Step-Down Functions in Airframe Learning Curves: Do They Exist?

Susan L. Moore

Jonathan D. Ritschel, Ph.D.

Edward D. White, Ph.D.

Clay M. Koschnick

Brandon M. Lucas

Shawn M. Valentine

Introduction

Defense cost analysts employ a multitude of techniques to estimate the cost of a weapon system. One of the most widely accepted and utilized techniques is learning curve analysis. Learning curves are traditionally used to estimate recurring costs in a production process (Mislick & Nussbaum, 2015). While previous researchers have studied learning curves along a multitude of dimensions, one area that lacks empirical examination in defense programs is the concept of a step-down function. This lack of empirical examination has led to some debate on whether step-down functions should be employed in cost estimates. We examine the evidence in military fighter airframes to shed light on the issue. Thus, the purpose of this article is four-fold: 1) empirically detect step-down functions in defense aircraft programs 2) examine the impact of weight normalization on step-down functions 3) analyze factors that impact step-down functions and 4) develop an empirically based Cost Estimating Relationship (CER) to predict first unit production costs based upon development unit cost data.

The Step-Down Function

The production of an end item often begins with prototype or development units. This presence of prototype or development units has created the idea of a step-down function. More specifically, step-down functions occur between the Engineering, Manufacturing, and Development (EMD) and the Production phases and are a method for estimating the theoretical first unit production cost based on development asset data (Mislick & Nussbaum, 2015). In learning curves, a step-down would appear as a downward shift on the graph with learning resuming at the same or a modified slope.

The theory undergirding step-down functions is that the development unit is a near production copy in design, physical, and performance characteristics, but is usually accomplished in an EMD environment rather than a production line set-up (Hardin & Nussbaum, 1994). Therefore, the cost to manufacture a development asset is expected to be more expensive than a production model (Hardin & Nussbaum, 1994). Mathematically, the ratio of the production phase first unit cost to the development first unit cost (or average development unit cost) is known as a step-down factor. The actual cost difference between the development first unit cost (or average development unit cost) and the production first unit cost is the step-down (Mislick & Nussbaum, 2015). When first unit production costs are greater than development costs it is called a step-up rather than a step-down.

The Government Accountability Office (GAO) identifies two main learning curve methodologies: Continuous and Step-Down (GAO, 2020). The Step-Down methodology is further broken into two subcategories consisting of Sequential and Disjoint theory. These three models are shown in Figure 1. For ease of visualization, the models are shown in log space.

Continuous learning curve theory is the traditional learning curve described by Wright (1936) for aircraft production but includes the developmental units as part of the curve as shown in Figure 1(a). As such, continuous learning curve theory assumes the same improvement slope in production as well as development. The production estimate can simply be calculated by continuing down the curve for the desired quantity (GAO, 2020).

The two subsets of Step-Down theory, Sequential and Disjoint, typically assume that the improvement slope remains the same in development and production but there is a step down in the value between the cost of the first development unit and the cost of the first production unit (GAO, 2020). Sequential theory states that the cost improvement continues when the first production unit equals the last development unit plus one. For example, if the last development unit is 10, then the first production unit would be 10 + 1 = 11. Initially, Sequential theory sounds like Continuous theory where you consider learning made in development and apply it to production units. Where it differs is that there is a discontinuity in the curve between development and production as shown in Figure 1(b).

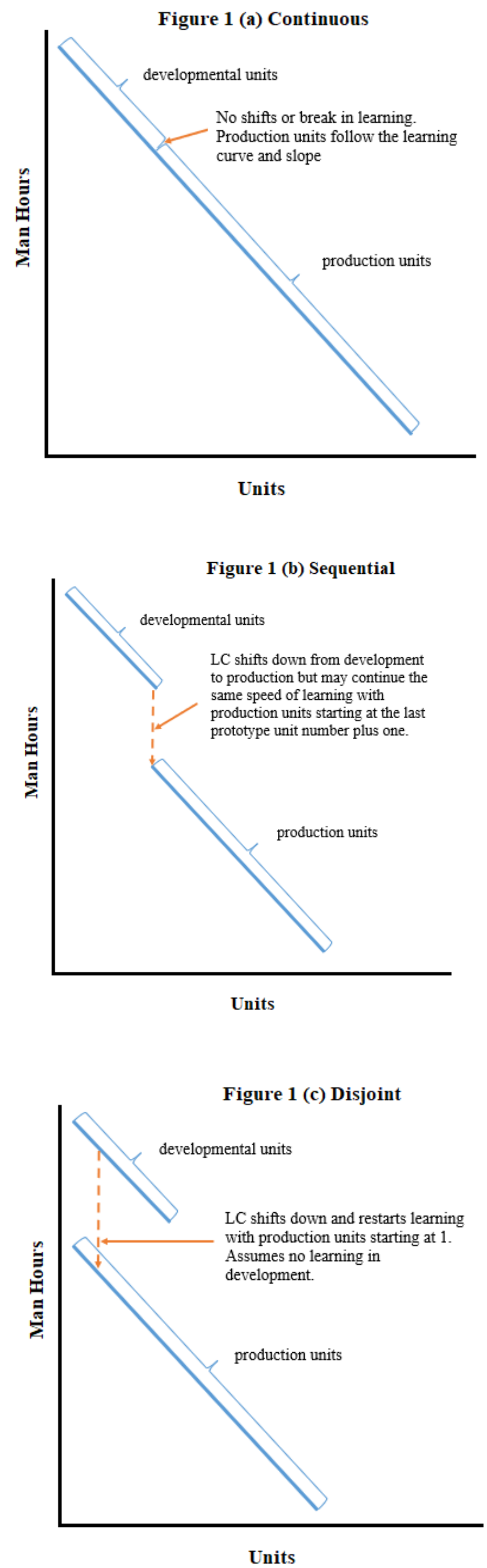

Figure 1. Step-Down Function Types in Log Space

Table 1: Summary of Step-Down Studies

Author	Data	Focus	Method	Conclusion
Waller (1976)	DoD electronics data set (data unavailable)	Formulate step-down factors using disjoint and sequential theories	Compared theoretical first production unit (T1s) developed using disjoint and sequential theories	Mixed results; There was no clear conclusion to whether disjoint or sequential theory was a better predictor for all
Federic (1979)	Same data set as Waller (1976) and included hypothetical data	Cost continuity does not have to be distinctly disjoint or sequential but a potential spectrum between the two theories	Cost improvement curves were fit using different points on the spectrum between production quantities only and prototype and production quantities together	Mixed results; There was no clear conclusion to whether disjoint or sequential theory was a better predictor for all
Hubach, Pehrsson, & Fox (1987)	8 Airframe systems, 6 engines, 6 avionics systems	Determine what is the appropriate range of production data to use in fitting cost improvement curves	Used Ordinary Least Squares (OLS) regression to calculate 6 curves for each system – 3 disjoint and 3 sequential curves.	Mixed results; Airframe did best under disjoint theory. Engine modeling was marginally better using sequential modeling. Avionics had inconclusive results
Malcolm (1991)	7 Marine amphibious assault vehicles	Focused on the relationship between development and production unit costs	Used OLS regression with disjoint and sequential theories	Sequential model was the most applicable to estimating the costs of the amphibious assault vehicles
Hardin & Nussbaum (1994)	Reviewed 9 step-up/ step-down studies	Analyzed the relationship between development and production costs and compared it to other step-down or step-up studies	N/A, no unique model development.	A general step-up or step-down factor can be applied to all types of systems, but the equation would have a much higher variance
Cherwonik et al. (2012)	2 assault vehicles	Use a reference point other than T1 to calculate the learning curve and examined the step-up/down factors	Created basic production learning curves and then calculated step-down factors for prototype to production units	Utilizing a prototype Average Unit Cost (AUC) compared to T1000 provided the least varied step factor
Bui (n.d.)	6 prototype aircraft airframes, 12 production aircraft airframes, and 6 tactical missiles	Analyzed production step-down factors for aircraft and tactical missile manufacturing experiences	Use OLS regression to generate learning curves and determine the step-down percentage from the calculated T1 and average prototype unit cost	Airframe step-down was between prototype and Full Scale Development (FSD) and had lower step-down factors. Missile step downs were higher but occurred between FSD and production

Disjoint theory restarts improvement at the first production unit and does not consider learning created during development phases to be significant (GAO, 2020). Disjoint theory has a curve displacement, but the improvement starts over at unit one rather than at the last development unit plus one as shown in Figure 1 (c). Because it restarts learning, disjoint theory usually results in significantly lower production estimates (GAO, 2020).

Previous Step-Down Function Studies

To the best of our knowledge, there are no previous step-down function studies published in peer-reviewed literature. However, we were able to find seven reports (or conference presentations) specifically related to step-down functions in defense programs. These reports are summarized in Table 1.

There are several key points from the studies in Table 1. First, note that all but one of the step-down studies are more than 25 years old. Second, the majority of the studies have very small sample sizes. For example, the most recent study from 2012 only examined two assault vehicles. Third, many of the studies, in addition to examining the step-down function itself, attempted to develop a Cost Estimating Relationship (CER) between development and first unit production costs. Thus, one of the goals of this article is to develop a new CER with more recent data from fighter aircraft for modern-day practitioner use.

Perhaps the most comprehensive study we discovered is by Hardin and Nussbaum (1994). They reviewed nine internal Navy studies that focused on missile systems, radar, and general electronics. Part of their study examined CERs developed for these disparate system types both individually and as an aggregated CER. Their conclusion was that there could be a general (i.e. aggregated) step-up or step-down factor CER that can be applied to all types of systems, but that equation would have a much higher variance. Therefore, in general, they recommend using system specific step-up/step-down factors in lieu of a CER that applies to all system types. This finding from Hardin and Nussbaum provides the motivation for our study focusing on fighter aircraft programs as a single system type.

Data

The data is primarily sourced from Contractor Cost Data Summary Reports, or DD 1921-2s (Progress Curve Reports), via the Life Cycle Management Center (LCMC) at Wright-Patterson Air Force Base, Ohio. The focus of this study is fighter airframes. The original dataset included 18 programs with 513 lots. The following four criteria had to be met for a program to be included in the final dataset:

1. Have at least one development lot
    a. If no development lot is listed based off DD 1921-2, the early lots can be deemed development if the absolute value of the airframe weight is at least 5% different than its successor *and* if the lot has less than three aircrafts manufactured
2. Have at least four or more production lots
3. Have direct man-hour data for each lot
    a. If less than 20% of direct man-hour production data was missing, the data was imputed
4. Is fighter airframe

The first inclusionary criterion is to ensure that there is adequate data to formulate a step-down factor. To maximize the programs that can be included, two sub-criteria had to be met. Both criteria were developed by reviewing the data available that had development lots identified and by speaking to a subject matter expert at the Air Force Life Cycle Management Center (S. Valentine, personal communication, October 27, 2021). The sub-criteria were purposely made to be conservative in nature to ensure that the inclusion of any lots in the final dataset would not skew the results.

The second inclusionary criterion is to ensure that there was adequate data to calculate the production theoretical first unit (T1). The third inclusionary criterion is that the programs must have complete direct man-hour data for each lot. To maximize the programs that could be included, programs with minimal missing direct man-hour data were reviewed. If the program's production lots were missing less than 20 percent of its direct man-hour data, a line was fitted to the available production lots and an equation was generated. The equation was used to calculate the missing production direct man-hours per aircraft. Table 2 shows the final dataset after employing the inclusion/exclusion criteria.

Table 3 shows the 10 programs analyzed in this article. The program names have been omitted and are designated as Program A, Program B, etc, as a precaution in protecting the data. Note the two programs with asterisks. First, Program F, had missing production man-hour data for two production lots. However, because Program F was missing less than 20 percent of the direct man-hour data, the hours were derived in accordance with inclusion criteria 3(a) above. Second, Program J's development lot was categorized in accordance with inclusion criteria 1(a) above. This program did not have a development lot annotated on the 1921-2 due to its unique acquisition history.

Table 2: Inclusion/exclusion criteria describing the establishment of the final analyzed dataset

	Number of programs	Number of entries
Original dataset	18	513
No development lots	4	69
Less than four production lots	1	6
Missing >20% direct man-hours for lot(s)	2	114
Not airframe	1	12
Remaining dataset	**10**	**312**

Table 3: Final Dataset

Aircraft	Service	Type	Dev Lot(s)	Prod Lot(s)	Total
Program A	Air Force	Fighter	1	8	9
Program B	Navy/ Marine	Fighter	1	50	51
Program C	Air Force	Fighter	1	13	14
Program D	Air Force	Fighter	2	55	57
Program E	Air Force	Fighter	7	33	40
Program F*	Air Force	Fighter	9	11	20
Program G	Navy	Fighter	3	67	70
Program H	Air Force	Fighter	1	12	13
Program I	Marines	Fighter	6	4	10
Program J**	Air Force	Fighter	2	26	28
				Total	**312**

Methods

The term *improvement curve* denotes that costs are used as the primary measure. While there are merits to employing improvement curves, using cost as the dependent variable has some well-documented limitations. These limitations include concerns with wrap rates (Mislick & Nussbaum, 2015), economies of scale (GAO, 2020), and escalation (Hogan et al., 2020). An alternative unit of measure for calculating learning curves is hours and is the approach this article takes. Practitioners have noted that as a program progresses and both cost and hours are provided by contractors, hours are the preferred learning curve unit of measure (S. Valentine, personal communication, August 25, 2021). In addition, negotiations between the government and contractor regarding the program learning curves are typically discussed from a man-hour perspective. These hours are reported on DD 1921-2s. More specifically, the direct man-hours per aircraft from the DD 1921-2 is the summation of four categories: engineering, tooling, quality control and manufacturing. Of these categories, only quality control and manufacturing are accounted for in learning curves and included in our dataset.

Determining Step-Down Factors

The first objective of this article is to determine whether there are step-down factors in the ten fighter aircraft programs. Two unique step-down factors are calculated for each program. The theoretical first production unit (T1) is calculated via OLS regression analysis. This calculated production T1 is then divided by the development first unit direct man-hours (EMD FUH) (note: for purposes of this research this is considered equivalent to development T1). Next, the production T1 is divided by the Engineering and Manufacturing Design (EMD) average unit hours (EMD AUH). The two equations are shown below:

$$\text{Step-Down Factor 1} = \text{Production T1}/\text{EMD FUH} \tag{1}$$

$$\text{Step-Down Factor 2} = \text{Production T1}/\text{EMD AUH} \tag{2}$$

These step-down factors will then be tested via a Sign test. The Sign test is the non-parametric equivalent of a paired *t* test where it tests for consistent differences of two groups using the median (Shier, 2004). The non-parametric test is required because the total amount of programs reviewed is less than 30 and a particular distribution cannot be assumed.

This Sign test will be conducted for both step-down factor calculations (EMD FUH and EMD AUH). The Sign test is based on the direction of the plus and minus sign of the observation and not on their numerical value. In other words, the Sign test will determine as a group of programs whether the actual first development lot direct man-hours or development average unit hours are statistically different than the calculated production T1. The hypotheses for the Sign test are as follows:

H_0: Difference in median of the signed differences = 0
H_a: Difference in median of the signed differences ≠ 0

Deriving a Cost Estimating Relationship (CER)

Previous step-down studies (Hardin & Nussbaum, 1994) and cost estimation textbooks (Mislick & Nussbaum, 2015) highlight the utility of a CER for practitioner use. The goal of these CERs is to provide a basis for determining an aircraft's production T1 when development data exists. Our first CER uses the development FUH as the independent variable (x) and the calculated production T1 as the dependent variable (y). The data will be fit to a linear (see Equation 3) and power (see Equation 4) function using a non-linear solver; the adjusted R^2 will be used to determine which equation best fits the data. These models will also be used for EMD AUH as the independent variable and the calculated production T1 as the dependent variable. Thus, a total of four regression models will be evaluated.

$$y = \beta_0 + \beta_1 x_1 + \varepsilon \tag{3}$$

Where:
y: Production T1
x_1: EMD FUH or EMD AUH

$$y = Ax^b + \varepsilon \tag{4}$$

Where:
y: Production T1
x: EMD FUH or EMD AUH

Impact of Weight Normalization on Learning Curves

Normalization by weight in learning curves is not a widespread approach. However, some practitioners advocate for it and there is precedence in the literature. For example, Alchian's (1950) study of 22 bomber, fighter, trainer, and transport airframes after World War II normalized the data using direct labor hours per pound. Therefore, we examine the impact of weight normalization in our dataset.

The normalization is accomplished by dividing the program's direct man-hours by the airframe weight. Next, we repeat the Sign tests as described in the *Determining Step-Down Factors* section previously with the newly normalized data. The results from the Sign test will indicate whether a step-down function exists in the data. Comparing the results from the Sign tests of the non-normalized to the normalized data will illuminate any impacts from weight normalization.

Results

We first developed step-down factors for the ten aircraft programs. Recall that two unique step-down factors (see Equations 1 and 2) are calculated for each program: one using EMD FUH and one using EMD AUH. A factor below 1 means that the EMD FUH or EMD AUH had higher direct man-hours per aircraft (step-down). A factor above 1 means that the EMD FUH or EMD AUH had lower direct man-hours per aircraft (step-up). Some programs have the same step-down factor for both EMD FUH and EMD AUH calculations due to only having one development lot. These programs are marked with an asterisk in Table 4.

Table 4: Step-Down Factors

Aircraft	Type	Step-Down EMD FUH	Step-Down EMD AUH
Program A*	Fighter	0.675	0.675
Program B*	Fighter	0.876	0.876
Program C*	Fighter	1.316	1.316
Program D	Fighter	0.631	0.822
Program E	Fighter	0.36	0.726
Program F	Fighter	0.168	0.212
Program G	Fighter	0.376	0.506
Program H*	Fighter	0.833	0.833
Program I	Fighter	0.391	0.493
Program J	Fighter	0.935	1.411
* Denotes only one development lot			

The step-down for the EMD FUH calculation ranged from 0.168 to 1.316. The range for EMD AUH is 0.212 and 1.411. These airframes have a mean of 0.656 and standard deviation of 0.344 for the EMD FUH and 0.787 mean and 0.364 standard deviation for EMD AUH. Only Program C and Program J had step-up factors.

Table 5: Sign Test Results

	EMD FUH	EMD AUH
Test Statistic	22.5	17.5
Prob > |z|	0.0039	0.084

The data in Table 4 appears to show a consistent step-down factor for both calculations. Testing for statistical significance of that observation is discerned by the Sign test. The null hypothesis is that there is no step-down factor, and the alternative hypothesis is that there is a step-down factor. This test uses an alpha of 0.1. Results of the Sign Test is shown in Table 5.

Both EMD FUH and EMD AUH rejected the null hypothesis. This means that, for the sample, there is a statistically significant step-down factor between development and production. These initial results indicate that practitioners developing estimates on fighter airframes should consider incorporating a step-down factor in their estimate.

Predicting Production T1: the Cost Estimating Relationship

Some previous step-down factor research (Malcolm, 1991; Hardin & Nussbaum, 1994) developed CERs for practitioner use in predicting production T1 values from EMD data. Thus, the next step of this research uses the step-down calculations from Table 4 to create CERs. Equation forms were limited to linear and power functions due to their prevalence in learning

Table 6: Cost Estimating Relationship

Step-Down Factor Type	Equation Type	Equation	Adjusted R^2
EMD FUH	Linear	$y = 0.3939x + 27.1994$	0.4983
EMD FUH	Power	$y = 0.0003x^{2.3502}$	0.4246
EMD AUH*	**Linear**	$\boldsymbol{y = 0.5842x + 28.847}$	**0.8149**
EMD AUH	Power	$y = 0.9083x^{0.9618}$	0.8222
*Recommended CER is bolded in the Table			

curve use. A total of four individual equations were created. See Table 6.

As shown in Table 6, we recommend using the EMD AUH linear CER. The adjusted R^2 is nearly equivalent between the linear and power AUH models. However, by choosing the linear CER, OLS regression can be utilized for further evaluation. This practical consideration trumps the minor decrease in adjusted R^2 incurred by selecting the linear model. The recommended CER has a coefficient of variation (CV) of 0.3222. This implies that the CER is a good starting point for a cost analyst to use, but some caution is advised due to the moderately high CV.

Weight Normalized Step-Down Factor Analysis

As previously discussed, weight normalization in learning curves is advocated by some practitioners and is also found in previous literature. Therefore, we normalize the data by weight and recalculated the step-down factors to determine the impacts. Note that one platform, Program G, is excluded from our original dataset due to lack of airframe weight information. The resultant step-downs are shown in Table 7.

The mean and standard deviation for EMD FUH is 0.689 and 0.345, respectively. These values are higher than the mean and standard deviation of the non-normalized data. Similarly, the mean and standard deviation for EMD AUH is 0.826 and 0.371, respectively. These EMD AUH results are also higher than the non-normalized data. These higher mean values indicate that normalizing by weights reduces the impact of a step-down factor. In other words, the reduction in hours for the first unit of production from its prototype development hours is less when normalized for weight than when it is not normalized for weight.

Table 7: Step-Down Factors Normalized by Weight

Aircraft	Type	Step-Down EMD FUH	Step-Down EMD AUH
Program A*	Fighter	0.67	0.67
Program B*	Fighter	0.909	0.909
Program C*	Fighter	1.296	1.296
Program D	Fighter	0.631	0.821
Program E	Fighter	0.371	0.749
Program F	Fighter	0.176	0.247
Program H*	Fighter	0.774	0.774
Program I	Fighter	0.401	0.505
Program J	Fighter	0.968	1.462
* Denotes only one development lot			

Next, the Sign Test is conducted for the weight normalized data. The alpha is 0.10 and the results of the Sign Test are shown in Table 8.

Table 8: Sign Test Results (Normalized by Weight Data)

	EMD FUH	EMD AUH
Test Statistic	16.5	9.5
Prob > |z|	0.0547	0.3008

The EMD FUH Sign Test rejects the null hypothesis. This suggests that for the sample, there is a statistically significant step-down factor between development and production. This finding is consistent with the finding from the non-normalized EMD FUH data in Table 5. However, EMD AUH fails to reject the null hypothesis of the Sign Test. This finding is contrary to the finding from the non-normalized EMD AUH data in Table 4 which rejected the null. The normalized EMD AUH result indicates that there is *not* a step-down between development and production. In other words, normalizing for weight matters in the EMD AUH calculations.

The contradictory findings of the Sign Test in Table 8 between EMD FUH and EMD AUH warrants further investigation. We hypothesize the difference may lie in "legacy" versus "modern" aircraft. The rationale is that touch labor in legacy aircraft was simpler, with machinists completing fewer complex tasks, in a pre-computer environment. To discern if this is the case, we divide the aircraft into the legacy and modern categories via subject matter expert inputs. Next, we run a Wilcoxon Rank Sum test. The Wilcoxon Rank Sum test is a non-parametric test that tests the locations of each set of data. If the distribution of each dataset is the same, then the location can be interpreted as the median (McDonald, 2014). The Wilcoxon Rank Sum test compares the following hypotheses:

Table 9: Wilcoxon Rank Sum Test Results

	Original Data		Normalized by Weight	
	EMD FUH	EMD AUH	EMD FUH	EMD AUH
Test Statistic	2.0254	2.2386	1.8371	2.327
Prob > |z|	0.0428	0.0252	0.0662	0.02

H_0: Median ranks are the same
H_a: Median ranks are different

The null hypothesis states that there is no difference in the step-down factor between the legacy and modern aircraft. The alternative hypothesis states that there is a difference in the step-down factors. The results of the Wilcoxon Rank Sum test are shown in Table 9.

As shown in Table 9, all tests reject the null hypothesis at an alpha level of 0.10. This indicates that there is a difference in the step-down factors across all four measures. These results lend credence to the aforementioned suggestion that the mixed results of Table 8 are likely due to differences in the six modern verses the four legacy aircraft. To corroborate these findings, we rerun the Sign Test previously performed, but this time we *only* include the modern fighter aircraft. The results of this new test rejected the null and supports our hypothesis. However, due to the low *n* value of six associated with just examining the modern fighter aircraft subset, the results of that Sign Test cannot be fully trusted. Therefore, while we mention this robustness check, we caution the reader that this result must be taken with a grain of salt, and therefore we do not show the actual test results.

In summary, there are three key points associated with weight-normalization. First, step-down factors exist even when normalizing by weight. Second, the impact of normalizing by weight, however, is to dampen the magnitude of the step-down factors. Lastly, when calculating weight normalized step-down factors, it is imperative to separate the modern from the legacy aircraft.

Factors that Impact Step-Down Functions

The final analysis examines the impact of factors that the GAO (2020) has identified as important to consider when developing a step-down factor in learning curves. The four factors identified by the GAO are:

A break from the last prototype unit to production
Similarity between prototype units and production units
The production rate
The extent to which the same facilities, processes, and people are used in development and production

Only five of the programs in our study had the requisite data to examine the four GAO criteria. Within those five programs, it was discovered that all five in the sample had similar development and production aircraft and had minimal changes in facilities, processes, and people. Those results effectively removed two of the four GAO considerations (numbers two and four) from the analysis. A regression was conducted with the two remaining factors (number one and three above). It found the production rate to be a significant factor with a *p*-value of 0.0058. The positive sign supports the intuition that the larger the production rate, the bigger the step-down factor will be in the learning curve.

We strongly caution that this result from the GAO influential factors is not conclusive. Our data sample of five was simply too small to draw any definitive conclusions. Additionally, as discussed above, our data only looked at two of the four factors. While we are encouraged that the singular result we found did align with the theory, future research with a more robust dataset is needed to have confidence in the result.

Discussion and Conclusion

The debate regarding step-down factors begins at the most fundamental level regarding whether they exist or not in defense aircraft. Our examination of fighter airframes provides empirical evidence that step-downs are present. In our dataset, the mean step-down factor was found to range from 0.656 (FUH) to 0.787 (AUH), which is a significant reduction in hours for the first unit of production from its development unit.

While we were able to detect the presence of step -down functions, we did not attempt to discern whether the step-down function is Sequential or Disjoint. Some of the prior non-peer reviewed studies from Table 1 attempted to delineate between the two. However, we did not believe the nature of our data lent to such a determination. As a result, the nature of the step-down function (Sequential or Disjoint) remains an open question.

A second issue that is debated is the impact of weight normalization on step-down functions. We find that weight normalization does have an impact in fighter airframes, but it only dampens the magnitude of the step-down rather than removing it fully. The magnitude of the mean differences are approximately 6% for both FUH and AUH calculations. This implies that those practitioners who choose to normalize by weight should show smaller hour reductions. Additionally, it is important to segregate the data between modern and legacy platforms if weight normalization is your preferred approach. Overall, the authors remain agnostic to whether practitioners choose to normalize by weight or not. We simply reiterate the step-downs will still occur in most cases, but to a lesser extent.

Lastly, we provide a recommended CER to estimate the theoretical first unit production cost with development data. Our recommended form is linear with average unit development hours as the independent variable. The simplicity of the CER and ease of implementation mirrors the prior DoD studies (Malcolm, 1991; Hardin and Nussbaum, 1994). Thus, we believe this has great potential for practitioner adoption.

In summary, this article is a significant step forward in understanding step-down functions in DoD programs. With advancement, however, comes limitations that merit acknowledgment. Specifically, the small sample sizes in our tests mutes the statistical results. In some cases, such as the examination of the four postulated GAO factors, the lack of data meant even exploratory examination was not possible. These limitations, however, present an opportunity for future researchers. While we focused solely on fighter aircraft airframes, there is the potential to replicate our analysis with other platform types. Similarly, as more data is collected, a more robust investigation into the factors that impact step-down functions can occur. All these endeavors can add to a fuller understanding of step-down functions in military systems. We hope this article provides a launching ground for these future research efforts.

References

Alchian, A.A. (1950). *Reliability of progress curves in airframe production, (Report Number RM-260-1)*. The RAND Corporation. https://www.rand.org/pubs/research_memoranda/RM260-1.html.

Bui, J. (n.d.). *Analyzing production step down with application to space hardware*. Institute for Defense Analyses.

Cherwonik, J., Craig, C., & Liang, W. (2012). *Learning curve step-down analysis: Abrams main battle tank and Bradley fighting vehicle final report*. Technomics Inc.

Frederic, B. (1979). *Further investigations of the relationship between hardware development costs and hardware production costs in military systems*. Tecolote Research Inc.

Government Accountability Office. (2020). *GAO-20-195G Cost estimating and assessment guide*. https://www.gao.gov/assets/gao-20-195g.pdf

Hardin III, P.L., & Nussbaum, D.A. (1994). *Analyses of the relationship between development and production costs and comparisons with other related step-up/step-down studies*. Naval Center for Cost Analysis.

Hogan, D.W., Elshaw, J.J., Koschnick, C.M., Ritschel, J.D., Badiru, A.B. & Valentine, S.M. (2020). Cost estimating using a new learning curve theory for non-constant production rates. *Forecasting,* 2(4): 429-451. https://doi.org/10.3390/forecast2040023

Hubach, S.O., Pehrsson, K.M., & Fox, T.B. (1987). *Prototype to production step-down model final report*. Management Consulting & Research Inc.

Malcolm, D.S. (1991). *Parametric cost estimation utilizing development-to-production relationship applied to the advanced amphibious assault vehicle*. [Master's Thesis, Naval Postgraduate School]. DTIC https://archive.org/details/DTIC_ADA245920

McClave, J.T., Benson, P.G., & Sincich, T. (2014). *Statistics for business and economics* (12th ed.). Pearson.

McDonald, J.H. (2014). *Handbook of biological statistics* (3rd ed.). Sparky House Publishing. Retrieved November 19, 2021, from http://www.biostathandbook.com/HandbookBioStatThird.pdf.

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

Shier, R. (2004). *Statistics: 2.1 the sign test - www.statstutor.ac.uk*. Statistics: 2.1 The Sign Test. Retrieved November 19, 2021, from https://statstutor.ac.uk/resources/uploaded/signtest.pdf.

Waller, W.E. (1976). *The relationship between development and production hardware costs in military weapons systems*. Tecolote Research, Inc.

Wright, T. P. (1936). Factors affecting the cost of airplanes. *Journal of the Aeronautical Sciences*, 3(4), 122-128. https://doi.org/10.2514/8.155

Captain Susan Moore, *is a cost analyst at the Space Systems Command, Los Angeles AFB, CA. She holds a BS in Business Administration with a concentration in Finance from the University of California, Riverside and a MS in Cost Analysis from the Air Force Institute of Technology (AFIT). Her primary research interests include learning curves, public choice, and cost analysis. (Email address: Susan.Moore.12@us.af.mil)*

Dr. Jonathan D. Ritschel *is an Associate Professor of Cost Analysis in the Department of Systems Engineering and Management at AFIT. He received his BBA in Accountancy from the University of Notre Dame, his MS in Cost Analysis from AFIT, and his PhD in Economics from George Mason University. Dr. Ritschel's research interests include public choice, cost analysis, and economic institutional analysis. (E-mail address: Jonathan.Ritschel@aft.edu)*

Dr. Edward D. White *is a Professor of Statistics in the Department of Mathematics and Statistics at AFIT. He received his BS in Mathematics from the University of Tampa, MAS from The Ohio State University, and PhD in Statistics from Texas A&M University. His primary research interests include statistical modeling, simulation, and data analytics. (E-mail address: Edward.White@aft.edu)*

Lt Col Clay M. Koschnick, PhD*, is an Assistant Professor of Systems Engineering in the Department of Systems Engineering and Management at AFIT. He received his BS in Operations Research from the United States Air Force Academy, his MS in Operations Research from the Georgia Institute of Technology, and his PhD in Industrial and Systems Engineering from the University of Florida. His research interests including engineering economics, decision analysis, and econometrics. (E-mail address: clay.koschnick@afit.edu)*

Dr. Brandon M. Lucas *is an Assistant Professor of Cost Analysis in the Department of Systems Engineering and Management at AFIT. He holds a BA in History from the University of Texas at Austin, a MA in International Relations and ME in Teacher Education from the University of Oklahoma, a MS in Cost Analysis from AFIT, and a PhD in Economics from George Mason University. Dr. Lucas' research interests include profit analysis, cost & economic analyses, and incentive structures. (E-mail address: Brandon.Lucas@aft.edu)*

Mr. Shawn M. Valenti*ne is an Operations Research Analyst at AFLCMC Cost Staff, Wright-Patterson Air Force Base, Ohio. He holds a BS in Actuarial Science from Ohio University and an MS in Financial Economics from Ohio University. He currently acts as the research technical lead at AFLCMC where he engages in Air Force research initiatives related to aircraft. (E-Mail address: shawn.valentine@us.af.mil)*