1932

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 074-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
03-15-2011	**Master's Thesis**	Sep 2009 - Mar 2011

4. TITLE AND SUBTITLE	
COMPARISON OF DEVELOPMENT TEST AND EVALUATION AND OVERALL PROGRAM ESTIMATE AT COMPLETION	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Rosado, William R., Captain, USAF	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Street, Building 642 WPAFB OH 45433-7765	8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/GCA/ENC/11-02
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of the Secretary of Defense, Cost Assessment and Program Evaluation Attn: Mr. Steve Miller Steven.Miller@osd.mil Room BE-829 1800 Defense Pentagon Washington, DC 20301-1800	10. SPONSOR/MONITOR'S ACRONYM(S) OSD CAPE 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Historically, cost growth regression models analyze aggregate, program-level information. Initiatives by the Office of Secretary of Defense, Cost Assessment and Program Evaluation (OSD CAPE) require direct, centralized reporting of the complete Work Breakdown Structure (WBS) Earned Value (EV) data. Centralized reporting allows access to unfiltered, unaltered, EV data for multiple programs. Using regression, we evaluate if WBS element Development Test and Evaluation (DT&E) EV data is related to program estimate at completion (EAC). Identifying a relationship provides evidence validating pertinence and reliability of low level EV data. Additionally, a relationship between a specific WBS element and program EAC establishes a basis for improved estimate development, and prediction capability. Our results show a strong relationship between DT&E and program EAC. Although limited by sample size and assumptions regarding DT&E commonality, our findings lead us to believe that there is potential for improved prediction models using low level WBS EV data.

15. SUBJECT TERMS

Development Test and Evaluation (DT&E), acquisition contracts, estimate at completion (EAC), work breakdown structure (WBS), Defense Cost and Resource Center (DCARC)

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Edward D. White, PhD (ENC)
U	U	U	UU	97	19b. TELEPHONE NUMBER *(Include area code)* 937-255-3636 extension 4540 Edward.White@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18