3-2001

Relating Initial Budget to Program Growth with Rayleigh and Weibull Models

Eric J. Unger

RELATING INITIAL BUDGET TO PROGRAM GROWTH WITH RAYLEIGH AND WEIBULL MODELS

THESIS

Eric J. Unger, Captain, USAF

AFIT/GAQ/EN/01M-3

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GAQ/EN/01M-3

RELATING INITIAL BUDGET TO PROGRAM GROWTH WITH RAYLEIGH AND WEIBULL MODELS

THESIS

Presented to the Faculty of the Graduate School of

Engineering and Management of the 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 of in Acquisition Management

Eric J. Unger, B.S.

Captain, USAF

March 2001

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AFIT/GAQ/EN/01M-3

RELATING INITIAL BUDGET TO PROGRAM GROWTH WITH RAYLEIGH AND WEIBULL MODELS

Eric J. Unger, B.S.

Captain, USAF

Approved:

Mark A. Gallagher	8 Mar 01
Lt Col Mark A. Gallagher (Chairman)	Date
William K. Stockman	8 MAR 01
Lt Col William K. Stockman (Member)	Date
Edward D. White	8 Mar 01
Capt Edward D. White (Member)	Date

Acknowledgements

In no small way, the work presented in this thesis benefited by the dedication, perseverance, support, and sacrifice of many people around me. My deepest gratitude goes to LtCol Mark Gallagher, who showed me a path through a very difficult journey. To call him a mentor is not sufficient to describe the dimensions of his help. I would like to acknowledge Dr. Tony White for his persistent encouragement and mathematical guidance. (No trend is my friend.) To LtCol William Stockman, I can only say that I stole the pig. Finally, I would like to thank Paul Porter for his humorous views on the serious nature of this topic.

[redacted]

EJU

Table of Contents

Page

List of Figures

Figure | Page

List of Tables

AFIT/GAQ/EN/01M-3

Abstract

Previous research on completed defense R&D shows that contract expenditures can be fit well with a Rayleigh model. With fixed outlay rates, as prescribed by the OSD comptroller, the budget profile must have most of the funds in the early years to produce Rayleigh-distributed expenditures. R&D programs with more delayed funding profiles may also produce expenditures that a Rayleigh model fits through schedule slips and cost overruns. This research tests how well the initial funding profile produces Rayleigh-distributed expenditures that can be related to the program's final cost overrun and schedule slips. Based only on the initial budget profile, we explain 53.4% of cost overruns and 50.5% of percent schedule slip in 37 completed programs.

Research Value

Our research indicates to the services the importance of budgeting funds appropriately in the program to reduce the frequency of cost overruns and schedule slips. The military services' financial communities, OSD Program Analysis and Evaluation, and OSD comptroller may use the method to ensure that R&D programs are adequately funded sufficiently early to reduce program cost and schedule growth.

I. Introduction

Problem Statement

Military Research and Development (R&D) programs have difficulties with cost overruns and schedule slips. This study hypothesizes a correlation between initial R&D program's budget profiles and program results. Based on evidence that historical expenditures follow the Rayleigh model, this study estimates a relationship between the initial budget profile and both the length of schedule slip and the amount of program cost overrun.

Background

Research and Development (R&D) cost growth is a persistent problem for the Department of Defense (DoD). As programs mature, their funding profiles tend to change significantly from the initial conception. The profile change indicates that the initial budget profile was inappropriate for the program. Funding is inappropriate if a program has adequate funding, but it becomes available in the wrong years. Moreover, if initial funding is inadequate or inappropriate to actual program funding needs, schedule slips and cost overruns may result.

Belcher and Dukovich state that constrained funding is one of the factors that causes productivity inefficiencies, resulting in schedule slips and greater program cost (Belcher, 1999). Several studies support the notion that the Rayleigh probability distribution is appropriate for modeling R&D budget profiles. Watkins (1982) and Abernathy (1984) concluded that the Raleigh fit R&D program cost data. Lee, Hogue, and Gallagher (1997) show that program expenditures fit the Rayleigh distribution. In the study, they devise two techniques for using a Rayleigh model to develop an efficient funding profile, given a point estimate of total development cost and desired program completion time. They base their method on the concept that a budget that supports expenditures that follow the Rayleigh model is an efficient program. Efficient implies minimal program cost and schedule increases.

The Rayleigh model is a function of time *t*, scale *d*, and shape parameter *a*. The Rayleigh probability density function is $f(t) = 2ate^{-at^2}$. The scale parameter *d* relates to total cost of the program and the shape parameter a relates to the program completion time. Figure 1-1 depicts the effect of the shape parameter *a*, with *d* and *t* identical for each curve.

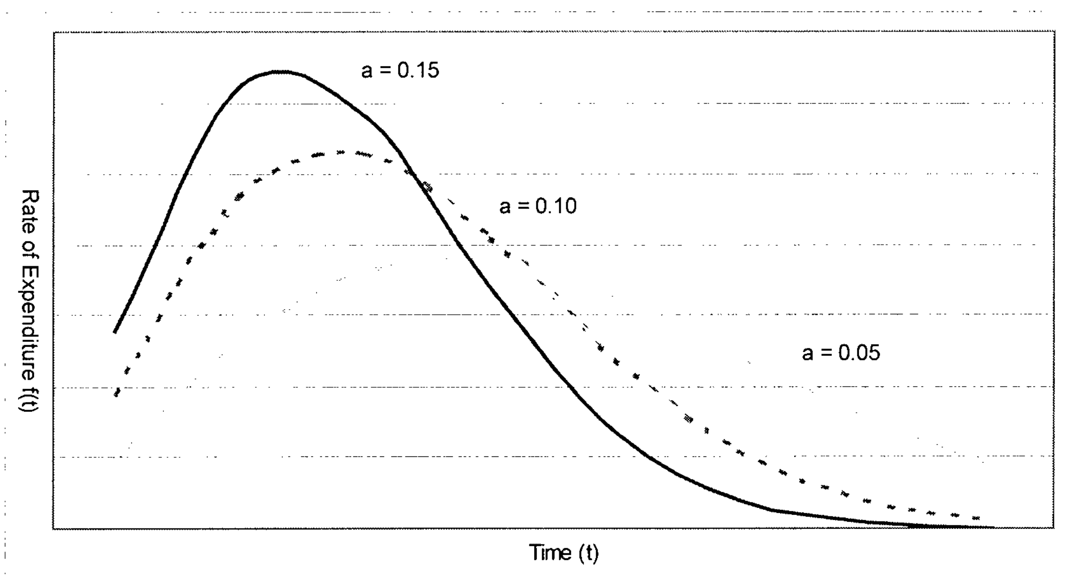


Figure 1-1. Rayleigh Probability Density Function

Lee, Hogue, and Gallagher (1997) develop budget profiles that support Rayleigh distributed expenditures. If the associated program expenditures follow a Rayleigh model, these budgets should minimize cost overruns and schedule slips. The focus of this research is to determine whether R&D budgeting contributes to cost overruns and delays by not having funding when needed.

Scope

Past research indicates that budgets supporting Rayleigh distributed expenditures are efficient. We test the impact of deviations from Rayleigh expenditures on program cost and schedule growth.

Research Approach

The change in shape of a budget profile and the resulting ending budget profile are dependent on many program factors. Program risk, program priority, annual budget, and program slips all play a role in defining how a budget profile develops over time. This study, however, seeks a correlation between a single indicator, the initial funding profile and the final expenditures. We approach confirming the relationship between initial funding profiles and Rayleigh expenditure profiles in two separate methods.

The first method assesses the cumulative expenditure distribution with a Rayleigh goodness of fit test. We evaluate the initial expenditures by determining whether there is sufficient statistical evidence to conclude they are Rayleigh using Anderson-Darling, Cramer-Von Mises, and Kolmolgorov goodness of fit statistics.

The second method determines Weibull parameters from the initial expenditures. The Weibull distribution, which is an extremely flexible distribution, includes the Rayleigh distribution.

While Abernathy (1984) and Lee, Hougue, and Gallagher (1997) showed that contractor costs follow the Rayleigh profile, contract costs represent a subset, albeit significant, of total program costs. In first step, we examine whether total program final expenditures also follow the Rayleigh profile. This relationship likely exists, as contractor costs are the major subset of total program costs. We use Selective Acquisition Reports (SAR) data, which provides funding data for large systems, to determine how well total program cost may be modeled by a Rayleigh distribution.

After establishing the quality of the Rayleigh model, we project final cost overruns and program slips from initial budget profiles for a variety of programs. We determine expected expenditure profiles given initial budget profiles from the SAR database with Office of the Secretary of Defense (OSD) Comptroller directed outlay rates. Also, we compare the Rayleigh expenditure profile to the expected expenditure profiles. The differences between the initial expected expenditure profiles and the Rayleigh derived profiles provide the information to explain schedule slips and cost overruns.

Research Benefits

This research evaluates historical program cost overruns and schedule slips based on the initial funding profile. Since this research assesses the impact of an initial R&D budget profile on program growth, the results are potentially useful to all levels of acquisition analysts and managers from the system program office to OSD Program Analysis and Evaluation (PA&E).

Chapter Two explores previous research which provides a foundation for the Rayleigh theory and subsequent corollaries. Chapter Three explains the methodology of evaluating historical DoD program data for its capacity to predict program growth and provides the results of the data analysis. Chapter Four provides conclusions and recommendations for further effort.

II. Literature Review and Concept Definition

Chapter Overview

Learning curve theory provides a typical approach to estimating certain components of production programs. The learning curve provides a mathematical relationship between the cost of the component and the quantity of components being produced. Generic learning curve theory states that as the quantity of components doubles, the cost per component decreases by a constant percentage. Along with providing a suitable estimate of cost, this approach can also provide an idea of when funding is required. Unfortunately, a standard analogous approach for estimating Research and Development (R&D) programs does not currently exist.

Establishing an appropriate funding profile for development programs is critical to their financial success. Inefficiently funded programs can either consume funding needed elsewhere, or cause program management crises in the form of funding shortfalls. In the absence of a standard model to establish development program costs, theorists developed mathematical models to derive the appropriate development funding profiles.

This chapter presents several mathematical models used to estimate development program effort. These models include the Beta model, the Sech-Squared Model, and the Rayleigh model. As the following effort is based on Rayleigh theory, the Rayleigh model will be explored in depth. Finally, we establish a relationship between the Rayleigh model and the Weibull distribution.

Each of the models provides an approach for modeling development program expenditures. Each of the models approximates constant years expenditures by applying a scaling factor, total expenditures in constant dollars, to the distribution. Therefore, an initial discussion of converting budget profiles to expenditures is necessary.

Conversion from Budgets to Expenditures

Multi-year appropriations and inflation affect the computation of DoD budgets. DoD must spend at least a certain percentage of the money allocated for R&D programs in a particular year, as prescribed by the Office of the Secretary of Defense Comptroller OSD(C). This percentage is called an outlay rate. Inflation also has a significant effect on the calculation of funding needed for outyear budgets. The combined effect of outlay rates and inflation implies that data in the form of budget profiles must be converted to expenditures in constant dollars to match the models. To reflect the expenditure amounts devoid of the effects of inflation, then year budget dollars need to be converted to constant year expenditures. To accomplish this conversion we use

$$O_i = B_i S_1 + B_{i-1} S_2 + B_{i-2} S_3 + \dots + B_{i-J} S_J \qquad (1)$$

where O_i represents the expenditure for year i in then year dollars, B_i the initial authorized budget amount, Si, the percentage outlay rate for a given year in the outlay profile, and J the number of years in the outlay profile. Since the intent is to outlay the entire budgeted amount,

$$\sum S_J \leq 1$$

Outlay rates may not sum to one, since DoD historically does not outlay its entire budget.

Since the outlays in (1) are still in then year dollars, they must be converted to constant year dollars to be properly used in the expenditure models. To remove inflation from each O_i, the annual expenditure must de deflated. Given the appropriate inflation factor,

$$O_i^* = O_i / c_i \tag{2}$$

where O_i^* is the expenditure for ith year in base-year dollars, c_i is the inflation factor of the ith year, and O_i is the expenditures for ith year in then-year dollars. The constant-year dollar expenditures are necessary for the models since they theoretically account for development effort, which inflation does not affect.

Using RDT&E budget data from the Air Force's E-3 Airborne Warning and Control System (AWACS) Radar System Improvement Program (RSIP), we show how to convert from the budget profile to outlays. Given the budget data, shown in the top row of Table 2-1, we can create an outlay profile in constant year dollars. We show then year dollars as TY$M and constant year dollars as CY$00M.

Table 2-1. Example Budget Conversion to Outlays

Fiscal Year	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Budget (TY$M)	44.2	63.7	68.8	116.3	72.1	30.9	0.0	0.0	0.0	0.0	0.0	0.0
Outlays (TY$M)	22.8	49.0	62.0	90.9	86.8	53.2	20.4	5.3	1.8	0.5	0.1	0.0
Expenditures (CY$00M)	29.0	59.9	72.6	103.6	96.3	57.8	21.7	5.6	1.8	0.5	0.1	0.0

Since RDT&E appropriations outlay over a period of six years we use (1) to create an outlay profile in CY$00M. Table 2-2 shows the Air Force outlay profile for Research and Development funds. The outlay process moves a portion of each year's

Table 2-2. Air Force RDT&E Outlay Rates

	Yr1	Yr2	Yr3	Yr4	Yr5	Yr6
AF Outlay Rates	0.516	0.366	0.071	0.027	0.008	0.001

budget dollars into later years to be expended. Thus, budget dollars spend in later fiscal years must reflect additional inflation. To determine the appropriate budget, we convert the constant year profile using (2) to a then year profile using the raw inflation index shown in Table 2-3.

Table 2-3. Air Force CY00 RDT&E Raw Inflation Index

	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
AF Inflation	0.787	0.818	0.854	0.877	0.901	0.919	0.937	0.955	0.975	0.982	0.990	1.000

The Beta Model

The Beta distribution, with its innate flexibility, can empirically fit manpower patterns, the analog for fitting development expenditures. The probability density function for the Beta is

$$\frac{dW(t)}{dt} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} t^{a-1} \bullet (1-t)^{b-1}, 0 < t < 1$$

where $\frac{dW(t)}{dt}$ represents that rate at which development problems are solved, $\Gamma(\bullet)$ is the gamma function, $a \geq 0$, $b \geq 0$, and t is time (Jarvis, 1999). However, there is little justification for its parameter inputs (Jarvis, 1999:8). In other words, the parameters do not relate to tangible constructs. With no method of estimating the model parameters a and b prior to program completion exists, the beta model can fit the program expenditure pattern only after completion of the program. Therefore, the beta model provides limited predictive ability.

The Sech-Squared Model

Like the Beta model, the Sech-Squared model focuses on the rate at which development problems are solved. The functional form of the Sech-Squared model is

$$f(t) = \frac{1}{4} \bullet \sec h^2 \left(\frac{(at + c)}{2} \right)$$

where a and c are parameters selected to fit expenditures and t is time.

Similar to the normal distribution, the traits of infinite tails and symmetry about a maximum characterize the Sech-Squared model (Parr, 1980:50). One of the primary problems of the Sech-Squared model derives from its infinite tail. Program initiation is particularly problematic to define. Parr defends the notion of an infinite tail in "negative" time (before zero on the time scale), by stating that practical considerations will establish the program start time (Parr, 1980:50). In practice, the lack of a defined program start could have significant impact on the predictive ability of the model. Compared to the Rayleigh, which starts at time zero and is skewed right with an infinite tail, the Sech-Squared model forces an amount of area under the curve before time zero.

Every model is an abstraction of reality. Significant considerations in building a model are its ease of use and its ability to supply useful results. While the Beta model and the Sech-Squared model can be used to predict expenditures, each is flawed in a way that makes its usefulness limited. While each model has benefits, the Rayleigh model demonstrates the greater functionality.

The Rayleigh Model

Norden related the pattern of manpower buildup and phase-out on development efforts to a family of mathematical functions, called the Rayleigh distribution. Norden developed the Rayleigh model based on skill level linearly increasing with time (Norden, 1970:122). Since the skill level is linear with respect to time, the rate of learning is constant. The rate of progress is proportional to the available skill level and the problems remaining to be solved (Jarvis, 1999:9).

Linear skills acquisition causes the initial ramp-up in the model. The Rayleigh model peak rate of expenditures is located at the 39th percentile in time. Since the rate of work performed is proportional to the work remaining, an exhaustion of the work causes the exponential tail. (Jarvis, 1999:13).

Putnam related the Rayleigh model to software development (Putnam, 1978:33). In testing the Rayleigh model against budgetary data for 50 Computer System Command systems, Putnam found that the Rayleigh approximated the software development well. In fact, Putnam checked the Rayleigh model against an additional 150 systems with good results. Putnam asserts that the Rayleigh model applies to software development because many software programs follow the characteristic growth to a peak, then exponential fall-off of the Rayleigh (Putnam, 1978:347).

Other researches have continued to apply the Rayleigh model to a wider variety of situations, most notably DoD contracts. Watkins (1982) applied the Rayleigh model to Cost/Schedule Status Report data on 30 DoD contracts. Watkins found statistically significant relationships for Rayleigh model between earned values inputs and total contract cost (Watkins, 1982:79). Later, Abernathy (1984) applied the Rayleigh model to

21 historical Navy contracts. He was able to adequately estimate the Rayleigh parameters to fit the contract data (Abernathy, 1984:37). While Abernathy did not meet his intended goal of using the Rayleigh model as a predictive tool, he established the model as having value for modeling DoD contract data. Lee, Hogue, and Hoffman (Lee, 1993) continued the application of the Rayleigh model to defense acquisition contracts. Their research showed that the Rayleigh model fit actual outlays of 20 defense development contracts.

An important aspect of the research completed on defense contract data is that it demonstrated that a scaled Rayleigh cumulative density function models development expenditures (Lee, 1997:16). Based on this assertion, Lee, Hogue, and Gallagher spread an R&D point estimate to a budget profile. Their contribution describes a method for developing a realistic budget profile, based on the Rayleigh model of expenditures (Lee, 1997). Gallagher and Lee (1996) subsequently develop and test a method of determining final program cost and completion time based on Rayleigh expenditures.

In summary, the mathematical relationship described in Norden's initial research provides a basis for many following efforts. Putnam applied the Rayleigh model to software development. Watkins and Abernathy substantiate the Rayleigh model's usefulness in modeling defense program expenditures. Finally, Lee, Hogue and Gallagher (1997) and Gallagher and Lee (1996) provide useful tools for defense development budget and schedule estimating.

Rayleigh Distribution. The Rayleigh model is a subset of the Weibull distribution, W(t) = $1 - e^{-\left(\frac{t}{\delta}\right)^{\beta}}$, with shape parameter $\beta = 2$. The concept that development projects involve solving a fixed number of problems, with linear learning, provides the foundation for the Rayleigh model. The rate at which problems are solved is proportional to skill level and proportion of unsolved problems remaining.

$$\frac{dW(t)}{dt} = p(t)[1 - W(t)]$$

where $p(t)$ is the skills level and $W(t)$ is the number of unsolved problems. Integration provides the total work completed in the following form

$$w(t) = 1 - e^{-\int_{\tau=0}^{\tau=t} p(\tau)d(\tau)}$$

Let $p(t) = ct$, where c represents linear learning

$$w(t) = 1 - e^{-0.5ct^2 \Big|_{\tau=0}^{\tau=t}}$$

$$= 1 - e^{-0.5ct^2}$$

$$= 1 - e^{-at^2}$$

with $a = 1/2c$, which gives the Rayleigh cumulative distribution function. The Rayleigh Rayleigh cumulative distribution function is generally presented in the functional form

$$F(t) = 1 - e^{-at^2} \tag{3}$$

Figure 2-1 shows the Rayleigh cumulative distribution function with $a = 0.94$. When scaled to a program's total cost, the cumulative distribution function illustrates cumulative expenditures over time.

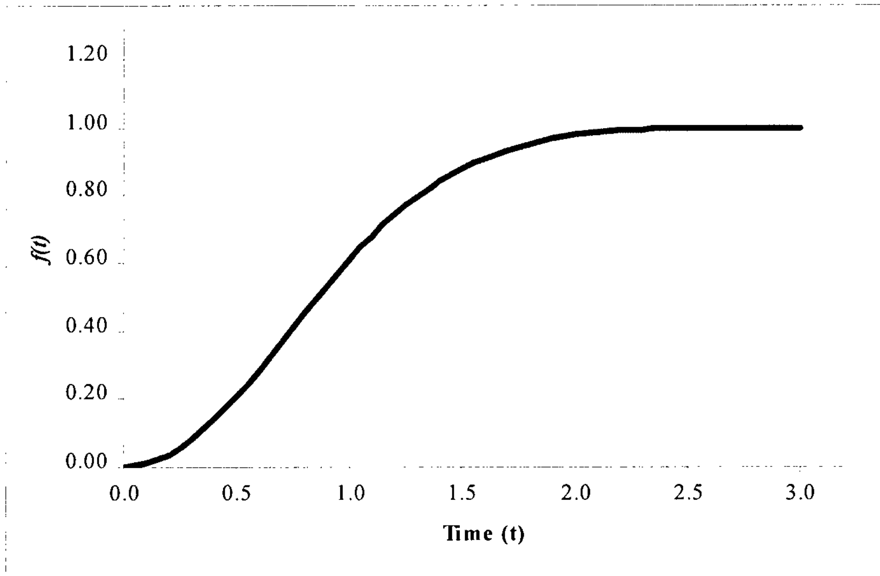


Figure 2-1. Rayleigh Cumulative Distribution Function

Rayleigh Probability Density Function. The derivative of the cumulative distribution (3) gives the probability density function. The Rayleigh probability density function is

$$f(t) = 2ate^{-at^2} \tag{4}$$

The scaled Rayleigh probability density function provides the rate of program expenditure for a given time. Figure 2-2 illustrates (4) for a value of a = 0.94.

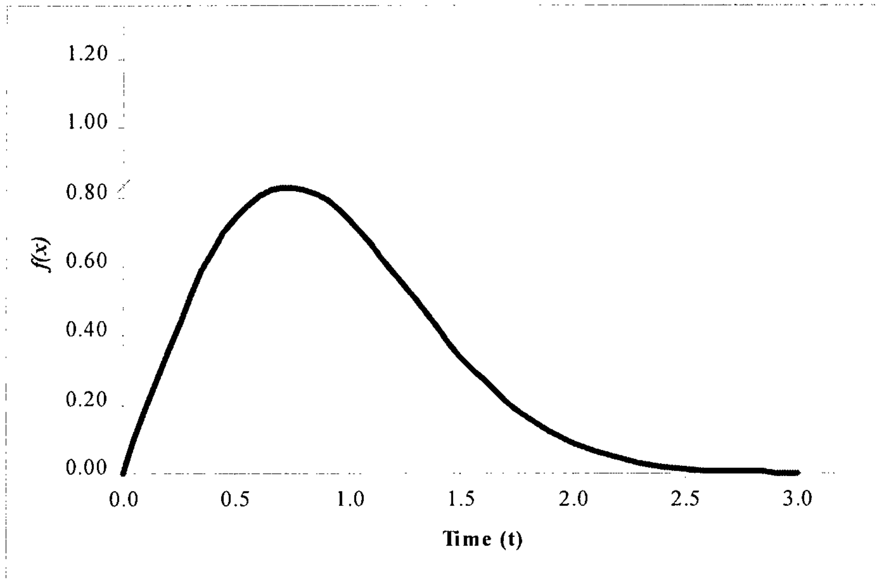


Figure 2-2. Rayleigh Probability Distribution Function

Infinite Tail. While the Rayleigh distribution's infinite tail is usually viewed as a flaw of the model, closer examination of the issue dismisses the problem. The normal distribution, the most commonly used distribution, shares the same characteristic of infinite tails. However, statisticians do not view this as a serious problem. For example, weight is often modeled with a normal distribution. Although this implies that a probability of negative weight and infinite weight exists, this is overlooked because of the incredible usefulness of the model. Furthermore, the infinite tail problem can be avoided through definition. While the tail stretches out to infinity, implying infinite program development time in the context of our intended use, Lee, Hogue, and Gallagher define program completion at 97% of the total program (Lee, 1997:6).

Rayleigh Parameters. Both the cumulative distribution function shown in Figure 2-1 and the probability density function in Figure 2-2 use the shape parameter a to determine the curve. The a parameter determines the steepness of the cumulative distribution function. Figure 2-3 depicts the effect of a change in the shape parameter *a* on the probability density function. Higher values of *a* create a more compressed curve, with an earlier peak time. This situation is analogous to programs with increased funding earlier in the program, producing earlier program completion.

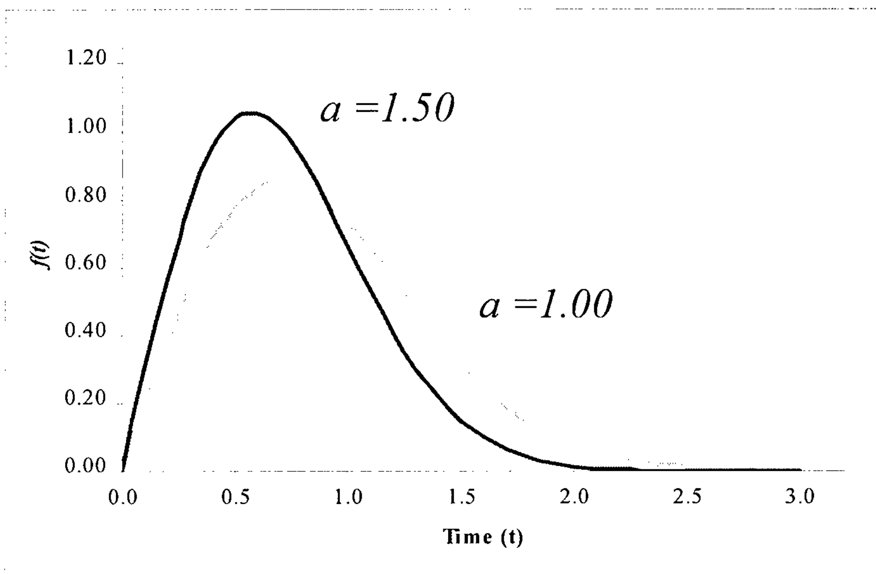


Figure 2-3. Effect of Rayleigh shape *a* change

The Rayleigh distribution is scaled from its value of [0,1] by multiplying by total program expenditures, *d*. Thus expenditures are defined as a function of time:

$$E(t) = d\,F(t) = d(1 - e^{-at^2}) \tag{5}$$

where *d* is the total cost of the program and the *t* is the time from inception of the R&D effort to completion (Lee, 1997:31).

<u>Parameter Determination</u>. Different methods can be used to derive the Rayleigh function parameters. The data available generally dictates what method must be used. The two following methods are characterized by the availability of program data to

determine the key time parameter a and cost parameter d. The remaining function parameters can be derived based on either the time that the expenditure rate is a maximum or the time that the expenditures are complete is known.

For an ongoing program, Lee, Hogue, and Gallagher indicate that if the time of the peak rate of expenditures, if known, provides a basis for estimating the Rayleigh parameter (Lee, 1997:6). Program peak time is reasonable easy to specify for certain types of programs. For example, peak time is normally associated with first flight on aircraft development programs (Lee, 1997:6) and full operational capability of software programs (Putnam, 1978:35).

From the Rayleigh probability density function in (4), the time of peak time may be determined by finding the maximum of the function. Taking the derivative of (4) produces

$$f'(t) = 2a\, e^{-at^2} + 2at(-2at)e^{-at^2}$$

Setting equal to 0 to find the maximum and designating t as t_p gives

$$0 = (2a - 4a^2 t_p^2)e^{-at^2}$$

Since e^{-at^2} is always positive, the coefficient of the exponential term must equal zero. The result is

$$t_p = \frac{1}{\sqrt{2a}} \text{ or equivalently } a = \frac{1}{2t_p^2}$$

We can derive the shape parameter a from the peak time, t_p, or vise versa since this is an one-to-one function.

In the absence of information about the programs peak time, an alternate method for deriving a is available. Often program mangers do not know when peak rate of expenditures will occur. The final time method is an option to predict peak time, when program final time is known. Lee, Hogue, and Gallagher define the final time, t_f, of a development project as the time when 97% of the cumulative Research and Development (R&D) constant year expenditures has been reached (Lee, 1997:32; Gallagher, 1996:52). For this analysis, final time t_f is defined as 0.97d, or the point at which the program is 97% complete. Setting the total expenditures from (5) equal to the defined final time gives

$$d(1 - e^{-at_f^2}) = 0.97d$$

Solving for t_f gives

$$t_f = \sqrt{\frac{\ln(0.03)}{-a}}$$

Since $a = \frac{1}{2}t_p^{-2}$, we can relate final time t_f to peak time t_p.

$$t_f = \sqrt{\frac{\ln(0.03)}{-0.5t_p^{-2}}}$$

$$t_f = t_p\sqrt{\ln(0.03)(-2)} = 2.65t_p$$

This relationship also equates to the constant that the peak expenditure rate occurs at 37% of total program time.

While the Rayleigh distribution has demonstrated many desirable characteristics, including its ability to model development expenditures, it proves inflexible for a broad assortment of development programs. The Rayleigh function forces a proportionate tail using the peak expenditure point as the start. However, programs where a proportionate tail is not derived from the point of peak expenditures exist. For example, a program may have a peak expenditure followed by a very short tail. The Rayleigh distribution is not useful for providing an accurate model, in this case. To better accommodate programs of a variety of shapes, the Weibull distribution can act as a surrogate for the Rayleigh distribution.

The Weibull Distribution

The Weibull distribution provides greater flexibility in modeling program expenditures than the Rayleigh model. While the Rayleigh model assumes program funding and manpower start with the inception of the program, many programs do not met this assumption. Some programs begin with one or more years of insignificant funding, while some programs may have large initial funding that tapers off. The three-parameter Weibull distribution handles either scenario. The Weibull accommodates insignificant initial funding with a location parameter. The Weibull, with a flexible shape parameter, accommodates other expenditure profiles not approximated well by the Rayleigh, such as most funding delays.

We derive the Weibull cost model from the assumption that the rate at which work is completed is a function of performance and remaining work. Define the percent work remaining as $w(t)$ and performance as $p(t)$, both at time t. Then

$$\frac{dw(t)}{dt} = p(t)[1 - w(t)],$$

which may be solved for $w(t)$. Let $z(t) = 1 - w(t)$, so $\frac{dz(t)}{dt} = \frac{-dw(t)}{dt}$ and

$$\frac{dz(t)}{dt} = -p(t)z(t).$$

Integrating, we obtain

$$\ln(z(t)) = -\int_{\tau=0}^{\tau=t} p(\tau)d\tau .$$

We evaluate both sides to the power of the base of the natural logarithm,

$$z(t) = e^{-\int_{\tau=0}^{\tau=t} p(\tau)d\tau} .$$

We substitute back in percent of work to obtain

$$w(t)=1-e^{-\int_{\tau=0}^{\tau=t} p(\tau)d\tau}$$

We define performance on the program for any given time as a constant multiplied by time to constant power, $p(t) = ct^b$. Since $\int_{\tau=0}^{\tau=t} p(\tau)d\tau = \int_{\tau=0}^{\tau=t} c\tau^b d\tau = \frac{c}{b+1}t^{b+1}$,

$$w(t)=1-e^{-\frac{c}{b+1}t^{b+1}} .$$

With linear growth in performance over time, $b = 1$ and $a = c/2$, we obtain the Rayleigh cumulative distribution function shown in (1). If performance improves with time to a power other than one, we have derived percent work complete according to a Weibull cumulative distribution function.

The Weibull cumulative distribution function is generally written as

$$F(t) = 1 - e^{-\left(\frac{t-\gamma}{\delta}\right)^{\beta}} \tag{6}$$

while the Weibull probability density function is written as

$$f(t) = \frac{\beta}{\delta}\left(\frac{t-\gamma}{\delta}\right)^{\beta-1} e^{-\left(\frac{t-\gamma}{\delta}\right)^{\beta}}$$

where t = time, γ = location, δ = scale, and β = shape (Hines, 1980:165). The shape parameter $\beta = b + 1 > 0$ and the scale parameter $\delta = (b+1/a)^{1/(b+1)} > 0$. For $\beta = 2$, the Weibull is the Rayleigh distribution.

When the Weibull location parameter $\gamma = 0$ and the Weibull shape parameter β = 2 , (3) and (6) are equivalent. The Rayleigh scale parameter in (3), a, is equal to $1/\delta^2$ in (6). The Weibull shape parameter allows skills acquisition at other than linear rates. Also, the Weibull time of peak expenditures does not fix the completion time of the program.

We represent the Weibull cost model as the cumulative distribution function in (6) multiplied by a cost scalar d with

$$E(t) = d\left[1 - e^{-\left(\frac{t-\gamma}{\delta}\right)^{\beta}}\right] \tag{7}$$

For given shape, scale, and location parameters, the Weibull cost model gives the cumulative cost of a program at specified time t.

Weibull Parameters. The flexibility of the Weibull distribution allows it to approximate a variety of different distributions. Changing the Weibull parameters produce noteworthy changes in the distribution. Figure 2-4 details the effect of changing the Weibull scale parameter δ while holding the other parameters constant. Increasing δ from 1.0 to 1.3, elongates the program. A scale increase implies that the program will take longer to accomplish.

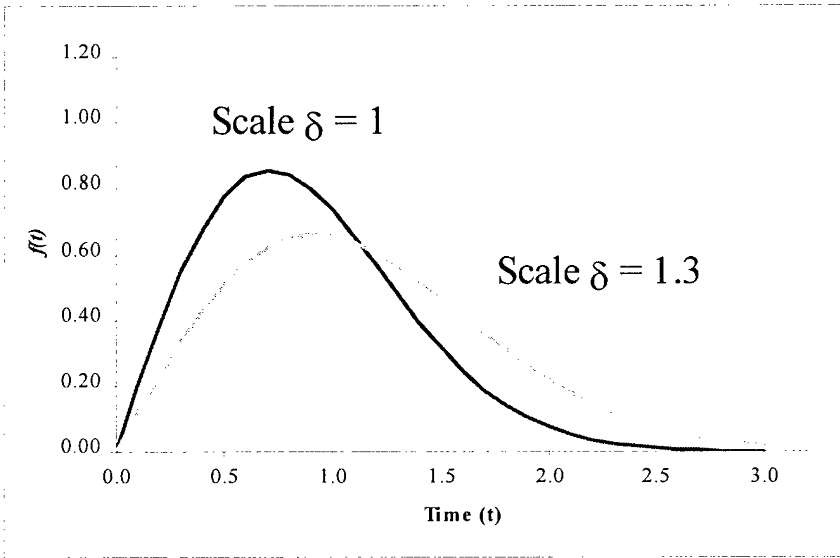


Figure 2-4. Effect of Weibull scale δ change, with $\beta = 2$, $\gamma=0$

The Weibull shape parameter is interesting because it has the greatest affect on the overall distribution appearance. Different shape parameter values allow the Weibull distribution to approximate other distributions. As noted earlier, a Weibull with a shape parameter $\beta = 2.0$ is a Rayleigh distribution. The Weibull with $\beta = 1$ is an exponential distribution. For $\beta = 3.4$, the Weibull approximates a normal distribution.

In programmatic terms, the shape parameter determines the location of peak spending for the program, for given location and scale parameters. This aspect of the model is important, since peak spending corresponds to a significant development event, such as first flight. Figure 2-5 shows the effect of changing the Weibull shape parameter β, holding the other parameters constant.

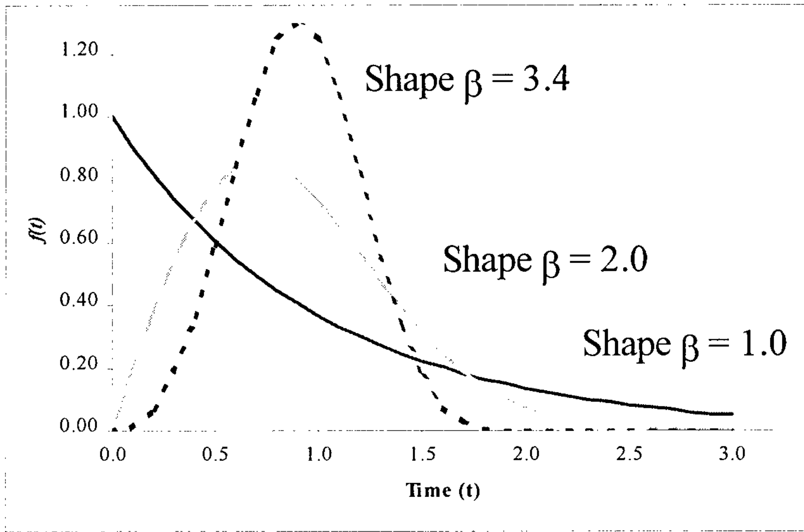


Figure 2-5. Effect of Weibull shape β change, with $\delta = 1$, $\gamma = 0$

The location parameter γ allows the Weibull distribution to shift along the x-axis. In certain program situations, the first several budget years contain very little funding. For program data, the location parameter identifies the length of time before the program significantly started. Figure 2-6 depicts the effect of changing the location parameter, with all other parameters held constant.

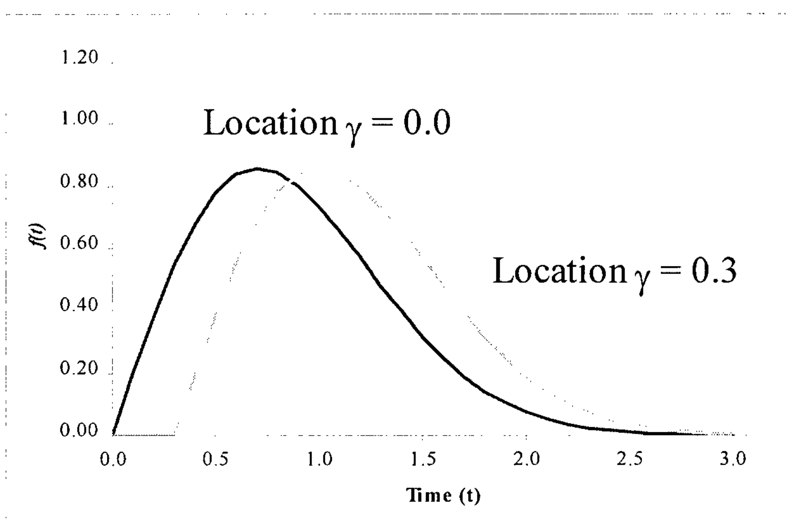


Figure 2-6. Effect of Weibull location γ change, with $\beta = 1$, $\delta = 2$

Factors Contributing to Development Cost

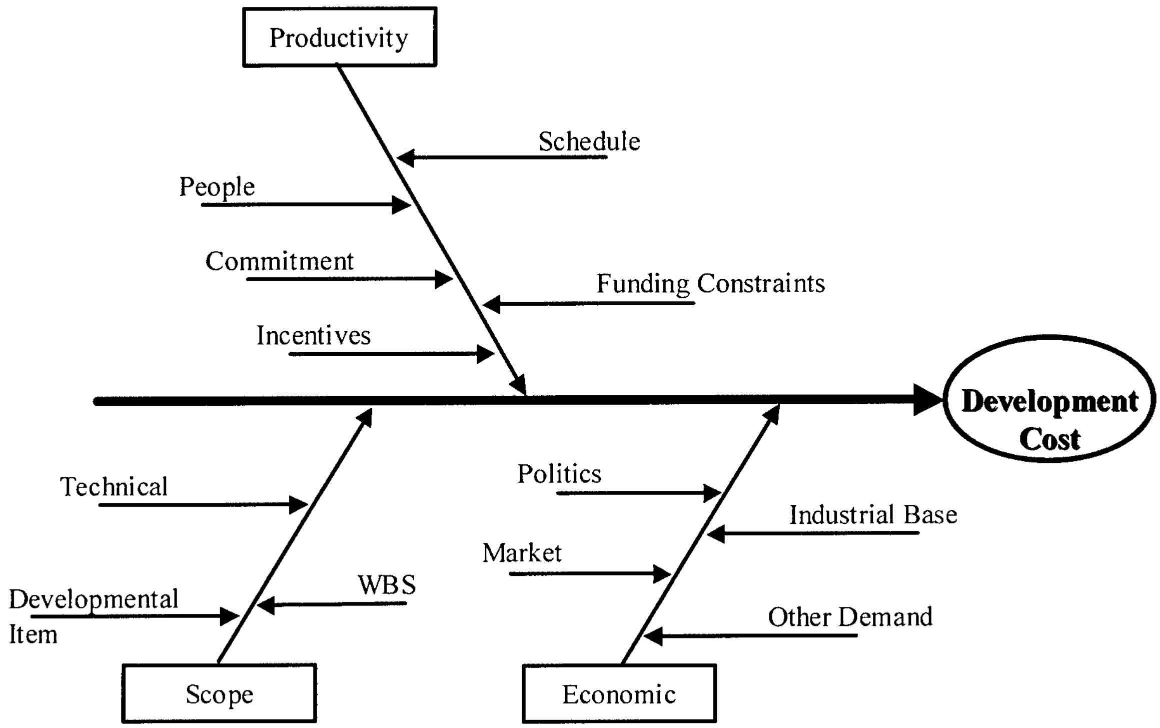


Figure 2-7. Factors that Contribute to Development Cost

A study by Belcher and Dukovich (2000) provides a macro perspective on the contributors to the cost of development programs. Based on expert opinion, Belcher and Dukovich determine a framework of how development programs incur cost. They cite three major areas that cause development cost: productivity, scope and economic. They refine the detail by providing several factors that span each major area. In total Belcher and Dukovich identify 12 factors that contribute to development cost.

In this study, we examine one of the factors described by Belcher and Dukovich. Our models focus only on the impact of the "Funding Constraints" factor on cost and schedule growth. We measure the funding constraints by how well the initial budget profile supports a Rayleigh distributed expenditures.

Chapter Summary

Based on research conducted over the past 30 years, the Rayleigh model provides a method for determining an expenditure profile for development programs. Norden described a useful mathematical relationship for development effort and manpower over time. Putnam applied the Rayleigh model to software development. Lee, Hogue, and Gallagher concluded the Rayleigh modeled outlays of a wide variety of defense development programs. Therefore, using the Rayleigh and its generalization, the Weibull function, is justified in modeling expenditure profiles. After describing the Rayleigh, its derivation, and its attributes, this chapter discusses the flexible Weibull distribution. We derive the applicability of the Weibull model to development programs, explain its parameters, and describe its ability to model a wide variety of funding profiles. Finally, we provide a summary view of the twelve factors that affect program growth while noting that we are only assessing the impact of one factor, funding constraints.

III. Research Methodology and Results

Chapter Overview

Previous research indicates that the Rayleigh cumulative distribution function models contract expenditures effectively. This research intends to expand the application of the model to include program level expenditures, a much broader scale. The focus of this chapter is to explain the methodology employed to test for a relationship between program initial funding profiles and growth.

The initial step of this research is to collect appropriate data. We collected program budget data from the Selected Acquisition Report, maintained by the Office of the Secretary of Defense Program Analysis and Evaluation (PA&E). We selected programs meeting certain criteria from this large data repository.

After translating the initial budget data into constant year expenditures, we estimated parameters for both a Weibull and Rayleigh cost models. We then evaluate whether or not the expenditures from the initial budget fits a Rayleigh distribution via the Cramer Von Mises, Anderson-Darling, and Kolmolgorov-Smirnov goodness of fit statistics. Each of these parameters and statistics is available as an independent variable for regression modeling.

We test for a relationship between program growth by creating a regression model. We create two separate models, one for cost growth and one for schedule growth, using multiple linear regression. We validate the resulting models by assessing the respective predictive abilities against programs not included in the initial dataset.

Program Data

We collected program funding data for evaluating the research hypothesis from numerous Selected Acquisition Reports (SARs). The full dataset consists of thirty-seven Army, Navy, Air Force, and Joint programs. The program total costs range in constant year dollars (CY00$) from $15.5M to $13,686.3M.

Selected Acquisition Report (SAR). The SAR provides standard, comprehensive summary reporting of cost, schedule, and performance information for DoD Acquisition Category (ACAT) IC and ID Major Defense Acquisition Programs (MDAPs) to Congress. The SAR compares the current estimate of total program acquisition cost, schedule, and performance data against a baseline to derive program cost variances and analysis. The SAR also includes unit cost reporting and Nunn-McCurdy unit cost breach information in accordance with Title 10 USC 2433 for programs beyond Milestone II (DAD, 1999:2.B.3.2). The SAR reports are prepared annually in conjunction with the President's budget. Quarterly exception reports are required only for those programs experiencing unit cost increases of at least 15 percent or schedule delays of at least six months. Programs submit Quarterly SARs for initial reports, final reports, and for programs rebaselined at major milestone decisions.

The total program cost estimates provided in the SARs include research and development, procurement, military construction, and acquisition-related operation and maintenance. Total program costs reflect actual costs to date as well as anticipated costs for future efforts. All estimates include allowances for anticipated inflation; the SAR presents them in millions of then year dollars (TY$M).

Over the past 3 decades, the SAR format changed several times. In 1983, a SAR Improvement Task Force significantly reduced the content of the SAR. However, in 1985 the FY86 Authorization Act restored the information that had been removed and added production rate and operating and support cost information. The FY87 Authorization Act provided for limited reporting for pre-Milestone II programs and relaxed some reporting criteria. The FY92 Authorization Act gave the Secretary of Defense the authority to waive selected acquisition reporting for certain programs and to change the content of the SAR as long as the appropriate House and Senate committees were notified in advance. The Federal Acquisition Streamlining Act of 1994 changed the baseline for unit cost reporting purposes from the prior President's budget to the approved acquisition program baseline and substituted procurement unit cost for current year procurement unit cost. In 1996, the Department made additional changes in SAR format and content, which reduced the volume of the SAR by about 20-30%.

Program Selection. The Office of the Secretary of Defense's (OSD) SAR archive provides funding data on hundreds of defense programs, dating back to the early 1970s. However, not all of the programs are useful to this study. The research requirements and the SAR database attributes limit the number of applicable programs.

To establish a proper funding profile, the identified program's budget needs to be represented in annual budget format. Before 1980, the SAR reported all budget information in aggregate. While interesting, the cumulative budget and expenditures do not provide adequate detail to construct a funding profile. Around 1982, OSD changed the SAR format to include annual budget and expenditure data. We selected only programs that reported in the annual budget format.

Since the goal of this study is to establish a relationship between initial budget and program growth, the dataset necessarily must include only completed programs. This requirement forces the final program SAR to be dated 1998 or earlier, since we used the 1999 version of the OSD SAR database. The combination of the first two criteria bound programs' development life between 1982 and 1998. We excluded all of the programs falling out of this range.

Joint programs are generally large programs with funding from various sources, including the services and other defense agencies. We prefer to evaluate each funding source of the program separately, to fit the funding profiles to the distributions at a lower level. However, the apparent financial management of joint programs, combined with the top-level SAR data collection precludes separate evaluation. Therefore, we treat joint programs as a single program by summing individual service contributions, avoiding the problems of discerning interservice funding transitions.

For each program, we collect the initial and final budget profile. Each budget profile includes the then year funding by year. The SAR presents funding in both then year and base year formats. However, we collect the then year information to apply conversion to constant year consistently, across programs. Table 3-1 shows an example budget profile for the Air Force's E-3 Airborne Warning and Control System (AWACS) Radar System Improvement Program (RSIP).

Table 3-1. RSIP Program Budget Profile

	1989	1990	1991	1992	1993	1994	Total
Budget (TY$M)	44.2	63.7	68.8	116.3	72.1	30.9	396.0

Table 3-2 shows the programs included in the study and their associated dollar cost overrun, schedule slip, initial budget dollars, and initial budget profile years. We show cost growth as the dollar difference between the initial program cost and final program cost, expressed in CY\$00M. We show schedule growth as a ratio. For example, if the original program was 5 years long and the final program was 7 years long, the schedule slip is 1.40. A schedule slip of 1.00 indicates that the final program was the same length as the initial program, indicating no growth. The 37 programs in Table 3-2 show a range of schedule growth of 0.67 to 3.20, with an average schedule growth of 1.26. The programs range in schedule growth from –\$688.62M to \$1199.33M, with an average cost growth of \$40.33M CY00.

Table 3-2. Included Programs

Program	Cost Overrun (CY00$M)	Schedule Slip Ratio	Cost Initial Budget (CY00$M)	Years Initial Budget
1 AAMRAAM-J	500.850	1.31	1230.4	13
2 ASPJ-J	42.912	1.13	752.7	15
3 B-1B	1199.329	1.43	3861.3	7
4 Battleship	2.619	1.00	32.2	6
5 IUS	21.987	1.60	1230.7	15
6 KC-135R	-3.694	1.00	162.9	8
7 Kiowa	-0.588	1.17	339.9	6
8 Lantirn	152.199	1.50	654.3	10
9 Trident-MSL	-688.615	1.23	13686.3	13
10 Trident-Sub	22.154	3.20	78.9	5
11 RPV(Aquila)	59.157	1.00	1123.9	14
12 MK48 ADCAP	101.786	1.23	1522.9	13
13 E-6(Tacamo)	32.614	1.13	497.5	8
14 Avenger	1.398	1.00	15.5	3
15 PLS	0.306	1.80	51.9	5
16 RSIP	27.620	1.33	449.8	6
17 Longbow	180.750	1.22	847.9	9
18 CMU	174.181	1.10	1511.1	20
19 AOE6	2.451	1.11	41.1	9
20 TRITAC	13.737	1.22	576.7	18
21 MLRS-TGW	-28.432	1.08	407.6	12
22 JSOW	31.381	1.44	404.2	9
23 ASAT	276.789	0.89	2112.4	19
24 ADDS	87.449	1.78	267.5	9
25 LCAC	12.892	1.71	37.8	7
26 LSD41	-5.515	1.00	94.4	11
27 MK 50	77.069	1.18	2129.8	17
28 Backscatter	-48.563	1.09	642.2	11
29 Peacekeeper	-576.644	1.57	9847.5	7
30 T46A	-57.526	0.80	546.0	10
31 Cvhelo	-3.120	1.00	81.6	4
32 TAO187Oiler	-1.890	0.67	24.7	9
33 FDS	479.552	1.30	931.6	10
34 ATARS	-7.870	1.00	203.0	9
35 SRAMII	-347.508	0.89	1337.4	9
36 ATACMS	-300.523	1.00	1058.3	11
37 TOW2	61.400	1.44	155.8	9

Conversion to Outlays

Theory and past research suggest that outlays should follow a Rayleigh distribution. We convert the program budget profile into an expenditure profile to which we fit Rayleigh and Weibull models. The conversion process takes two steps: converting budget dollars to outlays and adjusting for inflation. We depict the RSIP program budget in Figure 3-1.

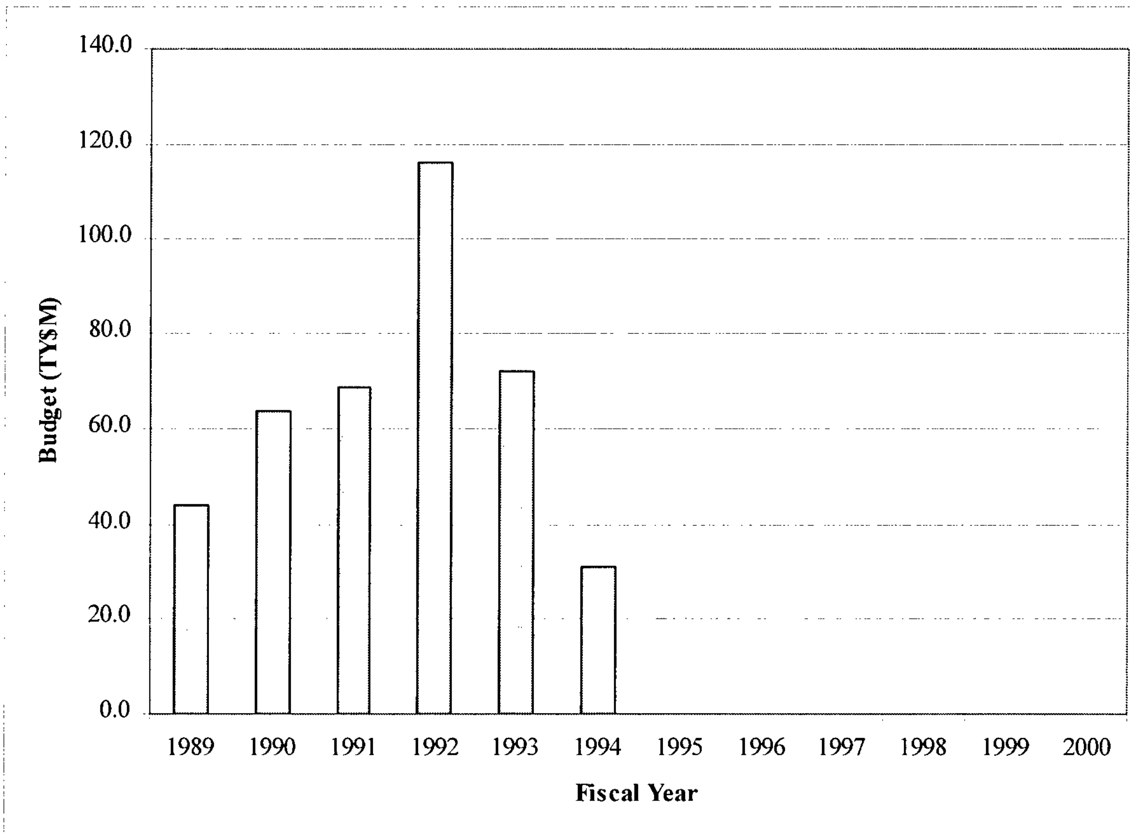


Figure 3-1. Then Year Budget Profile for RSIP Program

The first step in the translation to an appropriate outlay profile is to convert the then year budget profile into outlays. We apply the outlay rates the historical percents of funds obligated in the budget year and subsequent years. After multiplying the budget by the outlay rates, we summed the funds obligated from various budget years to obtain the amount obligated in each fiscal year, as shown in (1).

The OSD Comptroller, OSD(C), provides Congressionally approved outlay rates, also called spendout rates, annually for service and appropriation. OSD(C) develops outlay rates for every funding appropriation. These outlay rates reflect the percent of budget authority expected to be expended by year, based on historical experience. Outlay rates are implicitly considered in the budgetary concept of then-year dollars. OSD(C) disseminates outlay rates within DoD for development of weighted inflation indices. Historically, different commodities have had different rates of inflation. Within the DoD, the rates of inflation are linked to the source appropriation, such as Military Construction (MILCON) or Research Development Test and Evaluation (RDT&E). DoD weights the inflation by the outlay rates to form weighted composite inflation indices. In addition to different inflation rates, the appropriations do not all spent at the same rate. Military Personnel (MILPERS) costs are spent in the year appropriated, while Research Development Test and Evaluation costs may be spread out over as much as six years.

Table 3-3. Service R&D Outlay Rates (Percentages)

Air Force

	FY01	FY00	FY99	FY98	FY97	FY96	FY95	FY94	FY93	**Avg**
Yr1	59.5	58.8	59.1	50.7	45.8	46.3	46.5	46.5	50.8	51.56
Yr2	33.7	34.5	33.1	37.4	39.9	39.1	38.8	38.8	34.5	36.64
Yr3	3.6	3.6	5.3	6.8	8.9	8.9	8.8	8.9	9.5	7.14
Yr4	1.0	1.0	1.4	3.0	3.6	3.6	3.6	3.6	3.4	2.69
Yr5	0.3	0.3	0.4	0.8	1.1	1.1	1.1	1.1	1.1	0.81
Yr6	0.0	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.31
Yr7	0.0	0.0	0.2	0.0	0.1	0.0	0.2	0.0	0.2	0.08
	98.1	98.5	99.8	99.0	99.7	99.4	99.4	99.3	99.9	99.23

Army

	FY01	FY00	FY99	FY98	FY97	FY96	FY95	FY94	FY93	**Avg**
Yr1	57.5	56.8	58.0	58.0	58.0	57.0	57.0	55.0	55.0	56.92
Yr2	32.5	33.7	33.0	33.0	33.0	34.0	34.0	34.0	34.0	33.47
Yr3	6.3	5.0	5.3	5.3	5.3	5.3	5.3	7.3	7.3	5.82
Yr4	2.1	2.1	2.1	1.8	1.8	1.8	1.8	1.8	1.8	1.90
Yr5	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.80
Yr6	0.0	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.44
Yr7	0.0	0.0	0.2	0.0	0.2	0.0	0.2	0.0	0.2	0.09
	99.2	98.9	99.9	99.4	99.6	99.4	99.6	99.4	99.6	99.44

Navy

	FY01	FY00	FY99	FY98	FY97	FY96	FY95	FY94	FY93	**Avg**
Yr1	59.5	59.3	60.5	58.0	55.9	55.9	54.0	55.0	55.0	57.01
Yr2	31.4	33.6	32.5	33.1	31.5	31.5	32.4	33.4	33.4	32.53
Yr3	5.9	4.5	4.5	5.4	8.2	8.2	8.0	7.8	7.8	6.70
Yr4	1.9	1.0	1.0	2.0	2.0	2.0	2.0	1.3	1.3	1.61
Yr5	0.7	0.3	0.3	0.6	1.1	1.1	1.1	1.1	1.1	0.82
Yr6	0.0	0.2	0.2	0.2	0.4	0.4	0.4	0.4	0.4	0.29
Yr7	0.0	0.0	0.0	0.0	0.6	0.0	0.2	0.0	0.2	0.11
	99.4	98.9	99.0	99.3	99.7	99.1	98.1	99.0	99.2	99.08

Since the outlay rates vary slightly from year to year, we created an average outlay rate profile for each of the services. We apply the average outlay profile to each program for a given service. The composite outlay rate profile is the average of nine outlay profiles from FY93 to FY01. Table 3-3 shows the yearly outlay rates and the composite average used in the individual program outlay calculations.

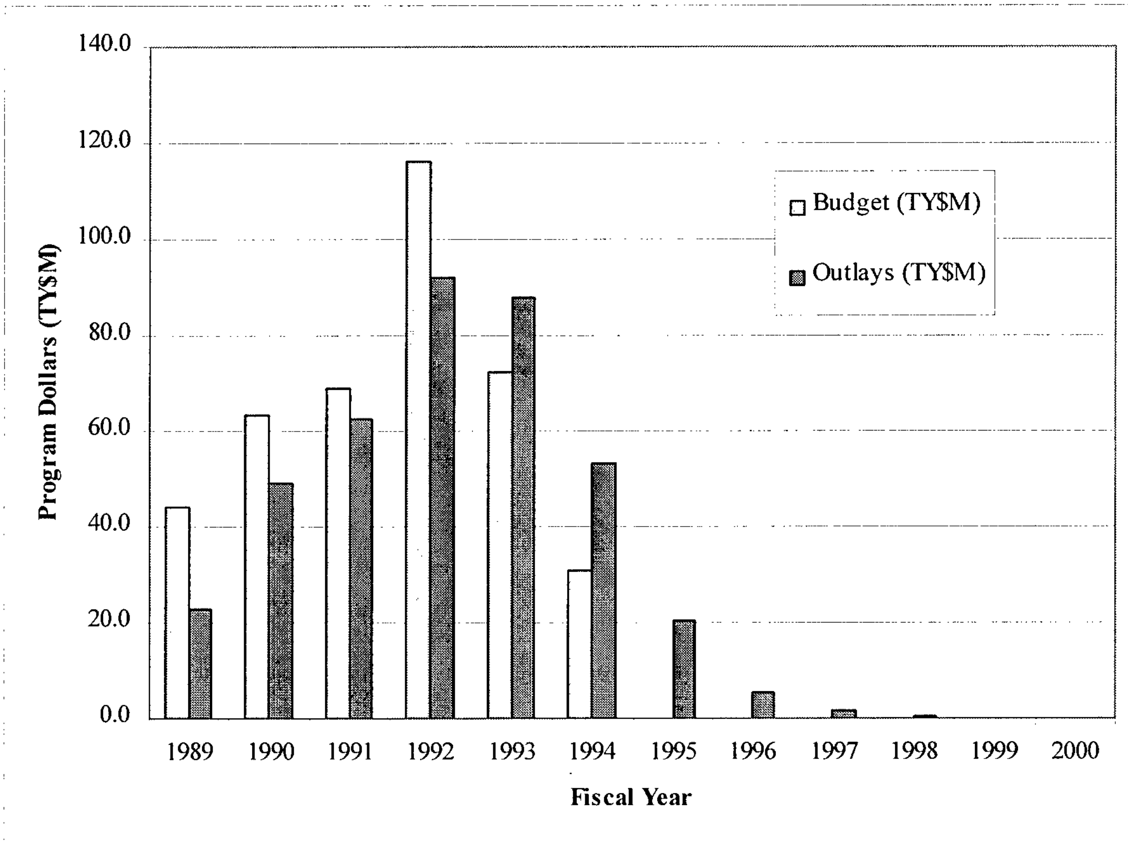


Figure 3-2. Then Year Budget and Outlay Profile for RSIP Program

For the RSIP budget example in Figure 3-1, we apply the Air Force average outlay rates and sum to determine outlays for each fiscal year. Figure 3-2 depicts the resulting outlays, also called expenditures, from the RSIP initial budget. Table 3-4 provides the actual results for the RSIP conversion to outlays.

Table 3-4. RSIP Budget to Outlays

Fiscal Year	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Budget (TY$M)	44.2	63.7	68.8	116.3	72.1	30.9	0.0	0.0	0.0	0.0	0.0	0.0
Outlays (TY$M)	22.8	49.0	62.0	90.9	86.8	53.2	20.4	5.3	1.8	0.5	0.1	0.0

Convert Then Year to Constant Year. After converting the then year budget profile to a then year outlay profile, we translate into the outlays into constant year dollars. This translation makes the funds spent in various years comparable within a program and across programs by eliminating the impact of inflation. For instance, $10 million dollars in 1990 is worth more than $10 million in 2000 dollars, because inflation has eroded the value of a dollar during that time. Also, the Rayleigh theory is based on linear skills acquisition, which does not have an inflation component. Thus, we must remove the effects of inflation from the outlays to properly fit the model.

Since all of the program data is in then year dollars, we convert each year of budget to constant year using the raw inflation index for the appropriate service and appropriation. Like outlay rates, OSD publishes inflation rates for each service annually. The services sometimes modify published rates to account for service peculiarities. Appendix A contains tables for the Air Force, Army and Navy raw and Weighted Inflation indices.

To apply the Rayleigh model properly, we convert then year outlay profiles to constant year 2000 profiles. Since each program's outlays are in then year dollars, a single application of the appropriate service inflation rates converts the outlays to constant year 2000 dollars. Table 3-5 shows results of the conversion to constant year dollars for the Air Force RSIP program.

Table 3-5. RSIP CY00$M Expenditures

Fiscal Year	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Budget (TY$M)	44.2	63.7	68.8	116.3	72.1	30.9	0.0	0.0	0.0	0.0	0.0	0.0
Outlays (TY$M)	22.8	49.0	62.0	90.9	86.8	53.2	20.4	5.3	1.8	0.5	0.1	0.0
Expenditures (CY$00M)	29.0	59.9	72.6	103.6	96.3	57.8	21.7	5.6	1.8	0.5	0.1	0.0

Figure 3-3 contrasts the initial then year budget profile to the constant year 2000 expenditure profile for the RSIP program.

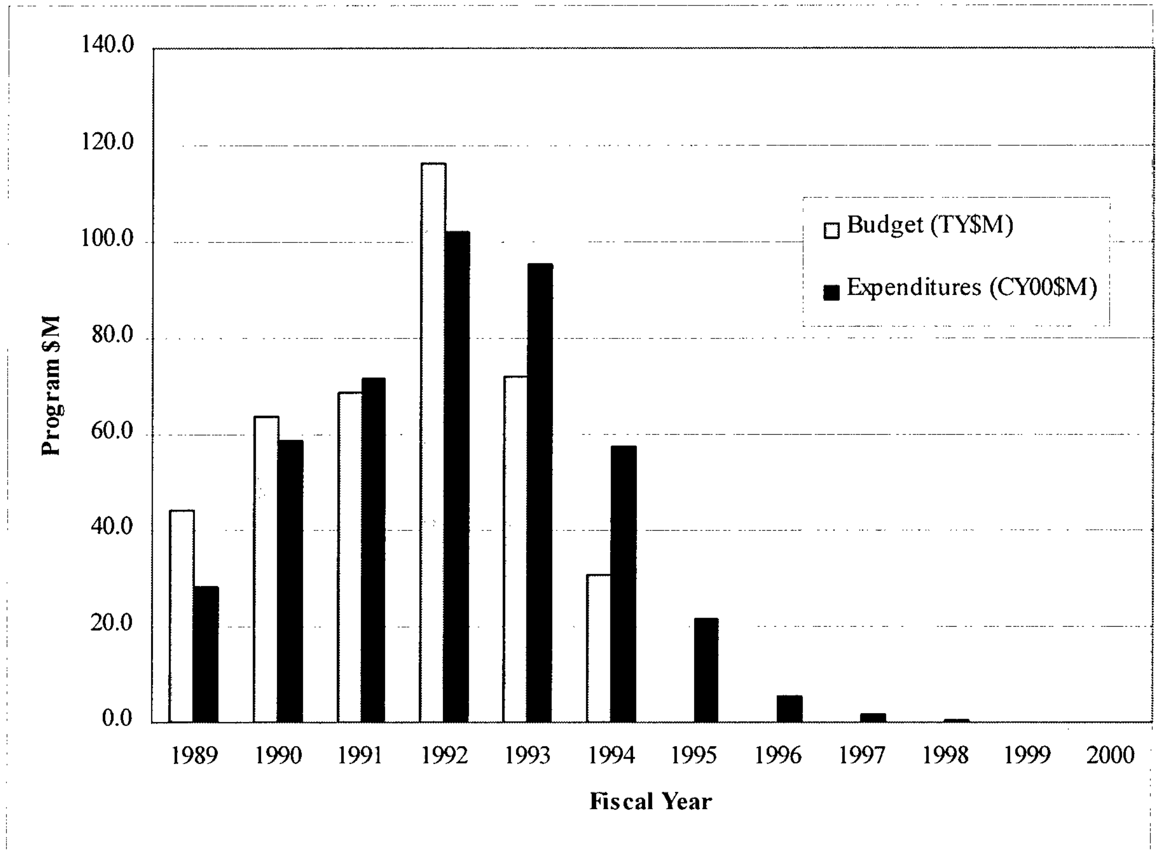


Figure 3-3. Constant Year Expenditure Profile for RSIP

With the expenditure profile in constant year 2000 dollars, we fit a Rayleigh and Weibull distribution to the expenditures using the least squares error technique to identify the respective parameters. The purpose of this is threefold. The Rayleigh parameters provide a means to evaluate whether the profile is Rayleigh using Goodness of Fit statistics. The Weibull parameters provide another means of evaluating whether the expenditures are Rayleigh, by examining the shape parameter. Finally, we use the Rayleigh parameters, Weibull parameters and Goodness of Fit statistics as a pool of independent variables to create a regression model. The regression model establishes the relationship between program growth and initial budget profile.

Parameter Estimation

Fitting the Rayleigh and Weibull functions to the expenditures involves identifying the parameters for each respective function that minimizes the sum of squared difference between the actual cumulative CY00\$M expenditures from (1) and (2) and the proposed theoretical expenditure model, either Rayleigh or Weibull given in (5) and (7), respectively.

For each year, we subtract the scaled Rayleigh cumulative distribution function (5) value from the actual expenditures. This difference is squared and added to each of the other squared differences. We use Solver™ to minimize the sum of the squared errors by changing the Rayleigh cumulative distribution function parameters. Let C_n represent the cumulative actual expenditures expressed in CY00\$M at the end of the nth fiscal year (time t_n). We represent the least squares estimation

$$\min \Sigma(\text{Error})^2 = \min \sum_{n=1}^{n=N} \left[C_n - d\left(1 - e^{-\left(\frac{t-\gamma}{\delta}\right)^{\beta}}\right)\right]^2 \qquad (8)$$

where $\beta = 2$ for the Rayleigh model and is allowed to vary for the Weibull approximation and N is the number of budget years plus outlay years minus one.

The Rayleigh has difficulty modeling programs with insignificant funding. Therefore, we allow γ to vary for the Rayleigh distribution to take advantage of the location parameter's ability to model programs with very low initial funding. Including the location parameter allows for better fit, and a lower rejection with the goodness of fit tests for programs being estimated with the Rayleigh.

Table 3-6 shows the iterations from budget to cumulative then year expenditures for the Air Force RSIP program. We estimate the parameters for the Rayleigh function and Weibull function that minimize the sum of squared error.

Table 3-6. Cumulative Expenditure Profile for the RSIP Program

Fiscal Year	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Budget (TY$M)	44.2	63.7	68.8	116.3	72.1	30.9	0.0	0.0	0.0	0.0	0.0	0.0
Outlays (TY$M)	22.8	49.0	62.0	90.9	86.8	53.2	20.4	5.3	1.8	0.5	0.1	0.0
Expenditures (CY$00M)	29.0	59.9	72.6	103.6	96.3	57.8	21.7	5.6	1.8	0.5	0.1	0.0
Cum Expenditures (CY$00M)	29.0	88.9	161.5	265.1	361.4	419.2	441.0	446.5	448.3	448.8	448.9	448.9

Table 3-7 depicts the minimization results for the RSIP program, when $\beta = 2$ in (13). For any other combination of parameters, the sum of squared error would be greater than 1136.97.

Table 3-7. RSIP Minimization of $\Sigma(\text{Error})^2$

	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Annual Rayleigh (CY00$M)	12.6	66.0	98.7	99.5	78.0	49.9	26.6	11.9	4.5	1.5	0.4	0.1
Cum Rayleigh (CY00$M)	12.6	78.6	177.3	276.8	354.8	404.7	431.2	443.1	447.7	449.2	449.6	449.7
Error²	267.7	106.0	249.2	136.8	43.3	211.8	94.8	11.4	0.4	0.1	0.4	0.6
										Σ Error² =		1136.97

Table 3-8 shows the initial parameter values, along with the parameters resulting from the minimization in (8). The start value for d is 1.05 times the initial total budget. We set δ to 7.0, since the parameter estimation failed for lower values. We set γ to zero, to reflect a presumed program start at time zero. We allow the Rayleigh to use a location parameter, similar to the Weibull, allowing for greater flexibility in fitting programs with insignificant initial funding.

Table 3-8. RSIP Rayleigh Parameters

Start Values	**Least Squares Rayleigh Parameters**
$d=$ 472.3	Cost Factor d = 449.70
$\delta =$ 7.0	Scale (>0) $\delta =$ 3.708 a = 0.073
	Shape $\beta = 2$
$\gamma =$ 0.0	Location (>t, else 0) $\gamma =$ 0.375

Figure 3-4 depicts the resulting Rayleigh model fit to both the cumulative and rate functions. While the sum of squared error technique produces the closest Rayleigh model for given expenditures, it does not have the flexibility to fit expenditures exactly.

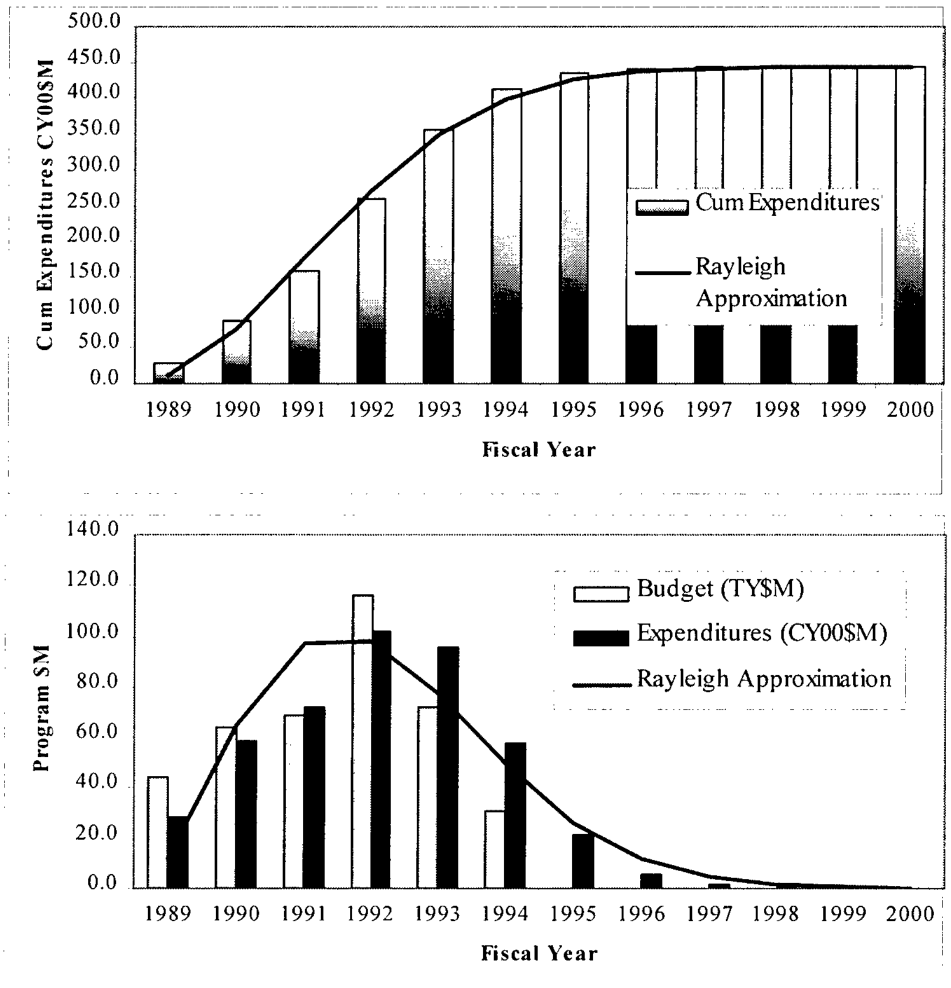


Figure 3-4. RSIP Cumulative (top) and Rate (bottom) Rayleigh Models

We use the same procedure to identify the parameters for the Weibull model. Table 3-9 shows the Weibull least squares parameters, with a sum of squared error of 636.17, for the RSIP program. Since the flexibility of the Weibull model allows for closer approximation of the expenditures, the sum of squared error technique produces a smaller value.

Table 3-9. RSIP Weibull Parameters

Start Values	**Least Squares Weibull Parameters**
$d =$ 472.3	Cost Factor $d =$ 449.29
$\delta =$ 5.0	Scale (>0) $\delta =$ 4.103
$\beta =$ 1.5	Shape (>0) $\beta =$ 2.359
$\gamma =$ 0.0	Location (>t, else 0) $\gamma =$ 0.000

Figure 3-5 depicts the resulting Weibull model fit to both the cumulative and rate functions. In conjunction with a smaller sum of squared error, the Weibull model has a closer visual fit to the expenditures than the Rayleigh model.

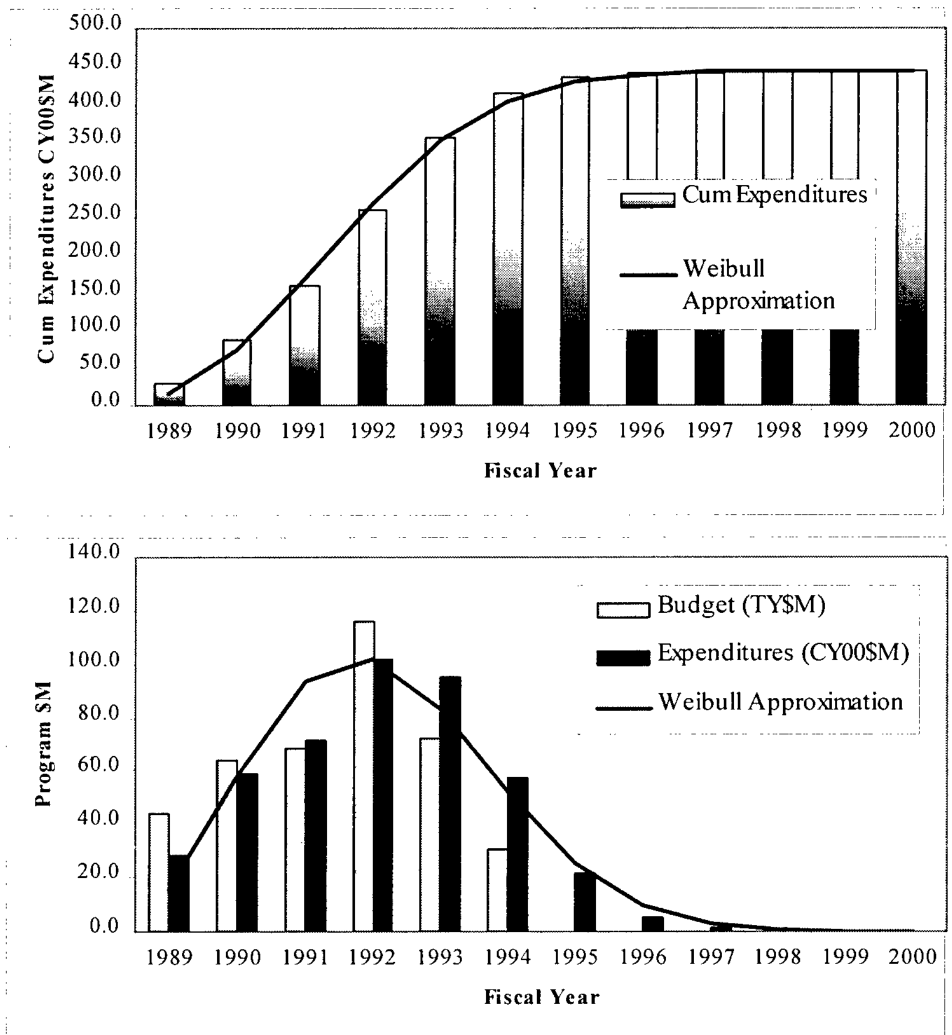


Figure 3-5. RSIP Cumulative (top) Rate (bottom) Weibull Models

Goodness of Fit

To evaluate if actual expenditures are Rayleigh, we assess the cumulative proportion of the actual expenditures against the Rayleigh cumulative distribution function. This comparison describes a statistical technique called "goodness of fit." We use three goodness of fit statistics.

The Kolmolgorov is simply the maximum absolute deviation between $F_n(x)$, which is the empirical distribution function (EDF), and $F_o(x)$, which is the hypothesized cumulative distribution function (CDF).

$$K = \max\left|F_n(x) - F_o(x)\right|$$

The Cramer-von Mises and the Anderson-Darling "goodness of fit" statistics may be represented in the form

$$Q = n \int_{x=-\infty}^{x=+\infty} [F_n(x) - F_o(x)]^2 \Psi(x) dF_o(x)$$

where $\Psi(x)$ is a weighting function, and n is the number of data points. When $\Psi(x) = 1$, the Cramer-von Misses statistic $W^2 = \text{Q}$. The Anderson-Darling (1952) goodness of fit statistic $\boldsymbol{A^2} = \text{Q}$ uses the weighting function

$$\Psi(x) = \frac{1}{F_o(x)(1 - F_o(x))}$$

Since the CDF ranges from zero at $x = 0$ to one at $x = \infty$, the denominator starts at zero increases and then decreases. Therefore, the weighting function, which is the inverse of the denominator, starts at infinity decreases and eventually increases back to infinity. As a result, the Anderson-Darling "goodness of fit" statistic A^2 weights the tails of the distribution much heavier than the center of the distribution.

Stephens (1974) presents computational formulas for these goodness of fit statistics that assume a sample of random variables at which cumulative probabilities increases to i/n, for $i = 1, \ldots n$, at the point of the ith random variable in ascending order. In contrast, we have annual cost reports that provided the cumulative percent at the end of each year. In this application, $F_n(i)$ is known for $i = 1,\ldots,n$. Following the derivation of the computational formula in Crown (1997), we develop the appropriate calculation formula for this application.

The Anderson-Darling goodness of fit statistic is defined as

$$A^2 = n \int_{x=-\infty}^{x=+\infty} \left[F_n(x) - F_o(x)\right]^2 \left[F_o(x)(1 - F_o(x))\right]^{-1} dF_o(x)$$

We define $u = F_o(x)$. If the density function corresponding $F_o(x)$ is continuous then the cumulative density function is strictly monotonically increasing, and hence invertible. Furthermore, the cumulative density function is limited from zero to one inclusive. We substitute, and rewrite by partial fractions, to obtain

$$A^2 = n \int_{u=0}^{u=1} \frac{\left[F_n(F_o^{-1}(u)) - u\right]^2}{u(1-u)} du$$

$$= n \int_{u=0}^{u=1} \frac{\left(F_n(F_o^{-1}(u))\right)^2 - 2F_n(F_o^{-1}(u))u + u^2}{u(1-u)} du$$

$$= n \int_{u=0}^{u=1} \left\{ \frac{\left(F_n(F_o^{-1}(u))\right)^2}{u} + \frac{\left(F_n(F_o^{-1}(u))\right)^2 - 2F_n(F_o^{-1}(u)) + u}{(1-u)} \right\} du$$

Expanding the numerator and completing the squares, we get

$$A^2 = n\int_{u=0}^{u=1}\left\{\frac{\left(F_n(F_o^{-1}(u))\right)^2}{u}+\frac{\left(F_n(F_o^{-1}(u))\right)^2-2F_n(F_o^{-1}(u))+1-1+u}{(1-u)}\right\}du$$

$$=n\int_{u=0}^{u=1}\left\{\frac{\left(F_n(F_o^{-1}(u))\right)^2}{u}+\frac{\left(F_n(F_o^{-1}(u))-1\right)^2}{(1-u)}-1\right\}du$$

We note that the empirical distribution function only changes at integer values, when the annual costs are reported. Hence, $F_n(i)$ is constant over the interval from i to i+1. At each point i, the function $F_n(i)$ has a jump discontinuity. If we set the limits of integration equal to the jump discontinuities, we may determine the value of the integral using the constant for that integral. Defining $F_o(0)=0$ and $F_o(n+1)=1$, we obtain

$$A^2 = n\left[\int_{u=0}^{u=F_o(1)}+\int_{u=F_o(1)}^{u=F_o(2)}+\cdots\int_{u=F_o(n)}^{u=1}\left\{\frac{\left(F_n(F_o^{-1}(u))\right)^2}{u}+\frac{\left(F_n(F_o^{-1}(u))-1\right)^2}{1-u}-1\right\}du\right]$$

$$=n\sum_{i=1}^{i=n+1}\int_{u=F_o(i-1)}^{u=F_o(i)}\left\{\frac{\left(F_n(F_o^{-1}(u))\right)^2}{u}+\frac{\left(F_n(F_o^{-1}(u))-1\right)^2}{1-u}-1\right\}du$$

$$=n\sum_{i=1}^{i=n+1}\left[\int_{u=F_o(i-1)}^{u=F_o(i)}\frac{\left(F_n(F_o^{-1}(u))\right)^2}{u}du+\int_{u=F_o(i-1)}^{u=F_o(i)}\frac{\left(F_n(F_o^{-1}(u))-1\right)^2}{1-u}du-\int_{u=F_o(i-1)}^{u=F_o(i)}du\right]$$

$$=n\sum_{i=1}^{i=n+1}\left[\left(F_n(i-1)\right)^2\int_{u=F_o(i-1)}^{u=F_o(i)}\frac{du}{u}+\left(F_n(i-1)-1\right)^2\int_{u=F_o(i-1)}^{u=F_o(i)}\frac{du}{1-u}-\int_{u=F_o(i-1)}^{u=F_o(i)}du\right]$$

$$=n\sum_{i=1}^{i=n+1}\left[(F_n(i-1))^2\ln(u)\Big|_{u=F_o(i-1)}^{u=F_o(i)}-(F_n(i-1)-1)^2\ln(1-u)\Big|_{u=F_o(i-1)}^{u=F_o(i)}-u\Big|_{u=F_o(i-1)}^{u=F_o(i)}\right]$$

$$=n\sum_{i=1}^{i=n+1}\left[(F_n(i-1))^2\left(\ln(F_o(i))-\ln(F_o(i-1))\right)-(F_n(i-1)-1)^2\left(\ln(1-F_o(i))-\ln(1-F_o(i-1))\right)\right]$$
$$-n\sum_{i=1}^{i=n+1}\left[F_o(i)-F_o(i-1)\right]$$

$$=n\sum_{i=1}^{i=n}\left[(F_n(i))^2\left(\ln(F_o(i+1))-\ln(F_o(i))\right)-(F_n(i-1)-1)^2\left(\ln(1-F_o(i))-\ln(1-F_o(i-1))\right)\right]-n$$

Law and Kelton (1991:392) present critical values for an adjusted Anderson-Darling test statistic. The adjusted test statistic is

$$\left(1+\frac{0.2}{\sqrt{n}}\right)A^2$$

For a Weibull distribution, they give the 90% critical value as 0.637. We reject the null hypothesis that the distribution is Weibull when the calculated Anderson-Darling test statistic exceeds this critical value.

With the data in ascending order, the formula for the Cramer-von Mises is

$$W^2 = \sum_{i=1}^{i=n}\left[F_o(x_i) - \frac{2i-1}{2n}\right]^2 + \frac{1}{12n}.$$

The expected value of the Cramer-von Misses statistic is 1/6, so some multiply it by 6. Watson (1961) proposed an adjustment to the Cramer-von Misses to correct for the sample mean. Again, we calculate the Cramer-von Mises statistic for data at fixed intervals.

$$W^2 = n\int_{u=0}^{u=1}\left[F_n(F_o^{-1}(u)) - u\right]^2 du$$

$$= n\left[\int_{u=0}^{u=F_o(1)} + \int_{u=F_o(1)}^{u=F_o(2)} + \cdots \int_{u=F_o(n)}^{u=1}\left[F_n(F_o^{-1}(u)) - u\right]^2 du\right]$$

$$= n\sum_{i=1}^{i=n+1}\int_{u=F_o(i-1)}^{u=F_o(i)}\left[F_n(F_o^{-1}(u)) - u\right]^2 du$$

Let $g(u) = F_n(F_0^{-1}(u) - u)$, where $F_0^{-1}(u)$ is consistent in any interval $[i-1, i]$ so $\frac{dg(u)}{du} = -1$ and $\int (\mathrm{g(u)})^2 du = \frac{-g(u)^3}{3} + C$

$$W^2 = n\sum_{i=1}^{i=n+1}\int_{u=F_0(i-1)}^{u=F_0(i)}\left[F_n(F_o^{-1}(u)) - u\right]^2 du$$

$$= n\sum_{i=1}^{i=n+1}\int_{u=F_0(i-1)}^{u=F_0(i)}\left[g(u)\right]^2 du$$

$$= n\sum_{i=1}^{i=n+1}\frac{-1}{3}(g(u))^3\Big|_{u=F_0(i-1)}^{u=F_0(i)}$$

$$= n\sum_{i=1}^{i=n+1} \frac{-1}{3}\left[\left(F_n(F_0^{-1}(u)) - u\right)^3\right]_{u=F_0(i-1)}^{u=F_0(i)}$$

$$= \frac{n}{3}\sum_{i=1}^{i=n+1}\left[(u - F_n(F_0^{-1}(u)))^3\right]_{u=F_0(i-1)}^{u=F_0(i)}$$

$$= \frac{n}{3}\sum_{i=1}^{i=n+1}\left[\left(F_0(i) - F_n(i-1)\right)^3 - \left(F_0(i-1) - F_n(i-1)\right)^3\right]$$

Bush and Moore present critical values for the Cramer-von Misses test statistic (Bush, 1983:2469). For a Weibull distribution with $\beta = 2$ (Rayleigh), they give the 90% critical values for various sample sizes. The critical values for sample sized of 10, 20, and 30 are 0.134, 0.139 and 0.141, respectively. The sample size equals the number of budget years plus outlay years minus one, and we linearly interpolate between the critical values as appropriate. We reject the null hypothesis that the distribution is Weibull when the calculated Cramer-von Misses test statistic, W^2 exceeds the critical value.

Goodness of Fit Results

When we use the standard service outlay rates, the initial program budgets generally support Rayleigh expenditure profiles. The Cramer-von Misses supports 59.5% and the Anderson-Darling 70.3% of the programs tested. The theoretical acceptance rate is 90%. Table 3-10 shows Rayleigh goodness of fit test results for all of the initial expenditure profiles tested. In our regression model, we include the test statistics from the initial expenditures as indicators of cost overruns and schedule slips.

Table 3-10 Initial Expenditure Profile Goodness of Fit Results

#	Program	Kolmogorov	Cramer-von Mises	Init CvM Rayleigh?	Anderson-Darling	Init A-D Rayleigh?
1	AAMRAAM-J	0.4025	0.1139	Accept	0.3923	Accept
2	ASPJ-J	0.2283	0.0968	Accept	0.3879	Accept
3	B-1B	0.1016	0.1275	Accept	0.5076	Accept
4	Battleship	0.0335	0.0998	Accept	0.5537	Accept
5	IUS	0.0995	0.0844	Accept	0.4357	Accept
6	KC-135R	0.4377	0.1434	REJECT	0.6196	Accept
7	Kiowa	0.3282	0.1632	REJECT	0.7328	REJECT
8	Lantirn	0.1445	0.1020	Accept	0.4110	Accept
9	Trident-MSL	0.8195	0.1591	REJECT	0.6747	REJECT
10	Trident-Sub	0.1853	0.1192	Accept	0.8904	REJECT
11	RPV(Aquila)	0.6349	0.1253	Accept	0.5350	Accept
12	MK48 ADCAP	0.1330	0.0900	Accept	0.4418	Accept
13	E-6(Tacamo)	0.4994	0.1602	REJECT	0.7226	REJECT
14	Avenger	0.0402	0.1842	REJECT	0.9884	REJECT
15	PLS	0.0934	0.2751	REJECT	1.5696	REJECT
16	RSIP	0.1207	0.1261	Accept	0.4956	Accept
17	Longbow	0.2359	0.0970	Accept	0.4056	Accept
18	CMU	0.4990	0.0705	Accept	0.2808	Accept
19	AOE6	0.1343	0.1652	REJECT	0.5992	Accept
20	TRITAC	0.1142	0.0503	Accept	0.2474	Accept
21	MLRS-TGW	0.6417	0.1174	Accept	0.3689	Accept
22	JSOW	0.3735	0.1248	Accept	0.5040	Accept
23	ASAT	0.6909	0.1076	Accept	0.4466	Accept
24	ADDS	0.0443	0.0866	Accept	0.4327	Accept
25	LCAC	0.4772	0.1862	REJECT	0.7363	REJECT
26	LSD41	0.0582	0.0770	Accept	0.6066	Accept
27	MK 50	0.3491	0.0751	Accept	0.3183	Accept
28	Backscatter	0.0435	0.0753	Accept	0.5358	Accept
29	Peacekeeper	0.0348	0.1804	REJECT	1.0162	REJECT
30	T46A	0.6746	0.2109	REJECT	1.0077	REJECT
31	Cvhelo	0.0425	0.1629	REJECT	0.9396	REJECT
32	TAO187Oiler	0.1658	0.4641	REJECT	7.7864	REJECT
33	FDS	0.5654	0.1501	REJECT	0.6265	Accept
34	ATARS	0.3818	0.1134	Accept	0.5824	Accept
35	SRAMII	0.5136	0.1362	REJECT	0.4807	Accept
36	ATACMS	0.7006	0.1627	REJECT	0.6737	REJECT
37	TOW2	0.0848	0.1129	Accept	0.7955	REJECT

The expenditures based on final budgets distributed Rayleigh for 56.8% based on Cramer-von Misses and 51.4% based on Anderson-Darling. We question our basic assumption that total program cost for R&D programs result in Rayleigh expenditures, based on these low hypothesis acceptance rates. Table 3-11 shows the Rayleigh goodness of fit test results for the final expenditure profiles.

Table 3-11. Final Expenditure Profile Goodness of Fit Results

#	Program	Kolmogorov	Cramer-von Mises	Final CvM Rayleigh?	Anderson-Darling	Final A-D Raylirgh?
1	AAMRAAM-J	0.5044	0.1485	REJECT	0.6565	REJECT
2	ASPJ-J	0.1490	0.0831	Accept	0.4875	Accept
3	B-1B	0.0395	0.0930	Accept	0.5520	Accept
4	Battleship	0.0295	0.1023	Accept	0.5342	Accept
5	IUS	0.0948	0.1311	Accept	0.8506	REJECT
6	KC-135R	0.4357	0.1453	REJECT	0.6292	REJECT
7	Kiowa	0.3049	0.1657	REJECT	0.7627	REJECT
8	Lantirn	0.1533	0.1051	Accept	0.5050	Accept
9	Trident-MSL	0.8082	0.1780	REJECT	0.8057	REJECT
10	Trident-Sub	0.1702	0.1810	REJECT	1.8601	REJECT
11	RPV(Aquila)	0.6166	0.1115	Accept	0.4702	Accept
12	MK48 ADCAP	0.0639	0.0880	Accept	0.6113	Accept
13	E-6(Tacamo)	0.4509	0.1467	REJECT	0.6786	REJECT
14	Avenger	0.0469	0.1800	REJECT	0.9372	REJECT
15	PLS	0.0984	0.3444	REJECT	2.4180	REJECT
16	RSIP	0.0356	0.0793	Accept	0.4527	Accept
17	Longbow	0.0822	0.0623	Accept	0.2289	Accept
18	CMU	0.4497	0.0634	Accept	0.2642	Accept
19	AOE6	0.0665	0.1400	REJECT	0.8563	REJECT
20	TRITAC	0.0858	0.0538	Accept	0.2943	Accept
21	MLRS-TGW	0.4975	0.0873	Accept	0.3815	Accept
22	JSOW	0.4943	0.1376	REJECT	0.6395	REJECT
23	ASAT	0.6586	0.0801	Accept	0.3096	Accept
24	ADDS	0.0551	0.0639	Accept	0.4575	Accept
25	LCAC	0.1306	0.1025	Accept	0.5828	Accept
26	LSD41	0.0445	0.0980	Accept	0.6382	REJECT
27	MK 50	0.3247	0.0726	Accept	0.3457	Accept
28	Backscatter	0.0428	0.0954	Accept	0.6647	REJECT
29	Peacekeeper	0.0371	0.2476	REJECT	1.4433	REJECT
30	T46A	0.7087	0.2113	REJECT	1.0080	REJECT
31	Cvhelo	0.0417	0.1710	REJECT	0.9907	REJECT
32	TAO187Oiler	0.1102	0.5247	REJECT	5.0072	REJECT
33	FDS	0.4766	0.1012	Accept	0.4135	Accept
34	ATARS	0.5331	0.1580	REJECT	0.5654	Accept
35	SRAMII	0.6240	0.1632	REJECT	0.7115	REJECT
36	ATACMS	0.5088	0.1112	Accept	0.4720	Accept
37	TOW2	0.1494	0.1128	Accept	0.7785	REJECT

Regression Analysis

Regression is a mathematical tool used to describe a future response. It is based upon the correlation of the independent and dependent variables. Correlation demonstrates a mathematical relationship between variables, but does not establish cause and effect relationships.

This regression analysis includes two types of independent variables, quantitative and qualitative (McClave et al., 1998:579). Quantitative variables are measured on a continuous numerical scale. Variables that represent categorical values are qualitative. We use qualitative variables to express a situation expressed in discrete terms. For example "Weibull Location Indicator" is a qualitative variable since it is a binary variable that describes whether the "Weibull Location" is greater than zero or not.

This study uses regression to establish a relationship between the independent variables collected through the cost model parameter estimation and goodness of fit processes and the dependent variables of cost growth and schedule growth. We fit the data to the linear model using a least squares approach.

Full Regression Model. The first step in creating the model is to determine relationships between the dependent variables and the independent variables. We identify the relationship graphically by plotting each of the independent variables against the dependent variable. We identify an approximate mathematical relationship by looking for general trends in the plots. The plots of our data in Appendix B identify possible linear relationships for certain variables, but no discernable higher order relationships.

The full model includes all possible combinations of variables, including categorical and interaction terms. The full model equation

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \ldots\ldots + \beta_i x_i + \varepsilon$$

where y is the dependent variable, the β_i represent parameters with unknown values and the x_i represent independent variables. The x_i may represent higher order or interaction terms such as x_k^2, or $x_m x_n$.

Since the Weibull and Rayleigh parameters describe similar models to the same dataset, a high degree of correlation exists between them. To avoid multicollinearity, a phenomenon that indicates multiple variables are explaining the same phenomena, we separate the Rayleigh parameters and Weibull parameters into different full regression models. Therefore, we evaluate four full models: Cost growth with Weibull parameters, cost growth with Rayleigh parameters, schedule growth with Weibull parameters, and schedule growth with Rayleigh parameters. In Table 3-12, we list the variables contained in our full models.

Table 3-12. Regression Full Model Variables

Cost Variables - Set 1	Cost Variables - Set 2
$ Cost Overrun (dependent)	$ Cost Overrun (dependent)
Yrs Initial SAR	Yrs Initial SAR
Weibull Cost Factor	Rayleigh Cost Factor
Weibull Scale	Rayleigh Location
Weibull Shape	Rayleigh Scale
Weibull Location	Kolmogorov
Kolmogorov	Cramer-von Mises
Cramer-von Mises	Anderson-Darling
Anderson-Darling	
Interaction Weibull Shape / Weibull Scale	

Scheduel Variables - Set 1	Schedule Variables - Set 2
% Schedule Overrun (dependent)	$ Cost Overrun (dependent)
Yrs Initial SAR	Yrs Initial SAR
Weibull Cost Factor	Rayleigh Cost Factor
Weibull Scale	Rayleigh Location
Weibull Shape	Rayleigh Scale
Weibull Location	Kolmogorov
Kolmogorov	Cramer-von Mises
Cramer-von Mises	Anderson-Darling
Anderson-Darling	
Interaction Weibull Shape / Weibull Scale	

Model Reduction. Model reduction is the process of eliminating variables that do not add explanatory capability to the model. We systematically remove variables from the full model based on their relative contribution to the regression model. This reduction iteratively removes the least significant variables, until the overall significance of the model is reduced. We test the each reduced model to determine if the reduced model is statistically equivalent to the model before the reduction.

We select variables for removal by examining their observed significance level or p-value. Statistical programs, like JMP™, provide a p-value for each variable that may be used to make this evaluation. If the p-value is greater than desired significance level α, the variable may be removed and the newly reduced model tested. (McClave et al., 1998: 452-453). We reduce the full model, until the reduced model is not statistically equivalent to the model before reduction.

From the full models in Table 3-12 we iteratively reduced the number of variables, by evaluating each variable's contribution to the model. We considered each variable's p-value, an indicator of the explanatory value of the variable. Generally, we eliminated variables with a p-value greater than 0.05, provided their removal does not adversely affect the overall statistics of the model, as determined by an F-test.

We use the F-test to determine if the reduced model is statistically equivalent to the model before reduction. The test statistic for the F-test is

$$F = \frac{(SSE_r - SSE_f)/(\beta_f - \beta_r)}{SSE_f / df_f}$$

where $SSEr$ is the reduced model sum of squared errors, SSE_f is the sum of squared errors for the model before reduction, β_f is the number of parameters in the model before reduction, β_r is the number of parameters in the reduced model, and df_f is the degrees of freedom for the full model before reduction. The rejection region for this test statistic is $F > F_\alpha$. In other words, if F is less than F_α, the reduced model is statistically as good as the model before reduction. F_α is based on $v_1 = (\beta_f - \beta_r)$ numerator degree of freedom (df) and $v_2 = df_f$ denominator df (McClave et al., 1998: 600). We chose α to be 0.05, for 95% confidence in our models.

After creating the model, we evaluated how well the model describes the relationship between dependent and independent variables using the coefficient of multiple determination, R^2. However, R^2 alone can be misleading, since R^2 depends on the model, sample size and sample. Adding more independent variables to the model always increases R^2. The measure of adjusted R^2 accounts for the number of explanatory variables relative to the number of observations. Therefore, we use adjusted R^2 generally for comparison between statistical models, while we use R^2 to evaluate the explanatory capacity of a given model. While we prefer to have R^2 as close to 1.0 as possible, using real world data can result in relatively low R^2.

Arriving at the reduced form of the model, we evaluate how well it subscribed to the basic error term assumptions of normality, constant variance and independence. Each of the models essentially passed the tests for the basic assumptions. We detail these results in the model discussion section.

We also evaluated each of the final models for outliers. Cook's distance or Cook's D Influence measures the overall influence of each observation on the regression coefficients, including the intercept. Cook proposes that the influence of the ith data point be measured by the distance D_i where

$$D_i = \{\mathbf{b} - \mathbf{b}(i)\}' \, \mathbf{X}' \, \mathbf{X}\{\mathbf{b} - \mathbf{b}(i)\}/(ps^2)$$

where $\mathbf{X}$ is n x p, $\mathbf{b}$ is the usual least squares estimator, $s2$ is the variance of the error, and $\mathbf{b}(i)$ is the least squares estimator obtained after the ith data point is removed. Cook's D provides useful information in identifying outliers, since it checks whether the deletion of one or more critical observations greatly affects the model (Draper, 1981:170). Data that have a significant influence on the overall model may be candidates for removal as outliers. On the several occasions when we identified an outlier, we removed the data point and recalculated the regression model.

The dependent variable of cost growth is the actual cost overrun in constant year 2000 dollars. For example, the RSIP program initially budgeted $449.8 CY$00M and the final program showed $477.4 million CY$00M, a cost growth is $27.6 CY$00M. All cost numbers in the dataset are in millions of constant year 2000 dollars. A negative cost growth indicates that the final program cost less than originally planned. The 37 programs show a range of cost growth from -$688.6M to $1199.3M, with an average cost growth of $40.3M. The final cost model is based on a dataset with B-1B data removed. We removed the B1-B data because the residual plot indicates it as an extreme outlier that significantly influences the model parameter estimates. The next highest cost growth to the B1-B's $1199.3 was the AAMRAAM-J's at $500.9 (see Table 3-2).

Final Cost Model. The cost growth model is

$ Cost overrun = -0.0558 Cost of Initial SAR + 23.92 Weibull Scale – 59.36

The cost model sensitivity shown in Table 3-13 indicates the total contribution of each variable to the predicted cost growth. Our model indicates that the Weibull Scale parameter has the most significant contribution towards cost growth on average. The scale parameter is offset by the contribution of the initial budget variable.

Table 3-13. Cost Model Sensitivity

Variable	Coeff		Inputs	Outputs
Cost Initial SAR	-0.055778	Min	15.482	-0.8635787
		Max	13686.313	-763.3951927
		Avg	1252.450	-69.8591463
		Med	521.738	-29.1014807
Weibull Scale	23.920513	Min	1.056	25.2610278
		Max	13.952	333.7271738
		Avg	5.760	137.7898914
		Med	5.295	126.6650761

Figure 3-6 shows the statistics for the final cost model. Each of the variables is significant; the p-value in the Parameter Estimate section is less than 0.05. The model is significant since the F-ratio probability has a value less than 0.05. The model explains 53.4% of the total cost growth variation in the dataset.

Response: $ Cost Overrun

Summary of Fit

RSquare	0.533828
RSquare Adj	0.505576
Root Mean Square Error	155.0953
Mean of Response	8.132778
Observations (or Sum Wgts)	36

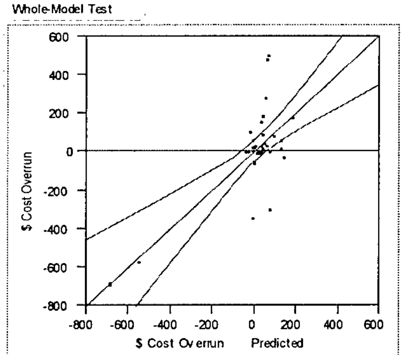
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-59.35861	57.61211	-1.03	0.3104
Cost Initial	-0.055778	0.009771	-5.71	<.0001
Wscale	23.920513	8.869031	2.70	0.0109

Effect Test

Source	Nparm	DF	Sum of Squares	F Ratio	Prob>F
Cost Initial	1	1	783842.68	32.5860	<.0001
Wscale	1	1	174979.07	7.2743	0.0109

Whole-Model Test

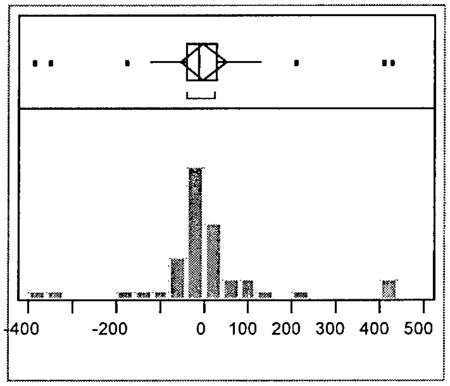


Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	909007.2	454504	18.8947
Error	33	793800.6	24055	Prob>F
C Total	35	1702807.9		<.0001

Figure 3-6. Final Cost Model Statistics

Residual $ Cost Overrun



Quantiles

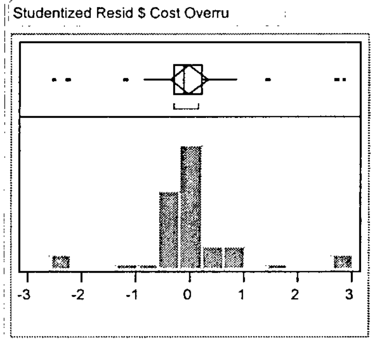
maximum	100.0%	429.34
	99.5%	429.34
	97.5%	429.34
	90.0%	156.08
quartile	75.0%	31.61
median	50.0%	-12.98
quartile	25.0%	-39.28
	10.0%	-138.80
	2.5%	-382.11
	0.5%	-382.11
minimum	0.0%	-382.11

Moments

Mean	-0.0000
Std Dev	150.5989
Std Error Mean	25.0998
Upper 95% Mean	50.9550
Lower 95% Mean	-50.9550
N	36.0000
Sum Weights	36.0000

Figure 3-7. Final Cost Model Residual Analysis

Studentized Resid $ Cost Overru



Quantiles

maximum	100.0%	2.8400
	99.5%	2.8400
	97.5%	2.8400
	90.0%	1.0354
quartile	75.0%	0.2148
median	50.0%	-0.1136
quartile	25.0%	-0.2752
	10.0%	-0.9327
	2.5%	-2.5285
	0.5%	-2.5285
minimum	0.0%	-2.5285

Moments

Mean	-0.00339
Std Dev	0.99568
Std Error Mean	0.16595
Upper 95% Mean	0.33350
Lower 95% Mean	-0.34028
N	36.00000
Sum Weights	36.00000

Test for Normality

Shapiro-Wilk W Test

W	Prob<W
0.859862	0.0002

Figure 3-8. Final Cost Model Studentized Residual Analysis

Figures 3-7 and 3-8 show the cost model residual plot and Studentized residual plot, respectively. The intent of these plots is to indicate whether the model satisfies the assumption that the residuals are normally distributed. As indicated by Figures 3-7 and 3-8, the residuals for the cost model are too heavy in the tails of the distribution to be considered normal. In other words, the presence of the relatively tall bars near three standard deviations indicates that the residuals may not be normally distributed. On the other hand, Figures 3-7 and 3-8 indicate the residuals are not highly skewed; they are fairly symmetric. The cost model, however, fails the requirement for normality by inspection of the residual plots and also according to the Shapiro-Wilk test. If the Shapiro-Wilk p-value reported, shown as Prob<W, is less than .05 (or some other alpha), then you conclude that the distribution is not normal. However, regression is robust with respect to this condition, so the model is still acceptable. Normality primarily drives confidence intervals. Since we do not make predictions with confidence intervals, we accept the model.

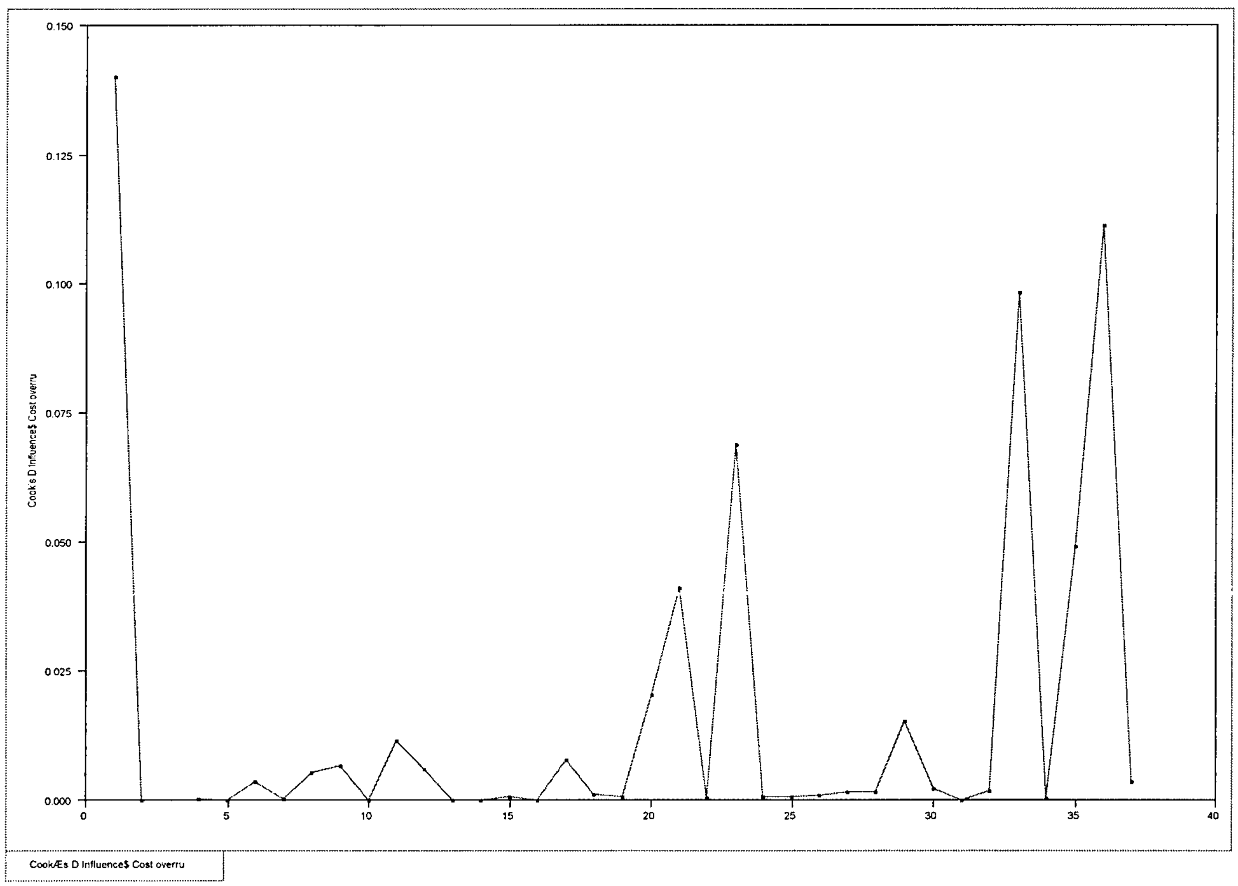


Figure 3-9. Final Cost Model Cook's D Plot

The Cook's D test for influential points reveals that no data in the final model significantly influence the results. Generally, we consider a point with a Cook's D value greater than 1.0 to be influential. Figure 3-9 shows that none of the data in the final model significantly influence the model.

Final Schedule Model. The final schedule model is based on a dataset with Trident Submarine removed. We removed the Trident Submarine program because it had a schedule slip of 3.2, a clear outlier since no other slips were nearly as large, as seen in Table 3-2. The schedule growth model is the following:

$$
\begin{gathered}
\%\ \text{Schedule Growth} = 0.00003\ \text{Initial Budget} - 0.047\ \text{Weibull Scale} \\
+ 0.22\ \text{Weibull Shape} - 2.25\ \text{Interaction Weibull Shape / Scale} \\
- 1.24\ \text{Kolmolgorov} + 8.69\ \text{Cramer Von Mises} \\
- 0.45\ \text{Anderson-Darling} + 1.54
\end{gathered}
$$

The schedule model sensitivity in Table 3-14 indicates the total contribution of each variable to the predicted schedule growth. Our model indicates that the Weibull Shape and Cramer-von Mises variables have the most significant contribution towards positive cost growth on average. The interaction term has the greatest offset contribution. The positive schedule growth represented by Cramer-von Mises variable is offset by both the Kolmolgorov and Anderson-Darling variable contributions. This model explains 50.5% of the variation in schedule growth for the 36 included programs.

Table 3-14. Schedule Model Sensitivity

Variable	Coeff		Inputs	Outputs
Cost Initial SAR	0.0000025	Min	15.482	0.0000387
		Max	13686.313	0.0342158
		Avg	1395.598	0.0034890
		Med	576.703	0.0014418
Weibull Shape	0.2207517	Min	1.089	0.2404470
		Max	4.582	1.0114757
		Avg	2.666	0.5884917
		Med	2.473	0.5459822
Weibull Scale	-0.047334	Min	1.572	-0.0743941
		Max	13.952	-0.6603806
		Avg	5.930	-0.2806735
		Med	5.317	-0.2516685
Interaction Shape / Scale	-2.24977	Min	0.241	-0.5417110
		Max	0.825	-1.8568785
		Avg	0.492	-1.1074555
		Med	0.509	-1.1445318
KS	-1.236105	Min	0.031	-0.0389020
		Max	0.821	-1.0147159
		Avg	0.310	-0.3826174
		Med	0.237	-0.2928824
CVM	8.69238	Min	0.050	0.4356635
		Max	0.272	2.3654553
		Avg	0.129	1.1186659
		Med	0.116	1.0116808
AD	-0.451847	Min	0.245	-0.1108735
		Max	1.557	-0.7036010
		Avg	0.599	-0.2707557
		Med	0.542	-0.2447851

Response: Sched Slip

Summary of Fit

RSquare	0.504712
RSquare Adj	0.38089
Root Mean Square Error	0.212852
Mean of Response	1.204167
Observations (or Sum Wgts)	36

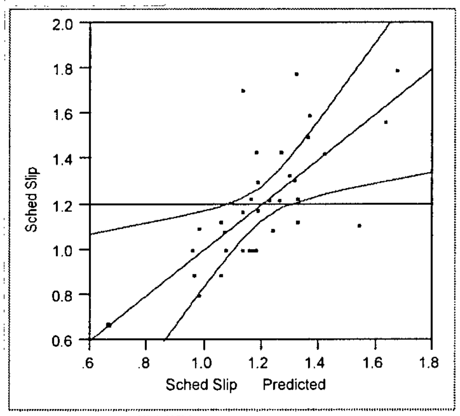
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1.5149604	0.288802	5.25	<.0001
Cost Initial	0.000025	0.000014	1.76	0.0888
W scale	-0.047334	0.039075	-1.21	0.2359
W Shape	0.2207517	0.126301	1.75	0.0915
W Interaction	-2.24977	0.804842	-2.80	0.0093
Kol	-1.236105	0.343533	-3.60	0.0012
CVM	8.69238	2.51408	3.46	0.0018
AD	-0.451847	0.118422	-3.82	0.0007

Effect Test

Source	Nparm	DF	Sum of Squares	F Ratio	Prob>F
Cost Initial	1	1	0.14087533	3.1094	0.0888
W scale	1	1	0.06648206	1.4674	0.2359
W Shape	1	1	0.13840519	3.0549	0.0915
W Interaction	1	1	0.35400653	7.8137	0.0093
Kol	1	1	0.58658455	12.9472	0.0012
CVM	1	1	0.54159610	11.9542	0.0018
AD	1	1	0.65958637	14.5585	0.0007

Whole-Model Test



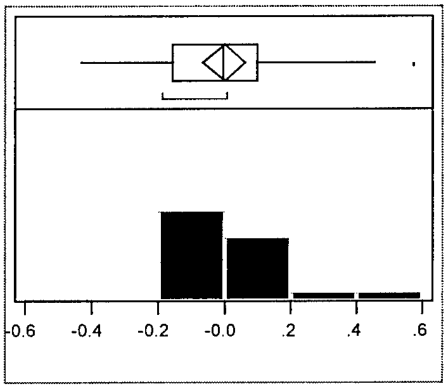
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	7	1.2927057	0.184672	4.0761
Error	28	1.2685693	0.045306	Prob>F
C Total	35	2.5612750		0.0034

Figure 3-10. Final Schedule Model Statistics

Figure 3-10 shows the statistics for the final schedule model. The inclusion of the interaction term "Weibull Shape/Weibull Scale" necessitates the inclusion of both the individual Weibull Shape and Weibull Scale parameters, although neither is significant to the 0.05 level. We included the "Cost Initial" variable because it affected the explanatory ability of the model significantly. All other variable have p-values less than 0.05, as shown in the Parameter Estimate Section. The model is significant since the F-ratio probability has a value less than 0.05. The model explains 50.5% of the total variation of the schedule growth.

Residual Sched Slip

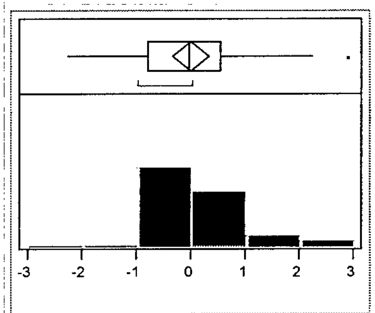


Quantiles		
maximum	100.0%	0.57517
	99.5%	0.57517
	97.5%	0.57517
	90.0%	0.24160
quartile	75.0%	0.10257
median	50.0%	-0.0045
quartile	25.0%	-0.1544
	10.0%	-0.1826
	2.5%	-0.4299
	0.5%	-0.4299
minimum	0.0%	-0.4299

Moments	
Mean	-0.00000
Std Dev	0.19038
Std Error Mean	0.03173
Upper 95% Mean	0.06442
Lower 95% Mean	-0.06442
N	36.00000
Sum Weights	36.00000

Figure 3-11. Final Schedule Model Residual Analysis

Studentized Resid Sched Slip



Quantiles		
maximum	100.0%	2.8801
	99.5%	2.8801
	97.5%	2.8801
	90.0%	1.1974
quartile	75.0%	0.5735
median	50.0%	-0.0234
quartile	25.0%	-0.7876
	10.0%	-0.9237
	2.5%	-2.2213
	0.5%	-2.2213
minimum	0.0%	-2.2213

Moments	
Mean	0.01326
Std Dev	0.98118
Std Error Mean	0.16353
Upper 95% Mean	0.34525
Lower 95% Mean	-0.31872
N	36.00000
Sum Weights	36.00000

Test for Normality	
Shapiro-Wilk W Test	
W	Prob<W
0.950232	0.1406

Figure 3-12. Final Schedule Model Studentized Residual Analysis

Figures 3-11 and 3-12 show the schedule model residual plot and Studentized residual plot, respectively. The intent of these plots is to indicate whether the model satisfies the assumption that the residuals are normally distributed. Figures 3-7 and 3-8 indicate that the residuals indicate are fairly symmetric and appear to be normally distributed. The schedule model passes the Shapiro-Wilk test for normality with a Prob<W of 0.1406. Therefore, the schedule model passes the normality assumption.

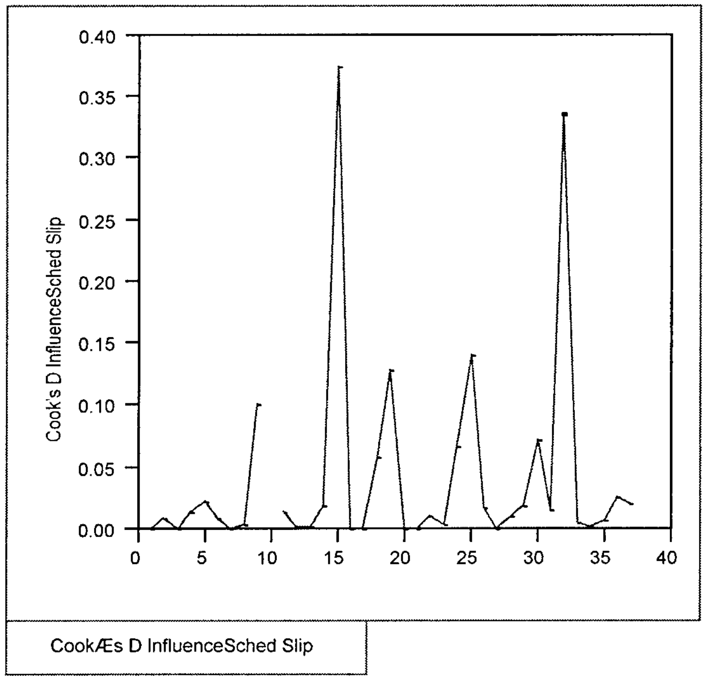


Figure 3-13. Final Schedule Model Cook's D Plot

The Cook's D test for influential points reveals that no data in the final model are significantly influencing the results. Generally, we consider a point with a Cook's D value greater than 1.0 to be influential. Figure 3-9 shows that none of the data in the final model significantly influence the model.

Regression Summary. Belcher and Dukovich identified 12 factors in three areas contributing to development program cost growth and schedule growth. (See Figure 2-7.) Our cost and schedule models each account for only a single of Belcher and Dukovich's factors, funding constraints. Yet, these models explain 53.4% of the variation in cost growth and 50.5% of the variation in schedule growth. Appendix B contains a summary of all of the results for each of the 37 programs.

Model Validation. We use validation to evaluate the model accuracy. Validation uses program information outside the data used to create the model to test the model's descriptive ability. We excluded each of the programs in Table 3-15 from the original dataset because they are still ongoing programs as of the date of our SAR data. Based on the estimated completion date and cost, we determined that each of the programs is greater than 80% complete, making them useful for model validation. We used estimated completion cost for the six programs shown in Table 3-15 to validate the model.

Table 3-15. Near Completed Programs for Validation

Program	Cost Initial	Yrs Initial SAR	Weibull Scale	Weibull Shape	Weibull Location	Weibull Interaction	Kol	CVM	AD
JTIDS-J	1741.92	21	12.69	3.56	1.61	0.28	0.6097	0.0856	0.4066
DSCSIII	328.26	15	1.92	1.04	1.56	0.54	0.1114	0.2905	1.6495
SADARM	812.57	9	4.96	2.42	0.00	0.49	0.1402	0.1162	0.4045
DMSP	685.88	17	6.60	1.49	1.61	0.23	0.0479	0.0637	0.4321
T45TS	784.28	12	6.46	3.32	2.16	0.51	0.7167	0.1467	0.5941
WISWAM	887.23	8	6.38	3.47	0.00	0.54	0.4402	0.1310	0.5603

We used a statistical technique to calculate the variance of the new observation $Y_{h(new)}$ corresponding to $\mathbf{X}_h$, the specified values of the X variables. The variance formula is

$$S^2(Y_{h(new)}) = MSE(1+\mathbf{X}'_h(\mathbf{X}'\mathbf{X})^{-1}\,\mathbf{X}_h$$

where X is the matrix of program input variables, $\mathbf{X}'$ is its transpose, $\mathbf{X}_h$ is the vector containing the values for the new data point variables, $\mathbf{X}'_h$ is its transpose, and MSE is the Mean Squared Error of the original regression.

Cost Model Validation. From the six programs shown in Table 3-15, we calculated cost growth using the cost model. Each of the six programs is within the range of data used to construct the regression model. Table 3-16 shows the comparison of actual to predicted cost growth. Table 3-16 also shows the variance associated with the regression prediction, along with "z" which represents the normalized value. The model appears to predict too little cost growth in four of six cases, with an overall average of $87.38 CY$00M. The bias appears to be about 0.56 standard deviations.

Table 3-16. Cost Model Validation

Program	Actual Cost growth	Predicted Cost growth	Actual Minus Predicted	Variance	z
JTIDS-J	-166.94	143.72	-310.66	168.57	-1.843
DSCSIII	547.71	-31.89	579.60	160.98	3.600
SADARM	-333.08	13.70	-346.78	157.44	-2.203
DMSP	93.76	46.48	47.28	157.41	0.300
T45TS	338.65	49.18	289.47	157.42	1.839
WISWAM	308.27	42.90	265.37	157.38	1.686

Schedule Model Validation. From the six programs shown in Table 3-15, we calculated schedule growth using the schedule model. Three of the six programs are within the range of data used to construct the regression model. At least one variable of the JTIDS, DSCSIII, and DMSP programs falls outside the range of the data used to construct the model. Table 3-17 shows the comparison of actual to predicted schedule growth, the variance associated with the regression prediction, and the normalized value. The model appears to predict too little schedule growth in five of six cases, with an overall average of 0.230. The bias appears to be about 0.30 standard deviations.

Table 3-17. Schedule Model Validation

Actual Sched growth	Pred Sched growth	Actual Minus Predicted	Variance	z
1.24	0.91	0.326	1.384	0.235
1.47	2.08	-0.609	0.385	-1.582
1.89	1.39	0.498	0.488	1.019
1.41	1.34	0.074	0.835	0.089
1.50	0.92	0.580	0.531	1.094
1.63	1.12	0.509	0.540	0.943

Model Confirmation of Theory

We hypothesize that Rayleigh budget profiles perform better than the profiles actually used. In order to test this hypothesis, we created Rayleigh efficient expenditure profiles using the method described by Lee, Hogue and Gallagher (1997). Given the final costs from the 37 initial programs, we create a Rayleigh expenditure profile.

We then applied the procedures outlined earlier in this chapter to identify Weibull parameters and goodness of fit associated with the Rayleigh expenditure profile. Table 3-18 identifies the parameters for the Rayleigh expenditures. Using the resulting parameters as input variables, we compare the expected performance of the Rayleigh expenditures to the actual program performance.

Table 3-19 shows the results of the comparison for the cost model. In seven of the 37 programs, the Rayleigh expenditures resulted in lower predicted cost overruns. In the thirty other programs, the Rayleigh expenditures resulted in greater cost growth. Table 3-20 shows the results of the comparison for the schedule model. In five of the 37 programs, the Rayleigh expenditures resulted in lower predicted schedule overruns. In the thirty-two other programs, the Rayleigh expenditures resulted in greater schedule growth.

Table 3-18. Rayleigh Efficient Profiles

#	Program	Cost Init SAR	Yrs Init SAR	W Shape	W Scale	W Shape / W Scale	W Locn	W Locn Dum	Kolmogorov	Cramer-von Mises	Anderson-Darling
		Rayleigh Efficient Profiles									
1	AAMRAAM-J	1230.4	13	2	9.62	0.21	0.00	1	0.00	0.022	0.123
2	ASPJ-J	752.7	15	2	10.69	0.19	0.00	1	0.00	0.018	0.100
3	B-1B	3861.3	7	2	6.41	0.31	0.00	0	0.00	0.050	0.271
4	Battleship	32.2	6	2	5.88	0.34	0.00	1	0.00	0.059	0.321
5	IUS	1230.7	15	2	10.69	0.19	0.00	0	0.00	0.018	0.100
6	KC-135R	162.9	8	2	6.95	0.29	0.00	1	0.00	0.043	0.232
7	Kiowa	339.9	6	2	5.88	0.34	0.00	1	0.00	0.059	0.321
8	Lantirn	654.3	10	2	8.02	0.25	0.00	1	0.00	0.032	0.176
9	Trident-MSL	13686.3	13	2	9.62	0.21	0.00	0	0.00	0.022	0.123
10	Trident-Sub	78.9	5	2	5.35	0.37	0.00	0	0.00	0.072	0.386
11	RPV(Aquila)	1123.9	14	2	10.16	0.20	0.00	0	0.00	0.020	0.111
12	MK48 ADCAP	1522.9	13	2	9.62	0.21	0.00	0	0.00	0.022	0.123
13	E-6(Tacamo)	497.5	8	2	6.95	0.29	0.00	1	0.00	0.043	0.232
14	Avenger	15.5	3	2	4.28	0.47	0.00	0	0.00	0.111	0.592
15	PLS	51.9	5	2	5.35	0.37	0.00	0	0.00	0.072	0.386
16	RSIP	449.8	6	2	5.88	0.34	0.00	0	0.00	0.059	0.321
17	Longbow	847.9	9	2	7.48	0.27	0.00	1	0.00	0.037	0.201
18	CMU	1511.1	20	2	13.36	0.15	0.00	0	0.00	0.011	0.054
19	AOE6	41.1	9	2	7.48	0.27	0.00	1	0.00	0.037	0.201
20	TRITAC	576.7	18	2	12.29	0.16	0.00	1	0.00	0.014	0.073
21	MLRS-TGW	407.6	12	2	9.09	0.22	0.00	1	0.00	0.025	0.138
22	JSOW	404.2	9	2	7.48	0.27	0.00	1	0.00	0.037	0.201
23	ASAT	2112.4	19	2	12.83	0.16	0.00	0	0.00	0.012	0.064
24	ADDS	267.5	9	2	7.48	0.27	0.00	1	0.00	0.037	0.201
25	LCAC	37.8	7	2	6.41	0.31	0.00	1	0.00	0.050	0.271
26	LSD41	94.4	11	2	8.55	0.23	0.00	1	0.00	0.028	0.155
27	MK 50	2129.8	17	2	11.76	0.17	0.00	1	0.00	0.015	0.082
28	Backscatter	642.2	11	2	8.55	0.23	0.00	1	0.00	0.028	0.155
29	Peacekeeper	9847.5	7	2	6.41	0.31	0.00	0	0.00	0.050	0.271
30	T46A	546.0	10	2	8.02	0.25	0.00	1	0.00	0.032	0.176
31	Cvhelo	81.6	4	2	4.81	0.42	0.00	1	0.00	0.088	0.473
32	TAO187Oiler	24.7	9	2	7.48	0.27	0.00	1	0.00	0.037	0.201
33	FDS	931.6	10	2	8.02	0.25	0.00	1	0.00	0.032	0.176
34	ATARS	203.0	9	2	7.48	0.27	0.00	0	0.00	0.037	0.201
35	SRAMII	1337.4	9	2	7.48	0.27	0.00	1	0.00	0.037	0.201
36	ATACMS	1058.3	11	2	8.55	0.23	0.00	1	0.00	0.028	0.155
37	TOW2	155.8	9	2	7.48	0.27	0.00	0	0.00	0.037	0.201

Table 3-19. Cost Model Confirmation Results

#	Program	Rayleigh Cost overrun	Actual $ Cost overrun	Delta	Rayleigh Better
1	AAMRAAM-J	102.16	500.850	-398.69	Y
2	ASPJ-J	154.31	42.912	111.40	
3	B-1B	-121.30	1199.329	-1320.63	Y
4	Battleship	79.49	2.619	76.87	
5	IUS	127.72	21.987	105.73	
6	KC-135R	97.77	-3.694	101.46	
7	Kiowa	62.33	-0.588	62.92	
8	Lantirn	95.94	152.199	-56.26	Y
9	Trident-MSL	-592.61	-688.615	96.01	
10	Trident-Sub	64.10	22.154	41.95	
11	RPV(Aquila)	120.89	59.157	61.73	
12	MK48 ADCAP	85.84	101.786	-15.94	Y
13	E-6(Tacamo)	79.11	32.614	46.49	
14	Avenger	42.07	1.398	40.67	
15	PLS	65.61	0.306	65.30	
16	RSIP	56.20	27.620	28.58	
17	Longbow	72.34	180.750	-108.41	Y
18	CMU	176.01	174.181	1.83	
19	AOE6	117.34	2.451	114.89	
20	TRITAC	202.55	13.737	188.82	
21	MLRS-TGW	135.26	-28.432	163.69	
22	JSOW	97.10	31.381	65.72	
23	ASAT	129.68	276.789	-147.11	Y
24	ADDS	104.72	87.449	17.27	
25	LCAC	91.97	12.892	79.08	
26	LSD41	139.95	-5.515	145.46	
27	MK 50	103.14	77.069	26.07	
28	Backscatter	109.34	-48.563	157.91	
29	Peacekeeper	-455.20	-576.644	121.44	
30	T46A	101.94	-57.526	159.46	
31	Cvhelo	51.16	-3.120	54.28	
32	TAO187Oiler	118.26	-1.890	120.15	
33	FDS	80.46	479.552	-399.09	Y
34	ATARS	108.32	-7.870	116.20	
35	SRAMII	45.05	-347.508	392.56	
36	ATACMS	86.19	-300.523	386.71	
37	TOW2	110.96	61.400	49.56	

Table 3-20. Schedule Model Confirmation Results

#	Program	Rayleigh Schedule Slip	Actual Schedule Slip	Delta	Rayleigh Better
1	AAMRAAM-J	1.56	1.308	0.25	
2	ASPJ-J	1.52	1.133	0.39	
3	B-1B	1.72	1.429	0.29	
4	Battleship	1.64	1.000	0.64	
5	IUS	1.53	1.600	-0.07	Y
6	KC-135R	1.61	1.000	0.61	
7	Kiowa	1.65	1.167	0.48	
8	Lantirn	1.59	1.500	0.09	
9	Trident-MSL	1.87	1.231	0.64	
10	Trident-Sub	1.67	3.200	-1.53	Y
11	RPV(Aquila)	1.54	1.000	0.54	
12	MK48 ADCAP	1.57	1.231	0.34	
13	E-6(Tacamo)	1.62	1.125	0.49	
14	Avenger	1.76	1.000	0.76	
15	PLS	1.67	1.800	-0.13	Y
16	RSIP	1.65	1.333	0.32	
17	Longbow	1.61	1.222	0.39	
18	CMU	1.46	1.100	0.36	
19	AOE6	1.59	1.111	0.48	
20	TRITAC	1.47	1.222	0.24	
21	MLRS-TGW	1.55	1.083	0.47	
22	JSOW	1.60	1.444	0.15	
23	ASAT	1.49	0.895	0.59	
24	ADDS	1.59	1.778	-0.18	Y
25	LCAC	1.62	1.714	-0.09	Y
26	LSD41	1.56	1.000	0.56	
27	MK 50	1.52	1.176	0.35	
28	Backscatter	1.58	1.091	0.48	
29	Peacekeeper	1.87	1.571	0.30	
30	T46A	1.59	0.800	0.79	
31	Cvhelo	1.71	1.000	0.71	
32	TAO187Oiler	1.59	0.667	0.92	
33	FDS	1.60	1.300	0.30	
34	ATARS	1.59	1.000	0.59	
35	SRAMII	1.62	0.889	0.73	
36	ATACMS	1.59	1.000	0.59	
37	TOW2	1.59	1.444	0.15	

We hypothesize that a Rayleigh budget profile performs better than the profile actually used. However, this hypothesis proved incorrect, as the majority of programs performed worse with a Rayleigh derived budget profile. This leads to the important question of what expenditure profile parameters provide the most efficient program execution, in terms of program growth.

Chapter Summary

This chapter describes the methodology and results associated with testing for a relationship between an initial budget profile and cost and schedule growth. We detail the procedure for converting budgets to outlays and describe the technique for estimating the Weibull and Rayleigh parameters for the outlays. We evaluate if the expenditures are Rayleigh using goodness of fit statistics. We test for a relationship between the above parameters and statistics and program growth with regression. Finally, we test the hypothesis that Rayleigh budget profiles perform better than the actual budget profiles.

IV. Conclusion and Recommendations

Chapter Overview

We present summary finding and conclusions in this chapter. We discuss the general conclusions of this research, including overall model performance and limitations. We also discuss the potential of future research, based on our findings.

Conclusion

We conclude that a statistically significant relationship exists between initial budget profile and cost and schedule growth. Based on the results of the regression models, we explain over fifty percent of the variation for both cost and schedule growth. Furthermore, the results of our model validation indicate that both the cost and schedule models perform acceptably. Five of the six test programs for cost and all six of the test programs for schedule resulted in less than three standard deviations difference from the expected result. Also, we observe from the validation results that both the cost model and schedule model tend to underestimate program growth. The cost model tends to underestimate by an average of $87.4M CY$00, while the schedule model is low by an average of 0.30 standard deviations.

The main model limitation derives from the type of data used to create the model. The SAR database presents budget data in a variety of formats, severely limiting the amount of useable data. Also, the data included in the model include total program budgets, a subset of the contract data used in previous studies. While using program budget data increases the usefulness and robustness of the model, top level data tends to be noisy. Finally, according to the Dukovich and Belcher development cost model, we examined only one of the many potential contributors to development cost. However, we explained 53.4% of the variation in cost growth and 50.5% of the variation in schedule growth. Ideally, we would like to be able to include more of the contributors to increase the accuracy of the models.

Recommendation

The Rayleigh expenditure profile does not appear to be the best, based on these regression models. However, this fact leads to the question of what parameter choices might lead to better performance than the actual performance. Future thesis efforts may identify the Weibull parameters that provide the most efficient program execution, in terms of program growth.

Appendix A. Service Inflation Indices

Appendix Overview

This appendix provides FY00 raw and weighted inflation indices for the Air Force, Army, and Navy.

Table A-1. Air Force Inflation Indices

Air Force FY00 Inflation Rates, RDT&E Appn 3600

FY	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Raw	0.2546	0.2676	0.2799	0.2923	0.3151	0.3491	0.3732	0.3985	0.4256	0.4614
Weighted	0.2594	0.2722	0.2849	0.3006	0.3246	0.3560	0.3865	0.4097	0.4438	0.4789

FY	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
Raw	0.5048	0.5648	0.6168	0.6470	0.6716	0.6944	0.7139	0.7332	0.7552	0.7869
Weighted	0.5327	0.5896	0.6304	0.6596	0.6852	0.7083	0.7255	0.7585	0.7788	0.8163

FY	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Raw	0.8184	0.8535	0.8774	0.9011	0.9192	0.9366	0.9554	0.9754	0.9822	0.9901
Weighted	0.8415	0.8731	0.8985	0.9176	0.9332	0.9508	0.9678	0.9806	0.9869	0.9961

FY	**2000**	2001	2002	2003	2004	2005	2006	2007	2008	2009
Raw	**1.0000**	1.0150	1.0302	1.0457	1.0666	1.0879	1.1097	1.1319	1.1545	1.1776
Weighted	**1.0082**	1.0235	1.0392	1.0569	1.0781	1.0996	1.1216	1.1440	1.1669	1.1903

FY	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Raw	1.2012	1.2252	1.2497	1.2747	1.3002	1.3262	1.3527	1.3798	1.4073	1.4355
Weighted	1.2141	1.2384	1.2631	1.2884	1.3142	1.3404	1.3672	1.3946	1.4225	1.4509

Table A-2. Army Inflation Indices

Army FY00 Inflation Rates, RDT&E

FY	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Raw	0.2577	0.2693	0.2796	0.2910	0.3082	0.3547	0.3750	0.4057	0.4346	0.4741
Weighted	0.2594	0.2722	0.2849	0.3006	0.3246	0.3560	0.3833	0.4168	0.4550	0.5023

FY	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
Raw	0.5243	0.5799	0.6240	0.6489	0.6736	0.6964	0.7160	0.7353	0.7574	0.7892
Weighted	0.5525	0.6021	0.6386	0.6637	0.6876	0.7093	0.7297	0.7524	0.7829	0.8144

FY	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Raw	0.8216	0.8569	0.8826	0.9038	0.9219	0.9394	0.9582	0.9754	0.9822	0.9901
Weighted	0.8452	0.8771	0.8976	0.9188	0.9358	0.9541	0.9710	0.9818	0.9896	0.9991

FY	**2000**	2001	2002	2003	2004	2005	2006	2007	2008	2009
Raw	**1.0000**	1.0150	1.0302	1.0457	1.0666	1.0879	1.1097	1.1319	1.1545	1.1776
Weighted	**1.0119**	1.0271	1.0432	1.0617	1.0829	1.1046	1.1267	1.1492	1.1722	1.1957

FY	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Raw	1.2012	1.2252	1.2497	1.2747	1.3002	1.3262	1.3527	1.3798	1.4073	1.4355
Weighted	1.2196	1.2440	1.2688	1.2942	1.3201	1.3465	1.3734	1.4009	1.4289	1.4575

Table A-3. Navy Inflation Indices

Navy FY00 Inflation Rates, RDT&EN

FY	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Raw	0.2625	0.2760	0.2887	0.3012	0.3253	0.3608	0.3847	0.4060	0.4336	0.4700
Weighted	0.2706	0.2831	0.2959	0.3133	0.3406	0.3706	0.3923	0.4198	0.4520	0.4992

FY	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989
Raw	0.5198	0.5749	0.6186	0.6489	0.6736	0.6965	0.7160	0.7353	0.7574	0.7892
Weighted	0.5520	0.6021	0.6339	0.6624	0.6863	0.7076	0.7281	0.7496	0.7747	0.8073

FY	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Raw	0.8208	0.8561	0.8800	0.9038	0.9219	0.9394	0.9582	0.9754	0.9822	0.9901
Weighted	0.8403	0.8705	0.8958	0.9165	0.9337	0.9517	0.9677	0.9795	0.9874	0.9966

FY	**2000**	2001	2002	2003	2004	2005	2006	2007	2008	2009
Raw	**1.0000**	1.0150	1.0302	1.0457	1.0666	1.0879	1.1097	1.1319	1.1545	1.1776
Weighted	**1.0090**	1.0243	1.0400	1.0580	1.0792	1.1008	1.1228	1.1452	1.1681	1.1915

FY	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Raw	1.2012	1.2252	1.2497	1.2747	1.3002	1.3262	1.3527	1.3798	1.4073	1.4355
Weighted	1.2153	1.2396	1.2644	1.2897	1.3155	1.3418	1.3687	1.3960	1.4240	1.4524

Appendix B. Data Plots

Appendix Overview

This appendix provides the plots of the independent variables against the dependent variables for both the cost and schedule models.

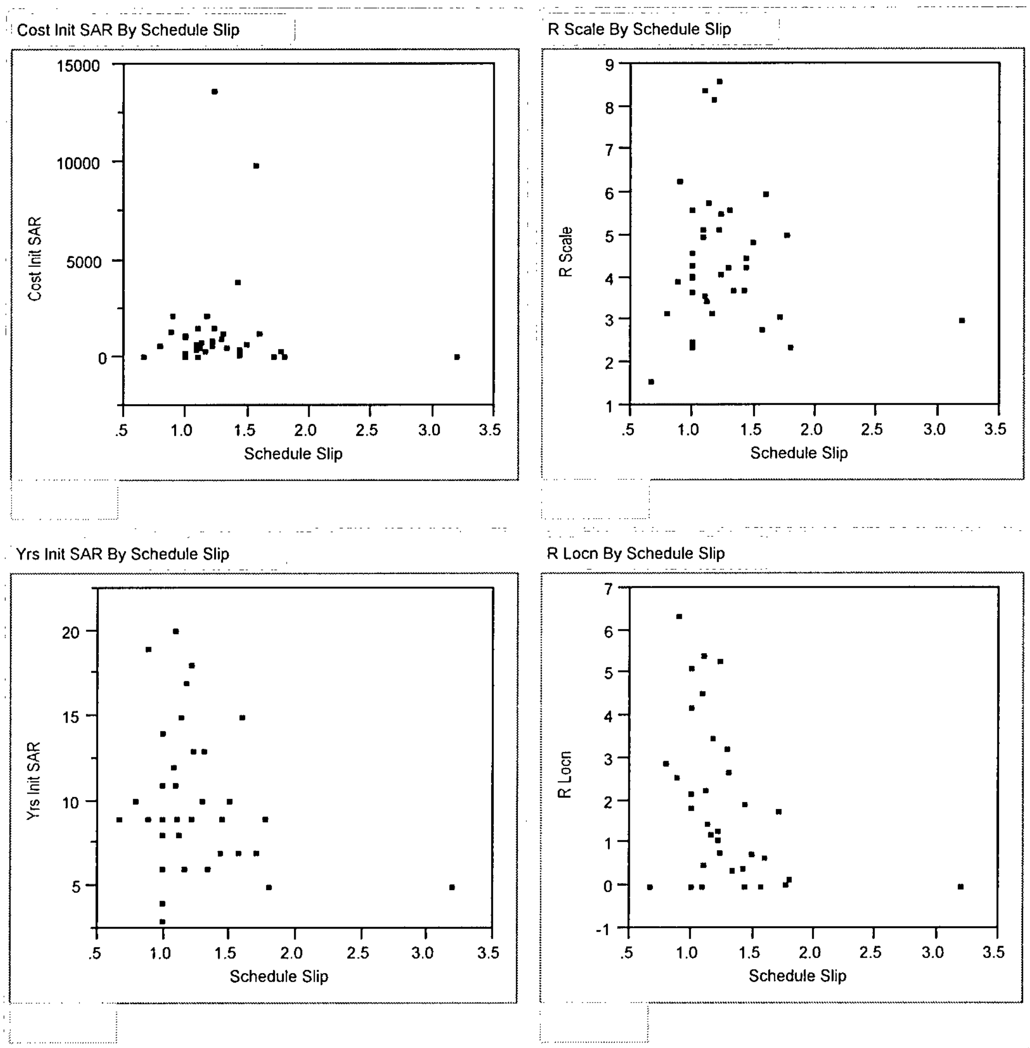


Figure B-1. Schedule Slip Versus Initial Budget Values and Rayleigh Parameters

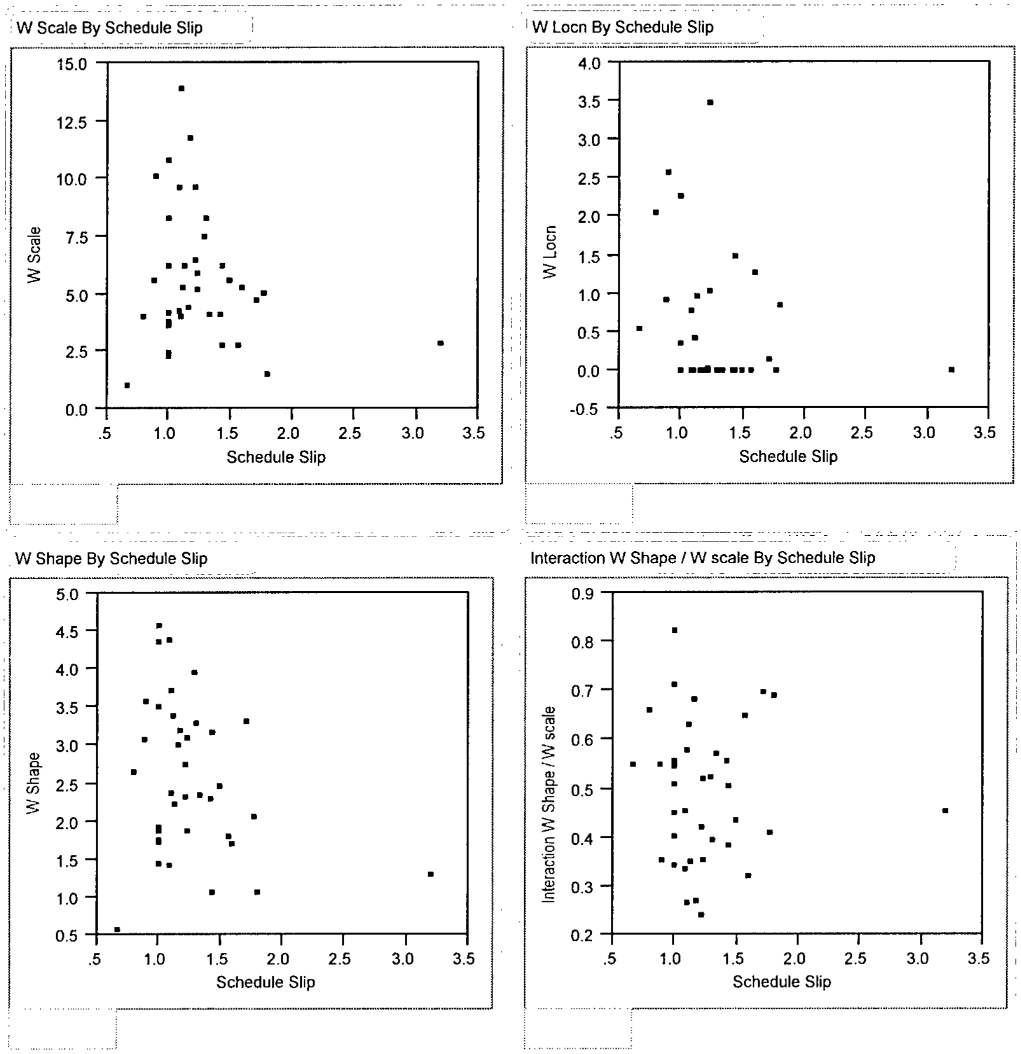


Figure B-2. Schedule Slip Versus Weibull Parameters

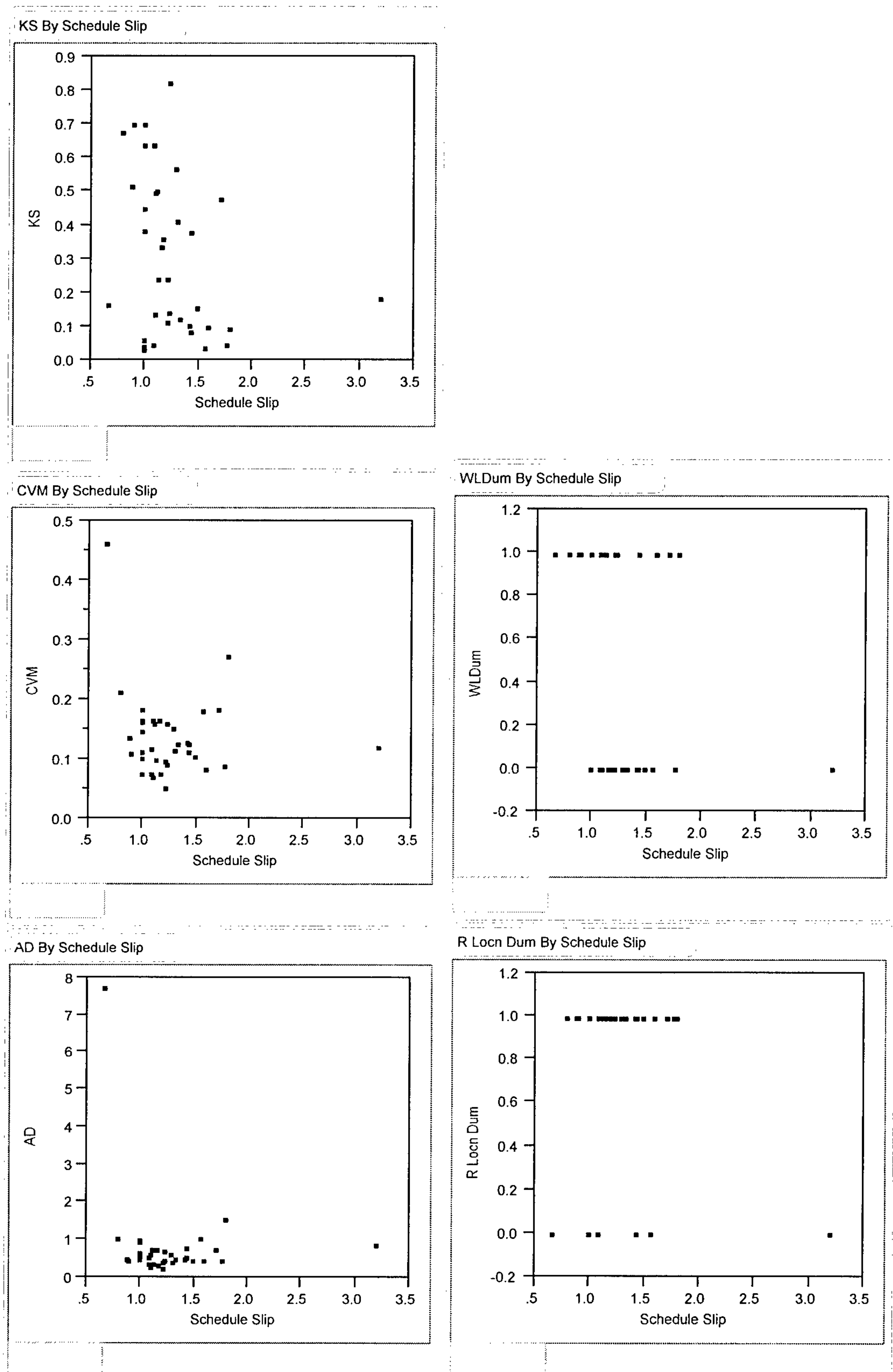


Figure B-3. Schedule Slip Versus Goodness of Fit Statistics and Dummy Variables

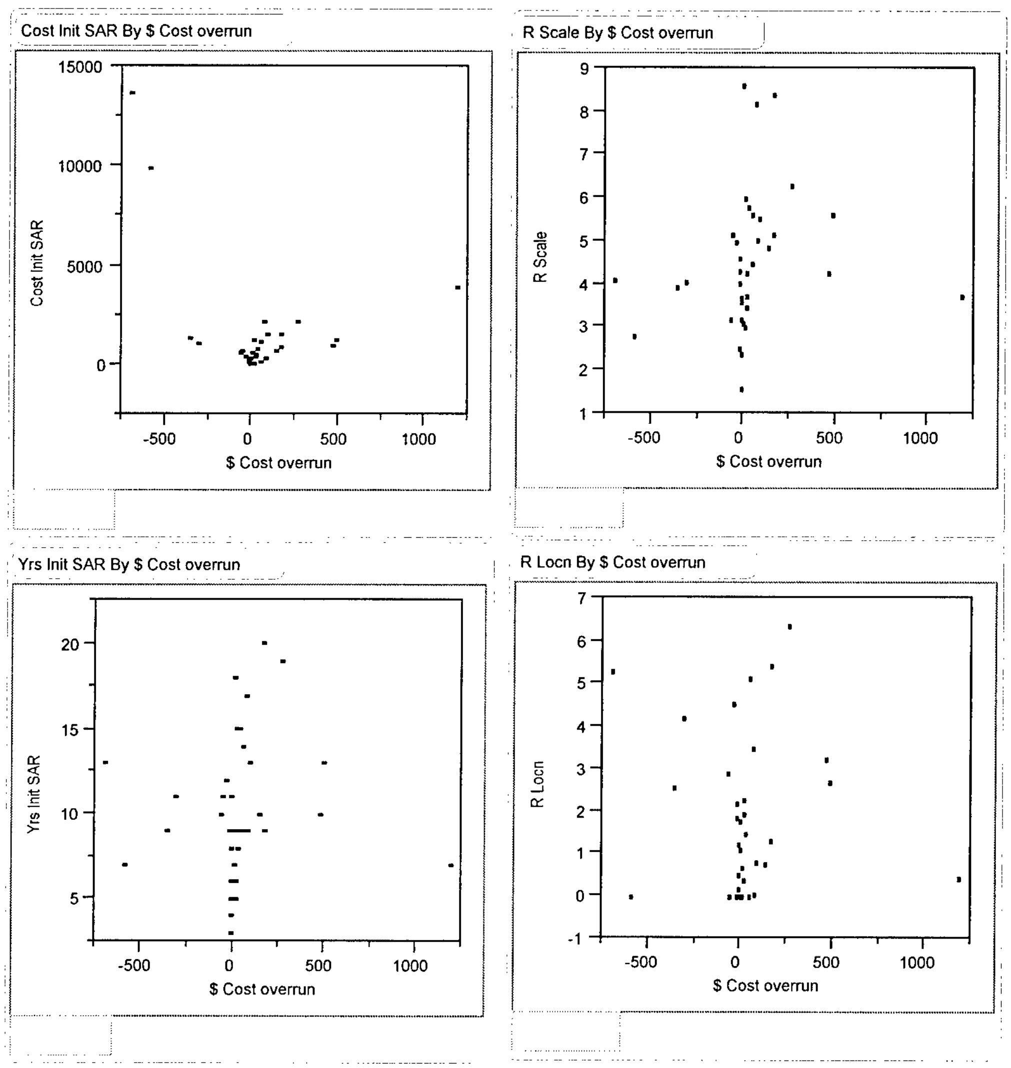


Figure B-4. Cost Growth Versus Initial Budget Values and Rayleigh Parameters

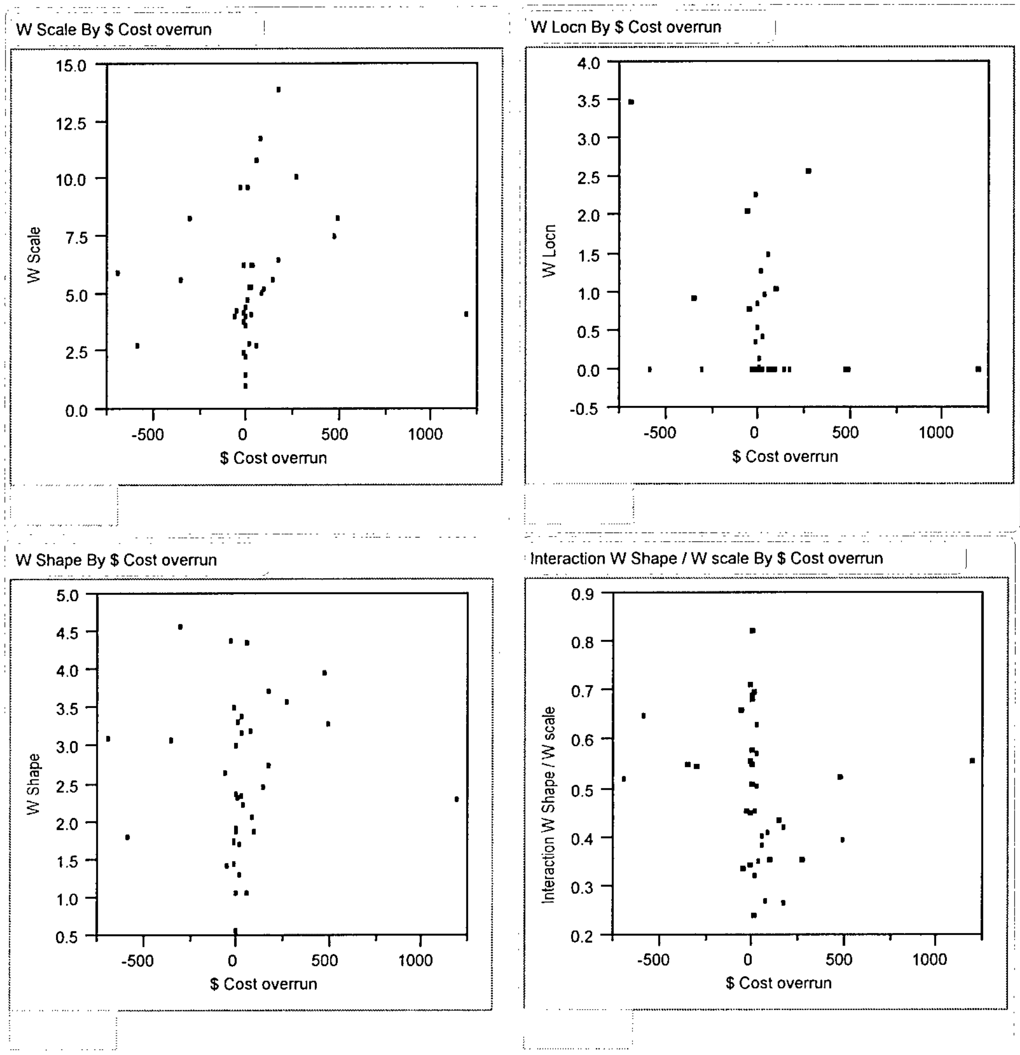


Figure B-5. Cost Growth Versus Weibull Parameters

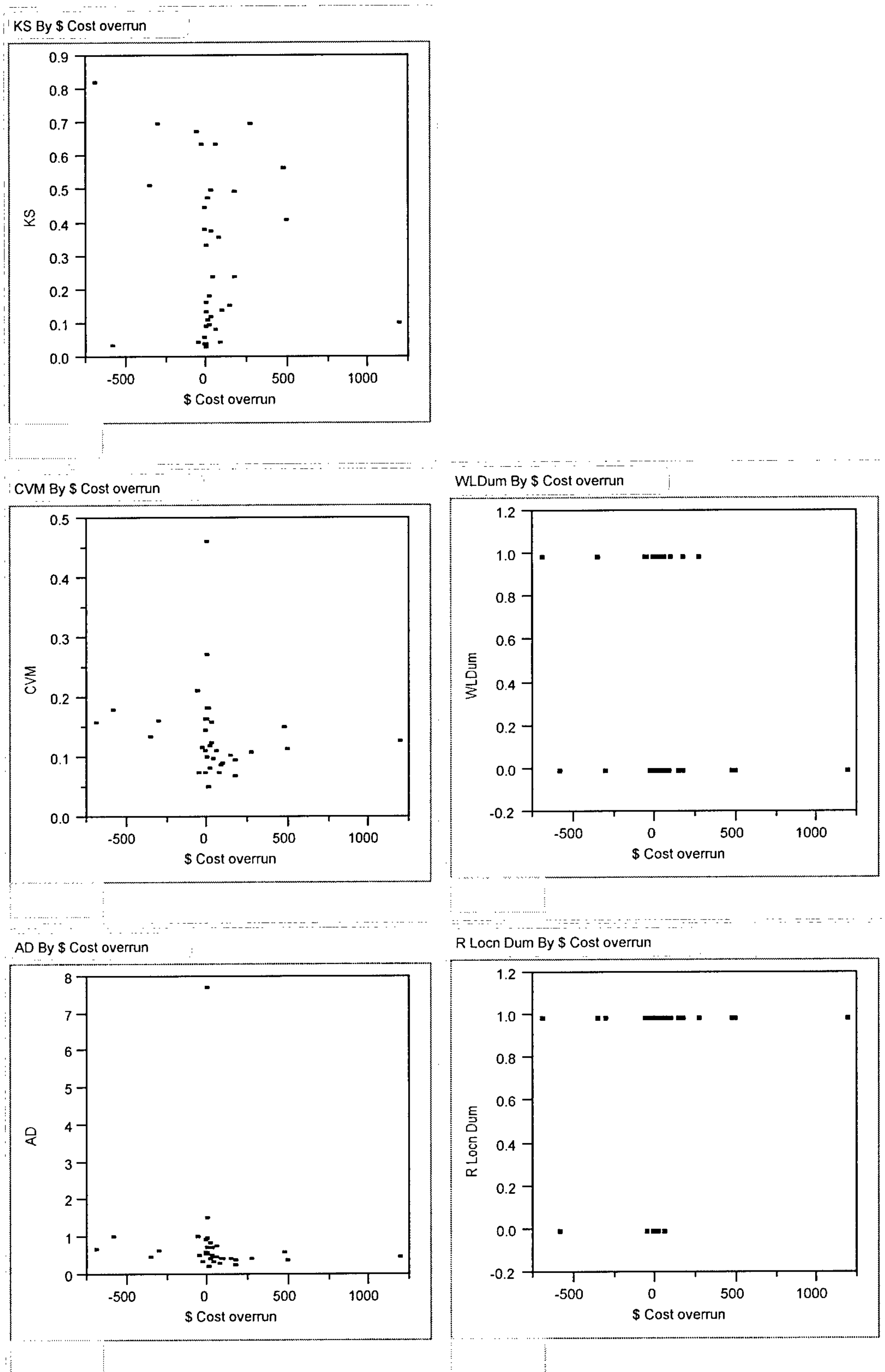


Figure B-6. Cost Growth Versus Goodness of Fit Statistics and Dummy Variables

Appendix C. Program Data

Appendix Overview

This appendix provides program data for the 37 programs included in the analysis.

Table C-1. Program and Cost Growth Data

#	Program	Service	Cost CY00$M Cost Initial SAR	Cost Final SAR	Schedule Years Yrs Initial SAR	Yrs Final SAR	Delta Cost Growth CY00$M	Schedule Slip
1	AAMRAAM-J	AF	1230.37	1731.21	13	17	500.85	1.31
2	ASPJ-J	N	752.68	795.60	15	17	42.91	1.13
3	B-1B	AF	3861.26	5060.59	7	10	1199.33	1.43
4	Battleship	N	32.17	34.79	6	6	2.62	1.00
5	IUS	AF	1230.65	1252.64	15	24	21.99	1.60
6	KC-135R	AF	162.85	159.16	8	8	-3.69	1.00
7	Kiowa	A	339.92	339.33	6	7	-0.59	1.17
8	Lantirn	AF	654.33	806.53	10	15	152.20	1.50
9	Trident-MSL	N	13686.31	12997.70	13	16	-688.61	1.23
10	Trident-Sub	N	78.87	101.02	5	16	22.15	3.20
11	RPV(Aquila)	A	1123.92	1183.08	14	14	59.16	1.00
12	MK48 ADCAP	N	1522.94	1624.72	13	16	101.79	1.23
13	E-6(Tacamo)	N	497.52	530.13	8	9	32.61	1.13
14	Avenger	A	15.48	16.88	3	3	1.40	1.00
15	PLS	A	51.86	52.16	5	9	0.31	1.80
16	RSIP	AF	449.77	477.39	6	8	27.62	1.33
17	Longbow	A	847.89	1028.64	9	11	180.75	1.22
18	CMU	AF	1511.09	1685.27	20	22	174.18	1.10
19	AOE6	N	41.14	43.59	9	10	2.45	1.11
20	TRITAC	AF	576.70	590.44	18	22	13.74	1.22
21	MLRS-TGW	A	407.59	379.16	12	13	-28.43	1.08
22	JSOW	N	404.15	435.53	9	13	31.38	1.44
23	ASAT	A	2112.45	2389.24	19	17	276.79	0.89
24	ADDS	A	267.53	354.97	9	16	87.45	1.78
25	LCAC	N	37.77	50.66	7	12	12.89	1.71
26	LSD41	N	94.40	88.88	11	11	-5.52	1.00
27	MK 50	N	2129.83	2206.90	17	20	77.07	1.18
28	Backscatter	AF	642.17	593.61	11	12	-48.56	1.09
29	Peacekeeper	AF	9847.51	9270.87	7	11	-576.64	1.57
30	T46A	AF	545.95	488.43	10	8	-57.53	0.80
31	Cvhelo	N	81.64	78.52	4	4	-3.12	1.00
32	TAO187Oiler	N	24.66	22.77	9	6	-1.89	0.67
33	FDS	N	931.65	1411.20	10	13	479.55	1.30
34	ATARS	AF	202.97	195.10	9	9	-7.87	1.00
35	SRAMII	AF	1337.36	989.85	9	8	-347.51	0.89
36	ATACMS	A	1058.33	757.81	11	11	-300.52	1.00
37	TOW2	A	155.79	217.19	9	13	61.40	1.44

Table C-2. Rayleigh and Weibull Initial Parameters

#	Program	Weibull W Cost Factor	Weibull W Scale δ	Weibull W Shape β	Weibull W Locn γ	Rayleigh R Cost Factor	Rayleigh R Scale δ	Rayleigh R Locn γ
1	AAMRAAM-J	1217.30	8.34	3.285	0.000	1221.10	5.65	2.610
2	ASPJ-J	745.43	6.36	2.240	0.908	746.15	5.82	1.426
3	B-1B	3821.83	4.13	2.306	0.000	3824.32	3.72	0.389
4	Battleship	31.95	3.65	1.864	0.000	31.92	3.66	0.000
5	IUS	1210.57	5.24	1.699	1.330	1208.64	5.94	0.673
6	KC-135R	160.50	6.28	3.480	0.000	160.82	4.04	2.165
7	Kiowa	338.55	4.42	3.012	0.000	338.96	3.20	1.180
8	Lantirn	647.12	5.64	2.450	0.000	647.97	4.89	0.715
9	Trident-MSL	13556.71	6.04	3.156	3.439	13581.81	4.12	5.283
10	Trident-Sub	78.33	2.90	1.320	0.000	77.91	3.00	0.000
11	RPV(Aquila)	1119.69	10.78	4.355	0.000	1123.94	5.59	5.070
12	MK48 ADCAP	1510.17	5.30	1.883	1.017	1509.45	5.57	0.766
13	E-6(Tacamo)	492.90	5.38	3.405	0.434	493.67	3.49	2.259
14	Avenger	15.49	2.35	1.947	0.000	15.48	2.35	0.000
15	PLS	52.00	1.56	1.089	0.860	51.82	2.36	0.159
16	RSIP	449.29	4.10	2.359	0.000	449.70	3.71	0.375
17	Longbow	849.88	6.52	2.764	0.000	851.52	5.15	1.312
18	CMU	1510.71	13.94	3.744	0.000	1520.90	8.38	5.419
19	AOE6	40.64	4.09	2.371	0.000	40.67	3.59	0.479
20	TRITAC	568.40	9.52	2.293	0.154	569.42	8.55	1.081
21	MLRS-TGW	407.90	9.64	4.416	0.002	409.23	4.98	4.540
22	JSOW	401.24	6.26	3.182	0.000	401.92	4.25	1.935
23	ASAT	2120.72	10.05	3.540	2.653	2130.58	6.31	6.296
24	ADDS	269.62	5.08	2.080	0.000	269.78	5.04	0.037
25	LCAC	37.61	4.64	3.249	0.244	37.65	3.06	1.768
26	LSD41	93.85	4.15	1.448	0.397	93.28	4.56	0.000
27	MK 50	2120.90	11.75	3.177	0.000	2131.96	8.23	3.416
28	Backscatter	638.60	4.32	1.452	0.787	635.66	5.16	0.000
29	Peacekeeper	9765.59	2.78	1.809	0.000	9754.73	2.78	0.000
30	T46A	541.16	4.04	2.663	2.076	541.61	3.18	2.899
31	Cvhelo	80.91	2.48	1.769	0.000	80.81	2.48	0.000
32	TAO187Oiler	24.53	1.05	0.580	0.553	23.95	1.56	0.000
33	FDS	924.49	7.55	3.973	0.000	926.63	4.25	3.211
34	ATARS	202.94	3.83	1.736	2.298	202.78	4.31	1.859
35	SRAMII	1336.37	5.55	3.060	0.996	1338.31	3.91	2.566
36	ATACMS	1066.90	8.36	4.587	0.000	1069.57	4.07	4.187
37	TOW2	156.94	2.77	1.079	1.509	155.95	4.44	0.000

Table C-3. Rayleigh Final Parameters

#	Program	RF Cost Factor	RF Scale δ	RF Locn γ
1	AAMRAAM-J	1715.27	5.21	3.183
2	ASPJ-J	786.46	6.71	1.002
3	B-1B	5010.30	4.82	0.000
4	Battleship	34.55	3.89	0.000
5	IUS	1219.03	6.05	0.618
6	KC-135R	157.14	4.00	2.144
7	Kiowa	338.27	3.30	1.112
8	Lantirn	798.04	5.43	0.905
9	Trident-MSL	12893.61	4.19	5.177
10	Trident-Sub	98.67	4.15	
11	RPV(Aquila)	1183.67	5.85	5.050
12	MK48 ADCAP	1600.74	6.57	0.255
13	E-6(Tacamo)	526.15	3.82	2.212
14	Avenger	16.90	2.47	0.000
15	PLS	51.87	2.61	0.000
16	RSIP	476.04	4.60	0.000
17	Longbow	1032.82	6.39	0.540
18	CMU	1699.12	9.28	5.320
19	AOE6	42.95	4.16	0.175
20	TRITAC	579.70	9.11	0.776
21	MLRS-TGW	381.26	6.29	4.074
22	JSOW	433.76	5.34	3.279
23	ASAT	2418.88	6.92	6.317
24	ADDS	355.27	6.55	0.000
25	LCAC	50.23	5.79	0.496
26	LSD41	87.98	4.24	0.000
27	MK 50	2207.08	8.66	3.283
28	Backscatter	586.93	4.86	0.000
29	Peacekeeper	9174.33	2.69	0.000
30	T46A	484.29	2.94	2.901
31	Cvhelo	77.71	2.42	0.000
32	TAO187Oiler	22.32	1.36	0.000
33	FDS	1405.29	5.78	3.388
34	ATARS	195.05	3.67	2.442
35	SRAMII	990.32	3.54	3.131
36	ATACMS	766.72	4.97	3.209
37	TOW2	217.48	6.77	0.000

Bibliography

Abernathy, T., An Application of the Rayleigh Distribution to Contract Cost Data. MS Thesis, Naval Postgraduate School, Monterey CA, 1984.

Belcher, Gerald J. and John A. Dukovich, "Improved Methods for Estimating Development Costs" Logistics Management Institute, 1999.

Draper, N. R. and H. Smith, Applied Regression Statistics, New York: John Wiley and Sons, 1981.

Gallagher, Mark A. and David A. Lee, "Final Cost Estimates for Research and Development Programs Conditioned on Realized Costs" Military Operations Research Journal: 51-65 (Summer 1996).

Jarvis, Will and Ed Pohl. Lecture Slides, Program Execution Efficiency and Resource Phasing. Department of Defense Cost Analysis Symposium (DODCAS), Washington DC, November 1999.

Lee, David A., M. R. Hogue, and Mark A. Gallagher, "Determining a Budget Profile from a R&D Cost Estimate" Journal of Cost Analysis: 29-41 (Fall 1997).

Norden, P. V., "Useful Tools for Project Management," Operations Research in Research and Development, B. V. Dean, Editor, John Wiley and Sons, 1963.

Outlay Rates to be Used for incremental Changes in BA Purchases (Table 5-11). National Defense Budget Estimates for FY2000 (The Greenbook), office of the Undersecretary of Defense (Comptroller). Department of Defense, Washington DC.

Parr, F. N. "An Alternative to the Rayleigh Curve Model for Software Development Effort," IEEE Transactions on Software Engineering, 6:291-296 (May 1980).

Putnam, Lawrence H., "A General Empirical Solution to the Macro Software Sizing and Estimation Problem," IEEE Transactions on Software Engineering, 4:345-361 (July 1978).

Watkins, Harry, An Application of Rayleigh Curve Theory to Contract Cost Estimation and Control. MS Thesis, Naval Postgraduate School, Monterey CA, 1982. (ADA 118213)

Acknowledgements

In no small way, the work presented in this thesis benefited by the dedication, perseverance, support, and sacrifice of many people around me. My deepest gratitude goes to LtCol Mark Gallagher, who showed me a path through a very difficult journey. To call him a mentor is not sufficient to describe the dimensions of his help. I would like to acknowledge Dr. Tony White for his persistent encouragement and mathematical guidance. (No trend is my friend.) To LtCol William Stockman, I can only say that I stole the pig. Finally, I would like to thank Paul Porter for his humorous views on the serious nature of this topic.

EJU

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 074-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From – To)
20-03-2001	Master's Thesis	Feb 2000 – Mar 2000

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
RELATING INITIAL BUDGET TO PROGRAM GROWTH WITH RAYLEIGH AND WEIBULL MODELS	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER ENR#00-73
Unger, Eric, J., Captain, USAF	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)	8. PERFORMING ORGANIZATION REPORT NUMBER
Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/ENS) 2950 P Street, Building 640 WPAFB OH 45433-7765	AFIT/GAQ/ENS/01M-03
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**
Dr R. P. Burke, Director of Operations Analysis & Procurement Planning Division, OSD PA&E (RA), 1800 Defense, Pentagon, Washington D.C. 20301-1800	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Previous research on completed defense R&D shows that contract expenditures can be fit well with a Rayleigh model. With fixed outlay rates, as prescribed by the OSD comptroller, the budget profile must have most of the funds in the early years to produce Rayleigh-distributed expenditures. R&D programs with more delayed funding profiles may also produce expenditures that a Rayleigh model fits through schedule slips and cost overruns. This research tests how well the initial funding profile produces Rayleigh-distributed expenditures that can be related to the program's final cost overrun and schedule slips. Based only on the initial budget profile, we explain 53.3% of cost overruns and 51.1% of percent schedule slip in 37 completed programs.

15. SUBJECT TERMS
Cost Models, Cost Analysis, Budget Estimates, Cost Estimates, Military Budgets

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Lt Col Mark A. Gallagher, ENS
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (Include area code)
U	U	U	UU	102	(937)255-6565 x4486 Mark.Gallagher@afit.edu

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