3-2020

An Analysis of Learning Curve Theory & Diminishing Rates of Learning

Dakotah W. Hogan

AN ANALYSIS OF LEARNING CURVE THEORY & DIMINISHING RATES OF LEARNING

THESIS

Dakotah W. Hogan, 1st Lt, USAF

AFIT-ENV-MS-20-M-212

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENV-MS-20-M-212

AN ANALYSIS OF LEARNING CURVE THEORY & DIMINISHING RATES OF LEARNING

THESIS

Presented to the Faculty

Department of Systems and Engineering Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Cost Analysis

Dakotah W. Hogan, BS

1st Lieutenant, USAF

March 2020

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENV-MS-20-M-212

AN ANALYSIS OF LEARNING CURVE THEORY & DIMINISHING RATES OF LEARNING

Dakotah W. Hogan, BS

1st Lieutenant, USAF

Committee Membership:

Dr. John J. Elshaw
Chair

Lt Col Clay M. Koschnick, PhD
Member

Dr. Jonathan D. Ritschel
Member

AFIT-ENV-MS-20-M-212

Abstract

Traditional learning curve theory assumes a constant learning rate regardless of the number of units produced; however, a collection of theoretical and empirical evidence indicates that learning rates decrease as more units are produced in some cases. These diminishing learning rates cause traditional learning curves to underestimate required resources, potentially resulting in cost overruns. A diminishing learning rate model, Boone's Learning Curve (2018), was recently developed to model this phenomenon. This research confirmed that Boone's Learning Curve is more accurate in modeling observed learning curves using production data of 169 Department of Defense end-items. However, further empirical analysis revealed deficiencies in the theoretical justifications of why and under what conditions Boone's Learning Curve more accurately models observations. This research also discovered that diminishing learning rates are present but not pervasive in the sampled observations. Additionally, this research explored the theoretical and empirical evidence that may cause learning curves to exhibit diminishing learning rates and be more accurately modeled by Boone's Learning Curve. Only a limited number of theory-based variables were useful in explaining these phenomena. This research further justifies the necessity of a diminishing learning rate model and proposes a framework to investigate learning curves that exhibit diminishing learning rates.

Acknowledgments

I would like to first thank my former supervisor, Mrs. Amy Knecht, and KC-46 Financial Management teammates for introducing me to cost analysis and teaching me how to cost estimate in the field. This basis of knowledge paved the way for my success in this program. I also express my sincere appreciation to my thesis advisor, Dr. John Elshaw, for his extensive support and guidance throughout all aspects of this effort. I thank the members of my thesis committee, Lt Col Clay Koschnick and Dr. Jonathan Ritschel, whose technical skills and background were instrumental throughout this endeavor. I also thank Mr. Shawn Valentine whose support in collecting the data and assisting in understanding the data was vital to the completion of this research. I also greatly appreciate the camaraderie and extensive insight from my AFIT GCA classmates. Most importantly, the constant love, support, and understanding of my wife was indispensable throughout my AFIT experience.

1st Lt Dakotah W. Hogan, USAF

Table of Contents

Page

List of Figures

Page

List of Tables

Page

I. Introduction

Background

The U.S. Government Accountability Office (GAO) critiqued the cost and schedule performance of the Department of Defense's $1.7 trillion portfolio of 86 major weapons systems in their 2018 "Weapons System Annual Assessment." The GAO cited realistic cost estimates as a reason for the relatively low cost growth of the portfolio in comparison to earlier portfolios (GAO, 2018, p. 1). Congress and its oversight committees maintain a watchful eye on the Department of Defense's complex and expensive weapons system portfolio. Inefficient programs are scrutinized and may be terminated if inefficiencies persist. Funding of inefficient programs will also lead to the underfunding of other programs. These terminated and underfunded programs may result in capabilities gaps that negatively impact our nation's defense.

A key to the efficient use of resources is accurately estimating the resources required to produce an end-item. A popular method of forecasting required resources is using learning curves. Learning curves predict the unit cost of an end-item using the item's sequential unit number in the production line with other fixed parameters. Learning curves are especially useful when estimating the required resources for complex products. The most popular learning curve models used in the Department of Defense are over 80 years old and may be outdated in today's technology-rich production environment. Additionally, several researchers to include founders of the theory have identified a specific shortfall of these traditional learning curve models to include Wright (1936), Asher (1956), and Badiru (2012), among others. These researchers have demonstrated both theoretically and empirically that the effects of learning do not continue in perpetuity: the effects of learning slow or cease over time.

A new model, Boone's Learning Curve constructed by Boone (2018) has been proposed to account for diminishing rates of learning as more units are produced. The purpose of this research is to

examine Boone's Learning Curve and assess if it is more accurate in explaining the required resources in comparison to the traditional learning curve theories. This research will compare each learning curve theory using a large number of programs with a wide range of attributes.

Research Objectives, Questions, & Hypotheses

The two learning curve models cited by the *GAO Cost Estimating and Assessment Guide* (2009) are Wright's Cumulative Average Learning Curve theory developed in 1936 and Crawford's Unit Learning Curve theory developed in 1947. Although both learning curve theories use the same general equation, the theories have different methodologies to estimate parameters and contrasting interpretations of outcomes. Wright's Learning Curve is shown in Equation 1.

$$\bar{Y} = Ax^b \tag{1}$$

Where:

$\bar{Y}$ = cumulative average cost of x units
A = theoretical cost to produce the first unit (T1)
x = cumulative number of units produced
b = $\frac{\ln \textit{Learning Curve Slope}}{\ln 2}$

Wright's Learning Curve yields a cumulative average cost per unit given the cumulative number of units produced, a learning curve slope, and a theoretical first unit cost. For example, with a learning curve slope of 80% and a first unit cost of 100 labor hours, the average cost of the first two units would be 80 labor hours or 60 labor hours for the second unit. Regardless of the number of units produced, there is a constant decrease in labor costs due to the rate of learning.

Boone's Learning Curve diverges from the assumption of constant learning by including a decay term in the exponent that decreases the coefficient "b" as the number of units produced increases. This transformation has the effect of attenuating cost efficiencies gained from learning as more units are produced. Boone's Learning Curve is shown in Equation 2.

$$\bar{Y} = Ax^{b/(1+\frac{x}{c})} \tag{2}$$

Where:

$\bar{Y}$ = cumulative average cost of x units
A = theoretical cost of the first unit (T1)
x = cumulative number of units produced
b = $\frac{\ln \textit{Learning Curve Slope}}{\ln 2}$
c = decay value (positive constant)

Boone’s Learning Curve is written in the original form in which it was proposed using Cumulative Average Learning Curve Theory. This research will examine whether Boone’s Learning Curve can be used with both Cumulative Average and Unit Learning Curve Theories. With a learning curve slope of 80%, first unit cost of 100 labor hours, and decay value of 100, Boone’s Learning Curve yields a cumulative average cost at the second unit of 80.35 labor hours or 60.70 labor hours for the second unit. In comparison to Wright’s Learning Curve using the same parameters, the effect of learning has decreased slightly in the production of unit two. What began as an 80% learning curve has decayed to an 80.35% learning curve for the second unit. The inclusion of the decay value increases the learning curve slope and hence decreases the rate of learning as more units are produced.

Boone (2018) demonstrated his learning curve using Unit theory[1] to be more accurate to a statistically significant degree in comparison to Crawford’s Learning Curve. Given these promising results, analysis of Boone’s Learning Curve should be expanded to provide a more robust examination. For instance, Boone (2018) analyzed cost improvement curve data for prime mission equipment (PME) elements in units of total dollars only. Boone (2018) did not investigate how his curve performed with Cumulative Average Theory. Boone (2018) did not compare the accuracy of his learning curve to the

[1] Boone (2018) cites Wright’s Cumulative Average Theory extensively and forms his model using Cumulative Average Theory. However, Boone performs analysis using Unit Theory that may not be evident without reviewing his results outside of his thesis.

traditional Unit Learning Curve theory in sections of the production cycle but only at an aggregate level. This research will investigate each of these topics to analyze Boone's Learning Curve further.

The primary goal for this research is to determine how accurate Boone's Learning Curve is in comparison to both Wright's and Crawford's learning curve theories. In order to determine how generalizable Boone's Learning Curve is, Cumulative Average Learning Curve Theory and Unit Learning Curve Theory will be examined using cost data in units of total dollars and labor hours. Also, system, engine, sub-system, and sub-component data will be analyzed along with historical and contemporary program data with a variety of attributes. These objectives will be the goal of the first phase of research and will seek to answer the following questions:

1. Does Boone's Learning Curve improve upon the traditional learning curve theories provided a wide variety of program attributes?
2. If Boone's Learning Curve improves upon the traditional models, how many observed learning curves are more accurately modeled by Boone's Learning Curve?
3. If Boone's Learning Curve improves upon the traditional models, what is the average amount of error reduction from the use of Boone's Learning Curve in comparison to the traditional learning curve theories?

A second phase will also be implemented to explore further diminishing rates of learning and how this phenomenon interacts with Boone's Learning Curve. This second phase will first seek to determine how common instances of diminishing rates of learning are when using the traditional learning curve theories. Investigating the presence of empirical instances of diminishing rates of learning will delineate opportunities where Boone's Learning Curve can significantly improve upon the traditional learning curve theories. Additionally, specific hypothesized program learning curve attributes will be investigated to determine how they affect the degree which Boone's Learning Curve improves upon the

traditional learning curve theories. These attributes will also be investigated to understand their effects on the presence of diminishing rates of learning when modeled using traditional learning curve theories. These objectives will be addressed through the following research questions:

4. Do program learning curves exhibit diminishing rates of learning when modeled using the traditional learning curve theories?
5. Does Boone's Learning Curve more accurately model program learning curves with diminishing rates of learning?
6. If Boone's Learning Curve improves upon the traditional models, in what segments of the learning curve does Boone's Learning Curve more accurately model observed learning curves?
7. If Boone's Learning Curve improves upon the traditional models, what program attributes affect the degree to which Boone's Learning Curve will more accurately model observed learning curves?
8. If observed learning curves exhibit diminishing rates of learning, what program attributes affect the degree to which diminishing rates of learning will occur as modeled by the traditional learning curves?

Methodology

In order to test the generalizability of Boone's Learning Curve in the first phase, a diverse set of DoD program data will be gathered from the DoD's Cost Assessment Data Enterprise (CADE) using Contractor Cost Data Reports 1921-1 and 1921-2. First, Cumulative Average Learning Curve Theory and Unit Learning Curve Theory will be used to calculate predicted learning curves. Following, Boone's Learning Curve will be used to calculate predicted learning curves using both the Cumulative Average Learning Curve and Unit Learning Curve methodologies. Non-linear solver optimization will be used to estimate learning curves parameters to form these predicted learning curves. The model error will be calculated using the predicted and observed learning curves in the form of root mean squared error

(RMSE) percentage difference and mean absolute percent error (MAPE) percentage difference for Boone's Learning Curve and the traditional learning curves. Statistical difference tests will then be used to compare Boone's Learning Curve and the traditional learning curves using these measures of error. When Boone's Learning Curve utilizes Cumulative Average Theory, Boone's Learning Curve will be compared to Wright's Cumulative Average Theory Learning Curve. When Boone's Learning Curve utilizes Unit Theory, Boone's Learning Curve will be compared to Crawford's Unit Theory Learning Curve.

In order to explore the presence of diminishing rates of learning and its interaction with how Boone's Learning Curve more accurately models observed learning curves, the amount of error in different segments of the learning curve will be investigated. The amount of error between Boone's Learning Curve and Crawford's Learning Curve as well as between Crawford's Learning Curve and observed learning curves will provide the data for these statistical tests. The learning curve segment on which these tests will focus is the fourth quarter because instances of diminishing rates of learning are common in this segment. Statistical tests will be utilized to compare error in the fourth quarter to determine if observed learning curves exhibit diminishing rates of learning. Additionally, these tests will provide evidence as to where in the learning curve Boone's Learning Curve more accurately models observed learning curves. Following, a confusion matrix will be utilized to determine if Boone's Learning Curve more accurately models programs with diminishing rates of learning. Lastly, regression analysis will be used to determine the effects, if any, hypothesized variables have on the degree to which Boone's Learning Curve will more accurately explain the observed learning curve. Regression analysis will be used once more to determine the effects, if any, hypothesized variables have on the degree to which diminishing rates of learning will occur in observed learning curves.

Implications

This research will examine whether Boone's Learning Curve is the solution to model learning curves with diminishing rates of learning. This research will also study if diminishing rates of learning are

systematically present in Department of Defense programs as well as what program attributes give rise to this phenomenon. Additionally, this research will investigate how Boone's Learning Curve more accurately models observed learning curves while exploring attributes that substantiate its use. In the end, Boone's Learning Curve may improve cost estimation accuracy in comparison to the traditional theories while promoting new research into learning decay.

II. Literature Review

Learning Curve Phenomenon

The concepts of learning and learning curves are intuitive: as a worker repetitively performs tasks to assemble a product, the worker will gain efficiencies. These efficiencies should decrease the time the worker spends on each unit as more units are produced. These efficiencies translate to a continuous reduction in labor hours and cost savings over time. The *GAO Cost Estimating and Assessment Guide* lists four reasons for these gains in efficiency that broaden the scope of learning. First, as more units are produced, workers tend to become "more physically and mentally adept" at performing tasks, and supervisors become more efficient at utilizing workers (GAO, 2009, p. 119). This aspect of learning is considered to be learning-by-doing at the laborer level and is termed autonomous learning (Levy, 1965). Second, the work environment is improved to include the "climate, lighting, and general working conditions" (GAO, 2009, p. 119). Third, the production process is changed to "optimize the placement of tools and material and simplify tasks" (GAO, 2009, p. 119). Lastly, market forces in the competitive business environment will require suppliers to improve efficiency to survive (GAO, 2009, p. 119). These last three aspects of learning are considered to be organizational learning by continuous improvement efforts termed induced learning (Levy, 1965; Dutton and Thomas, 1984).

Several terms are used to describe learning curves that include cost improvement curve, cost/quantity relationship, manufacturing process function, experience curve, and product improvement function (ICEAA Module 7, 2014, p. 7). The original and most generalized term "learning curve" will be used in this research. Although learning curves are used most popularly in aircraft manufacturing, the concept can be applied to "relatively large and complex products that require various types of fabrication and assembly skills" (Asher, 1956, p. 5). The *Air Force Cost Analysis Handbook* includes several specific situations in which learning curves apply that include "the manufacture of a complex end-item, limited changes to product characteristics or technology,

continuous manufacturing process, constant management pressure to improve, and consistent production rates" (Department of the Air Force, 2007, pp. 8-1 – 8-2). The Handbook also includes other criteria to include "a high proportion of manual labor, labor efficiency/job familiarization, standardization, specialization, and methods improvements, improved materiel flow and reduced scrap, improved production procedures, tools, and equipment, improved workflow and engineering support, and product redesign improvements" (Department of the Air Force, 2007, pp. 8-1 – 8-2). These situations describe aspects of the manufacturing process that enable organizational learning and allow for labor efficiencies, although learning can occur without all criteria being present.

Cumulative Average Learning Curve Theory

The concept of a learning curve was first formally recorded by Theodore Paul Wright in 1936 in his work "Factors Affecting the Cost of Airplanes." Wright identified the learning curve concept in a pre-World War II production environment of a small two-seater aircraft (ICEAA Module 7, 2014, p. 16). He observed that as a worker repeatedly performs the same task, the time required to complete that task will decrease at a constant rate. More specifically, Wright formulized that as the number of aircraft produced doubles, the cumulative average labor cost would decrease at a constant rate (Wright, 1936). This relationship is described in Equation 1 and is the Cumulative Average Theory widely in use. When learning curves utilize this Cumulative Average Theory, they are frequently called "Wright Curves" (ICEAA Module 7, 2014, p. 16).

The learning curve slope in the learning curve exponent "b" of Equation 1 defines how each doubling of produced units reduces cumulative average costs. For example, Wright used his empirical data to calculate a learning curve slope of 80%. Therefore, as the cumulative number of units doubles, the cumulative labor cost of the doubled units would be 80% of the original undoubled amount resulting in a 20% cumulative average reduction in labor cost (Wright, 1936). This 20% cumulative average reduction can also be called the rate of learning. Higher rates of learning lead to greater reductions in

labor costs. Although Wright’s model cited 80% as a universal learning curve slope, learning curve theory evolved to realize that other slopes are possible based on an end item's unique manufacturing characteristics (Jaber, 2006).

Wright’s Cumulative Average Learning Curve is cumbersome to use because cumulative average costs are calculated in place of the unit cost. Figure 1 depicts Wright’s Learning Curve with an 80% learning curve slope based on a first unit cost of 100. Figure 1 shows that as the cumulative number of units produced doubles, the cumulative average cost decreases by 20%. The corresponding unit cost is also displayed to highlight how the cumulative average cost differs from the unit cost. The unit cost is calculated by summing the cumulative costs up to but not including the unit number and subtracting it from the total cost. In Figure 1, the first unit cost for both cumulative average cost and unit cost is 100. For unit two, the cumulative average cost of the first two units is 80; this is a 20% reduction due to an 80% learning curve slope. The total cost of the first two units is 160. Because the first unit cost is known to be 100, the second unit cost can be calculated from the difference to be 60. These same calculations can be used to obtain the unit costs for the remaining units.

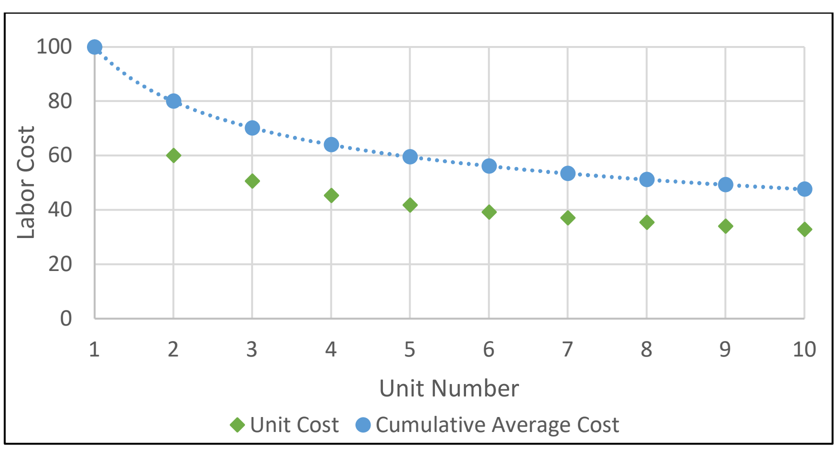


Figure 1: Wright’s Cumulative Average Theory at an 80% learning curve slope

Wright illustrated Equation 1 using a graph with vertical and horizontal axes displayed in logarithmic rather than linear scale. Wright illustrated his equation on this logarithmic graph in order to highlight the straight line representing a constant rate of learning (Wright, 1936). The same function can be graphed in linear scale by transforming Equation 1 into a log-linear form by taking the natural logarithm of both sides. This log-linear transformed equation is shown in Equation 3. The parameter definitions for Equations 1 and 3 are the same.

$$\ln \bar{Y} = \ln A + b \ln x \tag{3}$$

Equation 3 also allows analysts to apply linear regression analysis in order to estimate the parameters A and b from a set of cumulative average cost data (Mislick & Nussbaum, 2015, p. 185). Figure 2 displays Wright's Cumulative Average Theory at an 80% learning curve slope transformed into a log-linear form. The parameters are identical to the parameters of Figure 1 with a first unit cost of 100. The constant learning curve slope indicated by the linear slope in Figure 2 is a crucial concept of this traditional learning curve theory.

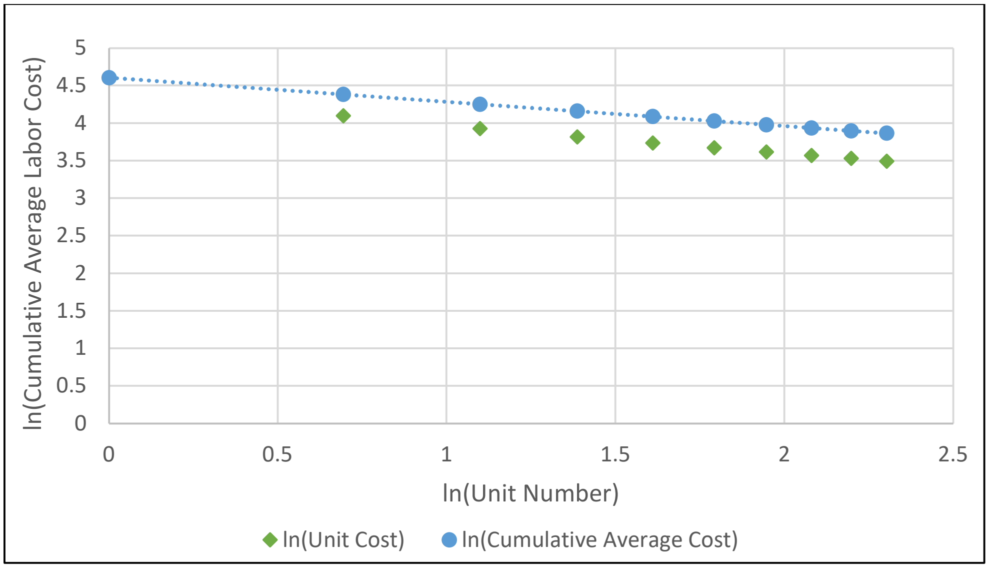


Figure 2: Wright's Cumulative Average Theory at an 80% learning curve slope in log-linear form

Unit Learning Curve Theory

Several years following Wright's Cumulative Average learning curve theory, J.R. Crawford formulated the Unit Learning Curve Theory, formally written in 1944. Together, these theories form the basis of the traditional learning curve theory. Crawford proposed his Unit Theory first in an undated manual prepared for Lockheed Aircraft Company personnel after realizing the difficulty of calculating unit costs from Cumulative Average Learning Curve Theory equations (Asher, 1956, pp. 21-22). As shown in Equation 4, Crawford's Learning Curve yields an estimated unit cost given the unit's sequential unit number within the production line, a learning curve slope, and a theoretical first unit cost. Crawford's Unit Theory is the same as Wright's aside from these differences in variable interpretation. Learning curves are often called "Crawford Curves" when they utilize Crawford's Unit Theory (ICEAA Module 7, 2014, p. 31).

$$Y = Ax^b \tag{4}$$

Where:

Y = cost of the xth unit
A = theoretical cost of the first unit (T1)
x = sequential unit number of the unit being calculated
b = $\frac{\ln Learning\ Curve\ Slope}{\ln 2}$

Using Crawford's Unit Theory with a learning curve slope of 80% and a first unit cost of 100 labor hours, the cost of the second unit is 80 labor hours. This 20% reduction in labor hours or rate of learning is due to the 80% learning curve slope. Figure 3 illustrates Crawford's Unit Theory using the same parameters from Wright's Cumulative Average Theory shown in Figures 1 and 2. The cumulative average costs are not shown in Figure 3 because these ate not germane to Unit Theory.

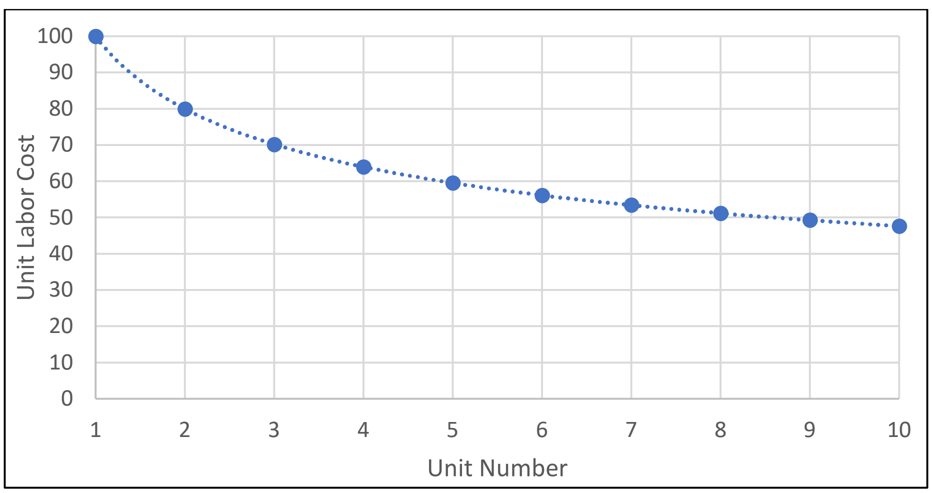


Figure 3: Crawford's Unit Theory at an 80% learning curve slope

Crawford's Unit Theory Learning Curve is transformed into a log-linear form using the same methodology used to derive Equation 3. Crawford's Learning Curve in log-linear form is shown in Equation 5 using the same parameter definitions from Equation 4.

$$\ln Y = \ln A + b \ln x \tag{5}$$

Crawford's Unit Theory Learning Curve is shown in logarithmic scale in Figure 4. Similar to Cumulative Average Theory, the constant learning indicated by the linear slope in Figure 4 is a vital concept of this theory: the rate of learning remains constant as the units double regardless of the number of units produced.

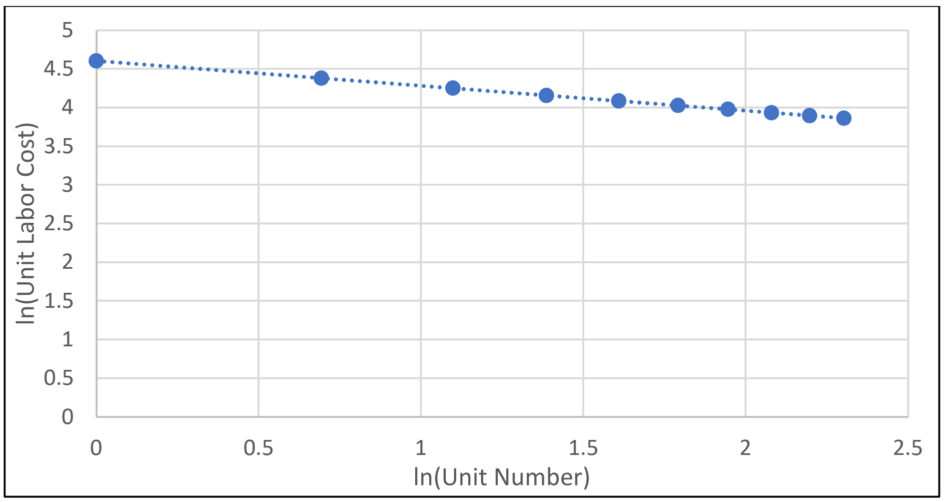


Figure 4: Crawford's Unit Theory at an 80% learning curve slope in log-linear form

Crawford's Unit Theory parameters are straightforward to estimate when data are available by each unit also call unitary data; however, manufacturers generally report cost data in the form of production lots that include the total lot cost and the number of units in that lot (Mislick & Nussbaum, 2015, p. 191). Unlike Wright's Cumulative Average Theory, Crawford's Unit Theory must utilize lot midpoints to estimate parameters when unitary cost data is unavailable. The algebraic lot midpoint is defined as "the theoretical unit whose cost is equal to the average unit cost for that lot on the learning curve" (Mislick & Nussbaum, 2015, p. 192). In other words, the lot midpoint is the unit that will divide the area under the learning curve evenly within the lot (Mislick & Nussbaum, 2015, p. 192). This lot

midpoint is used in the Unit Theory learning curve formula as the sequential unit number or independent variable “x” in Equations 4 and 5. The lot midpoint supplants using sequential unit numbers because sequential unit numbers are unavailable when using lot cost data. When the lot midpoint is the independent variable in Equation 4, the dependent variable will yield the average lot cost. The average lot cost results because this x-coordinate is the most representative point for the lot (ICEAA Module 7, 2014, p. 40).

Lot midpoints are calculated in a two-step approach due to the lack of a closed-form solution. A closed-form solution does not exist because the lot cost is a function of the learning curve exponent “b” from Equations 4 and 5 used to estimate the lot midpoint. However, the lot midpoint is also used to estimate the learning curve exponent “b.” The first step in calculating a lot midpoint utilizes a parameter-free approximation formula to estimate the lot midpoint. These lot midpoint estimates are then used to estimate the learning curve exponent “b.” The second step is to use a lot midpoint formula that includes an estimate of the learning curve exponent “b” and iterate until successive values of the estimated lot midpoints and “b” are sufficiently small (Mislick & Nussbaum, 2015, pp. 200-201). The parameter-free lot midpoint approximation is shown in Equations 6 (Mislick & Nussbaum, 2015, p. 193).

$$\text{Lot Midpoint} = \frac{F+L+2\sqrt{FL}}{4} \tag{6}$$

Where:

F = the first unit number in a lot
L = the last unit number in a lot

Several parameter lot midpoint approximations exist, but a simple and popular lot midpoint approximation is Asher’s Approximation shown in Equation 7. The same parameter definitions presented in Equation 6 also apply to Equation 7, and the learning curve exponent “b” is the same as shown previously in Equations 1-5 (Mislick & Nussbaum, 2015, p. 201)

$$\text{Lot Midpoint} \approx \left[\frac{(L+\frac{1}{2})^{b+1}-(F-\frac{1}{2})^{b+1}}{(L-F+1)(b+1)}\right]^{(\frac{1}{b})} \tag{7}$$

Comparison of Cumulative Average and Unit Theory

Cumulative Average Theory and Unit Theory will produce different predicted costs provided the same set of data despite all predicted costs being normalized to unit costs. Figure 5 demonstrates this point where Unit Theory was used to generate data using a first unit cost of 100 and a learning curve slope of 90%. The original Unit Theory data was converted to cumulative averages in order to estimate Cumulative Average Theory Learning Curve parameters. Cumulative Average Theory estimated a learning curve slope of 93% and a first unit cost of 101.24. These Cumulative Average Theory parameters were then used to predict cumulative average costs. These predicted costs were then converted to unit costs. This conversion allows for the Cumulative Average predictions to be directly compared to the original Unit Theory generated data.

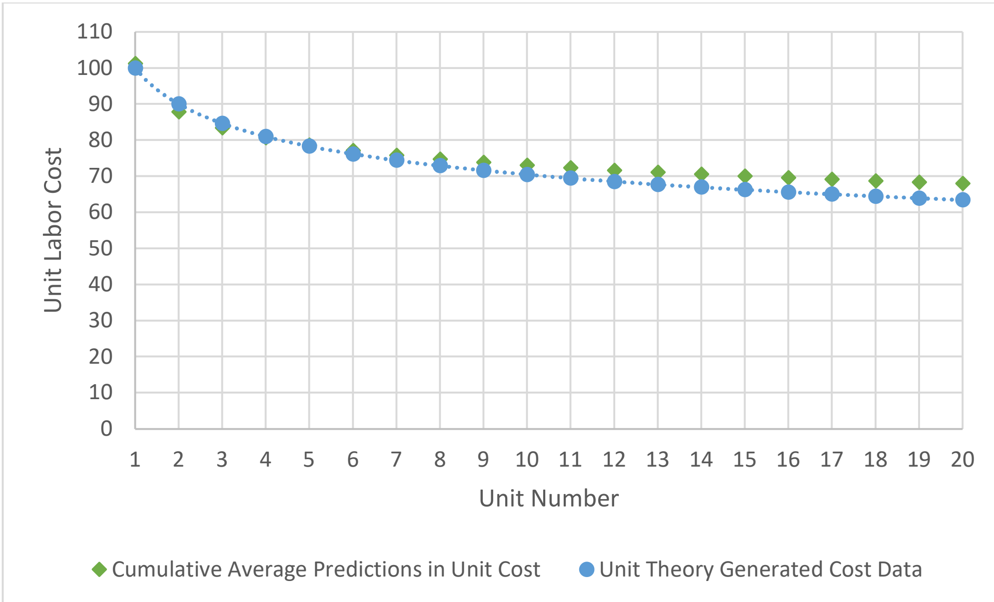


Figure 5: Cumulative Average & Unit Learning Curve Theory Comparison

As shown in Figure 5, the Cumulative Average Learning Curve predictions first overestimate, then underestimate, and ultimately overestimate the generated Unit Theory data for all remaining units. A similar case would occur if Cumulative Average Theory were used to generate data and Unit Theory learning curve parameters were estimated from this data. Figure 5 highlights that these two theories are inherently different due to differences that occur when estimating learning curves parameters using unit costs or cumulative average costs. Several factors can assist an analyst in deciding which theory to apply; however, solely relying on goodness-of-fit statistics will likely bias the decision toward Cumulative Average Theory (Mislick & Nussbaum, 2015, p. 215; Cullis, Coleman, & Braxton, 2008).

Frequently goodness-of-fit statistics to include the coefficient of determination (R^2) and standard error are used to determine which model best explains variation in a dataset. The coefficient of determination is the total variation of the dependent variable explained by the independent variable (Hilmer, 2014, p. 90). The standard error is the measure of how far on average the data tend to fall from the predicted learning curve (Hilmer, 2014, p. 91). These goodness-of-fit statistics can be used when

comparing models of the same units, which is not the case when comparing Cumulative Average and Unit Theory learning curves.

Researchers investigated this Cumulative Average Theory goodness-of-fit statistic bias and presented at a Society of Cost Estimating and Analysis (SCEA) conference (2008). The researchers used a methodology similar to that used to produce the example illustrated in Figure 5. Cumulative Average data were first generated, and a Unit Theory learning curve was fit to this data. Unit Theory data were then generated, and a Cumulative Average learning curve was fit to this data. The goodness-of-fit statistics reliably indicated the correct learning curve theory to model these perfect data. Next, artificial variation was injecting into the generated data, and the researchers repeated the process. When the researchers injected variation or error in the data, the goodness-of-fit statistics overwhelmingly favored selecting Cumulative Average Theory over Unit Theory even with small amounts of variation in the data (Cullis et al., 2008). In other words, Cumulative Average Theory Learning Curves will tend to have a higher coefficient of determination and lower standard error when compared to Unit Theory learning curves. This bias in the goodness-of-fit statistics is because Cumulative Average Theory is a cumulative running average, so the curves are generally smoother and closer to the data points (Mislick & Nussbaum, 2015, p. 215). Therefore, bias exists in favor of Cumulative Average Theory, so more subjective judgments are warranted to determine which learning curve theory to utilize.

The *GAO Cost Estimating and Assessment Guide* (2009) provides factors to consider when choosing which learning curve theory to model data. Analysts should review which theories were applied to analogous systems that are similar in form, fit, or function to the current system being considered (GAO, 2009, p. 369). Next, some industries have standards that prefer one theory over the other (GAO, 2009, p. 369). Experience should also be considered by reviewing which theory has been applied to the contractor in the past (GAO, 2009, p. 369). Lastly, some aspects of the production environment can indicate which theory is best to apply (GAO, 2009, p. 369). For example, Cumulative

Average theory is best when “the contractor is starting production with prototype tooling, has an inadequate supplier base, expects early design changes, is subject to short lead times,” or where there is a risk of concurrency between development and production phases (GAO, 2009, p. 369). In contrast, Unit Theory is more suited for contractors that are well-prepared to begin production (GAO, 2009, p. 369).

Other factors must be considered when deciding which learning curve theory to use. Cumulative Average learning curve theory will provide more conservative estimates and is less responsive to trends than the Unit learning curve theory (Mislick & Nussbaum, 2015, p. 215). For these reasons, Unit Learning Curve theory is frequently favored by government negotiators when negotiating contracts (Mislick & Nussbaum, 2015, p. 215). Cumulative Average Learning Curve theory also relies on continuous data and is unable to be calculated with missing prior data using traditional estimation techniques.

Lastly, when ordinary least squares (OLS) regression is used to estimate Cumulative Average learning curve parameters, the cumulative averaging technique violates the OLS regression assumption of independence. For OLS regression to provide an unbiased estimator, the data must be obtained through independent random sampling (Hilmer, 2014, pp. 111-112). In other words, the unit labor cost and its associated unit number cannot be statistically related to other observations of unit labor cost and their associated unit numbers. This assumption is violated due to the costs of earlier observations being a function of the costs of later observations from cumulative averaging calculations. This violation biases the learning curve parameters to produce expected values that are not equal to the population parameter being estimated (Hilmer, 2014, p. 109). Despite this violation, Cumulative Average Learning Curves estimated using OLS regression are widely used and remain a valid method for estimating learning curves.

Cost Accounting for Learning Curves

The fundamental aspects of traditional learning curve theory apply only to a subset of total program costs. Hence appropriate costs must be considered when applying the theory to yield viable parameter estimates and predictions. In a complex program, costs can be presented in units of hours or dollars and organized as recurring and non-recurring for various cost elements of the end-item or the program as a whole. For each cost element, labor costs are also categorized into further groups. The analyst must select the applicable subset of costs and consider their units when utilizing learning curve theory.

Costs are generally categorized as recurring and non-recurring costs. Non-recurring costs are one-time costs that are not directly attributable to the number of end-items being produced (Mislick & Nussbaum, 2015, p. 26). Recurring costs are costs that are incurred repeatedly for each unit produced (Mislick & Nussbaum, 2015, p. 26). At the basis of learning curve theory is the idea that costs vary as more units of an end-item are produced. Therefore, non-recurring costs are excluded from learning curve analysis due to the inability to relate these costs with the number of units produced (Mislick & Nussbaum, 2015, p. 180). Traditional research also has limited insight into how learning applies to non-recurring costs. For these reasons, learning curve analysis focuses solely on recurring costs in estimating learning curve parameters and predicting recurring costs (Mislick & Nussbaum, 2015, p. 180).

Manufacturing program costs are also organized broadly as labor, material, and overhead. T.P. Wright initially claimed that all these categories vary with the number of units produced, although he specifically focused on labor hour costs when forming his seminal theory (Wright, 1936). Due to this focus and the intuitive idea of learning at the laborer level, researchers have since focused solely on labor costs to include J.R. Crawford. Crawford exclusively studied labor learning and elaborated at length on how learning occurs from the laborer's perspective (Asher, 1956, p. 24). Both fundamental theorists also focused on the laborers who manufactured the aircraft by considering touch labor (Asher, 1956, pp.

16, 21). Additionally, cost estimating standard practice and guidance concerning learning curves also provides the basis for considering touch labor costs only (ICEAA Module 7, 2014, p. 7; Department of the Air Force, 2007, p. 8-1).

Defense contractors are often contractually required to submit costs incurred when producing large, complex end-items for the United States Government. These costs have historically been submitted on the Defense Department (DD) Form 1921 report series to include the Functional Cost-Hours Report DD Form 1921-1 and Progress Curve Report DD Form 1921-2. The Department of Defense transitioned to using the Cost and Hour Report "FlexFile" and Quantity Data reports on 15 May 2019 (Burke, 2019). However, historical program data will likely remain in the legacy 1921 series forms, and "FlexFile" and Quantity Data reports can easily be manipulated to legacy 1921 forms. The 1921-1 form is organized by work breakdown structure (WBS) elements that include various functional cost categories both in units of hours and dollars. Three broad functional cost categories: labor, material, and other costs are included in both forms of recurring and non-recurring costs. This form also has four functional labor categories that include manufacturing, tooling, engineering, and quality control labor. These four labor category costs, when summed with the material costs and other costs, comprise the total cost for each WBS element for recurring and non-recurring costs. This total cost is provided in units of dollars due to the underlying units of material and other costs. A document accompanies the 1921-1 to describe the elements of the form called a 1921-1 Data Item Description (DID). The 1921-1 DID defines these various functional cost categories to include the four labor categories whose definitions are useful in determining which categories pertain to learning curve analysis.

The definition for the manufacturing labor cost category most clearly aligns with the extant literature to be the focus as the pertinent labor cost category for learning curve research. According to the 1921-1 DID, the manufacturing labor category “includes the effort and costs expended in the fabrication, assembly, integration, and functional testing of a product or end item. It involves all the

processes necessary to convert raw materials into finished items" (1921-1 Data Item Description, p. 12). This manufacturing labor category aligns with the categories examined by Wright, which he called "assembly operations" (1936, p. 124), along with those cost categories Crawford studied, which he called "airframe-manufacturing processes" (Asher, 1956, p. 21). A RAND learning curve study also defined the direct labor used in the study as "those expended to manufacture the airframe and install the equipment required to transform the airframe into a complete, flyable airplane" (Asher, 1956, p. 49). Therefore, the manufacturing labor cost category as defined by the 1921-1 DID is associated with the types of labor costs studied by traditional learning curve theorists and succeeding research. However, data availability can prompt analysts and researchers to use total costs instead. Although these curves remain valid according to Wright (1936), they are composite learning curves with caveats to be discussed later.

The 1921-1 is organized into WBS elements defined for each program. A WBS element is a method to display, define, and organize the overall end-item into sub-products while maintaining their relationship with the end-item and other sub-products (Department of Defense, 2018, p. 4). For example, WBS elements for an aircraft program could include the airframe, wings, and engines, among other elements. These WBS elements are comprised of lower-level elements as well. For example, the airframe element may include elements such as the forward, middle, and aft airframe. WBS elements can also comprise activities instead of physical components such as testing the aircraft. Although some of these activities may experience efficiencies over time, traditional learning curve theory focuses exclusively on the production of physical components. WBS elements are frequently organized into various cost categories that can comprise elements suitable and unsuitable for learning curve analysis.

WBS elements are organized into various cost categories to include procurement costs, weapons system costs, and flyaway/rollaway/sailaway costs among others. Not all WBS elements and their respective cost categories are pertinent for learning curve analysis. The group of WBS elements in

which learning is relevant is prime mission equipment and its sub-elements. Prime mission equipment is all hardware and software WBS elements installed on the weapon system such as "propulsion equipment, electronics, armament, etc." (Flyaway Costs, n.d.). The prime mission equipment WBS aligns with those elements which experience learning according to the traditional learning curve theorists. The prime mission equipment WBS group excludes elements such as systems engineering and program management (SE/PM) and system test and evaluation (STE); these costs are instead included in flyaway costs (Flyaway Costs, n.d.). These latter elements are activities tangentially related costs that may not experience learning as theorized by the traditional learning curve literature. Recent learning curves research has considered equivalent WBS elements. Moore, Elshaw, Badiru, & Ritschel (2015) scoped their research to consider airframe costs, which is a sub-element of prime mission equipment due to the homogeneity of the programs they analyzed. Honious, Johnson, Elshaw, & Badiru (2016) also used airframe costs. Boone used prime mission equipment WBS elements to perform analysis due to the wide variety of programs researched (Boone, 2018, pp. 22-23).

Another cost accounting item to consider is whether to use hours or dollars as the units for labor cost. The total cost for each WBS element is provided in dollars due to material and other costs not having associated labor hours. Therefore, if the total WBS cost is used for analysis, units of dollars will be used. In contrast, the four labor categories to include manufacturing labor has both dollars and hours associated with each. Ideally, labor hours would be analyzed for a variety of reasons as discussed by the *Air Force Cost Analysis Handbook* (2007). First, labor dollars must be normalized to remove the effects of escalation (Department of the Air Force, 2007, p. 8-65). Escalation effects comprise economy-wide price changes as well as industry-specific price changes. Normalization removes these price variations that allow for labor costs to be compared costs across different fiscal years. The escalation indices used to normalize data are estimates of the escalation that the industry experienced that year.

Therefore, the use of escalation indices can inject error into the data. In contrast, if labor hour data were used, labor costs between years could easily be compared without normalization.

Furthermore, changes in labor rates can also bias the labor cost learning curve. Senior personnel are brought on to a program initially due to the initial complexity (Department of the Air Force, 2007, p. 8-65). Once the program stabilizes and production increases, the program usually transitions to more junior labor (Department of the Air Force, 2007, p. 8-65). This labor rate effect, when combined with the effect of normal learning, artificially steepens the learning curve (Department of the Air Force, 2007, p. 8-65). Therefore, the slopes of learning curves utilizing labor dollars will likely be steeper due to the influence of declining average labor rates as the workforce builds towards full-rate production (Department of the Air Force, 2007, p. 8-65). Although using labor cost data in dollars does not invalidate analysis and is frequently utilized due to data availability, labor hour data would ideally be used for these reasons.

In summary, the literature indicates using direct, recurring, manufacturing labor costs in the form of hours. These costs should be considered only for the WBS elements that include prime mission equipment and its lower-level elements. Using these specific WBS costs in the form of hours ensures alignment with the original costs and elements considered to be affected by learning in the traditional models. Although this review exclusively examined the DD 1921-1 form, the 1921-2 form reports the same cost data albeit in a different format for specific use in learning curve analysis. The methodology on which WBS elements and costs to consider is translatable between the two legacy forms along with the current Cost and Hour Report "FlexFile" and Quantity Data reports.

Variations to Traditional Learning Curve Theory

The traditional learning curve models assume a constant learning environment comprised of a stable production line and invariable end-item design. Due to the realities of changing production environments and modifications to the end-item configuration, many researchers have investigated

non-stable learning environments. These areas include production rate changes, changes to the end-item design during production, and breaks in production among other topics.

Production rates of end-items can vary as the program proceeds through the production lifecycle. Researchers in a 1974 RAND report first formally proposed that production rate effects can alter unit costs (Large, Hoffmayer, & Kontrovich). The researchers hypothesized that as more units are produced, fewer costs would be allocated to each unit due to fixed costs within the manufacturing process (Large et al., 1974). When fixed costs are allocated over more units, each unit will be less expensive. The researchers also hypothesized that as more units are produced, the contractor may be able to take advantage of volume discounts resulting in lower material costs per unit (Large et al., 1974). This modified learning curve equation, termed Unit Theory with Rate Adjustment, is shown in Equation 8 (Large et al., 1974).

$$Y = Ax^b R^c \tag{8}$$

Where:

Y = cost of the xth unit
A = theoretical cost to produce the first unit (T1)
x = sequential unit number of the unit being calculated
b = $\frac{\ln \mathit{Learning\ Curve\ Slope}}{\ln 2}$
R = Annual production rate
c = $\frac{\ln \mathit{Rate\ Slope}}{\ln 2}$

Despite these logical hypotheses, the researchers rarely found the rate term to be statistically significant. Several factors can cause the rate term to be not statistically significant such as how the contractor responds to the production rate changes. There are also statistical challenges when investigating rate effects due to the likely presence of multicollinearity. The independent variables “x” and “R” in Equation 8 do not make independent contributions to describe the dependent variable because there is no means to hold “x” constant while changing “R.” For these reasons, statistical analysis

is unable to discern the effects that either variable has on the dependent variable. The presence of multicollinearity tends to cause one or both independent variables to be not statistically significant when using linear regression analysis (Department of the Air Force, 2007, pp. 8-31 – 8-32).

Also, researchers have studied how production breaks alter the learning curve. Production breaks occur when the manufacturer of the end-item stops production for a period. During the time-lapse between the completion of a unit and the start of another unit, a loss of learning can occur. This learning loss results in an increased cost for the first unit following a production break and all subsequent units (Honious et al., 2016). One popular method to assess the loss of learning and subsequent unit costs is the Anderlohr Method (1969). This method identifies five factors that when weighted appropriately account for the amount of learning lost during the production break. The amount of lost learning is used to regress up the original learning curve before production resumes. Once production resumes, the manufacturing process resets to the cost of a previously produced unit and progresses back down the revised learning curve at the original learning rate for all units produced after the production break (Anderlohr, 1969).

The final area in which research has focused is changes to the end-item during production. These changes include additions, deletions, and substitutions of components of the end-item. These modifications can occur due to conventional engineering change orders to the configuration of the end-item or due to concurrency between development testing and production that reveals necessary design changes. The addition, deletion, and substitutions of components of an end-item during production can cause the following units' costs to differ significantly than what is predicted using traditional learning curve theory (Department of the Air Force, 2007, pp. 8-50 – 8-56). Additionally, configuration changes can also affect the rate at which the manufacturing process learns, which alters and often steepens the original learning curve slope (Honious et al., 2016). The learning curve slope is especially affected during

concurrency between development testing and production due to the continual flow of design changes, which tend to flatten the learning curve slope. (Department of the Air Force, 2007, p. 8-50 – 8-56).

This research into production rates, production breaks, and changes to end-items demonstrates the importance of a stable production environment with a constant end-item configuration. Without these tenets, traditional learning curve analysis becomes challenging due to confounding variables. The influence of confounding variables obscures how unit costs are related to the number of units produced.

Forgetting & Plateauing Phenomena

An implicit assumption in the traditional learning curve theories is that knowledge obtained through learning does not depreciate (Epple, Argote, & Devadas, 1991). However, empirical evidence demonstrates that knowledge depreciates in organizations (Argote, Beckman, & Epple, 1990; Argote, 1993). Argote et al. (1990) have shown that knowledge depreciation occurs at both the laborer level and the organizational level. Many variations of the traditional models make use of a concept of performance decay and forgetting to model non-constant rates of learning. Forgetting and its effects on lost learning can take many forms and is essential to consider in contemporary learning curve analysis.

Forgetting is the concept that laborers and the organization as a whole will experience a decline in performance over time resulting in non-constant rates of learning. Badiru (2012) theorizes that forgetting and resulting performance decay is a result of factors "including lack of training, reduced retention of skills, lapse in performance, extended breaks in practice, and natural forgetting" (p. 287). According to Badiru (2012), these factors may be caused by internal processes to include training policy or external factors to include breaks in production. Badiru (2012) lists three cases in which forgetting arises. First, forgetting may occur continuously as a worker or organization progresses down the learning curve due in part to natural forgetting (Badiru, 2012). In other words, the impact of forgetting may not wholly eclipse the impact of learning but will hamper the rate of learning while performance continues to increase at a slower rate. Second, forgetting may occur at distinct and bounded intervals such as

during a scheduled production break (Badiru, 2012). Third, forgetting may intermittently occur at random times and for stochastic intervals such as during times of employee turnover (Badiru, 2012). Figure 6 illustrates this third case where intermittent periods of forgetting degrade the regular learning curve path to result in a degraded performance curve. In this illustration, the learning curve is shown as an increase in performance rather than a decrease in time or cost (Badiru, 2012). Others have expanded on the causes of forgetting and have drawn similar conclusions to Badiru (2012) (Jaber, 2006; Glock, Grosse, Jaber, & Smunt, 2019; Jaber & Bonney, 1997).

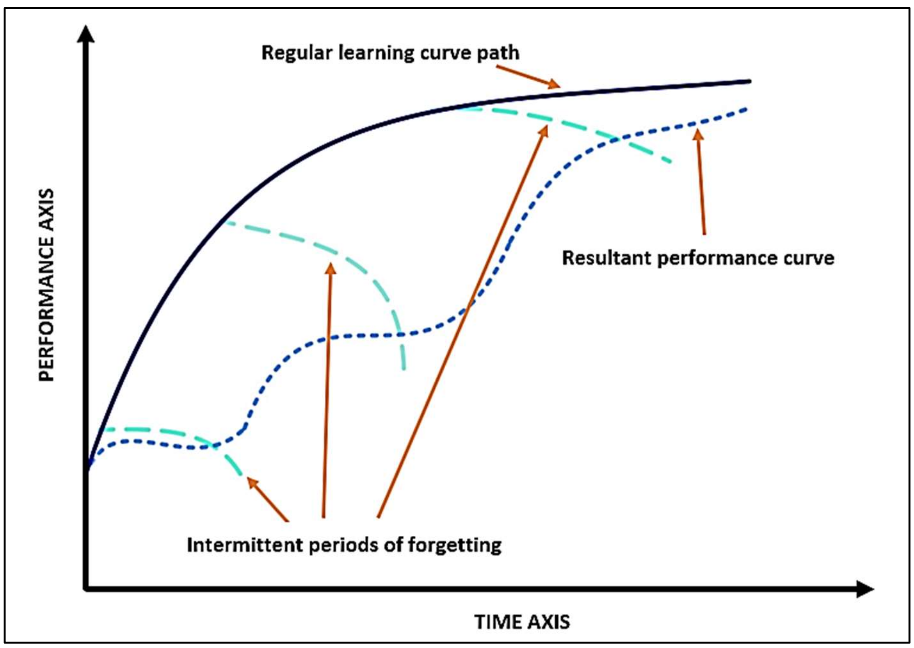


Figure 6: Effects of Forgetting on the Traditional Learning Curve (Adapted from Badiru, 2012)

This decline in performance decays the rate of learning that causes longer manufacturing times and higher costs than would be forecasted using traditional learning curve theory. Although forgetting is common when production breaks occur as previously discussed, forgetting can occur without interruptions to the production line as discussed by Argote et al. (1990) and Badiru (2012). Many contemporary learning curve models attempt to incorporate the concept of forgetting. Learning curves

that model Badiru's (2012) first case of forgetting will be discussed later in the Literature Review. Models that incorporate variations to traditional learning curves to include rates of production, breaks in production, configuration changes to end-items attempt to model Badiru's (2012) second case of forgetting. Badiru's (2012) third case is challenging to model due to the stochastic nature of when and for how long forgetting will occur.

The concept of forgetting and its impact on decaying and non-constant rates of learning has proven relevant in contemporary learning curve research. Several forgetting models have been developed to include the learn-forget curve model (LFCM) (Jaber & Bonney, 1996), the recency model (RCM) (Nembhard & Uzumeri, 2000), the power integration and diffusion (PID) model (Sikström & Jaber, 2002), and the Depletion-Power-Integration-Latency (DPIL) model (Sikström & Jaber, 2012) among others (Glock et al., 2019). However, these forgetting models focus solely on the phenomenon of forgetting due to interruptions of the production process and most directly model Badiru's (2012) second case of forgetting (Glock et al., 2019; Anzanello & Fogliatto, 2011; Jaber, 2006). Jaber (2006) states that "there has been no model developed for industrial settings that considers forgetting as a result of factors other than production breaks" (p. 30-13) and mentions this as a potential area of future research. Although forgetting models have emerged after Jaber (2006), a review of the popular forgetting models cited confirms Jaber's statement. Therefore, Badiru's (2012) first case of forgetting along the learning curve and its effect on the curve should be investigated.

A related concept to the forgetting phenomenon is the plateauing phenomenon. According to Jaber (2006), plateauing occurs when the learning process ceases. This ceasing of learning results in a flattening or partial flattening of the learning curve corresponding to rates of learning at or near zero. These near-zero rates of learning are in contrast to forgetting curves where rates of learning may become negative resulting in inverted learning curves. There remains debate as to when plateauing occurs in the production process or if learning ever ceases completely (Crossman, 1959; Asher, 1956;

Moore et al., 2015; Honious et al., 2016). Jaber (2006) provides several explanations to explain the plateauing phenomenon that include concepts related to forgetting. Baloff (1966, 1970) recognized that plateauing is more likely to occur when capital is used in the production process as opposed to labor. According to some researchers, plateauing can be explained by either having to process the efficiencies learned before making additional improvements along the learning curve or to forgetting altogether (Corlett & Morcombe, 1970). According to other researchers, plateauing can be caused by labor ceasing to learn or management's unwillingness to invest in capital to foster induced learning (Yelle, 1980). Related to this underinvestment to foster induced learning, management's doubt as to whether learning efficiencies related to learning can occur is cited as another hindrance to constant rates of learning (Hirschmann, 1964).

Li and Rajagopalan (1998) investigated these explanations and concluded that no empirical evidence supports or contradicts them while ascribing plateauing to depreciation in knowledge or forgetting. Jaber (2006) concludes that "there is no tangible consensus among researchers as to what causes learning curves to plateau" and alludes that this is a topic for future research (p. 30-9).

Empirical Evidence of the Forgetting & Plateauing Phenomena

Despite the controversy in the research surrounding forgetting and plateauing effects, empirical studies have shown learning curves to exhibit diminishing rates of learning. For instance, the plateauing phenomenon at the tail end of production was investigated by Harold Asher in a 1956 RAND study. The U.S. Air Force contracted RAND after the service realized traditional learning curves were underestimating labor costs at the tail end of production (Asher, 1956, p. 13). Asher intended to study if the traditional log-linear learning curves, as shown in Equations 3 and 5 were indeed a straight line as shown in Figure 2 and 4. The alternative hypothesis for these log-log transformed curves was a convex curve indicating diminishing rates of learning as the number of units doubled (Asher, 1956, p. 13). An example of a learning curve with a diminishing rate of learning is shown in Figure 7 in linear scale and

Figure 8 in logarithmic scale. The data for both curves were generated using Unit Theory. The first unit cost is 100 with an initial learning curve slope of 80% decaying at a rate of 0.25% with each additional unit; for example, the second unit's learning curve slope is 80.25%. As shown, this flattening effect is difficult to detect in linear scale but evident in logarithmic scale.

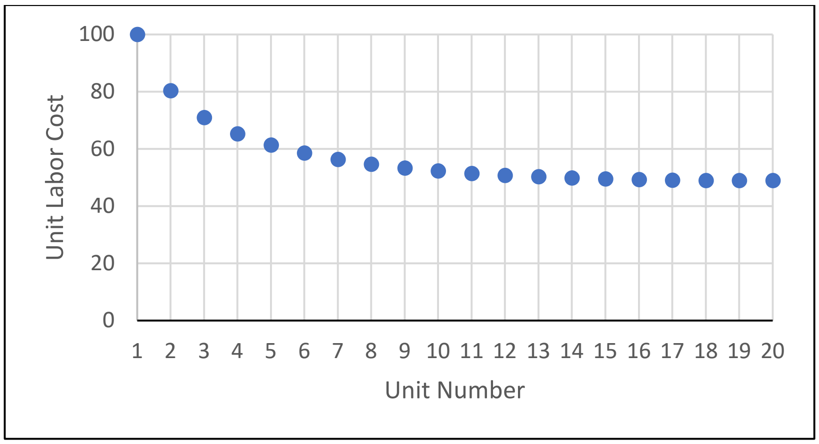


Figure 7: Unit Theory Learning Curve with a Decaying Learning Curve Slope in Unit Scale

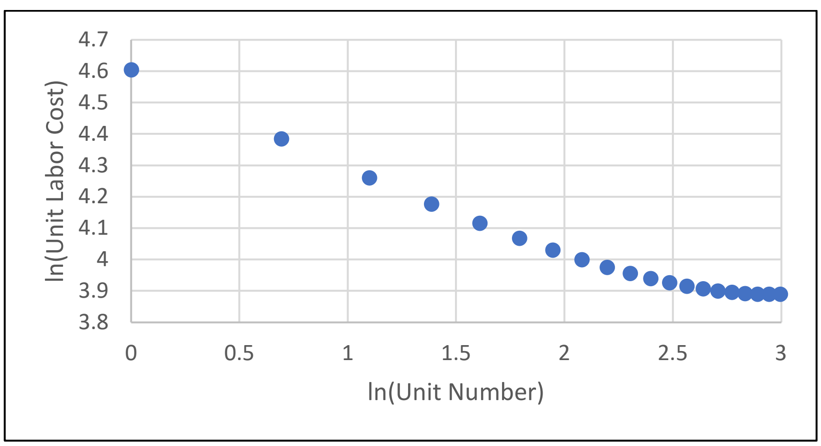


Figure 8: Unit Theory learning curve with a Decaying Learning Curve Slope in Logarithmic Scale

Asher established the basis for his hypothesis of convex log-linear transformed learning curves by first studying Wright's (1936) seminal work. Asher discovered that Wright realized his traditional curve will have various and increasing learning curve slopes when the total airplane cost is modeled by

summing the "labor, raw material, purchased material, and overhead" component curves (Wright, 1936, pp. 125-126). In other words, when the component cost curves for each cost category are aggregated and transformed to logarithmic scale, the composite curve will be convex with increasing learning curve slopes. Asher concludes the discussion of aircraft component cost curves stating that "if the component cost curves are different in slope, then the curve that represents the summation is convex" (Asher, 1956, p. 74). This explanation of plateauing due to composite curves is different from those offered by contemporary researchers.

Many researchers have neglected to research composite learning curves, possibly due to the traditional learning curve theories being unable to model the data using a constant learning curve slope. The *Air Force Cost Analysis Handbook* discusses performing analysis on composite curves stating that they are usually utilized "for high-level estimates and what-ifs, not for detailed analysis" (Department of the Air Force, 2007, p. 8-63). However, due to data availability, total WBS element cost data is often utilized for learning curve analysis. This total WBS element cost data is a composite curve due to the curve embodying all labor cost categories along with material and other costs. Asher advances this concept of overall composite curves to explain the plateauing effect further.

Asher investigated the alternate hypothesis of convex log-linear transformed learning curves by analyzing the learning curves of the various shops within a manufacturing department producing an aircraft. Asher utilized airframe cost data with the appropriate amount of detail to perform a learning curve analysis on the lower level job shops within the manufacturing department. He focused on using the direct labor cost category, which is equivalent to the manufacturing labor cost category used for modern learning curve analysis. According to Asher, the direct labor cost category includes "man-hours expended in each of the manufacturing operations, such as materials processing, sheet metal work, subassembly work, etc." (Asher, 1956, p. 90). Asher divided the eleven major kinds of manufacturing operations into four shop groups each with direct labor cost data (Asher, 1956, pp. 90-91).

Asher hypothesized that these shop group curves would differ and may themselves be convex in logarithmic scale because "airframe manufacturing consists of a large number of dissimilar labor operations, each requiring different amounts and kinds of equipment, skill, and training, and on which the workmen seem to experience different rates of learning" (Asher, 1956, p. 89). In other words, Asher investigated constituent direct labor cost curves in order to determine if they have different learning curve slopes or are themselves convex. If so, the manufacturing cost learning curve will be convex in logarithmic scale. Because this manufacturing cost learning curve is the level at which most learning curve analysis is performed due to data availability, this result would provide evidence for a convex learning curve in logarithmic scale and diminishing rates of learning.

Asher's results indicated that the manufacturing shop group learning curves had different learning curve slopes and were even convex in logarithmic scale. Asher claims the apparent convexity within the manufacturing shop group learning curves is due to the disparate operations within the job shops and stated that each had their unique learning curve (Asher, 1957, p. 94). Some of these shop group curves differed in where convexities occurred with some shop groups reaching an asymptote later in the production cycle at 1,000 units. However, other shop groups reached an asymptote relatively early in the production cycle at 100 units. Asher claims that a linear approximation is reasonable for a relatively small quantity of airframes produced; however, Asher claims that "it is unmistakable…that the shop-group curves for a particular model exhibit different slops over these linear segments." Asher concludes this part of his analysis by stating the convexities within job shop learning curves "cast doubt upon the validity of the hypothesis that the sum of the four shop group curves is linear" (Asher, 1956, pp. 97-98).

In summary, Asher empirically demonstrated convexity in the logarithmically transformed manufacturing cost curve by revealing that the constituent cost curves have different learning curve slopes and are themselves convex in logarithmic space. Because the manufacturing cost curve is usually

the lowest level of detail on which learning curve analysis is performed, the manufacturing cost curve will have diminishing rates of learning as cumulative output increases. Although the linear hypothesis may be appropriate for some amount of aircraft produced, it becomes less appropriate as more units are produced and higher-level composite curves are used for learning curve analysis. These conclusions demonstrate a demand for a learning curve formula with a non-constant, decaying learning curve slope.

In a more recent study, Bongers (2017) investigated if the Department of Defense's strategy in producing the F-35 Lightning II resulted in negative impacts on the learning curve during production due to its various models. While investigating this topic, Bongers (2017) also studied the F/A-18E/F Super Hornet and F-22A Raptor for comparison to the F-35 Lightning II. In order to investigate the learning curves of these aircraft, Bongers (2017) considers the impact of organizational forgetting and depreciation of knowledge by incorporating a coefficient for the depreciation of experience within the traditional learning curve model. This depreciation of experience formula is shown in Equation 9.

$$E_t = \lambda E_{t-1} + q_t \tag{9}$$

Where:

E_t = the cumulated experience
λ = measure of persistence in the stock of experience transmitted to the next period
q_t = experience gained between t and $t-1$, equivalent to the of units produced in that period
t = time period

Equation 9 was then nested into the traditional unit theory learning curve equation shown in Equation 5 in order to estimate parameters using OLS regression. Specifically, Equation 9 supplants the independent variable "x" in Equation 5. The final model resulted in Equation 10. Bongers (2017) original equation has been adapted to align with equation parameters stated prior. A time component was also included in a separate regression analysis by Bongers (2017).

$$\ln Y = \ln A + b \ln(\lambda E_{t-1} + q_t) \tag{10}$$

In Equation 10, the parameters "A,", "b," and "λ" are empirically estimated. If the coefficient estimate "λ" is zero, no transmission of experiences occurs between periods. If the coefficient estimate "λ" is one, all experience is transmitted between periods as theorized by the traditional learning curves. In Bongers (2017) results, the coefficient estimate "λ" was statistically significant to a level of 0.05 for only the F/A-18E/F Super Hornet. The coefficient estimate "λ" for the F/A-18E/F Super Hornet was 7.2% for the regression model, excluding the time component. This result indicates that on average merely 7.2% of experience in the form of produced units is transmitted from one period to the next. The coefficient estimate "λ" was 6.4% for the regression model including the time component. Bongers (2017) cites Darr, Argote, and Epple (1995) along with David and Brachet (2011) to demonstrate these results are consistent with previous research. These researchers showed a rapid depreciation of knowledge between production periods despite the absence of breaks in production (Darr et al., 1995; David & Brachet, 2011).

Both Asher (1956) and Bongers (2017) provide empirical evidence of forgetting and plateauing in Air Force aircraft production programs. Although contemporary researchers may provide various explanations for the causes of forgetting and plateauing, it is evident that some production programs do experience diminishing rates of learning as modeled by the traditional learning curve theories. Therefore, a learning curve model that incorporates non-constant rates of learning may be necessary to model some production program costs more accurately.

Forgetting & Plateauing Learning Curve Models

Wright's and Crawford's Learning Curve Theories provided the basis of the traditional approach that learning occurs at a constant rate as the number of units produced increases. Since this initial

discovery, several log-linear learning curve models were founded in attempts to more accurately model data from manufacturing processes. These contemporary models diverge from constant rates of learning by including adjustments in various forms. The six most popular models to include the traditional model are shown in Figure 9 in logarithmic scale. These illustrated models include the traditional log-linear model or Wright/Crawford Curves, the Plateau Model (Baloff, 1970), the Stanford-B Model (Chalmers & DeCarteret, 1949), the De Jong Model (De Jong, 1957), the S-Curve Model (Yelle, 1979), and Knecht's Upturn Model (Knecht, 1974). Each of these alternative learning curve models will be discussed briefly. Other notable univariate models such as Glover's Learning Formula, Pegel's Exponential Function, Yelle's Product Model, and the Multiplicative Power Model (Badiru, 1992) will not be discussed due to lack of relevance to the current topic.

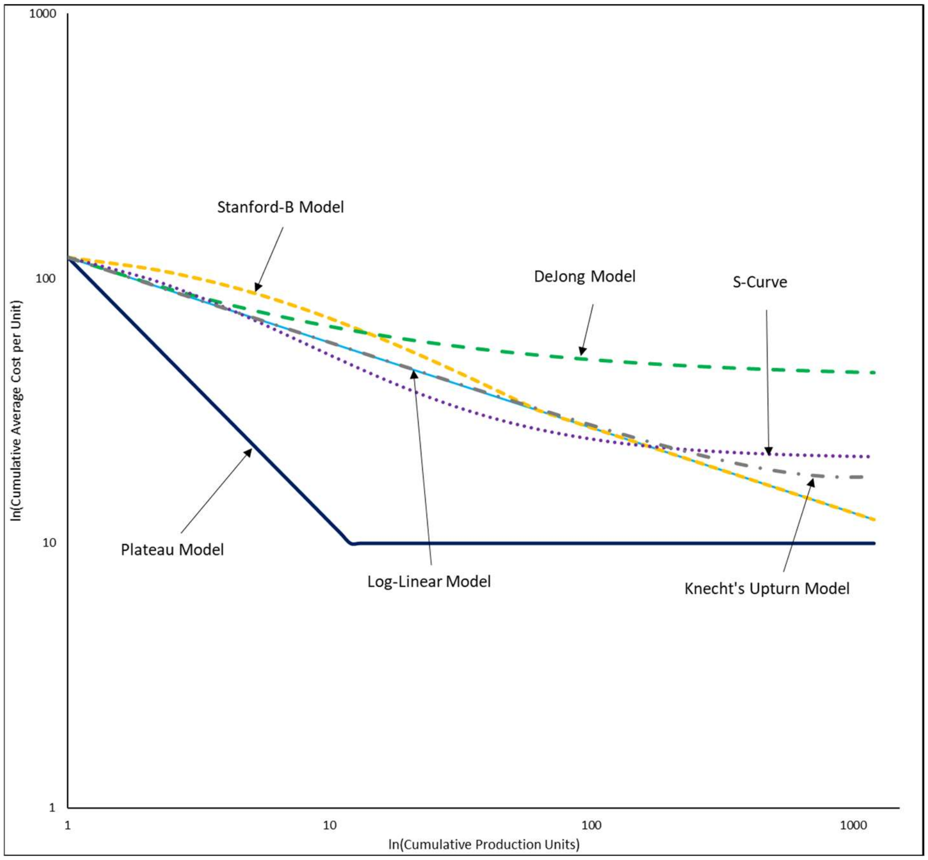


Figure 9: Comparison of Learning Curve Model on Logarithmic Scale (Adapted from Badiru, 1992)

N. Baloff proposed his Plateau Model in 1965 after recognizing that the traditional model's rate of learning remains constant over the number of units produced (Baloff, 1970). Baloff theorized this constant rate of learning is impractical due to production rates leveling off along with factors that may influence learning. To incorporate this theorized learning rate plateau, Baloff included a constant to flatten the learning curve for large numbers of produced units. After the plateau, the learning curve slope increases to 100%, where learning ceases or turns upward (Badiru, 1992).

G.R. Knecht proposed his Upturn Model in 1974 after investigating learning curves during long production runs (Knecht 1974; Anzanello and Fogliatto, 2011). Knecht's model behaves similarly to the traditional log-linear learning curves when few units are produced; however, his model decreases the learning rate as more units are produced (Glock et al., 2019). This deviation is due to the incorporation

of Euler’s number (e) raised to the product of a positive constant and the number of units produced. This transformation results in the convexity of the learning curve. Knecht’s positive constant affects the speed and degree of deviation from the traditional learning curve. The shape of Knecht’s Upturn Model visually approximates the shape of empirical data that exhibit forgetting and plateauing; however, contemporary research efforts have not investigated its utility for modeling forgetting and plateauing effects.

The Stanford-B Model was founded in 1949 after a study commissioned by the Department of Defense at the Stanford Research Institute (Chalmers & DeCarteret, 1949). The decision to pursue the study was due to visual inspection of World War II aircraft production data (Asher, 1956). The Stanford-B Model incorporates a constant parameter “B” that varies from one to ten to adjust the unit being estimated. This constant equals the equivalent units of prior experience by the manufacturer at the start of the production process. When “B” is equal to zero, the Stanford-B Model reduces to the traditional Log-Linear Model. The effects of this constant allow the Stanford-B learning curve to progress from a previous unit rather than starting the learning process at zero. This Stanford-B Learning Curve was found to more accurately model World War II aircraft production data (Badiru, 1992).

The De Jong Learning Formula was proposed in 1957 and incorporates a factor to account for the proportion of manual activity as opposed to machine activity in a manufacturing process (De Jong, 1957). This factor was termed the “incompressibility factor,” represented as “M.” The idea behind the incorporation of incompressibility is that capital equipment does not learn like human labor: the time to complete a task will remain constant or incompressible (Baloff, 1966; Baloff, 1970). When “M” is equal to zero, the De Jong Learning Formula reduces to the traditional Log-Linear Model representing a manufacturing process with exclusively human labor. When “M” is equal to one, the model reduces to the first unit cost; this simplification indicates that cost improvements do not occur because the manufacturing process comprises machine labor solely (Badiru, 1992). Therefore, the incompressibility

factor reduces the effects of learning and plateaus the learning curve. The learning curve becomes flatter as the process transitions from labor-intensive to capital-intensive. This relationship coincides with other models such as the Baloff's Plateau Model (Baloff, 1966; Baloff, 1970). Consequently, the De Jong Learning Formula is convex later in the production cycle in comparison to the constant sloped Log-Linear Model when graphed on logarithmic scales (Moore et al., 2015). This convexity in the De Jong Learning Formula indicates decreases in the rate of learning and decreases in cost efficiencies gained from learning.

In 1946, G.W. Carr proposed the S-Curve Model that combined the equivalent units factor of the Stanford-B Model and the incompressibility factor of the De Jong Learning Formula into a single model (Carr, 1946). The S-Curve models a gradual build-up at the early stages of production due to the inefficient state of production that is unfavorable for significant learning to occur. This stage is similar to the early stage of the Stanford-B model. This gradual build-up is followed by a period of peak performance where learning occurs at approximately a constant rate due to the stable, efficient manufacturing process. This stage is similar to the traditional Log-Linear Model. Following this region is a period of plateauing where learning is still occurring but at a diminishing rate. This diminishing rate is due to the effects of forgetting and inefficiencies in the manufacturing process late in production. This region is similar to the end of the De Jong Learning Formula (Moore et al., 2015).

These contemporary learning curves incorporate unique aspects of the manufacturing process to more accurately model the learning environment. Each contemporary learning curve adapts the traditional Log-Linear Model from a constant learning curve slope to learning curve slopes that can adjust to the manufacturing environment and the number of units produced. A potential drawback of these models is the requirement of additional parameters needed to adjust the learning curve slopes as more units are produced.

Research on Forgetting & Plateauing Learning Curve Models

Recent research has investigated whether the Stanford-B, De Jong, and S-Curve learning curve models more accurately predict program costs. Although these research efforts' findings were inconclusive considering their initial hypotheses, the findings provide insight into where in the production cycle and under what conditions the traditional learning curve models were less accurate than the plateauing and forgetting learning curve models.

Moore et al. (2015) used the Cumulative Average learning curve theory to compare how accurately the De Jong, Stanford-B, and S-Curve models predicted program cost data in comparison to the traditional Cumulative Average theory model. Moore et al. utilized the F-15 fighter aircraft variants A through E airframe cost data totaling 1,156 units to perform his research. Moore et al. used a Stanford-B prior experience constant of 10 when estimating the first and all other variants due to the presence of prototyping and previous variant production. Moore et al. varied the De Jong incompressibility factor from 0.00 to 0.20 in increments of 0.05 referencing sources to justify airframe manufacturing as being labor-intensive. The same Stanford-B prior experience constant and incompressibility factors were also used in the S-Curve model (Moore et al., 2015).

To obtain estimates for each contemporary model, Moore et al. (2015) began by estimating a Cumulative Average Theory learning curve for the F-15 variant A/B. He then maintained that learning curve slope, applied the respective factor(s) for each contemporary model, and estimated a new learning curve for the remaining F-15 variants C through E. Goodness-of-fit statistics were generated from each model, and the models were compared after considering statistical assumptions. This research indicated that the De Jong and S-Curve models were more accurate than Wright's Cumulative Average learning curve at an incompressibility factor of 0.05. However, as the incompressibility factor increases, the De Jong and S-Curve models were less accurate. As the incompressibility factor decreases, the contemporary models simplify to the traditional Cumulative Average learning curve model.

Therefore, the researchers found that the accuracy of the De Jong and S-Curve models in comparison to the traditional model was sensitive to the incompressibility factor; however, his research was inconclusive at determining a more accurate model for use in the Department of Defense (Moore et al., 2015).

Honious et al. (2016) also utilized the Cumulative Average learning curve theory to compare how well the De Jong and S-Curve formulas modeled program cost data in comparison to the traditional model. Honious et al. utilized an undisclosed Air Force program's airframe cost data with 95 units. Honious et al. focused on both total airframe cost and a subset of these costs: airframe integration, testing, and checkout (IA&CO). This subset of airframe costs was studied to justify low incompressibility values when using the De Jong and S-Curve models because integration, testing, and checkout are notoriously human labor-intensive processes. Honious et al. (2016) used a Stanford-B prior experience constant of 2 due to the equivalent units of experience from prototypes of the Air Force program before the first unit's production. Lastly, the researchers used an incompressibility factor from 0.05 (Honious et al., 2016).

To obtain their estimates for each contemporary model, Honious et al. (2016) first estimated a Cumulative Average Theory learning curve for the first early-fielded units. The researchers then maintained this learning curve slope, applied the respective factor(s) for each contemporary model, and estimated a new learning curve for the remaining units. Both the contemporary and traditional models were used to estimate the remaining units given the early-fielded unit's estimated parameters. Goodness-of-fit statistics were generated from each model, and the models were compared after considering statistical assumptions. Honious et al. (2016) then performed a sensitivity analysis by altering the incompressibility factor from 0.05 to 0.10. This research was performed for the total airframe cost and the airframe integration, testing, and checkout airframe cost. Honious et al. (2016) found that the De Jong and S-Curve models were more accurate than Wright's Cumulative Average

learning curve at low incompressibility factors of 0 to 0.10. However, similar to Moore et al. (2015), as the incompressibility factor increases, the De Jong and S-Curve models become less accurate. Although this research was inconclusive at determining an overall more accurate model, the research provides insights that Wright's Cumulative Average learning curve becomes less accurate at the tail-end of production when incompressibility values are low. Honious et al. (2016) explicitly reference a plateauing effect at the end of production runs. In other words, the airframe cost data these researchers analyzed tends to exhibit rates of learning less than the traditional rates of learning at the end of the production cycle. This plateauing of the learning curve slope results in a hypothesized slope below the actual learning curve data resulting in underestimated labor costs (Honious et al., 2016).

Both Moore et al. (2015) and Honious et al. (2016) illustrate that Wright's Cumulative Average Learning Curve is less accurate at low incompressibility factors where the proportion of human touch labor is high in the manufacturing process for their sets of data. Both researchers reference sources to suggest that the aircraft manufacturing industry tends to have a higher proportion of human touch labor. For example, Kronemer & Henneberger (1993) state that the aircraft manufacturing process is "fairly labor-intensive" despite the assembly of high-tech products. Kronemer & Henneberger (1993) provide three main reasons for this to include aircraft manufacturers building multiple models of the same aircraft that result in different configurations with each model requiring various tooling and assembly processes. Also, aircraft manufacturers produce relatively low quantities in comparison to other industries. These low quantities disincentivize investment in capital equipment due to the relatively high costs on a per-unit basis (Kronemer & Henneberger, 1993). Lastly, aircraft are complex and must be built to stringent standards, so manufacturers opt for skilled touch labor in place of advanced, costly machinery (Kronemer & Henneberger, 1993).

Although much has changed in the industry since 1993, the aircraft manufacturing process remains "quite a manual job, especially for low production parts and in what assembly is concerned,"

according to Vieira (2013, pp. 10, 14). Vieira cites the same reasons as Kronemer and Henneberger (1993) to explain the labor intensity of the sector despite these original observations being 20 years old. Viera (2013) and Kronemer & Henneberger (1993) provide evidence that the aircraft manufacturing industry is and will likely remain relatively labor-intensive in the near future. However, other researchers to include Sakinç (2016, p. 192) and Vértesy (2011, p. 239) provide evidence against the claim that the aircraft manufacturing industry is labor-intensive.

Another potential reason for this underinvestment in capital equipment is a lack of long-term commitment by the Department of Defense to procure end-items. Congressional fiscal policy and, in particular, the Department of Defense's Full Funding Policy requires the purchase of fully usable end-items (Department of Defense, 2017, p. 13). In other words, the Department of Defense cannot piecemeal fund an end-item in a fiscal year to procure the rest of that end-item in the following fiscal year. In order to implement this funding policy, aircraft production contracts and other complex end-item contracts are often awarded to contractors on a per-lot basis usually with several lots comprising the production phase of the program. Regulations also state that each lot of end-items must be delivered within 12 months by the contractor; therefore, the number of end-items that can be awarded is constrained (Department of Defense, 2017, p. 57). Although deviations occur to these policies due to long lead time components and multi-year procurements, these are exceptions to the policy. Due to this lot-based contracting strategy, the Department of Defense is not obligated nor able to award more than the current fiscal year's lot to the contractor to manufacture. This annual lot-based contracting structure creates a risk to the contractor of not recouping costs for specialized capital investments that require large numbers of produced units to justify. Therefore, government contractors are disincentivized from investing in specialized capital equipment due to the lack of long-term commitment by the Department of Defense as a result of funding policy (Baye, 2010, pp. 208-209). However, as the Department of Defense becomes increasingly committed to procuring a certain end-item, the contractor may begin to

invest in specialized capital equipment. This shift from labor to capital could lead to a high proportion of labor at the beginning of production and then to a high proportion of capital later. This transition from labor-intensive to capital-intensive production could lead to a steep learning curve at the beginning of production along with a flattening of the learning curve later in the production due to the inability of capital to learn and produce cost efficiencies related to learning (Baloff, 1966; Baloff, 1970).

Therefore, Moore et al. (2015) and Honious et al. (2016) conclude that there is potential for a more accurate learning curve model when the proportion of human labor is relatively high. The capital equipment dynamic in aircraft manufacturing processes will continue to evolve. However, this and other research provides evidence that the government-contracted aircraft manufacturing sector will likely lag in adopting machine labor when compared to other industries. This lag will result in a relatively high proportion of human labor. These insights provide further justification for investigating a more accurate learning curve model that diminishes its rate of learning.

Boone's Learning Curve: A Contemporary Forgetting & Plateauing Model

In an answer to these researchers findings, Boone (2018) developed a learning curve model with a rate of learning that diminishes as more units are produced. The traditional learning curve theories diminish the rate of cost reductions as more units are produced because costs will decrease at a constant rate only when the number of units produced doubles. Because the rate at which units double decreases as more units are produced, the rate of cost reductions will also decrease as more units are produced. However, the Literature Review cited various theoretical and empirical evidence indicating that the cost reductions that occur with each doubling of units may not be constant as the number of units produced increases. Therefore, Boone (2018) sought to attenuate the cost reductions that occur with each doubling of produced units by reducing the amount that each doubled unit's cost decreases as the number of units increases. This attenuation of cost reductions was accomplished by decreasing the rate of learning as the number of units increases.

Boone (2018) began by formulating a model where the exponent of the traditional learning curve equation is a function of the number of units produced. This amendment was intended to vary the learning curve exponent with the independent variable “x” in order to alter the degree of cost efficiencies experienced as the number of units produced changes. This fundamental model is presented in Equation 11.

$$\bar{Y} = Ax^{f(x)} \tag{11}$$

Where:

$\bar{Y}$= cumulative average cost of x units
A = theoretical cost to produce the first unit (T1)
x = cumulative number of units produced

Following, Boone (2018) devised a series of specific models that decreased the learning curve exponent “b” as the number of units produced “x” increased. Boone (2018) first created a model without an additional parameter, as shown in Equation 12, that aimed to reduce the learning curve exponent “b” directly by the unit number. However, Boone (2018) claimed that Equation 12 resulted in too drastic of changes to the exponent value and did not model data appropriately (Boone, 2018, p. 20).

$$\bar{Y} = Ax^{\frac{b}{x}} \tag{12}$$

Where:

$\bar{Y}$= cumulative average cost of x units
A = theoretical cost to produce the first unit (T1)
x = cumulative number of units produced
b = $\frac{\ln \textit{Learning Curve Slope}}{\ln 2}$

To temper the effect each additional unit has on the parameter “b,” a qualifier was added. This qualifier “c” was named Boone’s Decay Value with an initially studied value ranging from zero to 5,000

(Boone, 2018, p. 21). The resulting Boone's Learning Curve is shown in Equation 2. Boone found this curve was not only flatter near the end of production but was also steeper in the early stages in comparison to the traditional theory learning curve (Boone, 2018, p. 21). Holding the cumulative number of units produced "x" constant, as Boone's Decay Value approaches zero, the parameter "b" approaches zero representing a learning curve slope approaching 100%. As Boone's Decay Values approaches infinity, the parameter "b" remains unchanged, and Boone's Learning Curve simplifies to the traditional learning curve (Boone, 2018, p. 23).

Boone (2018) proceeded to test his model using Cumulative Average Theory against Wright's model. Boone (2018) utilized prime mission equipment cost data in units of total dollars for fighter, bomber, and cargo aircraft programs along with missile and munition programs (p. 21). He constrained his data to require at least five lots per program to prevent overfitting the data (Boone, 2018, p. 22). In total, 46 weapon system platforms were tested (Boone, 2018, p. 23). The OLS regression method is unable to estimate the parameters for Boone's Learning Curve because of the non-constant rate of learning; Boone's Learning Curve is convex in logarithmic scale. Instead, Boone utilized Microsoft Excel's Solver package to minimize the sum of squares errors (SSE) by iteratively adjusting the theoretical first unit cost "T1", the learning curve slope "b", and Boone's Decay Value "c" (Boone, 2018, p. 24). The conventional OLS methodology was maintained to estimate the parameters for Wright's Cumulative Average learning curve because this model remains linear in logarithmic scale (Boone, 2018, p. 24).

Microsoft Excel's Solver requires bounds for each parameter when solving for the combination of parameter values that minimize the SSE. Wright's Learning Curve parameters informed these bounds to assist in estimating Boone's Learning Curve parameters (Boone, 2018, pp. 24-25). Boone's theoretical first unit cost parameter minimum bound was equal to half of Wright's theoretical first unit cost and twice the value of Wright's first unit cost for the maximum bound (Boone, 2018, pp. 24-25). Boone's "b" parameter bounds were set between -3 and 3 times Wright's "b" specific to each estimated learning

curve (Boone, 2018, pp. 24-25). These two bounds' values varied for each learning curve estimated due to their dependence on Wright's Learning Curve parameters. Lastly, Boone's Decay Value was bound from zero to 5,000 for all estimated learning curves (Boone, 2018, pp. 24-25). The only limits that were found to be binding when solving for optimal values were the upper limit of Boone's Decay Value (Boone, 2018, p. 25).

Boone then estimated the parameters for his curve and Wright's curve for each of the 46 programs using Cumulative Average Learning Curve theory. He obtained goodness-of-fit statistics in the form of the SSE and MAPE for each estimate in order to compare the accuracy of both curves (Boone, 2018, pp. 25-26). Boone then performed a paired difference t-test for both SSE and MAPE statistics (Boone, 2018, pp. 25-26). Both the SSE and MAPE paired difference t-tests rejected the null hypothesis that the means were equal to zero (Boone, 2018, pp. 29-30). These tests indicate that Boone's Learning Curve more accurately explains the cost data in comparison to Wright's Learning Curve at a significance level (α) of 0.05. Therefore, Boone demonstrated his learning curve to be more accurate to a statistically significant degree in comparison to Wright's Learning Curve (Boone, 2018, pp. 30-31).

Boone did not assess the predictive capacity of his model in comparison to Wright's by estimating parameters for a subset of early units and then extrapolating to future lots using the same estimated parameters. This approach departs from previous research to include Moore et al. (2015) and Honious et al. (2016); however, this may be due to data availability as several lots are required for predictive analysis. Additionally, Boone did not hold constant the traditional learning curve parameters and estimate his new parameter, Boone's Decay Value, with these values fixed. Instead, Boone allowed all three of his parameters to change from Wright's estimated parameter values. Boone's methodology is also a departure from the previous methodology where Moore et al. (2015) and Honious et al. (2016) estimated the traditional learning curve, held the estimated parameters constant, and then added additional parameters. Boone's methodology may also be justified because Boone's Decay Value is a

result of estimating learning curves and requires a different learning curve slope and first unit cost be considered. This Decay Value is unlike the parameters Stanford-B parameter and incompressibility parameter “M” which are measurable values that describe the manufacturing process itself.

III. Phase 1

Methodology

Population and Sample

In order to test Boone's Learning Curve against the traditional learning curve theories, quantitative data from a diverse set of Department of Defense programs was gathered. The population studied is Department of Defense programs that have produced several complex end-items over time. These complex units can include but are not limited to aircraft, land vehicles, and missiles along with their complex sub-systems and sub-components. The data sample consists of programs with the necessary information required for learning curve analysis. This required program data included direct recurring labor costs in units of labor hours or total dollars per production lot along with the number of units per lot.

The direct recurring manufacturing labor cost category for each applicable WBS element was used to obtain labor hour data for each program. If this labor hour data were unavailable, the total recurring cost for each applicable WBS element was utilized instead. The total recurring dollar cost comprises the costs of all functional categories of labor along with materials costs and other costs for each WBS element. Unlike labor hours, costs in units of dollars must be normalized to be compared over time; therefore, all costs in units of dollars were normalized using escalation rates based on Producer Price Index (PPI) 3364 Aerospace Products and Parts. Removing the effects of escalation using PPI 3364 is common practice when normalizing costs in the aerospace industry. Additionally, total costs in units of dollars were provided and maintained in units of thousands of dollars. These labor cost data included costs at the prime mission equipment. WBS level prime mission equipment costs are directly related to touch labor and experience learning. Depending on data availability, additional elements below the prime mission equipment WBS elements were also analyzed to include engines and wings among other

sub-systems and sub-components. Therefore, one program may contribute several unique components for learning curve analysis.

This final sample included direct recurring cost data from bomber, cargo, and fighter aircraft along with missiles and munitions. The programs in this dataset are both historical and contemporary spanning 1957 to 2018 and include a variety of defense contractors. This diverse dataset tested the generalizability of Boone's Learning Curve due to the varying levels of analysis along with the multitude of platforms, contractors, and production periods that foster various learning environments. In total, data from 123 weapon system programs were gathered with 258 unique components. Learning curve analysis will be performed on these unique components.

Data Collection

This dataset was created using DD Form 1921-1 "Functional Cost-Hour Report" and DD Form 1921-2 "Progress Curve Report" data obtained from the CADE Defense Automated Cost Information Management System (DACIMS). CADE DACIMS is a repository of DoD weapons system program cost data available to DoD analysts. Some historical data was also extracted from the AFLCMC Cost Research Library.

Business rules were created to avoid overfitting the data and to ensure learning curve analysis was appropriate to model each component's cost. The first business rule omitted programs with production lots of four or fewer. This business rule is consistent with Boone (2018, p. 22) and limited the sample from 258 to 169 unique components. A second business rule was also necessary when performing Cumulative Average Theory analysis. Cumulative Average Theory relies on continuous data because each lot's cumulative average cost and cumulative quantity is a function of all previous lots' costs and quantities. Therefore, if a program's production lot was missing cost or quantity data, all lots after that missing lot were removed for that program; however, all lots before the missing lot were

retained[2]. These lot removals decreased the number of lots in the total program. This reduced number of lots warranted the complete removal of some programs by applying the first business rule. This second business rule limited the dataset to 140 unique components for Cumulative Average Theory analysis. Despite these business rules, there was not a systematic elimination of any characteristic of the program labor cost data; a diverse dataset remained with the previously stated attributes.

Data Analysis

Overview

This analysis will examine whether Boone's Learning Curve more accurately explains variability in program labor cost data than the traditional theories. Both Cumulative Average Theory and Unit Theory will be used to make these comparisons. In order to test these hypotheses, learning curve parameters were estimated using each program's labor cost data for Boone's Learning Curve and the traditional learning curve models. Next, parameters from Boone's Learning Curve and the respective traditional theory were used to predict a learning curve. These predicted learning curves were then compared to the observed data. In order to utilize Unit Learning Curve Theory, Boone's Learning Curve was adapted from its original Cumulative Average Theory form to Unit Learning Curve Theory form as shown in Equation 13.

---

[2] The researchers utilized the direct Cumulative Average Theory method instead of the iterative Cumulative Average Theory method. Although the iterative method can estimate learning curve parameters with missing lots, the direct method was utilized to align analysis with the original Cumulative Average Theory.

$$Y = Ax^{b/(1+\frac{x}{c})} \tag{13}$$

Where:

Y = cost of the xth unit
A = theoretical cost of the first unit (T1)
x = sequential unit number of the unit being calculated
b = $\frac{\ln \textit{Learning Curve Slope}}{\ln 2}$
c = decay value (positive constant)

When Boone's Learning Curve utilized Cumulative Average Theory, Boone's Learning Curve was compared to Wright's Learning Curve. When Boone's Learning Curve utilized Unit Theory, Boone's Learning Curve was compared to Crawford's Learning Curve.

Parameters for each learning curve were estimated using non-linear optimization techniques in Microsoft Excel. The traditional learning curve theories could be estimated using OLS regression. However, non-linear optimization was utilized to estimate the traditional curves for equitable comparison with Boone's Learning Curve. In contrast, Boone's Learning Curve required the use of non-linear optimization techniques. This requirement spawns from the fact that Boone's Learning Curve is not linear when logarithmically transformed due to the decaying learning curve slope. This non-linearity of Boone's Learning Curve precludes the parameters from being estimated using OLS regression.

The learning curve parameters for Boone's Learning Curve (i.e., "A", "b", and "c" from Equations 2 & 13) and the traditional theories (i.e., "A" and "b" from Equations 1 & 4) were estimated by minimizing the SSE using Excel's Generalized Reduced Gradient (GRG) Nonlinear Solver and Excel's Evolutionary Solver engines. The SSE is calculated by squaring the vertical difference of the observed data and predicted data for each lot and summing these squared differences across all lots. The SSE is calculated separately for Boone's Learning Curve and the traditional learning curves. For each model, Excel Solver is set to minimize the objective cell that is set as the SSE. The changing variable cells are the learning curve parameters specific to each learning curve model. These parameters are iteratively solved

for using optimization techniques specific to each engine for each learning curve model. When using Evolutionary Solver, it is also necessary to bound the changing variable cells and choose a set of values from which to begin the optimization process. Due to the inherent differences in both Cumulative Average Theory and Unit Learning Curve Theory, different specific processes were used to estimate parameters for each.

Cumulative Average Learning Curve Theory

The following process was implemented to estimate parameters for Wright's Learning Curve and Boone's Learning Curve using Cumulative Average Theory for each program.

1. Wright's Learning Curve parameters "A" and "b" were initially estimated using OLS regression.
    a. *Cumulative Average Cost* was the dependent variable, while *Cumulative Number of Units Produced* was the independent variable.
2. These initial learning curve parameter estimates were used as starting values to more precisely estimate Wright's Learning Curve parameters using GRG Non-Linear Solver. This process generated final estimates for Wright's Learning Curve parameters.
3. Boone's Learning Curve parameters "A," "b," and "c" were estimated using Excel's Evolutionary Solver. This process generated initial estimates for Boone's Learning Curve parameters.
    a. Final estimates for Wright's Learning Curve parameters were used to calculate the upper and lower bounds of Evolutionary Solver.
    b. The starting values were calculated from the upper and lower bounds.
4. The Evolutionary Solver learning curve parameter estimates for Boone's Learning Curve were used as starting values to more precisely estimate parameters using GRG Non-Linear Solver. This process produced final estimates for Boone's Learning Curve parameters.

When estimating Wright's Learning Curve parameters, the GRG Nonlinear Solver technique should produce SSE at least equal to the SSE using OLS regression. The GRG Nonlinear Solver technique was appropriate to estimate Wright's Learning Curve parameters because this technique is used to find locally optimal solutions of smooth and non-linear functions (Solver Technology – Smooth Nonlinear Optimization, 2012). Because OLS regression was used to provide starting values for GRG Nonlinear Solver, it was reasonable to assume that the global minimum is within this local region approximated using OLS regression. However, GRG Nonlinear Solver cannot guarantee that a global minimum is found. The GRG Nonlinear Solver Multistart method was also utilized to ensure this technique yielded a global minimum. However, the Multistart method failed to provide consistent, reliable parameter estimates, unlike GRG Nonlinear Solver.

When estimating Boone's Learning Curve, the GRG Nonlinear Solver optimized values for Wright's Learning Curve parameters were used to calculate bounds for Boone's Learning Curve. For the "A" parameter, the lower bound was half of Wright's "A" parameter, and the upper bound was twice that of Wright's "A" parameter. For bounds on Boone's Learning Curve "b" parameter, the absolute value of Wright's Learning Curve "b" parameter was multiplied by 3 to yield upper (positive) and lower (negative) bounds. Lastly, for Boone's Decay Value "c," bounds were set between 0 and 500,000 in contrast to Boone's original bounds of 0 and 5,000. This bound was increased in comparison to Boone (2018) due to the upper bound being binding in Boone's analyses. Except for Boone's Decay Value, these bounds are consistent with Boone (2018) and provide the optimization model with a restricted range to decrease the search time for an optimal solution but a broad enough range to not constrain the model. Similar to Boone (2018), none of these constraints were binding except for the Decay Value "c" upper bound despite relaxing this constraint in comparison to Boone (2018). Further relaxing this constraint would not have led to substantive changes to Boone's Learning Curve; as Boone's Decay Value "c" approaches infinity, Boone's Learning Curve transforms into Wright's Learning Curve.

In order to estimate Boone's Learning Curve, the starting values were set as the midpoint of the negative values of the slope parameter "b" because negative values are those that represent learning curve slopes below 100%. This starting point is reasonable because most programs in this analysis will experience cost efficiencies from learning. Thus, these programs will have learning curve slopes at or below 100% that translate to a negative learning curve exponent "b" parameter. The starting values for Boone's Learning Curve first unit cost "A" parameter were the midpoint of the upper and lower bounds. These two starting values and bounds depend on the parameters estimated by Wright's Learning Curve. The last starting value was set as the midpoint between the upper and lower bounds of the Boone's Decay Value "c," which was static at 250,000 due to the static bounds.

Once these starting values and bounds were set, Evolutionary Solver was used to estimate a globally optimal solution. The Evolutionary Solver technique was appropriate to estimate Boone's Learning Curve parameters because this technique is used to find a globally optimal solution of smooth and non-smooth functions (Solver Technology – Global Optimization, 2016). In contrast to estimating Wright's Learning Curve, a local region with the global minimum cannot be reliably approximated before using this optimization technique. This local region cannot be reliably approximated because OLS regression cannot be used to provide starting values for Boone's Learning Curve. Similar to GRG Nonlinear Solver, Evolutionary Solver cannot guarantee the parameter estimates produce a global minimum. However, using the Evolutionary Solver solution as starting values, the GRG Nonlinear Solver was then used to ensure the solution is locally optimal.

Unit Learning Curve Theory

Unit Learning Curve Theory parameter estimation maintained the Cumulative Average Learning Curve Theory parameter estimation methodology for calculating various bounds and starting values. Additionally, the justification for utilizing both Excel Solver engines also remains the same. However, the inclusion of lot midpoint calculations required different analysis techniques. The following process was

implemented to estimate parameters for Crawford's Learning Curve and Boone's Learning Curve using Unit Theory for each program.

1. Parameter-free lot midpoint approximations (Equation 6) were calculated for each production lot.
2. Crawford's Learning Curve parameters "A" and "b" were initially estimated using OLS regression.
    a. *Average Unit Cost* was the dependent variable while *Lot Midpoint*, calculated in Step 1, was the independent variable.
3. These initial learning curve parameter estimates were used as starting values to more precisely estimate Crawford's Learning Curve parameters using GRG Non-Linear Solver. This process generated intermediate estimates of Crawford's Learning Curve parameters.
4. The intermediate estimate of Crawford's Learning Curve "b" parameter was used to calculate a more precise set of lot midpoints using Asher's Approximation (Equation 7).
5. Applying these more precise lot midpoint approximations, Crawford's Learning Curve parameters "A" and "b" were more accurately estimated using GRG Nonlinear Solver.
    a. Steps 4 and 5 were repeated until the iterative process converged on a solution to produce final estimates of Crawford's Learning Curve parameters and lot midpoint approximations.
6. Parameter-free lot midpoint approximations (Equation 6) were used to estimate Boone's Learning Curve parameters "A," "b," and "c" using Excel's Evolutionary Solver. This process generated intermediate estimates of Boone's Learning Curve parameters.
    a. The final estimates for Crawford's Learning Curve parameters were used to calculate upper and lower bounds.
    b. The starting values were calculated from the upper and lower bounds.

7. The intermediate estimates of Boone's Learning Curve parameters "b" and "c" were used to approximate a more precise set of lot midpoints using Asher's Approximation adapted for Boone's Learning Curve (Equation 14).
8. Applying these more precise lot midpoint approximations, Boone's Learning Curve parameters were more accurately estimated using Evolutionary Solver. This process generated Evolutionary Solver parameter estimates of Boone's Learning Curve.
9. These Evolutionary Solver parameter estimates were used as starting values to improve further the accuracy of Boone's Learning Curve parameters estimates "A," "b," and "c" using GRG Non-Linear Solver.
    a. Steps 7, 8, & 9 were repeated until the iterative process converged on a solution to produce a final estimate of Boone's Learning Curve parameters and lot midpoint approximations.

For both Crawford's and Boone's Learning Curves, an iterative process is used to calculate precise parameter estimates and lot midpoint approximations. This iterative process was repeated until a solution converged. A solution converged when small changes (1×10^{-50}) in the learning curve exponent "b" parameter were calculated between iterations. This process of iterative solving was adapted from Hu and Smith's "Accuracy matters" (2013). For Boone's Learning Curve, a limit of 10 iterations was placed on the iterative process. This limit of 10 iterations was reached a limited number of times and still produced relatively small differences of Boone's Learning Curve exponent "b" between iterations.

In order to estimate lot midpoints and Boone's Learning Curve parameters, Asher's Approximation from Equation 7 was adapted to incorporate Boone's decaying learning curve slope. Asher's Approximation adapted for Boone's Learning Curve is shown in Equation 14. Equation 14 is undefined when Boone's Learning Curve "b" parameter is equal to zero, similar to Equation 7. Equation

14 uses previously calculated lot midpoints that are predicated from previously estimated Boone's Learning Curve parameters.

$$Lot\ Midpoint_i \approx \left[\frac{(L+\frac{1}{2})^{b'+1}-(F-\frac{1}{2})^{b'+1}}{(L-F+1)(b'+1)}\right]^{(\frac{1}{b'})} \tag{14}$$

Where:

F = the first unit number in a lot
L = the last unit number in a lot
$b' = \frac{b}{1+(\frac{LMP_{i-1}}{c})}$
i = the iteration number

Statistical Significance Testing

The estimated parameters for Boone's Learning Curve and the traditional learning curves were used to create predicted learning curves. These predicted curves were then compared to observed data. Total model error was calculated by comparing the difference between observations and predicted values to determine which learning curve theory more accurately explained variability in the data. Two measures were used to determine the overall model error. The first error measure was RMSE. RMSE is calculated by dividing the total SSE by the number of observations, in our case, the number of lots in that program (McClave, Benson, & Sincich, 2014, p. 14-25). RMSE is a scale-dependent measure. Additionally, RMSE is not robust to outliers; therefore, the greater the magnitude of an outlier from the average error values, the more influence this outlier will have on RMSE. RMSE was used instead of total SSE because RMSE transforms units from squared units to original units. This transformation eases interpretation. RMSE can be interpreted as the average amount of error of the model in the model's original units.

The second measure used to determine the overall model error was MAPE. MAPE is calculated by subtracting the predicted value from the observed value, dividing this difference by the observed value, taking the absolute value, and multiplying by 100%; these absolute percent errors are then summed over all observations and divided by the total number of observations (McClave et al., 2014, p. 14-25). MAPE provides a unitless measure of accuracy and can be interpreted as the average percentage the model is inaccurate. MAPE is robust to outliers, so the effects of outliers do not unduly influence this measure.

After calculating these measures of overall model error, a series of paired difference t-tests will be conducted to determine if reductions in error from Boone's Learning Curve are significantly different than zero or due to random chance. In order to conduct the first paired difference t-test, Boone's Learning Curve RMSE using Cumulative Average Theory will be subtracted from Wright's Learning Curve RMSE, and the difference will be divided by Wright's Learning Curve RMSE by observation. This calculation will yield a percentage difference rather than raw difference to compare programs of varying differences in magnitude equitably. Only programs with errors reported in total dollars will be examined first to examine the results of the different unit measures. The null hypothesis for this test is that the percentage difference in RMSE is equal to or less than zero. This null hypothesis represents that Boone's Learning Curve results in an equal amount of or more error in predicting observed values in comparison to Wright's Learning Curve. The alternative hypothesis is that the percentage difference is greater than zero. The alternative hypothesis represents that Boone's Learning Curve results in less error in predicting observed values than Wright's Learning Curve. This same methodology will be repeated five more times to examine both learning curve theories each with two measures of model error for programs examined using labor hours and total dollars. The attributes that warranted six paired difference t-tests be conducted are shown in Table 1. A significance value of 0.05 will be used to

determine if the results are statistically significant. JMP Pro Version 13 will be used to perform this and all the following statistical analyses.

Table 1: Set of Paired Difference Hypothesis Tests to be Conducted

Learning Curve Theory	**Error Measure**	**Units of Measure**	**Hypothesis Test**
Cumulative Average Theory	Root Mean Squared Error Percentage Difference	Total Dollars (K)	1
		Labor Hours	2
	Mean Absolute Percent Error Percentage Difference	Total Dollars (K) & Labor Hours Combined	3
Unit Theory	Root Mean Squared Error Percentage Difference	Total Dollars (K)	4
		Labor Hours	5
	Mean Absolute Percent Error Percentage Difference	Total Dollars (K) & Labor Hours Combined	6

An assumption to utilize the paired difference t-test is that the data are approximately normally distributed (McClave et al., 2014, p. 441-442). For hypothesis tests with a sample size greater than or equal to 30, the Central Limit Theorem guarantees this assumption is met due to the large sample (McClave et al., 2014, p. 441). For hypothesis tests with a sample size of less than 30, the distribution will be tested for normality using a Shapiro Wilk Test. If the Shapiro Wilk test rejects the null hypothesis of normality, the paired difference t-test cannot be used. Instead, a Wilcoxon Rank Sum test will be used because the assumptions for this test do not depend on the shape of the distribution (McClave et al., 2014, pp. 15-10 – 15-11).

Analysis & Results

Cumulative Average Theory Comparison

Wright's Learning Curve and Boone's Learning Curve using Cumulative Average Theory are compared in Table 2 for a random sample of 30 of the 140 total components tested. The full list of components is included in Appendix A. This Cumulative Average Theory sample included 118 components analyzed in units of total dollars and 22 components analyzed in units of labor hours. Each entry lists the program number, the number of production lots, the number of production units, the component estimated, and the units in which the program was analyzed. Additionally, each entry lists both error measures and their respective percentage differences between Wright's Learning Curve and Boone's Learning Curve. Positive percentage differences in either error measure indicate Boone's Learning Curve had less error than Wright's Learning Curve in explaining program costs. Negative percentage differences indicate that Boone's Learning Curve had more error than Wright's Learning Curve in explaining program costs.

Table 2: Sample of Comparisons using Cumulative Average Theory

Program	Number of Lots	Number of Units	Component Estimated	Units	Traditional RMSE	Boone RMSE	RMSE Percentage Difference	Traditional MAPE	Boone MAPE	MAPE Percentage Difference
Program 1	6	483	Airframe	Dollars	411.22	114.07	72.3%	2.8%	0.7%	74.7%
Program 7	7	110	Electronic Warfare (2)	Dollars	140.32	107.21	23.6%	1.2%	0.8%	27.5%
Program 10	10	3803	PME - Air Vehicle	Hours	24.45	14.01	42.7%	4.3%	2.0%	54.0%
Program 12	10	20	PME - Air Vehicle	Dollars	699.20	694.08	0.7%	5.8%	5.7%	1.0%
Program 16	9	76	PME - Air Vehicle	Dollars	436.29	144.41	66.9%	2.6%	1.0%	62.9%
Program 20	11	84	PME - Air Vehicle	Dollars	1568.74	1121.89	28.5%	1.7%	1.5%	7.8%
Program 21	6	326	PME - Air Vehicle	Dollars	5267.10	2408.78	54.3%	8.0%	4.2%	47.4%
Program 21	7	344	Airframe	Dollars	4819.45	2544.26	47.2%	9.1%	5.4%	40.4%
Program 21	14	453	PME - Air Vehicle	Hours	3493.62	3495.94	-0.1%	4.8%	4.8%	0.1%
Program 21	14	453	Airframe	Hours	4338.35	4339.68	0.0%	6.2%	6.2%	0.1%
Program 27	18	631	PME - Air Vehicle	Dollars	1669.56	913.34	45.3%	3.6%	1.9%	46.2%
Program 33	10	178	Airframe	Dollars	1906.94	1910.76	-0.2%	1.7%	1.7%	-0.2%
Program 33	10	178	Electronic Warfare (3)	Dollars	62.5	62.4	0.2%	5.7%	5.7%	0.1%
Program 34	6	67	PME - Air Vehicle	Hours	9058.6	9061.7	0.0%	4.4%	4.4%	0.0%
Program 34	6	201	Body	Dollars	1924.5	828.9	56.9%	19.0%	8.7%	54.0%
Program 34	6	67	Electrical	Dollars	50.7	50.7	-0.1%	1.9%	1.9%	-0.1%
Program 34	5	49	Empennage	Dollars	202.2	202.2	0.0%	4.1%	4.1%	0.0%
Program 34	6	67	EO/IR	Dollars	45.6	36.6	19.7%	1.2%	1.1%	13.1%
Program 35	5	50	Electronic Warfare (1)	Dollars	259.6	259.7	0.0%	3.2%	3.2%	0.0%
Program 35	5	50	Hydraulic	Dollars	58.2	58.2	0.0%	3.1%	3.1%	0.0%
Program 35	5	50	Radar	Dollars	256.8	256.9	0.0%	3.2%	3.2%	0.0%
Program 35	5	50	Surface Controls	Dollars	121.5	121.5	0.0%	2.6%	2.6%	0.0%
Program 36	13	1285	PME - Air Vehicle	Dollars	28.8	29.4	-2.1%	0.6%	0.6%	-2.2%
Program 38	6	52	PME - Air Vehicle	Dollars	253.6	154.9	38.9%	1.2%	0.7%	41.6%
Program 46	6	68	Airframe	Dollars	539.1	527.6	2.1%	2.3%	2.1%	10.9%
Program 46	6	68	Electronic Warfare (1)	Dollars	221.8	221.9	0.0%	5.4%	5.4%	0.0%
Program 46	6	68	EO/IR	Dollars	120.7	120.8	0.0%	15.7%	15.7%	0.0%
Program 47	9	36	Data Link (1)	Dollars	170.2	170.2	0.0%	17.7%	17.7%	0.0%
Program 55	9	677	PME - Air Vehicle	Dollars	74.8	74.8	0.0%	1.6%	1.6%	0.0%
Program 56	5	590	PME - Air Vehicle	Dollars	6.6	6.6	0.5%	0.2%	0.2%	6.3%

Boone's Learning Curve incorporates an additional parameter into the traditional learning curve equation. This additional parameter should be able to theoretically increase to such a degree to transform Boone's Learning Curve into Wright's Learning Curve. Therefore, Boone's Learning Curve should be able to explain program component learning curve costs to at least the same degree of accuracy as the traditional learning curve theories. Greater amounts of accuracy could also be explained by Boone's Learning Curve if its functional form allowed it to more accurately model learning curves in comparison to the traditional theory. Despite these theoretical explanations, Boone's Learning Curve had more error than Wright's Learning Curve for some observations in explaining program costs as indicated by negative percentage differences. These negative percentage differences occur because an upper bound was placed on Boone's Decay value that restricted Boone's Learning Curve from transforming into Wright's Learning Curve[3]. Therefore, some observations' percentage error differences

---

[3] The researchers tested this claim by selecting programs from the dataset, increasing the upper bound of Boone's Decay value to 500,000,000, and re-estimating the parameters. The error of the programs with an increased upper

are approximately equal to but not exactly zero. Bounds were set to determine if an observation's percentage error difference is approximately equal to zero to account for these insignificant percentage error differences. The data were visualized to set bounds in order to identify observations with percentage error differences of approximately zero. Observations with percentage error differences of approximately zero were defined as those within the bounds (-0.25%, 0.25%). These bounds were used by the researchers to distinguish between observations with approximately zero and non-zero percentage error differences in order to better inform the descriptive statistics.

Boone's Learning Curve had less error for approximately 41% of observations across all percentage difference error measures. Boone's Learning Curve error was approximately equal to Wright's Learning Curve for approximately 50% of observations across all error measures. Lastly, Boone's Learning Curve had more error for approximately 9% of observations. Appendix B provides descriptive statistics by group for each percentage difference error measure. The relative similarity of the percentage of observations within the approximately zero group among the three error measures provides further justification for these bounds.

The mean and median amounts of percentage error reduction and the accompanying standard deviations for the three error measures are shown in Table 3. Because Boone's Learning Curve is an improvement on Wright's for a limited number of observations, many instances of approximately zero percentage differences are included in the positively skewed distribution. These inclusions of approximately zero also have a substantial influence on the descriptive statistics. For this reason, Table 4 displays descriptive statistics given Boone's Learning Curve was an improvement on Wright's Learning Curve. The mean for the RMSE percentage difference from Table 4 for program components analyzed in

---

bound for Boone's Decay value were compared to the error of programs with the original upper bound of 500,000. Error decreased for programs with increased upper bounds for Boone's Decay value. These results indicate that Boone's Learning Curve error will converge to Wright's Learning Curve error as Boone's Decay value approaches infinity as predicted.

units of thousands of dollars [Dollars (K)] can be interpreted as Boone's Learning Curve was on average 44% more accurate than Wright's Learning Curve for the set of program components in which Boone's Learning Curve was an improvement on Wright's. However, the relatively high standard deviation of all error measures indicate that Boone's Learning Curve causes widely variable reductions in error in comparison to Wright's Learning Curve.

Table 3: Cumulative Average Theory Descriptive Statistics for All Programs

Error Measure	Program Units	Number of Observations	Mean	Median	Standard Deviation
Root Mean Squared Error Percentage Difference	Total Dollars (K)	118	19.3%	0.0%	28.9%
	Labor Hours	22	15.2%	0.0%	31.2%
Mean Absolute Percent Error Percentage Difference	Total Dollars (K) & Labor Hours Combined	140	18.6%	0.0%	29.5%

Table 4: Cumulative Average Theory Descriptive Statistics Given Boone's Learning Curve Improved Upon Wright's

Error Measure	Program Units	Number of Observations	Mean	Median	Standard Deviation
Root Mean Squared Error Percentage Difference	Total Dollars (K)	52	43.9%	44.5%	28.5%
	Labor Hours	8	45.4%	34.4%	34.4%
Mean Absolute Percent Error Percentage Difference	Total Dollars (K) & Labor Hours Combined	60	44.1%	44.3%	29.6%

Cumulative Average Theory: Statistical Difference Testing

The results of the paired difference t-tests for Cumulative Average Theory are shown in Figures 10, 11, & 12. Figure 10 displays the results for program components analyzed in underlying units of total dollars for the RMSE percentage difference measure. Figure 11 displays the results for program components analyzed in underlying units of labor hours for the RMSE percentage difference measure.

Lastly, Figure 12 displays the results for all program components for the MAPE percentage difference measure. Paired difference tests were used to determine if reductions in error from Boone's Learning Curve are statistically significant from zero or due to random chance. The applicable p-values within these figures were the Prob > t row because the rejection region is the positive side of the distribution; however, additional p-values are also provided for full disclosure. No outliers were present in any of the paired difference t-tests. An observation was identified as an outlier if the observation's value fell more than three interquartile ranges from the upper 90% and lower 10% quantiles.

For the RMSE percentage difference measure in underlying units of total dollars shown in Figure 10, the paired difference t-test resulted in a test statistic of 7.23 with a p-value of <0.001. At a significance level of 0.05, this test rejects the null hypothesis because the p-value is less than the significance level. This result indicates that Boone's Learning Curve reduces the amount of error in comparison to Wright's for the RMSE percentage difference measure in underlying units of total dollars to a statistically significant degree.

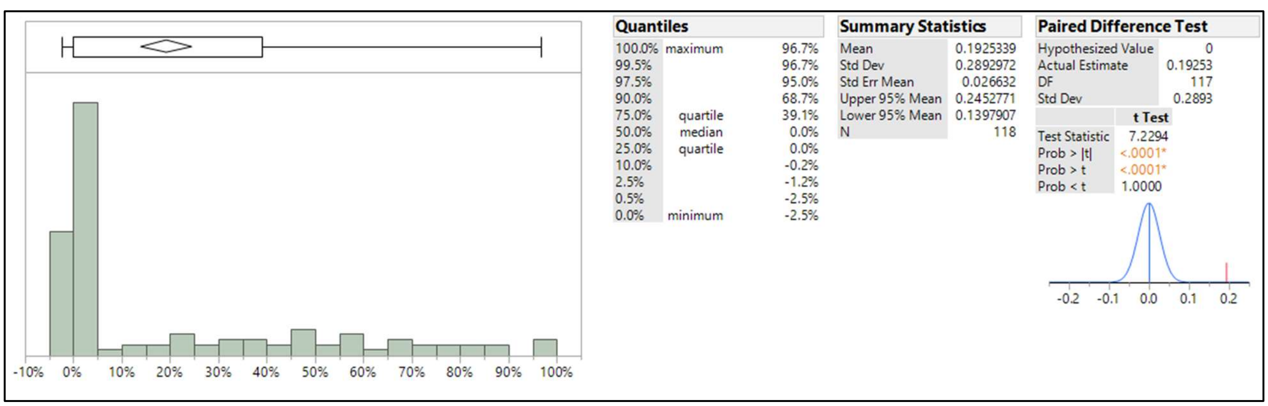


Figure 10: Cumulative Average Theory Percentage Difference in RMSE in Units of Dollars

For the RMSE percentage difference measure in underlying units of hours hypothesis test shown in Figure 11, the sample size is fewer than 30 observations. Due to this small sample size, the Central

Limit Theorem cannot be invoked, and the data cannot be assumed to be normally distributed (McClave et al., 2014, p. 441). The paired difference t-test assumes that the data are approximately normally distributed (McClave et al., 2014, p. 441-442). To statistically test if the data approximately fit a normal distribution, the Shapiro Wilk test was utilized. The Shapiro Wilk test statistic was 0.721 with a p-value of <0.001. This test rejects the null hypothesis that the data are normally distributed because the p-value is less than the significance level of 0.05. For these reasons, the Wilcoxon Signed Rank Test was used because it does not rely on normally distributed data. All assumptions of the Wilcoxon Signed Rank were met (McClave et al., 2014, p. 15-11). The Wilcoxon Signed Rank test resulted in a test statistic of 18.5 that corresponds to a p-value of 0.280. Because this p-value is greater than the significance level of 0.05, this test fails to reject the null hypothesis. This result indicates that Boone's Learning Curve does not reduce the amount of error in comparison to Wright's for the RMSE percentage difference measure in underlying units of hours to a statistically significant degree. However, small sample sizes can cause paired difference tests to have low power (Cohen, 1938). Hypothesis tests with low power can result in incorrectly failing to reject the null hypothesis (Cohen, 1938).

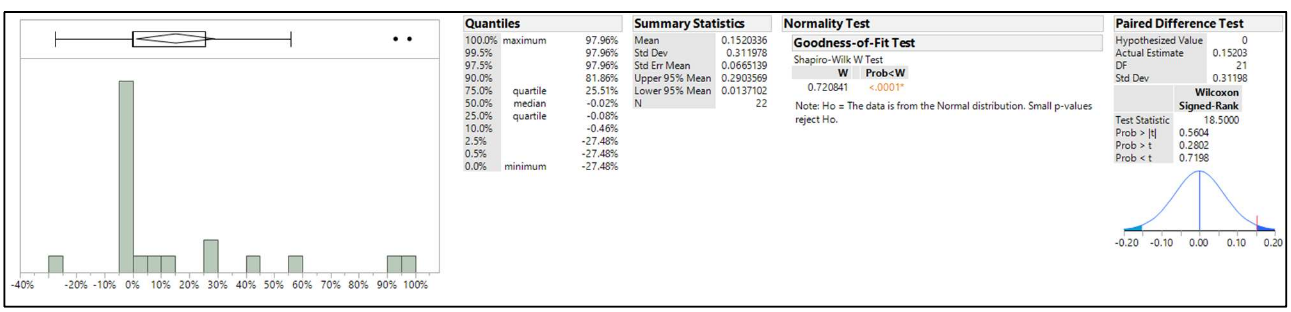


Figure 11: Cumulative Average Theory Percentage Difference in RMSE in Units of Hours

For the MAPE percentage difference measure shown in Figure 12, the paired difference t-test statistic was 7.45. This test statistic corresponded to a p-value of <0.0001. Because this p-value is less than the significance level of 0.05, this test rejects the null hypothesis. This result indicates that Boone's

Learning Curve reduces the amount of error in comparison to Wright's for the MAPE percentage difference measure to a statistically significant degree.

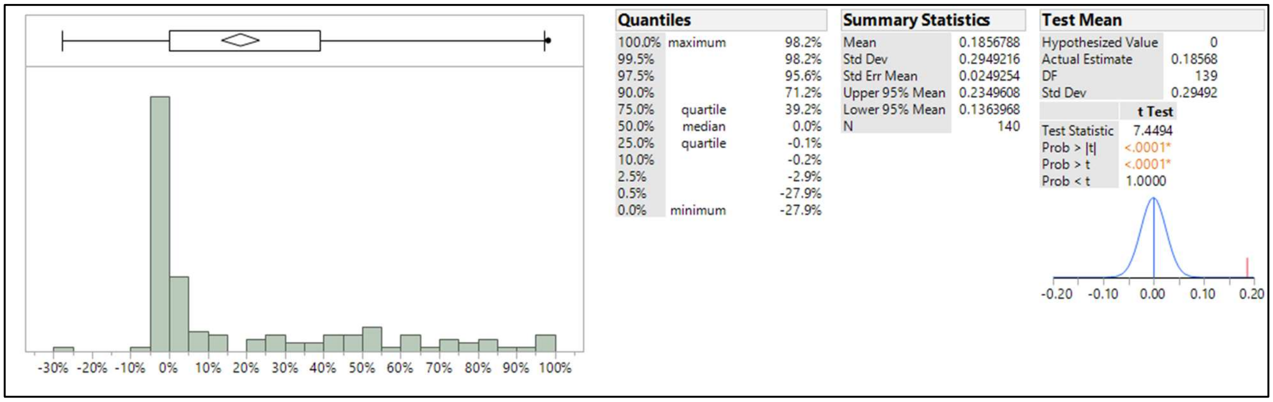


Figure 12: Cumulative Average Theory Percentage Difference in MAPE

Unit Theory Comparison

Crawford's Learning Curve and Boone's Learning Curve using Unit Theory are compared in Table 5 for a random sample of 30 of the 169 total program components tested. The full list of components is included in Appendix A. This Unit Theory sample included 141 components analyzed in units of total dollars and 28 components analyzed in units of labor hours. Each entry and all interpretations match that of the Cumulative Average Theory Comparison section.

Table 5: Sample of Comparisons using Unit Theory

Program	Number of Lots	Number of Units	Component Estimated	Units	Traditional RMSE	Boone RMSE	RMSE Percentage Difference	Traditional MAPE	Boone MAPE	MAPE Percentage Difference
Program 1	7	503	Airframe	Dollars	2383.2	857.9	64.0%	14.6%	4.9%	66.4%
Program 6	7	459	Electronic Warfare (1)	Dollars	20.9	20.9	0.0%	30.8%	30.8%	0.0%
Program 7	5	321	PME - Air Vehicle	Dollars	37.9	33.3	12.2%	3.8%	3.8%	1.1%
Program 8	6	98	Electronic Warfare (3)	Dollars	5.2	4.9	6.1%	4.8%	4.8%	1.4%
Program 8	6	98	Electronic Warfare (1)	Dollars	27.5	18.7	31.9%	10.2%	5.9%	42.5%
Program 10	9	1586	PME - Air Vehicle	Dollars	115.5	115.6	-0.2%	12.5%	12.5%	-0.2%
Program 13	10	3803	PME - Air Vehicle	Dollars	33.6	24.8	26.1%	10.3%	7.5%	27.1%
Program 15	7	11	Mission Computer (1)	Dollars	213.9	213.9	0.0%	11.6%	11.5%	0.6%
Program 18	11	83	PME - Air Vehicle	Dollars	82138.6	82143.3	0.0%	23.2%	23.2%	0.0%
Program 25	7	59	Electronic Warfare (1)	Dollars	1259.1	653.3	48.1%	16.1%	7.1%	55.6%
Program 27	14	453	PME - Air Vehicle	Hours	54142.9	53766.4	0.7%	59.9%	63.1%	-5.4%
Program 29	11	433	Electronic Warfare (1)	Dollars	57.5	57.5	0.0%	13.5%	13.5%	-0.1%
Program 30	5	469	PME - Air Vehicle	Dollars	1283.8	891.8	30.5%	13.5%	8.3%	38.3%
Program 31	10	59	PME - Air Vehicle	Dollars	11978.9	11979.3	0.0%	8.6%	8.6%	0.0%
Program 33	5	109	PME - Air Vehicle	Dollars	6824.7	6824.8	0.0%	28.2%	28.2%	0.0%
Program 34	18	631	PME - Air Vehicle	Dollars	6926.7	2799.9	59.6%	17.0%	6.6%	61.0%
Program 35	7	522	PME - Air Vehicle	Hours	4615.3	4458.5	3.4%	6.3%	6.1%	3.1%
Program 35	7	522	Airframe	Hours	6757.0	6280.7	7.0%	5.7%	5.4%	4.8%
Program 37	5	204	PME - Air Vehicle	Dollars	1468.7	921.0	37.3%	2.9%	1.9%	36.4%
Program 40	10	178	Electronic Warfare (1)	Dollars	1642.3	1643.0	0.0%	20.7%	20.7%	0.0%
Program 41	6	67	Mission Computer (1)	Dollars	1698.1	1542.4	9.2%	4.6%	3.7%	19.5%
Program 42	5	50	Alighting Gear	Dollars	78.6	77.4	1.5%	3.6%	3.5%	2.3%
Program 46	6	44	PME - Air Vehicle	Hours	7736.9	7255.3	6.2%	17.6%	16.7%	4.8%
Program 46	10	113	PME - Air Vehicle	Dollars	797.9	627.0	21.4%	3.8%	2.9%	22.7%
Program 54	9	134	PME - Air Vehicle	Dollars	1907.3	970.0	49.1%	11.8%	6.5%	44.9%
Program 57	6	68	Electronic Warfare (1)	Dollars	998.8	998.9	0.0%	58.9%	58.9%	0.0%
Program 57	6	68	Airframe	Dollars	1443.2	1285.1	11.0%	6.7%	5.4%	18.5%
Program 62	9	110	PME - Air Vehicle	Dollars	13027.5	13028.9	0.0%	24.0%	24.0%	0.0%
Program 67	9	677	PME - Air Vehicle	Dollars	273.5	273.5	0.0%	5.1%	5.1%	0.0%
Program 68	5	590	PME - Air Vehicle	Dollars	87.1	87.2	0.0%	2.8%	2.8%	0.0%

Similar to Cumulative Average Theory, Boone's upper limit of the Decay value precluded Boone's Learning Curve from converging to Crawford's Learning Curve. This bound resulted in insignificant percentage error differences between Boone's Learning Curve and Crawford's Learning Curve. Therefore, observations with percentage error differences of approximately zero were defined as those within the bounds (-0.25%, 0.25%). These bounds are the same as those used for Cumulative Average Theory. In order to ensure these bounds were justified, the distribution of the data around zero was reviewed. Boone's Learning Curve had less error for approximately 43% of observations across all percentage difference error measures in comparison to Crawford's Learning Curve. Boone's Learning Curve error was approximately equal to Crawford's Learning Curve for approximately 52% of observations across all error measures. Lastly, Boone's Learning Curve had more error for approximately 5% of observations. Appendix B provides descriptive statistics by group for each measure of percentage difference error.

The mean and median amounts of percentage error reduction and the accompanying standard deviations for the three error measures are shown in Table 6. Similar to Cumulative Average Theory analysis, many inclusions of approximately zero also have a strong influence on the descriptive statistics. For this reason, Table 7 displays descriptive statistics given that Boone's Learning Curve was an improvement on Crawford's Learning Curve. Similar to Cumulative Average Learning Curve analysis, the relatively high standard deviation of all error measures indicate that Boone's Learning Curve causes variable reductions in error in comparison to Crawford's Learning Curve.

Table 6: Unit Theory Descriptive Statistics for All Programs

Error Measure	Program Units	Number of Observations	Mean	Median	Standard Deviation
Root Mean Squared Error Percentage Difference	Total Dollars (K)	141	13.8%	0.0%	22.7%
	Labor Hours	28	6.0%	0.0%	14.8%
Mean Absolute Percent Error Percentage Difference	Dollars (K) & Labor Hours Combined	169	11.3%	0.0%	23.1%

Table 7: Unit Theory Descriptive Statistics Given Boone's Learning Curve Improved Upon Crawford's

Error Measure	Program Units	Number of Observations	Mean	Median	Standard Deviation
Root Mean Squared Error Percentage Difference	Total Dollars (K)	66	29.6%	26.7%	25.3%
	Labor Hours	12	14.1%	6.6%	20.3%
Mean Absolute Percent Error Percentage Difference	Dollars (K) & Labor Hours Combined	67	30.3%	24.5%	26.2%

Unit Theory: Statistical Difference Testing

The results of the paired difference testing for Unit Theory are shown in Figures 13, 14, & 15. These figures follow the same order as those provided in the Cumulative Average Theory analysis section. No outliers were present in any of the paired difference t-tests using the same methodology of

outlier identification used in the Cumulative Average Theory: Statistical Difference Testing sub-section. For the RMSE percentage difference measure in underlying units of total dollars shown in Figure 13, the standard paired difference t-test resulted in a test statistic of 7.23 with a p-value of <0.001. At a significance level of 0.05, this test rejects the null hypothesis because the p-value is smaller than the significance level. This result indicates that Boone's Learning Curve reduces the amount of error in comparison to Crawford's for the RMSE percentage difference measure in underlying units of total dollars to a statistically significant degree.

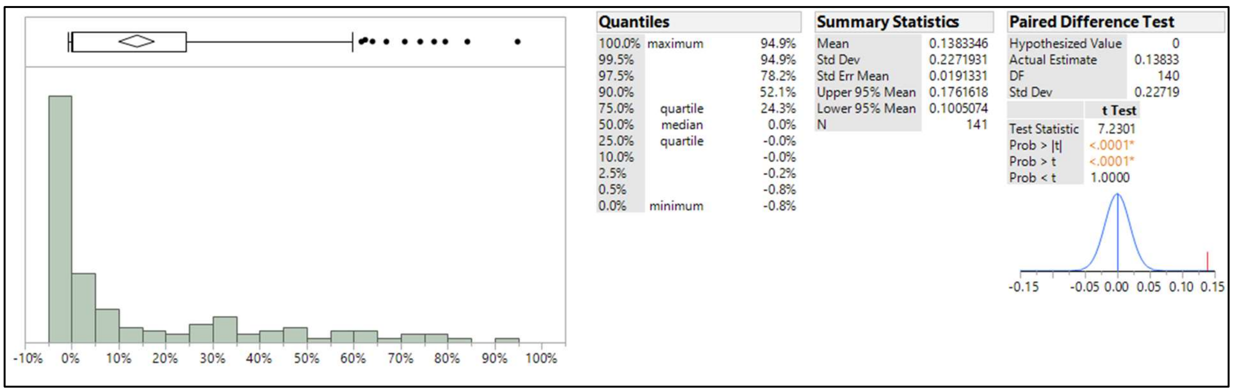


Figure 13: Unit Theory Percentage Difference in RMSE in Units of Dollars

For the RMSE percentage difference measure in underlying units of hours hypothesis test shown in Figure 14, the sample size is fewer than 30 programs. Similar to the Cumulative Average Theory analysis, a Shapiro Wilk test was utilized to test statistically if the data approximately fit a normal distribution. The Shapiro Wilk test statistic is 0.48 with a p-value of <0.001. This test rejects the null hypothesis that the data are normally distributed because the p-value is less than the significance level of 0.05. For these reasons, the Wilcoxon Signed Rank Test was used instead. All assumptions of the Wilcoxon Signed Rank were met (McClave et al., 2014, p. 15-11). The Wilcoxon Signed Rank test resulted in a test statistic of 74.00 that corresponds to a p-value of 0.0461. Because this p-value is less than the significance level of 0.05, this test rejects the null hypothesis. However, the two-tailed test statistic has a

p-value above the significance level and is provided for full disclosure. This result indicates that Boone's Learning Curve reduces the amount of error in comparison to Wright's for the RMSE percentage difference measure in underlying units of hours to a statistically significant degree.

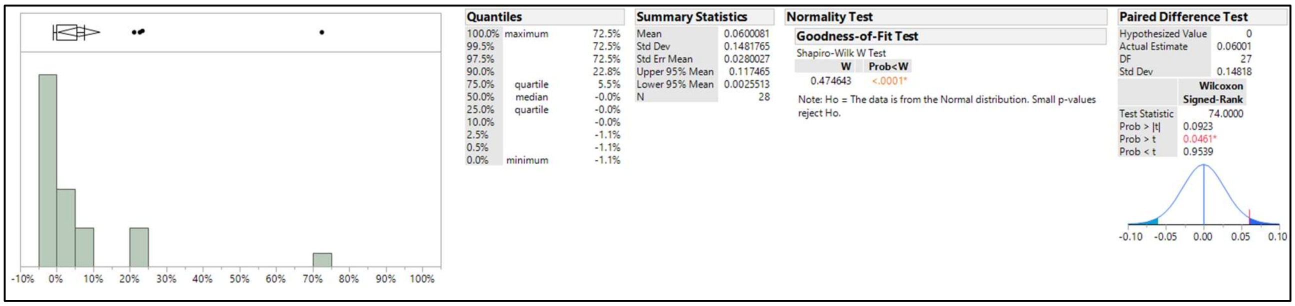


Figure 14: Unit Theory Percentage Difference in RMSE in Units of Hours

For the MAPE percentage difference measure shown in Figure 15, the paired different test statistic was 6.36. This test statistic corresponds to a p-value of <0.001. Because this p-value is less than the significance level of 0.05, this test rejects the null hypothesis. This result indicates that Boone's Learning Curve reduces the amount of error in comparison to Crawford's for the MAPE percentage difference measure to a statistically significant degree.

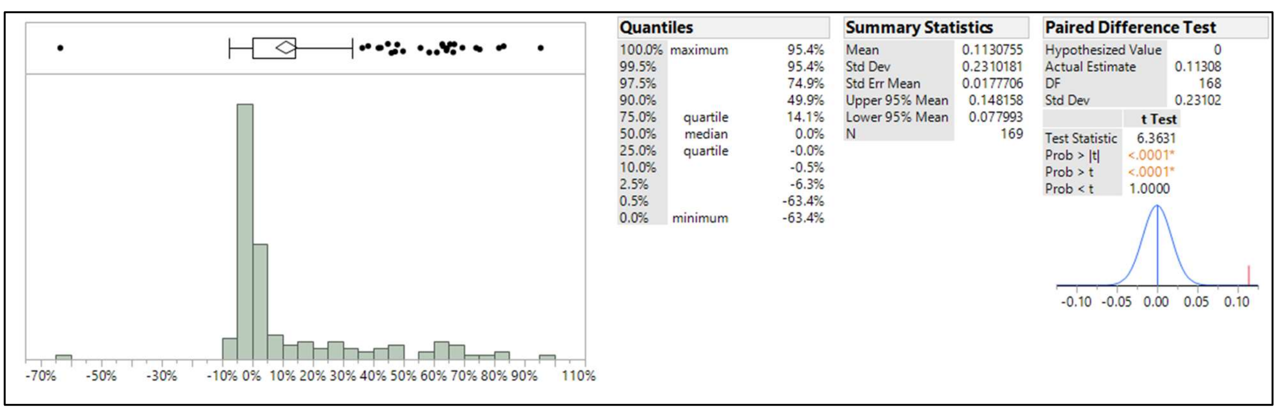


Figure 15: Unit Theory Percentage Difference in MAPE

Conclusions & Recommendations

Research Conclusions

A large, diverse dataset of Air Force production programs was used to test if Boone's Learning Curve more accurately explained error in comparison to the traditional learning curve theories. The direct recurring cost data from bomber, cargo, and fighter aircraft along with missiles and munitions programs in units of total dollars and labor hours were analyzed using Cumulative Average and Unit Learning Curve theories. Various components of these programs were analyzed from wings and data link systems to the airframes and air vehicles. Boone's Learning Curve was tested against both Cumulative Average and Unit Learning Curve theories using two different measures of model error that resulted in six paired difference tests. This methodology resulted in 998 total observations across all measures and ensured the generalizability of Boone's Learning Curve was tested.

Boone's Learning Curve improved upon the traditional learning curve theories for approximately 42% of the sampled program components while approximately equaling the traditional learning curve theory error for approximately 51% of program components. When Boone's Learning Curve improved upon the traditional learning curve theories, Boone's Learning Curve resulted in a range of mean percentage difference reductions of 14% to 45% across all measures. The standard deviations of these improvements were relatively high with coefficients of variation ($\frac{standard\ deviation}{mean}$) ranging from 65% to 144% across all measures. Absent additional analysis, these high amounts of variability make it challenging to conclude the degree to which Boone's Learning Curve will improve the accuracy of explaining program component costs in comparison to the traditional theories.

The paired difference tests between Boone's Learning Curve and the traditional theories indicate that Boone's Learning Curve reduces error to a statistically significant degree across a wide range of measures for a diverse set of Air Force and Navy programs. Table 8 summarizes the results for the six paired difference tests. Five of the six paired difference tests resulted in rejecting the null

hypothesis that Boone's Learning Curve had an equal amount or more error than the traditional theories at a significance level of 0.05. The subset of programs analyzed in units of hours for Cumulative Average Theory was a small sample; therefore, a failure to reject the null hypothesis may stem from the test's lack of power. The subset of programs analyzed in units of hours for Unit Theory was also a small sample; however, the test rejected the null hypothesis using a non-parametric test.

Table 8: Summary of Hypothesis Test Results

Learning Curve Theory	Error Measure	Units of Measure	Hypothesis Test	Results	Boone's Learning Curve Significatly Reduced Error
Cumulative Average Theory	Root Mean Squared Error Percentage Difference	Total Dollars (K)	1	Reject Null Hypothesis (p-value < 0.001)	Yes
		Labor Hours	2	Fail to Reject Null Hypothesis (p-value $= 0.280$)	Yes
	Mean Absolute Percent Error Percentage Difference	Total Dollars (K) & Labor Hours Combined	3	Reject Null Hypothesis (p-value < 0.001)	No
Unit Theory	Root Mean Squared Error Percentage Difference	Total Dollars (K)	4	Reject Null Hypothesis (p-value < 0.001)	Yes
		Labor Hours	5	Reject Null Hypothesis (p-value $= 0.046$)	Yes
	Mean Absolute Percent Error Percentage Difference	Total Dollars (K) & Labor Hours Combined	6	Reject Null Hypothesis (p-value < 0.001)	Yes

In summary, Boone's Learning Curve more accurately explained program cost data using both Cumulative Average Theory and Unit Theory in comparison to the traditional learning curves. Boone's Learning Curve modeled program component cost data more accurately than the traditional theories; however, it remains unclear which programs are more accurately modeled using Boone's Learning Curve and to what degree Boone's Learning Curve will more accurately model program component costs.

Research Limitations

Limitations of this research should be reviewed to draw appropriate conclusions and make suitable recommendations. This research sought to test the generalizability of Boone's Learning Curve using a variety of programs and their components. To achieve this goal, a large sample of Air Force and

Navy programs were utilized with a variety of attributes; however, small samples for components analyzed in units of labor hours precluded gaining compelling conclusions from all hypothesis tests. Additionally, program lot data was used instead of unitary data due to data availability. Although Boone's Learning Curve should perform just as well using either type of data, this research cannot conclusively state that Boone's Learning Curve will more accurately explain programs in unitary data. Lastly, the majority of data utilized were end-item components in units of total dollars. The total dollar cost includes all cost categories rather than solely labor costs. These data are not ideal when applying learning curve theory and may bias learning curves to display diminishing rates of learning. Despite these potential issues, total dollar cost data are regularly utilized by cost estimators in the field due to data availability. Therefore, the practical applications of this analysis remain valid despite the limitations of using imperfect total dollar cost data in learning curve analysis.

Recommendations for Future Research

Boone's Learning Curve was tested on programs whose lot costs were already known and whose parameters can be directly estimated. In other words, Boone's Learning Curve was tested against the traditional theories on how well it explained rather than predicted program costs. In practical use within the cost estimating community, Boone's Learning Curve should also be tested on how well it predicts rather than explains costs in comparison to the traditional theories. In order to utilize Boone's Learning Curve to predict costs, future research should investigate if Boone's Decay Value can be predicted using various attributes of a program. Tests should be performed on how well Boone's Learning Curve predicts costs for a program using prior lots of the same program or prior analogous programs in comparison to the traditional theories. Depending on how these prediction tests are implemented, these tests can indirectly examine if Boone's Learning Curve parameters quickly converge on stable values and if overfitting of the analogous learning curve data occurs among other topics. Lastly, additional labor hour

data should be collected and analyzed in order to dispel the potential bias of learning curves displaying diminishing rates of learning when analyzed in units of total dollars.

Summary

This research tested the generalizability of Boone's Learning Curve in explaining costs for up to 69 weapons system programs comprising 169 unique components. Among all analysis techniques, 998 observations were investigated. These programs had a variety of attributes to include programs analyzed in units of labor hours and total dollars. The programs spanned from 1957 to 2018 and were analyzed at various WBS elements that applied to learning curve theory analysis. Programs were analyzed using both Cumulative Average Theory and Unit Theory. Boone's Learning Curve was compared to the respective traditional theory using RMSE in units of total dollars and labor hours as well as in units of MAPE. This variety of attributes and tests allowed Boone's Learning Curve to be rigorously compared to the traditional learning curve theories. The researchers found that Boone's Learning Curve reduced error for RMSE and MAPE to a statistically significant degree although the amount of error reduction varied considerably. Further research should investigate the ability of Boone's Learning Curve to predict rather than explain program costs for applicability in the cost estimating community as well as which programs are better explained by Boone's Learning Curve.

IV. Phase 2

Methodology

Population and Sample

In order to further validate and understand Boone's Learning Curve, a subset of data from Phase 1 was utilized. The population remains all Department of Defense programs that have produced multiple complex end-items over time. The set of data sourced from Phase 1 consists of all program component learning curves analyzed using Unit Theory. This subset of data will be used to perform several statistical tests that would have been significantly compounded by the use of the two units of measures and two different learning curve theories. The justification for using this subset of data is Unit Theory is widely used and favored in the government (Mislick & Nussbaum, 2015, p. 215). Additionally, this dataset was the largest of the two learning curve theory datasets due to the additional assumptions of Cumulative Average Theory. This larger dataset provides more power for statistical tests (Cohen, 1938).

The dataset consists of several elements that were initially supplied with the data along with elements that were generated from the analysis in Phase 1. The elements that were initially provided with the dataset include each program component's observed lot costs, units produced per lot, and total quantified units among other attributes displayed in Table 9. In addition to these data, the predicted lot costs and empirically estimated learning curve parameter estimates for Boone's Learning Curve and Crawford's Learning Curve will also be utilized. These data were generated using the analysis techniques of Phase 1.

Table 9: Dataset Attributes and Values

Attribute	Values
Commodity	Aircraft, Aircraft subsystem, Engine, Helicopter, Missile, Unmanned Aerial Vehicle (UAV)
Component Estimated	Airframe, Alighting Gear, Auxiliary Power Plant, Avionics, Body, Data Link, Electrical, Electronic Warfare, Electro-Optical (EO), Electro-Optical Targeting System (EOTS), Electro-optical/Infrared (EO/IR), Empennage, Flyaway, Hydraulic, Mission Computer, Prime Mission Equipment, Radar, Surface Controls, Wing
Defense Contractor	26 defense contractors
First Year of Production	Various years spanning 1957 - 2018
Last Year of Production	Various years spanning 1957 - 2018
Number of Lots	Various numbers of lots spanning 1 to 21
Platform Type	Anti-submarine; Attack; Bomber; Command, Control, Information, Surveillance, & Reconnaissance (C2ISR); Combat Search & Rescue (CSAR); Electronic Warfare; Engine; Fighter; Missile; Multi-mission Maritime; Tanker; Trainer; Transport; UAV
Quantified Units	Various quantities spanning 11 to 10,035 units
Service	Air Force, Joint Air Force-Navy, Navy
Units of Measure	Dollars, Hours

In addition to this original dataset, 118 additional programs were added, each with a unique component. These additional programs comprise costs at the flyaway cost WBS level. The flyaway cost WBS level is the level directly above the prime mission equipment cost WBS level. The prime mission equipment cost WBS level was the highest level of analysis used in Phase 1 because traditional learning curve analysis cannot be directly applied to all constituent elements of the flyaway cost WBS level. However, learning curve analysis may be performed on program components at the flyaway cost WBS level due to data availability. Thus, the inclusion of program components at the flyaway cost WBS level will augment the research by providing a higher level of analysis above the prime mission equipment cost WBS level. This higher level of analysis will provide insight into how composite curves relate to diminishing rates of learning. These program component learning curves estimated at the flyaway cost WBS level will be included and excluded during different statistical tests to examine the sensitivity of results.

Data Collection

The dataset to include the 118 additional programs was extracted from the same source as Phase 1: CADE DACIMS and the Cost Research Library. The additional programs predicted lot costs and empirically estimated learning curve parameter estimates were generated from the analysis techniques in Phase 1.

The additional 118 program components increased the number of potential observations from 258 to 376. A business rule was implemented to exclude program components with lot numbers fewer than four. This business rule was necessary because programs lots will be divided into quarters; programs with less than four lots cannot be divided into quarters. This rule limited the number of observations to 185 when excluding program components estimated at the flyaway cost WBS level. The observations are limited to 261 when including program components estimated at the flyaway cost WBS level.

When performing statistical tests using Boone's Learning Curve, program components with lot numbers fewer than five will be excluded to prevent overfitting of Boone's Learning Curve and remain consistent with Phase 1. In these instances, the number of observations will be reduced to 169 when excluding program components estimated at the flyaway cost WBS level. The observations will be limited to 234 when including program components estimated at the flyaway cost WBS level.

Data Analysis

Analysis of Empirical Forgetting & Plateauing

This portion of the analysis will investigate if program components experience diminishing rates of learning towards the end of their production cycles using various methodologies. Boone's Learning Curve is predicated on the idea that observed learning curves experience a plateau towards the end of the production cycle when modeled using the traditional learning curves theories. In other words,

Boone's Learning Curve more accurately models observed learning curves in comparison to Crawford's Learning Curve due to the ability of Boone's Learning Curve to diminish its rate of learning as more units are produced. It is hypothesized that Crawford's Learning Curve fails to capture observed learning curves' diminishing rates of learning due to its assumption of constant rates of learning.

For example, a program from the dataset that displays diminishing rates of learning is illustrated in Figure 16. Crawford's Learning Curve imperfectly fits this observed learning curve as shown. Specifically, Crawford's Learning Curve underestimates the observed learning curve towards the end of production because this model is predicated on a constant rate of learning rate while the observed rate of learning diminishes. This constant rate of learning causes Crawford's Learning Curve predicted values to be lower than observed values in lots five and six. Although the difference in actual and predicted costs may seem insignificant in Figure 16, $2 million are underestimated in total that equates to 10% of the program's total costs across all six production lots. This example illustrates how small improvements in error can significantly improve a program's predicted costs and reduce cost overruns.

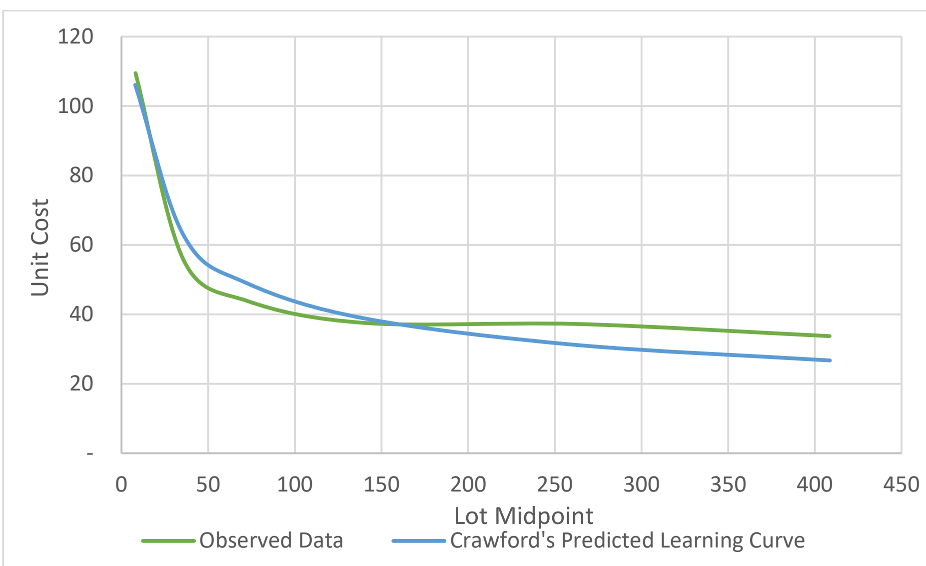


Figure 16: Example of a Program Displaying Plateauing Near the End of Production

In order to empirically investigate the claim that observed learning curves plateau in comparison to Crawford's Learning Curve, the observed unit cost per lot will be compared to Crawford's Learning Curve predicted unit cost per lot. Using these observed and predicted values per lot per program component, a percentage error will be calculated by subtracting the predicted value from the observed value and dividing by the observed value. These calculations will yield a series of percentage errors across all lots for each observed learning curve. Positive percentage errors will indicate that Crawford's Learning Curve underestimates, or is below, observed values. In contrast, negative percentage errors will indicate that Crawford's Learning Curve overestimates, or is above, observed values.

Next, each observed learning curve will be divided into quarters. This division will isolate the last quarter of each learning curve that is hypothesized to plateau. The methodology to divide an observed learning curve into quarters was employed because the number of lots for an observed learning curve varies. If instead a set number of lots were defined to capture the end of an observed learning curve production cycle, the various number of lots per observed learning curve would make comparisons incongruent. Therefore, the end of the production cycle is defined as the last quarter of the observed learning curve. The specific number of lots in the last quarter differs based on the observed learning curve's total number of lots. After these quarter divisions are accomplished, the mean percentage errors for all lots within each quarter will be calculated to yield a mean percentage error per quarter. Unlike the first and last quarters, the second and third quarters' lots will be combined into a single mean. This combination of the middle two quarters was necessary due to the inability to divide lots evenly for programs with an odd number of lots[4]. An example of this process is shown in Table 10 using an

[4] The total number of lots for each observation was divided by four, and the number of lots for the first and last quarter was calculated by rounding this quotient down. Therefore, the number of lots in the first and last quarter were equal, but the number of lots in the middle quarters may not be equal to the sum of the first and last quarter lots depending on the number of lots in the program.

observation from the dataset. In Table 10, Crawford's Learning Curve overestimates the actual data on average in the first and middle quarters while underestimating actual data in the fourth quarter.

Table 10: Example Calculation of Quarterly Mean Percentage Error

Lot Number	Percentage Error of Crawford's Learning Curve Prediction & Observed Data	Quarter	Mean Percentage Error
1	7.7%	1	-3.8%
2	-15.2%		
3	-17.2%	2	-4.9%
4	-18.4%		
5	7.8%	3	
6	8.1%		
7	9.2%	4	9.1%
8	9.0%		

Using these data, a t-test of the sample mean will be conducted to determine if using Crawford's Learning Curve systematically underestimates, or falls below, the observed learning curve in the last quarter. If Crawford's Learning Curve is systematically below the observed learning curve in the last quarter, this indicates that observed learning curves systematically plateau or experience diminishing rates of learning in comparison to Crawford's Learning Curve. If Crawford's Learning Curve fails to model observed learning curve plateaus accurately, Boone's Learning Curve has the potential to improve upon Crawford's Learning Curve for modeling plateaus.

These t-tests of the sample mean will be accomplished by comparing the mean percentage errors in the last quarter of all programs to zero. The null hypothesis for this test is Crawford's Learning Curve neither underestimates nor overestimates observed data in the last quarter. The alternative hypothesis is Crawford's Learning Curve systematically underestimates observed data in the last quarter. This alternative hypothesis would be supported if the global mean of the mean percentage errors in the last quarter were greater than zero to a statistically significant degree. Although there is a

focus on the plateauing phenomenon in the last quarter, the first and middle quarters will also be tested to gain additional insight on how Crawford's Learning Curve predictions compare to observed learning curves. A significance value of 0.05 will be used to determine if these results are statistically significant for this test and all following statistical tests. JMP Pro Version 13 will be used to perform this and all the following statistical analyses.

These same data will be converted to proportions to investigate further if plateauing occurs near the end of the production cycle. Specifically, the mean percentage error for each quarter will be converted to a dichotomous variable: if the mean percentage error is positive, the quarter will be labeled with a one and zero otherwise. For example, in Table 10 the first and middle quarters would be labeled with zeroes while the last quarter would be labeled with a one. Using these data, a hypothesis test of the sample proportions will be compared to a null hypothesis proportion of 0.5. The null hypothesis proportion of 0.5 assumes Crawford's Learning Curve predictions will randomly underestimate and overestimate observed learning curves. The alternative hypothesis is that Crawford's Learning Curve systematically underestimates observed data in the last quarter of programs. This alternative hypothesis would be indicated if the sample proportion was greater than 0.5 to a statistically significant degree. The first and middle quarters will also be assessed to gain additional insight.

Next, Boone's Learning Curve will also more accurately model program components costs that initially have high rates of learning that eventually diminish to low rates of learning at the end of the production cycle. If the observed data followed this pattern, Crawford's Learning Curve would underestimate observed data in both the first and last quarters of the program. Therefore, to test if observed data exhibit these characteristics, a dichotomous variable will be created and statistically tested. If Crawford's Learning Curve underestimates observed data in both the first and last quarters, the program component will be labeled with a one and zero otherwise. Investigating this proportion can

indicate how often Boone's Learning Curve can more accurately explain costs while further garnering an understanding of how Crawford's Learning Curve predictions compare to observed data.

Analysis of the Performance of Boone's Learning Curve

A further set of tests will be conducted to determine if Boone's Learning Curve improves upon Crawford's by more accurately modeling the plateauing phenomenon of observed learning curves. An alternative explanation is that Boone's Learning Curve improves upon Crawford's Learning Curve due to its additional empirically-estimated parameter. Because Boone's Learning Curve contains an additional parameter, it is expected to improve upon Crawford's Learning Curve at an aggregate level. However, if Boone's Learning Curve more accurately models observed learning curves that plateau and remains approximately equal to Crawford's Learning Curve in terms of error for observed learning curves that do not plateau, then Boone's Learning Curve provides inherent value in modeling the plateauing phenomenon. These tests will limit the number of observations to program components with lots greater than or equal to five to prevent overfitting of Boone's Learning Curve and remain consistent with Phase 1. Program component learning curves analyzed at the flyaway cost WBS level will be included in part of this analysis.

In order to test this hypothesis, a confusion matrix will be used first. A confusion matrix provides the number of observed learning curves that are predicted to be more accurately explained by Boone's Learning Curve and those that are not to the actual number of observed learning curves that are more accurately explained by Boone's Learning Curve and those that are not. An observed learning curve is predicted to be more accurately explained by Boone's Learning Curve if its learning curve plateaus in comparison to Crawford's Learning Curve. A learning curve plateaus in comparison to Crawford's Learning Curve if its last quarter mean percentage error is greater than zero as defined in the Analysis of Empirical Forgetting & Plateauing sub-section. The observed learning curves that were more accurately modeled by Boone's Learning Curve are defined as those with percentage error improvements

above 0.25% in comparison to Crawford’s Learning Curve. This cutoff is consistent with the groupings used in Phase 1 and was created from reviewing the data. Four categories exist within the confusion matrix to determine how the predicted counts and more accurately explained counts interact.

If an observed learning curve plateaus and is more accurately explained by Boone’s Learning Curve, this represents a true positive. If an observed learning curve does not plateau and was not more accurately explained by Boone’s Learning Curve, this represents a true negative. High counts in either true positive or true negative categories represent Boone’s Learning Curve modeling observed learning curves as hypothesized. The matrix also includes counts of observed learning curves that were more accurately explained by Boone’s Learning Curve but did not plateau. These counts are instances where Boone’s Learning Curve more accurately explained observed learning curves but for an unknown reason. A possible reason could be due to the additional parameter of Boone’s Learning Curve. These instances represent false positives. Lastly, the confusion matrix includes counts of observed learning curves that did plateau but were not more accurately explained by Boone’s Learning Curve. These counts represent false negatives. Low counts in either false positive or false negatives categories represent Boone’s Learning Curve modeling observed learning curves as hypothesized. An example of the confusion matrix format that will be utilized is shown in Table 11.

Table 11: Confusion Matrix Format

		Boone's Learning Curve More Accurately Explains	
		No	Yes
Observed Learning Curve Plateaus	No	True Negative	False Negative
	Yes	False Positive	True Positive

Another analysis technique will be used to investigate further where improvement in the learning curve occurs. Boone's Learning Curve error per lot will be compared to Crawford's Learning Curve error per lot to determine where in the production cycle Boone's Learning Curve improves upon Crawford's Learning Curve. If Boone's Learning Curve improves upon Crawford's Learning Curve in the last quarter only, then Boone's Learning Curve provides inherent value in explaining observed learning curves that plateau. However, if Boone's Learning Curve improves upon Crawford's Learning Curve across the entire production cycle, Boone's Learning Curve may fit better due to its additional, empirically-estimated parameter.

These comparisons will be accomplished using a similar methodology to that used in the Analysis of Empirical Forgetting & Plateauing sub-section. First, each theory's percentage errors per lot will be separately calculated by subtracting the predicted values from the observed values then dividing by the observed values. Next, Crawford's Learning Curve percentage error per lot will be compared to Boone's Learning Curve percentage error per lot for each observed learning curve. Using these percentage errors per lot by theory, a percentage error difference will be calculated by subtracting Boone's percentage error from Crawford's percentage error and dividing by Crawford's percentage error. These calculations will yield a series of percentage error differences across all lots for each observed learning curve. Positive percentage error differences will indicate that Boone's Learning Curve more accurately explains observed learning curves better than Crawford's Learning Curve. In contrast, negative percentage differences will indicate that Crawford's Learning Curve more accurately explains observed learning curves better than Boone's Learning Curve.

Next, each observed learning curve will be divided into quarters. This division will isolate each observed learning curve's last quarter that is hypothesized to plateau. This last quarter is the segment that Boone's Learning Curve is hypothesized to improve upon Crawford's Learning Curve significantly. The process of dividing the learning curves into quarters is consistent with that used in the Analysis of

Empirical Forgetting & Plateauing sub-section. After these divisions are completed, the mean percentage error differences for all lots within each quarter will be calculated to yield a mean percentage error difference per quarter. The second and third quarters' lots will be combined into a single average.

Statistical tests will then be used to determine where Boone's Learning Curve improves upon Crawford's Learning Curve to a statistically significant degree. A t-test of the sample mean will be conducted to determine if reductions in error from Boone's Learning Curve are different than zero to a statistically significant degree or due to random chance in the last quarter. The null hypothesis for this test is that the mean percentage error difference is equal to or less than zero in the fourth quarter. This null hypothesis represents that Boone's Learning Curve results in an equal amount of or more error in predicting observed values in comparison to Crawford's Learning Curve in the fourth quarter. The alternative hypothesis is that the mean percentage error difference is greater than zero. The alternative hypothesis represents that Boone's Learning Curve results in less error in predicting observed values than Crawford's Learning Curve in the last quarter. Each of these hypothesis tests will be conducted for the first, middle, and last quarters.

The dataset will be limited from the complete dataset and will consist of observed learning curves that were more accurately explained by Boone's Learning Curve to a significant degree on the aggregate level. Improvement to a significant degree is defined as Boone's Learning Curve resulting in a percentage error improvement above 0.25% in comparison to Crawford's Learning Curve as previously specified. The exclusion of observations that are not more accurately explained by Boone's Learning Curve is necessary due to the nature of the research question that asks how Boone's Learning Curve improves upon Crawford's Learning Curve. If Boone's Learning Curve did not improve upon Crawford's Learning Curve for an observation, it would be inappropriate to include this observations in the analysis.

The results of all three hypothesis tests per quarter will be used to perform a more aggregated hypothesis test. The null hypothesis is Boone's Learning Curve more accurately explains observed learning curves in comparison to Crawford's Learning Curve across all quarters. This null hypothesis represents Boone's Learning Curve more accurately explaining observed learning curves due to its additional parameter. The alternative hypothesis is Boone's Learning Curve more accurately explains observed learning curves in comparison to Crawford's Learning Curve in the fourth quarter only. This alternative hypothesis represents Boone's Learning Curve more accurately explaining the plateau effect in observed learning curves. The alternative hypothesis would indicate that Boone's Learning Curve provides inherent value in more accurately explaining the plateauing phenomenon in observed learning curves. Other alternative hypotheses are also possible if another section of the learning curve other than the last section or two of the three sections of the learning curve show improvement to a statistically significant degree; however, theoretical explanations are limited in these instances.

Regression Analysis of Forgetting & Plateauing

In the next part of this research, an OLS regression analysis will be used to determine what program characteristics, if any, can be used to determine the degree to which Boone's Learning Curve will more accurately explain the observed learning curve. Also, OLS regression analysis will be used to determine what program attributes, if any, affect the degree to which plateauing occurs as modeled by the Crawford's Learning Curve.

The percentage difference between Boone's Learning Curve MAPE and Crawford's Learning Curve MAPE, *Boone/Crawford MAPE Percentage Difference*, will be the first dependent variable for this OLS regression analysis. This dependent variable will address the seventh research question about explaining instances when Boone's Learning Curve will more accurately explain the observed learning curve. This dependent variable represents the amount of error that Boone's Learning Curve more accurately explained in comparison to Crawford's Learning Curve. Higher magnitudes of this measure

represent more accurate explanations of the observed learning curve by Boone's Learning Curve. By investigating the independent variables that explain this dependent variable, program characteristics will be explored that are related to Boone's Learning Curve more accurately explaining observed learning curves. The values of *Boone/Crawford MAPE Percentage Difference* comprise the same values as those calculated and examined in Phase 1.

The mean percentage error between Crawford's Learning Curve and observed data in the fourth quarter of each program, *Crawford MPE4Q*, will serve as the second dependent variable to investigate the eighth research question pertaining to investigating attributes that affect the degree to which plateauing occurs. The mean percentage error differences that comprise *Crawford MPE4Q* is the same as that calculated and examined in the Analysis of Empirical Forgetting & Plateauing and Analysis of the Performance of Boone's Learning Curve sub-sections. Greater values of this dependent variable indicate a plateauing of the learning curve in comparison to the traditional theory. This dependent variable provides a measure of plateauing indifferent to how Boone's Learning Curve models observed data. Several independent variables along with their hypothesized signs will be specified using the Literature Review.

References in the Literature Review discussed several program characteristics that can lead to or are responsible for plateauing of the learning curve as more units are produced. The characteristics that have been hypothesized to be positivity related to plateauing of the learning curve include 1) a higher proportion of capital to labor in the manufacturing process (Baloff, 1966, 1970; Department of the Air Force, 2007, pp. 8-68 – 8-69), 2) units of measure in total dollars rather than hours (Asher, 1956, pp. 97-98; Department of the Air Force, 2007, p. 8-63; Wright, 1936), 3) a higher level of aggregation of costs (Asher, 1956, pp. 97-98; Department of the Air Force, 2007, p. 8-63; Wright, 1936), and 4) the general presence of forgetting. The first hypothesized characteristic cannot be operationalized due to a lack of data. This lack of data is a commonly recognized problem when investigating this characteristic, as

discussed by Jaber (2006) and Dar-El (2000, p. 38). The latter three hypothesized characteristics can be operationalized directly or decomposed further.

The second hypothesized characteristic can be operationalized directly because the original data provided units of measure in either total dollars or labor hours. This variable will be annotated as *UnitsofMeasure*. A small number of programs were analyzed in units of hours that may complicate analysis. Additionally, other variables were not included with programs that were analyzed in units of hours to include the first and last years of production. This characteristic of the data limits the number of independent variables that can be regressed in the model simultaneously. The researchers have hypothesized and found evidence that learning curves in the form of total dollars will lead to greater plateauing of the learning curve because these learning curves are composed of constituent learning curves that likely vary in their rates of learning (Asher, 1956, pp. 97-98; Department of the Air Force, 2007, p. 8-63; Wright, 1936).

The third hypothesized characteristic was operationalized based on how aggregated each program component was in relation to other components. Four levels of aggregation were created based on a review of the data. The four levels of aggregation and their respective components were: 1) flyaway cost at the highest level of aggregation, 2) prime mission equipment at a lower level of aggregation, 3) major subcomponents such as the airframe, engine, body, and wing at another lower level of aggregation, and 4) subcomponents and subsystems such as alighting gear, avionics, and radars at the lowest level of aggregation. The list numbers for each level of aggregation correspond to the value placed in the *LevelofAggregation* independent variable for each observation. Converting the levels of aggregation from qualitative descriptions to categorical values that represent these descriptions is necessary for OLS regression analysis. The levels of aggregation are related to each component's WBS level; however, strict WBS level categorization was avoided due to the inability to match some WBS levels across programs along with a lack of data. Sources in the Literature Review hypothesize that as

the levels of aggregation increase from the least aggregated levels to the most aggregated levels, plateauing of the learning curve will be more prevalent (Asher, 1956, pp. 97-98; Department of the Air Force, 2007, p. 8-63; Wright, 1936). This independent variable was annotated as *LevelofAggregation*. A Tukey Test will determine if these groupings are different to a statistically significant degree. A set of dichotomous variables will be created using the output of this test to be utilized later in OLS regressions.

The fourth hypothesized characteristic, the presence of forgetting, will be decomposed in order to be operationalized. The Literature Review provided insight into which program characteristics contribute to forgetting (Argote et al., 1990, Badiru, 2012; Jaber, 2006). The most prominent causes of forgetting include significant breaks in production and design changes (Anderlohr, 1969; Argote et al., 1990). The data were reviewed to ensure programs with considerable perturbations to a stable production line or end-item did not enter the dataset. Learning curves with breaks in production would be inappropriate to estimate using traditional learning curve analysis; however, extensive records did not exist for all programs, and programs were not removed from the analysis without cause. Therefore, some programs may be included in the analysis that have breaks in production or design changes because records were unavailable to identify all these programs, so this aspect of forgetting cannot be operationalized with the present data.

The Literature Review also cited that as more units are produced or more time passes during production, forgetting is more likely to occur that may result in a plateauing of the learning curve (Badiru, 2012; Jaber 2006). The original data included variables to operationalize these concepts. The independent variable *UnitsProduced* will represent the number of units produced in a program components production cycle. The timespan of production was calculated by subtracting the last year of production by the first year of production to create the independent variable *TimeSpanned*.

An independent variable was also created from the original data to measure the compression of the production environment. As discussed in the Literature Review, laborers and the organization as a

whole must be provided the opportunity to process efficiencies gained from learning before making improvements (Badiru, 2012; Corlett & Morcombe, 1970; Jaber, 2006; Yelle, 1980). If units are produced too quickly, neither the laborers nor the organization can learn and implement processes to gain efficiencies; however, if units are produced too slowly, forgetting along the learning curve may occur. To capture this rate of production and its effect on forgetting, an independent variable, *UnitsProducedperLot*, was created by dividing *UnitsProduced* by the number of lots per program. The number of lots per program was used for this measure instead of dividing *UnitsProduced* by *TimeSpanned* due to the fewer number of observations that include the independent variable *TimeSpanned*.

Lastly, the number of units produced per lot and its effect on forgetting is likely related to the complexity of the end-item produced. For example, a program that produces 1,000 air vehicles in one lot is likely to experience forgetting differently than a program that produced 1,000 hydraulic components in one lot. The complete air vehicle is more complex than the hydraulic component and several hydraulic components comprise the air vehicle. A previously created variable will be utilized as a proxy variable for complexity, *LevelofAggregation*; therefore, a variable to measure the interaction of *UnitsProducedperLot* and *LevelofAggregation* will also be created. This interaction variable will be annotated as *UnitsProducedperLot***LevelofAggregation*.

All hypothesized independent variables will enter the OLS regression model simultaneously except for *UnitsofMeasure* and *TimeSpanned*. Some independent variables may not be statistically significant in the OLS regression model to a preferred degree; therefore, a methodology will be employed to remove independent variables that lack statistical significance. All independent variable coefficient estimate p-values must be below a significance level of 0.05. If any coefficient estimate p-values are above the significance level of 0.05, the independent variable with the greatest p-value will be singly removed from the analysis, and the regression will be reperformed until all coefficient estimate

p-values are below 0.05. Once all coefficient estimate p-values are below the significance level of 0.05, a comparison-wise error rate (α_C) will be calculated to ensure that the experiment-wise error rate (α_E) is below the significance level of 0.05. The comparison-wise error rate will be calculated using the Bonferroni correction by dividing the experiment-wise error rate of 0.05 by the number of independent variables in the OLS regression model (McClave et al., 2014). If coefficient estimate p-values remain above this comparison-wise error rate, the independent variable with the greatest p-value will be singly removed from the analysis, and the regression will be reperformed until all coefficient estimate p-values are below the comparison-wise error rate.

The models that will be created with the previously discussed independent variables rooted in the Literature Review use a theory-based approach to OLS regression modeling. Another methodology, a data-mining approach, will also be employed in order to utilize additional program attributes. These attributes may not have been cited as relevant to plateauing and forgetting phenomena in the Literature Review; however, they may control for additional variation in either dependent variable. By controlling for additional variability, the previously discussed theory-based independent variables may exhibit differing levels of significance or coefficient estimates. Therefore, the program attributes such as those discussed in Table 9 will be utilized in a separate OLS regression analysis. These additional independent variables include the categorical independent variables *Commodity*, *ComponentEstimated*, *DefenseContractor*, *FirstYearofProduction*, *PlatformType*, and *Service*.

The additional data-mining independent variables in addition to the theory-based independent variables will result in a total of 12 possible independent variables for the OLS regression model. In order to thoroughly test if these potential independent variables are significant for both dependent variables, a mixed stepwise regression analysis will be conducted using JMP Pro Version 13. The mixed stepwise regression will test the statistical significance of various combinations of independent variables and various groups of levels within each categorical variable for each dependent variable. A significance level

of 0.05 will be used initially as a threshold for independent variables to enter or exit the OLS regression model. After the stepwise regression calculates the statistically significant independent variables, a further test of the comparison-wise error rate will be used to ensure the experiment-wise error rate is below 0.05.

Most of the additional data-mining independent variables are categorical variables with a multitude of levels per variable. For example, *DefenseContractor* has 26 different levels. In order to utilize these categorical variables for OLS regression analysis, the stepwise function will test all combinations of levels for each independent categorical variable for their statistical significance to the overall OLS regression model. Each partitioned group will be represented using a dichotomous (dummy) variable. This dummy variable will be set to one if an observation falls into the partitioned group and zero otherwise. This mixed stepwise approach will allow the statistical significance of all possible groups of data-mining categorical variables to be tested alongside the theory-based independent variables within the OLS regression model.

Once these final models are created, statistical tests will be used to determine the validity of the regression model results. First, the assumption of normality will be assessed using the Shapiro-Wilk test along with a plot of the studentized residuals. These tests will ensure that the statistical inferences generated from the OLS regression are valid (Wooldridge, 2016, p. 149). The assumption of constant variance of the residuals (homoskedasticity) will also be tested using a Breusch-Pagan test and residual by predicted plot (Hilmer & Hilmer, pp. 262-265). These tests will ensure that the OLS regression model has minimum variance and estimated standard errors along with all measures of precision are correct (Hilmer & Hilmer, p. 258). All statistical tests will be conducted at a 0.05 significance level. Multicollinearity will also be assessed using the variance inflation factor (VIF) scores for multiple linear regression models (Gujarati & Porter, 2010, pp. 256-258). VIF scores above 10 indicate a strong linear relationship between independent variables. This linear relationship leads to multicollinearity that

results in coefficient estimate standard errors that are artificially large (Gujarati & Porter, 2010, pp. 256-258). Artificially large coefficient estimate standard errors result in coefficient estimates with wider confidence intervals, lower test statistics, and greater p-values (Gujarati & Porter, 2010, pp. 256-258). Influential data points will be investigated using Cook's Distance values. Influential data points with Cook's Distance values above 0.5 may skew the results of the OLS regression model. Lastly, studentized residuals will be reviewed to investigate if outliers are present that may bias results (Hilmer & Hilmer, p. 238). An outlier will be defined as any value falling three interquartile ranges below the 10% quantile and above the 90% quantile. Observations with suspect Cook's Distance values and studentized residual outliers will be investigated. These overall model tests will be conducted using JMP Pro Version 13 and RStudio.

In summary, several OLS regressions will be used to determine what program attributes, if any, are associated with better modeling from the use of Boone's Learning Curve as well as determine the degree to which plateauing is present in comparison to Crawford's Learning Curve. To investigate the seventh research question, *Boone/Crawford MAPE Percentage Difference* will be used as the dependent variable while the theory-based independent variables *UnitsofMeasure*, *LevelofAggregation*, *UnitsProduced*, *TimeSpanned*, *UnitsProducedperLot*, and *UnitsProducedperLot***LevelofAggregation* will be used. Another dependent variable, *Crawford MPE4Q*, will also be utilized to investigate the eighth research question. Next, additional data-mining independent variables will be tested in addition to these theory-based independent variables for each dependent variable. The additional data-mining independent variables *Commodity*, *ComponentEstimated*, *DefenseContractor*, *FirstYearofProduction*, *PlatformType*, and *Service* will be utilized with the theory-based independent variables using mixed stepwise OLS regression analysis.

Analysis & Results

Analysis of Empirical Forgetting & Plateauing

The t-tests of the sample mean for the mean percentage error differences in the first, middle, and last quarters of the program are shown in Table 12. The JMP Pro Version 13 output for these hypothesis tests are included in Appendix D. The tested dataset included all program components with four lots or more and initially excluded program components estimated at the flyaway cost WBS level. For these tests, if the mean percentage error differences in the first, middle, or last quarters of the program are greater than zero, Crawford's Learning Curve underestimates the observed learning curve. In contrast, if the mean percentage error differences in the first, middle, or last quarters of the program are less than zero, Crawford's Learning Curve overestimates the observed data. If Crawford's Learning Curve underestimates the observed learning curve in the last quarter, this indicates that observed learning curves systematically plateau towards the end of production in comparison to Crawford's Learning Curve. This systematic plateauing would provide support for the use of Boone's Learning Curve.

Table 12: Combined T-Tests of Sample Mean for Mean Percentage Error

Hypothesis Test: H_0: $\mu \leq 0$ H_A: $\mu > 0$							
Learning Curve Section	**Sample Mean ($\bar{x}$)**	**Standard Deviation (s)**	**Number of Observations**	**Test Statistic**	**P-Value**	**Result**	**Crawford's Learning Curve Systematically Underestimates**
First Quarter	-3.4%	25.7%	185	-1.808	0.964	Fail to Reject H_0	No
Middle Quarters	-13.0%	102.7%	185	-1.727	0.957	Fail to Reject H_0	No
Last Quarter	-26.1%	193.4%	185	-1.837	0.966	Fail to Reject H_0	No

Each of the t-tests of the sample mean fails to reject the null hypothesis that Crawford's Learning Curve systematically underestimates data in all quarters. These rejections of the null hypothesis are indicated by p-values that are greater than a significance level of 0.05. Each of these hypothesis tests indicates the opposite of the alternative hypothesis: Crawford's Learning Curve

systematically overestimates observed data in all quarters. This conclusion is indicated by the complement of the p-values ($1 -$ p-value) that are all less than a significance level of 0.05.

Several outliers were present in each of the t-tests of the sample mean. Although these outliers were valid for inclusion in the analysis, they were removed to determine their effects, if any, on the results. An observation was identified as an outlier if the observation's value fell more than three interquartile ranges from the upper 90% and lower 10% quantiles. When the six outliers were removed, the results of the t-tests of the sample mean did not change: Crawford's Learning Curve systematically overestimates data in all quarters to a statistically significant degree.

To further test the sensitivity of the results, program components estimated at the flyaway cost WBS level were added to the original dataset, and the t-tests of the sample mean were reevaluated. The results for each test remained unchanged. This expanded dataset also included outliers identified using the same outlier identification methodology. When the nine outliers were excluded, the t-test of the sample mean results in the first and middle quarters remained unchanged. However, the t-test of the sample mean in the fourth quarter became not statistically significant from zero with a test statistic of -1.2926 and a p-value of 0.0987. This latter test indicates that Crawford's Learning Curve mean percentage error is not different from zero to a statistically significant degree. In other words, Crawford's Learning Curve does not systematically overestimate or underestimate data in the fourth quarter of a program component's production when a higher aggregated program component is included. These nine additional t-tests of the sample mean are also included in Appendix D.

The addition of program components estimated at the flyaway cost WBS level did not significantly alter the results of various t-tests of the sample mean except in one instance. This result indicates that the influence of program components estimated at higher levels of aggregation may not have as pronounced of an effect on plateauing of the learning curve than initially hypothesized. Although these results indicate that Crawford's Learning Curve overestimates data in all quarters on

average, the variability for each measure remains high with coefficients of variation above 100%. Much like Boone's Learning Curve, these high amounts of variability make it challenging to conclude the degree to which Crawford's Learning Curve will overestimate or underestimate program component costs. These results highlight the importance of investigating if program characteristics exist to determine if a program component's learning curve is likely to plateau. However, these tests indicate that Crawford's Learning Curve does not systematically underestimate data in the last quarter to a statistically significant degree.

Next, these mean percentage error of Crawford's Learning Curve in the fourth quarter were analyzed using dichotomous values. The combined sample proportion hypothesis tests for the first, middle, and last quarters of the program are shown in Table 13. The JMP Pro Version 13 output for these hypothesis tests are included in Appendix E. The tested dataset included all program components with four lots or more and initially excluded program components estimated at the flyaway cost WBS level. For these tests, if the mean percentage error differences in the first, middle, or last quarters of the program are greater than zero, Crawford's Learning Curve underestimates the observed data. In these instances, the observation was coded with a one. If the mean percentage error differences in the first, middle, or last quarters of the program were less than zero, Crawford's Learning Curve overestimated the observed data. In these instances, the observation was code with a zero. Each of these hypothesis tests for the different sections of the learning curve compares the sample proportion of "success" or observations coded with a one to that of the null hypothesis of 0.5.

Table 13: Combined Sample Proportion Hypothesis Tests

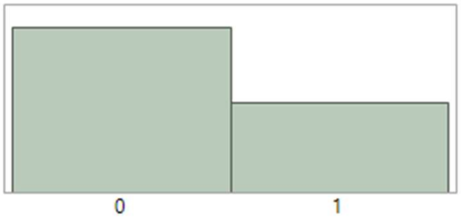
Hypothesis Test: H_0: $p \leq 0.5$ H_A: $p > 0.5$							
Learning Curve Section	**Sample proportion ($\hat{p}$)**	**Standard Deviation (σ_{p_0})**	**Number of Observations**	**Test Statistic**	**P-Value**	**Result**	**Crawford's Learning Curve Systematically Underestimates**
First Quarter	0.503	0.037	185	0.073	0.471	Fail to Reject H_0	No
Middle Quarters	0.378	0.037	185	-3.308	0.999	Fail to Reject H_0	No
Last Quarter	0.476	0.037	185	-0.662	0.745	Fail to Reject H_0	No

The sample proportion hypothesis tests failed to reject the null hypothesis that the proportion of programs that Crawford's Learning Curve underestimates is greater than 0.5 in any quarter. These rejections of the null hypothesis are indicated by p-values that are greater than a significance level of 0.05. These results confirm the results for the t-tests of the sample mean for the mean percentage error differences. For the middle quarters' sample proportion hypothesis test, the test indicates the opposite of the alternative hypothesis. This test indicates that the proportion of programs that Crawford's Learning Curve overestimates is different from random chance to a statistically significant degree when isolated to the middle quarters. This conclusion is indicated by the p-value's complement (1 - 0.9994 = 0.0003) that is less than a significance level of 0.05. To further test the sensitivity of the results, program components estimated at the flyaway cost WBS level were added to the original dataset and sample proportion hypothesis tests were reevaluated. The results for each test remained unchanged, and all three additional sample proportion hypothesis tests are also included in Appendix E. Despite these results, Crawford's Learning Curve underestimates observed learning curves in approximately half of all observed learning curves tested. This proportion emphasizes the opportunity for Boone's Learning Curve to improve upon Crawford's Learning Curve.

Lastly, the sample proportion of observed learning curves that experience high rates of learning in the first quarter that eventually diminish to low rates of learning in the last quarter are shown in Figure 17. These phenomena are indicated when Crawford's Learning Curve underestimates the observed learning curve in the first and last quarter. At first glance, the proportion of observed learning

curves that experience high rates of learning and the proportion of observed learning curves that diminish to low rates of learning can be treated as two independent events. The proportion of both these independent events coinciding can be calculated using the multiplicative law of probability; therefore, the multiplicative law of probability would calculate a null proportion of 0.25 because each independent event has a probability of 0.5 (McClave et al., pp. 158-160).

However, this calculation assumes that the two events are independent. OLS regression and GRG Nonlinear Solver were used to calculate Crawford's Learning Curve along with its errors or residuals. In OLS regression, the sum of residuals is equal to zero (Hilmer & Hilmer, p. 80). This calculation causes the two events to be dependent rather than independent. Because independent events are an assumption of the multiplicative law of probability, this law cannot be used to determine a null hypothesis and perform a formal hypothesis test (McClave et al., pp. 158-160). Therefore, we are unable to calculate the expected proportion of programs that are underestimated in the first and last quarter due features of the tool, OLS regression, used to model learning curves. However, benefit exists in understanding how the dependence of these two events results in proportions that are different than what could be expected from independent events. In other words, an instance of a high rate of learning in the first quarter may positively or negatively affect the probability of a diminishing rate of learning in the last quarter. If an instance of a high rate of learning in the first quarter increases the probability of a diminishing rate of learning in the last quarter, Crawford's Learning Curve would systematically underestimate the observed learning curve in both quarters. This systematic underestimation would provide support for Boone's Learning Curve.



Frequencies

Level	Count	Prob
0	120	0.64865
1	65	0.35135
Total	185	1.00000

N Missing 0
2 Levels

Test Probabilities

Level	Estim Prob	Hypoth Prob
0	0.64865	0.75000
1	0.35135	0.25000

Binomial Test	Level Tested	Hypoth Prob (p1)	p-Value
Ha: Prob(p > p1)	1	0.25000	0.0014*

Figure 17: Proportion of Programs with High Rates of Learning then Diminishing Rates of Learning

The sample proportion hypothesis test indicated that approximately 35% of programs experience instances of high rates of learning in the first quarter that eventually diminishes to low rates of learning in the last quarter. The dependence of these two events coinciding is positive because this proportion is greater than a proportion of 0.25 that would apply to independent events. This sample proportion is also greater than 0.25 to a statistically significant degree because the p-value of the hypothesis test is less than a significance level of 0.05. These results indicate that Crawford's Learning Curve systematically underestimates the observed learning curve in the first and last quarters more than can be expected from independent events. Although this systematic underestimation in both quarters may solely be due to the tool used to model learning curves, this inference along with the relatively high proportion of observed learning curves that are underestimated in both quarters highlights opportunities for Boone's Learning Curve to improve upon Crawford's Learning Curve.

Analysis of the Performance of Boone's Learning Curve

The confusion matrix shown in Table 14 investigates if Boone's Learning Curve improves upon Crawford's by more accurately modeling the plateauing phenomenon of observed learning curves. The confusion matrix can investigate if instances of plateauing are associated with instances of improvement in Boone's Learning Curve using counts of when an observed learning curve plateaus and counts of when Boone's Learning Curve significantly improves the observed learning curve. Table 14 excludes program components estimated at the flyaway cost WBS level. An additional confusion matrix including

program components estimated at the flyaway cost WBS level is displayed in Appendix F although the results are similar.

Table 14: Confusion Matrix of Counts & Percentages excluding Flyaway Components

		Boone's Learning Curve More Accurately Explains: No	Boone's Learning Curve More Accurately Explains: Yes
Observed Learning Curve Plateaus	No	77 46%	14 8%
	Yes	25 15%	53 31%

This table indicates that Boone’s Learning Curve more accurately explains observed learning curves when the observed learning curve plateaus for 31% observed learning curves. Additionally, Boone’s Learning Curve does not more accurately explain observed learning curves when the observed learning curve does not plateau for 46% observed learning curves. Therefore, Boone’s Learning Curve more accurately modeled observed learning curves based on if those observed learning curves plateaued for 77% of observations.

In contrast, Boone’s Learning Curve fails to more accurately explain observed learning curves when the observed learning curve plateaus for 15% observed learning curves. These instances may represent Crawford’s Learning Curve explaining the observed learning curves to a significant degree of accuracy, and Boone’s Learning Curve not improving upon Crawford’s Learning Curve to a significant enough degree. Lastly, Boone’s Learning Curve more accurately explained observed learning curves when the observed learning curve did not plateau for 8% observed learning curves. These instances represent when Boone’s Learning Curve more accurately explained the observed learning curve for an unknown reason. Most likely, Boone’s Learning Curve more accurately explained the observed learning

curve due to its additional empirically-estimated parameter. Therefore, Boone's Learning Curve did not perform as hypothesized for 23% of observed learning curves.

These results indicate that Boone's Learning Curve improves upon observed learning curves based in part on if the observed learning curve plateaus as predicted. However, approximately a quarter of the observations indicate Boone's Learning Curve did not perform as predicted by failing to improve upon the observed learning curves that plateau or more accurately explaining observed learning curves when the observed learning curve does not plateau. These results provide mixed conclusions as to if Boone's Learning Curve improves upon Crawford's Learning Curve by more accurately modeling the plateauing phenomenon of observed learning curves. To further investigate, a series of hypothesis tests will be conducted to determine where in the learning curve Boone's Learning Curve improves upon Crawford's Learning Curve.

The following hypothesis tests indicate wherein the program Boone's Learning Curve significantly improves upon Crawford's Learning Curve. The hypothesis tests of the mean percentage error differences between Boone's and Crawford's Learning Curves for the first, middle, and last quarters of the program are shown in Table 15. The JMP Pro Version 13 output for these hypothesis tests are included in Appendix G. For these tests, if the mean percentage error difference is greater than zero, Boone's Learning Curve more accurately explains the observed learning curve in comparison to Crawford's Learning Curve. If the mean percentage error difference is less than zero, Boone's Learning Curve fails to more accurately explain the observed learning curve in comparison to Crawford's Learning Curve.

Table 15: Combined Percentage Error Difference Hypothesis Tests between Boone's & Crawford's Learning Curves Given Boone's Learning Curve Significantly Improved

Hypothesis Test: H_0: $\mu \leq 0$ H_A: $\mu > 0$							
Learning Curve Section	**Sample Mean ($\bar{x}$)**	**Standard Deviation (s)**	**Number of Observations**	**Test Statistic**	**P-Value**	**Result**	**Boone's Learning Curve Systematically Improves**
First Quarter	-20.4%	367.4%	169	-0.721	0.764	Fail to Reject H_0	No
Middle Quarters	-22.5%	125.9%	169	-2.320	0.989	Fail to Reject H_0	No
Last Quarter	-40.1%	295.5%	169	-1.762	0.960	Fail to Reject H_0	No

Each of the t-tests of the sample mean fails to reject the null hypothesis that Boone's Learning Curve improves upon Crawford's Learning Curve in any quarter. This is in stark contrast to the results included in Phase 1. These rejections of the null hypothesis are indicated by p-values that are greater than a significance level of 0.05. The hypothesis tests for the middle quarters and last quarter indicate the opposite of the alternative hypothesis: Boone's Learning Curve performs systematically worse than Crawford's Learning Curve in the middle and last quarters to a statistically significant degree. This conclusion is indicated by the complement of the p-values (1 − p-value) for these quarters that are less than a significance level of 0.05.

However, these contrasting results may be caused by the presence of outliers in the hypothesis tests. Outliers were removed to determine their effects on the results. As previously defined, an observation was identified as an outlier if the observation's value fell more than three interquartile ranges from the upper 90% and lower 10% quantiles. When the 11 outliers were removed, the results of the t-tests of the sample mean changed significantly for the three tests. These hypothesis tests of the mean percentage error differences between Boone's and Crawford's Learning Curves for the first, middle, and last quarters of the program excluding outliers are shown in Table 16. The JMP Pro Version 13 output for these hypothesis tests is also included in Appendix G.

Table 16: Combined Percentage Error Difference Hypothesis Tests between Boone's & Crawford's Learning Curves Given Boone's Learning Curve Significantly Improved, Excluding Outliers

Hypothesis Test: H_0: $\mu \leq 0$ H_A: $\mu > 0$							
Learning Curve Section	**Sample Mean ($\bar{x}$)**	**Standard Deviation (s)**	**Number of Observations**	**Test Statistic**	**P-Value**	**Result**	**Boone's Learning Curve Systematically Improves**
First Quarter	13.4%	49.3%	158	3.425	<0.001	Reject H_0	Yes
Middle Quarters	-10.1%	54.1%	158	-2.347	0.990	Fail to Reject H_0	No
Last Quarter	3.6%	58.2%	158	0.778	0.219	Fail to Reject H_0	No

The t-test of the sample mean in the first quarter rejects the null hypothesis that Boone's Learning Curve does not improve upon Crawford's Learning Curve. This result indicates that Boone's Learning Curve systematically improves upon Crawford's Learning Curve in the first quarter to a statistically significant degree. This rejection of the null hypothesis is indicated by p-values that are less than a significance level of 0.05. In contrast, the t-tests of the sample mean fails to reject the null hypothesis that Boone's Learning Curve improves upon Crawford's Learning Curve in the middle quarters and last quarter. These failures to reject the null hypothesis are indicated by p-values that are greater than a significance level of 0.05. The middle quarters' p-value results in rejecting the opposite of the alternative hypothesis because the complement of this p-value $(1 - 0.9899 = 0.0101)$ is less than a significance level of 0.05. This result indicates that Boone's Learning Curve is systematically worse at explaining observed data in comparison to Crawford's Learning Curve in the middle quarters to a statistically significant degree.

Using these results, an aggregated hypothesis test will be conducted. For this aggregated hypothesis test the null hypothesis is that Boone's Learning Curve more accurately explains observed learning curves in comparison to Crawford's Learning Curve across all quarters due to its additional empirically-estimated parameter. The main alternative hypothesis is Boone's Learning Curve more accurately explains observed learning curves in comparison to Crawford's Learning Curve in the fourth quarter only. These latter hypothesis test results indicate that Boone's Learning Curve does not improve upon Crawford's Learning Curve to a statistically significant degree in all quarters; therefore, this null

hypothesis is rejected. However, the main alternative hypothesis is also rejected because Boone's Learning Curve does not improve upon Crawford's Learning Curve in the last quarter to a statistically significant degree. Surprisingly, Boone's Learning Curve more accurately explains observed learning curves in comparison to Crawford's Learning Curve only in the first quarter to a statistically significant degree.

These results suggest that Boone's Learning Curve improves upon Crawford's Learning Curve more than it simply having an additional empirically-estimated parameter. However, Boone's Learning Curve was not shown to model the plateau effect in the last quarter to a statistically significant degree in contrast to the main alternative hypothesis. Therefore, an alternative hypothesis to explain these results is Boone's Learning Curve may more accurately explain the first quarter of a learning curve to a statistically significant degree due to its ability to model observed data with high rates of learning at the beginning of the learning curve. Further, a competing alternative hypothesis is that observed learning curves were improved in different sections of the learning curves while performing worse in others in comparison to Crawford's Learning Curve. These error improvement tradeoffs resulted in significant improvement in the use of Boone's Learning Curve on the aggregate; however, these improvements in different sections of the learning curve would have caused large variations of error improvement in each section. These large variations of error improvement within each section could significantly obscure the statistical results due to the high standard deviations and coefficients of variation included in each hypothesis test. Further, these high standard deviations and coefficients of variation are present despite limiting the dataset that further strengthens this alternative hypothesis.

Regression Analysis of Forgetting & Plateauing

Before performing OLS regression analysis, additional independent variables were created. First, the categorical independent variable *LevelofAggregation* was converted to a series of dummy variables. For example, a dummy variable was created to represent when *LevelofAggregation* was equal to the

highest level: flyaway costs. When *LevelofAggregation* is equal to the flyaway cost level of aggregation, this dummy variable was coded one and zero otherwise. This dummy variable was named *LevelofAggregation = 1*. This process was repeated for the second and third levels of aggregation. The fourth level of aggregation was not converted to a dummy variable to serve as the omitted or base group. Next, another dummy variable was created to convert the categorical variable *UnitsofMeasure* to dichotomous values. For example, if *UnitsofMeasure* is equal to Total Dollars, this dummy variable, *UnitsofMeasure = Dollars*, was coded one and zero otherwise. The independent variable *UnitsProducedperLot***LevelofAggregation* was also created by multiplying the individual independent variables to create an interaction term.

Next, a statistical test was used to determine if the means of the levels of aggregation within the categorical independent variable *LevelofAggregation* are different from one another to a statistically significant degree. Tukey's HSD Test was used to conduct this analysis. The results of this test foretold if the *LevelofAggregation* dummy variables will be statistically significant in the OLS regression model. Additionally, Tukey's HSD Test was used to indicate if two or more levels of aggregation should be grouped within one dummy variable. The main dependent variable, *Boone/Crawford MAPE Percentage Difference*, served as the response variable for this test.

The results of Tukey's HSD Test are shown in Table 17. The mean of level one was different from the mean of level four to a statistically significant degree. This result is indicated by the lack of a letter connecting the two levels of aggregation and the non-overlapping 95% confidence intervals. However, the means of all other levels were not different from one another to a statistically significant degree. These conclusions indicate that the OLS regression results may not be statistically significant for some of these dummy variables. Despite these results, the means of the highest level of aggregation and the lowest level of aggregation decrease and are different to a statistically significant degree. These results also indicate that Boone's Learning Curve more accurately models highly aggregated learning curves in

comparison to Crawford's Learning Curve as hypothesized. Each of the previously created *LevelofAggregation* dummy variables still entered the OLS regression model simultaneously with other independent variables as previously planned. This methodology allowed for other independent variables to control for effects within the *LevelofAggregation* dummy variables that may result in at least one of the dummy variables being statistically significant.

Table 17: Tukey's HSD Test of *LevelofAggregation*

Level			Least Sq Mean	Lower 95% CI	Upper 95% CI
1	A		0.1892	0.1315	0.2468
2	A	B	0.1432	0.0931	0.1933
3	A	B	0.1109	0.0214	0.2003
4		B	0.0679	0.0058	0.1300

Levels not connected by same letter are significantly different.

The results of the first set of OLS regressions are shown in Table 18. Regressions one and two used *Boone/Crawford MAPE Percentage Difference* as the dependent variable while using the *LevelofAggregation* dummy variables, *UnitsProduced*, *TimeSpanned*, *UnitsProducedperLot*, and *UnitsProducedperLot***LevelofAggregation* as independent variables. *UnitsofMeasure = Dollars* and *TimeSpanned* cannot enter into the OLS regression together due to data limitations; therefore, *TimeSpanned* was used as an independent variable in regressions one and two instead of *UnitsofMeasure = Dollars*. Regressions three through six used *Boone/Crawford MAPE Percentage Difference* as the dependent variable while using the *LevelofAggregation* dummy variables, *UnitsofMeasure=Dollars*, *UnitsProduced*, *UnitsProducedperLot*, and *UnitsProducedperLot***LevelofAggregation* as independent variables. Therefore, *TimeSpanned* was excluded from regressions three through six while *UnitsofMeasure=Dollars* was included.

Table 18: *Boone/Crawford MAPE Percentage Difference* Regression Results

Dependent Variable: *Boone/Crawford MAPE Percentage Difference*	Regression Number					
	1	2	3	4	2	3
Intercept	0.152 (0.036) <0.001	0.155 (0.036) <0.001	-0.041 (0.053) 0.432	0.026 (0.044) 0.560	0.026 (0.045) 0.562	-0.046 (0.051) 0.366
UnitsofMeasure = Dollars			0.133 (0.050) 0.008	0.104 (0.049) 0.033	0.123 (0.047) 0.010	0.138 (0.047) 0.004
LevelofAggregation = 1	0.112 (0.039) 0.005		0.098 (0.039) 0.012	0.059 (0.035) 0.096		
LevelofAggregation = 2	0.105 (0.040) 0.009		0.086 (0.037) 0.021			
LevelofAggregation = 1 or 2		0.108 (0.033) 0.001				0.091 (0.032) 0.005
TimeSpanned	-0.011 (0.004) 0.005	-0.011 (0.004) 0.005				
Adjusted R^2	0.067	0.071	0.049	0.031	0.024	0.053
Number of observations	203	203	234	234	234	234

Each row below the header shows various OLS regression statistics by each independent variable. Each column represents a different OLS regression run with some independent variables entering or leaving among the different regression runs. Within each cell, the coefficient estimate is listed on top, the standard error in the middle in parentheses, and the p-value at the bottom. If a cell is left blank, that independent variable was not included in that OLS regression number. Only regression runs with statistically significant independent variables were reported with some exceptions. The adjusted coefficient of determination (Adjusted R^2) is shown as a goodness-of-fit statistic to compare models with different numbers of independent variables (Hilmer & Hilmer, pp. 161-162). The Adjusted R^2 penalizes the standard R^2 according to the number of independent variables within the OLS regression model (Hilmer & Hilmer, pp. 161-162). Lastly, the number of observations is shown that differs when

TimeSpanned is included in the model. This same format will be used for the remainder of the OLS regression analysis.

The OLS regression results in Table 18 demonstrate that few hypothesized independent variables were statistically significant in the OLS regression model. Concentrating on regression number one, the two dummy variables for *LevelofAggregation* that represent the highest and second-highest levels of aggregation were statistically significant. *TimeSpanned* is also statistically significant. These variables are statistically significant because their p-values are less than the comparison-wise error rate of 0.016. The coefficient estimates on the two dummy variables for *LevelofAggregation* indicate that Boone's Learning Curve is associated with a decrease of 26.4% and a 25.7% MAPE when modeling learning curves at the highest and second-highest levels of aggregation, respectively, in comparison to Crawford's Learning Curve. These values account for the intercept and are valid while holding all else constant. The sign and magnitude of these error reductions are as hypothesized. The coefficient estimates on *TimeSpanned* indicate that with each additional year, Boone's Learning Curve is associated with an increase of 1.1% in terms of MAPE in comparison to Crawford's Learning Curve holding all else constant. This result is in contrast to the hypothesized sign that proposed a longer timespan would lead to more instances of forgetting and hence more plateauing of the observed learning curve. This plateauing was hypothesized to lead to Boone's Learning Curve more accurately explaining observed learning curves; however, these results do not support this hypothesis.

Despite these significant independent variables, the Adjusted R^2 for regression number one remains relatively low at 0.067. This adjusted R^2 can be interpreted as 6.7% of the variation in *Boone/Crawford MAPE Percentage Difference* is explained by the independent variables while penalizing for the number of independent variables in the OLS regression. This relatively low adjusted R^2 signifies that a large amount of variability remains to be explained in the data. This high amount of variability illustrates that there remains limited insight into the instances where Boone's Learning Curve more

accurately models observed learning curves. This high amount of variability of the response variable is consistent with other statistical tests performed prior to OLS regression analysis.

Using the results of Table 18 regression number one, an additional independent variable was created to investigate the effects of combining the two significant dummy variables for *LevelofAggregation*. This dummy variable, *LevelofAggregation = 1 or 2*, is coded one if *LevelofAggregation* was equal to the highest or second-highest level of aggregation and zero otherwise. This new dummy variable supplanted the two previously significant dummy variables for *LevelofAggregation* in regression number two. This new dummy variable was statistically significant in the OLS regression model. The OLS regression results remained similar although the Adjusted R^2 increased slightly. This increased Adjusted R^2 is likely due to the decrease of an independent variable in the OLS regression model.

Several statistical tests were conducted to validate the OLS regression results and statistical inferences in Table 18 regressions one and two. First, regression number one will be reviewed. Regression number one had a Shapiro-Wilk test statistic of 0.829 and a p-value of <0.001. Because the p-value is below the significance level of 0.05, this OLS regression model failed the Shapiro Wilk test of normality. A plot of the studentized residuals confirmed this conclusion. Despite these results, asymptotic properties of large samples can render statistical inferences to be valid despite the distribution of the residuals (Wooldridge, 2016, p. 149-150). The asymptotic properties of large samples are valid contingent on the constant variance of the residuals and theoretical assumptions (Wooldridge, 2016, p. 152). The Breusch-Pagan test of constant variance of the residuals resulted in a test statistic of 7.264 and a p-value of 0.064. Because this p-value is above the significance level of 0.05, the Breusch-Pagan indicated constant variance of the residuals. A plot of the residuals by predicted values confirmed this conclusion. This statistical test validates the use of asymptotic properties of large samples. Multicollinearity was not present as indicated by VIF scores for all independent variables that were less

than 1.3. Cook's Distance values indicated that no observations had a disproportionate effect on the results of the model because all values were less than 0.1. Lastly, outliers were not present in the studentized residuals. These various tests in conjunction with the theoretical underpinnings of these models indicated that the OLS regression results for regressions one and two along with their statistics inferences were valid.

In Table 18, regression two has near-identical conclusions for the OLS regression one model tests except for the Breusch-Pagan test. The Breusch-Pagan test of constant variance of the residuals resulted in a test statistic of 7.007 and a p-value of 0.03. Because this p-value is below the significance level of 0.05, the Breusch-Pagan indicated non-constant variance (heteroskedasticity) of the residuals. A plot of the residuals by predicted values confirmed this conclusion. Failure of this statistical test invalidated the use of asymptotic properties of large samples. Therefore, these tests indicated that the statistical inferences generated from regression two may have been invalid, and the model may not have minimum variance among all unbiased estimators.

The results of the second set of OLS regressions are shown in Table 18 regressions three through six. This set of OLS regressions used *Boone/Crawford MAPE Percentage Difference* as the dependent variable while using *LevelofAggregation* dummy variables, *UnitsofMeasure=Dollars*, *UnitsProduced*, *UnitsProducedperLot*, and *UnitsProducedperLot***LevelofAggregation* as independent variables. *TimeSpanned* was excluded in order to include *UnitsofMeasure = Dollars*. Concentrating on regression number three, the p-value for *LevelofAggregation = 2* was greater than the comparison-wise error rate of 0.016; therefore, this dummy variable was removed from the OLS regression model. Following in regression number four, the previously statistically significant dummy variable *LevelofAggregation = 1* became not statistically significant and was removed from the OLS regression model. In regression number five, *UnitsofMeasure = Dollars* was the sole statistically significant independent variable from the original pool of independent variables. The coefficient estimate on this independent variable

suggests that Boone's Learning Curve is associated with a 14.9% decrease in MAPE when modeling observed learning curves with units of measure in total dollars rather than labor hours in comparison to Crawford's Learning Curve. This value accounts for the intercept and is valid while holding all else constant. The coefficient estimate on the intercept indicates that Boone's Learning Curve is associated with a 2.6% decrease in MAPE in comparison to Crawford's Learning Curve for observed learning curves with units of measure in hours holding all else constant. The relatively low adjusted R^2 signifies that merely 2.4% of the variation in *Boone/Crawford MAPE Percentage Difference* is explained by *UnitsofMeasure = Dollars*.

In Table 18, the additional dummy variable, *LevelofAggregation = 1 or 2*, was entered into the model in regression number six. This dummy variable was statistically significant to a comparison-wise error rate of 0.025 with *UnitsofMeasure = Dollars*. The coefficient estimate and statistical significance of *UnitsofMeasure = Dollars* remained relatively stable with the introduction of this dummy variable. The coefficient estimate on *LevelofAggregation = 1 or 2* indicates a 4.5% decrease in MAPE on average when modeling observed learning curves with the two highest levels of aggregation with units of measure in hours in comparison to Crawford's Learning Curve. This value accounts for the intercept and is valid while holding all else constant. In another interpretation, the coefficient estimate on *LevelofAggregation = 1 or 2* indicates an 18.3% decrease in MAPE on average when modeling observed learning curves with the two highest levels of aggregation in comparison to Crawford's Learning Curve for learning curves with units of measure in total dollars, accounting for the intercept and holding all else constant. These findings are relatively consistent with the hypothesized sign and are similar to the results in regressions one and two. The coefficient estimate on the intercept indicates that Boone's Learning Curve is associated with a 4.6% increase in MAPE in comparison to Crawford's Learning Curve for observed learning curves with units of measure in hours and in the lowest two levels of aggregation, holding all

else constant. This finding is also consistent with hypothesized signs. Lastly, In regression number six, the adjusted R^2 doubled to 5.3% although this amount of explained variability remains relatively low.

Statistical tests were conducted to validate the OLS regression results and statistical inferences in Table 18 regressions five and six. First, regression number five will be reviewed. Regression number five had a Shapiro-Wilk test statistic of 0.764 and a p-value of <0.001. Because the p-value is below the significance level of 0.05, this OLS regression model failed the Shapiro Wilk test of normality. The Breusch-Pagan test of constant variance of the residuals resulted in a test statistic of 0.875 and a p-value of 0.350. Because this p-value is greater than the significance level of 0.05, the Breusch-Pagan indicated constant variance of the residuals. A plot of the residuals by predicted values is inconclusive of this finding. This statistical test validated the use of asymptotic properties of large samples. Multicollinearity could not be present with a single independent variable. Cook's Distance values indicated that no observations had undue effects on the results of the model because all values fell below 0.2. Lastly, no outliers were present in the studentized residuals. Table 18 regression six has identical conclusions for these tests, and multicollinearity was not present because all VIF scores below 1.5. These various tests in conjunction with the theoretical underpinnings of these models indicated that the OLS regression results for regressions one and two along with their statistics inferences were valid.

Next, these prior theory-based independent variables were combined with the data-mining independent variables, and a stepwise regression was used with *Boone/Crawford MAPE Percentage Difference* as the dependent variable. The independent variables for this stepwise regression are *UnitsofMeasure*, *LevelofAggregation*, *UnitsProduced*, *TimeSpanned*, *UnitsProducedperLot*, *UnitsProducedperLot*LevelofAggregation*, *Commodity*, *ComponentEstimated*, *DefenseContractor*,

FirstYearofProduction, *PlatformType*, and *Service*. The results for this stepwise regression are displayed in Table 19.

Table 19: *Boone/Crawford MAPE Percentage Difference* Stepwise Regression Results

Dependent Variable: *Boone/Crawford MAPE Percentage*	Partioned Levels	Regression Statistics
Intercept	N/A	1.034 (0.121) <0.001
TimeSpanned	N/A	-0.011 (0.003) 0.003
Contractor Group 1	6 defense contractors	-0.146 (0.031) <0.001
Contractor Group 2	15 defense contractors	-0.457 (0.102) <0.001
Contractor Group 3	5 defense contractors	-0.355 (0.099) <0.001
PlatformType Group 1	C2ISR, CSAR, Electronic Warfare, Engine, Fighter, Missile, Multi-mission Maritime , Tanker, Transport, Trainer, UAV	-0.323 (0.079) <0.001
PlatformType Group 2	Attack, Anti-submarine, Bomber	-0.296 (0.088) 0.001
Adjusted R^2		0.390
Number of observations		205

The results of the stepwise regression in Table 19 indicate that the theory-based hypothesized variable *TimeSpanned* is statistically significant when controlled for using the data-mining hypothesized variables. The stepwise regression found that three groups of contractors were statistically significant along with two groups of platform types. Each of the five statistically significant groups are represented using a dummy variable that is coded one if the observed learning curve has an attribute in the partitioned level shown in Table 19 column two and zero otherwise. The three *Contractor Groups* share some contractors; therefore, an omitted group is still present in the OLS regression. Each of the coefficient estimates can be interpreted similarly to prior regressions. Unfortunately, using the data-mining hypothesized variables did not reveal any additional statistically significant theory-based

independent variables. However, more variability was explained from the introduction of data-mining independent variables: the adjusted R^2 did increase significantly to 39%. Lastly, the coefficient estimate on *TimeSpanned* is similar to that in Table 18 regression number two. This similarity indicates that *TimeSpanned* is consistently estimated across multiple independent variables and further justifies its coefficient estimate and value in explaining variability in *Boone/Crawford MAPE Percentage Difference.*

Statistical tests were conducted to validate the OLS regression results and statistical inferences in Table 19. The stepwise regression had a Shapiro-Wilk test statistic of 0.895 and a p-value of <0.001. Because the p-value is below the significance level of 0.05, this OLS regression model failed the Shapiro Wilk test of normality. The Breusch-Pagan test of constant variance of the residuals resulted in a test statistic of 25.676 and a p-value of <0.001. Because this p-value is less than the significance level of 0.05, the Breusch-Pagan indicated non-constant variance of the residuals. A plot of the residuals by predicted values confirmed this finding. This statistical test failed to validate the use of asymptotic properties of large samples. Multicollinearity was suspect in two independent variables: *Contractor Group 2* and *Contractor Group 3* because the VIF scores were 8.121 and 6.702. Multicollinearity was likely because both independent variables shared some of the same defense contractors. Because these VIF scores remain below 10, the *Contractor Group 2* and *Contractor Group 3* remained in the OLS regression model. All other VIF scores were below 4.2. Cook's Distance values indicate that no observations had undue effects on the results of the model because all values fell below 0.160. Lastly, no outliers were present in the studentized residuals. Therefore, these tests indicated that the statistical inferences generated from the OLS regression may have been invalid, and the model may not have minimum variance among all unbiased estimators.

The following two sets of regressions investigate aspects of observed learning curves that are associated with their plateau in comparison to Crawford's Learning Curve. The dependent variable for these tests is Crawford's mean percentage error in the fourth quarter, *Crawford MPEQ4*. This variable

provides a measure of plateauing indifferent to how Boone's Learning Curve models observed data. Greater values of this dependent variable indicate greater plateauing of the observed learning curve in comparison to Crawford's Learning Curve.

The results of the third set of OLS regressions are shown in Table 20. This set of OLS regressions excluded *UnitsofMeasure = Dollars* in order to include *TimeSpanned*. Regression number one indicates that the three *LevelofAggregation* dummy variables are individually statistically significant to a level of 0.05; however, *LevelofAggregation = 3* fails to be significant at a comparison-wise error rate of 0.166. Once *LevelofAggregation = 3* is removed, the remaining *LevelofAggregation* dummy variables become not statistically significant over regressions two and three. This results in none of the original independent variables being statistically significant in the model to an experiment-wise error rate of 0.05.

Table 20: *Crawford MPEQ4* Regression Results excluding *UnitsofMeasure*

Dependent Variable: *Crawford MPEQ4*	Regression Number				
	1	2	3	4	5
Intercept	-0.767 (0.229) 0.001	-0.545 (0.188) 0.004	-0.283 (0.133) 0.034	-0.545 (0.188) 0.004	-0.110 (0.043) 0.012
LevelofAggregation = 1	0.777 (0.312) 0.013	0.555 (0.284) 0.052	0.293 (0.252) 0.247		
LevelofAggregation = 2	0.736 (0.294) 0.013	0.514 (0.264) 0.053			
LevelofAggregation = 1 or 2				0.532 (0.234) 0.024	0.097 (0.054) 0.072
LevelofAggregation = 3	0.683 (0.401) 0.090				
Adjusted R^2	0.021	0.013	0.002	0.018	0.010
Number of observations	234	234	234	234	232

When the additional dummy variable, *LevelofAggregation = 1 or 2*, was entered into the model in regression number four, this variable was statistically significant to a level of 0.05. The coefficient estimate on this dummy variable suggests that on average Crawford's Learning Curve models observed learning curves at the two highest levels of aggregation with -1.3% mean percentage error in the fourth quarter holding all else constant. This negative coefficient estimate indicates Crawford's Learning Curve overestimates rather than underestimates observed learning curves in the fourth quarter when hypothesized aspects of plateauing are present. This finding is in contrast to hypotheses and previous OLS regression results because Crawford's Learning Curve is hypothesized to underestimate observed learning curves when aspects of plateauing are present. In regression number four, the adjusted R^2 remains low at 1.8% indicating that the sole independent variable explains minimal variability in the plateauing of observed learning curves.

Statistical tests were conducted to validate the OLS regression results and statistical inferences of Table 20 regression number four. Regression number four had a Shapiro-Wilk test statistic of 0.222 and a p-value of <0.001. Because the p-value is below the significance level of 0.05, this OLS regression model fails the Shapiro Wilk test of normality. The Breusch-Pagan test of constant variance of the residuals resulted in a test statistic of 3.655 and a p-value of 0.056. Because this p-value is greater than the significance level of 0.05, the Breusch-Pagan indicates constant variance of the residuals. A plot of the residuals by predicted values supports this conclusion. The Breusch-Pagan test validated the use of asymptotic properties of large samples. Multicollinearity cannot be present with a single independent variable. Cook's Distance values indicate that one observations had undue influence on the results of the model while another was suspect because their values were 0.84 and 0.48. These same observations were also outliers when analyzed using studentized residuals.

The two influential and outlier observations were investigated along with their influence on the OLS regression coefficient estimates. Both observations were valid for inclusion in the dataset but had

significantly negative MAPE values over 900 times the standard deviation from the mean. When both observations were removed from the model, the OLS regression results changed drastically as shown in Table 20 regression number five. Moreover, the p-value on *LevelofAggregation = 1 or 2* became not statistically significant.

With these alterations to the dataset, there were no hypothesized independent variables that explained any variation of *Crawford MPEQ4* to a statistically significant degree. The removal of these influential and outlier observations did not result in changes to final regression results using any other combination of independent variables. The final set of OLS regressions that excluded *TimeSpanned* in order to include *UnitsofMeasure = Dollars* converged to the results in Table 20; therefore, both regressions share the same final results that no independent variables were statistically significant.

Lastly, the theory-based independent variables were combined with the data-mining independent variables, and a stepwise regression used *Crawford MPEQ4* as the dependent variable. The independent variables for this stepwise regression are *UnitsofMeasure*, *LevelofAggregation*, *UnitsProduced*, *TimeSpanned*, *UnitsProducedperLot*, *UnitsProducedperLot*LevelofAggregation*, *Commodity*, *ComponentEstimated*, *DefenseContractor*, *FirstYearofProduction*, *PlatformType*, and *Service*. The stepwise regression initially displayed statistically significant results for a single level of the data-mining variable *PlatformType*. However, OLS regression model tests indicated that one observation had a high amount of influence with a Cook's Distance value of approximately 2.9. Once this observation was removed, the dummy variable representing *PlatformType* became not statistically significant. The stepwise regression was reperformed with this excluded observation. After several stepwise regression runs, four observations were excluded due to their Cook's Distance values being above 0.5. Eventually, a model with statistically significant variables was created as shown in Table 21.

Table 21: *Crawford MPEQ4* Stepwise Regression Results

Dependent Variable: *Crawford MPEQ4*	Partioned Levels	Regression Statistics
Intercept	N/A	0.245 (0.061) <0.001
PlatformType Group 1	Trainer, Transport	0.483 (0.093) <0.001
PlatformType Group 2	C2ISR, CSAR, Electronic Warfare, Engine, Fighter, Missile, Multi-mission Maritime, Tanker, UAV	-0.269 (0.065) <0.001
Component Estimated Group	Auxiliary Power Plant, Mission Computer, Wing, Airframe, Surface Controls	-0.155 (0.069) 0.016
Adjusted R^2		0.123
Number of observations		230

The results of the stepwise regression in Table 21 indicate that no theory-based independent variables were statistically significant when additional variation was explained using the data-mining independent variables. The stepwise regression found that two groups of *PlatformType* were statistically significant along with one group of *ComponentEstimated*. Each of the coefficient estimates can be interpreted similarly to prior regressions. Despite the lack of presence of statistically significant theory-based independent variables, the introduction of data-mining independent variables increased the amount of variability explained to 12.3% as indicated by the Adjusted R^2.

Statistical tests were conducted to validate the OLS regression results and statistical inferences in Table 21. The stepwise regression had a Shapiro-Wilk test statistic of 0.723 and a p-value of <0.001. Because the p-value is below the significance level of 0.05, this OLS regression model failed the Shapiro Wilk test of normality. The Breusch-Pagan test of constant variance of the residuals resulted in a test statistic of 3.112 and a p-value of 0.375. Because this p-value is greater than the significance level of 0.05, the Breusch-Pagan indicated constant variance of the residuals. A plot of the residuals by predicted values confirmed this finding. This statistical test validated the use of asymptotic properties of large samples. Multicollinearity was not present as indicated by all VIF scores being below 1.6. Cook's

Distance values indicate that no observations had undue effects on the results of the model because all values fell below 0.37. Lastly, no outliers were present in the studentized residuals. Therefore, these tests indicated that the statistical inferences generated from the OLS regression are valid. Despite these statistically significant and valid results, four observations were excluded from the stepwise regression before any statistically significant variables could be found. Additionally, these results provide little value in uncovering statistically significant theory-based independent variables.

Conclusions & Recommendations

Research Conclusions

The diverse dataset utilized in Phase 1 was augmented with observed learning curves estimated at the flyaway cost level to form the basis of the dataset used in Phase 2. Phase 2 utilized Unit Theory learning curve predictions generated from analyses in Phase 1. Observed learning curve data, Crawford's Learning Curve data, and Boone's Learning Curve data were used to answer a variety of research questions. In summary, these analyses sought to determine 1) if the plateauing and forgetting phenomena occur using traditional learning curve theory, 2) how Boone's Learning Curve models observed learning curve data in comparison to traditional learning curve theory, and 3) which program attributes, if any, can explain the programs that are best modeled by Boone's Learning Curve as well as explain the prevalence of plateauing and forgetting phenomena.

In order to understand if the plateauing and forgetting phenomena occur using the traditional learning curve theory, the mean percentage error of Crawford's Learning Curve in the fourth quarter was isolated. These fourth quarter mean percentage errors indicated that Crawford's Learning Curve did not systematically underestimate observed learning curves. In fact, the statistical tests across all quarters indicated Crawford's Learning Curve systematically overestimated observed learning curves with a high amount of variability. These findings suggest that the plateauing and forgetting phenomena are not systematically present in observed data when modeled with Crawford's Learning Curve. This conclusion limits instances where Boone's Learning Curve can more accurately model observed learning curve data in comparison to Crawford's Learning Curve.

These mean percentage error of Crawford's Learning Curve in the fourth quarter were then analyzed using proportions. Crawford's Learning Curve underestimated 47.6% of the observed learning curves tested. Therefore, these hypothesis tests of proportions also failed to indicate that Crawford's Learning Curve systematically underestimates observed learning curve data in the fourth quarter. Lastly,

for the proportion analysis, the proportion of learning curves that Crawford's Learning Curve underestimated in both the first and last quarter was investigated. These instances represent learning curves with high rates of learning decaying to low rates of learning; Boone's Learning Curve models these observed learning curves exceptionally well. The proportion of learning curves that experienced high rates of learning decaying to low rates of learning was 35% when modeled using Crawford's Learning Curve. This proportion highlights that if an observed learning curve is experiencing high rates of learning in the first quarter, it is also more likely to experience low rates of learning in the last quarter when modeled with the Crawford's Learning Curve. Although these results failed to show a systematic presence of plateauing, they emphasize opportunities for Boone's Learning Curve to improve upon Crawford's Learning Curve.

Next, Boone's Learning Curve was analyzed to determine how it models observed learning curve data in comparison to the traditional learning curve theory. Because Boone's Learning Curve contains an additional, empirically-estimated parameter, it is expected to improve upon Crawford's Learning Curve at an aggregate level. However, if Boone's Learning Curve more accurately models observed learning curves that plateau and remains approximately equal to Crawford's Learning Curve in terms of error for observed learning curves that do not plateau, then Boone's Learning Curve provides inherent value in modeling the plateauing phenomenon. To investigate this, a confusion matrix was used to determine how learning curves that plateau interact with learning curves that are more accurately explained by Boone's Learning Curve. The confusion matrix indicated that Boone's Learning Curve more accurately modeled observed learning curves based on if those observed learning curves plateaued for 77% of observations. These results provide mixed conclusions as to if Boone's Learning Curve improves upon Crawford's Learning Curve by more accurately modeling the plateauing phenomenon of observed learning curves or because of its an additional, empirically-estimated parameter.

To further investigate how Boone's Learning Curve models observed learning curves, the MAPE percentage differences between Boone's Learning Curve and Crawford's Learning Curve in the first, middle, and last quarters were investigated. Hypothesis tests were conducted to determine in which quarter Boone's Learning Curve improved upon Crawford's Learning Curve to a statistically significant degree. Outliers were excluded, and the dataset was limited to observations in which Boone's Learning Curve was a significant improvement in error. The hypothesis tests indicated that Boone's Learning Curve improved upon Crawford's Learning Curve to a statistically significant degree in the first quarter only. Furthermore, the tests indicated that the error improvements in the last quarter were not statistically different from zero. The middle quarters' hypothesis test indicated that Boone's Learning Curve was systematically worse at explaining observed data in comparison to Crawford's Learning Curve. Each of these tests had a high amount of variability. These results suggest that Boone's Learning Curve improves upon Crawford's Learning Curve more than it merely having an additional empirically-estimated parameter. However, Boone's Learning Curve was not shown to model the plateau effect in the last quarter to a statistically significant degree. There is a possibility observed learning curves improved in different sections of the learning curves while performing worse in others in comparison to Crawford's Learning Curve. These error improvements in different sections of the learning curve would have caused substantial variations of error in each section. These substantial variations of error improvement within each section could significantly obscure the statistical results due to the variability included in each hypothesis test.

Regression analysis was also used to investigate which program attributes, if any, can explain the programs that are best modeled by Boone's Learning Curve as well as explain the plateauing and forgetting phenomena. For the first part of this regression analysis, the percentage difference between Boone's Learning Curve MAPE and Crawford's Learning Curve MAPE was used as a dependent variable. A variety of independent variables were created using the Literature Review and operationalized using

available data. The independent variables representing the level of aggregation of the learning curve was significant in the model. The regression analysis indicated that Boone's Learning Curve MAPE would decrease 26% on average when modeling learning curves at the highest (flyaway cost) and second highest (air vehicle cost) levels of aggregation, respectively, in comparison to Crawford's Learning Curve. This finding is consistent with the research hypotheses. The independent variable that measured the amount of time that spanned between the first and last years of production was also statistically significant although the coefficient estimate's sign was opposite of the hypothesized sign. In a separate regression, the independent variable representing the learning curve units of measure in total dollars was also statistically significant. This independent variable indicated that Boone's Learning Curve is associated with a 14.9% decrease in error on average when modeling observed learning curves with units of measure in total dollars rather than hours in comparison to Crawford's Learning Curve. Despite these pool of significant variables, several hypothesized were not significant in the model. Additionally, the Adjusted R^2 for these models was relatively low at 6.7% and 5.3%, respectively. These Adjusted R^2 values highlight that a large amount of variability remains to be explained in the dependent variable. These theory-based independent variables were also tested using data-mining independent variables using a mixed stepwise regression. This methodology was employed to potentially expose statistically significant theory-based independent. Some groups of data-mining independent variables were statistically significant and able to explain more variability in the dependent variable; however, no additional statistically significant theory-based independent variables were revealed. Furthermore, the stepwise regression model results and statistical inferences may have been invalid due to not passing overall model tests.

For the second part of this regression analysis, the percentage error between Crawford's Learning Curve and observed data in the fourth quarter of the observed learning curve served as a dependent variable. This variable provided a measure of plateauing indifferent to how Boone's Learning

Curve models observed data. After testing various independent variables and validating the model with statistical tests, no hypothesized independent variables were significant in the model. This finding suggests that either the operationalization of plateauing was inappropriate, there exists too much variability in the data to explain instances of plateauing, or the hypothesized independent variables are unsuitable or were not appropriately operationalized. These theory-based independent variables were also tested using a mixed stepwise regression in order to potentially expose other statistically significant theory-based independent. Some groups of data-mining independent variables were statistically significant and able to explain more variability in the dependent variable; however, no statistically significant theory-based independent variables were revealed. Furthermore, the second set of stepwise regression model results and statistical inferences may have also been invalid due to not passing overall model tests.

Research Limitations

Limitations of Phase 2 should be reviewed to determine the bearing of results and suitability of conclusions. Throughout this analysis, Unit Theory using Crawford's Learning Curve was used solely to compare to observed learning curves and Boone's Learning Curve. This limitation was necessary to conduct a wide variety of tests; however, these results cannot be extended to Cumulative Average Theory using Wright's Learning Curve. Furthermore, program lot data was used instead of unitary data due to data availability. This use of program lot data may have caused data to be aggregated at too high of a level for forgetting and plateauing effects within observed learning curves to be apparent. Lastly, several methods exist to operationalize forgetting and plateauing effects of observed learning curves; however, only one method, percent error in the fourth quarter of Crawford's Learning Curve, was utilized in this research. Various methods also exist to operationalize hypothesized variables to explain the forgetting and plateauing effects. Therefore, other methods of operationalizing these variables

should be explored before claiming the forgetting and plateauing phenomena are not related to the theoretically-based hypothesized variables.

Recommendations for Future Research

In addition to the Recommendations for Future Research from Phase 1, future research should focus on attempting to explain when forgetting and plateauing phenomena occur. Several theory-based and data-mining independent variables were tested using OLS regression; however, very little of the variability in either dependent variable was explained. Other variables not included in this research should be operationalized and tested in order to gain a better understanding of the forgetting and plateauing phenomena. Additionally, there may be a benefit in performing case studies to understand further how Boone's Learning Curve more accurately explains observed learning curves. When research was performed at the aggregate level with a large dataset, there remained too much variability in the data to gather meaningful conclusions.

Bongers (2017) utilized a learning curve equation with a nested forgetting term to test the presence of forgetting in a limited number of Department of Defense programs. This equation and its forgetting term should be tested on a larger number of Department of Defense programs to operationalize the effects of forgetting differently than used in this research. This initiative will provide further evidence as to if the effects of forgetting are present in Department of Defense programs. Related to analysis performed by Bongers (2017), there may be benefit in analyzing learning curves using a time-series analysis when modeling learning curves. A time-series analysis of learning curves could shift the paradigm to reveal new insights and avenues of research.

Lastly, Boone's Learning Curve was a specific formulation of the general adapted learning curve model shown in Equation 11. This general model formulized that the learning curve exponent was some function of the independent variable. Other specific models created from this general model should be explored further to accurately model instances of diminishing rates of learning.

Summary

This research sought to determine 1) if the plateauing and forgetting phenomena occur using traditional learning curve theory, 2) how Boone's Learning Curve models observed learning curve data in comparison to traditional learning curve theory, and 3) which program attributes, if any, can explain the programs that are best modeled by Boone's Learning Curve as well as explain the plateauing and forgetting phenomena. The various statistical results indicate that the plateauing and forgetting phenomena do not systematically occur using the traditional Unit Learning Curve theory. Despite these results, there remain instances where plateauing and forgetting occur that provide opportunities for Boone's Learning Curve to improve upon Crawford's Learning Curve.

Additionally, results provide mixed conclusions as to if Boone's Learning Curve improves upon Crawford's Learning Curve by more accurately modeling the plateauing phenomenon of observed learning curves or because of its an additional, empirically-estimated parameter. Tests also did not show that Boone's Learning Curve improved upon Crawford's Learning Curve to a statistically significant degree in the last quarter but instead in the first quarter. This finding brings into question if Boone's Learning Curve does in fact more accurately model the plateauing phenomenon.

Lastly, few hypothesized independent variables were significant in explaining when Boone's Learning Curve more accurately models observed learning curves, and a significant amount of variability in the data remained. Despite these results, OLS regression analysis confirmed that the level of aggregation, the units of measure, and the time span of a program explain a small amount of variability of the degree to which Boone's Learning Curve more accurately models observed learning curves. The OLS regression results for explaining plateauing independent of Boone's Learning Curve indicated that no hypothesized independent variables were significant in the model even when non-theoretically based independent variables were included in the OLS regression model.

These results indicate that Boone's Learning Curve may provide value by more accurately explaining observed learning curves. However, the theoretical explanations as to where Boone's Learning Curve improves upon the traditional learning curve theory and under what circumstances Boone's Learning Curve is more appropriate to use than the traditional learning curve theory remains to be discovered.

Appendix A: Learning Curve Error Comparisons using Cumulative Average and Unit Theories

Table A1: Error Comparison using Cumulative Average Theory for All Programs

Program	Number of Lots	Number of Units	Component Estimated	Units	Traditional RMSE	Boone RMSE	RMSE Percentage Difference	Traditional MAPE	Boone MAPE	MAPE Percentage Difference
Program 1	6	483	PME - Air Vehicle	Dollars	557.89	111.69	80.0%	3.6%	0.7%	80.9%
Program 1	6	483	PME - Air Vehicle	Hours	15.51	0.32	98.0%	27.2%	0.5%	98.2%
Program 1	6	483	Airframe	Dollars	411.22	114.07	72.3%	2.8%	0.7%	74.7%
Program 1	6	483	Airframe	Hours	21.66	1.51	93.0%	31.0%	1.7%	94.6%
Program 2	5	638	PME - Air Vehicle	Dollars	129.77	6.47	95.0%	2.6%	0.1%	95.6%
Program 3	5	500	PME - Air Vehicle	Dollars	1630.26	291.09	82.1%	20.8%	3.9%	81.5%
Program 4	19	205	PME - Air Vehicle	Dollars	581.70	581.84	0.0%	3.1%	3.1%	0.0%
Program 4	19	205	Airframe	Dollars	545.98	546.44	-0.1%	3.2%	3.2%	-0.1%
Program 5	7	459	PME - Air Vehicle	Dollars	400.84	44.72	88.8%	2.7%	0.3%	88.2%
Program 5	7	459	Electronic Warfare (1)	Dollars	4.78	3.24	32.3%	7.2%	4.8%	33.7%
Program 6	6	98	PME - Air Vehicle	Dollars	99.32	32.22	67.6%	1.1%	0.3%	69.4%
Program 6	6	98	Electronic Warfare (1)	Dollars	12.74	1.68	86.8%	3.6%	0.6%	82.4%
Program 6	6	98	Electronic Warfare (2)	Dollars	15.05	13.34	11.4%	2.3%	2.0%	12.9%
Program 6	6	98	Electronic Warfare (3)	Dollars	1.77	1.06	40.3%	1.3%	0.8%	39.6%
Program 7	7	110	PME - Air Vehicle	Dollars	144.98	98.31	32.2%	1.0%	0.7%	32.6%
Program 7	7	110	Electronic Warfare (1)	Dollars	8.39	3.59	57.2%	2.7%	1.0%	61.3%
Program 7	7	110	Electronic Warfare (2)	Dollars	140.32	107.21	23.6%	1.2%	0.8%	27.5%
Program 7	7	110	Electronic Warfare (3)	Dollars	0.92	0.92	0.0%	0.5%	0.5%	-0.1%
Program 7	7	110	Electronic Warfare (4)	Dollars	140.74	111.32	20.9%	1.3%	1.0%	24.2%
Program 7	7	110	Electronic Warfare (5)	Dollars	21.28	21.04	1.1%	2.2%	2.1%	5.2%
Program 8	8	3529	PME - Air Vehicle	Dollars	27.72	23.62	14.8%	1.4%	1.3%	7.8%
Program 8	8	3529	PME - Air Vehicle	Hours	0.10	0.13	-27.5%	1.1%	1.3%	-27.9%
Program 9	9	3798	PME - Air Vehicle	Dollars	166.55	170.74	-2.5%	8.4%	8.8%	-3.7%
Program 10	10	3803	PME - Air Vehicle	Dollars	8.00	4.83	39.6%	2.5%	1.2%	51.7%
Program 10	10	3803	PME - Air Vehicle	Hours	24.45	14.01	42.7%	4.3%	2.0%	54.0%
Program 11	6	180	PME - Air Vehicle	Dollars	514.03	508.41	1.1%	0.9%	0.8%	4.2%
Program 12	10	20	PME - Air Vehicle	Dollars	699.20	694.08	0.7%	5.8%	5.7%	1.0%
Program 12	10	20	PME - Air Vehicle	Hours	1042.53	906.48	13.1%	9.5%	8.4%	11.8%
Program 12	7	11	Mission Computer (1)	Dollars	44.33	44.34	0.0%	2.5%	2.5%	0.0%
Program 13	5	100	PME - Air Vehicle	Dollars	53386.73	21143.70	60.4%	12.8%	4.8%	62.1%
Program 13	5	100	Airframe	Dollars	6569.73	6577.96	-0.1%	3.7%	3.7%	0.0%
Program 14	5	275	PME - Air Vehicle	Dollars	3114.02	145.54	95.3%	3.8%	0.2%	95.5%
Program 15	10	77	PME - Air Vehicle	Dollars	44385.98	44390.21	0.0%	9.5%	9.5%	0.0%
Program 15	12	83	PME - Air Vehicle	Hours	79241.95	79247.53	0.0%	6.5%	6.5%	0.0%
Program 15	11	83	Airframe	Dollars	39624.41	39628.03	0.0%	10.6%	10.6%	0.0%
Program 15	10	68	Mission Computer (1)	Dollars	1959.32	1959.39	0.0%	17.0%	17.0%	0.0%
Program 16	9	76	PME - Air Vehicle	Dollars	436.29	144.41	66.9%	2.6%	1.0%	62.9%
Program 17	5	50	PME - Air Vehicle	Dollars	13023.63	13029.76	0.0%	2.8%	2.8%	-0.1%
Program 18	9	31	PME - Air Vehicle	Dollars	2942.49	2941.89	0.0%	1.0%	0.9%	0.0%
Program 19	6	98	PME - Air Vehicle	Dollars	313.32	313.44	0.0%	0.5%	0.5%	-0.1%
Program 20	11	84	PME - Air Vehicle	Dollars	1568.74	1121.89	28.5%	1.7%	1.5%	7.8%
Program 20	7	59	Electronic Warfare (1)	Dollars	452.77	142.98	68.4%	4.6%	1.3%	71.5%
Program 20	11	84	Electronic Warfare (2)	Dollars	98.75	76.54	22.5%	3.4%	3.6%	-6.3%
Program 20	7	59	Electronic Warfare (5)	Dollars	562.52	517.37	8.0%	1.8%	1.8%	1.7%
Program 21	6	326	PME - Air Vehicle	Dollars	5267.10	2408.78	54.3%	8.0%	4.2%	47.4%
Program 21	7	344	Airframe	Dollars	4819.45	2544.26	47.2%	9.1%	5.4%	40.4%
Program 21	7	344	Avionics	Dollars	763.21	429.87	43.7%	6.6%	3.9%	40.8%
Program 21	14	453	PME - Air Vehicle	Hours	3493.62	3495.94	-0.1%	4.8%	4.8%	0.1%
Program 21	14	453	Airframe	Hours	4338.35	4339.68	0.0%	6.2%	6.2%	0.1%
Program 22	8	538	PME - Air Vehicle	Hours	856.69	857.66	-0.1%	2.5%	2.6%	-0.1%
Program 22	8	538	Airframe	Hours	5608.46	5609.69	0.0%	15.8%	15.9%	-0.1%
Program 23	5	469	PME - Air Vehicle	Dollars	637.47	339.27	46.8%	5.4%	2.9%	47.3%
Program 24	10	59	PME - Air Vehicle	Dollars	3032.51	3033.05	0.0%	2.2%	2.2%	0.0%
Program 25	9	348	PME - Air Vehicle	Dollars	117.82	118.06	-0.2%	0.9%	0.9%	-0.2%
Program 26	5	109	PME - Air Vehicle	Dollars	3247.44	1676.81	48.4%	11.0%	6.0%	45.7%
Program 26	5	109	PME - Air Vehicle	Hours	607.07	453.47	25.3%	5.7%	4.2%	25.9%
Program 27	18	631	PME - Air Vehicle	Dollars	1669.56	913.34	45.3%	3.6%	1.9%	46.2%
Program 28	6	425	PME - Air Vehicle	Dollars	319.98	322.00	-0.6%	0.9%	0.9%	-0.6%
Program 28	7	522	PME - Air Vehicle	Hours	1776.09	1785.61	-0.5%	1.8%	1.8%	-0.1%
Program 28	7	522	Airframe	Hours	1389.91	1393.86	-0.3%	1.2%	1.2%	-0.2%
Program 29	9	358	PME - Air Vehicle	Hours	610.63	611.08	-0.1%	0.9%	0.9%	0.4%
Program 29	9	358	Airframe	Hours	4804.76	2124.23	55.8%	7.3%	2.9%	60.1%
Program 30	5	204	PME - Air Vehicle	Dollars	513.53	212.75	58.6%	1.2%	0.5%	56.1%
Program 31	5	605	PME - Air Vehicle	Dollars	1482.62	629.10	57.6%	6.1%	2.9%	53.1%
Program 32	5	870	PME - Air Vehicle	Dollars	61.31	61.60	-0.5%	0.4%	0.4%	-0.3%
Program 33	10	178	PME - Air Vehicle	Dollars	7093.55	7101.08	-0.1%	3.5%	3.5%	-0.1%
Program 33	10	178	PME - Air Vehicle	Hours	8131.11	8144.11	-0.2%	2.9%	2.9%	-0.1%
Program 33	10	178	Airframe	Dollars	1906.94	1910.76	-0.2%	1.7%	1.7%	-0.2%
Program 33	10	712	Body	Dollars	232.17	234.86	-1.2%	1.5%	1.6%	-1.3%

Appendix A (continued): Learning Curve Error Comparisons using Cumulative Average and Unit Theories

Table A1(continued): Error Comparison using Cumulative Average Theory for All Programs

Program	Number of Lots	Number of Units	Component Estimated	Units	Traditional RMSE	Boone RMSE	RMSE Percentage Difference	Traditional MAPE	Boone MAPE	MAPE Percentage Difference
Program 33	10	178	Alighting Gear	Dollars	76.6	76.6	0.0%	7.9%	7.9%	0.0%
Program 33	10	178	Auxiliary Power Plant	Dollars	90.7	90.7	-0.1%	3.9%	3.9%	-0.1%
Program 33	10	178	Electronic Warfare (1)	Dollars	775.5	776.1	-0.1%	6.5%	6.5%	-0.1%
Program 33	10	178	Electronic Warfare (2)	Dollars	360.1	273.4	24.1%	58.3%	46.0%	21.2%
Program 33	10	178	Electronic Warfare (3)	Dollars	62.5	62.4	0.2%	5.7%	5.7%	0.1%
Program 33	10	178	Empennage	Dollars	352.2	352.3	0.0%	5.1%	5.1%	-0.1%
Program 33	10	178	Hydraulic	Dollars	22.7	22.7	-0.1%	2.2%	2.2%	-0.1%
Program 33	10	178	Wing	Dollars	296.5	296.9	-0.1%	2.3%	2.3%	-0.1%
Program 34	6	67	PME - Air Vehicle	Dollars	11059.1	11061.2	0.0%	6.6%	6.6%	0.0%
Program 34	6	67	PME - Air Vehicle	Hours	9058.6	9061.7	0.0%	4.4%	4.4%	0.0%
Program 34	6	67	Airframe	Dollars	2798.1	2004.6	28.4%	2.8%	1.7%	37.9%
Program 34	6	201	Body	Dollars	1924.5	828.9	56.9%	19.0%	8.7%	54.0%
Program 34	6	67	Alighting Gear	Dollars	316.5	166.9	47.3%	17.2%	8.3%	51.9%
Program 34	6	67	Electrical	Dollars	50.7	50.7	-0.1%	1.9%	1.9%	-0.1%
Program 34	6	67	Electronic Warfare (1)	Dollars	428.3	428.4	0.0%	5.3%	5.3%	0.0%
Program 34	5	49	Empennage	Dollars	202.2	202.2	0.0%	4.1%	4.1%	0.0%
Program 34	6	67	EO/IR	Dollars	45.6	36.6	19.7%	1.2%	1.1%	13.1%
Program 34	6	67	EOTS	Dollars	347.6	347.7	0.0%	6.5%	6.5%	0.0%
Program 34	6	67	Hydraulic	Dollars	122.3	101.5	17.0%	8.4%	6.2%	26.8%
Program 34	6	67	Mission Computer (1)	Dollars	484.8	484.9	0.0%	0.9%	0.9%	-0.2%
Program 34	6	67	Surface Controls	Dollars	196.0	196.0	0.0%	4.9%	4.9%	0.0%
Program 34	6	67	Wing	Dollars	998.4	998.6	0.0%	3.3%	3.3%	-0.1%
Program 35	5	41	PME - Air Vehicle	Dollars	3578.6	3579.8	0.0%	1.5%	1.5%	0.0%
Program 35	5	41	PME - Air Vehicle	Hours	2003.7	2004.7	0.0%	1.1%	1.1%	0.0%
Program 35	5	50	Airframe	Dollars	609.3	610.4	-0.2%	0.6%	0.6%	-0.3%
Program 35	5	150	Body	Dollars	235.8	156.5	33.6%	1.9%	1.4%	28.0%
Program 35	5	50	Alighting Gear	Dollars	13.2	13.2	-0.1%	0.5%	0.5%	0.0%
Program 35	5	50	Electronic Warfare (1)	Dollars	259.6	259.7	0.0%	3.2%	3.2%	0.0%
Program 35	5	50	EO/IR	Dollars	121.6	121.7	0.0%	1.3%	1.3%	-0.1%
Program 35	5	50	EOTS	Dollars	177.9	177.9	0.0%	2.8%	2.8%	-0.1%
Program 35	5	50	Hydraulic	Dollars	58.2	58.2	0.0%	3.1%	3.1%	0.0%
Program 35	5	50	Radar	Dollars	256.8	256.9	0.0%	3.2%	3.2%	0.0%
Program 35	5	50	Surface Controls	Dollars	121.5	121.5	0.0%	2.6%	2.6%	0.0%
Program 35	5	50	Wing	Dollars	1213.5	1213.6	0.0%	3.8%	3.8%	0.0%
Program 36	13	1285	PME - Air Vehicle	Dollars	28.8	29.4	-2.1%	0.6%	0.6%	-2.2%
Program 37	6	432	PME - Air Vehicle	Dollars	791.3	793.8	-0.3%	3.4%	3.4%	-0.4%
Program 38	6	52	PME - Air Vehicle	Dollars	253.6	154.9	38.9%	1.2%	0.7%	41.6%
Program 38	6	44	PME - Air Vehicle	Hours	831.5	614.2	26.1%	1.3%	0.8%	42.8%
Program 39	19	1023	PME - Air Vehicle	Dollars	19.3	19.3	-0.2%	0.7%	0.7%	-0.2%
Program 40	5	1725	PME - Air Vehicle	Dollars	19.2	0.6	96.7%	2.0%	0.1%	97.0%
Program 41	10	16	PME - Air Vehicle	Dollars	14787.6	14787.8	0.0%	5.2%	5.2%	0.0%
Program 41	10	16	Data Link (1)	Dollars	138.8	138.8	0.0%	3.7%	3.7%	0.0%
Program 42	11	203	PME - Air Vehicle	Dollars	1000.0	1000.1	0.0%	7.0%	7.0%	0.0%
Program 42	11	899	Electronic Warfare (1)	Dollars	67.5	67.7	-0.2%	13.9%	13.9%	-0.5%
Program 43	11	203	PME - Air Vehicle	Dollars	1121.7	1121.9	0.0%	5.5%	5.5%	0.0%
Program 43	13	251	PME - Air Vehicle	Hours	1944.2	1762.2	9.4%	3.4%	3.2%	6.1%
Program 44	5	136	PME - Air Vehicle	Dollars	57.1	16.3	71.4%	1.1%	0.3%	71.4%
Program 45	9	155	PME - Air Vehicle	Dollars	149.6	149.7	-0.1%	0.3%	0.3%	-0.1%
Program 46	6	68	PME - Air Vehicle	Dollars	3435.9	3436.0	0.0%	1.7%	1.7%	0.1%
Program 46	6	68	PME - Air Vehicle	Hours	2286.4	2286.6	0.0%	2.6%	2.6%	0.0%
Program 46	6	68	Airframe	Dollars	539.1	527.6	2.1%	2.3%	2.1%	10.9%
Program 46	6	68	Data Link (1)	Dollars	44.0	44.0	0.0%	3.0%	3.0%	0.0%
Program 46	6	68	Electronic Warfare (1)	Dollars	221.8	221.9	0.0%	5.4%	5.4%	0.0%
Program 46	6	68	Electronic Warfare (2)	Dollars	220.0	220.0	0.0%	6.5%	6.5%	0.0%
Program 46	6	68	Electronic Warfare (3)	Dollars	17.7	8.8	50.4%	2.2%	1.0%	54.6%
Program 46	6	68	Electronic Warfare (4)	Dollars	530.0	530.0	0.0%	5.2%	5.2%	0.0%
Program 46	6	68	EO/IR	Dollars	120.7	120.8	0.0%	15.7%	15.7%	0.0%
Program 46	6	68	Mission Computer (1)	Dollars	477.9	478.0	0.0%	4.3%	4.3%	0.0%
Program 47	9	36	PME - Air Vehicle	Dollars	1039.4	1039.4	0.0%	2.5%	2.5%	0.0%
Program 47	9	36	PME - Air Vehicle	Hours	8278.7	8278.6	0.0%	15.5%	15.5%	0.0%
Program 47	9	36	Data Link (1)	Dollars	170.2	170.2	0.0%	17.7%	17.7%	0.0%
Program 48	5	179	PME - Air Vehicle	Dollars	1858.3	391.3	78.9%	3.1%	0.6%	79.4%
Program 49	6	180	PME - Air Vehicle	Dollars	435.3	99.8	77.1%	4.4%	1.0%	76.5%
Program 50	5	488	PME - Air Vehicle	Dollars	349.3	350.7	-0.4%	3.3%	3.4%	-0.8%
Program 51	6	663	PME - Air Vehicle	Dollars	5.6	3.6	36.6%	0.6%	0.4%	24.8%
Program 52	5	380	PME - Air Vehicle	Dollars	456.9	454.6	0.5%	9.0%	8.9%	0.3%
Program 53	6	749	PME - Air Vehicle	Dollars	37.2	36.6	1.7%	0.5%	0.5%	4.3%
Program 54	8	194	PME - Air Vehicle	Dollars	28.8	28.8	-0.1%	0.6%	0.6%	-0.1%
Program 55	9	677	PME - Air Vehicle	Dollars	74.8	74.8	0.0%	1.6%	1.6%	0.0%
Program 56	5	590	PME - Air Vehicle	Dollars	6.6	6.6	0.5%	0.2%	0.2%	6.3%
Program 57	5	579	PME - Air Vehicle	Dollars	22.8	22.8	-0.1%	0.8%	0.8%	0.0%

Table A2: Error Comparison using Unit Theory for All Programs

Program	Number of Lots	Number of Units	Component Estimated	Units	Traditional RMSE	Boone RMSE	RMSE Percentage Difference	Traditional MAPE	Boone MAPE	MAPE Percentage Difference
Program 1	7	503	Airframe	Hours	4.6	3.5	23.4%	7.1%	5.0%	28.7%
Program 1	6	483	PME - Air Vehicle	Hours	5.4	1.5	72.5%	11.3%	2.9%	74.0%
Program 1	7	503	PME - Air Vehicle	Dollars	2260.6	517.0	77.1%	12.9%	3.2%	75.2%
Program 1	7	503	Airframe	Dollars	2383.2	857.9	64.0%	14.6%	4.9%	66.4%
Program 2	5	638	PME - Air Vehicle	Dollars	315.4	195.3	38.1%	5.8%	4.3%	26.3%
Program 3	5	500	PME - Air Vehicle	Dollars	2984.5	1120.2	62.5%	49.4%	17.6%	64.4%
Program 4	7	357	Airframe	Dollars	2662.2	2664.3	-0.1%	13.1%	13.2%	-0.1%
Program 4	9	424	PME - Air Vehicle	Dollars	9323.3	4999.8	46.4%	37.9%	14.1%	62.8%
Program 5	19	205	Airframe	Dollars	2446.1	2445.8	0.0%	12.6%	12.6%	-0.3%
Program 5	19	205	PME - Air Vehicle	Dollars	3228.6	3228.9	0.0%	12.4%	12.4%	0.0%
Program 6	7	459	Electronic Warfare (1)	Dollars	20.9	20.9	0.0%	30.8%	30.8%	0.0%
Program 6	7	459	PME - Air Vehicle	Dollars	1439.9	738.1	48.7%	11.3%	5.9%	47.2%
Program 7	5	321	PME - Air Vehicle	Dollars	37.9	33.3	12.2%	3.8%	3.8%	1.1%
Program 8	6	98	Electronic Warfare (3)	Dollars	5.2	4.9	6.1%	4.8%	4.8%	1.4%
Program 8	6	98	Electronic Warfare (2)	Dollars	84.2	70.3	16.5%	11.1%	10.6%	4.7%
Program 8	6	98	PME - Air Vehicle	Dollars	375.2	339.5	9.5%	4.2%	3.7%	13.4%
Program 8	6	98	Electronic Warfare (1)	Dollars	27.5	18.7	31.9%	10.2%	5.9%	42.5%
Program 9	7	110	Electronic Warfare (5)	Dollars	102.9	99.2	3.5%	9.7%	10.4%	-6.6%
Program 9	7	110	Electronic Warfare (3)	Dollars	6.4	6.4	0.0%	4.7%	4.7%	0.0%
Program 9	7	110	Electronic Warfare (4)	Dollars	653.6	653.6	0.0%	6.2%	6.2%	0.0%
Program 9	7	110	Electronic Warfare (2)	Dollars	709.4	709.4	0.0%	6.1%	6.1%	0.0%
Program 9	7	110	PME - Air Vehicle	Dollars	668.5	668.5	0.0%	5.1%	5.1%	0.0%
Program 9	7	110	Electronic Warfare (1)	Dollars	31.6	29.1	8.0%	8.7%	8.0%	8.3%
Program 10	9	1586	PME - Air Vehicle	Dollars	115.5	115.6	-0.2%	12.5%	12.5%	-0.2%
Program 10	10	1796	PME - Air Vehicle	Hours	150.8	150.9	0.0%	12.5%	12.5%	-0.1%
Program 11	8	3529	PME - Air Vehicle	Hours	0.9	0.7	21.2%	27.5%	44.9%	-63.4%
Program 11	8	3529	PME - Air Vehicle	Dollars	97.1	97.5	-0.4%	10.1%	10.4%	-2.1%
Program 12	16	7891	PME - Air Vehicle	Hours	520.1	525.6	-1.1%	86.2%	86.2%	0.0%
Program 12	21	10035	PME - Air Vehicle	Dollars	243.8	239.2	1.9%	30.1%	28.8%	4.2%
Program 13	6	3385	EO	Dollars	12.1	9.4	22.5%	10.7%	9.6%	10.0%
Program 13	10	3803	PME - Air Vehicle	Dollars	33.6	24.8	26.1%	10.3%	7.5%	27.1%
Program 13	10	3803	PME - Air Vehicle	Hours	130.1	100.5	22.7%	21.5%	17.1%	20.7%
Program 14	6	180	PME - Air Vehicle	Dollars	2249.4	1008.9	55.2%	6.4%	2.3%	64.2%
Program 15	10	20	PME - Air Vehicle	Hours	3430.3	3430.4	0.0%	41.5%	41.5%	0.0%
Program 15	10	20	PME - Air Vehicle	Dollars	3013.9	3013.9	0.0%	17.4%	17.4%	0.0%
Program 15	7	11	Mission Computer (1)	Dollars	213.9	213.9	0.0%	11.6%	11.5%	0.6%
Program 16	5	100	Airframe	Dollars	10807.3	7455.4	31.0%	7.0%	4.1%	41.8%
Program 16	5	100	PME - Air Vehicle	Dollars	137225.9	81884.9	40.3%	51.7%	26.9%	48.0%
Program 17	5	275	PME - Air Vehicle	Dollars	8837.5	1396.3	84.2%	17.6%	3.3%	81.6%
Program 18	12	83	PME - Air Vehicle	Hours	266012.8	266015.3	0.0%	39.3%	39.3%	0.0%
Program 18	11	83	Airframe	Dollars	89956.0	89961.1	0.0%	39.1%	39.1%	0.0%
Program 18	10	68	Mission Computer (1)	Dollars	4143.0	4143.2	0.0%	68.2%	68.2%	0.0%
Program 18	11	83	PME - Air Vehicle	Dollars	82138.6	82143.3	0.0%	23.2%	23.2%	0.0%
Program 19	5	45	Airframe	Dollars	501.2	501.2	0.0%	53.9%	53.9%	0.0%
Program 19	5	45	PME - Air Vehicle	Dollars	649.0	649.0	0.0%	17.6%	17.6%	0.0%
Program 19	5	45	Mission Computer (1)	Dollars	61.7	59.7	3.2%	9.8%	9.7%	1.2%
Program 20	9	76	PME - Air Vehicle	Dollars	1108.7	522.5	52.9%	7.2%	3.6%	49.9%
Program 21	5	50	PME - Air Vehicle	Dollars	24625.3	6362.0	74.2%	7.4%	2.3%	69.5%
Program 22	9	31	PME - Air Vehicle	Dollars	16636.3	16636.4	0.0%	6.6%	6.6%	0.0%
Program 23	5	14	PME - Air Vehicle	Dollars	14475.8	14476.0	0.0%	8.7%	8.7%	0.0%
Program 24	6	98	PME - Air Vehicle	Dollars	2259.9	2260.1	0.0%	3.3%	3.3%	0.0%
Program 25	7	59	Electronic Warfare (5)	Dollars	2808.4	2805.2	0.1%	14.8%	15.4%	-4.0%
Program 25	11	84	PME - Air Vehicle	Dollars	5083.2	4228.8	16.8%	8.7%	9.2%	-5.2%
Program 25	11	84	Electronic Warfare (2)	Dollars	248.9	248.6	0.1%	13.9%	14.3%	-2.9%
Program 25	7	59	Electronic Warfare (1)	Dollars	1259.1	653.3	48.1%	16.1%	7.1%	55.6%
Program 26	7	344	Airframe	Dollars	11474.7	8294.9	27.7%	22.7%	21.5%	5.3%
Program 26	7	344	Avionics	Dollars	2218.8	2102.8	5.2%	29.5%	26.9%	8.8%
Program 26	7	344	PME - Air Vehicle	Dollars	12898.4	8742.1	32.2%	20.7%	16.9%	18.4%
Program 27	14	453	PME - Air Vehicle	Hours	54142.9	53766.4	0.7%	59.9%	63.1%	-5.4%
Program 27	14	453	Airframe	Hours	70415.0	69426.8	1.4%	58.8%	59.1%	-0.5%
Program 28	8	538	PME - Air Vehicle	Hours	3828.8	3829.8	0.0%	9.8%	9.9%	0.0%
Program 28	8	538	Airframe	Hours	3865.3	3866.2	0.0%	7.6%	7.6%	0.0%
Program 29	8	529	Hydraulic	Dollars	156.9	156.4	0.3%	22.3%	22.9%	-2.8%
Program 29	12	477	Airframe	Dollars	6490.2	5974.2	7.9%	14.2%	14.4%	-1.8%
Program 29	12	477	Wing	Dollars	712.3	712.7	-0.1%	27.8%	27.8%	-0.1%
Program 29	11	433	Electronic Warfare (1)	Dollars	57.5	57.5	0.0%	13.5%	13.5%	-0.1%
Program 29	8	309	Electrical	Dollars	230.6	230.7	-0.1%	8.2%	8.2%	0.0%
Program 29	12	1045	Body	Dollars	1922.2	1826.7	5.0%	26.0%	25.9%	0.7%
Program 29	5	177	Empennage	Dollars	32.3	22.0	31.8%	6.1%	4.6%	24.5%
Program 29	12	477	PME - Air Vehicle	Dollars	8218.5	5525.3	32.8%	15.0%	10.2%	32.0%
Program 29	8	309	Alighting Gear	Dollars	205.7	42.2	79.5%	11.6%	2.0%	83.1%
Program 30	5	469	PME - Air Vehicle	Dollars	1283.8	891.8	30.5%	13.5%	8.3%	38.3%
Program 31	10	59	PME - Air Vehicle	Dollars	11978.9	11979.3	0.0%	8.6%	8.6%	0.0%
Program 32	9	348	PME - Air Vehicle	Dollars	430.6	430.8	0.0%	3.5%	3.5%	-0.1%
Program 33	5	109	PME - Air Vehicle	Hours	993.9	994.0	0.0%	9.5%	9.5%	0.0%
Program 33	5	109	PME - Air Vehicle	Dollars	6824.7	6824.8	0.0%	28.2%	28.2%	0.0%
Program 34	18	631	PME - Air Vehicle	Dollars	6926.7	2799.9	59.6%	17.0%	6.6%	61.0%
Program 35	6	425	PME - Air Vehicle	Dollars	1135.8	1137.5	-0.2%	3.5%	3.5%	-0.2%
Program 35	7	522	PME - Air Vehicle	Hours	4615.3	4458.5	3.4%	6.3%	6.1%	3.1%
Program 35	7	522	Airframe	Hours	6757.0	6280.7	7.0%	5.7%	5.4%	4.8%
Program 36	9	358	PME - Air Vehicle	Hours	5118.7	5120.1	0.0%	6.8%	6.8%	0.0%
Program 36	9	358	Airframe	Hours	12155.2	11257.1	7.4%	15.5%	14.3%	7.6%
Program 37	5	204	PME - Air Vehicle	Dollars	1468.7	921.0	37.3%	2.9%	1.9%	36.4%
Program 38	5	605	PME - Air Vehicle	Dollars	2641.9	1527.7	42.2%	14.9%	8.1%	46.0%

Table A2 (continued): Error Comparison using Unit Theory for All Programs

Program	Number of Lots	Number of Units	Component Estimated	Units	Traditional RMSE	Boone RMSE	RMSE Percentage Difference	Traditional MAPE	Boone MAPE	MAPE Percentage Difference
Program 39	5	870	PME - Air Vehicle	Dollars	310.9	311.5	-0.2%	2.3%	2.3%	-0.2%
Program 40	10	178	Electronic Warfare (3)	Dollars	751.2	551.9	26.5%	69.7%	74.7%	-7.1%
Program 40	10	712	Body	Dollars	617.6	577.6	6.5%	4.8%	5.1%	-7.6%
Program 40	10	178	Airframe	Dollars	4251.9	4226.4	0.6%	4.8%	4.9%	-1.0%
Program 40	10	178	Electronic Warfare (2)	Dollars	721.7	721.7	0.0%	393.4%	393.4%	0.0%
Program 40	10	178	Electronic Warfare (1)	Dollars	1642.3	1643.0	0.0%	20.7%	20.7%	0.0%
Program 40	10	178	PME - Air Vehicle	Hours	13454.5	13466.8	-0.1%	6.0%	6.0%	-0.1%
Program 40	10	178	Auxiliary Power Plant	Dollars	385.1	385.1	0.0%	24.9%	24.9%	0.0%
Program 40	10	178	PME - Air Vehicle	Dollars	12231.7	12236.6	0.0%	7.9%	7.9%	0.0%
Program 40	10	178	Alighting Gear	Dollars	233.6	233.6	0.0%	30.1%	30.1%	0.0%
Program 40	10	178	Wing	Dollars	607.4	607.6	0.0%	6.2%	6.2%	0.0%
Program 40	10	178	Empennage	Dollars	702.1	702.1	0.0%	17.4%	17.4%	0.0%
Program 40	10	178	Hydraulic	Dollars	72.2	70.2	2.8%	9.0%	8.8%	2.2%
Program 41	6	67	PME - Air Vehicle	Hours	12741.5	12743.8	0.0%	9.5%	9.5%	0.0%
Program 41	5	49	Empennage	Dollars	242.2	242.2	0.0%	5.8%	5.9%	0.0%
Program 41	6	67	PME - Air Vehicle	Dollars	16643.9	16645.6	0.0%	10.7%	10.7%	0.0%
Program 41	6	67	Surface Controls	Dollars	281.7	281.7	0.0%	7.7%	7.7%	0.0%
Program 41	6	67	EOTS	Dollars	442.3	442.4	0.0%	9.5%	9.5%	0.0%
Program 41	6	67	Wing	Dollars	1927.0	1927.3	0.0%	7.4%	7.4%	0.0%
Program 41	6	67	Electrical	Dollars	57.2	57.2	0.0%	2.1%	2.1%	0.0%
Program 41	6	67	Electronic Warfare (1)	Dollars	547.3	547.3	0.0%	8.1%	8.1%	0.0%
Program 41	6	67	Hydraulic	Dollars	281.5	274.6	2.4%	19.4%	19.0%	2.0%
Program 41	6	67	Mission Computer (1)	Dollars	1698.1	1542.4	9.2%	4.6%	3.7%	19.5%
Program 41	6	67	Airframe	Dollars	6877.8	5547.4	19.3%	8.7%	6.4%	26.8%
Program 41	6	67	Alighting Gear	Dollars	582.3	521.1	10.5%	28.3%	25.0%	11.6%
Program 41	6	67	EO/IR	Dollars	233.0	89.4	61.6%	9.3%	3.1%	66.8%
Program 41	6	201	Body	Dollars	3431.8	2343.2	31.7%	42.6%	29.9%	29.7%
Program 42	5	41	PME - Air Vehicle	Dollars	8498.6	8499.6	0.0%	6.2%	6.2%	0.0%
Program 42	5	41	PME - Air Vehicle	Hours	15696.5	15696.9	0.0%	10.7%	10.7%	0.0%
Program 42	5	50	EOTS	Dollars	593.3	593.3	0.0%	11.6%	11.6%	0.0%
Program 42	5	50	EO/IR	Dollars	578.4	578.4	0.0%	7.5%	7.5%	0.0%
Program 42	5	50	Hydraulic	Dollars	297.0	297.0	0.0%	15.4%	15.4%	0.0%
Program 42	5	50	Surface Controls	Dollars	424.9	424.9	0.0%	11.0%	11.0%	0.0%
Program 42	5	50	Radar	Dollars	733.8	733.8	0.0%	10.9%	10.9%	0.0%
Program 42	5	50	Airframe	Dollars	5222.7	5222.8	0.0%	5.9%	5.9%	0.0%
Program 42	5	50	Electronic Warfare (1)	Dollars	746.5	746.5	0.0%	10.7%	10.7%	0.0%
Program 42	5	50	Wing	Dollars	3726.6	3726.7	0.0%	16.5%	16.5%	0.0%
Program 42	5	50	Alighting Gear	Dollars	78.6	77.4	1.5%	3.6%	3.5%	2.3%
Program 42	5	150	Body	Dollars	1588.5	892.1	43.8%	12.6%	8.7%	30.8%
Program 43	13	1285	PME - Air Vehicle	Dollars	88.1	88.8	-0.8%	1.9%	1.9%	-1.0%
Program 44	6	432	PME - Air Vehicle	Dollars	1621.0	1623.3	-0.1%	10.0%	10.0%	-0.2%
Program 45	9	63	PME - Air Vehicle	Dollars	2152.3	1557.1	27.7%	9.5%	6.4%	33.2%
Program 46	6	44	PME - Air Vehicle	Hours	7736.9	7255.3	6.2%	17.6%	16.7%	4.8%
Program 46	10	113	PME - Air Vehicle	Dollars	797.9	627.0	21.4%	3.8%	2.9%	22.7%
Program 47	19	1023	PME - Air Vehicle	Dollars	115.2	115.2	0.0%	4.3%	4.2%	0.2%
Program 48	5	1725	PME - Air Vehicle	Dollars	59.8	3.1	94.9%	6.8%	0.3%	95.4%
Program 49	10	16	Data Link (1)	Dollars	470.3	470.3	0.0%	20.4%	20.4%	0.0%
Program 49	10	16	PME - Air Vehicle	Dollars	41008.9	41009.2	0.0%	14.1%	14.1%	0.0%
Program 50	7	577	PME - Air Vehicle	Dollars	1674.7	1224.7	26.9%	5.5%	4.6%	15.7%
Program 51	12	244	PME - Air Vehicle	Hours	625.6	612.8	2.0%	191.4%	191.8%	-0.2%
Program 52	11	899	Electronic Warfare (1)	Dollars	90.1	90.2	-0.1%	29.2%	29.3%	-0.1%
Program 52	11	203	PME - Air Vehicle	Dollars	2995.1	2992.0	0.1%	24.9%	23.6%	5.2%
Program 53	13	251	PME - Air Vehicle	Hours	4585.2	4585.2	0.0%	6.7%	6.7%	0.0%
Program 53	11	203	PME - Air Vehicle	Dollars	2459.9	2460.0	0.0%	9.6%	9.6%	0.0%
Program 54	11	184	PME - Air Vehicle	Hours	7010.4	7010.7	0.0%	18.0%	18.0%	0.0%
Program 54	9	134	PME - Air Vehicle	Dollars	1907.3	970.0	49.1%	11.8%	6.5%	44.9%
Program 55	5	136	PME - Air Vehicle	Dollars	321.6	277.7	13.7%	5.5%	4.7%	14.8%
Program 56	9	155	PME - Air Vehicle	Dollars	1356.5	1356.6	0.0%	3.9%	3.9%	0.0%
Program 57	6	68	EO/IR	Dollars	326.0	326.0	0.0%	1261.8%	1261.8%	0.0%
Program 57	6	68	PME - Air Vehicle	Dollars	8574.7	8470.9	1.2%	4.3%	4.3%	-0.5%
Program 57	6	68	Electronic Warfare (1)	Dollars	998.8	998.9	0.0%	58.9%	58.9%	0.0%
Program 57	6	68	Electronic Warfare (2)	Dollars	750.2	750.2	0.0%	31.3%	31.3%	0.0%
Program 57	6	68	Data Link (1)	Dollars	94.8	94.8	0.0%	7.2%	7.2%	0.0%
Program 57	6	68	Electronic Warfare (4)	Dollars	1156.3	1156.3	0.0%	12.2%	12.2%	0.0%
Program 57	6	68	Mission Computer (1)	Dollars	1030.6	1030.6	0.0%	13.0%	13.0%	0.0%
Program 57	6	68	PME - Air Vehicle	Hours	6435.9	6435.0	0.0%	12.3%	12.3%	0.3%
Program 57	6	68	Airframe	Dollars	1443.2	1285.1	11.0%	6.7%	5.4%	18.5%
Program 57	6	68	Electronic Warfare (3)	Dollars	53.4	21.8	59.1%	7.2%	3.0%	58.5%
Program 58	9	36	PME - Air Vehicle	Hours	60347.2	60347.3	0.0%	78.2%	78.2%	0.0%
Program 58	9	36	Data Link (1)	Dollars	227.8	227.8	0.0%	29.3%	29.3%	0.0%
Program 58	9	36	PME - Air Vehicle	Dollars	4570.2	4570.2	0.0%	10.9%	10.9%	0.0%
Program 58	5	18	EO/IR	Dollars	3488.4	3469.8	0.5%	28.8%	28.7%	0.3%
Program 59	5	179	PME - Air Vehicle	Dollars	4583.3	1334.5	70.9%	8.1%	2.8%	65.4%
Program 60	6	180	PME - Air Vehicle	Dollars	1010.5	333.9	67.0%	12.4%	4.6%	63.1%
Program 61	5	488	PME - Air Vehicle	Dollars	502.3	486.5	3.1%	9.2%	7.7%	16.3%
Program 62	6	78	PME - Air Vehicle	Hours	6027.1	5952.3	1.2%	33.8%	34.3%	-1.6%
Program 62	6	97	Airframe	Hours	2648.5	2649.0	0.0%	20.5%	20.5%	0.0%
Program 62	9	110	PME - Air Vehicle	Dollars	13027.5	13028.9	0.0%	24.0%	24.0%	0.0%
Program 63	6	663	PME - Air Vehicle	Dollars	23.2	21.1	9.2%	2.9%	2.6%	11.6%
Program 64	5	380	PME - Air Vehicle	Dollars	1520.9	1521.2	0.0%	57.4%	57.4%	0.0%
Program 65	6	749	PME - Air Vehicle	Dollars	116.6	115.9	0.6%	1.7%	1.8%	-5.1%
Program 66	8	194	PME - Air Vehicle	Dollars	128.3	119.3	7.0%	2.6%	2.4%	8.6%
Program 67	9	677	PME - Air Vehicle	Dollars	273.5	273.5	0.0%	5.1%	5.1%	0.0%
Program 68	5	590	PME - Air Vehicle	Dollars	87.1	87.2	0.0%	2.8%	2.8%	0.0%
Program 69	5	579	PME - Air Vehicle	Dollars	305.7	305.8	0.0%	9.5%	9.5%	0.0%

Appendix B: Descriptive Statistics of Boone's Learning Curve Using Cumulative Average Theory by Group

Table B1: Cumulative Average Theory Percentage Error Differences for RMSE in Units of Dollars

Group	Below Approximately Zero	Approximately Zero	Above Approximately Zero
Number of Observations	7	59	52
Percent of Observations	5.9%	50.0%	44.1%
Mean	-1.1%	0.0%	43.9%
Median	-0.6%	0.0%	44.5%
Standard Deviation	0.9%	0.1%	28.5%

Table B2: Cumulative Average Theory Percentage Error Differences for RMSE in Units of Hours

Group	Below Approximately Zero	Approximately Zero	Above Approximately Zero
Number of Observations	3	11	8
Percent of Observations	13.6%	50.0%	36.4%
Mean	-9.4%	-0.1%	45.4%
Median	-0.5%	0.0%	34.4%
Standard Deviation	15.6%	0.0%	34.4%

Table B3: Cumulative Average Theory Percentage Error Differences for MAPE

Group	Below Approximately Zero	Approximately Zero	Above Approximately Zero
Number of Observations	11	69	60
Percent of Observations	7.9%	49.3%	42.9%
Mean	-0.6%	0.0%	49.8%
Median	-0.8%	0.0%	44.3%
Standard Deviation	8.1%	0.1%	29.6%

Appendix C: Descriptive Statistics of Boone's Learning Curve Using Unit Theory by Group

Table C1: Unit Theory Percentage Error Differences for RMSE in Units of Dollars

Group	Below Approximately Zero	Approximately Zero	Above Approximately Zero
Number of Observations	2	73	66
Percent of Observations	1.4%	51.8%	46.8%
Mean	-0.6%	0.0%	29.6%
Median	-0.6%	0.0%	26.7%
Standard Deviation	0.3%	0.0%	25.3%

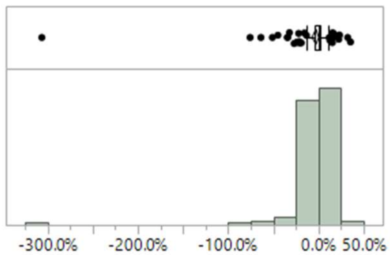
Table C2: Unit Theory Percentage Error Differences for RMSE in Units of Hours

Group	Below Approximately Zero	Approximately Zero	Above Approximately Zero
Number of Observations	1	15	12
Percent of Observations	3.6%	53.6%	42.9%
Mean	-1.1%	0.0%	14.1%
Median	-1.1%	0.0%	6.6%
Standard Deviation	N/A	0.0%	20.3%

Table C3: Unit Theory Percentage Error Differences for MAPE

Group	Below Approximately Zero	Approximately Zero	Above Approximately Zero
Number of Observations	18	84	67
Percent of Observations	10.7%	49.7%	39.6%
Mean	-6.6%	0.0%	30.3%
Median	-2.9%	0.0%	24.5%
Standard Deviation	14.4%	0.1%	26.2%

Appendix D: Test Results of Sample Means of Crawford's Learning Curve Error by Quarter



Quantiles

100.0%	maximum	35.3%
99.5%		35.3%
97.5%		22.5%
90.0%		9.2%
75.0%	quartile	2.2%
50.0%	median	0.0%
25.0%	quartile	-3.9%
10.0%		-12.2%
2.5%		-47.2%
0.5%		-306.3%
0.0%	minimum	-306.3%

Summary Statistics

Mean	-0.034187
Std Dev	0.2572047
Std Err Mean	0.0189101
Upper 95% Mean	0.0031213
Lower 95% Mean	-0.071496
N	185

Test Mean

Hypothesized Value	0
Actual Estimate	-0.0342
DF	184
Std Dev	0.2572
t Test	
Test Statistic	-1.8079
Prob > |t|	0.0723
Prob > t	0.9639
Prob < t	0.0361*

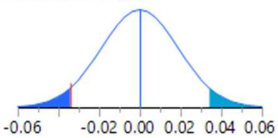
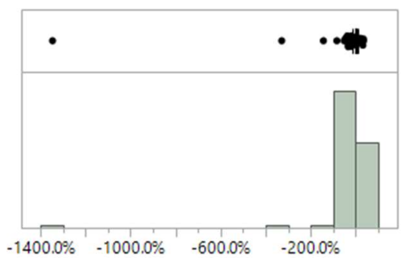


Figure D1: T-Test of Sample Mean for First Quarter Mean Percentage Error



Quantiles

100.0%	maximum	34.2%
99.5%		34.2%
97.5%		18.6%
90.0%		6.1%
75.0%	quartile	1.5%
50.0%	median	-1.3%
25.0%	quartile	-5.5%
10.0%		-16.7%
2.5%		-63.2%
0.5%		-1345.8%
0.0%	minimum	-1345.8%

Summary Statistics

Mean	-0.130342
Std Dev	1.026694
Std Err Mean	0.075484
Upper 95% Mean	0.0185835
Lower 95% Mean	-0.279268
N	185

Test Mean

Hypothesized Value	0
Actual Estimate	-0.1303
DF	184
Std Dev	1.02669
t Test	
Test Statistic	-1.7268
Prob > |t|	0.0859
Prob > t	0.9571
Prob < t	0.0429*

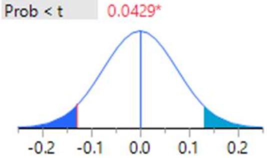
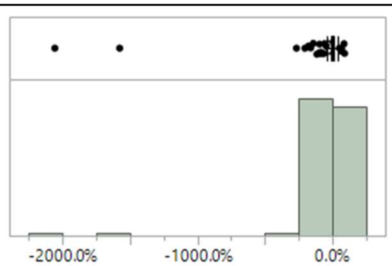


Figure D2: T-Test of Sample Mean for Middle Quarters Mean Percentage Error



Quantiles

100.0%	maximum	89.1%
99.5%		89.1%
97.5%		55.3%
90.0%		20.7%
75.0%	quartile	9.2%
50.0%	median	-0.8%
25.0%	quartile	-11.6%
10.0%		-35.2%
2.5%		-186.6%
0.5%		-2054.1%
0.0%	minimum	-2054.1%

Summary Statistics

Mean	-0.261128
Std Dev	1.9339689
Std Err Mean	0.1421882
Upper 95% Mean	0.0194008
Lower 95% Mean	-0.541657
N	185

Test Mean

Hypothesized Value	0
Actual Estimate	-0.2611
DF	184
Std Dev	1.93397
t Test	
Test Statistic	-1.8365
Prob > |t|	0.0679
Prob > t	0.9661
Prob < t	0.0339*

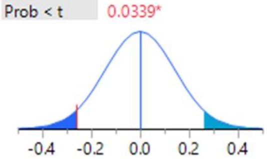


Figure D3: T-Test of Sample Mean for Last Quarter Mean Percentage Error

Appendix D (continued): Test Results of Sample Means of Crawford's Learning Curve by Quarter

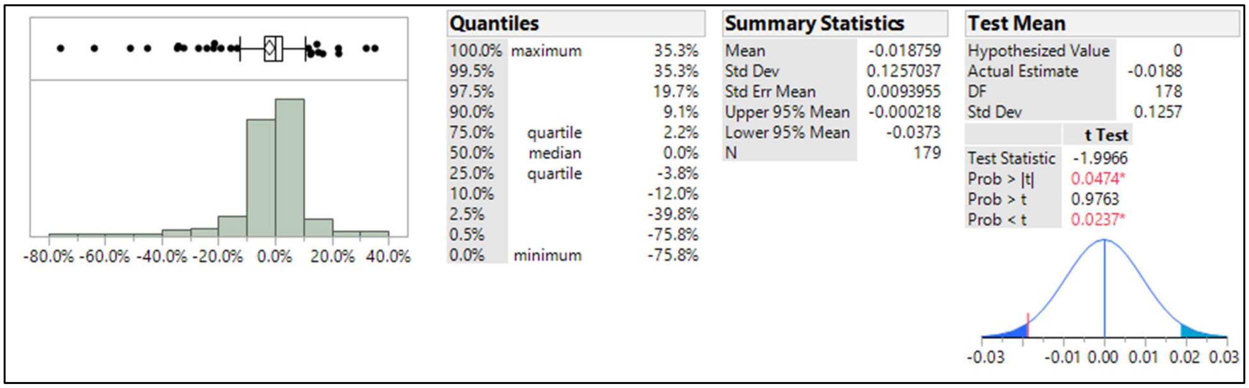


Figure D4: T-Test of Sample Mean for First Quarter Mean Error Percentage Excluding Outliers

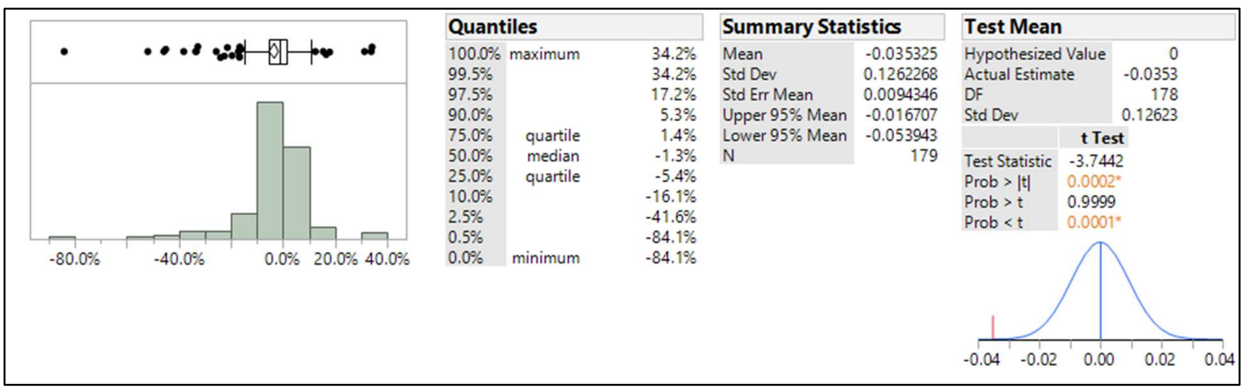


Figure D5: T-Test of Sample Mean for Middle Quarters Mean Error Percentage Excluding Outliers

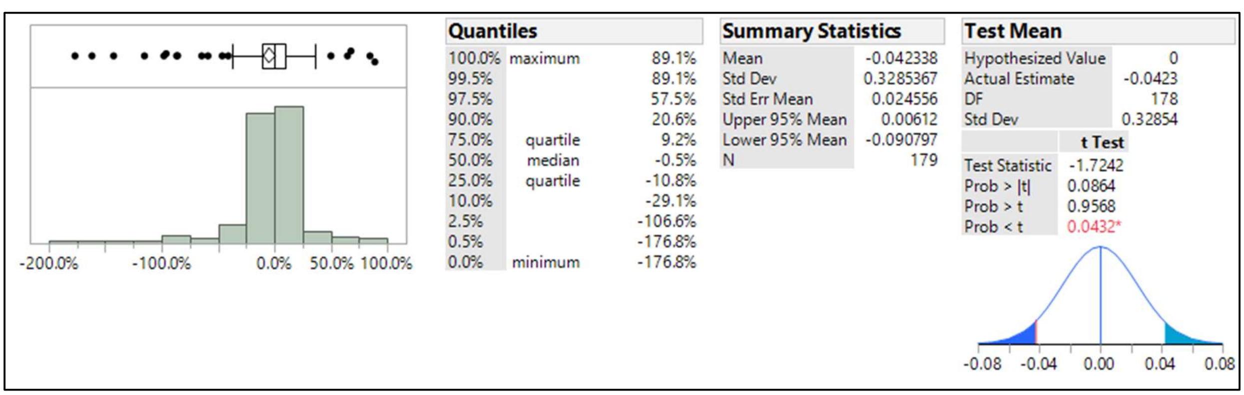


Figure D6: T-Test of Sample Mean for Last Quarter Mean Error Percentage Excluding Outliers

Appendix D (continued): Test Results of Sample Means of Crawford's Learning Curve by Quarter

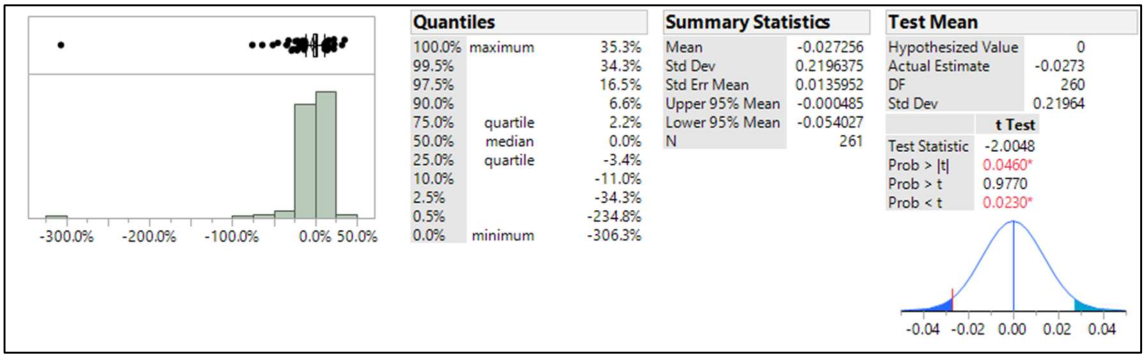


Figure D7: T-Test of Sample Mean for First Quarter Mean Error Percentage Including Flyaway Program Components

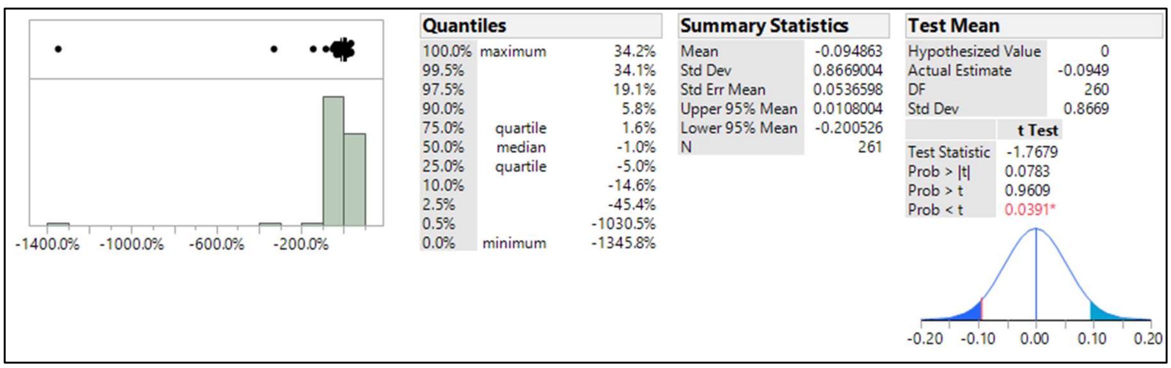


Figure D8: T-Test of Sample Mean for Middle Quarters Mean Error Percentage Including Flyaway Program Components

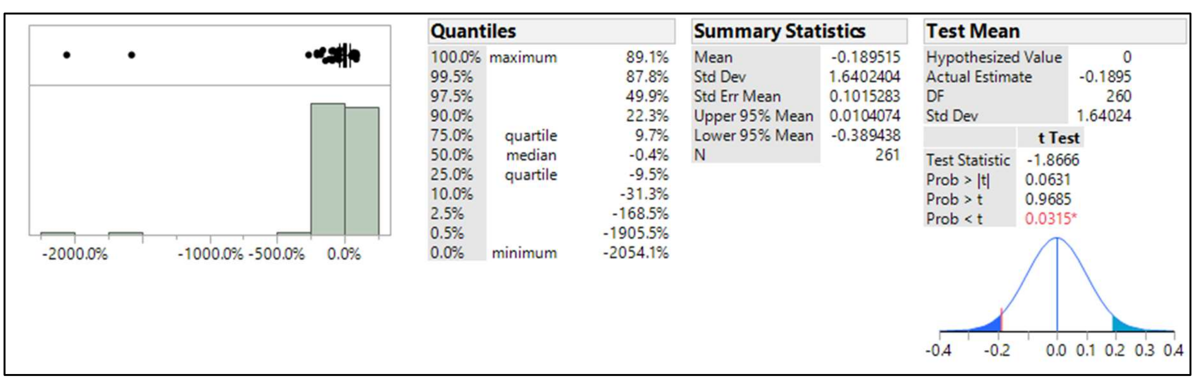


Figure D9: T-Test of Sample Mean for Last Quarter Mean Error Percentage Including Flyaway Program Components

Appendix D (continued): Test Results of Sample Means of Crawford's Learning Curve by Quarter

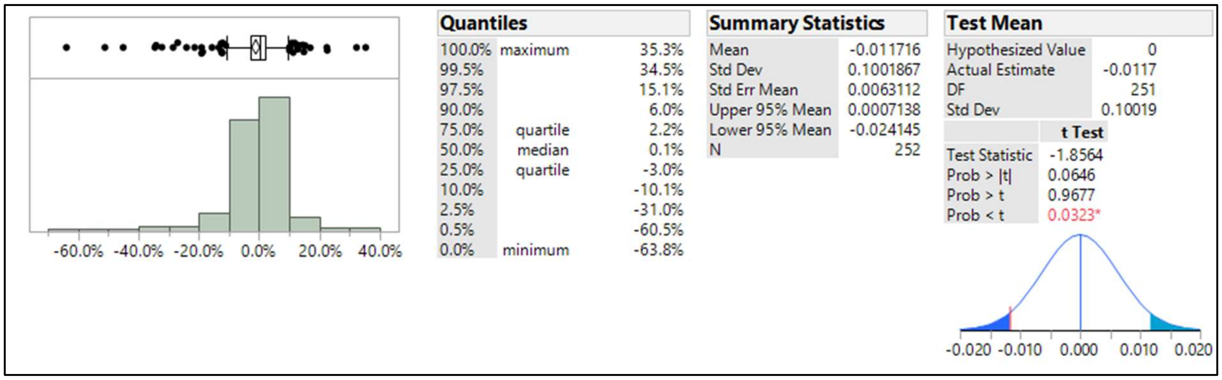


Figure D10: T-Test of Sample Mean for First Quarter Mean Error Percentage Including Flyaway Program Components and Excluding Outliers

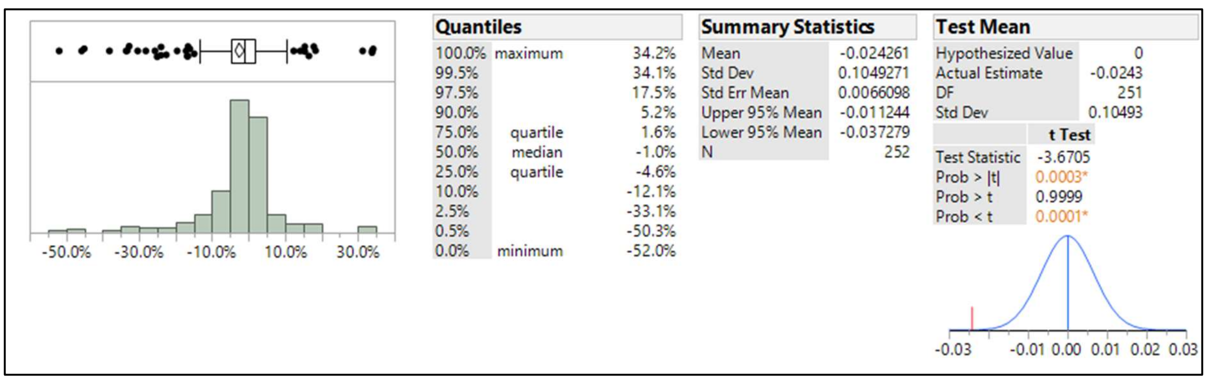


Figure D11: T-Test of Sample Mean for Middle Quarters Mean Error Percentage Including Flyaway Program Components and Excluding Outliers

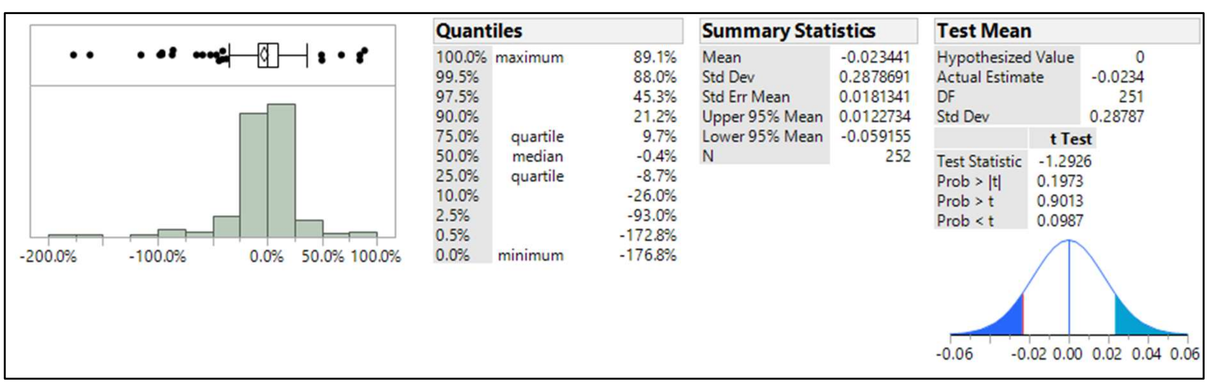


Figure D12: T-Test of Sample Mean for Last Quarter Mean Error Percentage Including Flyaway Program Components and Excluding Outliers

Appendix E: Test Results of Sample Proportions of Crawford's Learning Curve by Quarter

Frequencies

Level	Count	Prob
0	92	0.49730
1	93	0.50270
Total	185	1.00000

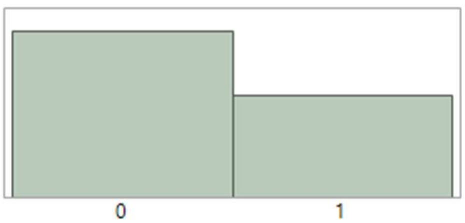
N Missing 0
2 Levels

Test Probabilities

Level	Estim Prob	Hypoth Prob
0	0.49730	0.50000
1	0.50270	0.50000

Binomial Test	Level Tested	Hypoth Prob (p1)	p-Value
Ha: Prob(p > p1)	1	0.50000	0.5000

Figure E1: First Quarter Sample Proportion Hypothesis Test



Frequencies

Level	Count	Prob
0	115	0.62162
1	70	0.37838
Total	185	1.00000

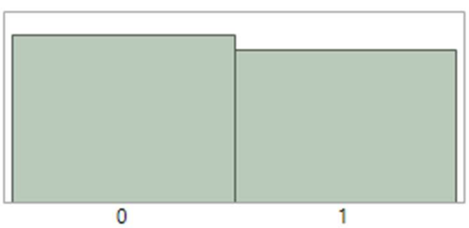
N Missing 0
2 Levels

Test Probabilities

Level	Estim Prob	Hypoth Prob
0	0.62162	0.50000
1	0.37838	0.50000

Binomial Test	Level Tested	Hypoth Prob (p1)	p-Value
Ha: Prob(p > p1)	1	0.50000	0.9997

Figure E2: Middle Quarters Sample Proportion Hypothesis Test



Frequencies

Level	Count	Prob
0	97	0.52432
1	88	0.47568
Total	185	1.00000

N Missing 0
2 Levels

Test Probabilities

Level	Estim Prob	Hypoth Prob
0	0.52432	0.50000
1	0.47568	0.50000

Binomial Test	Level Tested	Hypoth Prob (p1)	p-Value
Ha: Prob(p > p1)	1	0.50000	0.7688

Figure E3: Last Quarter Sample Proportion Hypothesis Test

Appendix E (continued): Test Results of Sample Proportions of Crawford's Learning Curve by Quarter

Frequencies

Level	Count	Prob
0	128	0.49042
1	133	0.50958
Total	261	1.00000

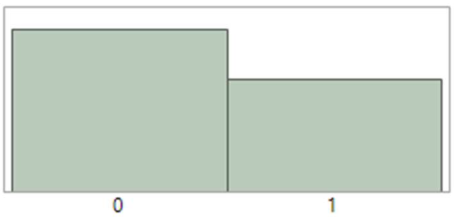
N Missing 0
2 Levels

Test Probabilities

Level	Estim Prob	Hypoth Prob
0	0.49042	0.50000
1	0.50958	0.50000

Binomial Test	Level Tested	Hypoth Prob (p1)	p-Value
Ha: Prob(p > p1)	1	0.50000	0.4023

Figure E4: First Quarter Sample Proportion Hypothesis Test including Flyaway Program Components



Frequencies

Level	Count	Prob
0	154	0.59004
1	107	0.40996
Total	261	1.00000

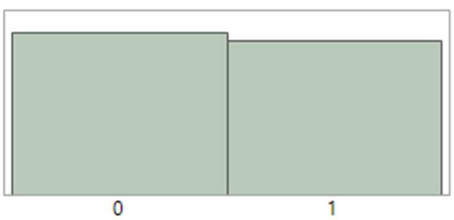
N Missing 0
2 Levels

Test Probabilities

Level	Estim Prob	Hypoth Prob
0	0.59004	0.50000
1	0.40996	0.50000

Binomial Test	Level Tested	Hypoth Prob (p1)	p-Value
Ha: Prob(p > p1)	1	0.50000	0.9986

Figure E5: Middle Quarters Sample Proportion Hypothesis Test including Flyaway Program Components



Frequencies

Level	Count	Prob
0	134	0.51341
1	127	0.48659
Total	261	1.00000

N Missing 0
2 Levels

Test Probabilities

Level	Estim Prob	Hypoth Prob
0	0.51341	0.50000
1	0.48659	0.50000

Binomial Test	Level Tested	Hypoth Prob (p1)	p-Value
Ha: Prob(p > p1)	1	0.50000	0.6897

Figure E6: Last Quarter Sample Proportion Hypothesis Test including Flyaway Program Components

Appendix F: Confusion Matrix for Boone's Learning Curve including Flyaway Components

Table F1: Confusion Matrix of Counts & Percentages including Flyaway Components

		Boone's Learning Curve More Accurately Explains	
		No	Yes
Observed Learning Curve Plateaus	No	99 42%	22 9%
	Yes	30 13%	83 35%

Appendix G: Test Results of Percentage Error Differences Between Boone's and Crawford's Learning Curves Error by Quarter

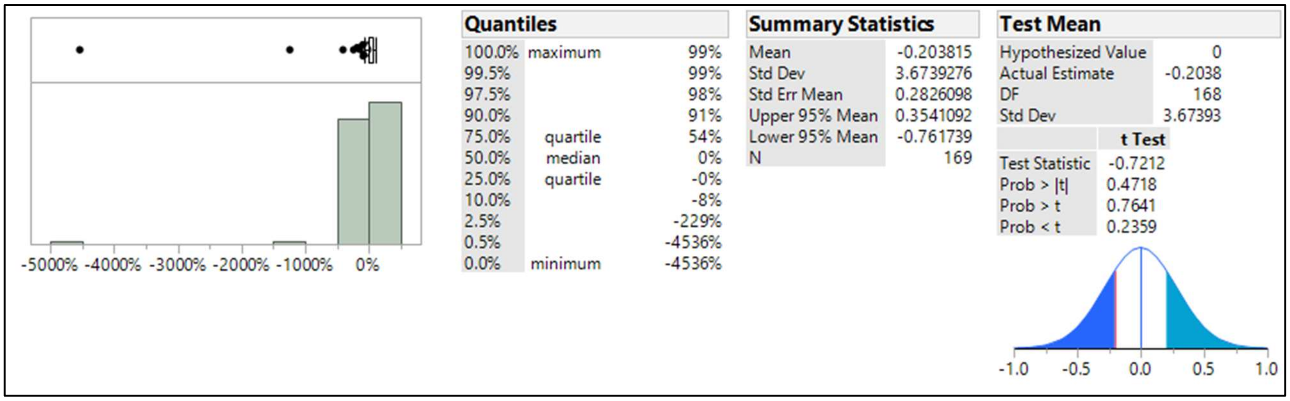


Figure G1: First Quarter Percentage Error Difference between Boone's & Crawford's Learning Curves Given Boone's Learning Curve Significantly Improved

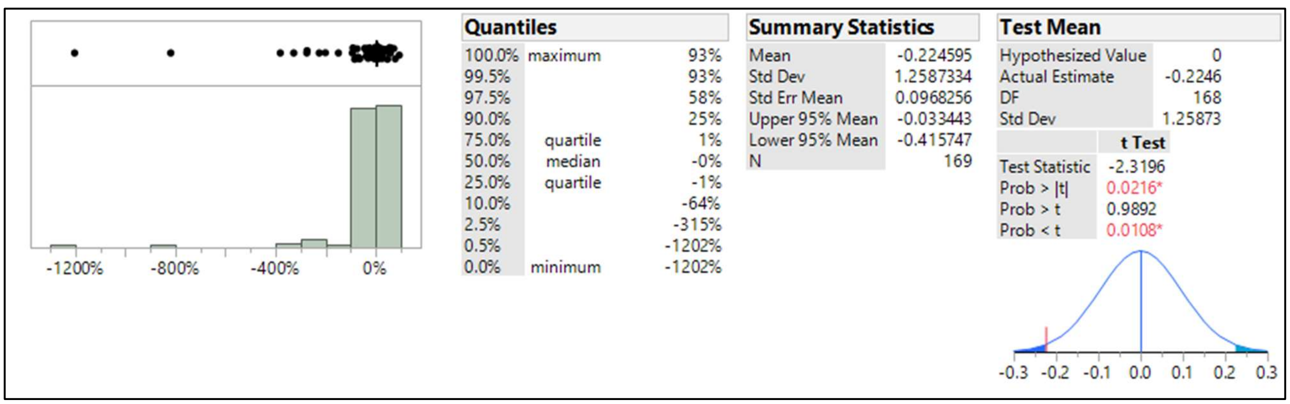
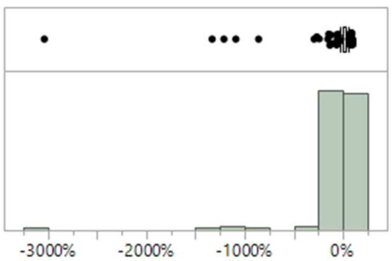


Figure G2: Middle Quarters Percentage Error Difference between Boone's & Crawford's Learning Curves Given Boone's Learning Curve Significantly Improved

Appendix G (continued): Test Results of Percentage Error Differences Between Boone's and Crawford's Learning Curve Error by Quarter



Quantiles

100.0%	maximum	98%
99.5%		98%
97.5%		94%
90.0%		77%
75.0%	quartile	22%
50.0%	median	-0%
25.0%	quartile	-0%
10.0%		-62%
2.5%		-1033%
0.5%		-3036%
0.0%	minimum	-3036%

Summary Statistics

Mean	-0.400501
Std Dev	2.95463
Std Err Mean	0.2272792
Upper 95% Mean	0.0481898
Lower 95% Mean	-0.849193
N	169

Test Mean

Hypothesized Value	0
Actual Estimate	-0.4005
DF	168
Std Dev	2.95463

t Test

Test Statistic	-1.7622
Prob > |t|	0.0799
Prob > t	0.9601
Prob < t	0.0399*

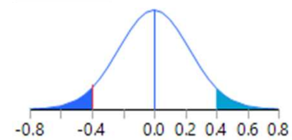
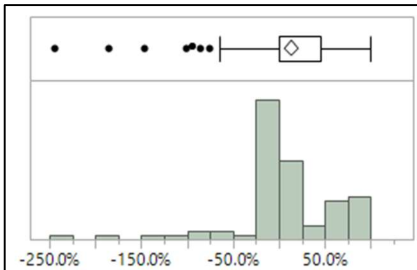


Figure G3: Last Quarter Percentage Error Difference between Boone's & Crawford's Learning Curves Given Boone's Learning Curve Significantly Improved



Quantiles

100.0%	maximum	99.0%
99.5%		99.0%
97.5%		98.3%
90.0%		83.6%
75.0%	quartile	45.7%
50.0%	median	0.0%
25.0%	quartile	-0.1%
10.0%		-6.2%
2.5%		-101.9%
0.5%		-244.0%
0.0%	minimum	-244.0%

Summary Statistics

Mean	0.1343194
Std Dev	0.4929238
Std Err Mean	0.0392149
Upper 95% Mean	0.2117762
Lower 95% Mean	0.0568625
N	158

Test Mean

Hypothesized Value	0
Actual Estimate	0.13432
DF	157
Std Dev	0.49292

t Test

Test Statistic	3.4252
Prob > |t|	0.0008*
Prob > t	0.0004*
Prob < t	0.9996

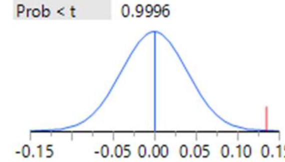
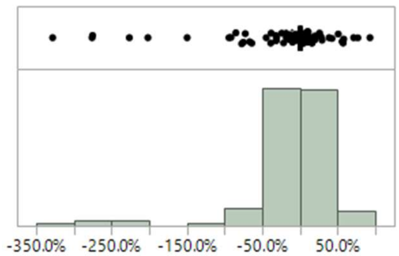


Figure G4: First Quarter Percentage Error Difference between Boone's & Crawford's Learning Curves Given Boone's Learning Curve Significantly Improved, Excluding Outliers

Appendix G (continued): Test Results of Percentage Error Differences Between Boone's and Crawford's Learning Curve Error by Quarter



Quantiles

100.0%	maximum	92.9%
99.5%		92.9%
97.5%		58.5%
90.0%		17.2%
75.0%	quartile	0.7%
50.0%	median	-0.0%
25.0%	quartile	-1.1%
10.0%		-39.3%
2.5%		-227.5%
0.5%		-328.1%
0.0%	minimum	-328.1%

Summary Statistics

Mean	-0.100972
Std Dev	0.540713
Std Err Mean	0.0430168
Upper 95% Mean	-0.016006
Lower 95% Mean	-0.185939
N	158

Test Mean

Hypothesized Value	0
Actual Estimate	-0.101
DF	157
Std Dev	0.54071

	t Test
Test Statistic	-2.3473
Prob > |t|	0.0202*
Prob > t	0.9899
Prob < t	0.0101*

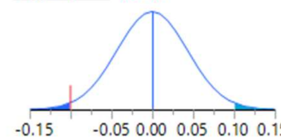
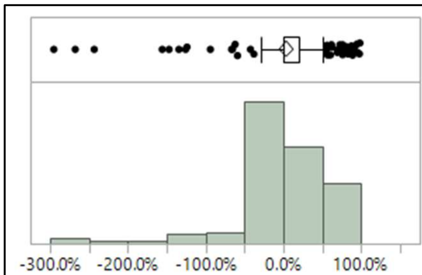


Figure G5: Middle Quarters Percentage Error Difference between Boone's & Crawford's Learning Curves Given Boone's Learning Curve Significantly Improved, Excluding Outliers



Quantiles

100.0%	maximum	98.3%
99.5%		98.3%
97.5%		94.7%
90.0%		76.9%
75.0%	quartile	20.4%
50.0%	median	-0.0%
25.0%	quartile	-0.2%
10.0%		-30.4%
2.5%		-158.0%
0.5%		-295.2%
0.0%	minimum	-295.2%

Summary Statistics

Mean	0.0360017
Std Dev	0.5817165
Std Err Mean	0.0462789
Upper 95% Mean	0.1274112
Lower 95% Mean	-0.055408
N	158

Test Mean

Hypothesized Value	0
Actual Estimate	0.036
DF	157
Std Dev	0.58172

	t Test
Test Statistic	0.7779
Prob > |t|	0.4378
Prob > t	0.2189
Prob < t	0.7811

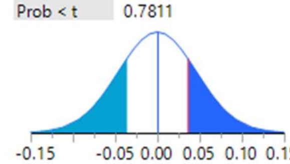


Figure G6: Last Quarter Percentage Error Difference between Boone's & Crawford's Learning Curves Given Boone's Learning Curve Significantly Improved, Excluding Outliers

Bibliography

1921-1 Data Item Description (pp. 1–15). (n.d.). Retrieved from https://cade.osd.mil/content/cade/files/csdr/dids/archive/1921-1.DI-FNCL-81566B.pdf

Anderlohr, G. (1969). Determining the cost of production breaks. *Management Review*, 58(12), 16-19.

Anzanello, M. J., & Fogliatto, F. S. (2011). Learning curve models and applications: Literature review and research directions. *International Journal of Industrial Ergonomics*, *41*(5), 573-583.

Argote, L. (1993). Group and organizational learning curves: Individual, system and environmental components. *British journal of social psychology*, *32*(1), 31-51.

Argote, L., Beckman, S. L., & Epple, D. (1990). The persistence and transfer of learning in industrial settings. *Management science*, *36*(2), 140-154.

Asher, H. (1956). *Cost-quantity relationships in the airframe industry* (Doctoral dissertation, The Ohio State University). Retrieved from https://etd.ohiolink.edu/!etd.send_file?accession=osu1485795295566357&disposition=inline

Badiru, A. B. (1992). Computational Survey of Univariate and Multivariate Learning Curve Models. *IEEE Transactions on Engineering Management*, *39*(2), 176–188. Retrieved from https://doi.org/10.1109/17.141275

Badiru, A. B. (2012). Half-life learning curves in the defense acquisition life cycle. *Defense ARJ*, *19*(3), 283–308. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA582714

Baloff, N. (1966). Startups in machine-intensive production systems. *Journal of Industrial Engineering*, *17*(1), 25.

Baloff, N. (1970). Startup management. *IEEE Transactions on Engineering Management, EM-17*(4), 132–141

Baye, M. R. (2010). *Managerial economics and business strategy* (7th ed.). New York: McGraw-Hill Irwin.

Boone, E. R. (2018). *An Analysis of Learning Curve Theory and the Flattening Effect at the End of the Production Cycle* (No. AFIT-ENV-MS-18-M-181). AIR FORCE INSTITUTE OF TECHNOLOGY WRIGHT-PATTERSON AFB OH WRIGHT-PATTERSON AFB United States.

Burke, R. P. Department of Defense. (2019). *Implementation of Cost and Hour Report (FlexFile) and Quantity Data Reports Within the Cost and Software Reporting (CSDR) System.* Washington, D.C.

Carr, G. W. (1946). Peacetime cost estimating requires new learning curves. *Aviation*, *45*(4), 220-228.

Chalmers, G., & DeCarteret, N. (1949). Relationship for determining the Optimum Expansibility of the Elements of a Peacetime Aircraft Procurement Program.

Cohen, J. (1938). Quantitative methods in psychology. *Nature*, *141*(3570), 613. https://doi.org/10.1038/141613a0

Corlett, E. N., & Morecombe, V. J. (1970). Straightening Out Learning Curves. *Personnel Management*, *2*(6), 14-19.

Crawford, J. R. (1944). Estimating, Budgeting and Scheduling. *Lockheed Aircraft Co., Burbank, California*.

Crossman, E. R. F. W. (1959). A theory of the acquisition of speed-skill. *Ergonomics*, *2*(2), 153-166.

Cullis, Bethia; Coleman, Richard; Braxton, Peter; McQueston, J. (2008). CUMAV or Unit? Is Cum Average vs. Unit Theory a Fair Fight? *SCEA Conference*, 1–17. ICEAA. *Published presentation by Lockheed Martin Corporation*.

Dar-El, E. M. (2013). *Human learning: From learning curves to learning organizations* (Vol. 29). Springer Science & Business Media.

Darr, E. D., Argote, L., & Epple, D. (1995). The acquisition, transfer, and depreciation of knowledge in service organizations: Productivity in franchises. *Management science*, *41*(11), 1750-1762.

David, G., & Brachet, T. (2011). On the determinants of organizational forgetting. *American Economic Journal: Microeconomics*, *3*(3), 100-123.

De Jong, J. R. (1957). The effects of increasing skill on cycle time and its consequences for time standards. *Ergonomics*, *1*(1), 51-60.

Department of Defense (2017). *Financial Management Regulation* (DoD 7000.14-R, Vol 2A). Washington, D.C.: Undersecretary of Defense (Comptroller). Retrieved from https://comptroller.defense.gov/FMR/fmrvolumes.aspx

Department of Defense (2018). *WORK BREAKDOWN STRUCTURES FOR DEFENSE MATERIEL ITEMS.* Washington D.C. U.S. Department of the Air Force.

Department of the Air Force (2007). *Air Force Cost Analysis Handbook.* Washington D.C. U.S. Department of the Air Force.

Dutton, J. M., & Thomas, A. (1984). Treating progress functions as a managerial opportunity. *Academy of management review*, *9*(2), 235-247.

Epple, D., Argote, L., & Devadas, R. (1991). Organizational learning curves: A method for investigating intra-plant transfer of knowledge acquired through learning by doing. *Organization Science*, *2*(1), 58-70.

Flyaway Costs. (n.d.). In *Defense Acquisition University* Glossary. Retrieved from https://www.dau.edu/glossary/Pages/Glossary.aspx

Glock, C. H., Grosse, E. H., Jaber, M. Y., & Smunt, T. L. (2019). Applications of learning curves in production and operations management: A systematic literature review. *Computers & Industrial Engineering*, *131*, 422-441.

Government Accountability Office (2009). *GAO Cost estimating and assessment guide: Best practices for developing and managing capital program costs* (Report No. GAO-09-3SP). Washington D.C. Government Accountability Office.

Government Accountability Office (2018). *Weapons System Annual Assessment: Knowledge Gaps Pose Risks to Sustaining Recent Positive Trends* (Report No. GAO-18-360SP). Washington D.C. Government Accountability Office.

Gujarati, D. N., & Porter, D. C. (2010). *Essentials of econometrics* (4th ed.). New York: McGraw-Hill/Irwin.

Hilmer, C. E., & Hilmer, M. J. (2014). *Practical Econometrics*. McGraw-Hill Education.

Hirschmann, W. B. (1964). Profit from the learning-curve. *Harvard Business Review*, *42*(1), 125-139.

Honious, C., Johnson, B., Elshaw, J., & Badiru, A. (2016). The Impact of Learning Curve Model Selection and Criteria for Cost Estimation Accuracy in the DoD (SYM-AM-16-075). *Acquisition Research Symposium: Vol. 2* (pp. 421-486). Monterey: Naval Postgraduate School.

Hu, S. P., & Smith, A. (2013). Accuracy matters: Selecting a lot-based cost improvement curve. *Journal of Cost Analysis and Parametrics*, *6*(1), 23-42.

International Cost Estimating and Analysis Association. (2013). *Cost estimating body of knowledge training, Unit III, Module 7. Learning curve analysis: How to account for cost improvement.*

Jaber, M. Y. (2006). Learning and forgetting models and their applications. *Handbook of industrial and systems engineering*, *30*(1), 30-127.

Jaber, M. Y., & Bonney, M. (1996). Production breaks and the learning curve: the forgetting phenomenon. *Applied mathematical modelling*, *2*(20), 162-169.

Knecht, G. R. (1974). Costing, technological growth and generalized learning curves. *Journal of the Operational Research Society*, *25*(3), 487-491.

Kronemer, A., & Henneberger, J. E. (1993). Productivity in aircraft manufacturing. *Monthly Lab. Rev.*, *116*, 24.

Large, J. P., Hoffmayer, K., & Kontrovich, F. (1974). *Production rate and production cost* (Vol. 1609, No. PA/E). RAND CORP SANTA MONICA CALIF.

Levy, F. K. (1965). Adaptation in the production process. *Management Science*, *11*(6), B-136.

Li, G., & Rajagopalan, S. (1998). A learning curve model with knowledge depreciation. *European Journal of Operational Research*, *105*(1), 143-154.

McClave, J. T., Benson, P. G., & Sincich, T. (2014). *Statistics for business and economics*. Boston: Pearson.

Mislick, G. K., & Nussbaum, D. A. (2015). *Cost estimation: methods and tools*. John Wiley & Sons.

Moore, J. R., Elshaw, J. J., Badiru, A. B., & Ritschel, J. D. (2015). Acquisition challenge: the importance of incompressibility in comparing learning curve models. *Defense ARJ, 22*(4), 416-449.

Nembhard, D. A., & Uzumeri, M. V. (2000). Experiential learning and forgetting for manual and cognitive tasks. *International journal of industrial ergonomics*, *25*(4), 315-326.

Sakinç, M. E. (2016). *Innovation or Financialization?: The Evolution of the Systems-Integration Business Model at Airbus and Boeing* (Doctoral dissertation, Bordeaux).

Sikström, S., & Jaber, M. Y. (2002). The power integration diffusion model for production breaks. *Journal of Experimental Psychology: Applied*, *8*(2), 118.

Sikström, S., & Jaber, M. Y. (2012). The Depletion–Power–Integration–Latency (DPIL) model of spaced and massed repetition. *Computers & Industrial Engineering*, *63*(1), 323-337.

Solver Technology - Global Optimization. (2016, September 12). Retrieved from https://www.solver.com/global-optimization-technology.

Solver Technology - Smooth Nonlinear Optimization. (2012, March 12). Retrieved from https://www.solver.com/smooth-nonlinear-technology#Generalized Reduced Gradient Method.

Vértesy, D. (2011). Interrupted innovation. *Emerging economies in the structure of the global aerospace industry. Maastricht University, Maastricht*, 309.

Vieira, P. A. G. (2013). *Current airframe manufacturing technologies in the aeronautical industry and trends for future developments* (Doctoral dissertation, Universidade da Beira Interior).

Wooldridge, J. M. (2016). *Introductory econometrics: a modern approach* (6th ed.). Boston: Cengage Learning.

Wright, T. P. (1936). Factors affecting the cost of airplanes. *Journal of the aeronautical sciences*, *3*(4), 122-128.

Yelle, L. E. (1979). The learning curve: Historical review and comprehensive survey. *Decision sciences*, *10*(2), 302-328.

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
26-03-2020	Master's Thesis	September 2018 – March 2020

4. TITLE AND SUBTITLE
An Analysis of Learning Curve Theory & Diminishing Rates of Learning

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Hogan, Dakotah, W., 1st Lt, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way, Building 640
WPAFB OH 45433-8865

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT-ENV-MS-20-M-212

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
1 University Circle, Monterey, CA 93943
831-656-3566
pao@nps.edu

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

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

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Statement A. Approved for Public Release; Distribution Unlimited.

13. SUPPLEMENTARY NOTES
This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT
Traditional learning curve theory assumes a constant learning rate regardless of the number of units produced; however, a collection of theoretical and empirical evidence indicates that learning rates decrease as more units are produced in some cases. These diminishing learning rates cause traditional learning curves to underestimate required resources, potentially resulting in cost overruns. A diminishing learning rate model, Boone's Learning Curve (2018), was recently developed to model this phenomenon. This research confirmed that Boone's Learning Curve is more accurate in modeling observed learning curves using production data of 169 Department of Defense end-items. However, further empirical analysis revealed deficiencies in the theoretical justifications of why and under what conditions Boone's Learning Curve more accurately models observations. This research also discovered that diminishing learning rates are present but not pervasive in the sampled observations. Additionally, this research explored the theoretical and empirical evidence that may cause learning curves to exhibit diminishing learning rates and be more accurately modeled by Boone's Learning Curve. Only a limited number of theory-based variables were useful in explaining these phenomena. This research further justifies the necessity of a diminishing learning rate model and proposes a framework to investigate learning curves that exhibit diminishing learning rates.

15. SUBJECT TERMS
Learning curves, diminishing rates of learning, forgetting curves, plateauing curves

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Dr. John Elshaw, AFIT/ENV
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	UU	161	(937)255-3636, ext 3363

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