




# ACQUISITION CHALLENGE: The Importance of INCOMPRESSIBILITY in Comparing Learning Curve Models

 *Capt Justin R. Moore, USAF, John J. Elshaw, Adedeji B. Badiru, and Lt Col Jonathan D. Ritschel, USAF*

---

The Department of Defense (DoD) cost estimating methodology currently employs T. P. Wright's 75-plus-year-old learning curve formula. The goal of this research was to examine alternative learning curve models and determine if a more reliable and valid cost estimation method exists, which could be incorporated within the DoD acquisition environment. This study tested three alternative learning models (the Stanford-B model, DeJong's learning formula, and the S-Curve model) to compare predicted against actual costs for the F-15 A-E jet fighter platform. The results indicate that the S-Curve and DeJong models offer improvement over current estimation techniques, but more importantly—and unexpectedly—highlight the importance of incompressibility (the amount of a process that is automated) in learning curve estimating.

---

**Keywords:** *cost estimation, Stanford-B, DeJong, S-Curve, Wright's Learning Curve, learning curve*

In 2008, the U.S. economy took a plunge that affected every industry from the real estate market to automobile manufacturers. This crash led to tightened budgets throughout the country, and many companies looked to operate more efficiently with less capital. That economic turmoil is reflected in the Department of Defense (DoD) through funding cuts and shrinking budgets at every level. The Budget Control Act of 2011, approved by Congress, places emphasis on commanders and managers using funds more efficiently.

On a micro level, the scrutiny of program cost estimates places more pressure on estimators than ever before. Due to the fact that sequestration cuts and their subsequent effects will continue seemingly over the next decade, cost estimators and the accuracy of acquisition cost estimates play a more important role than ever before in acquisition programs. Cost estimates are no longer just a box to check at milestone reviews; they now provide leverage for managers and valuable information in balancing budgets.

***Due to the fact that sequestration cuts and their subsequent effects will continue seemingly over the next decade, cost estimators and the accuracy of acquisition cost estimates play a more important role than ever before in acquisition programs.***

## Background

The Budget Control Act of 2011, which calls for a \$1.5 trillion deficit reduction over the next 10 years, has created a fiscally constrained environment in which competition for congressional funding is higher than ever before. On an organizational level, DoD acquisition programs have seen budget cuts up to 10 percent, changes in acquisition schedule, reduction in the number of systems purchased, and an increased scrutiny over cost estimates.

One way to assist cost estimators, and consequently decision makers, is to provide them with the most current and appropriate tools to calculate accurate and reliable predictions. However, conventional learning curve methodology has been in practice since the pre-World War II build-up in the 1930s, and those historical techniques may be outdated in today's fast-paced, technological environment.

Over the past two decades a new methodology, rooted in the concept of forgetting curves, has emerged and may provide a more accurate tool for assessing learning curves. Forgetting is becoming more widely accepted, but its application to learning curves in manufacturing is scarce. This research will incorporate contemporary learning curve models to cost estimates within large DoD acquisition programs.

The concept of learning and the application of learning curves are widely used in everything from industrial manufacturing to avionics software development. The footprint of the learning phenomenon applies throughout both public and private business sectors. In recent years, the concept of forgetting has been introduced, which unlike Wright's (1936) model, does not assume a constant learning rate. Learning curves are widely used and even expected throughout the DoD cost estimating community. Air Force guidance on learning curve theory and application primarily originates from the *Air Force Cost Analysis Handbook* (AFCAH, 2008), Chapter 8. This resource primarily focuses on two learning curve theories: unit theory and cumulative average theory. This research does not intend to discredit the use of learning curves, but rather incorporates and assesses contemporary methodology within the confines of major acquisition programs.

## Theory Review

Learning curve models came into use by manufacturing practitioners in the late 1930s. At the height of the pre-World War II build-up, aircraft production costs were as important as developing and producing the aircraft themselves. T. P. Wright (1936) first identified the existence of the learning relationship. He correctly theorized that as a worker performs the same task multiple times, the time required to complete that task will decrease at a constant rate. The workers are learning from previous experience and thus becoming more efficient in completing the task. Wright also identified the *80 percent learning effect* in aircraft production. He believed that organizations would observe a learning rate of 80, or a 20 percent production

improvement, as the number of units produced doubled (Wright 1936). This rule has been changed and modified over time to fit different applications; however, it remains the standard in many industries.

While a vast collection of theory and studies exists relating to learning curves, very little attention has been given to the performance degradation due to the impact of forgetting (Badiru, Elshaw, & Everly, 2013). We define forgetting as the process of unlearning and the loss of knowledge, particularly through the passage of time. Forgetting is simply the concept that workers will inevitably see a decline in performance (from many potential sources) while still theoretically moving along the learning curve (Badiru, 1995). The incorporation of forgetting is a critical piece of learning curve theory because it helps explain variance in the process that otherwise may be unaccounted for.

***Forgetting is simply the concept that workers will inevitably see a decline in performance (from many potential sources) while still theoretically moving along the learning curve (Badiru, 1995).***

The classical learning curve model, often referred to as Wright's Learning Model, gives mathematical representations of Wright's basic learning theory. The model shown in Equation (1) follows the assumption that as the quantity produced doubles, the cost will decrease at a constant rate.

$$T_x = T_1 x^b \quad (1)$$

Where:

$T_x$  = the cumulative average time (or related cost) after producing  $x$  units

$T_1$  = hours required to produce (theoretical) first unit

$x$  = cumulative unit number

$b = \log R / \log 2 = \text{learning index}$

*Note: R in the term above = learning rate (a decimal)*

J. R. Crawford (1944) adopted a similar learning curve approach in the *individual unit* model that he introduced in a training manual at Lockheed Martin. Crawford's model uses the same basic formula as Wright's model, but attempts to estimate *individual* times (or related cost) to produce a given unit by changing which variables are input into the model.

Both unit theory and cumulative average approaches are used in acquisition cost estimating, depending on the amount and validity of historical program data. However, contractor reports often come in the form of lots. This form of data is usually more advantageous when using a cumulative average learning curve. The DoD Basis of Cost Estimating illustrates how such data can be used as a lot average in the cumulative average learning curve theory rather than finding a theoretical lot midpoint as with the unit theory (DoD, 2007).

[A]pply the Cum Avg formulation to contractor lot information, add the hours/costs for a given lot to the hours/costs of all previous lots. The hour/cost plot value (Y axis) of a given lot is the total hours/costs through that lot divided by the last unit number of that lot, while the unit plot point (X axis) is the last unit number of that lot. Lot midpoints are not used with the Cum Avg formulation. (p. 8-21)

Furthermore, Hu and Smith (2013) identify a method for plotting and predicting learning curves using lot data, "If the cumulative average costs for all consecutive lots are present, then the direct approach can be applied to the lot data with the last unit in the lot as the lot plot point (LPP)" (p. 28). This LPP is the same as the unit plot point described in the AFCAH and provides a means for plotting lot data against individual units (on the X axis) to determine the learning parameters. Hu and Smith describe this process saying, "T1, b, and other exponents can be obtained directly from the ordinary least squares (OLS) method by regressing [cumulative average costs] vs. cumulative quantities" (p. 28).

Since Wright's initial theory, several other models have been adopted in learning curve literature. One of the earliest modifications to the learning curve model came along with introduction of the Stanford-B model shown in Equation (2).

$$T_i = T_1 (x + B)^{-b} \quad (2)$$

Where:

$T_i$  = the cumulative average time (or related cost) after producing  $x$  units

$T_1$  = hours required to produce (theoretical) first unit

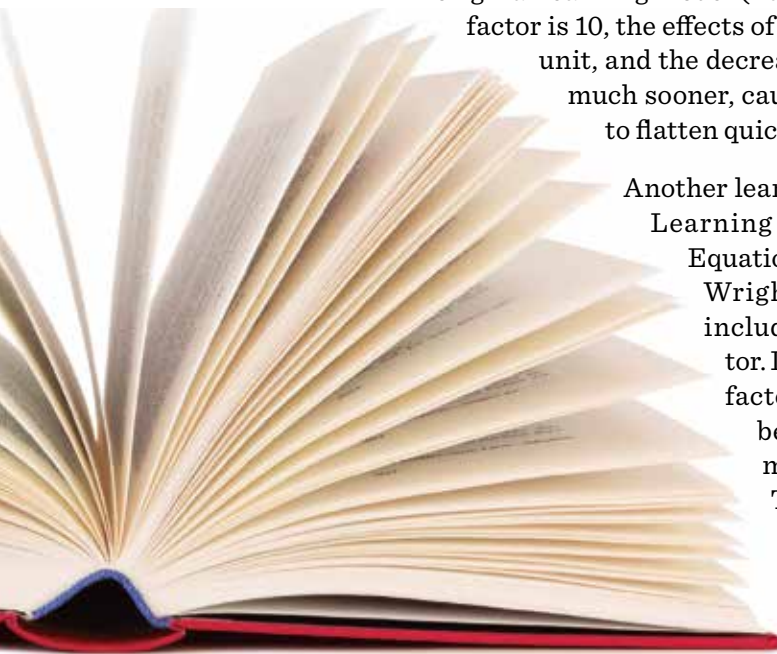
$x$  = cumulative unit number

$b = \log R / \log 2$  = learning index

$B$  = equivalent experience units (a constant); slope of the asymptote of the curve.

This model is attributed to Louis E. Yelle (1979) during a government-funded research initiative at Stanford. It introduces the equivalent experience unit parameter to Wright's original equation. This parameter, represented by  $B$ , is a constant from 0 to 10, accounting for the number of units produced prior to start of production of the first unit, and is the slope of the asymptote of the learning curve. If this factor is 0, the model reverts to Wright's original learning model (Badiru, 2012). Conversely, if the factor is 10, the effects of learning will begin at the 11th unit, and the decrease in performance will occur much sooner, causing the learning curve slope to flatten quickly.

Another learning curve model is DeJong's Learning Formula. DeJong's model in Equation (3) is also a derivation from Wright's original function, which includes an incompressibility factor. Denoted by the constant  $M$ , this factor represents the relationship between manual processes and machine-dominated processes. The incompressibility factor is



a constant between 0 and 1, in which a value of 0 implies a fully manual operation and a value of 1 denotes a completely machine-dominated operation (Badiru et al., 2013).

$$T_x = T_1 [M + (1 - M)x^{-b}] \quad (3)$$

Where:

$T_i$  = the cumulative average time (or related cost) after producing  $x$  units

$T_1$  = hours required to produce (theoretical) first unit

$x$  = cumulative unit number

$b = \log R / \log 2 =$  learning index

$M =$  incompressibility factor (a constant)

Wright's original model, which inherently assumes an incompressibility factor of 0, fails to account for a major percentage of the production industry that uses automated manufacturing technology.

The S-Curve model accounts for both the prior experience and incompressibility factors together. Carr (1946) believed that there was an error in Wright's constant learning assumption and hypothesized that the effects of learning and thus performance followed the S-Curve shape. The S-Curve model assumes a gradual build-up in the early stages of production followed by a period of peak performance. This build-up is typically attributed to personnel and procedural changes as well as time needed for new machinery set-ups that occur early in the production process. Towill and Cherrington (1994) used the theory hypothesized by Carr to develop a model that follows an S-shaped pattern. The S-Curve model shown in Equation (4) assumes that learning takes the S-shaped curve often seen in a cumulative normal distribution.

$$T_x = T_1 + M(x + B)^{-b} \quad (4)$$

Where:

$T_i$  = the cumulative average time (or related cost) after producing  $x$  units



$T_1$  = hours required to produce (theoretical) first unit

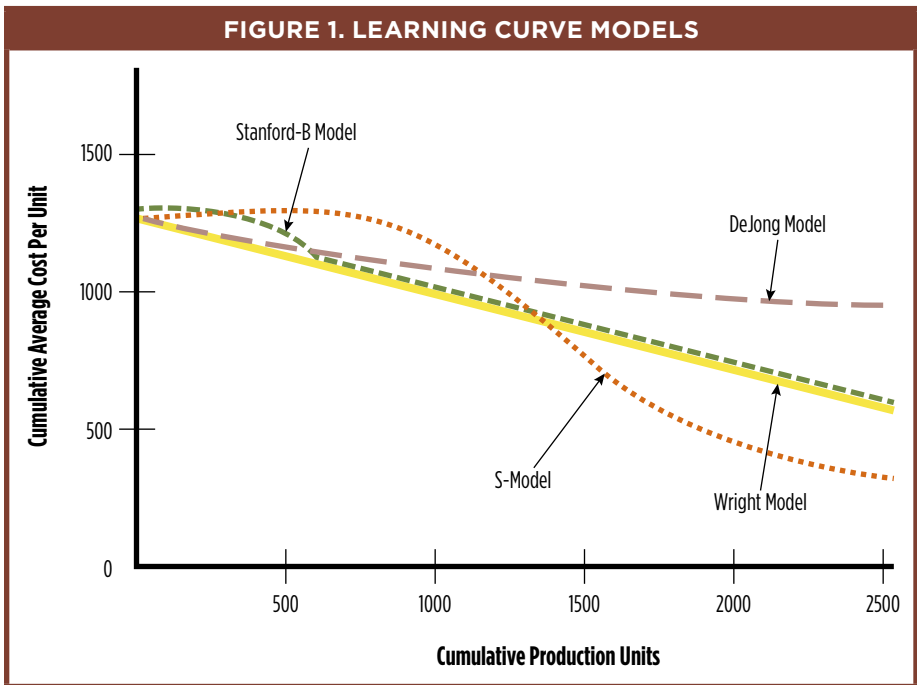
$x$  = cumulative unit number

$b = \log R / \log 2 =$  learning index

$M =$  incompressibility factor (a constant)

$B =$  equivalent experience units (a constant)

Figure 1 contains a graphical comparison of these three models. These models have specific, easily identifiable parameters that are more conducive for cost estimators to put to practical use. The goal is to make the estimator's calculations more reliable and avoid a series of equations that decision makers must interpret.



Note. Adapted from Badiru, 1992.

## Hypotheses Development

### Wright's Learning Curve

The status quo for the learning curve models is Wright's Learning Curve (WLC) model, which takes the form  $T_x = T_1 x^{-b}$ . The two parameters that must be determined to perform an estimate are  $T_1$  and  $b$ . In common cost estimating practices,  $b$  and  $T_1$  are determined through a linear regression on a plot of the natural log of cumulative unit number  $[\ln(x)]$  against the natural log of the actual reported costs  $[\ln(y)]$ . This regression will determine whether the cumulative average or unit learning curve theory should be applied to the data. The regression providing the most accurate fit according to the  $R^2$  value will determine whether unit theory or cumulative average theory will be used for the remainder of the study. Once a theory is selected, the corresponding regression equation will be used to determine the parameters of the model.  $R^2$  is a simple goodness-of-fit measure that represents the amount of variance between the independent and dependent variables expressed as a percentage. In other words, it represents the amount of variability that can be explained by the model (McClave, Benson, & Sincich, 2011). From the linear regression,  $b$  is simply the slope of the line and  $T_1$  is derived by taking the natural log of the y-intercept. Once these two parameters are determined for Wright's model, they remain constant for the other three models used in this analysis.

### Stanford-B Model

The first model selected for comparison was the Stanford-B model. The Stanford-B model is a relatively older application of the learning curve using the equation  $T_i = T_1 (x + B)^{-b}$ . The point of interest where this model differs from Wright's is the equivalent experience unit constant represented by the constant  $B$ . The  $B$  constant falls between 0 and 10 and represents the equivalent units of previous experience at the start of the production process. If more than 10 units have been produced, then the constant remains at 10. This parameter accounts for how many times the process has already been completed and adjusts the learning curve based on that number. The Stanford-B model is only a slight derivation from Wright's traditional learning curve model, and when  $B$  is equal to the first unit produced, then the models are identical (Badiru et al., 2013). Properly applying previous experience into the model is the key to using this equation, and for this study  $B$  is represented by the number of previous units produced. This can be in the form of prototypes, test aircraft, or any other relevant production unit that was not part of the F-15 A/B production lines. Twenty test units were produced beginning in 1970, which will be counted for prior experience, and therefore the factor  $B$  will be 10. This prior experience unit constant

of 10 will remain consistent when used in the S-Curve model described in the following section. With  $B$  determined, the data are incorporated into the model to estimate the total lot costs for the 15 remaining F-15 C/D and E lots. The residuals from these estimates, when compared to the actual lot costs, are then compared to each of the other three models to determine if one is a better fit than the others.

### DeJong's Model

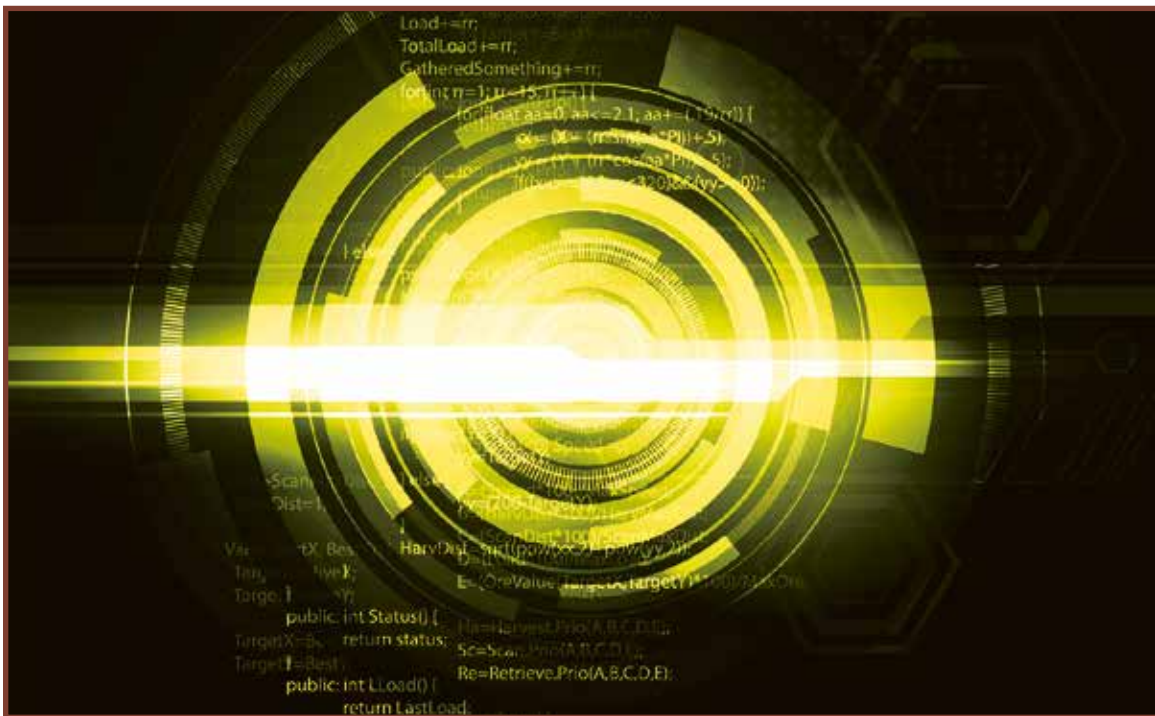
The second model used for comparison was the DeJong Learning Formula. DeJong's model is essentially a simple power function, similar to Wright's model, which accounts for the percentage of the task that requires mechanical activity to the amount that is touch labor. The effects of learning are typically only seen in touch, or human, labor because oftentimes, very few improvements in machine efficiency are observed over time. The basic form of this learning curve is  $T_i = T_1 + Mx^{-b}$ . Unlike previous models, DeJong's model incorporates the incompressibility factor ( $M$ ); however, there is no equivalent experience constant. The incompressibility factor,  $M$ , is a constant between 0 and 1 where 0 represents a fully manual process and 1 represents a machine-dominated process (Badiru et al., 2013). Aircraft production falls somewhere between 0 and 1, but there is no precedent set for application to aircraft production. A U.S. Bureau of Labor Statistics report from June 1993 gives the following description of the industry: "[A]lthough the industry assembles a high-tech product, its assembly process is fairly labor-intensive, with relatively little reliance on high-tech production techniques" (Kronemer & Henneberger, 1993). This report indicates that the highly specialized process of aircraft production, similar to that of high-end performance automobiles, supports a proper application of  $M$  closer to 0 than 1. Where exactly that number falls is undefined and leads to some subjectivity. To avoid any biases that may skew the results and apply robustness to the analysis, the application of the constant will start at 0.0 and move to 0.2 in increments of 0.05, resulting in five sets of analyses. This range of incompressibility factors will remain consistent in the application of the S-Curve model as well.

### S-Curve Model

The third and final model used for comparison in this study is the S-Curve model, which was developed by Towill and Cherrington in 1994. The S-Curve model is a combination of the Stanford-B model and DeJong's model. As mentioned earlier, this model is based on the assumption of gradual build-up early on in the production process (a period of steady learning), and then a flattened portion at the top of the S-Curve called the slope of diminishing returns, which is often attributed to forgetting. The basic

S-Curve model,  $T_i = T_1 + M(x + B)^{-b}$ , uses the same previous experience unit constant,  $B$ , and incompressibility factor,  $M$ , as the Stanford-B and DeJong models, respectively. Three of the four variables on the right side of the equation ( $T_1$ ,  $b$ ,  $M$  and  $B$ ) must be known to make an assumption about the fourth (Badiru et al., 2013). In this study, we will use the same known  $T_1$ ,  $b$ , and  $B$  used in the prior equations to make an educated assumption about  $M$  as described in the DeJong model discussed earlier. The S-Curve model is a very strong representation of how forgetting will affect the rate of learning and is a sound model to use in testing the theory.

Towill and Cherrington (1994) identify three primary sources for estimating error, the first being errors due to inevitable fluctuations in performance that occur naturally. Estimators have little if any control over this source. The second is psychological, physiological, or environmental causes that affect deterministic errors. These can be accounted for by estimators, but again this lies largely outside of their control. The final source for prediction error is modelling error, meaning that the form of the model used may be inappropriate and therefore not fit the trend line of the data. This research will address the third issue and attempt to determine the most appropriate model form that fits defense aircraft over a production life.



The premise for this study is that at least one of the alternative learning curve models is a more accurate predictor of actual production costs than traditional learning models. This theory is founded on the belief that forgetting occurs in airframe production, and models that do not assume a constant rate of learning will provide a more accurate estimate. The research hypothesis for this theory is that there is a significant difference between the Mean Average Percent Error (MAPE) of the predicted lot costs between the four models. MAPE is a measure of variation that takes the average of the absolute values from the error of each prediction. The absolute value is taken to avoid any cancelling out of positive and negative error values. The smaller the MAPE, the more accurate and reliable the estimates.

Addressing the issue identified by Towill and Cherrington (1994) led to the necessity for this line of research. This study will compare three modern learning curve models (Stanford-B, DeJong, and S-Curve) to Wright's learning curve and attempt to determine if one is more accurate than the others. The previous discussion leads to the following hypotheses:

H1: One or more of the four models compared will have a MAPE significantly different from the others.

H2: One or more of the modern learning curve models will be significantly more accurate than Wright's learning model in predicting aircraft costs.

H3: The S-Curve model will have the lowest MAPE and prove to be the most accurate predictor of aircraft costs over time.

The null hypothesis ( $H_0$ ) for the first hypothesis in this study is that  $\mu_1 = \mu_2 = \mu_3 = \mu_4$ , meaning all of the MAPEs are the same, as contrasted against the alternative hypothesis ( $H_a$ ) that at least one of the models has a mean that is different. If the null hypothesis can be rejected and the evidence supports a significant difference, then it will be necessary to test each of the new learning models against the conventional model. The second null hypothesis mathematically states that  $\mu_i = \mu_1$  where  $i = 2, 3, 4$  to be tested against the  $H_a: \mu_i > \mu_1$ . These individual hypotheses test whether each of the modern learning curve models has a MAPE significantly lower than the conventional model. One final test will be to investigate the third hypothesis and determine which of these models that has displayed significantly smaller mean errors from the conventional model is the best predictor. The third null hypothesis states that  $\mu_i = \mu_j$ , where  $i$  and  $j$  are both significantly

lower than  $\mu_1$ , to be tested against the  $H_a: \mu_1 > \mu_1$ . That analysis will provide an answer to the initial inquiry of this research of determining if an alternative *best fit* model is more accurate than Wright's model.

## Methods

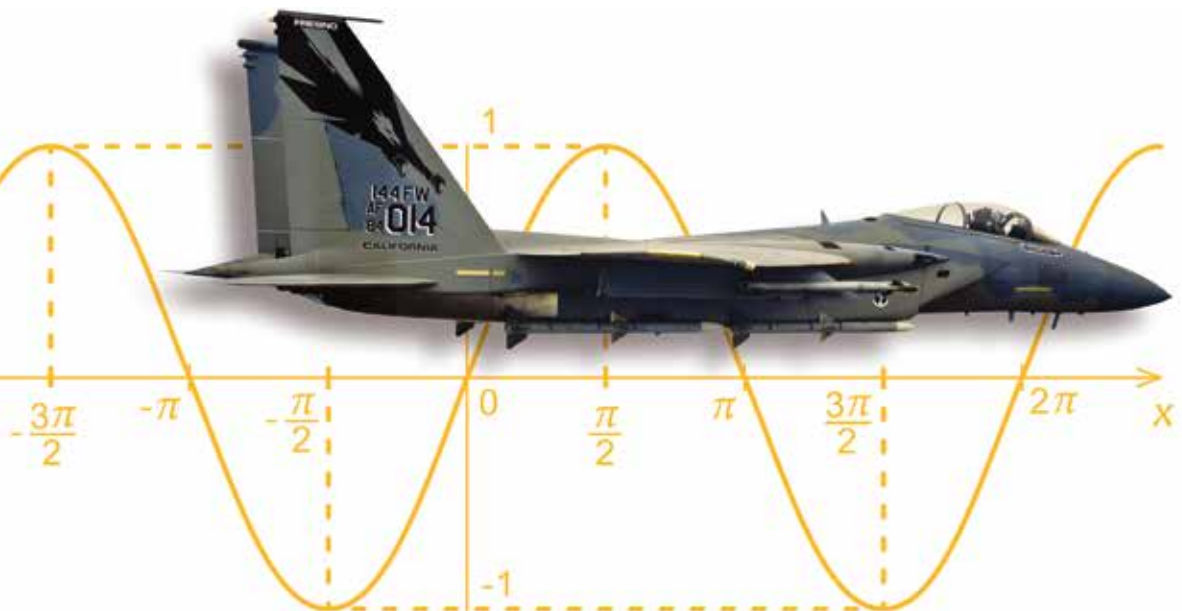
The initial task is to determine which of the models should be used in comparison to conventional learning curves, and how to improve upon conventional learning curve application. Several learning and forgetting curve models were identified for application in this study, but the three models selected are based on a literature review and subject matter expert (SME) opinion from cost analysts. These SMEs confirmed the Stanford-B model, DeJong's Learning Formula, and the S-Curve model are applicable to cost estimation and should be examined in the DoD environment. Additionally, they agreed the conventional Wright's model lacks the application of key factors such as prior experience and incompressibility that affect learning. Accounting for these previously unrecognized factors may reduce the amount of estimating error for airframe costs. In the DoD environment, an error reduction of a modest 5 percent could greatly enhance our ability to understand the cost overruns over the life of a program. The three models discussed in this article account for one or more forgetting factors, which can be easily assessed by cost estimators and quickly incorporated into current estimation techniques. The applicability and ease of use are other primary factors behind the selection of the models reviewed in this study. Providing a model that takes hours or days of secondary analysis and data collection is of little practical value to estimators, even if it proves more accurate. The following section explains how those models will be applied to the data in this study, which methods will be used to compare them, and how the data are analyzed in this research.

***In the DoD environment, an error reduction of a modest 5 percent could greatly enhance our ability to understand the cost overruns over the life of a program.***

## Data

Airframe costs were chosen for this analysis for a number of reasons. First, using airframe costs allows for the assumption of homogeneity over multiple model types. One can safely assume that the F-15 A/B, C/D, and E all have similar if not identical airframes, making it easier to compare the costs and examine the learning process. Also, in Foreign Military Sales (FMS) to the allies of the United States, the airframe of the aircraft typically does not change despite changes to avionics or electronics systems. Also, Badiru et al. (2013) state, “as rapid emergence of new technology necessitates that airframe designs and manufacturing processes be upgraded frequently... the opportunity for forgetting clearly increases.” Therefore, the application of airframe costs to this study will provide results consistent with that theory.

After some initial investigation, fighter aircraft became the primary platform type for this analysis for a multitude of reasons, the first reason being that several years of production data exist and hundreds of units were produced for these aircraft. Note that over 1,150 aircraft were produced in a 20-year span for the F-15 alone. Bailey (1989) stated that forgetting is a function of both the amount of learning and the passage of time. This makes the analysis of aircraft production cycles spanning over several years a prime candidate to exhibit the declining performance rate attributed to forgetting. The second reason is that the Air Force has several models of fighters (F-15



A-E and F-18 A-F, to name a few) in its inventory—all of which are variants of the same basic airframe, making the assumption for comparison of airframe costs from model to model possible. The final reason for choosing fighters was the ability to work face to face with cost estimators from the program offices located at Wright-Patterson Air Force Base, Ohio. Their assistance as SMEs would prove invaluable in verifying our assumptions and verifying the parameter estimates for our models.

The initial pool of aircraft data collected for analysis consisted of five fighters: the Air Force F-15, F-16, and F-22; the Navy F/A-18; and the joint (Air Force, Navy, and Marines) F-35. We eliminated the F-35 from analysis due to too few data points available. The F-22 was eliminated from consideration because it had two primary contractors: Lockheed Martin Aeronautics and Boeing Defense, Space, and Security. These two contractors both contributed components to the airframe production, making it difficult to measure and assess the effects of learning since production processes were not consistent between the two companies. For this reason, it does not provide a suitable comparison to other aircraft being tested. The F-16 was a prime candidate for analysis given the long production life and model upgrade, but relevant airframe data were incomplete or missing altogether in some cases. The F/A-18 had sufficient available data, but the program switched primary contractors, making it difficult to homogeneously compare the costs over that transition. This left the F-15 as the primary platform for analysis based on production history and availability of relevant airframe costs.

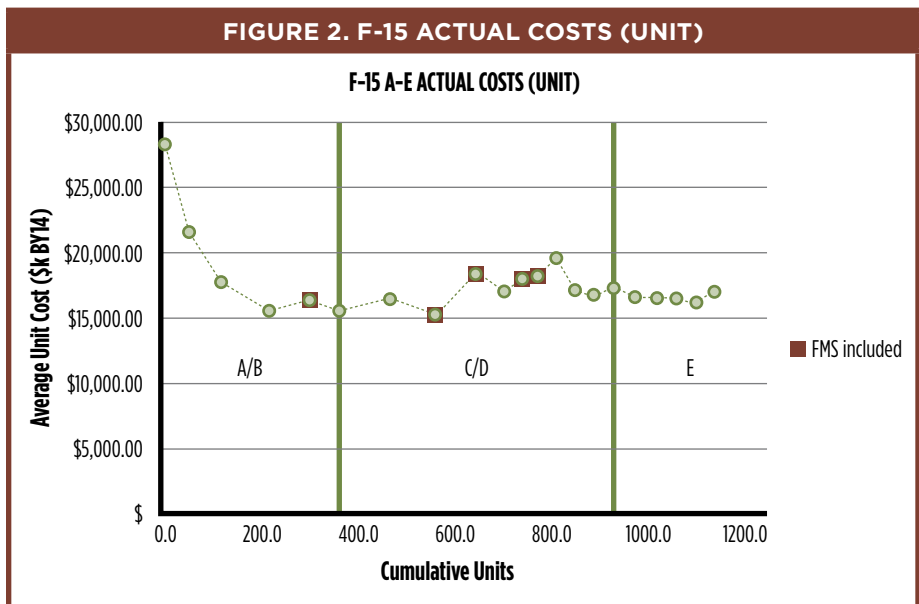
F-15 airframe costs were acquired from two databases. The F-15 A-D airframe lot averages were acquired from the *Cost Estimating System, Aircraft Cost Handbook*, published in 1987 by the Delta Research Corporation. This handbook includes all 19 lot purchases from 1970–1985 and details the quantity produced as well as the total airframe costs (minus administrative costs). These data were presented in Base Year 1987 dollars (BY\$87), meaning that the values for each year are set at a fixed price as if all of the funds were expended in 1987 (DoD, 2007). Summarized, this statement means that each of the values was initially represented at its equivalent purchasing power in the year 1987.

The F-15E data were taken directly from the Joint Cost Analysis Research Database (JCARD) system. These data were much more detailed and included five of the six lot purchases, with Lot 1 data missing. The system had data broken out into each cost element (including airframe) and the total quantity produced. The JCARD data were in Then Year dollars (TY\$), which are BY\$ inflated/deflated to represent the purchasing power of the funds if



they were expended in that given year (DoD, 2007). Both the F-15 A-D BY\$87 values and the F-15E TY\$ values are standardized in this research to a Base Year 2014 (BY\$14) value using the 2014 Office of the Secretary of Defense (OSD) Inflation Tables. The OSD Inflation Tables are published every year, and this research was begun in 2014 so those tables have been used to avoid crossing over to and from inflation tables. This step ensures that all dollar amounts are compared on a level plane and also represent a dollar value that is relevant to today’s economy.

The unit theory data of the entire F-15 A-E data set are shown in Figure 2. The data indicate that the later stages of the production cycle show possible signs of forgetting. The average unit cost is actually increasing towards the end of production rather than decreasing as would be predicted by Wright’s learning theory. The F-15 data appear to show significant signs of declining performance over the program’s life cycle in the sharp flattening trend in the data. After the production of around 600 units, the effects of learning nearly come to a complete stop and, in some cases, the costs actually increase over time.



Note. Average Unit Cost reflects actual cost per unit to the government for airframe only.

The goal of this study is to identify a model, or models, which more accurately predict the decline in performance over time and provide more accurate estimates for airframe costs than Wright’s contemporary model.

For this research, the F-15 A/B lots will be treated as historical data, and each of the models will be used to estimate the costs for the C/D and E lots based on that data. This scenario allows for the simulation of a *real-world* cost estimating scenario rather than a controlled study where the data are treated in a way that is beneficial to the researcher.

### Analysis Methods

Once the data are standardized to BY\$14 averages, the estimates from each of the models will be recorded using one of the four models described. There will also be data collected for cumulative units and lot number. An error term is calculated, which is the difference between the actual and predicted (Unit or Cumulative Average Theory) values. Absolute error (Abs Error) is simply the absolute value of the error, and absolute percent error (Abs PE) is the absolute error divided by the actual cost.

Once the data are coded, the next step is to perform the analysis and test the hypotheses. For the overall research hypothesis  $\mu_1 = \mu_2 = \mu_3 = \mu_4$ , the set of percent errors will be compared using an analysis of variance (ANOVA) method, as well as the Kruskal-Wallis (KW) test. These tests produce an *F*-statistic falling within a Chi-distribution and a resultant *p*-value that will either support or not support the null hypothesis based on the given confidence level. The null hypothesis is that all of the sample means are the same while the alternative hypothesis is that at least one of the sample means is different. The KW test is used to determine whether multiple samples arise from the same distribution and have the same parameters (Kruskal & Wallis, 1952). An *F*-test from the initial ANOVA and KW test, both performed in SPSS Statistics software, will provide insight into the first hypothesis. If the *F*-statistic is significant, then at least one of the sample means is different.

To test the second hypothesis (that at least one of the models is more accurate), this research will use Dunnett's test performed in SPSS. Dunnett's test is used to compare multiple sample means to one value held as the control (Everitt & Skrondal, 2010). Wright's learning curve model, the status quo, will be used as the control for this study, and the significance will be used to test if any of the other models' MAPE values are less than (<) the control. If the assumption for equal variance is not met, Dunnett's T3 test will be used for comparing the sample means. The T3 is similar to Dunnett's test described earlier, but it uses each sample as a control individually to compare against the other values.

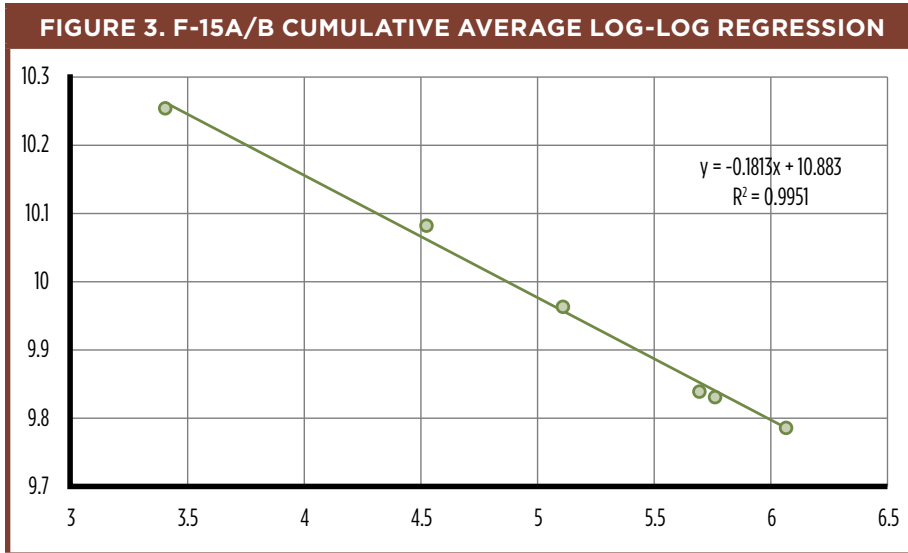
The final test will be to analyze which model is most accurate given significant results from previous tests. This analysis will be conducted through a simple paired difference  $t$  test—again performed in SPSS. A paired difference experiment uses a probability distribution when comparing two sample means and produces a  $t$  statistic that falls within a student  $t$  distribution that can either reject or fail to reject the null hypothesis, depending on the desired confidence level (McClave et al., 2011). If the assumption for equal variances is not met and the T3 test is used, information regarding which models are significantly different will be found in the T3 test, and there will be no need for paired  $t$  tests.

For this analysis, an  $F$ -statistic (or  $t$ -statistic) with a resulting  $p$  value  $< 0.05$  will support rejection of the null hypotheses and support the alternative hypothesis that the mean values between the models are different. A  $p$  value, or observed significance level (McClave et al., 2011), is defined as: “the probability (assuming  $H_0$  is true) of observing a value of the test statistic that is at least as contradictory to the null hypothesis, and supportive of the alternative hypothesis, as the actual one computed from the sample data.”

In other words, the  $p$  value is the chance of having an actual result that is contradictory to the sample result. By rejecting the null hypothesis, the data are essentially demonstrating a 95 percent chance that the means of the two populations are different.

### **F-15 C-E Analysis**

**Unit Theory and Cumulative Average Theory.** The first step of the analysis was to identify which learning theory was most appropriate for the given data. For the F-15 data using an  $M$  value of 0.20, a log-log regression was run against the A/B model data, using both the unit theory and cumulative average theory to predict the learning parameters for the C/D and E models used in the analysis. Figure 3 shows the regression using the cumulative average theory, which produced an  $R^2$  value of 0.9951. The cumulative average  $R^2$  value for the A/B model was slightly higher than the 0.9735 value produced using the unit theory data. This indicates that the cumulative average theory should be used for estimating the C-E model costs, and the lot-plot point assumption holds for the data.



These results also provide the basic parameters for all four learning models used in the study. The learning rate factor,  $b$ , is the slope of the linear regression line, which in this case is  $-0.1813$ . This value indicates a learning curve slope of 88.19 percent ( $LCS=2^b$ ). Figure 3 also provides information about the  $T_1$  value that is used in the analysis. The intercept of the linear regression equation is the natural log of the theoretical unit 1,  $T_1$ , value. By raising the mathematical constant  $e$  to the value of the intercept (10.883), one can determine the average cost of the theoretical first unit; in this case, that value is \$53,263.

**Assumption Parameters.** The next step was to populate the data tables so that the comparative analysis could be performed. Table 1 shows the Absolute Percent Error (APE) values for all 15 lots calculated using each of the four learning models with an incompressibility factor of 0.1. As the table shows, Wright's Curve and the Stanford-B models initially have the lowest MAPE of the four models, but analysis must be conducted to determine whether the data reflect a significant difference. That analysis can then be applied to a range of incompressibility factors to determine how sensitive the results are to a change in that factor.

TABLE 1. F-15 APE VALUES FOR EACH MODEL

M = 0.1				
Lot	WLC	Stanford-B	DeJong	S-Curve
7	0.0549032	0.0509017	0.2716447	0.2680433
8	0.0927225	0.0892703	0.3285742	0.3254672
9	0.1085792	0.1085792	0.0904993	0.0882712
10	0.0530433	0.0554482	0.1634820	0.1613176
11	0.1172022	0.1193309	0.0873964	0.0854805
12	0.1272667	0.2192897	0.0771023	0.0752816
13	0.1958247	0.1975958	0.0049876	0.0065815
14	0.0816980	0.0836323	0.1387508	0.1370100
15	0.0764948	0.0783588	0.1476580	0.1459804
16	0.1119286	0.1136465	0.1059919	0.1044458
17	0.0813009	0.0829968	0.1468597	0.1453335
18	0.0823053	0.0839250	0.1482298	0.1467721
19	0.0880680	0.0896143	0.1433682	0.1419766
20	0.0824747	0.0839757	0.1525089	0.1511580
21	0.1269814	0.1283646	0.0984203	0.0971754
<b>AVG</b>	<b>0.0987196</b>	<b>0.0996620</b>	<b>0.1403659</b>	<b>0.1386863</b>

Note. WLC = Wright's Learning Curve.

To analyze the samples, certain assumptions must be tested. The assumption of normality was not met, meaning that nonparametric tests must be used for comparing the means. Kurtosis is a measure of the *peakedness* of the distribution, and the high kurtosis values from the data set imply the data are non-normal and result in a sharply peaked distribution. All of the samples also have a skewness greater than 1, so normality cannot be assumed. The KW test must be used to determine whether the sample distributions are significantly different and if at least one sample has a median different from the others.

The tests for equal variances were not uniform through the range of incompressibility factors, and therefore certain values were tested using the more conservative Dunnett T3 test (if variances are unequal) rather than the Dunnett test (if variances are assumed equal), which only uses one control.

Regardless of which means comparison was used, the results indicate which models are significantly different from the WLC status quo. The results of all five tests are summarized in Table 2.

<b>TABLE 2. MAPE COMPARISON RESULTS</b>					
	<b>M = 0.0</b>	<b>M = 0.05</b>	<b>M = 0.10</b>	<b>M = 0.15</b>	<b>M = 0.20</b>
<b>WLC</b>	N/A	N/A	N/A	N/A	N/A
<b>Stanford-B</b>	X	X	X	X	X
<b>DeJong</b>	X	-	X	+	+
<b>S-Curve</b>	X	-	X	+	+

*Note.* MAPE = Mean Average Percent Error; WLC = Wright's Learning Curve  
 X indicates model is not significantly different from WLC  
 (+) indicates model is statistically less accurate than WLC (Higher MAPE)  
 (-) indicates model is statistically more accurate than WLC (Lower MAPE)

When the factor was held at 0.0 or 0.1, there was no statistical difference between the models, and these results reject all of the hypotheses. On the contrary, when the factor is held at 0.05, the DeJong and S-Curve models are more accurate, and these findings support all three of the hypotheses. When the incompressibility factor rises to 0.15 and 0.20, Wright's model holds as the most accurate. Results for all five means' comparison tests are displayed in the Appendix. In all cases, no statistical difference was shown between Wright's model and the Stanford-B model, and the same was true when comparing the S-Curve model and DeJong's model. This illustrates that in high production volumes, such as the 1,100-plus F-15s produced, incompressibility becomes much more significant than the prior experience units factor.

## Results

The results of this research are inconclusive regarding an answer to the overarching research question of whether a more accurate learning curve model is available for DoD use than Wright's original formulation. However, the results do provide some insight into the effects of learning and where to go from here. The findings also emphasize the importance of incompressibility ( $M$ ) in the learning process. Slight changes in the assumed incompressibility of the process led to drastically different results as to which model was most accurate.

The first hypothesis from this research was that at least one of the models would have a MAPE value statistically different from the others. This was not the case when the incompressibility factor was assumed to be 0.0 or 0.1, but the hypothesis holds for values of 0.05, 0.15, and 0.20. These results indicate that, although not uniformly, there does appear to be evidence that at least two of the models display a statistical difference. This result is important because it sets up the framework to be able to test the other hypotheses in the study.

The second hypothesis was that at least one model would have a MAPE value statistically lower than Wright's model. This hypothesis held only when the incompressibility factor was assumed to be 0.05; in all of the other cases, no statistical difference was calculated at 0.1, and the models were actually less accurate than Wright's model when  $M = 0.15$  and  $0.20$ . This finding indicates that as the process becomes more automated, Wright's curve actually performs better. These results do not fully support the second hypothesis, but do illustrate potential for learning curve improvement if an actual, universal incompressibility factor is found to be somewhere between 0.0 and 0.1. Post hoc analysis found that the S-Curve and DeJong models switch from being statistically more accurate to having no significant difference in MAPE value somewhere between 0.05 and 0.06. The follow-on research section will provide potential impacts of a statistically supported incompressibility factor and how that factor could potentially support the findings from these results.

***The findings of this research lead to two additional theoretical questions: why were the results so sensitive to the incompressibility factor, and what conclusions can be drawn about the application of modern learning models in DoD acquisition?***

The final part of this analysis was to test which model was the most accurate between the four. The third hypothesis from this research was that the S-Curve model would be the most accurate because it accounts for the slow decline in performance over time due to forgetting. As with the second hypothesis, this hypothesis is only partially supported when the incompressibility factor is assumed to be 0.05 and rejected by the other results. At 0.05, both the DeJong and S-Curve models are more accurate than Wright's

model; however, neither the DeJong nor S-Curve proved to be more accurate than the other. These results lead to inconclusive outcomes about which model is best, but again point to the importance of the incompressibility factor when determining best model fit.

The findings of this research lead to two additional theoretical questions: why were the results so sensitive to the incompressibility factor, and what conclusions can be drawn about the application of modern learning models in DoD acquisition? While the second question will be addressed at the end of this section, the first question may be due to the data itself. The incompressibility factor essentially represents the amount of potential learning that is lost for each unit due to automated production processes. If an incompressibility factor is 0.3, then only 70 percent of the potential learning can be achieved. When compounded over several lots and units (over 1,100 units for the F-15 A-E), a small shift in that percentage can result in a massive change in the cost of the units at the end of the production process.

This sensitivity affirms the need for additional research into incompressibility factors within the DoD and defense contractors in general. As mentioned earlier, the production of an aircraft is not unlike the production of a high-end sports car. The level of precision and craftsmanship required eliminates the use for certain automated processes that may be present in an assembly line at Ford or Toyota. Given this dynamic, assuming the real incompressibility factor is somewhere between 0.0 and 0.1 is not implausible. Follow-up investigation, involving inquiries to top practitioners and SMEs in the learning curve field, supports the belief that the percentage of automation is very, very small in an aircraft production environment. Additionally, different defense contractors may use various production processes that result in different incompressibility factors and thus increase the sensitivity of the costs to those factors. This is yet another reason for future incompressibility research that will be described later in this section.

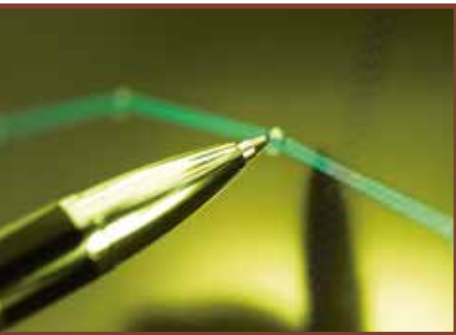
These results also indicate that learning is affected much more by incompressibility than prior experience units. The prior experience units parameter ( $B$ ) was the differentiating parameter between the WLC and Stanford-B model, as well as the difference between DeJong's learning formula and the S-Curve model. One explanation for this result may be the large number of units produced for the F-15. When examining over 1,100 units, a change to a mere 10 of the units will have a very limited impact on the outcome. However, if the same prior experience units' factor was applied to a smaller production line such as the 21 original units of the B-2 bomber, the difference may become very significant. In all five cases, there was no



statistical difference between the model and its close relative, meaning that the maximum change in  $B$  of 10 had no impact on the long-term estimates of the models. Therefore, it is safe to assume that simply adding a prior experience units' factor alone provides no value to the estimate if the production number is high, but the interaction between prior units and incompressibility could be very significant.

### Significance of Research

The results discussed in the previous section indicate that there is potential for a more accurate model in predicting the effects of learning within DoD acquisition. This study was unique in two primary areas. First, it investigated defense aircraft costs where past studies had primarily investigated commercial aircraft or components; and second, due to its nature, DoD cost estimating examines costs from an external perspective rather than internal. Therefore, the availability and accuracy of data may lead to more assumptions than prior studies.



Despite these intricacies, a few major conclusions can be drawn from the results. The first is that there is potential with two of the alternative learning curve models to increase estimate accuracy using learning curves by up to 5 percent over the entire production cycle based on the results for an incompressibility factor of 0.05. Post hoc analysis indicated that the largest difference between the Wright and S-Curve models—just over 5.2 percent—was seen at 0.04. While this percentage may seem small, for the more than \$20 billion production cycle of the F-15 A-E airframes, this percentage could

reduce error in the estimation process by as much as \$1 billion simply by changing the estimating tool. This research does not go so far as to say current cost estimating methodology is wrong; cost estimates are just that—estimates. This research suggests and hopes to provide the foundation for ways to improve current learning curve methodology. Determining which model is most appropriate is an area that requires more analysis. Thus far, the S-Curve and DeJong models appear to be worthy candidates. Further analysis incorporating incompressibility could reveal more information related to the application of the S-Curve and DeJong models, and consequently, the theory of forgetting within DoD methodology.

While the findings of this study do not support all of the hypotheses of this research or indicate which model is the best predictor of future costs, they do open up a dialogue for future change in DoD acquisition methodology.

These results stress the importance of incompressibility in learning and the potential for improvement based on that significance. Data collected during the initial production run of a weapon system could be used as a baseline to establish an incompressibility factor that is specifically tailored to that weapon system and production environment. Future research into incompressibility in aircraft production and comparative research into additional airframes as well as any of the dozens of other learning models available may help provide decision makers with additional information, and hopefully increase the accuracy of cost estimates as a whole. Additionally, the use of an incompressibility factor should not be limited to aircraft, as every weapon system production process utilizes some form of automated manufacturing. One of the primary contributions of this research is to highlight the importance of incompressibility and the relationship it has with the production process. Recognizing that each weapon system may have a unique incompressibility factor and incorporating this into estimation techniques should greatly improve cost estimates across weapon systems.

### **Assumptions and Limitations**

As always, there are limitations to this research and the methods used to test the hypotheses. One limitation to this study was the amount of data available for analysis. While some of the results from the analysis appear to be inconclusive, the data presented in this analysis are only a small fraction of all aircraft programs, and an even smaller portion of DoD programs as a whole. The Air Force Life Cycle Management Center/Financial Management Mission Execution Directorate (AFLCMC/FZ) has access only to programs under their control, and only data from those programs that reported on learning curves. These factors will limit the number of aircraft available for future analysis. A larger data set would have been preferred, but in this case the sample was limited to the data available. Follow-on analysis of incompressibility and additional Air Force and DoD programs are necessary before generalization of the findings can be made.

Another limitation is the accuracy of the data reported as actual costs. The accuracy, or lack thereof, in updating actual values for estimates has long been an issue in DoD, and has just recently been brought to light in an effort to clean up data repositories. However, the fact that many of the programs are under AFLCMC/FZ local control and span multiple decades should help to mitigate some of the uncertainty of the results.

One other potential limitation was the use of the lot plot point with the cumulative average theory. Lot data are often used in DoD cost estimates due to the nature of contractor reports, but that type of analysis has not been applied to the additional models used in this analysis. However, the

methods used were backed up by the AFCAH as well as other studies into learning curves. This methodology, in addition to the fact that lot data are widely used throughout the DoD, should reduce the effect the lot plot point assumption has on the results while simultaneously making them more generalizable to individual unit data.

## Recommendations for Future Research

This research answered several questions about the effects of learning in DoD, but there are still more questions that need to be addressed. Further, it sought to determine whether any alternative learning models are more accurate than Wright's model, which is commonly used throughout defense acquisition programs today. This study took steps toward accomplishing that goal and found that the S-Curve and DeJong models may be more accurate if the incompressibility factor for aircraft production is found to be between 0.0 and 0.5. However, the evidence is inconclusive as to which model is the most accurate, and results are extremely dependent on the assumptions made. Additional research into incompressibility factors would prove valuable to this learning curve analysis and paramount to any additional research using these models. As mentioned earlier, one of the major assumptions in this study was in the use of an incompressibility range from 0.0 to 0.2. Future research into what incompressibility factor should be used for aircraft production would provide insight into which models may be more appropriate, and also provide further insight into the validity of these results. Also, analysis into how incompressibility factors change between different defense contractors or how different platform types affect the production process could provide even more accuracy in future research. Clarifying these uncertainties will help produce more accurate and useful cost estimates using the models described in this article.

Future research should also look to broaden the scope of the programs used in this analysis. This research focused on fighter aircraft, and the initial pool of six was trimmed down to one aircraft. Follow-on studies should attempt to incorporate the findings in additional platforms such as bombers, cargo/tanker, and unmanned aircraft. Also, the use of additional models that do not rely on an incompressibility factor may provide more robust results. Results from the analysis of the F-15 should not necessarily be generalized to all aircraft as a whole. Further analysis may shed light on which models perform best on which aircraft or whether there is a single model that can be generalized to all platforms.

## Summary

When this research began, the goal was to find out whether a more accurate learning curve model for use in DoD exists. The AFLCMC cost staff supported the effort to find a way to improve current learning curve methodology in defense acquisition. Through the efforts of this research and the findings entailed within, there is evidence to support the hypothesis that at least one of the models may be more accurate than Wright's original model. This research found that both the DeJong and S-Curve models are statistically more accurate than the status quo when the incompressibility factor is somewhere between 0.0 and 0.5. However, if the factor is assumed to be .01 or higher, then Wright's model is the most accurate and the additional models do not improve on the current methodology. The results as to which model is the most accurate are inconclusive and do not support nor disprove the hypothesis that the S-Curve model is the most accurate of the four. At a minimum, this research provides the foundation for further research into additional types of aircraft as well as an applicable incompressibility factor that may indicate which model is the most accurate. Only then can the alternative models be considered for DoD methodology.

***One premise behind this research is that the current DoD learning curve methodology using Wright's 75-plus-year-old model should not be accepted as the status quo for the sake of simplicity or nostalgia.***

One premise behind this research is that the current DoD learning curve methodology using Wright's 75-plus-year-old model should not be accepted as the status quo for the sake of simplicity or nostalgia. If a more accurate learning model exists that can be applied to cost estimating within the DoD, it should be investigated and considered. This research illustrates the point that additional models are available. Some are more accurate in certain cases, and would undoubtedly provide the foundation for future research in defense acquisition, which can hopefully increase the accuracy and reliability of cost estimates and result in a more efficient use of government funding.

## References

- AFCAH. (2008). *Air Force cost analysis handbook*. Washington, D.C.: Author.
- Badiru, A. B. (1992). Computational survey of univariate and multivariate learning curve models. *IEEE Transactions on Engineering Management*, 39(2), 176-188.
- Badiru, A. (1995). Multivariate analysis of the effects of learning and forgetting on product quantity. *International Journal of Production Research*, 33(3), 777-794.
- Badiru, A. (2012). Half-life learning curves in the defense acquisition life cycle. *Defense ARJ*, 19(3), 283-308.
- Badiru, A., Elshaw, J., & Everly, M. (2013). Half-life learning curve computations for airframe life-cycle costing of composite manufacturing. *Journal of Aviation and Aerospace Perspectives*, 3(2), 6-37.
- Bailey, C. D. (1989). Forgetting and the learning curve: A laboratory study. *Management Science*, 35(3), 340-352.
- Carr, G. W. (1946). Peacetime cost estimating requires new learning curves. *Aviation*, 45, 76-77.
- Crawford, J. R. (1944). *Learning curve, ship curve, ratios, related data*. Burbank, CA: Lockheed Aircraft Corporation.
- Delta Research Corporation. (1987). Cost estimating system. *Aircraft cost handbook* (Vol. 2). Arlington, VA: Author.
- Department of Defense (DoD). (2007). *Basis of cost estimating*. Washington DC: Office of the Secretary of Defense.
- Everitt, B., & Skrondal, A. (2010). *The Cambridge dictionary of statistics* (4th ed.). Cambridge, UK: The Cambridge University Press.
- 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.
- Kronemer, A., & Henneberger, J. E. (1993). Productivity in aircraft manufacturing. *Monthly Labor Review*, 116(6), 24-33.
- Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583-621.
- McClave, J. T., Benson, P. G., & Sincich, T. T. (2011). *Statistics for business and economics* (11th ed.). Boston, MA: Pearson Education.
- Towill, D. R., & Cherrington, J. E. (1994). Learning curve models for predicting the performance of AMT. *International Journal of Advanced Manufacturing Technology*, 9, 195-203.
- Wright, T. P. (1936). Factors affecting the cost of airplanes. *Journal of the Aeronautical Sciences*, 3(4), 122-128. doi: 10.2514/8.155
- Yelle, L. E. (1979). The learning curve: Historical review and comprehensive study. *Decision Science*, 10(2), 302-328.

## Appendix

### Dunnnett T3 Test Results

DUNNETT T3 TEST (M = 0.0)						
(I) Model	(J) Model	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-.00094	.01299	1.000	-.0375	.0356
	3.00	.00000	.01288	1.000	-.0363	.0363
	4.00	-.00111	.01299	1.000	-.0377	.0355
2.00	1.00	.00094	.01299	1.000	-.0356	.0375
	3.00	.00094	.01299	1.000	-.0356	.0375
	4.00	-.00017	.01309	1.000	-.0371	.0367
3.00	1.00	.00000	.01288	1.000	-.0363	.0363
	2.00	-.00094	.01299	1.000	-.0375	.0356
	4.00	-.00111	.01299	1.000	-.0377	.0355
4.00	Dimension 3 1.00	.00111	.01299	1.000	-.0355	.0377
	2.00	.00017	.01309	1.000	-.0367	.0371
	3.00	.00111	.01299	1.000	-.0355	.0377

Note. Dunnnett *t* tests treat one group as a control and compare all other groups against it.

DUNNETT T3 TEST (M = 0.05)						
(I) Model	(J) Model	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2.00	Dimension 3 1.00	.00094	.01784	1.000	-.0421	.0440
3.00	1.00	-.04616*	.01784	.033	-.0892	-.0031
4.00	1.00	-.04670*	.01784	.030	-.0898	-.0036

\* The mean difference is significant at the 0.05 level.

DUNNETT T3 TEST (M = 0.1)						
(I) Model	(J) Model	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-.00094	.01299	1.000	-.0375	.0356
	3.00	-.04165	.02199	.343	-.1055	.0222
	4.00	-.03997	.02178	.376	-.1032	.0232
2.00	1.00	.00094	.01299	1.000	-.0356	.0375
	3.00	-.04070	.02204	.369	-.1047	.0233
	4.00	-.03902	.02184	.404	-.1024	.0243
3.00	1.00	.04165	.02199	.343	-.0222	.1055
	2.00	.04070	.02204	.369	-.0233	.1047
	4.00	.00168	.02814	1.000	-.0776	.0810
4.00	1.00	.03997	.02178	.376	-.0232	.1032
	2.00	.03902	.02184	.404	-.0243	.1024
	3.00	-.00168	.02814	1.000	-.0810	.0776

DUNNETT T3 TEST (M = 0.15)						
(I) Model	(J) Model	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-.00094	.01299	1.000	-.0375	.0356
	3.00	-.15035*	.02337	.000	-.2185	-.0822
	4.00	-.14856*	.02328	.000	-.2164	-.0807
2.00	1.00	.00094	.01299	1.000	-.0356	.0375
	3.00	-.14941*	.02343	.000	-.2176	-.0812
	4.00	-.14762*	.02333	.000	-.2156	-.0797
3.00	1.00	.15035*	.02337	.000	.0822	.2185
	2.00	.14941*	.02343	.000	.0812	.2176
	4.00	.00179	.03037	1.000	-.0838	.0874
4.00	1.00	.14856*	.02328	.000	.0807	.2164
	2.00	.14762*	.02333	.000	.0797	.2156
	3.00	-.00179	.03037	1.000	-.0874	.0838

\* The mean difference is significant at the 0.05 level.

DUNNETT T3 TEST (M = 0.2)							
(I) Model	(J) Model	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval		
					Lower Bound	Upper Bound	
1.00	2.00	-.00094	.01299	1.000	-.0375	.0356	
	3.00	-.25972*	.02454	.000	-.3314	-.1880	
	4.00	-.25804*	.02445	.000	-.3295	-.1866	
2.00	1.00	.00094	.01299	1.000	-.0356	.0375	
	3.00	-.25877*	.02459	.000	-.3306	-.1870	
	4.00	-.25709*	.02451	.000	-.3286	-.1855	
3.00	1.00	.25972*	.02454	.000	.1880	.3314	
	2.00	.25877*	.02459	.000	.1870	.3306	
	4.00	.00168	.03216	1.000	-.0889	.0923	
4.00	1.00	.25804*	.02445	.000	.1866	.3295	
	2.00	.25709*	.02451	.000	.1855	.3286	
	3.00	-.00168	.03216	1.000	-.0923	.0889	

\* The mean difference is significant at the 0.05 level.



## Biographies



**Capt Justin R. Moore, USAF**, is a cost analyst at the U.S. Air Force Cost Analysis Agency. His professional experience to date includes aircraft production and software estimating. Capt Moore holds an MS in Cost Analysis from the Air Force Institute of Technology, and a BS in Management from the U.S. Air Force Academy.

*(E-mail address: [justin.r.moore52.mil@mail.mil](mailto:justin.r.moore52.mil@mail.mil))*



**Dr. John J. Elshaw** is an assistant professor of systems engineering in the Department of Systems Engineering and Management at the Air Force Institute of Technology. He holds a BS in Accounting from The University of Akron, an MBA from Regis University, and a PhD in management with a specialization in Organizational Behavior and Human Resource Management from Purdue University. He is a graduate of the United States Air Force Squadron Officer School and Air Command and Staff College.

*(E-mail address: [john.elshaw@afit.edu](mailto:john.elshaw@afit.edu))*



**Dr. Adedeji B. Badiru** is dean of the Graduate School of Engineering and Management at the Air Force Institute of Technology. He is a registered professional engineer, a Fellow of the Institute of Industrial Engineers, and a Fellow of the Nigerian Academy of Engineering. He holds a BS in Industrial Engineering; an MS in Mathematics and an MS in Industrial Engineering from Tennessee Technological University; and a PhD in Industrial Engineering from the University of Central Florida.

*(E-mail address: [adedeji.badiru@afit.edu](mailto:adedeji.badiru@afit.edu))*



**Lt Col Jonathan D. Ritschel, USAF**, is an assistant professor and director, Cost Analysis Program, in the Department of Systems Engineering and Management at the Air Force Institute of Technology (AFIT). His research interests include public choice, the effects of acquisition reforms on cost growth in DoD weapon systems, research and development cost estimation, and economic institutional analysis. Lt Col Ritschel holds a PhD in Economics from George Mason University, a BBA in Accountancy from the University of Notre Dame, and an MS in Cost Analysis from AFIT.

*(E-mail address: [jonathan.ritschel@afit.edu](mailto:jonathan.ritschel@afit.edu))*