3-2004

# Logistic and Multiple Regression: A Two-Pronged Approach to Accurately Estimate Cost Growth in Major DoD Weapon Systems

Matthew B. Rossetti

**LOGISTIC AND MULTIPLE REGRESSION:**

**A TWO-PRONGED APPROACH TO ACCURATELY ESTIMATE**

**COST GROWTH IN MAJOR DoD WEAPON SYSTEMS**

THESIS

Matthew B. Rossetti, B.A.

First Lieutenant, USAF

AFIT/GCA/ENC/04-04

**DEPARTMENT OF THE AIR FORCE**
**AIR UNIVERSITY**

**AIR FORCE INSTITUTE OF TECHNOLOGY**

**Wright-Patterson Air Force Base, Ohio**

AFIT/GCA/ENC/04-04

# LOGISTIC AND MULTIPLE REGRESSION: A TWO-PRONGED APPROACH TO ACCURATELY ESTIMATE COST GROWTH IN MAJOR DoD WEAPON SYSTEMS

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

Matthew B. Rossetti, B.A.

First Lieutenant, USAF

March 2004

AFIT/GCA/ENC/04-04

LOGISTIC AND MULTIPLE REGRESSION:

A TWO-PRONGED APPROACH TO ACCURATELY ESTIMATE

COST GROWTH IN MAJOR DoD WEAPON SYSTEMS


Matthew B. Rossetti, B.A.
First Lieutenant, USAF


Approved:

| | |
|---|---|
| _______________//signed//_______________<br>Dr. Edward D. White (Chairman) | _30 Jan 2004_<br>date |
| _______________//signed//_______________<br>Michael A. Greiner, Major, USAF (Member) | _30 Jan 2004_<br>date |
| _______________//signed//_______________<br>Michael J. Seibel (Member) | _30 Jan 2004_<br>date |

## Acknowledgments

I could not have completed this endeavor had it no been for the undying support of my wife. Her patience and understanding are limitless. I thank my son, for reminding me when it was time to go fishing and I thank my Labrador retriever for begging me to take her for walks so I could clear my head.

I sincerely thank Dr. Edward White, my thesis advisor, for all his support and direction during this research. His guidance was indispensable. I thank my committee members, Mr. Michael Seibel and Major Michael Greiner for their constructive guidance on this thesis. Finally, I would like to thank the US Air Force for the opportunity to come to AFIT and learn.

Matthew B. Rossetti

## Table of Contents

Page

Page

## List of Figures

Page

## List of Tables

AFIT/GCA/ENC/04-04


## Abstract

With the ever changing threat to the security of the United States and a perpetually shrinking budget to provide this security, the defense acquisition community finds itself in the position of having to do more with less. For this reason, elected representatives, as well as higher ranking members of the Department of Defense (DoD) pay close attention to the cost performance of major defense acquisition programs.

We build on the previous research conducted by Captains Sipple, Bielecki, and Moore, who effectively demonstrate the use of a two-step logistic and multiple regression methodology to predict cost growth. This research confirms the usefulness of this two-step procedure for assessing cost growth in major DoD weapon systems.

We compile programmatic data from the Selected Acquisition Reports (SARs) between 1990 and 2002 for programs covering all defense departments. Our analysis concentrates on cost growth in the procurement appropriations of the Engineering and Manufacturing Development phase of acquisition. We investigate the use of logistic regression in cost growth analysis to predict whether or not cost growth will occur in a program. If applicable, the multiple regression step is implemented to predict how much cost growth will occur. Our study focuses on the *estimating* and *support* SAR cost variance categories within the procurement appropriations. We study each of these categories individually for significant cost growth characteristics and develop predictive models for each.

# LOGISTIC AND MULTIPLE REGRESSION: A TWO-PRONGED APPROACH TO ACCURATELY ESTIMATE COST GROWTH IN MAJOR DoD WEAPON SYSTEMS

## I. Introduction

### General Issue

Defense spending has undergone great change in the last 20 years. During the Reagan Administration of the 1980s, the Cold War saw high levels of defense spending. In 1985, the United States spent over $245 billion for national defense, a significant 25.9% of the President's Budget (OMB, 2004: 73, 78). The arms race with the former Soviet Union kept funding for weapon system acquisition flowing with relative ease.

As time passed, however, defense spending became heavily scrutinized as public perception of waste and excessive funding grew. In the years following the Cold War, particularly under the Clinton Administration of the 1990s, the United States experienced record-setting reductions in defense spending. By 2002, the budget for national defense hovered around $332 billion, a mere 16.5% of the President's Budget (OMB, 2004: 75, 80).

Unfortunately, global threats to the security of the United States have not declined in the past 20 years, merely changed form. This puts the defense acquisition community in the position of having to find ways to do more with less. For this reason, elected representatives, as well as higher ranking members of the Department of Defense pay close attention to the cost performance of major defense acquisition programs (MDAPs).

With each new administration, a movement to reform the Department of Defense's (DoD) major acquisitions programs and processes begins. This movement has gained serious momentum over the past decade. Major weapon systems being completed over budget and behind schedule is the motivation behind the current movement.

Cost growth in the procurement of major weapon systems can be attributed to poor program management or contractor inefficiencies, however, it mainly stems from risk and uncertainties about the program (Bielecki, 2003:2). In a 1993 RAND study, Drezner and others sought to characterize cost growth (variance between initial and final contract baselines) against a wide variety of factors. In general, they found that during the time period between McNamara's reforms 1965 and 1990, cost growth hovered at around 20 percent, on average.

In the last 15 years, the DoD has seen more reforms such as the Packard Commission of 1986, the Goldwater-Nichols Act of 1987, and the Acquisition Reform movement. In spite of claims that these reforms would lead to cost reductions, Air Force cost overruns grew another 9.9 percent (Suddarth, 2002:7). This 29.9 average cost growth is confirmed by the Assistant Secretary of the Air Force (Acquisition), Dr. Marvin Sambur, and the Deputy Chief of Staff for Air and Space Operations, Lieutenant General Ronald Keys, during their statement before the House Armed Services Committee on April 2, 2003 where they stated that for the Air Force, program execution problems had resulted in average cost growth of 30% for acquisition programs (Sambur/Keys, 2003).

In order for the DoD to retain its credibility with Congress and the American taxpayer, this cost growth must be slowed, contained, and reduced. DoD program managers must concern themselves with accurately identifying the cost risks associated with potential cost increases in their program cost estimates. To control cost growth, managers must focus on accurately assigning dollar values to risks, so that the original estimate from which we calculate cost growth is more accurate (Bielecki, 2003:2)

**Specific Issue**

The primary objective of weapon system cost estimating is to provide decision makers with an accurate estimate of the resources required to complete a project. To this end, cost estimators have many methodologies at their disposal: analogy, engineering, actual, and parametric.

The highly subjective analogy method compares a new system with an existing system for which there are accurate cost and technical data, and is most often used early in the program when little is known about the specific system being developed. Later in the program, the engineering estimate, commonly referred to as the "bottom up" method, is used when the scope of work is well defined and a comprehensive Work Breakdown Structure (WBS) is in place. Actual costs are used whenever they are available, but they are rarely available in the early stages of a program.

The parametric (statistical) method is used to analyze our data during this research. This method allows the cost estimator to objectively analyze large databases of historical data and make inferences about the relationship of the cost risk associated with one or more program parameters. The parametric technique is used early in the program to estimate cost risks throughout the life cycle of a program using statistical regression techniques to develop cost estimating relationships (CER).

Using regression to predict whether or not a program experiences cost growth, and the magnitude of that growth (should it occur) are the key focuses of this research. This study builds upon the thesis work of Bielecki (2003), Moore (2003), and Sipple (2002) to provide the cost estimating community a model to accurately estimate cost risk of the *estimating* and *support* cost variance categories of the procurement appropriations during the engineering and manufacturing development (EMD) phase of defense acquisition programs.

## Scope and Limitations of the Study

Fundamental to any discussion of cost growth is the Selected Acquisition Report (SAR); "Since 1969, Congress has required DoD to submit SARs on its major acquisition programs" (Calcutt, 1993:3). They are readily available and contain relatively reliable data on cost growth. As SARs are historically the foundation from which cost growth is analyzed, they are also the source of data for this study. The SAR contains the following three cost estimates useful for analyzing program cost growth:

- <u>Planning Estimate (PE)</u>: This is the DoD estimate normally made during the Concept Exploration and Definition phase of the acquisition cycle (Calcutt, 1993:3).
- <u>Development Estimate (DE)</u>: This is the estimate established at Milestone II, which begins the Engineering and Manufacturing (EMD) phase of the acquisition cycle (Calcutt, 1993:3).
- <u>Current Estimate (CE)</u>: This is the most up-to-date estimate of what the program will cost at completion (Calcutt, 1993:3).

The SAR reports cost variances in base year and then year dollars (allowing for analysis between programs on a constant dollar basis) and classified into one of the following seven categories:

1. Economic: changes in price levels due to the state of the national economy
2. Quantity: changes in the number of units produced
3. Estimating: changes due to refinement of estimates
4. Engineering: changes due to physical alteration
5. Schedule: changes due to program slip/acceleration
6. Support: changes associated with support equipment
7. Other: changes due to unforeseen events

(Drezner, 1993:7)

The security classification of some of the programs will limit our research. Any program with a confidential or higher classification will not be looked at in this study. Given that this type of information is not classified as confidential or higher on the vast majority of Major Defense Acquisition Programs (MDAPs), this limitation is viewed as having negligible impact on the utility of the model we build. Other limitations exist within the SAR which are discussed further in Chapter 3.

For the purposes of this research, cost growth is measured as a positive percentage increase from the DE to the latest CE as reported in the SAR. Furthermore, this research excludes cost growth due to changes in the economy and adjustments to quantity (the first two categories of cost growth reported in the SAR) since these two categories are beyond the control of the cost estimator.

Since we build upon the research previously fielded by Sipple, Bielecki, and Moore, we employ the same framework and methodologies initiated by Sipple and expanded by Bielecki and Moore. The difference being that this study focuses on the *estimating* and *support* cost variance categories of the procurement appropriation during the EMD phase of defense acquisition programs. In particular, this research builds logistic and multiple regression models with predictor variables from the EMD phase that predict whether or not a program experiences cost growth (logistic) and, if it exists, how much it experiences (multiple). Additionally, we utilize the database developed by Sipple (2002), update it to contain the latest CE (2002 data) of each program, if applicable, and add any new programs that are at least three years into the EMD phase (mature program).

**Research Objectives**

The purpose of this research is twofold. First, logistic regression (yes or no response) will be used to ascertain if there are certain parameters within the program that are able to predict if a program will experience cost growth in the *estimating* and *support* cost variance categories of the procurement appropriation during the EMD phase of program development. Second, if cost growth is present, multiple regression will be used to determine how much growth occurs.

**Chapter Summary**

This research expands the cost estimating methodology originally developed by Sipple, and further developed by Bielecki and Moore. Our specific goal provides the cost estimating community an effective model to estimate the cost risk associated with a program early in its development, and the overall goal reduces the DoD cost growth rate from its current levels. We continue with Sipple's two step methodology — analyzing SAR historical data with logistical and multiple regression to successfully predict cost growth in the EMD phase of program development. In the following chapter we present an overview of the acquisition process and its environment, examine cost risk and the effect it has on our study, and finally, investigate past research in cost growth.

# II. Literature Review

## Chapter Overview

This chapter establishes a historical framework from which to base our methodology and develop our models. First, we discuss the acquisition process, past and present, and how that process affects our approach in this study. Next, we look at the acquisition environment to familiarize ourselves with the increasing importance of these types of models. Cost risk and its considerations are addressed after the environment has been established. We conclude the chapter with a review of recent studies that have relevance to ours.

## The Acquisition Process

Being that this research focuses on a very specific portion of the overall acquisition process, we begin this chapter with a brief overview of how that process works and where our focus lies. To this end, we start with Department of Defense Instruction (DoDI) 5000.2 Operation of the Defense Acquisition System, which

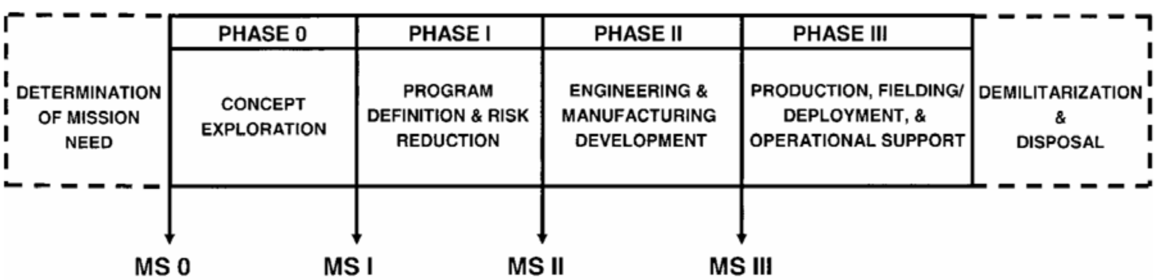



**Figure 2.1 – Old Acquisition Milestones and Phases (DoDI 5000.2, 2000:1)**

"Establishes a simplified and flexible management framework for translating mission needs and technology opportunities, based on approved mission needs and requirements, into stable, affordable, and well-managed acquisition programs that include weapon systems and automated information systems." (DoDI 5000.2, 2003:1).

Figure 2.1 is a graphical representation of what the Defense Acquisition Management Framework looked like prior to a January 2001 change to DoDI 5000.2. We include this past business practice because the SAR data in our database is based on this format. The process consists of four milestones (MS 0-MS III) and four phases (PHASE 0-PHASE III), described below. This information was extracted from the DoD 5000.2, prior to the Jan 2001 change.

- *Approval to conduct concept studies (MS 0)*- The Milestone Decision Authority (MDA) approves short-term concept studies and the PHASE 0 exit criteria.
- *Concept Exploration (PHASE 0)*- Evaluate the feasibility of alternative concepts, determine the most promising concepts and solutions.
- *Approval to begin new acquisition program (MS I)*- MDA approves the Acquisition Strategy, Cost as an Independent Variable (CAIV) objectives, initial Program Management Baseline (APB) and PHASE I exit criteria.
- *Program Definition and Risk Reduction (PHASE I)*- Design the system, demonstrate critical processes and technologies, and develop prototypes.
- *Approval to enter Engineering and Manufacturing Development (EMD) (MS II)*- Approval of Acquisition Strategy, CAIV objectives, updated

APB, Low-Rate Initial Production (LRIP) quantities, live-fire and Test and Evaluation (T&E) waiver (if applicable) and PHASE II exit criteria.

- *Engineering and Manufacturing Development (PHASE II)*- Mature and finalize selected design, validate manufacturing and production processes and test and evaluate the system.
- *Production or fielding development approval (MS III)*- Approval of Acquisition Strategy, production (weapon systems), deployment (information systems), updated APB and PHASE III exit criteria.
- *Production, Fielding or Deployment and Operational Support (PHASE III)*- Produce system, field it, monitor mission performance, support fielded system, modify or upgrade as required.

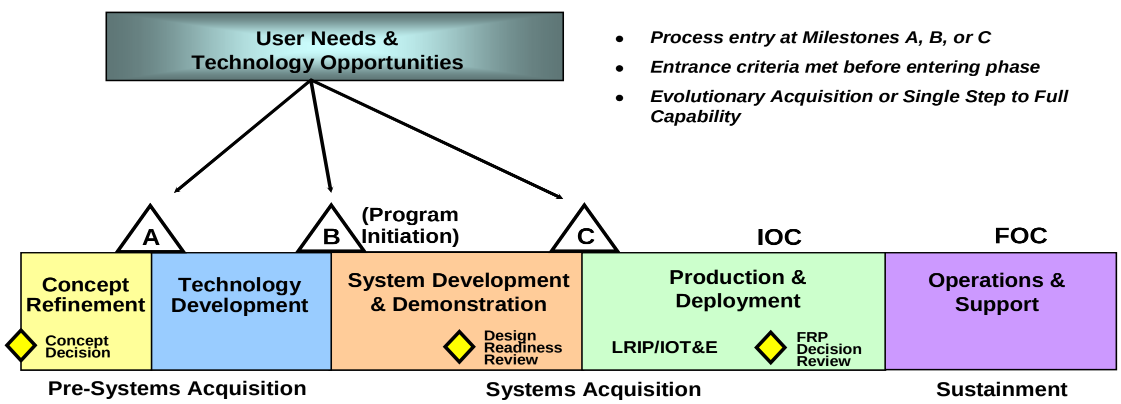



**Figure 2.2 – New Acquisition Milestones and Phases (DoDI 5000.2, 2001:1)**

Figure 2.2 is a graphical representation of what the Defense Acquisition Management Framework looks like now, due to the aforementioned change to the DoDI 5000.2 in January of 2001. It replaces the traditional milestones with an ABC format and labels the phases by name (as opposed to numbering or lettering them). The following is a brief overview of the new framework, taken from the current DoD 5000.2.

- *Concept Refinement Phase*- Refine the initial concept and develop a Technology Development Strategy (TDS). This phase cannot begin until the MDA makes a Concept Decision and does not mean that a new acquisition program has been initiated.
- *Milestone A*- MDA approves the TDS.
- *Technology Development Phase*- Reduce technology risk and determine the appropriate set of technologies that will be integrated into the full system. This process is iterative in that it assesses the viability of available technologies and refines user requirements simultaneously.
- *Milestone B*- The acquisition program has officially started. For programs using Evolutionary Acquisition (which will be described in more detail later in this chapter), each increment will have its own Milestone B. This is where the PM and MDA prepare and approve an Acquisition Strategy.
- *System Development and Demonstration*- Develop full or increment of capability, reduce integration and manufacturing risk, ensure operational supportability, implement human systems integration, and design for producibility.
- *Milestone C*- MDA commits the DoD to production and authorizes entry into LRIP, production and limited deployment for operational testing.
- *Production and Deployment Phase*- Achieve operational capability that satisfies mission needs, either incrementally or fully.

- *Operations and Support Phase*- The two major components of this phase are sustainment and disposal. The purpose being to ensure the system continues to perform its mission and is ultimately disposed of properly.

As you can see, we did not go into as much detail on the new acquisition framework as we did on the old. The reason for this is simple: our study is based on the old phases and milestones because all of our historical data (from the SARs) is based on the old process. It is also important to note at what point we focus on in the acquisition process. Figure 2.3 indicates the focus of our research.

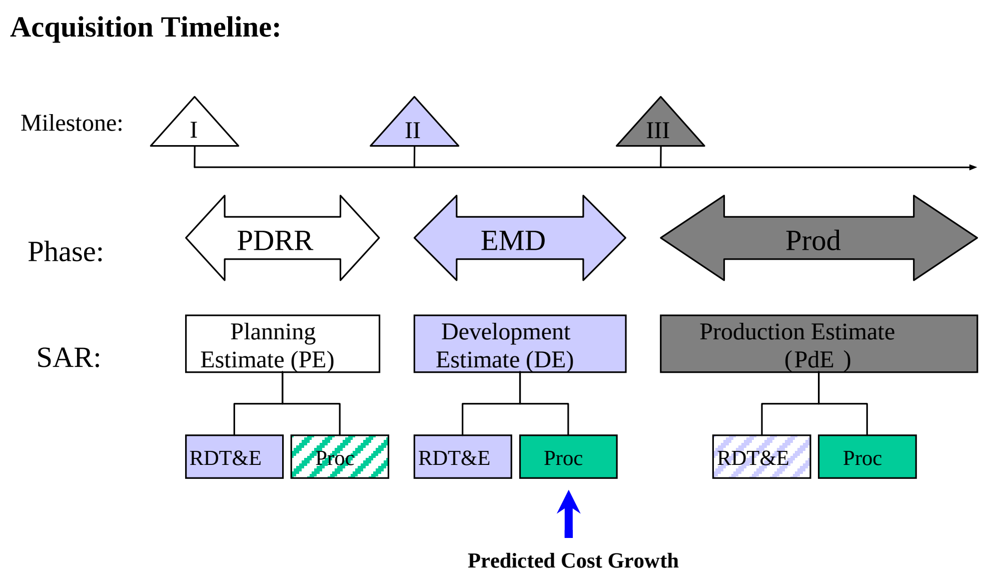



**Figure 2.3 – Acquisition Timeline (Dameron, 2001:4)**

Later in this chapter, we review the thesis work on this subject of our predecessors (Sipple, Bielecki and Moore). Sipple focuses on the *engineering* cost variance (CV) category and Bielecki on the *estimating, schedule, support,* and *other* categories of the RDT&E appropriation. While these studies target specific CV categories, Moore targets the overall procurement appropriation in the EMD phase. Our research focuses on the individual CV categories of *estimating* and *support*. We make the assumption that the cost estimator is more concerned with specific areas of cost growth.

**The Acquisition Environment**

The acquisition process is under great scrutiny as evidenced by the sweeping changes in the overall acquisition framework in January of 2001. The changes, however, do not stop there. The latest initiative to revamp the current acquisition process is traced back to September 2002 when the Secretary of Defense issued an unsigned memorandum stating that the current regulations were "overly prescriptive and do not constitute an acquisition policy environment that fosters efficiency, creativity and innovation." As a result, said the memo, the 5000 series, which includes versions 5000.1 and 5000.2, would be "cancelled ... effective immediately." (Erwin, 2002)

On 12 May of this year (2003), DoD Directives 5000.1 and 5000.2, were signed by the Deputy Secretary of Defense and replaced the same directives previously dated October 23, 2000. One of the policies instituted by this directive is that of cost and affordability:

> All participants in the acquisition system shall recognize the reality of fiscal constraints. They shall view cost as an independent variable(CAIV), and the DoD Components shall plan programs based on realistic projections of the dollars and manpower likely to be available in future years (DoD Directive 5000.1, 2003:4).

This policy indicates the importance of CAIV to program management and signifies the extent to which the OSD believes cost estimation should be used in budgeting. Realistic projections become extremely important in that appropriated funds are scarce and under heavy supervision by multiple stakeholders. In addition, when taken into account the number of government civil servants, military officers and enlisted troops that it takes to make funding changes, it is fair to assume that administrative costs due to poor planning are high, and could be reduced with more accurate initial estimates. For these reasons, each program manager must strive to get their cost estimations right, more often than not, so they can maintain their programs' credibility with DoD executives, Congress, and the American public.

The seriousness of this acquisition reform movement is echoed yet again in April 2003 when Dr. Marvin Sambur, Assistant Secretary of the Air Force (Acquisition), and the Deputy Chief of Staff for Air and Space Operations, Lieutenant General Ronald Keys, state before the House Armed Services Committee:

> In the past, we have designed our programs with a 60-70% confidence level of meeting cost, schedule, and performance goals. In order to be credible both to the warfighters and Congress, I have implemented a 90% confidence level in meeting our requirements. By demanding collaboration between all the parties, we can ensure the right trade-offs are made throughout the acquisition process to meet the required goals. It is imperative that, both the warfighting and acquisition communities work together to make tradeoffs of non-critical elements within programs to buy down risk, throughout the acquisition cycle. Bottom line: credibility means delivering what we promise, on time and on budget (Sambur/Keys, 2003).

Clearly, a major concern in the acquisition community is that of credibility and fiscal responsibility, and it would be difficult to have one and not the other. To obtain this credibility, the pressure is on the cost estimator to accurately predict the costs associated with the program at all phases of the system life cycle. This is no easy challenge. The methods available to the estimator range from subjective methods (quick and easy) to objective methods (time consuming and complex), both of which have their strengths and weaknesses, and both must address risk.

**Cost Risk**

"Risk: Minimizing the possibility that something goes wrong" (Cancian, 1995:191). Cancian's definition may appear oversimplified, but it's a great place to start. As cost estimators, much of the risk we encounter involves uncertainty. Uncertainty about the countless variables we identify, and uncertainty about the variables we fail to identify. These uncertainties have great potential to make "something go wrong" in our estimates. This is especially true when attempting to estimate the cost of a system that has not yet been built, or is in the process of being built.

A cost estimator must first identify and consider all areas of uncertainty associated with a system and related future events. Once identified and estimated, the cost risk is translated into a dollar figure which can then be used by decision makers. The Air Force Materiel Command (AFMC) *Financial Management Handbook* confirms "program risk refers to the uncertainties and consequences of future events that may affect a program", and goes on to say that "risk is the summation of probable effects of unknown elements in technical, schedule, or cost related activities within the program." The latter of these three risk parameters asks the question: "can the program as presently structured technically and with respect to schedule, be completed for the budgeted amount of money?" (*AFMC Financial Management Handbook*, 1998:11-20).

In the case of the Air Force's most expensive acquisition program, the Advanced Tactical Fighter (a.k.a. the F-22 Raptor), the answer to this question has historically been "no". This program is an excellent example of how uncertainty creates risk. Although there are countless factors (especially in the EMD phase) that can be held responsible for F-22 program cost growth, a very interesting uncertainty is worth mentioning. According to a 1999 GAO report, "A factor the Air Force did not consider in its estimate of potential cost growth was the possibility that the F-22 program may have to absorb a higher share of the manufacturing plant's overhead costs if the contractor does not sell enough C-130J aircraft, which are produced at the same plant as the F-22." (GAO/NSIAD-99-55, 1999:5). Ironically, this is a factor that the Air Force would have easily been able to predict (since C-130J is also a DoD acquisition program) had they realized its potential impact on cost growth.

The F-22 program is also an excellent example of what could be argued is a program's biggest risk of all: being cut. Funding instability is a fact of life that the F-22 has been dealing with for years. This is because "as threats began to change, developmental challenges arose, and total ownership costs continued to mount, it was unlikely to be overlooked as a prime source of funding for other 'must pay' bills." (Myers, 2002:322). The truth of this statement is easily reflected in the Defense Subcomittee's rationale behind their $1.8B cut in the 2000 Department of Defense Appropriations Bill:

> It is clear from a larger perspective, the F-22 is consuming resources that could be used to address other critical strategic concerns such as emerging threats from chemical/biological/nuclear terrorism, information warfare, and cruise missiles. (Defense Subcommittee, 2000)

The bottom line is that a cost analyst must deal with countless unforeseen events in order to protect their program's funding, and thus, the program itself. The AFMC Financial Management Handbook discusses three methods the analyst can use to approximate the likelihood of a certain event occurring: *a posteriori*, (after the fact), *a priori* (a prediction based upon theoretical probability distributions), or *subjective judgment* (*AFMC Financial Management Handbook*, 1998:11-21). No matter which method the estimator chooses, the end product will depend largely on the skill of the estimator, the level of accuracy required, the level of detail needed, and the time required (and available) to complete the estimate. These are also the factors that will determine how well an analyst mitigates risk when applying their chosen methodology.

We mentioned in Chapter 1 that the cost estimating community has different cost estimating methodologies at their disposal including, but not limited to, analogy, engineering, actual, and parametric. These methods are widely accepted and practiced in both the DoD and civilian sectors. Figure 2.4 shows the techniques recognized by the Ballistic Missile Defense Organization (BMDO) cost estimating community. These techniques are also widely accepted and practiced in most cost estimating communities. It is interesting to note that as the level of detail and difficulty of gathering the data increase, the techniques exhibit a diminishing level of precision.

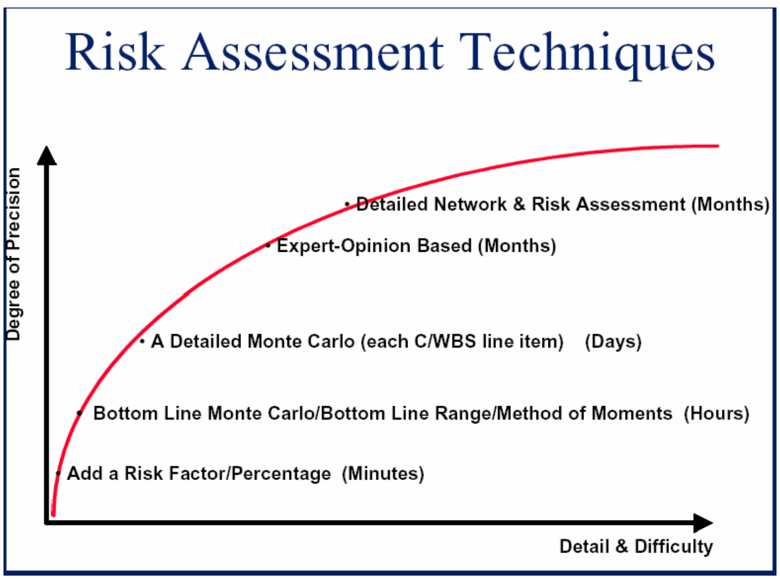



**Figure 2.4 – Risk Assessment Techniques (Coleman, 2000:4-9)**

In conclusion, risk needs to be addressed up front and early and the cost estimator's role in this process is crucial. This philosophy is made very clear by the Air Force Materiel Command (AFMC) *Financial Management Handbook*:

> Because resources are limited, considerable time and effort in planning for future acquisitions is necessary. The central issue in such planning usually concerns resource allocation. Cost analysis supports acquisition decisions required to allocate financial resources among alternative systems. The acquisition process revolves around the cost estimate - budgets are based on estimates and future cost performance is measured against estimates. Cost estimating must be accurate if the operation of the Planning, Programming, and Budgeting System (PPBS) is to be realistic, and effective decision making is to take place (*AFMC Financial Management Handbook*, 1998:11-2)

**Past Research in Cost Growth**

A benefit to doing continuing research on three comprehensive studies on cost growth is that the previous authors: Sipple, Bielecki, and Moore, provide us with an exhaustive review of the pertinent literature on cost growth from 1974 through 2001. Sipple's review of the literature was thorough enough that the follow-on work performed by Bielecki and Moore provides us with no relevant studies outside of their own findings. The important thing to note here is that the unique two-step methodology adopted by Sipple to identify and then quantify cost growth is tangent to existing studies on predicting cost growth.

Sipple provides us with twelve relevant studies on this matter, see Table 2.1. For a complete review of the studies listed refer to Sipple (2002). These studies influenced Sipple in his development and creation of the predictor variables used in both the logistical and ordinary least squares (OLS) models.

| Author (Year) |
| --- |
| IDA (1974) |
| Woodward (1983) |
| Obringer (1988) |
| Singleton (1991) |
| Wilson (1992) |
| RAND (1993) |
| Terry & Vanderburgh (1993) |
| BMDO (2000) |
| Christensen & Templin (2000) |
| Eskew (2000) |
| NAVAIR (2001) |
| RAND (2001) |

**Table 2.1 – Sipple Thesis (Sipple, 2002:20-44)**

*Sipple Thesis*

Where Sipple's methodology differs from previous studies is that Sipple looks at predicting cost growth in the EMD phase of the system life cycle instead of attempting to predict overall cost growth for an entire system life cycle. This approach affords us the ability to break down the cycle into its different phases: PDRR, EMD, and Prod and further into the appropriations contained in each and study the effects that over 75 predictor variables have on these appropriations given a particular phase. Sipple is also unique in that he recognizes that the Y response variable (*Engineering percent*) exhibits a mixed distribution. "About half of the distribution is continuous, while the other half is massed at one value, zero—indicating no cost growth. This mixed distribution scenario generally calls for splitting the data into two sets" (Sipple, 2002:58). We will utilize these same variables and two-step methodology in our approach to predict cost growth in the *estimating* and *support* cost variance categories of the procurement appropriations during the EMD phase of program development.

The goal of Sipple's research is to predict cost growth in the EMD Phase as it relates to RDT&E appropriations in the SAR *engineering* cost variance category. Sipple collects SAR data and builds a database of over 75 predictor variables using 115 major acquisition programs. He then uses logistic regression to first identify if cost growth exists. If it exists, OLS regression is implemented to indicate how much cost growth will occur. "Sipple demonstrates through the use of four regression models (A, B, C, D) that the combination of logistic and multiple regression produce similar predictive results as a traditional single-step multiple regression cost estimating methodology. However, the two-step methodology is preferred to the single-step methodology because of the stronger statistical foundation achieved with the two-step method" (Bielecki, 2003:21).

> We build four regression models that we briefly introduce in this paragraph. We build one logistic model using 90 data points. This model predicts whether a program will have engineering cost growth in RDT&E dollars. To simplify our analysis, we call this Model A. We then build three multiple regression models. We call Model B the model that we build from the 47 of the 90 data points that do have cost growth. We apply a log transformation to the response variable in this model to correct for heteroscedasticity in the residual plot. We build Model C as an alternative to Model B. Model C is the same as Model B except that we do not transform the response variable. Model D represents what would happen if we skip logistic regression and use stepwise and multiple regression on all 90 data points (ignoring the problems of heteroscedasticity in the residuals, and ignoring the fact that we do not desire to predict negative cost growth) (Sipple, 2002:72).

Upon validation of the four models using the 20 percent test set, Sipple found that both Models A and B accurately predicted the existence of cost growth and the amount of cost growth with about a 70 percent accuracy rate. Model A utilizes seven out of 78 possible predictor variables, while Model B uses three. Model C does fairly well at predicting the validation data. Using an 80 percent confidence bound, Model C contains

73 percent of the data, however, due to the violation of the OLS assumptions, it is unknown whether or not this confidence bound is a true 80 percent.

Comparing Model D to Model B, Sipple found that “Model B produces higher $R^2$ values than Model D…Model B yields more predictive ability for the number of variables, and none of Model D’s versions can compare to the versions of Model B above two predictor variables” (Sipple, 2002:104).

It would appear that the two-step methodology employing Models A and B is superior than using a one model approach. The C and D Models seem to perform well, but their lack of conformity with underlying regression assumptions greatly reduces the ability of the user to accurately interpret their results (Sipple, 2002:113).

*Bielecki Thesis*

Employing the same methodology and underlying philosophy, Bielecki carries Sipple’s work forward to research cost growth in the four remaining SAR cost variance categories: *schedule*, *estimating*, *support*, and *other*. Bielecki employs logistic and multiple regression to build models aimed at identifying cost growth characteristics in each category as they relate to RDT&E appropriations in the EMD phase of the system life cycle.

Bielecki also finds that the distribution for each cost growth category are mixed — indicating the need for the two-step approach. In addition, he observes that the *other* and s*upport* categories do not contain enough data to support a inferential statistical analysis. Therefore, Bielecki limits his study to the remaining two categories: *schedule* and *estimating*.

As Sipple does before him, Bielecki builds a family of logistic and multiple regression models for each category and picks the best one for each. The best logistic regression model submitted for each category validates at 85.71 percent and 78.26 percent for the *schedule* and *estimating* categories respectively. Using an 80 percent prediction bound, the best multiple regression model submitted for each category validates at 80.00 percent and 100 percent for the *schedule* and *estimating* categories, respectively.

*Moore Thesis*

Unlike Sipple and Bielecki, Moore's research does not focus on a specific SAR cost variance category. Instead, Moore focuses on the procurement appropriations and any cost growth associated with them in the EMD phase of the system life cycle as he states this is the "next logical level" (Moore, 2003:16).

When Moore performs a preliminary analysis of his data, he found that the distribution for procurement cost growth during the EMD phase exhibits identical characteristics to those found by Sipple (Moore, 2003:21). Meaning that there is a mixed distribution and the two-step methodology will be used.

The logistic regression model Moore submits for validation accurately predicts four out of the four data points available for validation. Of the 25 data points randomly selected for validation, only four of them contained the variable *FUE-based Maturity*. Upon further validation, the model was found to accurately predict 37 out of the 39 data points used to build the model. Therefore, the variable, *FUE-based Maturity*, turns out to be the '600-pound gorilla' that predicts the presence of cost growth accurately about 95% of the time. The multiple regression model Moore submits for validation also contains a ‘600-pound gorilla’, *FUE-based Length of EMD*, and accurately predicts 100 percent of the predicted data points, using an 80 percent prediction interval (Moore, 2003:47).

*OSD CAIG Study*

In addition to the above three theses, we find one additional study by the Office of the Secretary of Defense Cost Analysis Improvement Group (OSD CAIG) to be relevant to our study and therefore include it in our literature review.

The study, *Cost Growth of Major Defense Programs,* is the culmination of 10 years of research between the OSD CAIG, NAVSHIPSO and AT&T. This study uses the SARs of 286 programs as its source of data. When bumped up against the study criteria: unclassified, milestone II captured, three years of data past milestone II, and data complete; these 286 programs are reduced to 142 and are entered into the database.

They define cost growth as the “difference between today’s estimate and a baseline estimate caused by:”

- Poor initial estimates
  - Ill defined programs
- Different program than originally conceived
  - Different procurement quantities
  - Requirement changes
- Inefficiencies
  - Too many people
  - Too much money
  - Lack of focus

- Other

(*Cost Growth of Major Defense Programs*, 2003:6)

The main objective of the study is to identify how much of cost growth is attributable to: 1) decisions: discretionary changes to the system relative to the description at milestone II , and 2) mistakes: changes not attributable to discretionary changes post milestone II. Also, a main objective is to establish a historical record for comparison between systems (*Cost Growth of Major Defense Programs*, 2003:10). The results of the study follow:

- Cost growth appears to have a correlation with commodity
- Cost estimating assumptions account for majority of mistakes cost growth
- Under estimating engineering effort is major source of RDT&E growth
- Nearly half of perceived cost growth is content change (i.e. decisions)
- Procurement cost growth is primarily due to optimistic learning curves
- Majority of systems do not have significant growth
- Higher cost systems appear to have less growth

(*Cost Growth of Major Defense Programs*, 2003:66).

Note that this study, like Sipple, Bielecki, and Moore's, evaluate cost growth as of the EMD phase of the system life cycle. Where this study differs is that the OSD and company do not focus on a single SAR cost variance category or a single appropriation. Instead, they seek to categorize cost growth into one of two categories: decisions or mistakes. From the results of their study we take away their finding that cost estimating assumptions account for the majority of cost growth in the mistakes category. This is consistent with most of our research as it reemphasizes the importance of generating accurate cost estimates up front and early in the acquisition process.

**Chapter Summary**

In this chapter, we discuss how the current acquisition process works as compared to how it used to work and explain why our study needs to analyze the old business practices. We also explore why accurate cost estimating is critical in today's acquisition environment, with heavy oversight, multiple stakeholders, scarce funding and numerous worldwide threats and ways to mitigate them. Upon examining the current acquisition environment we point out how risk is inherent in cost estimating due to countless unknowns, and that it is crucial to discover and address these unknowns up front and early. Finally we highlight the relevant findings of recent studies in this area in order that we may approach our own research with an arsenal of "lessons learned".

# III. Methodology

## Chapter Overview

This chapter presents the procedures used to perform our research. We also discuss our database to include the data collection process, as well as list and explain the response and predictor candidate variables. We provide and discuss the results of the exploratory data analysis on our response variables. Lastly, we state our methodology for performing both the logistic and multiple regression models.

## Database

For this study we employ a slightly modified version of the database originally built by Sipple during his research. These modifications affect some of the predictor variables and are discussed in detail later in this chapter. The database is a culmination of information from the SARs and the 1996 RAND report, *The Defense System Cost Performance Database: Cost Growth Analysis Using Selected Acquisition Reports.* For insight into the foundation of the database and a comprehensive look into the use of SARs as a historical source of data in analyzing cost growth, to include their limitations, see Sipple, 2002.

## Data Collection

This research utilizes the database originally composed by Sipple (2002). We begin our data collection with a thorough review of Sipple's database. Sipple builds the database with individual program SAR reports beginning in the year 1990 and ending in 2000. Bielecki and Moore add to the database all programs fitting the entry criteria with a SAR date in 2001. In order for a program to be entered into the database, it must be at least three years into the EMD phase (mature program). After combing through the most current SAR (2002) database, we find four programs that meet the entry criteria, and we add them to the database. To keep the data consistent we omit any programs that meet the criteria, but use the A, B, and C milestone labeling scheme as opposed to the I, II, and III labeling scheme.

Once all new programs are added we scrub each program listed in the database by validating each predictor variable against the information listed in each SAR and RAND report. This involves printing off and indexing each program SAR and visually inspecting each data point for each program. Also, the following information: *prototype, prototype phase, modification, weapon type, whether or not the program had a MS I,* and *service* is checked against the RAND report.

The most obvious change to the database is the addition of the indexing or numbering system assigned to all of the programs and predictor variables. We place a number in front of each program data point as well as each predictor variable. By assigning a sequential numbering system to each program SAR and predictor variable, we are able to quickly look up all data pertaining to a given program without 'thumbing' through 135 SARs. It also aids during our model building in that when we add and remove variables during the logistic model building series, we are easily alerted to any omitted predictor variables.

**Response Variables**

As mentioned in Chapter 1, the SAR reports cost variance in seven categories: *economic, quantity, estimating, engineering, schedule, support,* and *other*. Our research focuses on predicting and quantifying cost growth in the *estimating* and *support* categories of the procurement appropriation. Since we are dealing with a mixed distribution, a distribution with both continuous and discreet data, we have two response variables for each cost variance category.

The logistic regression response variables: *Estimating Cost Growth? Procurement* and *Support Cost Growth? Procurement* are expressed as a binary variable where a value of '1' indicates that we estimate a program will experience cost growth, while a '0' indicates that we estimate it will not.

The multiple regression response variables: *Cost Variance - Procurement % Estimating* and *Cost Variance - Procurement % Support* are expressed as percentages, rather than dollar amounts. The percentage-based variable is preferred since it eliminates the need to quantify between programs and it normalizes programs of different sizes for comparison purposes (Bielecki, 2003:35).

**Predictor Variables**

The predictor variables that Sipple (2002) gathered are not exhaustive, but endow us with a plethora of proven predictors of cost growth. Sipple groups the predictor variables into five categories: program size, physical type of program, management characteristics, schedule characteristics, and other characteristics. We keep these same categories; however, we modify some of the subcategories by removing, changing, or adding variables.

The first major change we make to the list of predictor variables is to remove any variable that has 37 data points or less. This is done because once we remove 20% of the data points for our validation subset, we are left with less that 30 data points to build our models.

The following variables are removed for the reason given:

- *Maturity from MSII in mos*
  - For some programs in which the latest SAR date is after MSIII this variable artificially adds months into the EMD phase

- *Actual Length of EMD using FUE-MSII in mos/FUE-based Maturity of EMD%*
  - FUE and IOC are interchangeable terms, therefore, we eliminate both variables containing FUE and use only variables containing IOC dates

- *MSIII Complete*?
  - We are concerned only with the EMD phase of the life cycle. This variable is removed because it will always be '0' during this phase

- *RAND Concurrency Measurement Interval & RAND Concurrency Measurement Interval %*
  - Both of these are removed because MS IIIA indicates that the program is in the procurement phase; as our model is focused on programs within EMD, this variable does not apply

- *Class at Least S*
  - This variable appears to indicate whether a program has a security classification of secret or higher. Since we are dealing with only secret or lower data, this variable does not apply

- *Terminated?*
  - Removed because our research applies to a living program; if the program is terminated then the need for a prediction is not applicable

- *Qty in PE*
  - Removed because it had only seven fields with data

The names of many of the variables are changed for semantic reasons; however, the following variable is re-formulated for the reason given:

- *Maturity of EMD %*
  - A new formula is developed to prevent programs from being more than 100% complete. With the old formula an EMD phase could be more that 100% complete. Now any EMD phase greater that 100% is simply 100%

The following variables are added:

- *6 ACAT* – discrete variable to indicate the ACAT level
- *7 ACAT 1*? – binary variable: 1 for yes and 0 for no
- 21 # of Svs = discrete variable to indicate the number of services involved in the program
- *28 Service = Marines Only* – binary variable: 1 for yes and 0 for no
- *60 LRIP Qty Planned* – continuous variable to indicate the quantity in the baseline estimate
- *61 LRIP Qty Current Estimate* – continuous variable to indicate the quantity as currently estimated in the latest SAR
- *77 LRIP Planned?* – binary variable: 1 for yes and 0 for no; indicates if the program had LRIP planned
- *78 % R&D of Total Program* – continuous variable calculated by dividing *52 Length of R&D in Funding Yrs* by *48 Funding YR Total Program Length*
- *79 % Proc of Total Program* – continuous variable calculated by dividing *51 Length of Prod in Funding Yrs* by *48 Funding YR Total Program Length*
- *80 Length of R&D Funding > 12 yrs?* – binary variable which indicates if *52 Length of R&D in Funding Yrs* exceeds 12 years: 1 for yes 0 for no
- *81 Length of Proc Funding > 11 yrs?* – binary variable which indicates if *51 Length of Prod in Funding Yrs* exceeds 11 years: 1 for yes 0 for no
- *82 R&D Funding Yr Maturity % > 75%?* – binary variable which indicates if *53 R&D Funding Yr Maturity %* exceeds 0.75: 1 for yes 0 for no
- *83 Proc Funding Yr Maturity % > 40%?* – binary variable which indicates if *54 Proc Funding Yr Maturity %* exceeds 0.4: 1 for yes 0 for no
- *84 Funding Yrs of R&D Complete < 9?* – binary variable which indicates if *49 Funding Yrs of R&D Completed* is less that 9 years: 1 for yes 0 for no
- *85 Funding Yrs of Proc Complete < 5?* – binary variable which indicates if *50 Funding Yrs of Prod Completed* is less that 5 years: 1 for yes 0 for no

Listed below are the categories and subcategories of the all the predictor variables used for this research:

*Program Size Variables*

- *1 Total Cost CY $M 2002* – continuous variable which indicates the total cost of the program in CY $M 2002
- *2 Total Quantity* – continuous variable which indicates the total quantity of the program at the time of the SAR date; if no quantity is specified, we assume a

quantity of one (or another appropriate number) unless the program was terminated

- *3 Unit Cost* – continuous variable that equals the quotient of the total cost and total quantity variables above
- *4 Qty Planned for R&D* – continuous variable which indicates the quantity in the baseline estimate
- *5 Qty Currently Estimated for R&D* – continuous variable that indicates the quantity that was estimated in the Planning Estimate
- *6 ACAT* –continuous variable to indicate the ACAT level
- *7 ACAT 1*? –binary variable: 1 for yes and 0 for no

*Physical Type of Program*

- Domain of Operation Variables

    - *8 Air* – binary variable: 1 for yes and 0 for no; includes programs that primarily operate in the air; includes air-launched tactical missiles and strategic ground-launched or ship-launched missiles
    - *9 Land* – binary variable: 1 for yes and 0 for no; includes tactical ground-launched missiles; does not include strategic ground-launched missiles
    - *10 Space* – binary variable: 1 for yes and 0 for no; includes satellite programs and launch vehicle programs
    - *11 Sea* – binary variable: 1 for yes and 0 for no; includes ships and ship-borne systems other than aircraft and strategic missiles

- Function Variables

    - *12 Electronic* – binary variable: 1 for yes and 0 for no; includes all computer programs, communication programs, electronic warfare programs that do not fit into the other categories
    - *13 Helo* – binary variable: 1 for yes and 0 for no; helicopters; includes V-22 Osprey
    - *14 Missile* – binary variable: 1 for yes and 0 for no; includes all missiles
    - *15 Aircraft* – binary variable: 1 for yes and 0 for no; does not include helicopters
    - *16 Munition* – binary variable: 1 for yes and 0 for no
    - *17 Land Vehicle* – binary variable: 1 for yes and 0 for no
    - *18 Space (Rand)* –binary variable: 1 for yes and 0 for no
    - *19 Ship* – binary variable: 1 for yes and 0 for no; includes all watercraft
    - *20 Other* – binary variable: 1 for yes and 0 for no; any program that does not fit into one of the other function variables

*Management Characteristics*

- Military Service Management

- 21 # of Svs = continuous variable to indicate the number of services involved in the program
- *22 Svs > 1* – binary variable: 1 for yes and 0 for no; number of services involved at the date of the SAR
- *23 Svs > 2* – binary variable: 1 for yes and 0 for no; number of services involved at the date of the SAR
- *24 Svs > 3* – binary variable: 1 for yes and 0 for no; number of services involved at the date of the SAR
- *25 Service = Navy Only* – binary variable: 1 for yes and 0 for no
- *26 Service = Joint* – binary variable: 1 for yes and 0 for no
- *27 Service = Army Only* – binary variable: 1 for yes and 0 for no
- *28 Service = Marines Only* – binary variable: 1 for yes and 0 for no
- *29 Service = AF Only* – binary variable: 1 for yes and 0 for no
- *30 Lead Svc = Army* – binary variable: 1 for yes and 0 for no
- *31 Lead Svc = Navy* – binary variable: 1 for yes and 0 for no
- *32 Lead Svc = DoD* – binary variable: 1 for yes and 0 for no
- *33 Lead Svc = AF* – binary variable: 1 for yes and 0 for no
- *34 AF Involvement* – binary variable: 1 for yes and 0 for no
- *35 N Involvement* – binary variable: 1 for yes and 0 for no
- *36 MC Involvement* – binary variable: 1 for yes and 0 for no
- *37 AR Involvement* – binary variable: 1 for yes and 0 for no

- Contractor Characteristics

  - *38 Lockheed-Martin* – binary variable: 1 for yes and 0 for no
  - *39 Northrup Grumman* – binary variable: 1 for yes and 0 for no
  - *40 Boeing* – binary variable: 1 for yes and 0 for no
  - *41 Raytheon* – binary variable: 1 for yes and 0 for no
  - *42 Litton* – binary variable: 1 for yes and 0 for no
  - *43 General Dynamics* – binary variable: 1 for yes and 0 for no
  - *44 No Major Defense Contractor* – binary variable: 1 for yes and 0 for no; a program that does not use one of the contractors mentioned immediately above = 1
  - *45 More than 1 Major Defense Contractor* – binary variable: 1 for yes and 0 for no; a program that includes more than one of the contractors listed above = 1
  - *46 Fixed-Price EMD Contract* – binary variable: 1 for yes and 0 for no

*Schedule Characteristics*

- RDT&E and Procurement Maturity Measures

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

- *48 Funding YR Total Program Length* – continuous variable which indicates the total number of years for which the program has either RDT&E funding or procurement funding budgeted
- *49 Funding Yrs of R&D Completed* – continuous variable which indicates the number of years completed for which the program had RDT&E funding budgeted
- *50 Funding Yrs of Prod Completed* – continuous variable which indicates the number of years completed for which the program had procurement funding budgeted
- *51 Length of Prod in Funding Yrs* – continuous variable which indicates the number of years for which the program has procurement funding budgeted
- *52 Length of R&D in Funding Yrs* – continuous variable which indicates the number of years for which the program has RDT&E funding budgeted
- *53 R&D Funding Yr Maturity %* – continuous variable which equals *49 Funding Yrs of R&D Completed* divided by *52 Length of R&D in Funding Yrs*
- *54 Proc Funding Yr Maturity %* – continuous variable which equals *50 Funding Yrs of Prod Completed* divided by *51 Length of Prod in Funding Yrs*
- *55 Total Funding Yr Maturity %* – continuous variable which equals *Maturity (47 Funding Yrs complete)* divided by 48 *Funding YR Total Program Length*

o EMD Maturity Measures

- *56 Actual Length of EMD* – continuous variable calculated by subtracting the earliest MS II date from the latest MS III date indicated
- *57 Maturity of EMD %* – continuous variable calculated by dividing *Maturity from MS II (current calculation in months)* by *56 Actual Length of EMD*
- *58 Time From MSII to IOC in months* – continuous variable calculated by subtracting the earliest MS II date from the IOC date
- *59 Maturity of EMD at IOC %* – continuous variable calculated by dividing *Maturity from MS II (current calculation in months)* by *57 Time From MSII to IOC in months*
- *60 LRIP Qty Planned* – continuous variable to indicate the quantity in the baseline estimate
- *61 LRIP Qty Current Estimate* – continuous variable to indicate the quantity as currently estimated in the latest SAR

o Concurrency Indicators

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

- *63 Proc Funding before MS III* – binary variable:  1 for yes and 0 for no

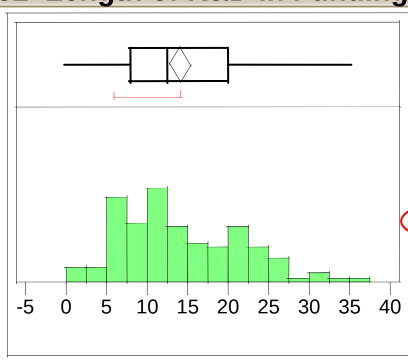
*Other Characteristics*

- *64 # Product Variants in this SAR* – continuous variable which indicates the number of versions included in the EMD effort that the current SAR addresses
- *65 Class – S* – binary variable:  1 for yes and 0 for no; security classification Secret
- *66 Class – C* – binary variable:  1 for yes and 0 for no; security classification Confidential
- *67 Class – U* – binary variable:  1 for yes and 0 for no;  security classification Unclassified
- *68 Risk Mitigation* – binary variable:  1 for yes and 0 for no; indicates whether there was a version previous to SAR or significant pre-EMD activities
- *69 Versions Previous to SAR* – binary variable:  1 for yes and 0 for no; indicates whether there was a significant, relevant effort prior to the DE; a pre-EMD prototype or a previous version of the system would apply
- *70 Modification* – binary variable:  1 for yes and 0 for no; indicates whether the program is a modification of a previous program
- *71 Prototype* – binary variable:  1 for yes and 0 for no; indicates whether the program had a prototyping effort
- *72 Dem/Val Prototype* – binary variable:  1 for yes and 0 for no; indicates whether the prototyping effort occurred in the PDRR phase
- *73 EMD Prototype* – binary variable:  1 for yes and 0 for no; indicates whether the prototyping effort occurred in the EMD phase
- *74 PE?* – binary variable:  1 for yes and 0 for no; indicates whether the program had a Planning Estimate
- *75 Significant pre-EMD activity immediately prior to current version* – binary variable:  1 for yes and 0 for no; indicates whether the program had activities in the schedule at least six months prior to MSII decision
- *76 Program have a MS I?* – binary variable:  1 for yes and 0 for no
- *77 LRIP Planned?* – binary variable:  1 for yes and 0 for no; indicates if the program had LRIP planned
- *78 % R&D of Total Program* – continuous variable calculated by dividing *52 Length of R&D in Funding Yrs* by *48 Funding YR Total Program Length*
- *79 % Proc of Total Program* – continuous variable calculated by dividing *51 Length of Prod in Funding Yrs* by *48 Funding YR Total Program Length*
- *80 Length of R&D Funding > 12 yrs?* – binary variable which indicates if *52 Length of R&D in Funding Yrs* exceeds 12 years: 1 for yes 0 for no
- *81 Length of Proc Funding > 11 yrs?* – binary variable which indicates if *51 Length of Prod in Funding Yrs* exceeds 11 years: 1 for yes 0 for no
- *82 R&D Funding Yr Maturity % > 75%?* – binary variable which indicates if *53 R&D Funding Yr Maturity %* exceeds .75: 1 for yes 0 for no
- *83 Proc Funding Yr Maturity % > 40%?* – binary variable which indicates if *54 Proc Funding Yr Maturity %*  exceeds .4: 1 for yes 0 for no

- *84 Funding Yrs of R&D Complete < 9?* – binary variable which indicates if *49 Funding Yrs of R&D Completed* is less that 9 years: 1 for yes 0 for no
- *85 Funding Yrs of Proc Complete < 5?* – binary variable which indicates if *50 Funding Yrs of Prod Completed* is less that 5 years: 1 for yes 0 for no

Of the last eight variables that are added to the database, the final six are computed by 'discretizing' the continuous variables for which they represent. By discretizing we mean to take a continuous variable and turn it in to a binary variable. For example, this is done by first running a distribution of the variable *52 Length of R&D in Funding Yrs* in JMP® and analyzing the quantiles for the median value (see Figure 3.1).

**52 Length of R&D in Funding Yrs**




**Quantiles**

| | | |
|---|---|---|
| 100.0% | maximum | 35.000 |
| 99.5% | | 35.000 |
| 97.5% | | 31.550 |
| 90.0% | | 24.100 |
| 75.0% | quartile | 20.000 |
| 50.0% | median | 12.500 |
| 25.0% | quartile | 8.000 |
| 10.0% | | 6.000 |
| 2.5% | | 2.000 |
| 0.5% | | 0.000 |
| 0.0% | minimum | 0.000 |

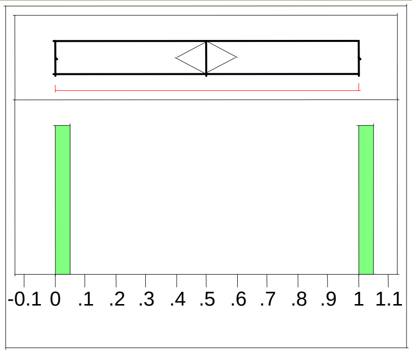
**Moments**

| | |
|---|---|
| Mean | 14.111111 |
| Std Dev | 7.624767 |
| Std Err Mean | 0.7336935 |
| upper 95% Mean | 15.565573 |
| lower 95% Mean | 12.656649 |
| N | 108 |

**Figure 3.1 – Histogram of variable *52 Length of R&D in Funding Yrs***

The aim is to establish a logical cut-off point at which the binary responses of the new variable, *80 Length of R&D Funding > 12 yrs?*, in this example, are approximately equal (see Figure 3.2). The median is the best starting point to find the logical cut-off point. From there, the cut-off point can be 'tweaked' in either direction until an approximately equal split is obtained. In this example, the median value of 12 appears to do the trick.

**80 Length of R&D Funding > 12 yrs?**




**Quantiles**

| | | |
|---|---|---|
| 100.0% | maximum | 1.0000 |
| 99.5% | | 1.0000 |
| 97.5% | | 1.0000 |
| 90.0% | | 1.0000 |
| 75.0% | quartile | 1.0000 |
| 50.0% | median | 0.5000 |
| 25.0% | quartile | 0.0000 |
| 10.0% | | 0.0000 |
| 2.5% | | 0.0000 |
| 0.5% | | 0.0000 |
| 0.0% | minimum | 0.0000 |

**Moments**

| | |
|---|---|
| Mean | 0.5 |
| Std Dev | 0.502331 |
| Std Err Mean | 0.0483368 |
| upper 95% Mean | 0.5958221 |
| lower 95% Mean | 0.4041779 |
| N | 108 |

**Figure 3.2 – Histogram of variable *80 Length of R&D Funding > 12 yrs?***

## Model Building

Now that the database is complete we begin to build our regression models. The first step to building successful models is to set aside part of the database for validation. We choose 20% of the database for validation. To ensure bias is not present in our 80% model building subset or the 20% model validation subset, we add a random number column to our database, sort on this column, then remove the last 20% of the data points. For this database, this gives us an 80% model database with 108 data points and a 20% validation database with 27 data points.

### *Preliminary Data Analysis*

Once the database is partitioned the next step is to ensure that the response variables used in our multiple regression models have an underlying distribution that is reasonably continuous. To confirm this distribution, we run a histogram of our *Cost Variance - Procurement % Estimating* and *Cost Variance - Procurement % Support* response variables in JMP® using the data from the 80% subset. Looking at Figure 3.3 we find mixed distributions. These distributions are identical to those identified by

Sipple, Bielecki, and Moore during their research.  They exhibit the same characteristics — continuous with a discrete mass, or ‘spike’, around zero.

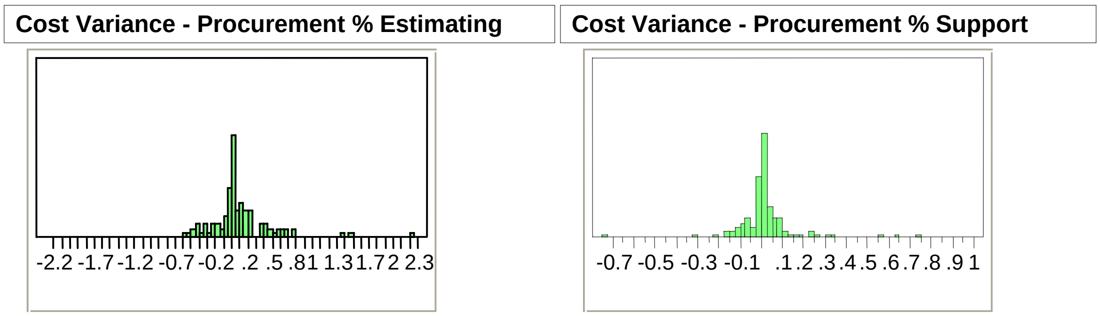



**Figure 3.3 – OLS Response Variable Histograms**

“This situation necessitates that we split the data into two separate sets to accurately model the individual effects of both the discrete and continuous data components.  As demonstrated by Sipple (2002), a two-step cost growth model produces statistically equivalent results as a single-step regression model however; the two-step model is statistically more reliable due to the validity of its underlying assumptions.  For these reasons, we adopt this two-step methodology (Bielecki, 2003:47).”

The first part of the two step methodology, logistic regression, utilizes the entire data set by assigning a ‘1’ to any positive percentage and a ‘0’ to any zero or negative percentage.  The second step, OLS regression, uses only the positive percentages of the data set.   Only positive percentages are used because they represent the positive cost growth that cost estimators and program managers are concerned with.

Now that we have established that our overall data mimics that of Sipple, Bielecki, and Moore, we confirm that the OLS data set (only the positive percentages) are reasonably continuous.  Note in Figure 3.4 that the variables are reasonably continuous and maintain a log-normal distribution as indicated by the p-values exceeding our alpha level of 0.05 with 0.0821 and 0.15 for the *Estimating* and *Support* variables respectively. These distributions are indicative of the distributions first identified by Sipple, and later confirmed by Bielecki and Moore. Note that represented in these histograms are 61 and 53 data points for the *Estimating* and *Support* variables respectively.

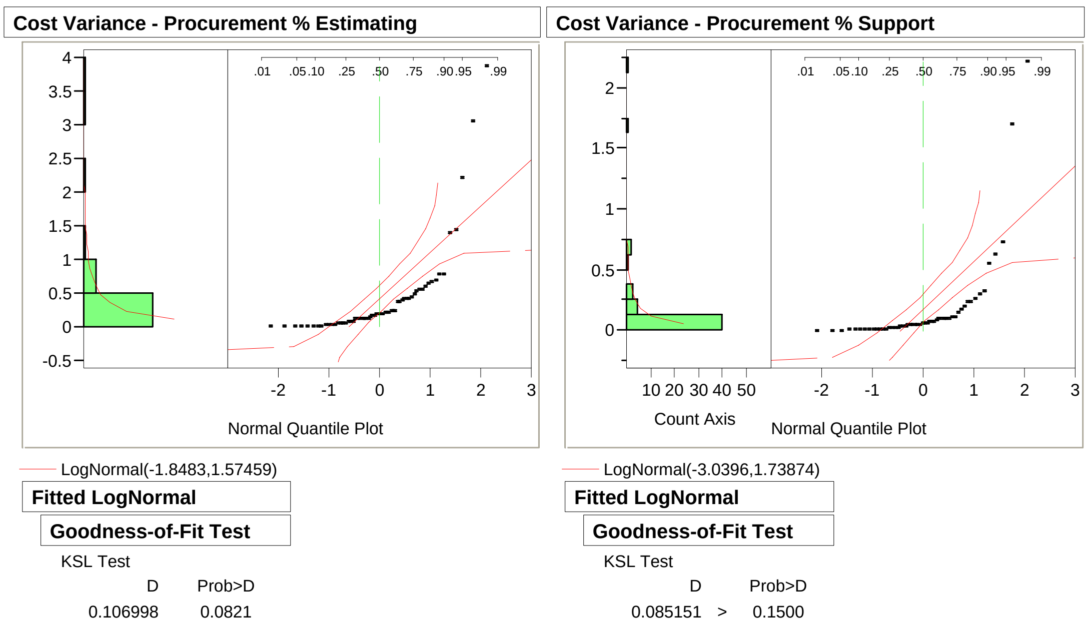



**Figure 3.4 – Log-normal histograms of OLS response variables**

The histograms in Figure 3.5 show the same log-normal distribution as the prior research and suffice the OLS basic assumption of having to be reasonably continuous. Due to the fact that all three researchers before us corrected this log-normal distribution in order to satisfy constant variance in the residuals once their models are built, we will begin with the assumption that we must correct for constant variance by transforming our OLS response variables by applying a natural log.

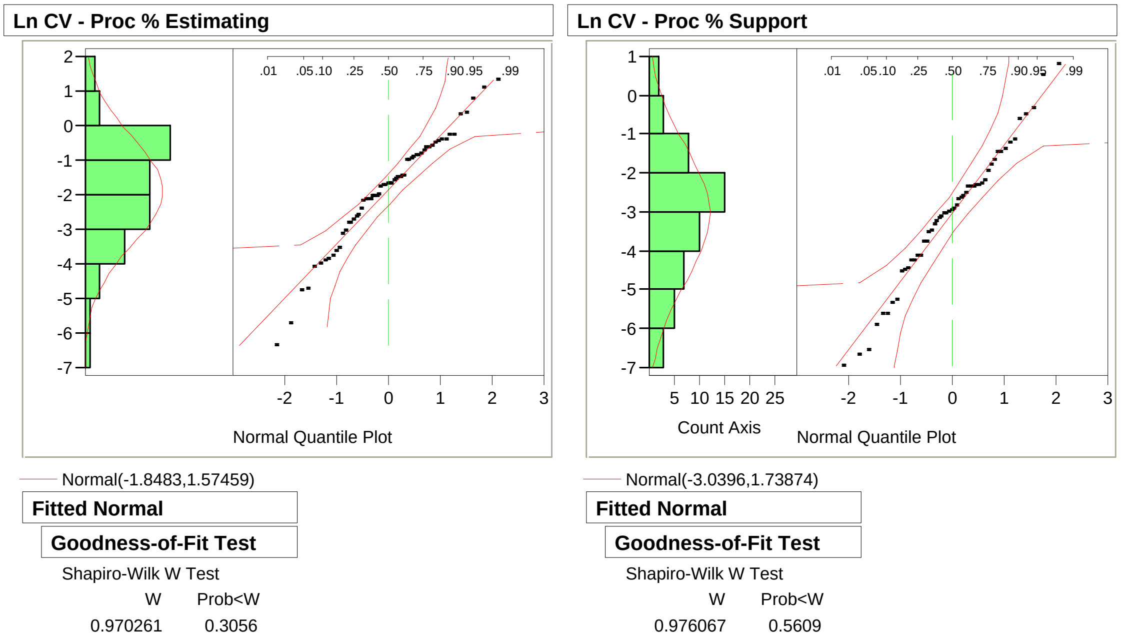



**Figure 3.5 – Normal histograms of OLS response variables**

The histograms in Figure 3.5 reveal an approximately normal distribution as evidenced by the p-value exceeding our alpha level of 0.05 with 0.3056 and 0.5609 for the *estimating* and s*upport* variables respectively.

*Logistic Regression*

We use logistic regression to analyze whether some event will occur or not. In our case we want to know if a program will experience cost growth in the *estimating* and *support* cost variance categories of procurement appropriations during the EMD phase of the system life cycle. To this end, the binary responses are coded ‘1’ for any positive program cost growth percentages and ‘0’ for any zero or negative percentages.

We use JMP® statistical software to build our logistic regression models. Since JMP® version 4 does not contain an automated method such as stepwise to build logistic models, we follow the methodology established by Sipple (2002):

> "..we manually compute thousands of individual regressions, recording our results on spreadsheets. We start with one-predictor models of all possible variables. Then we regress using all combinations of two-predictor models and record the results. We continue this process, eventually whittling down the best combinations for use at the next level in order to cut down on the amount of regressions necessary. We stop when we reach a model for which the gain of adding another variable does not warrant the additional complexity of the model that another variable adds. We intend to find several candidate models for each number of predictors and then narrow down to the best one for each number of predictors and validate the model using about 20 percent of the data that we set aside for validation (Sipple, 2002:70)."

Our initial criterion for allowing a variable to enter a model is that each variable must have an individual p-value less than 0.04. This is more of a guideline than a cold hard fact. As the model progresses from one to two to three variables, etc., natural cut-offs within the data are used to advance the 'best' models forward to the next level. This is accomplished by analyzing the average of the sum of the individual p-values, the R squared (U), the number of observations, and the area under the receiver operating curve (ROC) simultaneously for each model. For an in-depth description of each of these performance measures see Sipple (2002). This natural cut-off approach is used to prevent us from blindly pick the 'top 10' or 'top 8' models where the last 3 or 4 of these 'top' models may have performance measure far from the top 5 or 6 models, see Table 3.1 for an illustration. As seen in Table 3.1, all models are sorted by each performance measure then ranked using a consecutive number from 1 to n, where n is the number of total models built for that level, (i.e. all two variables, all three variables, etc.). Table 3.1 is an excerpt of all two variable models. Note the natural break in the results. In this case the top six models are the 'best' models and are carried forward to begin building the three variable models. The next four model scores are an average of 4.19 points above the sixth model.

| Estimating % Two-Variable Models | | | | | | Performance Measures | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| Variables | | R Sq (U) | Obs | P-Value | AUC | R Sq (U) | Obs | P-Value | AUC | Score |
| 7 | 77 | 0.1439 | 108 | 0.0067 | 0.7173 | 1 | 1 | 1 | 9 | 3 |
| 38 | 51 | 0.1283 | 101 | 0.017 | 0.74203 | 4 | 4 | 8 | 2 | 4.5 |
| 7 | 9 | 0.13 | 105 | 0.0135 | 0.71296 | 2 | 2 | 4 | 12 | 5 |
| 7 | 78 | 0.1272 | 108 | 0.0235 | 0.75131 | 5 | 1 | 14 | 1 | 5.25 |
| 9 | 51 | 0.119 | 105 | 0.0139 | 0.72889 | 9 | 2 | 6 | 4 | 5.25 |
| 38 | 81 | 0.1194 | 101 | 0.0107 | 0.7185 | 8 | 4 | 2 | 7 | 5.25 |
| 51 | 77 | 0.1101 | 108 | 0.0179 | 0.71747 | 14 | 1 | 9 | 8 | 8 |
| 7 | 38 | 0.1285 | 101 | 0.0236 | 0.71232 | 3 | 4 | 15 | 13 | 8.75 |
| 38 | 77 | 0.1194 | 101 | 0.0138 | 0.69657 | 7 | 4 | 5 | 25 | 10.25 |
| 9 | 77 | 0.1057 | 105 | 0.0162 | 0.70167 | 16 | 2 | 7 | 18 | 10.75 |
| 48 | 77 | 0.1029 | 108 | 0.0202 | 0.70771 | 19 | 1 | 12 | 14 | 11.5 |
| 38 | 48 | 0.1039 | 101 | 0.0282 | 0.71651 | 18 | 4 | 18 | 10 | 12.5 |
| 77 | 13 | 0.1125 | 108 | 0.0332 | 0.70649 | 13 | 1 | 26 | 15 | 13.75 |
| 9 | 27 | 0.1043 | 103 | 0.0119 | 0.67644 | 17 | 3 | 3 | 38 | 15.25 |
| 7 | 46 | 0.1163 | 98 | 0.0377 | 0.70281 | 10 | 5 | 32 | 16 | 15.75 |
| 51 | 64 | 0.1005 | 108 | 0.0286 | 0.69759 | 20 | 1 | 19 | 23 | 15.75 |
| 77 | 47 | 0.0947 | 108 | 0.0287 | 0.7016 | 27 | 1 | 20 | 19 | 16.75 |
| 46 | 51 | 0.1146 | 98 | 0.0687 | 0.73639 | 11 | 5 | 49 | 3 | 17 |
| 48 | 44 | 0.0942 | 101 | 0.0387 | 0.72448 | 28 | 4 | 33 | 5 | 17.5 |
| 9 | 48 | 0.0948 | 105 | 0.0308 | 0.70093 | 26 | 2 | 22 | 20 | 17.5 |
| 7 | 24 | 0.123 | 108 | 0.039 | 0.69114 | 6 | 1 | 35 | 29 | 17.75 |
| 51 | 7 | 0.1137 | 108 | 0.0679 | 0.71346 | 12 | 1 | 48 | 11 | 18 |
| 9 | 15 | 0.0939 | 105 | 0.0197 | 0.67944 | 29 | 2 | 11 | 36 | 19.5 |
| 38 | 13 | 0.1076 | 101 | 0.0315 | 0.67544 | 15 | 4 | 24 | 39 | 20.5 |
| 77 | 50 | 0.0924 | 108 | 0.0342 | 0.69899 | 32 | 1 | 28 | 22 | 20.75 |
| 46 | 48 | 0.0949 | 98 | 0.0726 | 0.71854 | 25 | 5 | 50 | 6 | 21.5 |
| 81 | 44 | 0.0969 | 101 | 0.0338 | 0.6882 | 23 | 4 | 27 | 32 | 21.5 |
| 48 | 64 | 0.0882 | 108 | 0.0247 | 0.68748 | 39 | 1 | 16 | 33 | 22.25 |

**Natural Break**

**Lower is better!**

**Table 3.1 – Example of model ranking (Two-Variable)**

The models that possess the best average sum of these performance measures are advanced to the next round of model building. This 'best' model is our 'kernel" model or our full model — meaning it possesses the core variables with the best predictive value. This full model represents our final candidate model. The full model is then subjected to analysis. We fine tune the kernel variables contained in the full model by mathematically combining the variables to include higher order terms, removing variables, seeing if there are any interactions between variables, and finally, retesting any excluded variables. An example of fine tuning is to remove each predictor variable one at a time and rerun the model and note the effects. Our end goal is to build one model for each cost variance category that is both parsimonious and robust. This parsimonious model becomes our final model. We then submit this final model for validation using the 20% validation subset database we created from the master database.

*Multiple Regression*

The second step of our research uses multiple regression to predict how much cost growth a program has once our logistic model predicts that growth will occur. Again, we use JMP® to build our multiple regression models.

Using the transformed response variables discussed in the preliminary data analysis section, we regress the candidate predictor variables using the same procedure outlined for building our logistic models. Even though JMP® has a stepwise function to help build statistically significant models, we find this function unable to produce significant results with such a large amount of predictor variables. Therefore, we pursue the same 'Darwinist' approach in selecting our candidate variable models as we did for our logistic models. This methodology selects only the strongest, most statistically significant, models to be carried forward for each successive generation of model building, and culminates with only those combinations of variables (models) surviving which have the most value in predicting cost growth. (Bielecki, 2003:52).

We narrow our results to the best model for each number of predictors by adding or removing variables to the model until the number of variables equals approximately one tenth of the number of data points used in the model; this ensures we do not over-fit the model to the data (Bielecki, 2003:71).

As in the logistic regression method, we fine tune the variables within the kernel model and note the effects on the measurement parameters. With the same end goal in mind, we submit this final model for validation using the 20% validation subset database we created from the master database.

**Chapter Summary**

This chapter details the research methodology used during this study. We examine our database, describe the data collection process, and chronicle the candidate response and predictor variables. We discuss the preliminary data analysis on our OLS response variables and the need for the two step methodology using logistic and multiple regression. Finally, we examine the process used to build both logistic and multiple regression models. We introduce the results of our model building process in the next chapter.

# IV. Results and Discussion

## Chapter Overview

This chapter lays out the results of our logistic and multiple regression analysis. We begin with the logistic regression models followed by the multiple regression models for each cost variance category with the *estimating* response first and *support* response second. We walk through the methodology laid out in Chapter 3 and evaluate the statistical significance and robustness of each model. We discuss the final models submitted for validation, and finally validate each model to ensure each model is universal, accurate, and practical.

## Preliminary Findings

Upon initial building of our regression models we find some predictor variables exist that contribute no value to our models — see Appendix A. The highlighted variables represent all predictor variables that, when regressed on their respective response variables, have either an individual p-value greater than 0.3 or sum to greater that 0.3 in all two variable models. More importantly, they are present in more than 50% of the two variable models. Therefore, all predictor variables that are present in more that 50% of all two variable models are removed from further model building. Once we build our final model, each removed variable is put back into the model to ensure it adds no value to the model.

In addition to the variables in Appendix A that are removed, we discover that redundancy exists between some of the predictor variables. After ranking our 'best' two variable models, we find that variables *6 ACAT?* and 7 *ACAT 1?*are nearly identical.

Upon investigation we find variable 6 indicates what ACAT level the program is, 1, 2, or 3, and 7 indicates whether or not the program is ACAT 1. We see in the ranking that variable 7 consistently has a lower sum of p-values and $R^2$ (U), (except for one instance with variable 7 where the sum of p-values for 6 are slightly better than that of 7 and 51), and the area under the curve is nearly identical. Therefore, to reduce the number of models built, save time, and remove redundancy, we remove variable 6 from our already built models and preclude variable 6 from further model testing. This discovery leads us to run a pairwise correlation (using JMP®) among all predictor variables to see if redundancy exists among other variables. Table 4.1 depicts all variables with a correlation of greater than 0.9.

| Variable | by Variable | Correlation |
|---|---|---|
| 7 ACAT 1? | 6 ACAT | -0.934 |
| 22 Svs>1 | 21 # of Svs | 0.9083 |
| 49 Funding Yrs of R&D Completed | 47 Maturity (Funding Yrs complete) | 0.9312 |
| 23 Svs >2 | 21 # of Svs | 0.9386 |
| 5 Qty currently estimated for R&D | 4 Qty planned for R&D | 0.9735 |
| 72 Dem/Val Prototype? | 71 Prototype? | 1 |

**Table 4.1 – Redundant Predictor Variables**

Table 4.1 indicates that only two variables, 71 and 72, are identical—shown by the correlation of 1; however, based on the behavior of variables 6 and 7, which have a correlation of -0.934, we remove and keep the following predictor variables from further model building:

| Remove | Keep |
|---|---|
| 5  Qty currently estimated for R&D | 4  Qty planned for R&D |
| 6  ACAT | 7  ACAT 1? |
| 22  Svs>1 | 21  # of Svs |
| 23  Svs >2 | 47  Maturity (Funding Yrs complete) |
| 49  Funding Yrs of R&D Completed | 71  Prototype? |
| 72  Dem/Val Prototype? | |

**Table 4.2 – Predictor Variables Removed and Kept**

Unlike the variables listed in Appendix A, which are specific to each response variable, the variables found to be redundant in Table 4.2 are removed from building either model, and are not re-entered into our final models.

In addition, we find the predictor variables represented in Table 4.3 to be common in all models built for each response variable. We recommend that further studies in this area omit these predictor variables. These variables provide no statistical significance in any of the models built during this analysis.

| Common Bad Variables |
|---|
| 10 Space |
| 28  Service = Marines only |
| 31  Lead Svc = Navy |
| 43  General Dynamics |
| 45  More than 1 Major Defense Contracto |
| 55  Total Funding Yr Maturity % |
| 63  Proc Funding before MS III? |
| 71  Prototype? |
| 82  R&D Funding Yr Maturity % > 75%? |

**Table 4.3 – Bad Predictor Variables Common to All Response Variables**

### Logistic Regression Results — *Estimating* Response

We use the methodology described in Chapter 3 to build both of our logistic models. In all, we build over 3,000 logistic regression models for the *estimating* and *support* response variables not including the models built when reducing the full model for parsimonious purposes.  We find that as we proceed to build the best model by adding each predictor variable to the 'best' one-variable, two-variable, three-variable model, etc., there are some predictor variables that tend to show up in the best models at each level until, finally, there are no predictor variables left that dramatically improve the performance of the best model.  In essence we see the 'best' predictor variables 'bubble' to the top of each round of model building.

We believe that the model weighting method we use based on the performance measures: $R^2$ (U), Number of Observations, Sum of All individual P-Values, and Area Under the Receiver Operating Curve (AUC), afford us with the best opportunity to come up with this best model.  To illustrate this 'bubbling' phenomenon see Table 4.4 below.

| Estimating % N-Variable Models | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| # of Variables in Model | | | | | | | | | | Model Value | | | | Weighted Performance Measures | | | | Total |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | R Sq (U) | Obs | P-Value | AUC | R Sq (U) | Obs | P-Value | AUC | Score |
| 51 | | | | | | | | | | 0.0625 | 108 | 0.004 | 0.66777 | 3 | 1 | 3 | 2 | 2.25 |
| 7 | | | | | | | | | | 0.0884 | 108 | 0.0009 | 0.63987 | 1 | 1 | 1 | 6 | 2.25 |
| 77 | | | | | | | | | | 0.0589 | 108 | 0.0041 | 0.64039 | 4 | 1 | 4 | 5 | 3.5 |
| 48 | | | | | | | | | | 0.0472 | 108 | 0.0117 | 0.64597 | 8 | 1 | 7 | 3 | 4.75 |
| 9 | | | | | | | | | | 0.0505 | 105 | 0.0083 | 0.62222 | 7 | 3 | 5 | 8 | 5.75 |
| 81 | | | | | | | | | | 0.0418 | 108 | 0.0143 | 0.61999 | 10 | 1 | 8 | 9 | 7 |
| 46 | | | | | | | | | | 0.0555 | 98 | 0.0086 | 0.61607 | 5 | 8 | 6 | 10 | 7.25 |
| 78 | | | | | | | | | | 0.032 | 108 | 0.0359 | 0.62871 | 13 | 1 | 11 | 7 | 8 |
| 13 | | | | | | | | | | 0.0548 | 108 | 0.0296 | 0.57953 | 6 | 1 | 10 | 19 | 9 |
| 38 | | | | | | | | | | 0.0443 | 101 | 0.0208 | 0.60108 | 9 | 6 | 9 | 12 | 9 |
| 7 | 77 | | | | | | | | | 0.1439 | 108 | 0.0067 | 0.7173 | 1 | 1 | 1 | 9 | 3 |
| 38 | 51 | | | | | | | | | 0.1283 | 101 | 0.017 | 0.74203 | 4 | 4 | 8 | 2 | 4.5 |
| 7 | 9 | | | | | | | | | 0.13 | 105 | 0.0135 | 0.71296 | 2 | 2 | 4 | 12 | 5 |
| 7 | 78 | | | | | | | | | 0.1272 | 108 | 0.0235 | 0.75131 | 5 | 1 | 14 | 1 | 5.25 |
| 9 | 51 | | | | | | | | | 0.119 | 105 | 0.0139 | 0.72889 | 9 | 2 | 6 | 4 | 5.25 |
| 38 | 81 | | | | | | | | | 0.1194 | 101 | 0.0107 | 0.7185 | 8 | 4 | 2 | 7 | 5.25 |
| 9 | 51 | 15 | | | | | | | | 0.177 | 105 | 0.0099 | 0.77611 | 8 | 2 | 1 | 6 | 4.25 |
| 7 | 77 | 38 | | | | | | | | 0.1986 | 101 | 0.0177 | 0.77392 | 1 | 4 | 4 | 8 | 4.25 |
| 38 | 51 | 77 | | | | | | | | 0.1874 | 101 | 0.0184 | 0.77671 | 2 | 4 | 6 | 5 | 4.25 |
| 9 | 51 | 15 | 38 | | | | | | | 0.2493 | 101 | 0.0137 | 0.82117 | 1 | 3 | 1 | 1 | 1.5 |
| 7 | 77 | 38 | 67 | | | | | | | 0.2389 | 101 | 0.0375 | 0.8112 | 2 | 3 | 2 | 3 | 2.5 |
| 9 | 51 | 15 | 38 | 7 | | | | | | 0.2855 | 101 | 0.0561 | 0.83971 | 1 | 1 | 2 | 1 | 1.25 |
| 9 | 51 | 15 | 38 | 77 | | | | | | 0.2847 | 101 | 0.0512 | 0.83373 | 2 | 1 | 1 | 2 | 1.5 |
| 7 | 77 | 38 | 67 | 9 | | | | | | 0.2747 | 101 | 0.0722 | 0.82157 | 3 | 1 | 3 | 3 | 2.5 |
| 9 | 51 | 15 | 38 | 7 | 77 | | | | | 0.3249 | 101 | 0.0927 | 0.85805 | 1 | 1 | 2 | 1 | 1.25 |
| 7 | 77 | 38 | 67 | 9 | 15 | | | | | 0.3144 | 101 | 0.1079 | 0.8451 | 2 | 1 | 1 | 2 | 1.5 |
| 7 | 77 | 38 | 67 | 9 | 15 | 51 | | | | 0.3493 | 101 | 0.1697 | 0.87002 | 2 | 1 | 2 | 2 | 1.75 |
| 7 | 77 | 38 | 67 | 9 | 15 | 51 | 44 | | | 0.3707 | 101 | 0.2376 | 0.88796 | 1 | 1 | 1 | 1 | 1 |
| 7 | 77 | 38 | 67 | 9 | 15 | 51 | 44 | 2 | | 0.4081 | 100 | 0.2642 | 0.90228 | 1 | 1 | 1 | 1 | 1 |
| 7 | 77 | 38 | 67 | 9 | 15 | 51 | 44 | 2 | 39 | 0.4526 | 100 | 0.2304 | 0.91922 | 2 | 1 | 1 | 2 | 1.5 |
| 7 | 77 | 38 | 67 | 15 | 44 | 39 | 1/3 | ln 51 | | 0.6113 | 86 | 0.1478 | 0.95197 | | | N/A | | |

**Table 4.4 – Illustration of Predictor Variable 'Bubbling'**

Our best *estimating* model is depicted on the last line of Table 4.4. This table shows the best models from each level, or generation, of the process. The best one-variable models are at the top followed by the best two-variable models, followed by the best three-variable models, etc., until we arrive at the best reduced model. The bold lines indicate the natural cut-off point in the results of each successive generational round of model building. All other models are not shown due to simplicity of illustrative purposes. We can see in Table 4.4 that the highlighted variables that end up in our final reduced model surface in all rounds of model building beginning with the best one-variable models, and their appearance increases at each round until all but one float to the top and enter the final model. This 'bubbling' phenomenon is shown here as an example of what was common during all model building including ordinary least squares and will not be illustrated for each response variable.

The following table summarizes the best model at each round of our logistic model building process for the *estimating* response.

| Logistic (Estimating) Best Models | | | | |
|---|---|---|---|---|
| # Variables | R-Sq (U) | Obs | P-Value | AUC |
| 1 | 0.0884 | 108 | 0.0009 | 0.63987 |
| 2 | 0.1439 | 108 | 0.0067 | 0.7173 |
| 3 | 0.177 | 105 | 0.0099 | 0.77611 |
| 4 | 0.2493 | 101 | 0.0137 | 0.82117 |
| 5 | 0.2855 | 101 | 0.0561 | 0.83971 |
| 6 | 0.3249 | 101 | 0.0927 | 0.85805 |
| 7 | 0.3593 | 100 | 0.1162 | 0.87576 |
| 8 | 0.3707 | 101 | 0.2376 | 0.88796 |
| 9 | 0.4081 | 100 | 0.2642 | 0.90228 |
| Full (10) | 0.4526 | 100 | 0.2304 | 0.91922 |
| Next Best (11) | 0.4871 | 100 | 0.2549 | 0.92901 |
| Reduced (9) | 0.6113 | 86 | 0.1478 | 0.95197 |

**Table 4.5 – Best Logistic *Estimating* Models For Each Generation**

With the performance measures for each best model stated in Table 4.5, we decide to illustrate and discuss in the following graphs the relative changes of each performance measure as the number of variables increase. We begin our discussion with the relative change in $R^2$ (U), and continue with the data point to variable ratio, relative change in p-value, and relative change in AUC.

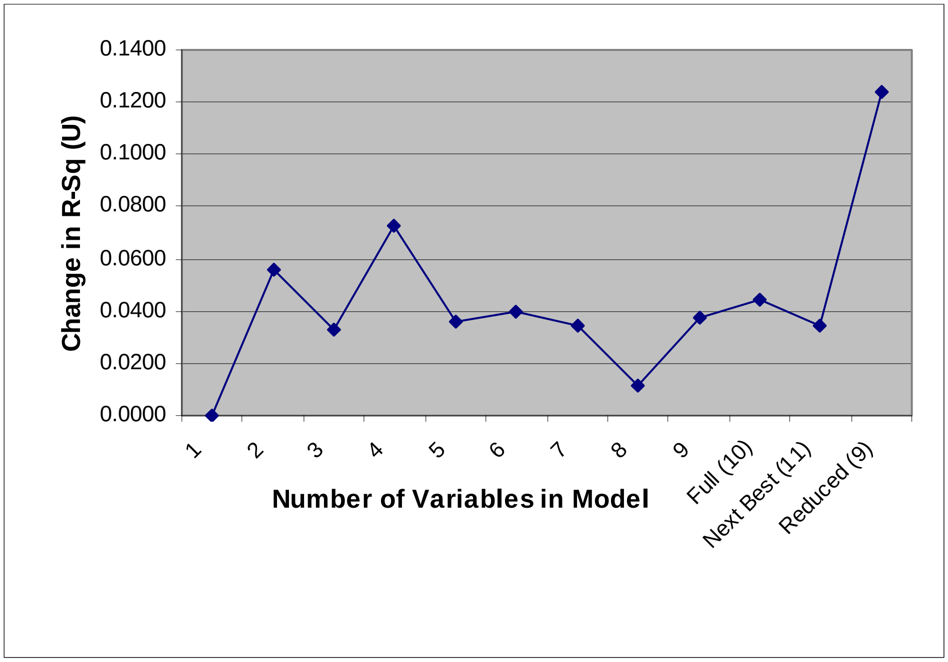



**Figure 4.1 – Relative Change in $R^2$ (U) - *Estimating* Models (Logistic)**

We see in Figure 4.1 that $R^2$ (U) changes sporadically as the number of variables per model increase. With the exception of our eight variable model, we the changes in $R^2$ (U) decrease from our one to five variable models, then, more or less, plateau from the five to eleven variable models. The next best model improves to 0.4871 from 0.4526, or a change of 0.0345; however, when we look at all of the performance measure together, we do not feel that the 0.0345 increase warrants the complexity inherent with the addition of too many variables, thus we keep the ten variable model as our full model. After fine tuning the variables in the full model we arrive at our final reduced model with our highest $R^2$ (U) value and largest relative change.

Before we fine tune our full model we look for the $R^2$ (U) to taper off or ‘plateau’ which indicates the amount of certainty explained by the model has more or less reached its peak. We say more or less because we could, theoretically, keep adding variables to the model and $R^2$ (U) would more than likely keep going up — increasing at a decreasing rate. Unlike in OLS regression where there is an adjusted $R^2$ wherein your model is ‘penalized’ for including too many variables, logistic regression has no such performance measure, which is why the next performance measure we look at is the ratio of data points to variables per model.

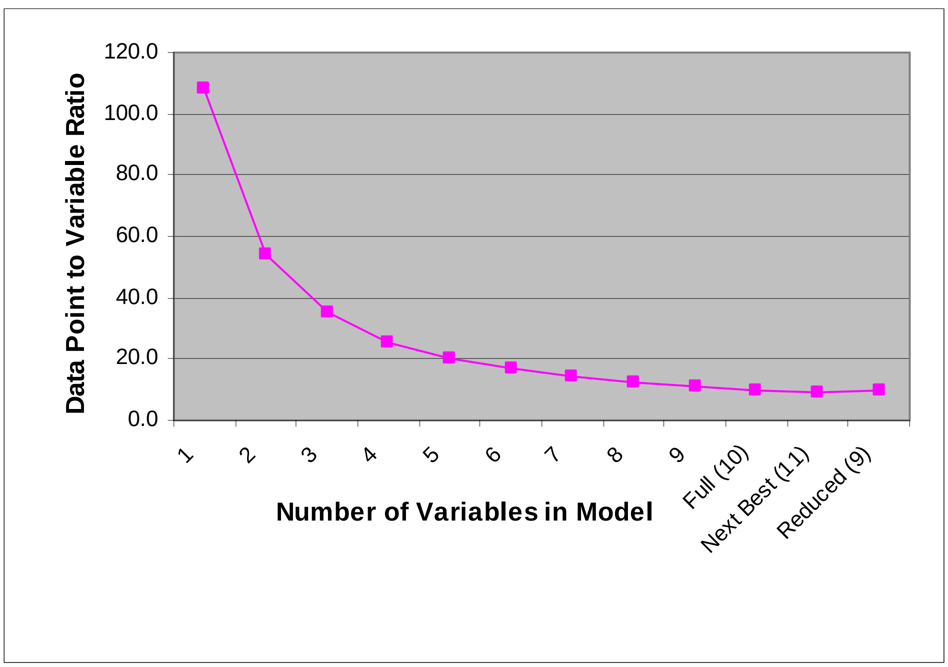



**Figure 4.2 – Relative Change in Number of Observations - *Estimating* Models (Logistic)**

Figure 4.2 graphically displays the data point to variable ratio. We are extremely suspect of any ratio less than 10:1, and we attempt to keep a 10:1 ratio if at all possible.

> The number of data points plays a particularly important role, because the higher the number of data points, the more of our population we capture in our sample. Thus, our sample becomes more representative of the population. In addition, the larger the sample size, the more predictor variables we can add before the model becomes invalid statistically. According to Neter et al., a model should have at least six to ten data points for every predictor used. Thus, in this study, if a model falls below ten data points per predictor, then we carefully consider the additional benefits to the model gained by adding the variable (Neter, 1996:437) (Sipple, 2002:76).

As we see in Figure 4.2 the ratio of data points to variables per model sharply decreases as we add variables then plateaus at around ten to one. When we reduce the full model we lose 14 data points (86 data points total); however, we also reduce the number of variables in the model to 9. This gives us a 9.6:1, or an approximate 10:1 ratio. In effect, we have a parsimonious model with the most statistically significant predictor variables.

Next we look at the p-values associated with each best model. As we state in Chapter 3, we use the sum of all individual p-values in each model when we weight them against one another. The reasoning for this is that the whole model chi-squared test does not assure us that every independent predictor variable is significant, only that the whole model has statistical significance as a predictive model. When our models contain greater than three or four variables the whole model chi-squared p-value is $< 0.0001$ for all models. Thus, the whole model p-value is an indiscriminant performance measure.

Figure 4.3 displays the change in the sum of individual p-values as we progressively build our model. Our goal is to have the lowest p-values both individually and collectively for our model. We see the change in p-value for each model from model one to four as more or less unchanging. The next three models begin with a slight increase then a gradual decrease. From there, there is a relatively large increase from our seven to our eight variable model then a decreasing trend down to the reduced model. The increase from our full model to our next best model throws up a 'red flag' indicating that we are starting to over fit our data set, so we stop at our ten variable model and reduce from there.

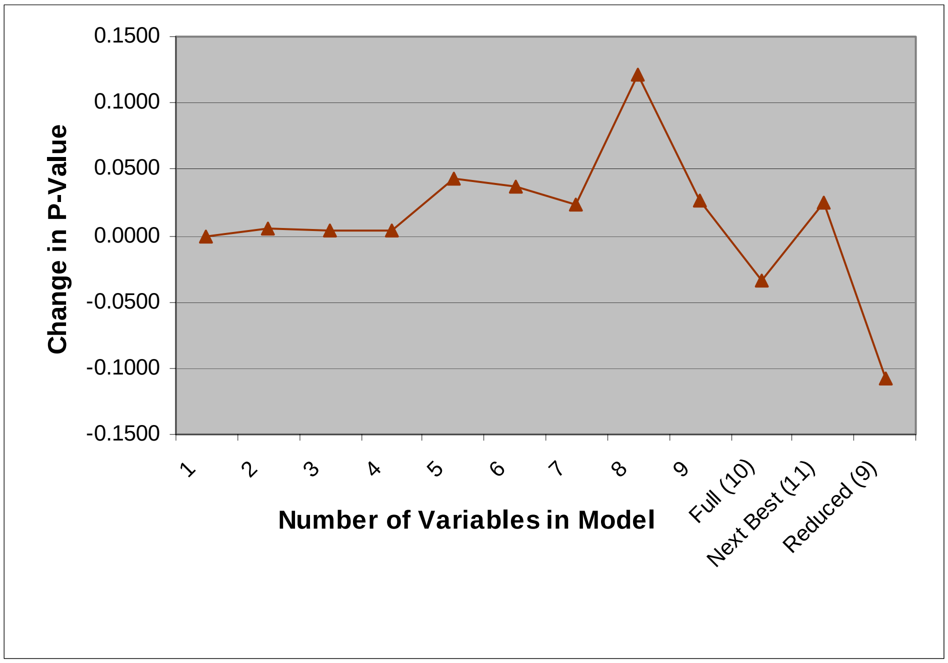



**Figure 4.3 – Relative Change in P-Value - *Estimating* Models (Logistic)**

Lastly, we look at the area under the curve (AUC). For a detailed explanation of this measure see Sipple (2002) and Bielecki (2003). Generally, the higher the AUC the more accurate our model is at predicting cost growth.

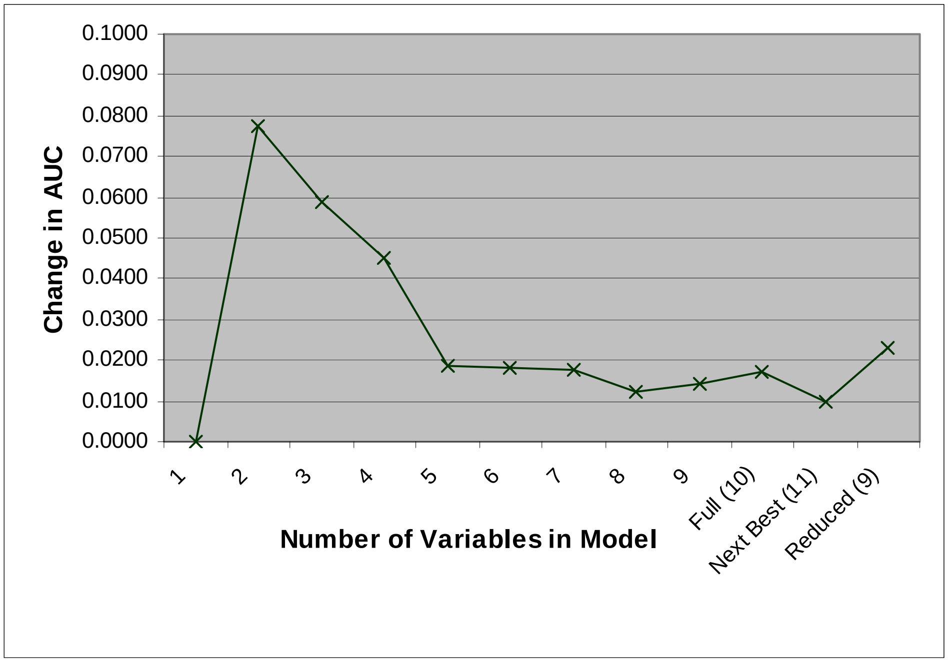



**Figure 4.4 – Relative Change in AUC - *Estimating* Models (Logistic)**

In Figure 4.4 we see the change in AUC increase relatively substantially from our one to two variable models then sharply decline to the five variable model where the change then levels out to the full model. When we add one more variable to our 10 variable model, we see a decrease in AUC. This decrease, together with the decrease in $R^2$ (U), decrease in data point to variable ratio, and increase in p-values, indicate to us that the eleven variable model offers no performance over our ten variable model. Thus, we reduce the ten variable model to find that all performance measures increase dramatically.

See Appendix B for complete results and JMP® output of both full and reduced logistic — *estimating* models. Below are the parameter estimates, Figure 4.5, of the reduced model and the ensuing probability formula, Figure 4.6 which we submit for validation. Note that the numbers in parentheses in the formula of Figure 4.6 are actually the numbers of the predictor variables themselves not constants. In this formula '$P_{est}$' is the probability of a zero or one. JMP® uses a cut-off of 50 % to determine whether a program has cost growth. Above 50% is coded a one and below 50% is coded a zero.

**Parameter Estimates**

| Term | Estimate | Std Error | ChiSquare | Prob>ChiSq |
|---|---|---|---|---|
| Intercept | 3.74251185 | 1.9561775 | 3.66 | 0.0557 |
| 7 ACAT 1? | -4.3368579 | 1.3497976 | 10.32 | 0.0013 |
| 77 LRIP Planned? | -2.4954635 | 1.1182264 | 4.98 | 0.0256 |
| 38 Lockheed-Martin | -2.8377295 | 1.2719104 | 4.98 | 0.0257 |
| 67 Class - U | 3.15286508 | 1.2494315 | 6.37 | 0.0116 |
| 15 Aircraft | 4.38455975 | 1.5281374 | 8.23 | 0.0041 |
| 44 No Major Defense Contractor | 4.15463156 | 1.352822 | 9.43 | 0.0021 |
| 39 Northrop Grumman | 5.14122691 | 1.9324404 | 7.08 | 0.0078 |
| 1 / Variable # 3 | 0.58771192 | 0.2620326 | 5.03 | 0.0249 |
| ln(Variable # 51) | -1.6495495 | 0.8215535 | 4.03 | 0.0447 |

For log odds of 0/1

**Figure 4.5 – Parameter Estimates - *Estimating* Model (Logistic)**

$$P_{est} := \frac{e^{-(x)}}{1 + e^{-(x)}}$$

Where:

$$x := 3.7425 - 4.3369 \cdot (V7) - 2.4955 \cdot (V77) - 2.8377 \cdot (V38) + 3.1529 \cdot (V67) + 4.3846 \cdot (V15) + 4.1546 \cdot (V44) + 5.1412 \cdot (V39) + 0.5877 \cdot \left(\frac{1}{V3}\right) - 1.6496 \cdot (\ln 51)$$

**Figure 4.6 – Probability Formula - *Estimating* Model (Logistic)**

### Logistic Regression Results — *Support* Response

Now that our model building methodology and weighting criteria are fully understood we begin our discussion of our logistic regression — *support* response with a summary of the best models at each round of our logistic model building process.

| Logistic (Support) Best Models | | | | |
|---|---|---|---|---|
| # Variables | R Sq (U) | Obs | P-Value | AUC |
| 1 | 0.0959 | 108 | 0.0007 | 0.67667 |
| 2 | 0.1463 | 108 | 0.0096 | 0.73431 |
| 3 | 0.2265 | 93 | 0.0333 | 0.81512 |
| 4 | 0.2889 | 93 | 0.0338 | 0.84372 |
| 5 | 0.3595 | 90 | 0.0709 | 0.87723 |
| 6 | 0.4028 | 90 | 0.1382 | 0.89038 |
| 7 | 0.4266 | 90 | 0.2519 | 0.90179 |
| 8 | 0.4566 | 90 | 0.3043 | 0.9127 |
| Full (9) | 0.4896 | 90 | 0.2318 | 0.93105 |
| Next Best (10) | 0.5121 | 90 | 0.353 | 0.93353 |
| Reduced (9) | 0.4896 | 90 | 0.2318 | 0.93105 |

**Table 4.6 – Best Logistic *Support* Models For Each Generation**

With the performance measures for each best model stated in Table 4.6, we illustrate and discuss in the following graphs the relative changes of each performance measure as the number of variables increase. Note that our full and reduced models are the same model. Again, we begin our discussion with the relative change in $R^2$ (U), and continue with the data point to variable ratio, relative change in p-value, and AUC.

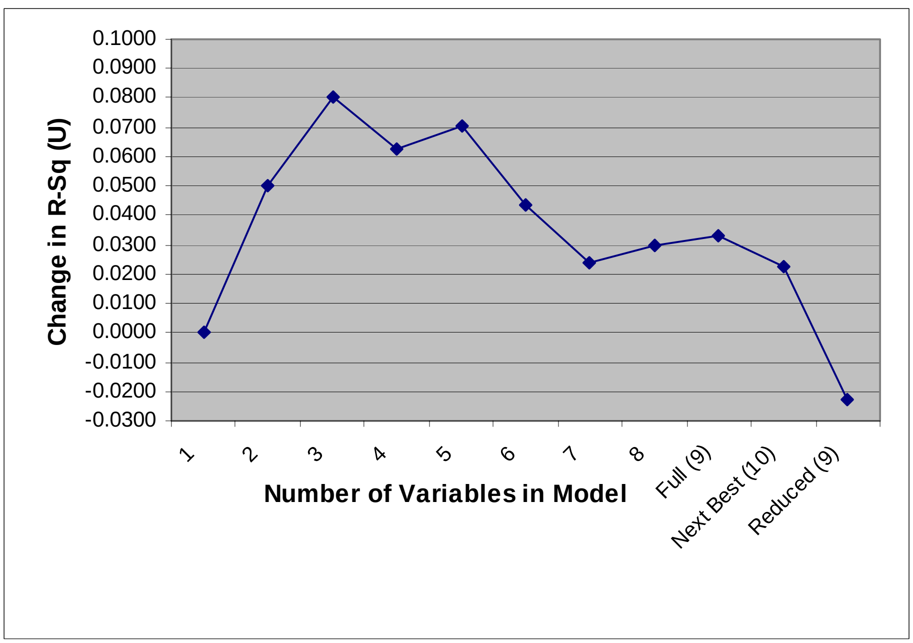



**Figure 4.7 – Relative Change in R Sq (U) - *Support* Models (Logistic)**

In Figure 4.7 we see the relative change in $R^2$ (U) expectedly increase at a decreasing rate with the addition of each new variable. Note the sharp decrease from the next best model to the reduced model; however, keep in mind that the full and reduced models are one in the same. The reasons for not selecting the next best model as our full model are apparent in the discussions of the following performance measures.

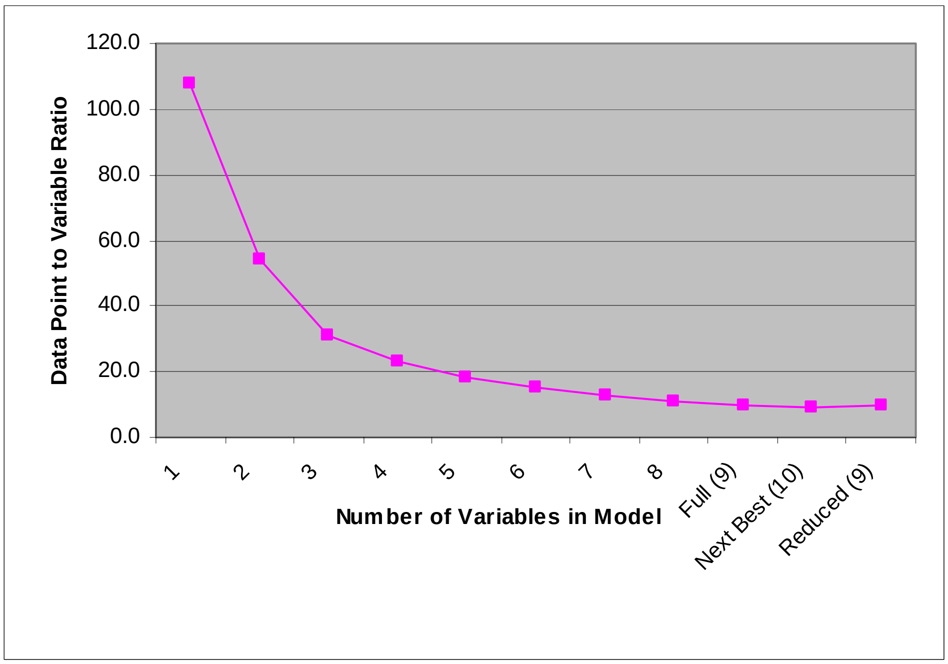



**Figure 4.8 – Relative Change in Number of Observations - *Support* Models (Logistic)**

As is the case in our logistic-*estimating* response, Figure 4.8 shows the ratio of data points to variables per model sharply decreases as we add variables then plateaus at around ten to one. Even thought the ten variable model gives us a 9:1 ratio, which is within the acceptable range of 6 to 10 as defined by Neter (1996), we are suspect of any ratio that falls below ten to one, thus we lean towards our nine variable reduced model.

The relative change of the next performance measure, sum of individual p-values, is displayed in Figure 4.9 below. As we expect, the p-values increase as variables are added, then begin to decrease as each independent predictor variable adds predictive statistical significance to the whole model.

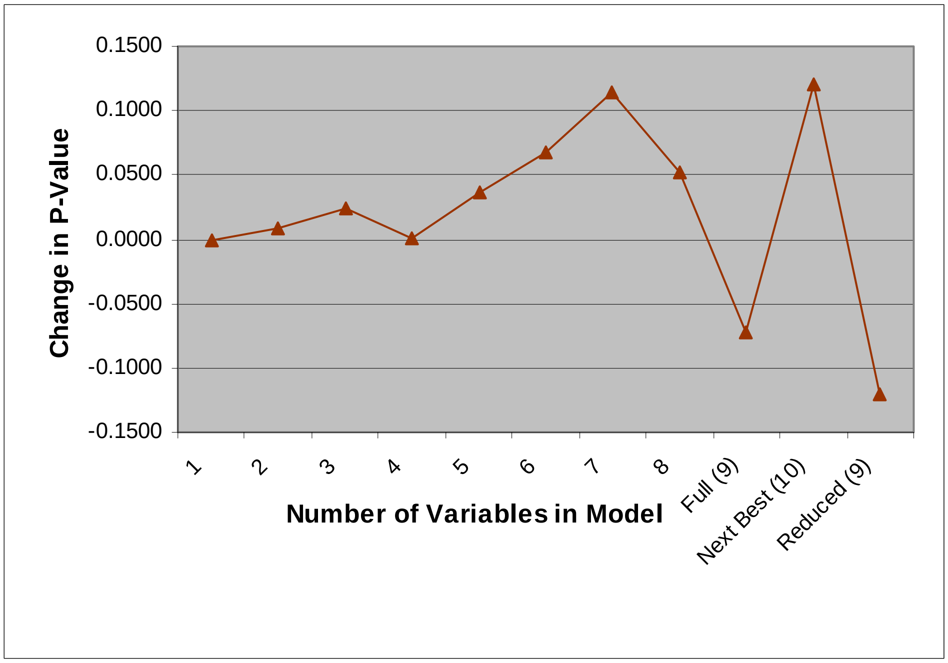



**Figure 4.9 – Relative Change in P-Value - *Support* Models (Logistic)**

When the tenth variable is added to the model the p-values increase by 0.1212 to 0.3530. This increase concerns us because the larger the sum of the p-values, the less predictive the model is. This huge increase in p-value is the leading reason we choose not to accept the ten variable model. We look at our last performance measure, the area under the curve to make our final determination.

In Figure 4.10, we see the AUC call to mind the law of diminishing returns. The AUC increases only 0.0025 from the nine to ten variable models. This is not a large enough increase for us to sacrifice what little parsimoniousness we attain with the nine variable model, thus we abandon the ten variable model in favor of the nine variable model.  All attempts to reduce the nine variable model are unsuccessful.  Therefore, our full model and reduced model are the same.

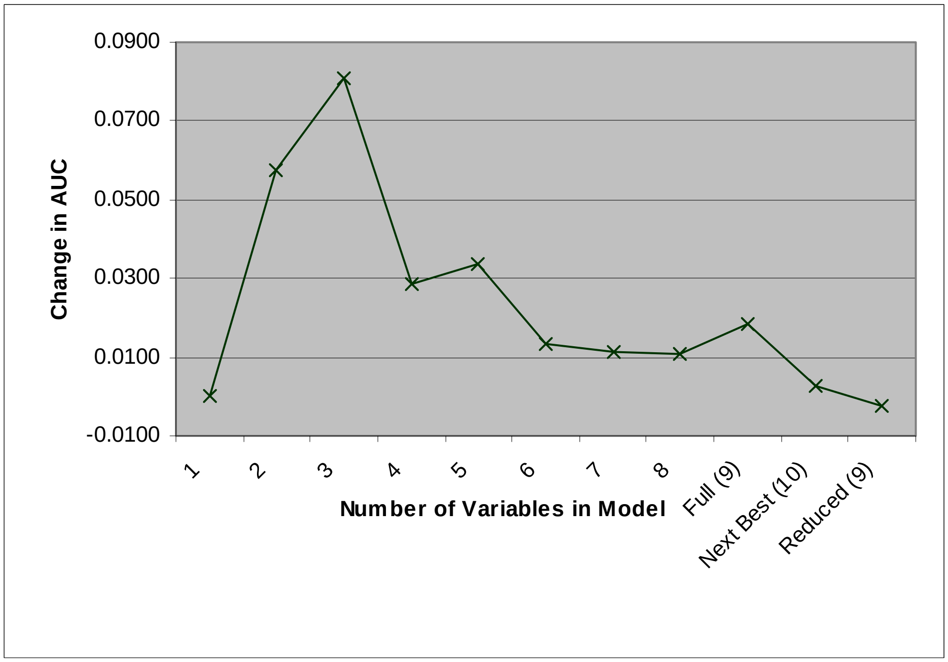



**Figure 4.10 – Relative Change in AUC - *Support* Models (Logistic)**

See Appendix C for complete results and JMP® output of our logistic – *support* model.  Below are the parameter estimates, Figure 4.11, of the reduced model and the ensuing probability formula, Figure 4.12, which we submit for validation.  Note again that the numbers in parentheses in the formula of Figure 4.12 are actually the numbers of the predictor variables themselves not constants.  In this formula '$P_{sup}$' is the probability of a zero or one.  JMP® uses a cut-off of 50 % to determine whether a program has cost growth.  Above 50% is coded a one and below 50% is coded a zero.

**Parameter Estimates**

| Term | Estimate | Std Error | ChiSquare | Prob>ChiSq |
|---|---|---|---|---|
| Intercept | 2.69155828 | 1.3671233 | 3.88 | 0.0490 |
| 50 Funding Yrs of Proc Completed | -0.2905103 | 0.0886537 | 10.74 | 0.0010 |
| 76 Program have a MS I? | 2.63525559 | 0.956811 | 7.59 | 0.0059 |
| 18 Space (RAND) | 6.58080015 | 2.7361572 | 5.78 | 0.0162 |
| 46 Fixed-Price EMD Contract? | 2.30444445 | 0.9818671 | 5.51 | 0.0189 |
| 66 Class - C | -6.2703953 | 1.9867388 | 9.96 | 0.0016 |
| 13 Helo | -3.2106087 | 1.8037027 | 3.17 | 0.0751 |
| 35 N involvement | 2.580247 | 1.0233947 | 6.36 | 0.0117 |
| 62 Proc Started based on Funding Yrs? | -2.8842632 | 1.4347345 | 4.04 | 0.0444 |
| 21 # of Svs | -0.8539183 | 0.448616 | 3.62 | 0.0570 |

For log odds of 0/1

**Figure 4.11 – Parameter Estimates - *Support* Model (Logistic)**

$$P_{sup} := \frac{e^{-(x)}}{1 + e^{-(x)}}$$

Where:

$$x := 2.6916 - 0.2905\cdot(V50) + 2.6353\cdot(V76) + 6.5808\cdot(V18) + 2.3044\cdot(V46) - 6.2704\cdot(V66) - 3.2106\cdot(V13) + 2.5802\cdot(V35) - 2.8843\cdot(V62) - 0.8539\cdot(V21)$$

**Figure 4.12 – Probability Formula - *Support* Model (Logistic)**

### Multiple Regression Results — *Estimating* Response

Since we use the same methodology to build our ordinary least squares (OLS) models as we do our logistic models we do not discuss the step-by-step process as we do in our first logistic model above. However, we do comment on the difference in performance measurements we use to weight the OLS models versus the logistic models.

We still use the data point to variable ratio and sum of individual p-values as performance measures for the same reasons; however, instead of an $R^2$ (U) and the area under the receiver-operating curve (AUC), we use $R^2$ and adjusted $R^2$. As we mentioned earlier, the adjusted $R^2$ penalizes the model for adding too many independent variables. By penalize we mean that the adjusted $R^2$ weighs the model by the number of independent variables and number of observations included in the model. While $R^2$ is a measure of the amount of variation explained by our model, adjusted $R^2$ is not — instead, it is a value that allows us to compare our models to one another.

> In theory, using an infinite number of independent variables to explain the change in a dependent variable would result in an $R^2$ of one. In other words, the $R^2$ value can be manipulated and should be suspect. The adjusted $R^2$ value is an attempt to correct this shortcoming by adjusting both the numerator and denominator by their respective degrees of freedom (see Figure 4.13 below). Unlike the $R^2$, the adjusted $R^2$ can decline in value if the contribution to the explained deviation by the additional variable is less than the impact on the degrees of freedom. This means that the adjusted $R^2$ will react to alternative equations for the same dependent variable in a manner similar to the Standard Error of the Estimate (SEE); i.e., the equation with the smallest SEE will most likely also have the highest adjusted $R^2$ (Jensen, 2003).

$$\mathrm{Adj}R^2 := 1 - \left(1 - R^2\right)\left(\frac{n-1}{n-k-1}\right)$$

where: $n$ = number of observations
$k$ = number of independent variables

**Figure 4.13 – Formula for Calculating Adjusted $R^2$**

We begin our discussion of our OLS regression model — *estimating* response with a summary of the best models at each round of our model building process (see Table 4.7). With the performance measures for each best model stated in Table 4.7, we illustrate and discuss in the following graphs the relative changes of each performance measure as the number of variables increase. We begin our discussion with the relative change in the difference between $R^2$ and adjusted $R^2$, and continue with data point to variable ratio, and relative change in p-value.

| OLS (Estimating) Best Models | | | | |
|---|---|---|---|---|
| # Variables | R Sq | Adj R Sq | Obs | P-Value |
| 1 | 0.168298 | 0.154201 | 61 | 0.001 |
| 2 | 0.281471 | 0.255342 | 58 | 0.0159 |
| 3 | 0.361658 | 0.320027 | 50 | 0.0235 |
| 4 | 0.407226 | 0.362488 | 58 | 0.0508 |
| Full (5) | 0.48983 | 0.431856 | 50 | 0.1029 |
| Next Best (6) | 0.517803 | 0.447238 | 48 | 0.1929 |
| Reduced (4) | 0.578606 | 0.538473 | 47 | 0.0242 |

**Table 4.7 – Best OLS *Estimating* Models for Each Generation**

The difference between $R^2$ and adjusted $R^2$ is shown in Figure 4.14. We want an adjusted $R^2$ as close to the $R^2$ value as possible while also maximizing our other performance measures. In Figure 4.14 we see the difference, or 'gap', between the $R^2$ and adjusted $R^2$ steadily increase with the addition of each variable into the model up to our next best six variable model. Note that the gap between the two measurements is better for the reduced model than for both the full and next best models. This decreased difference or shorter gap is what we desire.

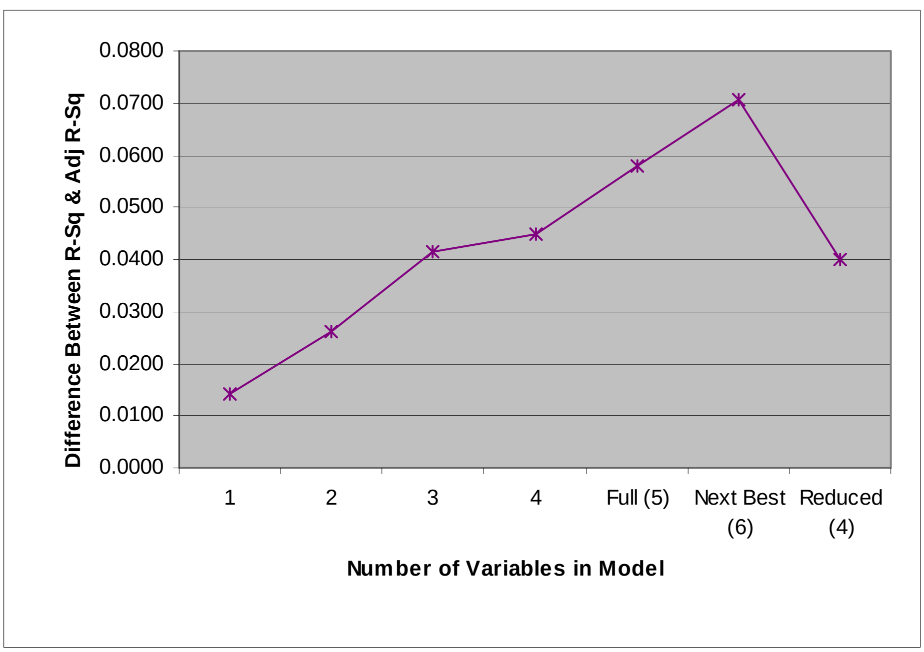



**Figure 4.14 – Relative Change Between $R^2$ and adjusted $R^2$ - *Estimating* Models (OLS)**

Next we evaluate the ratio of data points to variables. In Figure 4.15 we see the ratio drop as variable are added until we reach a ten to one ratio for the full model and an eight to one ratio for the next best. The reduced model has 47 data points and four predictor variables which gives us a data point to variable ratio of 11.75:1. This is a welcomed improvement over the full model ratio of 10:1.

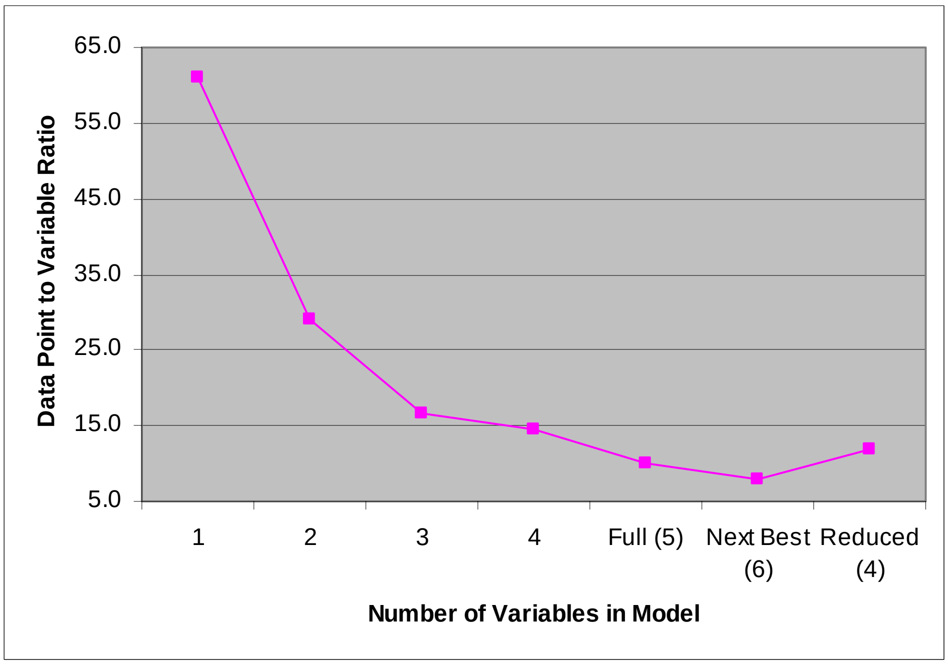



**Figure 4.15 – Ratios of Data Points to Variables - *Estimating* Models (OLS)**

Finally, we observe the change in the sum of individual p-values. Figure 4.16 shows the p-values increase as we add variables to our model. It also shows the dramatic decrease we achieve when we reduce our full model. This large decrease in p-values indicates that the variables contained in our reduced model are very significant and highly predictive.

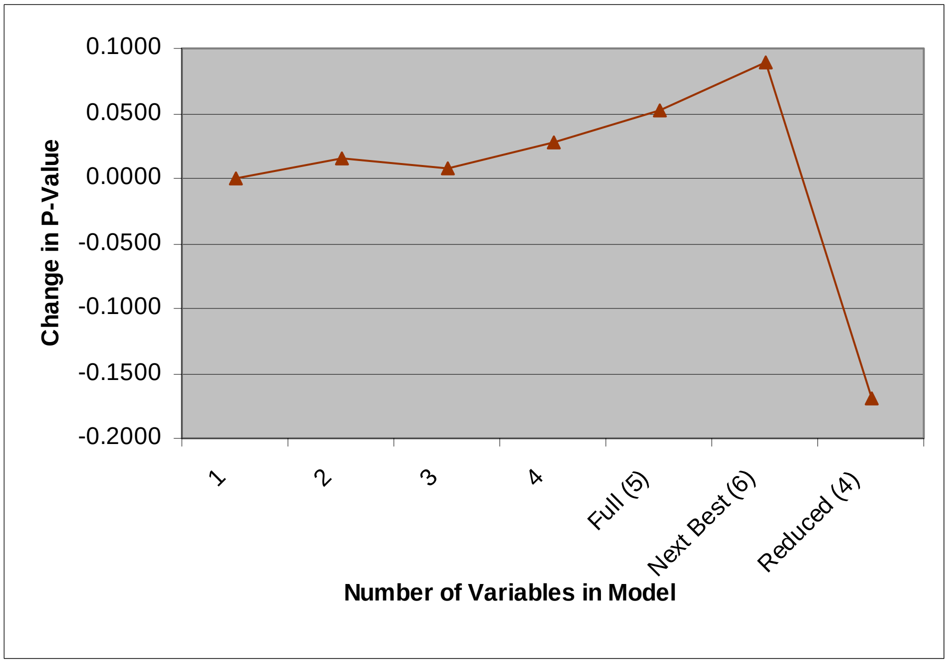



**Figure 4.16 – Relative Change in P-Value - *Estimating* Models (OLS)**

Based on these performance measures, we are confident in the predictive capability and statistical soundness of our reduced model. At this point, we must test the assumptions of the residuals of multiple regression model to see if they are satisfied by this reduced model. We do not display the tests of assumptions for the full model; however, both full models are subjected to all of the following tests, except the Breush-Pagan test for constant variance (a visual inspection of the residual plot is done instead), and meet the assumptions.

The first assumption we must satisfy is that of independence. Since we obtain and use only the most recent SAR as data for each program, we assume independence is met. Next we perform a Shapiro-Wilk goodness-of-fit test for normality. Using an alpha of 0.05, the output from JMP® in Figure 4.17 shows that our residuals do meet the assumption of normality with a p-value of 0.44 which is above our stated alpha of 0.05.

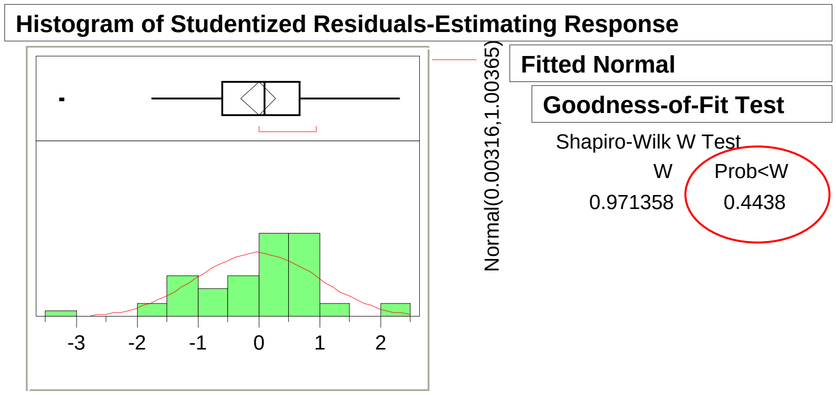



**Figure 4.17 – Shapiro-Wilk Test for Normality – *Estimating* (Reduced) Model (OLS)**

Finally, we perform a Breusch-Pagan test for constant variance of the residuals. Using Microsoft Excel® we calculate a p-value of 0.841237. This high p-value, which is above our alpha of 0.05, indicates that our residuals indeed pass the Breusch-Pagan test for constant variance.

In addition to the assumption tests, we also ensure that our model contains no influential data points. For this we use JMP® to run an overlay plot of the Cook's Distance values.

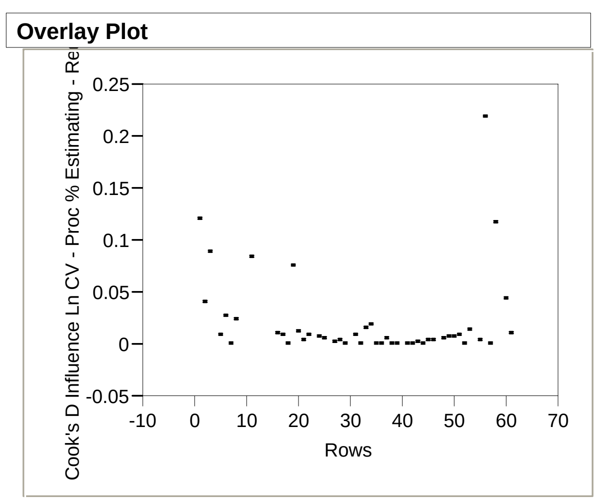



**Figure 4.18 – Cook's Distance Overlay Plot for Influential Data Points – *Estimating* (Reduced) Model (OLS)**

A Cook's Distance greater than 0.5 indicates that an influential outlier exists (Neter, 1996:381). Consequently, we would remove any outliers above 0.05 to see the effect on our model. In some cases, removal of the influential outlier may cause other influential outliers to surface causing subsequent removal of these outliers. Figure 4.18 shows no data points above 0.25, thus our model does not contain any influential data points.

Therefore, based on the successes of these tests and the overall performance measures above, we are confident in the predictive capability of our model and submit this four variable model for validation.

See Appendix D for complete results and JMP® output of our OLS – *estimating* model. Below are the parameter estimates, Figure 4.19, of the reduced model and the ensuing linear regression formula, Figure 4.20, which we submit for validation. Also, the variance inflation factors (VIF) scores are displayed in Figure 4.19. Variance inflation is the consequence of multicollinearity. In a regression model we expect a high variance explained ($R^2$). The higher the variance explained is, the better the model is. However, if collinearity exists among our predictor variables, then most likely the variance, standard error, and parameter estimates are all inflated. In other words, the high $R^2$ may not be the result of good independent predictors, but a result of a mis-specified model that carries mutually dependent and thus redundant predictors. The VIF is common way for detecting multicollinearity. The general rule of thumb is that the VIF should not exceed ten (Yu, 2004). As we see in Figure 4.19, all of our VIF scores are well below ten. Again, the numbers in parentheses in the formula of Figure 4.20 are actually the numbers of the predictor variables themselves not constants. In this formula '$Y_{est}$' gives us the estimated percentage of cost growth for the *estimating* cost variance category.

| **Parameter Estimates** | | | | | |
|---|---|---|---|---|---|
| Term | Estimate | Std Error | t Ratio | Prob>\|t\| | VIF |
| Intercept | -4.803647 | 0.46946 | -10.23 | <.0001 | . |
| 62 Proc Started based on Funding Yrs? | 2.1386646 | 0.45508 | 4.70 | <.0001 | 1.0490251 |
| (Variable #58 * Variable # 73)^2 | 0.0000926 | 0.000026 | 3.62 | 0.0008 | 1.0705798 |
| 81 Length of Proc Funding > 11 yrs? | 1.1384232 | 0.356188 | 3.20 | 0.0026 | 1.0603576 |
| 2 Total Quantity | 0.0000186 | 0.000008 | 2.40 | 0.0207 | 1.0138227 |

**Figure 4.19 – Parameter Estimates – *Estimating* (Reduced) Model (OLS)**

$$Y_{est} := e^{x}$$

Where:

$$x := -4.8036 + 2.1387 \cdot (V62) + 0.0001 \cdot (V58 \cdot V73)^2 + 1.1384 \cdot (V81) + 0.00002 \cdot (V2)$$

**Figure 4.20 – Linear Regression Equation – *Estimating* (Reduced) Model (OLS)**

### Multiple Regression Results — *Support* Response

We begin our discussion of our OLS regression model — *support* response with a summary of the best models at each round of our logistic model building process (see Table 4.8).

| **Logistic (Support) Best Models** | | | | |
|---|---|---|---|---|
| **# Variables** | **R Sq** | **Adj R Sq** | **Obs** | **P-Value** |
| 1 | 0.176725 | 0.160583 | 53 | 0.0017 |
| 2 | 0.319518 | 0.290561 | 50 | 0.0037 |
| 3 | 0.400248 | 0.360264 | 49 | 0.0152 |
| Full (4) | 0.472743 | 0.42481 | 49 | 0.0364 |
| Next Best (5) | 0.512596 | 0.455921 | 49 | 0.0887 |
| Reduced (4) | 0.542253 | 0.492767 | 42 | 0.0179 |

**Table 4.8 – Best OLS *Support* Models For Each Generation**

With the performance measures for each best model stated in Table 4.8, we illustrate and discuss in the following graphs the relative changes of each performance measure as the number of variables increase. We begin our discussion with the relative change in the difference between $R^2$ and adjusted $R^2$, and continue with data point to variable ratio, and relative change in p-value.

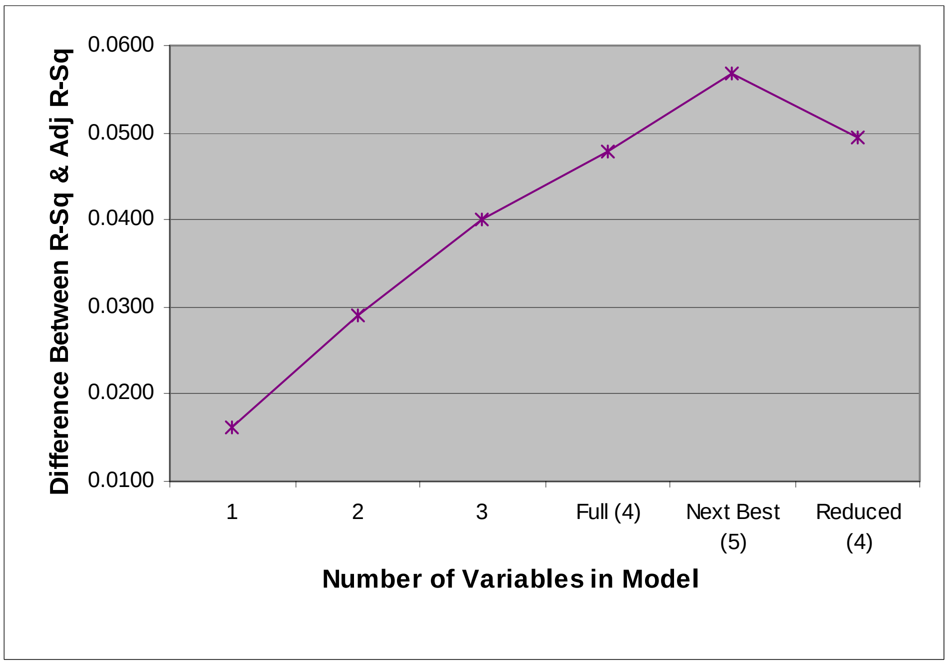



**Figure 4.21 – Relative Change Between $R^2$ and adjusted $R^2$ - *Support* Models (OLS)**

The difference between $R^2$ and adjusted $R^2$ is shown in Figure 4.22. Again, we want an adjusted $R^2$ as close to the $R^2$ value as possible while also maximizing our other performance measures. Therefore, we look for this distance to be minimized. In Figure 4.21 we see the difference, or 'gap', between the $R^2$ and adjusted $R^2$ steadily increase with the addition of each variable into the model up to our next best five variable model. Upon reduction of the full model we see the gap between the two performance measures shorten. Although the gap of our reduced model is slightly more than that of our next best model, it is still smaller than that of our full model. We look at the remaining two performance measures to make our final determination.

Next we evaluate the data point to variable ratio in our models. In Figure 4.22 we see the ratio drop as variable are added until we reach a 12.3:1 ratio for the full model and a 9.8:1 ratio for the next best. The reduced model has 42 data points and four predictor variables which gives us a data point to variable ratio of 10.5 to 1. This is above our ten to one cut-off so we move on to the final performance measure, p-value.

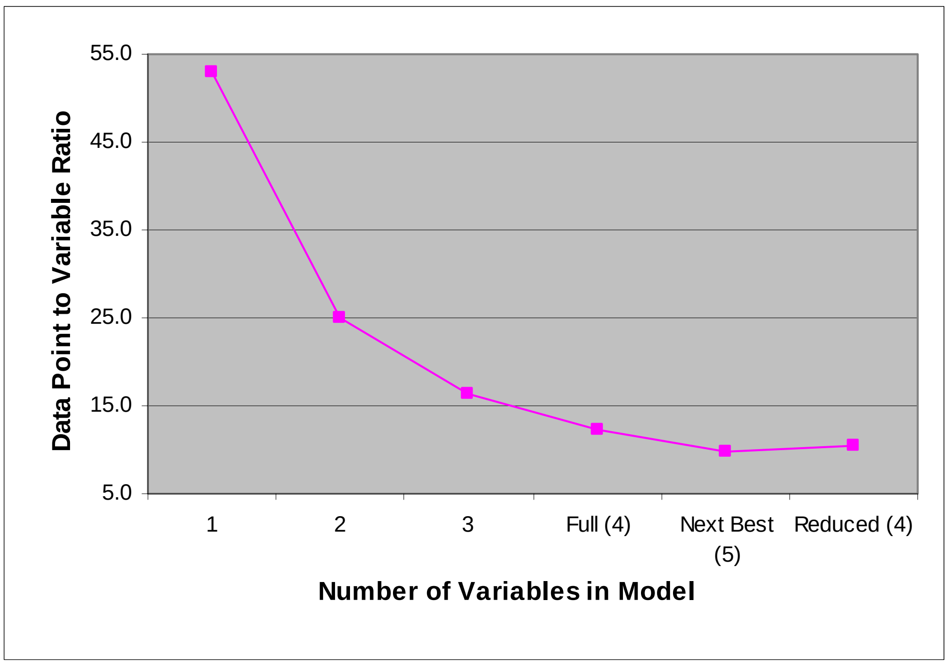



**Figure 4.22 – Ratios of Data Points to Variables – *Support* Models (OLS)**

Finally, we observe the relative change in the sum of individual p-values. Figure 4.23 shows the p-values increase as we add variables to our model. It also shows the dramatic decrease we achieve when we reduce our full model. This large decrease in p-values indicates that the variables contained in our reduced model are more significant than both our full and next best models.

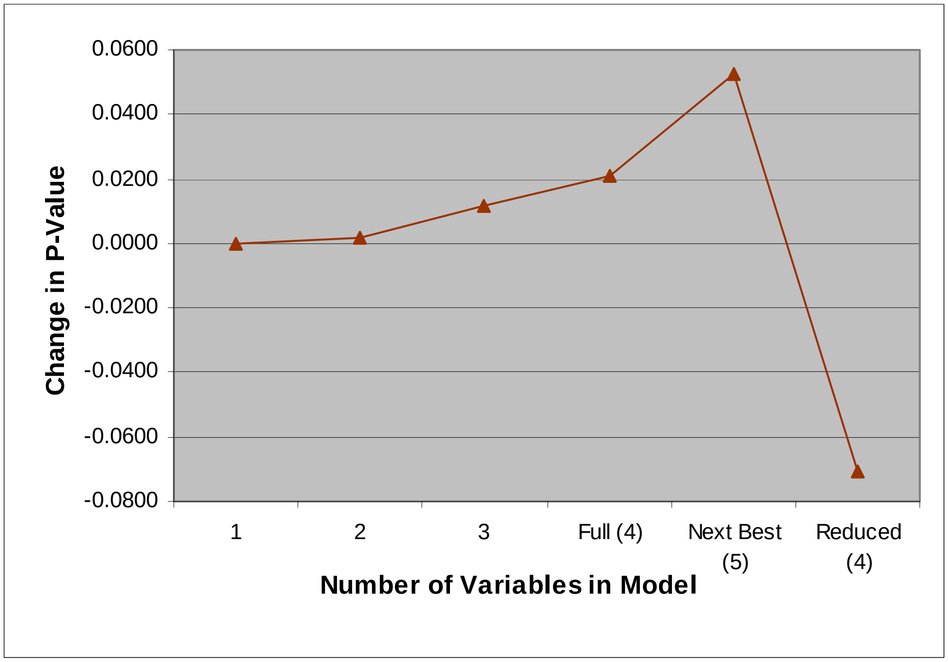



**Figure 4.23 – Relative Change in P-Value – *Support* Models (OLS)**

Based on these performance measures, we are confident in the predictive capability and statistical soundness of our reduced model. At this point, we must test the assumptions of the residuals of multiple regression model to see if they are satisfied by this reduced model. Again we do not display the tests of assumptions for the full model; however, they are performed and are met.

The assumption of independence is the same as that of the OLS – *estimating* model above. Next we perform a Shapiro-Wilk goodness-of-fit test for normality. Using an alpha of 0.05, the output from JMP® in Figure 4.24 shows that our residuals do meet the assumption of normality with a p-value of 0.45 which is above our stated alpha of 0.45.

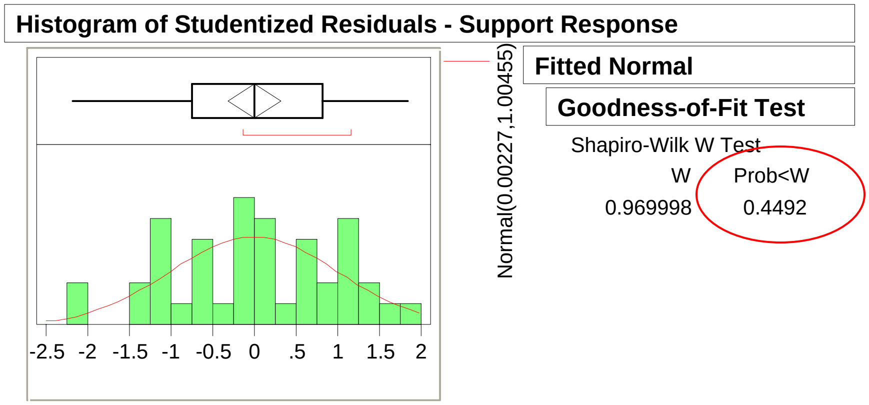



**Figure 4.24 – Shapiro-Wilk Test for Normality –** ***Support*** **(Reduced) Model (OLS)**

Finally, we perform a Breusch-Pagan test for constant variance of the residuals. Using Microsoft Excel® we calculate a p-value of 0.890527. This high p-value, which is above our alpha of 0.05, indicates that our residuals indeed pass the Breusch-Pagan test for constant variance.

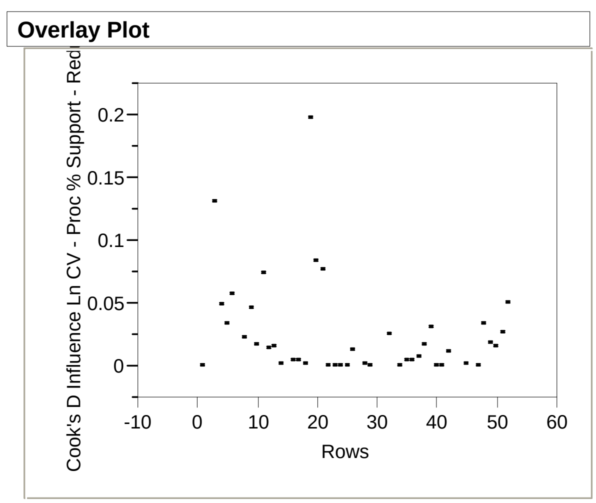



**Figure 4.25 – Cook's Distance Overlay Plot for Influential Data Points**
***Support*** **(Reduced) Model (OLS)**

In addition to the assumption tests, we also ensure that our model contains no influential data points. For this we use JMP® to run an overlay plot of the Cook's

Distance values. Figure 4.25 shows no data points above 0.25, thus our model does not contain any influential data points. Therefore, based on the successes of these tests and the overall performance measures above, we are confident in the predictive capability of our model and submit this four variable model for validation.

See Appendix E for complete results and JMP® output of our OLS – *support* model. Below are the parameter estimates, Figure 4.26, of the reduced model and the ensuing linear regression formula, Figure 4.27, which we submit for validation. Note that the variance influence factors are well below ten which indicates little or no multicollinearity. The numbers in parentheses in the formula of Figure 4.27 are actually the numbers of the predictor variables themselves not constants. '$Y_{sup}$' gives us the estimated percentage of cost growth for the *support* cost variance category.

**Parameter Estimates**

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| | VIF |
|---|---|---|---|---|---|
| Intercept | -3.064493 | 0.284403 | -10.78 | <.0001 | . |
| 26 Service = Joint | -1.35354 | 0.513573 | -2.64 | 0.0122 | 1.2030299 |
| 19 Ship | -2.491327 | 0.777774 | -3.20 | 0.0028 | 1.1868498 |
| 12 Electronic | -1.37066 | 0.42391 | -3.23 | 0.0026 | 1.0848091 |
| Variable # 58 * Variable # 80 | 0.0148537 | 0.003674 | 4.04 | 0.0003 | 1.0709069 |

**Figure 4.26 – Parameter Estimates – *Support* (Reduced) Model (OLS)**

$$Y_{sup} := e^{x}$$

Where:

$$x := -3.0645 - 1.3535 \cdot (V26) - 2.4913 \cdot (V19) - 1.3706 \cdot (V12) + 0.0149 \cdot (V58 \cdot V80)$$

**Figure 4.27 – Linear Regression Equation – *Support* (Reduced) Model (OLS)**

## Validation

*Logistic Regression Model – **Estimating** Response*

For validation, we add back the 20% validation set we create prior to model building to the 80% model building set. Once they are merged we run our model against the entire 135 data points and save the functionally predicted values ('0' or '1') for each of the validation data points. We then compare these predicted values to the actual values. JMP® computes the predicted values by assessing the probability of having cost growth based upon the factors in the specific model, wherein a '1' (yes, there is cost growth) is assigned to any point with a probability of 0.5 or greater and a '0' (no cost growth exists) otherwise.

Table 4.9 details the validation percentage of the logistic regression model – *estimating* response. The model validates our 20% validation data set at 65.2%. This is well below our expected validation percentage of 95.2% using the AUC as a guide.

Upon initial investigation of Figure 4.9 we see four programs did not validate due to missing data points within the program data. This leaves us with 23 programs to validate. Our model predicts 15 of these 23 programs correctly. The nine programs predicted incorrectly are highlighted. Of these nine, five of them predicted a '1', or that the program would have cost growth, but the actual response was a '0', or that the program did not have cost growth. This is somewhat reassuring in that our model will trigger the program manager to budget for expected cost growth in the estimating cost variance category approximately 22% of the time, but will not experience cost growth due to estimating.

| Logistic - Estimating Cost Growth (20%) | | | | Probability that response is a: | |
|---|---|---|---|---|---|
| Program # | Actual Response | Calculated Response | Validated Correctly? | 1 | 0 |
| 132 | 0 | . | N/A | N/A | N/A |
| 73 | 0 | . | N/A | N/A | N/A |
| 4 | 0 | 1 | n | 0.97578342 | 0.02421658 |
| 98 | 1 | 1 | y | 0.90825811 | 0.09174189 |
| 71 | 1 | 0 | n | 0.35604297 | 0.64395703 |
| 89 | 1 | 1 | y | 0.52290843 | 0.47709157 |
| 36 | 0 | 0 | y | 0.00647761 | 0.99352239 |
| 29 | 1 | 1 | y | 0.96745085 | 0.03254915 |
| 70 | 1 | 1 | y | 0.8383921 | 0.1616079 |
| 87 | 0 | 0 | y | 0.47868218 | 0.52131782 |
| 85 | 1 | 1 | y | 0.96445291 | 0.03554709 |
| 16 | 1 | 1 | y | 0.99291854 | 0.00708146 |
| 117 | 0 | 1 | n | 0.99845821 | 0.00154179 |
| 13 | 1 | . | N/A | N/A | N/A |
| 46 | 1 | 1 | y | 0.99967494 | 0.00032506 |
| 31 | 0 | 0 | y | 0.00152028 | 0.99847972 |
| 72 | 1 | 0 | n | 0.08563746 | 0.91436254 |
| 110 | 1 | 1 | y | 0.99773197 | 0.00226803 |
| 109 | 0 | 1 | n | 0.65689323 | 0.34310677 |
| 75 | 0 | 0 | y | 0.00378066 | 0.99621934 |
| 107 | 1 | 0 | n | 0.3627366 | 0.6372634 |
| 33 | 1 | 1 | y | 0.97029883 | 0.02970117 |
| 10 | 0 | . | N/A | N/A | N/A |
| 48 | 1 | 1 | y | 0.72022701 | 0.27977299 |
| 39 | 0 | 1 | n | 0.87806222 | 0.12193778 |
| 56 | 1 | 1 | y | 0.99503405 | 0.00496595 |
| 124 | 0 | 1 | n | 0.94947478 | 0.05052522 |
| Count | 27 | 23 | | | |
| | # Validated Correctly | | 15 | | |
| | Validation Percentage | | 65.2% | | |

**Table 4.9 – Validation of Logistic Regression Model – *Estimating* Response**

Due to the low validation percentage of 65.2% we perform a validation of our model on the 80% model building data set. We do this because we want to see if our 20% validation data set is representative of our entire database. Upon validating our model on 100% of the data set we find that our model correctly predicts cost growth in 89 out of 109 data points for a validation percentage of approximately 82%. This is much closer to our expected AUC percentage of 95.2%. Note that 4 out of the 89 that are predicted correctly are borderline probabilities, meaning they are within plus or minus .05 of the 0.5 cut-off used by JMP® to categorize them as having cost growth.

Because there is a difference between the two validations, we run distributions of each predictor variable in our reduced model from both the 20% and 80% data sets. Upon investigation of these distributions we find that one variable, *15 Aircraft*, exhibits a large enough difference in their means that we conclude the validation set is non-representative of the model building set (see Figure 4.28).

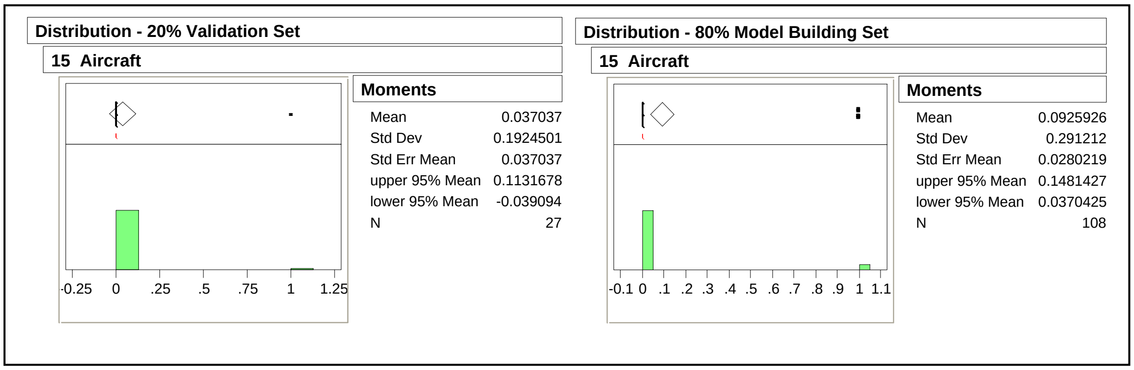



**Figure 4.28 – Variable Distribution Difference – *Estimating* Response**

As we see in Figure 4.28, the mean of the variable in the 20% data set is 0.037 which represents 1 out of 27 data points that is an aircraft, while the mean of the same variable in the 80% data set is 0.093 which represents 10 out of 108 data points. This difference is large enough that we feel it explains the poor validation we observe with our 20% data set. Due to increase in validation we observe against our entire data set, we are confident that our logistic regression model – *estimating* response will correctly predict cost growth in the *estimating* cost variance category at least 82% of the time.

*Logistic Regression Model – **Support** Response*

Table 4.10 details the validation percentage of the logistic regression model – *support* response. The model validates our 20% validation data set at 58.3%. This is well below our expected validation percentage of 93.1% using the AUC as a guide.

| Logistic - Support Cost Growth (20%) | | | | Probability that response is a: | |
|---|---|---|---|---|---|
| Program # | Actual Response | Calculated Response | Validated Correctly? | 1 | 0 |
| 132 | 0 | . | N/A | N/A | N/A |
| 73 | 0 | 0 | y | 0.00115578 | 0.99884422 |
| 4 | 1 | 0 | n | 0.29255221 | 0.70744779 |
| 98 | 0 | 1 | n | 0.90654619 | 0.09345381 |
| 71 | 1 | 0 | n | 0.41764394 | 0.58235606 |
| 89 | 1 | 0 | n | 0.00989476 | 0.99010524 |
| 36 | 0 | . | N/A | N/A | N/A |
| 29 | 1 | 1 | y | 0.73359907 | 0.26640093 |
| 70 | 0 | 0 | y | 0.21017694 | 0.78982306 |
| 87 | 1 | 0 | n | 0.03605327 | 0.96394673 |
| 85 | 0 | 0 | y | 0.0072219 | 0.9927781 |
| 16 | 1 | 1 | y | 0.99806722 | 0.00193278 |
| 117 | 0 | 0 | y | 0.01124753 | 0.98875247 |
| 13 | 0 | 0 | y | 0.00143575 | 0.99856425 |
| 46 | 1 | 0 | n | 0.10695023 | 0.89304977 |
| 31 | 0 | 1 | n | 0.83893971 | 0.16106029 |
| 72 | 1 | 0 | n | 0.41764394 | 0.58235606 |
| 110 | 1 | 0 | n | 0.01421111 | 0.98578889 |
| 109 | 0 | 0 | y | 0.01491129 | 0.98508871 |
| 75 | 0 | 1 | n | 0.81835701 | 0.18164299 |
| 107 | 1 | 1 | y | 0.90454787 | 0.09545213 |
| 33 | 0 | 1 | y * | 0.5407222 | 0.4592778 |
| 10 | 0 | 0 | y | 0.01352457 | 0.98647543 |
| 48 | 1 | 1 | y | 0.81835701 | 0.18164299 |
| 39 | 0 | 0 | y | 0.10695023 | 0.89304977 |
| 56 | 1 | 0 | y * | 0.49351519 | 0.50648481 |
| 124 | 0 | . | N/A | N/A | N/A |
| Count | 27 | 24 | | | |
| | # Validated Correctly | | 14 | | |
| | Validation Percentage | | 58.3% | | |

**Table 4.10 – Validation of Logistic Regression Model – *Support* Response**

Upon initial investigation of Table 4.11 we see three programs did not validate due to missing data points within the program data. This leaves us with 24 programs to validate. Our model predicts 14 of these 24 programs correctly. Note that the two data points with a 'y *' are borderline probabilities and are included as predicted correctly. The ten programs predicted incorrectly are highlighted. Unlike the estimating model above that predicts cost growth when there is none approximately 22% of the time, our support model predicts no cost growth when there is growth present approximately 29% of the time.

Due to the low validation percentage of 58.3% we perform the same validation of our model on 100% of the data set and, also, run distributions of the predictor variables. Upon validating our model on 100% of the data set we find that our model correctly predicts cost growth in 90 out of 114 data points for a validation percentage of approximately 80%. This is much closer to our expected AUC percentage of 93.1%. Note that 3 out of the 90 that are predicted correctly are borderline probabilities.

When we compare the distributions of each predictor variable from both sets of data we find three of the predictor variables non-representative in the 20% validation set. Table 4.11 outlines the variable and the differences in their means for each data set.

| **Log - *Support*** | **Difference in Mean** | |
|---|---|---|
| | **20%** | **80%** |
| 13 Helo | 0.037 | 0.111 |
| 18 Space (RAND) | 0.111 | 0.046 |
| 66 Class C | 0.030 | 0.129 |

**Table 4.11 – Variable Distribution Differences – *Support* Response**

For the *13 Helo* variable, only one program in the validation set is a ‘1’ while 13 ‘helos’ are represented in the 80% data set. For the *18 Space (RAND)* variable, 3 programs in the validation set are a coded as ‘1’ while only 5 are represented in the 80% data set. For the *66 Class C* variable, one program is coded as a ‘1’ while 14 are represented in the 80% data set. These differences are large enough to explain the poor validation we observe with our 20% data set. Due to increase in validation we observe against our entire data set, we are confident that our logistic regression model – *support* response will correctly predict cost growth correctly in the *support* cost variance category at least 80% of the time.

*Ordinary Least Squares Model – **Estimating** Response*

For multiple regression validation, we use the same 20% validation data set, which we used for logistic regression validation. The OLS validation consists of combining the validation data set with our working data set, and saving the predicted values for each individual model to be validated. JMP® computes the predicted value by fitting the specified model parameters with the values of the validation set. We then calculate a 80 percent upper prediction bound, back-transform the log normal *Y* response to a percentage, and assess the accuracy of the model's prediction capability. We gauge the accuracy by comparing the actual percentage cost growth (*Y* response un-transformed) to the upper prediction bound. A success is recorded when the prediction bound contains the actual value, or stated another way, if the actual value is less than the prediction bound (Bielecki, 2002:70)

Unlike Bielecki and Moore, who use an 80% prediction bound, we use a 90% prediction bound due to Dr. Sambur's vision of institutionally implementing a 90% confidence level to meet cost requirements (see *The Acquisition Environment*, Chapter 2).

| OLS - *Estimating* Cost Growth | | | |
|---|---|---|---|
| Program # | Actual Response | Upper Bound | Validated Correctly? |
| 13 | 0.00027 | . | N/A |
| 29 | 0.04790 | 1.49921 | y |
| 33 | 0.18153 | 0.30274 | y |
| 98 | 0.20895 | 0.92770 | y |
| 71 | 0.26153 | 0.30228 | y |
| 107 | 0.26682 | 0.30733 | y |
| 46 | 0.39296 | 0.92057 | y |
| 89 | 0.39598 | 2.11896 | y |
| 72 | 0.41449 | 14.94475 | y |
| 70 | 0.52910 | 0.30343 | n |
| 85 | 0.73965 | 0.92188 | y |
| 48 | 0.99298 | 0.92054 | n |
| 16 | 1.00096 | 1.22253 | y |
| 110 | 1.18442 | . | N/A |
| 56 | 4.05634 | 0.92072 | n |
| Count | 15 | 13 | |
| | # Validated Correctly | | 10 |
| | Validation Percentage | | 76.9% |

**Table 4.12 – Validation of Multiple Regression Model – *Estimating* Response**

Table 4.12 details the validation percentage of the multiple regression model – *estimating* response. The model validates our validation data set at 76.9%. Out of a possible 13 data points, 10 are below the prediction bound. We consider this to be successful because we expect to see approximately 90% of the validation data points to fall below the prediction bound. To further validate our model we validate the entire data set and find the validation percentage to be 91.7%, or 55 out of 60 possible data points fall below the prediction bound. Thus, we are confident that this model will correctly predict the amount of cost growth for the *estimating* cost variance category.

*Ordinary Least Squares Model –* ***Support*** *Response*

Table 4.13 details the validation percentage of the multiple regression model – *support* response. The model validates our validation data set at 72.7%.

| Logistic - *Support* Cost Growth | | | |
|---|---|---|---|
| Program # | Actual Response | Upper Bound | Validated Correctly? |
| 48 | 0.00562 | 0.11615 | y |
| 46 | 0.01701 | 0.18432 | y |
| 72 | 0.02483 | 1.07941 | y |
| 29 | 0.03003 | 0.11383 | y |
| 107 | 0.05202 | 0.23089 | y |
| 71 | 0.05729 | 0.53772 | y |
| 89 | 0.06082 | 0.26458 | y |
| 110 | 0.15563 | . | N/A |
| 16 | 0.26119 | 0.47181 | y |
| 4 | 0.31513 | 0.23089 | n |
| 87 | 0.36141 | 0.01749 | n |
| 56 | 0.61879 | 0.35615 | n |
| Count | 12 | 11 | |
| | # Validated Correctly | | 8 |
| | Validation Percentage | | 72.7% |

**Table 4.13 – Validation of Multiple Regression Model –** ***Support*** **Response**

Out of a possible 11 data points, 8 are below the prediction bound. We consider this to be successful because we expect to see approximately 90% of the validation data points to fall below the prediction bound. To further validate our model we validate the entire data set and find the validation percentage to be 88.7%, or 47 out of 53 possible data points fall below the prediction bound. Thus, we are confident that this model will correctly predict the amount of cost growth for the *support* cost variance category.

**Chapter Summary**

This chapter reports the results of both logistic and multiple regression models for the *estimating* and *support* cost variance categories. We identify some redundant predictor variables and other predictor variables that provide no statistical significance to each cost variance category. As we detail the findings of our model building we discuss the performance measures and weighting process used to select the best models. Finally, we validate each model to asses its accuracy and usefulness.

Our analysis shows that both logistic regression models contain predictor variables that are not fully represented in the validation data set; however, upon further validation of the model building data set, the models perform well at predicting cost growth in both categories. We also show that both reduced OLS regression models are very accurate at predicting the amount of cost growth in each category. In the next chapter we entertain a final discussion and application of all of the models presented in this chapter to include a comparison of our models to those developed by Moore (2003).

## Conclusions

### Chapter Overview

This chapter reviews the pressures that exist in the DoD acquisition environment of major weapons systems procurement which underscore the necessity of this research (Bielecki, 2003:76). We summarize the pressures placed on the cost estimating community, and discuss the limitations of extrapolating our research findings to other areas of cost research. We look at our additions to the exhaustive literature review performed by Sipple (2002), and review the methodology used in this research. We restate our findings and use the current F-22 program to further validate the accuracy of our models. Finally, we explore recommendations of and possible follow-on theses to this research.

### Restatement of the Problem

Defense spending has undergone great change in the last 20 years—large increases during the Reagan Administration of the 1980s, and record setting reductions under the Clinton Administration of the 1990s. The threat to the security of the United States, however, has not declined; merely changed form. This puts the defense acquisition community in the position of having to find ways to do more with less. For this reason, elected representatives, as well as higher ranking members of the Department of Defense pay close attention to the cost performance of major defense acquisition programs. This scrutiny is the cause behind Dr. Marvin Sambur's new policy of meeting cost and performance goals with a 90% confidence level.

Our research gives the cost estimating community quantitative tools to aid the estimator in achieving these levels. The models provided by our research will enable the cost estimator to estimate cost growth early in the Engineering and Manufacturing Development (EMD) phase of a program. This ability allows the program manager to budget dwindling resources with greater confidence; thereby promoting greater credibility of the Department of Defense (DoD) acquisition community to the American public.

## Limitations

Through our research we aim to predict the presence and magnitude of cost growth in the procurement appropriations *estimating* and *support* categories during the EMD phase of a program life cycle. Our models are built from historical selected acquisitions reports (SAR)s of DoD programs from the years 1990 to 2002. Only programs with a development estimate (DE) are entered into our database, and we focus exclusively on procurement appropriations. Therefore, the use and application of our models are limited by these boundaries, and we caution the reader against extrapolating our resulting models beyond these bounds.

## Review of Literature

We add to the exhaustive literature review accomplished by Sipple (2002) with the review of Sipple (2002), Bielecki (2003), and Moore (2003). That is to say, that this follow-on research is bench-marked against these three using Sipple's predictor variables, procedures, and overall methodology. In addition to the above three theses, we find and add a study, *Cost Growth of Major Defense Programs*, by the Office of the Secretary of Defense Cost Analysis Improvement Group (OSD CAIG) to our literature review.

This study, like ours and that of our predecessors, evaluate cost growth as of the EMD phase of the system life cycle. This study is different in that the OSD does not focus on a single SAR cost variance category or a single appropriation. Instead, they seek to categorize cost growth into one of two categories: *decisions* or *mistakes*. From their results we take away their finding that cost estimating assumptions account for the majority of cost growth in the *mistakes* category.

**Review of Methodology**

We utilize the logistic and multiple regression two-step methodology introduced by Sipple (2002) to predict cost growth during our research. This two-step process first uses logistic regression to establish whether or not a program will experience cost growth. If it does experience such growth, then multiple regression is used to predict the percentage of cost growth for that program. Our research focuses strictly on the *estimating* and *support* cost variance categories of procurement appropriations in the EMD phase of program development.

We update and use the database originally created by Sipple (2002). This database is comprised of major acquisitions programs from all service components, which use a DE baseline estimate. The database contains both RDT&E and procurement dollar programs that have an EMD phase of development between 1990 and 2001, to which we add calendar year 2002 programmatic SAR data. We convert all programmatic dollar amounts into a common base year (2002) and compute our response variables.

Our database contains 135 potential data points of which 80% is used to develop our models and 20% is used to validate our models.

Before we develop our multiple regression models for both cost variance categories, we transform the Y response using a natural logarithm to ensure that the underlying assumption of heteroscedasticity (constant variance) in the residual plots is met. From here we begin to build our models by regressing each predictor variable on each response variable one at a time until the following performance measures are maximized and the most parsimonious model is achieved:

| **Model** | **Performance Measures** |
|---|---|
| Logistic Regression | Sum of the individual p-values<br>R-Squared (U)<br>Number of observations<br>Area under the receiver operating curve (ROC) |
| Multiple Regression | Sum of the individual p-values<br>R-Squared<br>Adjusted R-Squared<br>Number of observations |

Each model is then validated using the 20% validation data set that is set aside before model development.

**Restatement of Results**

Our research yields one logistic regression model and one multiple regression model for each (*estimating* and *support*) cost variance category. The validation percentage or accuracy rate of each model is detailed in Table 5.1.

| Accuracy Rate of Each Model | | | |
|---|---|---|---|
| **Model** | **Cost Variance Category** | **20% Validation Rate** | **100% Validation Rate** |
| Logistic | *Estimating* | 65.2% | 81.70% |
| | *Support* | 58.3% | 78.90% |
| Multiple | *Estimating* | 76.9% | 91.7% |
| | *Support* | 72.7% | 88.7% |

**Table 5.1 – Validation Rate of Regression Models – *All* Responses**

Upon investigation of the low validation rates among the logistic models we find one predictor variable contained in the validation set (V15) is non-representative of the 80% database for the logistic *estimating* model, and three predictor variables contained in the validation dataset (V13, V18, V66) are non-representative of the 80% database for the logistic *support* model. However, based on the validation rates of the 100% dataset, we are confident that both logistic regression models will correctly predict cost growth in both cost variance categories. Since both multiple regression models validation rates encapsulate the 90% upper prediction bound, we are confident that both multiple regression models will correctly predict cost growth in both cost variance categories.

*F-22 Validation*

To see how our models fare with an on going high profile program, we collect data on the F-22 Raptor program and put our models to the test. We plug the necessary predictor variables into the formulas for the *estimating* response as outlined in Figures 4.11 and 4.27 and find that our logistic model predicts that there is a 0.9943 probability, or 99.4% chance, that the F-22 program will experience cost growth in the *estimating* cost variance category. Furthermore, our multiple regression model yields the amount of cost growth to be 70.1%. Comparing these results to the actual results in our database, we find that there is indeed cost growth for this category and the amount of that cost growth is 13.15%. Our multiple regression model predicts the amount of cost growth in excess of what is computed by the database. With a predicted amount of cost growth of 70% we expect the cost estimator and program manager to be suspect of this predicted value and not rely these results. At this point the cost estimator should find alternate methods of predicting cost growth.

We continue by plugging the necessary predictor variables into the formulas for the *support* response as outlined in Figures 4.18 and 4.33 and find that our logistic model predicts that there is a 0.395 probability, or 39.5% chance, that the F-22 program will experience cost growth in the estimating cost variance category. Since 39.5% is below the 50% cut-off, this result is coded as a '0', thus, our formula predicts that the F-22 program will not experience cost growth in the *support* category. When compared to the actual results we find that this is indeed the case. Our Excel® database computes a negative percentage for this category (-4.8%) and, therefore, the program does not experience cost growth in this category. Since there is no cost growth for the *support* category, we do not use the multiple regression equation to predict the magnitude of cost growth. Our results for this scenario leave us confident that our models can and will accurately predict the presence of cost growth for both categories.

**Prior Research Comparison**

We would be negligent if we did not take this opportunity to discuss how our models compare to the models developed by Moore (2003) during his research in this area. Moore developed one logistic and one multiple regression model for all procurement dollars in the engineering and manufacturing development (EMD) phase of the acquisition program life cycle. Both of his models include data from the *quantity*, *schedule*, *engineering*, *estimating*, *support*, and *other* cost variance categories. In contrast, our models are built using the piecemeal approach started by Sipple (2002), and continued by Bielecki (2003) wherein each cost variance category has its own logistic and own multiple regression model.

By comparison, if we use Moore's logistic model on our F-22 data we find that his model estimates a 99.1% probability that cost growth will be present somewhere in the procurement appropriations of the EMD phase. We estimate that there will be cost growth in the *estimating* cost variance category, but not in the *support* category. To continue, Moore estimates the percentage of cost growth in the overall procurement appropriations of the EMD phase to be 51%, whereas we estimate cost growth to be 7.1% in the *estimating* category only.

Looking at the percentage of cost growth data for all categories as computed by our MS Excel® database for the F-22, we find that including the *quantity* category there is -28% cost growth, or no cost growth. If the *quantity* percentage is removed the overall cost growth for this program is approximately 9%.

Which is better? The answer to this question is ultimately left up to the program manager. We believe that using a logistic and multiple regression model for each cost variance category allows the cost estimator to be able to pinpoint cost growth down to a particular category. By knowing which category contains cost growth the cost estimator and program manager can focus on finding and fixing the cause specific to that category. This opportunity is not available with the overall approach used by Moore.

### Possible Follow-on Theses

The database used in this research is by no means complete. We promote further additions to this database in both programmatic data and potential predictor variables. The larger the database, the more useful it will become in other cost related research. Some possible related areas of research include:

- Allow data to build under the new A B C Acquisitions Milestone Phases, then expand the database and perform the same methodology.
- Explore a way to convert the old I II III Milestone phased data into the new A B C phased data.
- Take the *quantity* cost variance data out of Moore's models and see if there is a change.
- Identify programs that did not have significant overruns and evaluate their risk estimating methodology to see if there is a best methodology (Sipple, 2002:121).
- Create a program utilizing the CERs developed from this and other analyses (Sipple, 2002:121).
- Explore the applicability of our results to the Monte Carlo simulation technique of risk analysis (Sipple, 2002:121).
- Compare individual and overall RDT&E cost growth with individual and overall procurement cost growth. Identify trends, accuracy and root causes within each category (Bielecki, 2003:83).

### Recommendations

Our results further validate the ability of the two-step regression approach to accurately predict cost growth. This is no more evident than in our F-22 validation example above. Logistic regression saves us the trouble of having to gather the necessary data to predict cost growth for the *support* category.

This research continues to demonstrate the effectiveness of logistic regression and multiple regression to predict cost growth in large DoD programs. We believe the ability of these models to correctly predict the presence and amount of cost growth warrant their implementation for use across the DoD in estimating major weapons system program costs. We further submit that use of logistic regression has a wider place within the DoD community that is as yet unrecognized (Bielecki, 2003:82).

We also recommend that separate models be used for each cost growth category as opposed to an overall model. These category specific models will enable the cost estimator to keep his or her program manager better informed on the issue of cost growth by accurately detecting cost growth in each category.

# Appendix A

## Predictor Variables Removed From (Logistic) *% <u>Estimating</u>* Models

| Predictor Variables as Regressed on *Cost Variance - Procurement % Estimating* | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Individually** | **6 +** | **7+** | **9 +** | **38 +** | **46 +** | **48 +** | **51 +** | **77 +** | **78 +** | **81 +** | **Count** | **%** |
| | 1 | 1 | | | | 1 | 1 | | | 1 | 5 | 45% |
| 10 Space | | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 91% |
| 11 Sea | 11 | 11 | 11 | 11 | 11 | 11 | 11 | | 11 | 11 | 10 | 91% |
| 14 Missile | | | 14 | 14 | 14 | | | 14 | 14 | | 6 | 55% |
| 16 Munition | | | 16 | 16 | | 16 | 16 | | | 16 | 6 | 55% |
| 17 Land Vehicle | 17 | | | | 17 | | 17 | 17 | | 17 | 6 | 55% |
| 18 Space (RAND) | | | 18 | 18 | 18 | 18 | 18 | | | 18 | 7 | 64% |
| 19 Ship | 19 | 19 | 19 | | 19 | 19 | 19 | | 19 | 19 | 9 | 82% |
| 2 Total Quantity | | | 2 | 2 | 2 | | | | 2 | | 5 | 45% |
| | 20 | 20 | 20 | 20 | 20 | | | 20 | 20 | | 7 | 64% |
| | | | | | | | | | 21 | | 1 | 9% |
| 22 Svs>1 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 11 | 100% |
| | | | | | | | | | 23 | 23 | 2 | 18% |
| 25 Service = Navy only | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 11 | 100% |
| 26 Service = Joint | 26 | 26 | 26 | 26 | 26 | 26 | 26 | 26 | 26 | 26 | 11 | 100% |
| | | | | | 27 | | | 27 | 27 | | 3 | 27% |
| 28 Service = Marines only | 28 | 28 | 28 | 28 | 28 | 28 | 28 | 28 | 28 | 28 | 11 | 100% |
| 29 Service = AF only | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 11 | 100% |
| | | | | 3 | 3 | | | 3 | 3 | | 4 | 36% |
| 30 Lead Svc = Army | 30 | 30 | | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 10 | 91% |
| 31 Lead Svc = Navy | 31 | 31 | 31 | | 31 | 31 | 31 | 31 | 31 | 31 | 10 | 91% |
| 32 Lead Svc = DoD | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 11 | 100% |
| | | | | | | | 34 | | 34 | 34 | 3 | 27% |
| 35 N involvement | 35 | 35 | | 35 | 35 | 35 | 35 | 35 | 35 | 35 | 10 | 91% |
| 36 MC involvement | 36 | 36 | 36 | 36 | 36 | 36 | 36 | 36 | 36 | 36 | 11 | 100% |
| 37 AR involvement | 37 | 37 | | 37 | 37 | 37 | 37 | 37 | 37 | 37 | 10 | 91% |
| | | | 4 | 4 | | | | | | | 2 | 18% |
| 40 Boeing | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 11 | 100% |
| 41 Raytheon | 41 | 41 | 41 | 41 | 41 | 41 | 41 | 41 | 41 | 41 | 11 | 100% |
| 42 Litton | 42 | 42 | 42 | 42 | 42 | 42 | 42 | 42 | 42 | 42 | 11 | 100% |
| 43 General Dynamics | 43 | 43 | | 43 | 43 | 43 | 43 | 43 | 43 | 43 | 10 | 91% |
| 45 More than 1 Major Defense Contracto | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 11 | 100% |
| | 47 | 47 | | | | 47 | 47 | | | 47 | 5 | 45% |
| | | | | | | | 48 | | | 48 | 2 | 18% |
| | 49 | 49 | | | | 49 | 49 | | | 49 | 5 | 45% |
| | | | 5 | 5 | | | | | | | 2 | 18% |
| | 50 | 50 | | | | 50 | 50 | | | 50 | 5 | 45% |
| | | | | | | 51 | | | | 51 | 2 | 18% |
| 52 Length of R&D in Funding Yrs | 52 | 52 | 52 | | 52 | | 52 | 55 | | 52 | 8 | 73% |
| 53 R&D Funding Yr Maturity % | 53 | 53 | 53 | 53 | 53 | 53 | | | 53 | 53 | 9 | 82% |
| 54 Proc Funding Yr Maturity % | 54 | 54 | 54 | 54 | 54 | 54 | 54 | | | 54 | 9 | 82% |
| 55 Total Funding Yr Maturity % | 55 | 55 | 55 | 55 | 55 | 55 | 55 | | 55 | 55 | 10 | 91% |
| 56 Actual Length of EMD | 56 | 56 | 56 | 56 | 56 | 56 | 56 | 56 | 56 | 56 | 11 | 100% |
| 57 Maturity of EMD % | 57 | 57 | 57 | 57 | | | 57 | 57 | 57 | | 8 | 73% |
| | | | 58 | 58 | | 58 | | | 58 | | 4 | 36% |
| | | | 59 | 59 | 59 | | | | 59 | | 4 | 36% |
| | | 6 | | | | | | | | | 1 | 9% |
| 60 LRIP Qty Planned | 60 | 60 | 60 | 60 | 60 | 60 | 60 | | 60 | 60 | 10 | 91% |
| | | | 61 | 61 | | 61 | | | 61 | | 4 | 36% |
| | 62 | 62 | | | 62 | 62 | 62 | | | 62 | 6 | 55% |
| 63 Proc Funding before MS III? | 63 | 63 | | | | 63 | 63 | 63 | 63 | 63 | 8 | 73% |
| | | | | | | | | | 64 | | 1 | 9% |
| 66 Class - C | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 11 | 100% |
| | | | | | 68 | | | | | | 1 | 9% |
| 69 Versions Previous to SAR | 69 | | 69 | 69 | 69 | | | 69 | 69 | 69 | 8 | 73% |
| | 7 | | | | | | | | | | 1 | 9% |
| | | | 70 | | 70 | | | 70 | 70 | 70 | 5 | 45% |
| 71 Prototype? | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 11 | 100% |
| 72 Dem/Val Prototype? | 72 | 72 | 72 | 72 | 72 | 72 | 72 | 72 | 72 | 72 | 11 | 100% |
| 73 EMD Prototype? | 73 | 73 | 73 | 73 | 73 | 73 | 73 | 73 | 73 | 73 | 11 | 100% |
| | 74 | 74 | 74 | 75 | 74 | 74 | 74 | 74 | 74 | 74 | 10 | 91% |
| 75 Significant pre-EMD activity immedi | 75 | 75 | 75 | | 75 | 75 | 75 | 75 | 75 | 75 | 10 | 91% |
| 76 Program have a MS I? | 76 | 76 | 76 | 76 | 76 | 76 | 76 | 76 | 76 | 76 | 11 | 100% |
| | | | | | | | 79 | | | 79 | 2 | 18% |
| | | | 8 | | | | | | | | 1 | 9% |
| | 80 | 80 | | | | 80 | 80 | | | 80 | 5 | 45% |
| | | | | | | 81 | 81 | | | | 2 | 18% |
| 82 R&D Funding Yr Maturity % > 75%? | 82 | 82 | 82 | 82 | 82 | | | | 82 | 82 | 8 | 73% |
| 83 Proc Funding Yr Maturity % > 40%? | | | 83 | 83 | 83 | 83 | 83 | 83 | 83 | 83 | 9 | 82% |
| 84 Funding Yrs of R&D Complete < 9? | 84 | 84 | 84 | 84 | 84 | 84 | 84 | | | 84 | 9 | 82% |
| 85 Funding Yrs of Proc Complete < 5? | 85 | 85 | | 85 | 85 | 85 | 85 | | | 85 | 8 | 73% |

# Appendix A (cont.)

## Predictor Variables Removed From (Logistic) *% <u>Support</u>* Models

| Predictor Variables as Regressed on *Cost Variance - Procurement % Support* | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Individually** | **50 +** | **79 +** | **51 +** | **54 +** | **47 +** | **85 +** | **35 +** | **70 +** | **81 +** | **Count** | **%** |
| 1 Total Cost CY $M 2002 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 10 | 100% |
| 10 Space | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 100% |
| 11 Sea | 11 | | | 11 | 11 | 11 | 11 | 11 | 11 | 8 | 80% |
| 12 Electronic | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 12 | 10 | 100% |
| | 14 | 14 | 14 | | | | | | 14 | 4 | 40% |
| 15 Aircraft | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 10 | 100% |
| 16 Munition | 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 | 16 | 10 | 100% |
| 17 Land Vehicle | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 17 | 10 | 100% |
| | | | | | | | | 18 | 19 | 2 | 20% |
| 19 Ship | 19 | 19 | 19 | 19 | 19 | 19 | | 19 | | 8 | 80% |
| 2 Total Quantity | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 10 | 100% |
| 20 Other | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 10 | 100% |
| | | | | 21 | | | 21 | 21 | | 3 | 30% |
| 24 Svs>3 | 24 | 24 | 24 | 24 | 24 | 24 | 24 | 24 | 24 | 10 | 100% |
| 25 Service = Navy only | 25 | | 25 | 25 | 25 | 25 | 25 | 25 | 25 | 9 | 90% |
| | | | | 26 | | | 26 | 26 | | 3 | 30% |
| | | | | | | | 27 | | 27 | 2 | 20% |
| 28 Service = Marines only | 28 | 28 | 28 | 28 | 28 | 28 | 28 | 28 | 28 | 10 | 100% |
| 29 Service = AF only | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 29 | 10 | 100% |
| 30 Lead Svc = Army | | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 9 | 90% |
| 31 Lead Svc = Navy | | | | 31 | 31 | 31 | 31 | 31 | | 6 | 60% |
| 32 Lead Svc = DoD | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 32 | 10 | 100% |
| 33 Lead Svc = AF | 33 | 33 | 33 | 33 | 33 | 33 | 33 | 33 | 33 | 10 | 100% |
| 34 AF involvement | 34 | 34 | 34 | 34 | 34 | 34 | 34 | 34 | 34 | 10 | 100% |
| 36 MC involvement | 36 | 36 | 36 | 36 | 36 | 36 | 36 | 36 | 36 | 10 | 100% |
| 37 AR involvement | 37 | 37 | 37 | 37 | 37 | 37 | 37 | 37 | 37 | 10 | 100% |
| 38 Lockheed-Martin | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 10 | 100% |
| 39 Northrop Grumman | 39 | 39 | 39 | 39 | 39 | 39 | 39 | 39 | 39 | 10 | 100% |
| 4 Qty planned for R&D | 4 | 4 | | 4 | 4 | | 4 | 4 | 4 | 8 | 80% |
| 40 Boeing | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 10 | 100% |
| | | | | | | | 41 | | | 1 | 10% |
| 42 Litton | 42 | 42 | 42 | 42 | 42 | 42 | 42 | 42 | 42 | 10 | 100% |
| 43 General Dynamics | 43 | 43 | 43 | 43 | 43 | 43 | 43 | 43 | 43 | 10 | 100% |
| 44 No Major Defense Contractor | 44 | 44 | 44 | 44 | 44 | 44 | 44 | | 44 | 9 | 90% |
| 45 More than 1 Major Defense Contracto | 45 | | | 45 | 45 | 45 | | | 45 | 6 | 60% |
| | 47 | | | 47 | | 47 | | | 47 | 4 | 40% |
| | 48 | | 48 | | 48 | 48 | | | 48 | 5 | 50% |
| 5 Qty currently estimated for R&D | | | | | | | | | | 1 | 10% |
| | | | 50 | 50 | 50 | 50 | | | 50 | 5 | 50% |
| | 51 | 51 | | | | | | | 51 | 3 | 30% |
| | 52 | | 52 | 52 | 52 | 52 | | | 52 | 6 | 60% |
| 53 R&D Funding Yr Maturity % | 53 | | 53 | 53 | 53 | 53 | 53 | 53 | 53 | 9 | 90% |
| | 54 | | | | 54 | 54 | | | | 3 | 30% |
| 55 Total Funding Yr Maturity % | | | 55 | | | 55 | 55 | 55 | 55 | 6 | 60% |
| | 56 | 56 | | 56 | | 56 | | | | 4 | 40% |
| 57 Maturity of EMD % | | 57 | 57 | | | 57 | 57 | 57 | 57 | 7 | 70% |
| 58 Time from MSII to IOC (in months) | 58 | 58 | 58 | 58 | | 58 | 58 | 58 | 58 | 9 | 90% |
| 59 Maturity of EMD at IOC% | 59 | 59 | 59 | 59 | 59 | 59 | 59 | 59 | 59 | 10 | 100% |
| 60 LRIP Qty Planned | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 60 | 10 | 100% |
| 61 LRIP Qty Current Estimate | 61 | 61 | 61 | 61 | 61 | 61 | 61 | 61 | 61 | 10 | 100% |
| | 62 | 62 | 62 | 62 | | 62 | | | | 5 | 50% |
| 63 Proc Funding before MS III? | 63 | | 63 | 63 | 63 | 63 | 63 | 63 | 63 | 9 | 90% |
| 64 # Product variants in this SAR | | 64 | 64 | 64 | 64 | 64 | 64 | 64 | 64 | 9 | 90% |
| | 66 | | | 66 | 66 | | | | | 3 | 30% |
| 67 Class - U | 67 | 67 | 67 | 67 | | 67 | 67 | 67 | 67 | 9 | 90% |
| | | | | | | | | 68 | | 1 | 10% |
| | | 69 | | | | | 69 | 69 | | 3 | 30% |
| | 7 | | 7 | | 7 | 7 | | | 7 | 5 | 50% |
| 71 Prototype? | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 71 | 10 | 100% |
| 72 Dem/Val Prototype? | | | | | | | | | | 1 | 10% |
| | | | | | | | 73 | 73 | | 2 | 20% |
| | 74 | | | 74 | | | | | | 2 | 20% |
| 75 Significant pre-EMD activity immedi | 75 | | 75 | 75 | 75 | 75 | 75 | 75 | 75 | 9 | 90% |
| | | | | | | | 76 | | | 1 | 10% |
| | | 77 | | | | | | | | 1 | 10% |
| | | 78 | 78 | | | | | | | 2 | 20% |
| | | | 79 | | | | | | 79 | 2 | 20% |
| 8 Air | | | | | 8 | | 8 | 8 | 8 | 5 | 50% |
| | 80 | | 80 | 80 | 80 | 80 | | | 80 | 6 | 60% |
| | 81 | 81 | 81 | | 81 | | | | | 4 | 40% |
| 82 R&D Funding Yr Maturity % > 75%? | 82 | | | 82 | 82 | 82 | 82 | 82 | 82 | 9 | 90% |
| | | | | | 83 | 83 | | | | 2 | 20% |

## Appendix A (cont.)

## Predictor Variables Removed From (Multiple) *% Estimating* Models

| Predictor Variables as Regressed on *Ln CV - Proc % Estimating* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Individually | 62 + | 85 + | 51 + | 81 + | 60 + | 57 + | Count | % |
| 1 | 1 | 1 | 1 | 1 | 1 | | 6 | 86% |
| 3 | 3 | 3 | | 3 | 3 | 3 | 6 | 86% |
| 4 | 4 | 4 | | 4 | | 4 | 5 | 71% |
| 5 | | | | | | | 1 | 14% |
| 6 | | | | | | | 1 | 14% |
| | 7 | 7 | 7 | 7 | | 7 | 5 | 71% |
| 8 | 8 | 8 | 8 | 8 | 8 | 8 | 7 | 100% |
| 9 | 9 | 9 | 9 | | 9 | 9 | 6 | 86% |
| 10 | 10 | 10 | 10 | 10 | 10 | 10 | 7 | 100% |
| | | | | | 11 | 11 | 2 | 29% |
| 12 | 12 | 12 | 12 | 12 | 12 | 12 | 7 | 100% |
| | | | | 13 | 13 | 13 | 3 | 43% |
| 14 | 14 | 14 | 14 | 14 | 14 | 14 | 7 | 100% |
| | 15 | | | | 15 | 15 | 3 | 43% |
| | | 16 | | | | | 1 | 14% |
| 17 | 17 | 17 | 17 | 17 | 17 | 17 | 7 | 100% |
| 18 | 18 | 18 | 18 | 18 | 18 | 18 | 7 | 100% |
| 19 | 19 | 19 | | | 19 | 19 | 5 | 71% |
| 20 | | 20 | 20 | 20 | 21 | 20 | 6 | 86% |
| 21 | 21 | 21 | 21 | 21 | | 21 | 6 | 86% |
| 22 | 24 | 24 | 24 | 24 | 24 | 24 | 7 | 100% |
| 23 | | | | | | | 1 | 14% |
| 24 | | | | | | | 1 | 14% |
| 25 | | 25 | 25 | 25 | 25 | 25 | 6 | 86% |
| 26 | 26 | 26 | 26 | 26 | 26 | 26 | 7 | 100% |
| | | | | | 27 | 27 | 2 | 29% |
| 28 | 28 | | 28 | 28 | 28 | | 5 | 71% |
| | | | | 29 | 29 | | 2 | 29% |
| | | | | | 30 | 30 | 2 | 29% |
| 31 | 31 | 31 | 31 | | 31 | 31 | 6 | 86% |
| | 33 | | 33 | 33 | | | 3 | 43% |
| | 34 | 34 | 34 | 34 | | | 4 | 57% |
| 35 | | 35 | | | 35 | 35 | 4 | 57% |
| 36 | 36 | 36 | 36 | 36 | 36 | 36 | 7 | 100% |
| 38 | 38 | 38 | 38 | | | 38 | 5 | 71% |
| 39 | | 39 | 39 | 39 | 39 | | 5 | 71% |
| 40 | 40 | 40 | 40 | 40 | | 40 | 6 | 86% |
| 41 | 41 | 41 | 41 | 41 | 41 | 41 | 7 | 100% |
| 42 | 42 | 42 | 42 | 42 | 42 | 42 | 7 | 100% |
| | 43 | 43 | | | 43 | 43 | 4 | 57% |
| 44 | 44 | 44 | 44 | 44 | 44 | 44 | 7 | 100% |
| 45 | 45 | | 45 | 45 | 45 | 45 | 6 | 86% |
| | 46 | | 46 | | 46 | | 3 | 43% |
| 47 | 47 | 47 | 47 | 47 | 47 | 47 | 7 | 100% |
| | | 48 | | 48 | 48 | | 3 | 43% |
| 49 | | | | | | | 1 | 14% |
| | 50 | 50 | 50 | 50 | 50 | 50 | 6 | 86% |
| | | | | 51 | | | 1 | 14% |
| 52 | 52 | 52 | 52 | 52 | 52 | 52 | 7 | 100% |
| | | 53 | | | 53 | 53 | 3 | 43% |
| | 54 | 54 | | | 54 | 54 | 4 | 57% |
| 55 | 55 | 55 | | 55 | 55 | 55 | 6 | 86% |
| | | | | | 56 | | 1 | 14% |
| | 57 | 57 | | | 57 | | 3 | 43% |
| | | | 58 | 58 | 58 | | 3 | 43% |
| | 59 | 59 | 59 | | 59 | | 4 | 57% |
| | | | | | | 60 | 1 | 14% |
| | | | | 61 | 61 | 61 | 3 | 43% |
| | | | | | | 62 | 1 | 14% |
| | 63 | | 63 | 63 | 63 | | 4 | 57% |
| | 64 | 64 | 64 | 64 | | 64 | 5 | 71% |
| 65 | 65 | 65 | 65 | 65 | 65 | 65 | 7 | 100% |
| 66 | | | 66 | 66 | 66 | 66 | 5 | 71% |
| 67 | 67 | 67 | 67 | 67 | 67 | 67 | 7 | 100% |
| | | | 68 | | | | 1 | 14% |
| 69 | 69 | 69 | | 69 | 69 | 69 | 6 | 86% |
| 70 | 70 | 70 | 70 | 70 | 70 | 70 | 7 | 100% |
| 71 | 71 | 71 | 71 | 71 | 71 | 71 | 7 | 100% |
| 72 | | | | | | | 1 | 14% |
| | | | | | 73 | | 1 | 14% |
| 74 | 74 | 74 | | | | 74 | 4 | 57% |

# Appendix A (cont.)

## Predictor Variables Removed From (Multiple) *% **Support*** Models

| Predictor Variables as Regressed on *Ln CV - Proc % Support* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Individually | 19 + | 11 + | 26 + | 42 + | 64 + | 17 + | Count | % |
| 1 | 1 | | 1 | 1 | 1 | 1 | 6 | 86% |
| 2 | | | 2 | 2 | 2 | 2 | 5 | 71% |
| 3 | 3 | 3 | 3 | 3 | 3 | 3 | 7 | 100% |
| 4 | | | 4 | 4 | 4 | 4 | 5 | 71% |
| 7 | | 7 | 7 | 7 | | | 4 | 57% |
| | 8 | 8 | 8 | 8 | | | 4 | 57% |
| 9 | 9 | 9 | 9 | 9 | 9 | 9 | 7 | 100% |
| 10 | 10 | 10 | 10 | 10 | 10 | 10 | 7 | 100% |
| | | | | 11 | | | 1 | 14% |
| | | | 12 | | | | 1 | 14% |
| 13 | 13 | 13 | 13 | 13 | 13 | 13 | 7 | 100% |
| 14 | 14 | 14 | 14 | 14 | 14 | 14 | 7 | 100% |
| | | | | | 15 | | 1 | 14% |
| 16 | 16 | 16 | 16 | 16 | 16 | 16 | 7 | 100% |
| 18 | 18 | 18 | 18 | 18 | 18 | 18 | 7 | 100% |
| | | | | 19 | | | 1 | 14% |
| 20 | | | | | | | 1 | 14% |
| | | | 24 | | | 24 | 2 | 29% |
| 25 | | | 25 | | 25 | | 3 | 43% |
| 27 | 27 | 27 | 27 | 27 | | 27 | 6 | 86% |
| 28 | 28 | 28 | 28 | 28 | 28 | 28 | 7 | 100% |
| 29 | 29 | 29 | 29 | 29 | 29 | | 6 | 86% |
| | 30 | 30 | 30 | 30 | | 30 | 5 | 71% |
| 31 | 31 | 31 | 31 | 31 | 31 | 31 | 7 | 100% |
| 32 | | | | | | | 1 | 14% |
| 33 | | | 33 | 33 | | 33 | 4 | 57% |
| 34 | | | 34 | 34 | 34 | 34 | 5 | 71% |
| | 35 | 35 | 35 | 35 | | | 4 | 57% |
| | | | 36 | | | | 1 | 14% |
| 37 | 37 | 37 | 37 | 37 | 37 | 37 | 7 | 100% |
| 38 | 38 | 38 | 38 | 38 | 38 | 38 | 7 | 100% |
| 39 | 39 | 39 | 39 | 39 | 39 | 39 | 7 | 100% |
| 40 | 40 | 40 | 40 | 40 | 40 | 40 | 7 | 100% |
| 41 | 41 | 41 | 41 | 41 | 41 | 41 | 7 | 100% |
| | 42 | 42 | | | | | 2 | 29% |
| 43 | | | 43 | 43 | 43 | | 4 | 57% |
| 44 | 44 | 44 | 44 | 44 | 44 | 44 | 7 | 100% |
| 45 | 45 | 45 | 45 | 45 | 45 | 45 | 7 | 100% |
| 46 | 46 | 46 | 46 | 46 | 46 | 46 | 7 | 100% |
| | | | | | 47 | | 1 | 14% |
| | | | | | 48 | | 1 | 14% |
| | | | | | 50 | | 1 | 14% |
| | | | | | 51 | 51 | 2 | 29% |
| | | | | | 52 | | 1 | 14% |
| 53 | | | 53 | | 53 | 53 | 4 | 57% |
| | | | 54 | | 54 | 54 | 3 | 43% |
| 55 | | | 55 | | 55 | 55 | 4 | 57% |
| 56 | 56 | 56 | | 56 | 56 | 56 | 6 | 86% |
| 58 | 58 | 58 | | 58 | | 58 | 5 | 71% |
| 59 | 59 | 59 | 59 | 59 | 59 | 59 | 7 | 100% |
| 60 | 60 | 60 | 60 | 60 | 60 | 60 | 7 | 100% |
| 61 | 61 | 61 | 61 | 61 | 61 | 61 | 7 | 100% |
| | | | | | 62 | | 1 | 14% |
| 63 | 63 | 66 | 63 | 63 | 63 | 63 | 7 | 100% |
| | 65 | | | | | 65 | 2 | 29% |
| 66 | 66 | | | 66 | 66 | 66 | 5 | 71% |
| 67 | | | 67 | | 67 | 67 | 4 | 57% |
| 68 | 68 | 68 | 68 | 68 | 68 | 68 | 7 | 100% |
| 69 | 69 | 69 | | 69 | 69 | 69 | 6 | 86% |
| 70 | 70 | 70 | | 70 | 70 | 70 | 6 | 86% |
| 71 | 71 | 71 | 71 | 71 | 71 | 71 | 7 | 100% |
| 73 | 73 | 73 | 73 | 73 | | 73 | 6 | 86% |
| 74 | 74 | 74 | 74 | 74 | 74 | 74 | 7 | 100% |
| | 75 | 75 | 75 | 75 | | | 4 | 57% |
| 76 | 76 | 76 | 76 | 76 | | | 5 | 71% |
| 77 | | 77 | 77 | 77 | 77 | | 5 | 71% |
| 78 | 78 | 78 | | 78 | 78 | 78 | 6 | 86% |
| 79 | | 79 | 79 | 79 | 79 | 79 | 6 | 86% |
| | | | | | 80 | | 1 | 14% |
| | | | | | 81 | 81 | 2 | 29% |
| 82 | 82 | | 82 | | 82 | 82 | 5 | 71% |

# Appendix B

## Logistic Regression – <u>Full</u> Model – *Estimating* Response

**Nominal Logistic Fit for Estimating Cost Growth? Procurement**

**Whole Model Test**

| Model | -LogLikelihood | DF | ChiSquare | Prob>ChiSq |
|---|---|---|---|---|
| Difference | 30.925806 | 10 | 61.85161 | <.0001 |
| Full | 37.405685 | | | |
| Reduced | 68.331491 | | | |

| | |
|---|---|
| RSquare (U) | 0.4526 |
| Observations (or Sum Wgts) | 100 |

Converged by Gradient

**Lack Of Fit**

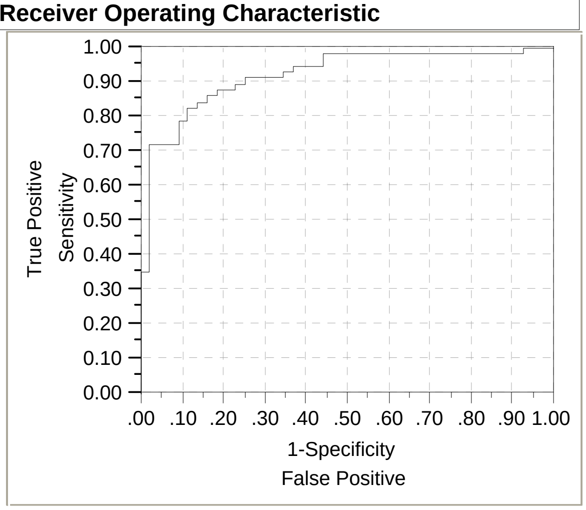
| Source | DF | -LogLikelihood | ChiSquare |
|---|---|---|---|
| Lack Of Fit | 89 | 37.405685 | 74.81137 |
| Saturated | 99 | 0.000000 | Prob>ChiSq |
| Fitted | 10 | 37.405685 | 0.8589 |

**Parameter Estimates**

| Term | Estimate | Std Error | ChiSquare | Prob>ChiSq |
|---|---|---|---|---|
| Intercept | 1.66953814 | 0.8953001 | 3.48 | 0.0622 |
| 7 ACAT 1? | -2.0385841 | 0.8208517 | 6.17 | 0.0130 |
| 77 LRIP Planned? | -1.8962074 | 0.6954414 | 7.43 | 0.0064 |
| 38 Lockheed-Martin | -1.9469863 | 0.8269847 | 5.54 | 0.0186 |
| 67 Class - U | 1.54974656 | 0.7168706 | 4.67 | 0.0306 |
| 9 Land | 1.22320931 | 0.6802552 | 3.23 | 0.0722 |
| 15 Aircraft | 3.01262894 | 1.0899479 | 7.64 | 0.0057 |
| 51 Length of Proc in Funding Yrs | -0.1413412 | 0.0557439 | 6.43 | 0.0112 |
| 44 No Major Defense Contractor | 1.93784885 | 0.8055776 | 5.79 | 0.0161 |
| 2 Total Quantity | 0.00003553 | 0.0000169 | 4.44 | 0.0351 |
| 39 Northrop Grumman | 2.77113551 | 1.2057433 | 5.28 | 0.0215 |

For log odds of 0/1

**Receiver Operating Characteristic**




Area Under Curve = 0.91922

## Appendix B (cont.)

## Logistic Regression – <u>Reduced</u> Model – *Estimating* Response

**Nominal Logistic Fit for Estimating Cost Growth? Procurement**

**Whole Model Test**

| Model | -LogLikelihood | DF | ChiSquare | Prob>ChiSq |
|---|---|---|---|---|
| Difference | 35.003898 | 9 | 70.0078 | <.0001 |
| Full | 22.259748 | | | |
| Reduced | 57.263647 | | | |

| | |
|---|---|
| RSquare (U) | 0.6113 |
| Observations (or Sum Wgts) | 86 |

Converged by Gradient

**Lack Of Fit**

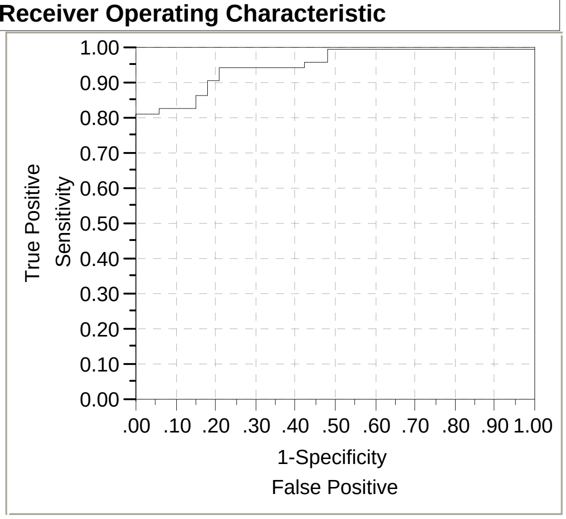
| Source | DF | -LogLikelihood | ChiSquare |
|---|---|---|---|
| Lack Of Fit | 76 | 22.259748 | 44.5195 |
| Saturated | 85 | 0.000000 | Prob>ChiSq |
| Fitted | 9 | 22.259748 | 0.9985 |

**Parameter Estimates**

| Term | Estimate | Std Error | ChiSquare | Prob>ChiSq |
|---|---|---|---|---|
| Intercept | 3.74251185 | 1.9561775 | 3.66 | 0.0557 |
| 7 ACAT 1? | -4.3368579 | 1.3497976 | 10.32 | 0.0013 |
| 77 LRIP Planned? | -2.4954635 | 1.1182264 | 4.98 | 0.0256 |
| 38 Lockheed-Martin | -2.8377295 | 1.2719104 | 4.98 | 0.0257 |
| 67 Class - U | 3.15286508 | 1.2494315 | 6.37 | 0.0116 |
| 15 Aircraft | 4.38455975 | 1.5281374 | 8.23 | 0.0041 |
| 44 No Major Defense Contractor | 4.15463156 | 1.352822 | 9.43 | 0.0021 |
| 39 Northrop Grumman | 5.14122691 | 1.9324404 | 7.08 | 0.0078 |
| 1 / Variable # 3 | 0.58771192 | 0.2620326 | 5.03 | 0.0249 |
| ln(Variable # 51) | -1.6495495 | 0.8215535 | 4.03 | 0.0447 |

For log odds of 0/1

**Receiver Operating Characteristic**




Area Under Curve = 0.95197

# Appendix C

## Logistic Regression – Full and Reduced Model – *Support* Response

### Nominal Logistic Fit for Support Growth? Procurement

#### Whole Model Test

| Model | -LogLikelihood | DF | ChiSquare | Prob>ChiSq |
|---|---|---|---|---|
| Difference | 30.443590 | 9 | 60.88718 | <.0001 |
| Full | 31.739508 | | | |
| Reduced | 62.183098 | | | |

| | |
|---|---|
| RSquare (U) | 0.4896 |
| Observations (or Sum Wgts) | 90 |

Converged by Gradient

#### Lack Of Fit

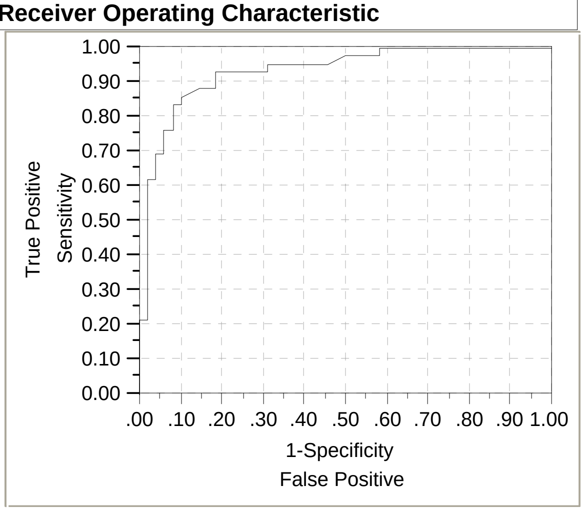
| Source | DF | -LogLikelihood | ChiSquare |
|---|---|---|---|
| Lack Of Fit | 69 | 27.920423 | 55.84085 |
| Saturated | 78 | 3.819085 | Prob>ChiSq |
| Fitted | 9 | 31.739508 | 0.8734 |

#### Parameter Estimates

| Term | Estimate | Std Error | ChiSquare | Prob>ChiSq |
|---|---|---|---|---|
| Intercept | 2.69155828 | 1.3671233 | 3.88 | 0.0490 |
| 50 Funding Yrs of Proc Completed | -0.2905103 | 0.0886537 | 10.74 | 0.0010 |
| 76 Program have a MS I? | 2.63525559 | 0.956811 | 7.59 | 0.0059 |
| 18 Space (RAND) | 6.58080015 | 2.7361572 | 5.78 | 0.0162 |
| 46 Fixed-Price EMD Contract? | 2.30444445 | 0.9818671 | 5.51 | 0.0189 |
| 66 Class - C | -6.2703953 | 1.9867388 | 9.96 | 0.0016 |
| 13 Helo | -3.2106087 | 1.8037027 | 3.17 | 0.0751 |
| 35 N involvement | 2.580247 | 1.0233947 | 6.36 | 0.0117 |
| 62 Proc Started based on Funding Yrs? | -2.8842632 | 1.4347345 | 4.04 | 0.0444 |
| 21 # of Svs | -0.8539183 | 0.448616 | 3.62 | 0.0570 |

For log odds of 0/1

#### Receiver Operating Characteristic




Area Under Curve = 0.93105

# Appendix D

## Ordinary Least Squares Regression – <u>Full</u> Model – *Estimating* Response

### Whole Model

#### Summary of Fit

| | |
|---|---|
| RSquare | 0.48983 |
| RSquare Adj | 0.431856 |
| Root Mean Square Error | 1.216127 |
| Mean of Response | -1.8356 |
| Observations (or Sum Wgts) | 50 |

#### Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Model | 5 | 62.48005 | 12.4960 | 8.4492 |
| Error | 44 | 65.07442 | 1.4790 | Prob > F |
| C. Total | 49 | 127.55447 | | <.0001 |

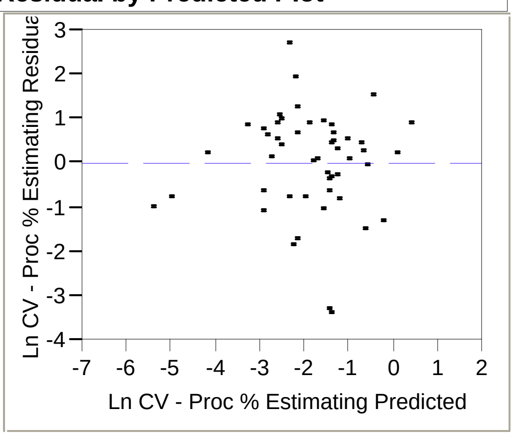
#### Lack Of Fit

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Lack Of Fit | 43 | 64.968942 | 1.51091 | 14.3237 |
| Pure Error | 1 | 0.105483 | 0.10548 | Prob > F |
| Total Error | 44 | 65.074424 | | 0.2071 |
| | | | | Max RSq |
| | | | | 0.9992 |

#### Parameter Estimates

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| | VIF |
|---|---|---|---|---|---|
| Intercept | -6.186931 | 0.780911 | -7.92 | <.0001 | . |
| 62 Proc Started based on Funding Yrs? | 2.0620946 | 0.509358 | 4.05 | 0.0002 | 1.0560508 |
| 58 Time from MSII to IOC (in months) | 0.0123957 | 0.005496 | 2.26 | 0.0291 | 1.0890115 |
| 77 LRIP Planned? | 0.8359001 | 0.356835 | 2.34 | 0.0237 | 1.0486372 |
| 37 AR involvement | 0.7832285 | 0.349319 | 2.24 | 0.0300 | 1.0049238 |
| 81 Length of Proc Funding > 11 yrs? | 0.9420611 | 0.390057 | 2.42 | 0.0199 | 1.0801587 |

#### Residual by Predicted Plot



## Appendix D (cont.)

## Ordinary Least Squares Regression – <u>Reduced</u> Model – *Estimating* Response

### Whole Model

#### Summary of Fit

| | |
|---|---|
| RSquare | 0.578606 |
| RSquare Adj | 0.538473 |
| Root Mean Square Error | 1.084489 |
| Mean of Response | -1.7421 |
| Observations (or Sum Wgts) | 47 |

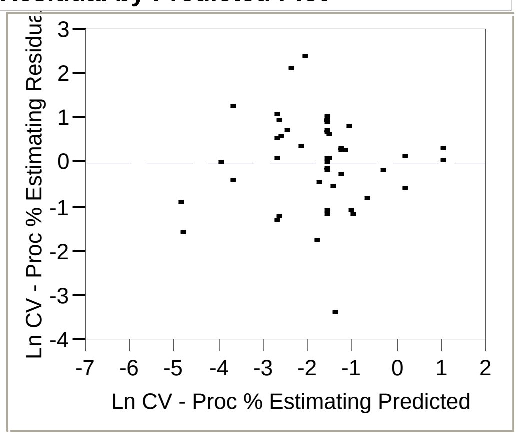
#### Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Model | 4 | 67.82556 | 16.9564 | 14.4173 |
| Error | 42 | 49.39689 | 1.1761 | Prob > F |
| C. Total | 46 | 117.22245 | | <.0001 |

#### Parameter Estimates

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| | VIF |
|---|---|---|---|---|---|
| Intercept | -4.803647 | 0.46946 | -10.23 | <.0001 | . |
| 62 Proc Started based on Funding Yrs? | 2.1386646 | 0.45508 | 4.70 | <.0001 | 1.0490251 |
| (Variable #58 * Variable # 73)^2 | 0.0000926 | 0.000026 | 3.62 | 0.0008 | 1.0705798 |
| 81 Length of Proc Funding > 11 yrs? | 1.1384232 | 0.356188 | 3.20 | 0.0026 | 1.0603576 |
| 2 Total Quantity | 0.0000186 | 0.000008 | 2.40 | 0.0207 | 1.0138227 |

#### Residual by Predicted Plot



# Appendix E

## Ordinary Least Squares Regression – Full Model – *Support* Response

### Whole Model

#### Summary of Fit

| | |
|---|---|
| RSquare | 0.472743 |
| RSquare Adj | 0.42481 |
| Root Mean Square Error | 1.289852 |
| Mean of Response | -3.16676 |
| Observations (or Sum Wgts) | 49 |

#### Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Model | 4 | 65.63480 | 16.4087 | 9.8627 |
| Error | 44 | 73.20355 | 1.6637 | Prob > F |
| C. Total | 48 | 138.83835 | | <.0001 |

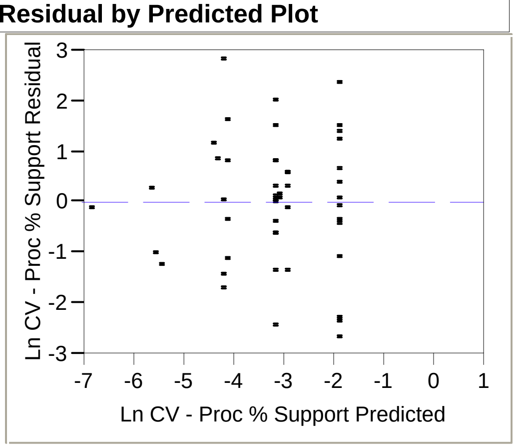
#### Lack Of Fit

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Lack Of Fit | 7 | 5.226751 | 0.74668 | 0.4064 |
| Pure Error | 37 | 67.976797 | 1.83721 | Prob > F |
| Total Error | 44 | 73.203548 | | 0.8922 |
| | | | | Max RSq |
| | | | | 0.5104 |

#### Parameter Estimates

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| | VIF |
|---|---|---|---|---|---|
| Intercept | -3.132919 | 0.31017 | -10.10 | <.0001 | . |
| 26 Service = Joint | -1.228773 | 0.48973 | -2.51 | 0.0159 | 1.1473671 |
| 19 Ship | -2.472786 | 0.714491 | -3.46 | 0.0012 | 1.1271722 |
| 12 Electronic | -1.04533 | 0.424996 | -2.46 | 0.0179 | 1.1299646 |
| 80 Length of R&D Funding > 12 yrs? | 1.2928744 | 0.378091 | 3.42 | 0.0014 | 1.0310846 |

#### Residual by Predicted Plot



## Appendix E (cont.)

## Ordinary Least Squares Regression – Reduced Model – *Support* Response

### Whole Model

#### Summary of Fit

| | |
|---|---|
| RSquare | 0.542253 |
| RSquare Adj | 0.492767 |
| Root Mean Square Error | 1.191581 |
| Mean of Response | -3.11065 |
| Observations (or Sum Wgts) | 42 |

#### Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Model | 4 | 62.23369 | 15.5584 | 10.9577 |
| Error | 37 | 52.53503 | 1.4199 | Prob > F |
| C. Total | 41 | 114.76872 | | <.0001 |

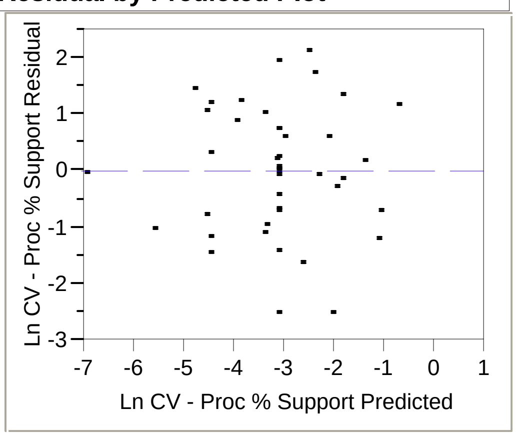
#### Lack Of Fit

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Lack Of Fit | 24 | 36.650887 | 1.52712 | 1.2498 |
| Pure Error | 13 | 15.884148 | 1.22186 | Prob > F |
| Total Error | 37 | 52.535035 | | 0.3455 |
| | | | | Max RSq |
| | | | | 0.8616 |

#### Parameter Estimates

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| | VIF |
|---|---|---|---|---|---|
| Intercept | -3.064493 | 0.284403 | -10.78 | <.0001 | . |
| 26 Service = Joint | -1.35354 | 0.513573 | -2.64 | 0.0122 | 1.2030299 |
| 19 Ship | -2.491327 | 0.777774 | -3.20 | 0.0028 | 1.1868498 |
| 12 Electronic | -1.37066 | 0.42391 | -3.23 | 0.0026 | 1.0848091 |
| Variable # 58 * Variable # 80 | 0.0148537 | 0.003674 | 4.04 | 0.0003 | 1.0709069 |

#### Residual by Predicted Plot



## Bibliography


Air Force Materiel Command. *AFMC Financial Management Handbook*. Wright-Patterson AFB OH: HQ AFMC, December 2001.

Bielecki, John V. *Estimating Engineering and Manufacturing Development Cost Risk Using Logistic and Multiple Regression*. MS thesis, AFIT/GAQ/ENC/03-02. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2003 (ADA413231).

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

Cancian, Mark. "Acquisition Reform: It's Not as Easy as it Seems," *Acquisition Review Quarterly,* (Summer 1995).

Department of Defense. *Department of Defense Directive:I Defense Acquisition System*. DoD 5000.1. Washington: GPO, May 2003

Department of Defense. *Department of Defense Instruction Operation of the Defense Acquisition System*. DoD 5000.2. Washington: GPO, May 2003

Department of Defense. "F-22 Fighter Program," *Department of Defense Appropriations Bill*, 2000. http://www.house.gov/appropriations_democrats/view_def_2.htm., Aug 2003.

Drezner, J. A., J. M. Jarvaise, R. W. Hess, P. G. Hough, and D. Norton. *An Analysis of Weapon System Cost Growth*. Santa Monica CA: RAND, 1993 (MR-291-AF).

Erwin, Sandra I. "Pentagon's Reform Plans Puzzle Experts." *National Defense Magazine.* November 2002. http://www.nationaldefensemagazine.org/article.cfm?Id=950., Aug 2003.

Grossman, Elaine. "Defense Budgets Include $7.4 Billion Boost For Realistic Costing." Early Bird Article. Inside the Pentagon, February 7, 2002.

Jensen, Arthur N. "Adjusted R_Squared", *College of Business Administration-California State University, Sacramento,* August 2003. http://www.csus.edu/indiv/j/jensena/mgmt105/adjustr2.htm., Nov 2003.

JMP® Version 4.0.4, (Academic), CD-ROM. Computer software. SAS Institute Inc., Cary NC, 2001.

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

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

Microsoft® Excel 2000. Version 9.0, IBM, 2.35K, disk. Computer software. Microsoft® Corporation, Redmond WA, 2000.

Moore, Gary W. *Estimating Procurement Cost Growth Using Logistic and Multiple Regression*. MS thesis, AFIT/GAQ/ENC/03-02. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2003 (ADA413830).

Myers, Dominique. "Acquisition Reform—Inside the Silver Bullet, A Comparative Analysis—JDAM versus F-22," *Acquisition Review Quarterly*, 313-332, (Fall 2002).

Neter, John, Michael H. Kutner, Christopher J. Nachtsheim, and William Wasserman. *Applied Linear Statistical Models*. Boston: McGraw-Hill, 1996.

Office of Management and Budget. *Historical Tables, Budget of the United States Government: Fiscal Year 2004*. Washington: GPO, 2003.

Sipple, Vincent P. *Estimating Engineering Cost Risk Using Logistic and Multiple Regression*. MS thesis, AFIT/GAQ/ENC/02-02. Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 2002 (ADA400576).

Sambur, Marvin Dr., Keys, Ronald E. *Statement Before the Subcommittee on Tactical and Land Forces House Armed Services Committee United States House of Representatives Concerning TACAIR Modernization.* Washington, April 2003.

Suddarth, Steven C. "Solving the Great Air Force Systems Irony," *Aerospace Power Journal*, (Spring 2002).

United States General Accounting Office. *F-22 Aircraft: Issues in Achieving Engineering and Manufacturing Development Goals.* Washington: GPO, March 1999 (GAO/NSIAD-99-55).

Yu, Alex Dr. "Multi-collinearity, Variance Inflation and Orthogonalization in Regression", *Arizona State University College of Education,* http://seamonkey.ed.asu.edu/~alex/computer/sas/collinear_VIF.html., January 2004.

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



| **1. REPORT DATE** *(DD-MM-YYYY)* | **2. REPORT TYPE** | **3. DATES COVERED** *(From – To)* |
|---|---|---|
| 10-03-2004 | **Master's Thesis** | Jun 2003 – Mar 2004 |

**4. TITLE AND SUBTITLE**
LOGISTIC AND MULTIPLE REGRESSION: A TWO-PRONGED APPROACH TO ACCURATELY ESTIMATE COST GROWTH IN MAJOR DoD WEAPON SYSTEMS

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**

Rossetti, Matthew B., First Lieutenant, USAF

**5d. PROJECT NUMBER**

**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 Way
WPAFB OH 45433-7765

**8. PERFORMING ORGANIZATION REPORT NUMBER**

AFIT/GCA/ENC/04-04

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
ASC/FMCE, AFMC
Attn: Mr. Michael J. Seibel
1865 4th St Rm 134 DSN: 986-5478
WPAFB OH 45433-7123 e-mail: Michael.Seibel@wpafb.af.mil

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION/AVAILABILITY STATEMENT**

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

This research confirms the usefulness of the logistic and multiple regression two-step procedure for assessing cost growth in major DoD weapon systems. We compile programmatic data from the Selected Acquisition Reports (SARs) between 1990 and 2002 for programs covering all defense departments. Our analysis concentrates on cost growth in the procurement appropriations of the Engineering and Manufacturing Development phase of acquisitions. We investigate the use of logistic regression in cost growth analysis to predict whether or not cost growth will occur in a program. If applicable, multiple regression is implemented to predict how much cost growth will occur. Our study focuses on the estimating and support SAR cost variance categories within the procurement appropriations. We study each of these categories individually for significant cost growth characteristics and develop predictive models for each.

**15. SUBJECT TERMS**
Logistic Regression, Multiple Regression, Cost Variance, Cost Growth, Selected Acquisition Report, SAR, DoD Cost Growth, Inferential Statistics, Cost Growth in DoD Acquisition Programs, Predicting Cost Growth

| **16. SECURITY CLASSIFICATION OF:** | | | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON** |
|---|---|---|---|---|---|
| **a. REPORT** | **b. ABSTRACT** | **c. THIS PAGE** | | | Dr. Edward D. White (ENC) |
| | | | | | **19b. TELEPHONE NUMBER** *(Include area code)* |
| **U** | **U** | **U** | **UU** | 114 | (937) 255-3636; e-mail: Edward.White@AFIT.edu |

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