Journal of Cost Analysis and Parametrics, 5:52–61, 2012
ISSN: 1941-658X print
DOI: 10.1080/1941658X.2012.682943

ICEAA
www.iceaaonline.com

Comparison of Cumulative Average to Unit Learning Curves: A Monte Carlo Approach

TREVOR MILLER, AUSTIN DOWLING, DAVID YOUD,
ERIC UNGER, and EDWARD WHITE

Air Force Institute of Technology, WPAFB, Ohio

Cumulative average and unit cost learning curve methodologies dominate current learning curve theory. Both models mathematically estimate the structure of costs over time and under particular conditions. While cost estimators and industries have shown preferences for particular models, this article evaluates model performance under varying program characteristics. A Monte Carlo approach is used to perform analysis and identify the superior method for use under differing programmatic factors and conditions. Decision charts are provided to aide analysts' learning curve model selection for aircraft production and modification programs. Overall, the results indicate that the unit theory outperforms the cumulative average theory when more than 40 units exist to create a prediction learning curve or when the data presents high learning and low variation in the program; however, the cumulative average theory predicts unit costs with less error when few units to create the curve exists, low learning occurs, and high variation transpires.

Introduction

Background

Learning curves greatly impact the cost estimate of a project or program; therefore, choosing the correct learning curve proves imperative for an accurate estimate. The differences between the cumulative average and the unit theory are evaluated to determine the effect each has on cost estimates of programs. This study identifies decision points for analysts to use when working with learning curves and provides a table to aide an analyst's decision between learning curves.

Since the start of manufacturing, laborers depict a learning effect that improves their efficiency in producing a good over time. The learning effect derives from laborers becoming more efficient as they repeat a task and translates into reduced productions costs for subsequent units (Malashevitz, Williams, & Kankey, 2004). In addition to laborer learning, management also reduces unit costs through improved processes and tools. This myriad of actions leads to per unit cost reduction throughout the production cycle (Stump, 2002). This article generically refers to these cost reductions as "cost improvement." General cost improvement theory states that as the number of units produced double, the cost of producing units reduces at a steady percentage, defined as the learning curve scope (Contract Pricing Reference Guide, 2011). For example, with a learning curve slope (LCS) of 90% and as the quantity doubles from 100 to 200, the unit cost of the 200th unit will decrease to 90% of the unit cost of the 100th unit.

Address correspondence to Trevor Miller, Air Force Institute of Technology, 2950 Hobson Way, WPAFB, OH 45433. E-mail: tmill88@gmail.com

Today, there are many learning curve models, including Pegels's exponential function, Levy's adaptation function, and the Stanford-B model (Belkaoui, 1986). Pegels's exponential function uses an exponent to develop its learning curve, while the Stanford-B model adds a "B-factor" to the learning curve equation in an attempt to quantify program difficulties beyond managerial control (Belkaoui, 1986). While many different models like these are available, two log-linear models, unit cost and cumulative average, remain predominant in learning curve theory. As demonstrated in the following section, log-linear models, derived from early attempts to quantify learning, use a power function to derive their learning curves (Belkaoui, 1986). The cumulative average model, developed by Wright in 1936, while working in the airline industry, initially measures the learning effect (Jensen, 1991). This model combines each unit cost with the cost of the prior unit(s) to arrive at an average cost per unit produced (Defense Acquisition University, 2008, p. 6). The averaging causes smoothing of significant cost variances between units and thus is less susceptible to the effects of these variances when estimating the cost of subsequent units (Malashevitz et al., 2004).

Crawford developed the second primary methodology, the unit cost model. Crawford, researching for the Stanford Research Institute, developed the unit cost model while updating the cumulative average model using World War II aircraft production (Contract Pricing Reference Guide, 1996). Rather than using an average cost of all units produced, the unit cost model provides a unique cost for each unit. This method provides no way of smoothing significant cost variations between units and, thus, works better for production of items with little expected variation (Malashevitz et al., 2004). This method provides direct attention to the variation within data of past production units. While the two models have differing approaches, they provide similar results.

Given equal starting costs and LCSs, unit cost and cumulative average curves reveal similar shapes, as shown in Figure 1, but the unit cost model realizes learning, and thus cost savings, more quickly than the cumulative average model (Shea & Thomson, 1994). Different studies debate the merits and superiority of learning curve methods; Jensen (1991) found that each model can mirror the other's results given specific input parameters. Additional studies indicate that each methodology provides distinct benefits; generally, the unit cost method provides better results when comparing specific production units or when the data exists in a per unit basis. When data exists in historical blocks, the cumulative average method eases the implementation process (Stump, 2002). Stump's assertions highlighted that neither model proves superior under all circumstances.

While neither learning curve method establishes superiority, both provide similar and repeatable results. For improved accuracy, Anderson (2003) stipulated the same use of a

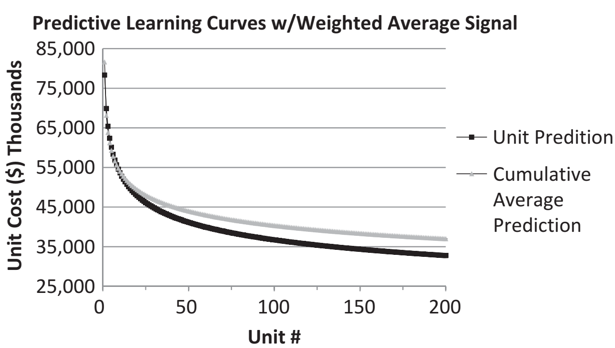

FIGURE 1 Unit and cumulative average predition learning curves versus generated data.

learning curve method throughout the program's lifecycle. While maintaining consistency provides accurate results, the goal is to identify programmatic qualities where one method is preferable to the other. A further goal is to develop a guideline for cost estimators to use when choosing the learning curve model with the lowest mean absolute percent error (MAPE), which is described further in the Methodology section, given the particular characteristics of the acquisition program.

Objective

The present research focuses on identifying the program characteristics, or factors, under which one learning curve method is superior to the other. This article defines superior as providing more accurate results more than half the time with a 5% error margin; therefore, a learning theory must perform better 55% of the time. This 5% margin of error enables confident distinguishing between performances of the two methods. Using this criterion, decision charts are developed to aid analysts' choice between learning curve methodologies under differing conditions. To generate the analysis, enough data must be obtained; however, insufficient real-world data exists of large acquisition programs to develop the analysis. This problem is solved by generating synthetic data that mimic possible real-world program characteristics.

Generating synthetic data that mirror reality requires real-world data. The supplied data identify the distribution of errors between each model's predicted values and the actual values of the real-world program. This error distribution is used to verify that the synthetic data reflects actual data. One hundred forty-six data points were obtained from seven Air Force aircraft production and modification programs from Air Force Material Command, ASC/FMCE. These programs include: C-5 AMP, B-2 Link 16, B-2 UHF SATCOM, B-2 RMP, F/A-18E/F, C-17A, and F-22, totaling 146 data points. The use of these data is described further in the next section.

Methodology

Overview

Generated data and an error distribution from previously mentioned aircraft programs are combined to evaluate the performance of the unit and cumulative average learning theories using Monte Carlo simulation methods. By controlling for the three influencing factors (LCS, overall variance, and number of actual used to produce a learning curve), the circumstances in which one theory performs better (contains less error) than the other theory are determined. The decision criteria are based on the percent of time within the simulation that one theory outperforms the other theory. After conducting the simulations, a table is created as a guideline for reference to a reader or analyst.

Data Generation

Synthetic data are created to represent actual cost of units produced, which represent a combination of two learning curve formulas, unit and cumulative average, which are manipulated through simulation. It is found that when using data generated from either a cumulative average curve or a unit cost curve, even with proper error distribution, the data biases the decision toward the curve from the synthetic generation. To attenuate this issue, the two formulas are evenly weighted to create unbiased data points, and control factors are changed to adjust the curve to match actual data. To calculate the unit cost learning curve, the following equation, found in Stump (2002), is used:

$$Unit\ Cost = First\ Unit\ Cost * \ Unit\ \#^{-\ln(LCS)}.$$

Leaving all other factors of the learning curve generations random, the LCS is changed. When calculating the cumulative average slope, the following equation, also found in Stump (2002), is used:

$$Avg.\ Cost\ First\ Unit\ Cost * \left((Unit\ \#)^{-\ln(LCS_c)+1} - (Unit\ \# - 1)^{-\ln(LCS_c)}\right).$$

To calculate the LCS for the cumulative average curve, the values from the unit portion of the data are used to ensure that the LCS values match the cumulative average portion of the generation.

Error Generation

As previously addressed, with access to unit cost data for U.S. Air Force aircraft, the actual data points provided are used to calculate the error distribution of the weighted average of the unit and cumulative average data points. The error distribution is determined by using the previously mentioned synthetic data to predict the actual data and calculate the residuals from the prediction. Once the residuals are calculated, the data are standardized using $\frac{(X_i - \bar{X})}{S_x}$, where X_i is the individual residual of the prediction from the actual, $\bar{X}$ is the mean residual of the prediction from the actual, and S_x represents the sample standard deviation of the residuals.

With the standardized data, the residuals from different aircraft programs are compared. Using the weighted average equation, the values are predicted for the actual data. After calculating the predicted values, the residuals of the different aircraft programs are standardized. Crystal Ball's® (2011) Fit Distribution command is used to determine the distribution to model the error as accurately as possible. The Fit Distribution command uses the Kolmogorov-Smirnov (KS), Anderson-Darling, and chi-squared tests, as well as others, and compares the test statistics of each model for each test. The Fit Distribution command chooses the model that produces the greatest amount of the best test statistic values for the varying tests to generate the standardized error. This distribution creates the individual error values for our simulation.

Multiplying each of the individual errors by a constant, the amount of the error is controlled, depending on the simulation conditions, to create the individual error. To ensure constant variance, the simulation's individual error is multiplied by the previous unit cost, which creates a constant percentage difference throughout the units produced. For example, given these conditions: first unit cost of $100,000, individual error value for the second unit of 0.5, and an overall variance factor of 0.1, the error amount for the next unit equates to $5000. Because the individual error takes different values for each unit based upon the error distribution, one unit's error could be positive, while the next could be negative.

Determining Decision Factors

After generating the data and applying noise, or error, to the data, analysis is made of which factors impact the decision between the cumulative average and unit cost models. To choose between the cumulative average and unit cost methods, the values predicted for both and their performance on the MAPE is judged, described by $\sum_i^n \left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| * 100$, where Y_i is the individual unit's cost, and $\hat{Y}_i$ is the individual prediction of the unit's cost with either the unit or cumulative average theory. The MAPE measures the absolute percent error of a prediction to the actual data. The following factors are considered: first unit cost, number of actual data points used to generate the learning curve, overall variance (amount of error), total number of units produced, and the LCS of the data.

To preliminarily determine the factors, a minimum of 5000 simulations is run, holding one of the previously mentioned factors constant, at different levels, while other factors remain random. By holding one factor constant, the level of impact is determined on the percent of time the cumulative average outperforms the unit theory at different levels of that factor. This information is used to decide factors to control in the simulation, delineating factors that do not impact the decision between the cumulative average and unit theory.

The results are then graphed from the analysis of each factor to determine if changing the factor affects the decision to choose between the cumulative average and unit theory. If changing the factor changes the decision to choose between the cumulative average and unit theory by more than 5%, the factor was manipulated. For example, as shown in these results, the magnitude of the overall variance impacts the percent of time that one learning curve would be chosen over the other. If the factor does not affect the decision, determined by response changes of less than 5%, then it does not matter if the variable remains constant or is kept random. The first unit cost is modeled as a random variable to emulate different real costs of programs, but the total number of units is kept constant in order to improve the efficiency of the simulations.

Validating Error Distribution

To ensure that the synthetic data statistically approximate the actual aircraft data, the KS test is used as a validation test. The KS test determines whether data match a distribution by measuring the distances between the data and the distribution. The null hypothesis of the KS test states that the distribution of one dataset represents the distribution of the other dataset's distribution, while the alternate hypothesis assumes that the two distributions do not replicate each other. Using the KS test tests if the simulated error distribution statistically differs from the error distribution of the actual data. If the distributions do not match at the 0.05 alpha, the results from that simulation are not used. Using this test on each simulation and only analyzing the passing results ensures that simulations approximate the actual data distributions.

Simulation

To ensure that the simulations match reality, Crystal Ball® (2011) is used to conduct the Monte Carlo simulations. A range of conditions is tested with a combination of the levels within the dependent factors. Table 1 depicts the three different factors—LCS, variance factor, and units used for LCS—with four, five, and four levels, respectively, corresponding to the factors. A program with 500 units is simulated; the overall variance, LCS, and number of units within the bins given in Table 1 are controlled using a uniform distribution to create the predicted values. For the cost of the first unit, a uniform random number between \$10,000 and \$100,000,000 is generated. For example, 1 level combination contains between 5 and 10 units to create a learning curve, an overall variance between 0% and 20%, and an LCS between 80 and 85. Therefore, with 4 LCSs, 5 variance factors,

TABLE 1 Three factors—LCS, variance factor, and number of units used to create LCS—and their subsequent levels for generating the random data

LCS	Variance factor	Units used for LCS
80–85	0.0–0.2	5–10
85–90	0.2–0.4	11–20
90–95	0.4–0.6	21–30
95–99	0.6–0.8	31–40
	0.8–1.0	

and 4 units used to create a learning curve as possibilities, 80 different bins exist. Fifty thousand simulations exist for each combination, totaling 4,000,000 individual simulations, to determine the percent of time with that combination of factors that the cumulative average theory outperforms the unit cost theory.

Results

We find that three of the five factors—LCS, variance, and number of data used to generate the LCS—develop differences in the decision. The conditions are tested by fixing the variable to a certain value and keeping all other variables uniformly random, then running 50,000 simulations to determine the amount of time the cumulative average contains less error. As seen in the trend in Figure 2, the number of units to create the learning curves impacts the percentage of time that the cumulative average learning curves outperform the unit cost theory. For example, with 5 units, the cumulative average contains less error over 50% of the time; while with 100 units, the cumulative average contains less error around 15% of the time. Therefore, the number of units to produce the predicted values proves significant in method selection.

Figure 3 shows the difference in the percent of time that the cumulative average learning curves contain less errors than the unit learning curves when the amount of variation changes. A trend exists and indicates that the overall variance proves to be another important factor in determining which method best predicts costs, with a change from less than 5% to about 25% of the time that the cumulative average contains less errors. When testing the LCS, it is found that the LCS impacts the decision between the unit cost or cumulative

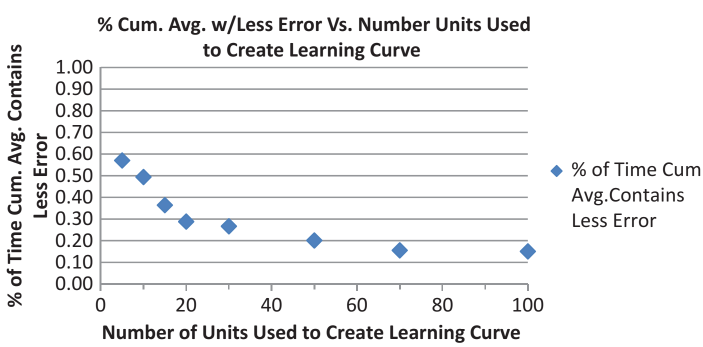

FIGURE 2 Percent of time cumulative average's MAPE is less than unit's versus number of units used to create prediction learning curve (color figure available online).

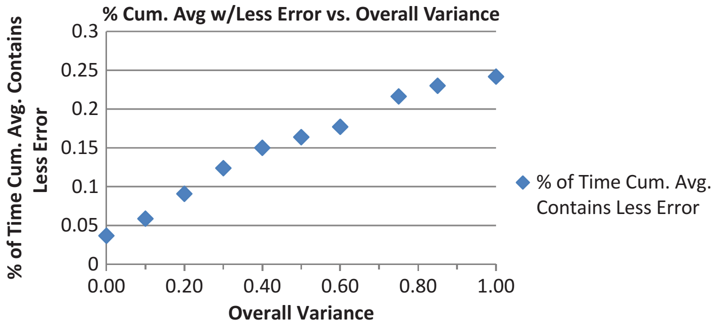

FIGURE 3 Percent of time cumulative average's MAPE is less than unit's versus overall variance (color figure available online).

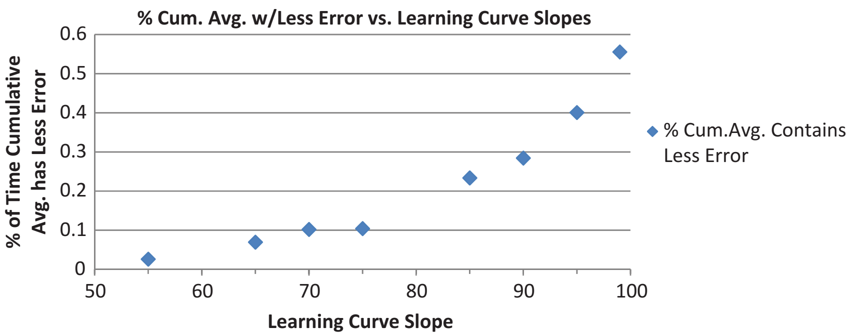

FIGURE 4 Percent of time cumulative average's MAPE is less than unit's versus LCSs (color figure available online).

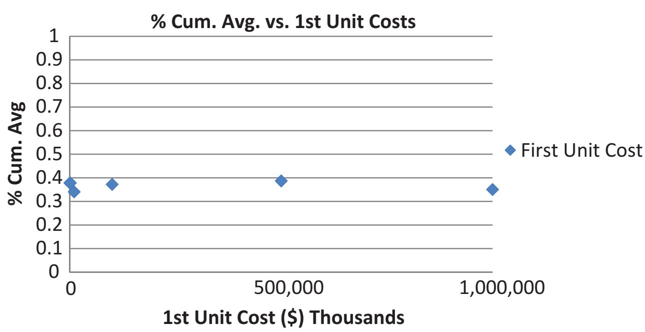

FIGURE 5 Percent of time cumulative average's MAPE is less than unit's versus first unit costs (color figure available online).

average methods, as seen in Figure 4, with a change from about 5% of the time to about 25% of the time that the cumulative average contains less error. The first unit cost and total number of units produced did not affect the amount of time that the cumulative average contains less error, as shown in Figures 5 and 6, respectively.

Results of these simulations provide the information as to which individual factors to alter in the final simulations described in the Methodology section. In general, the unit cost method provides better MAPE results when producing a greater numbers of units. The unit

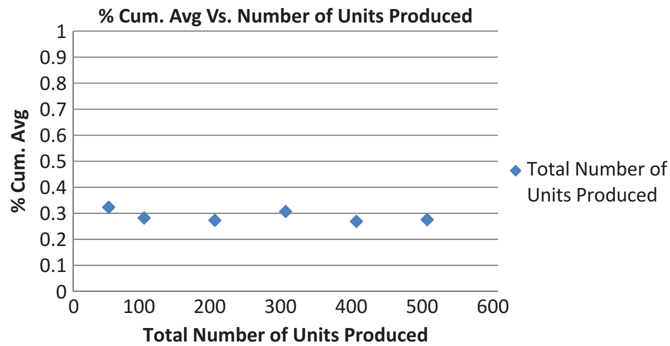

FIGURE 6 Percent of time cumulative average's MAPE is less than unit's versus number of units produced (color figure available online).

cost method also provides better results when greater learning occurs. Finally, when overall variance reaches the extremes, the unit cost method produces more accurate results.

Values in Table 2 indicate the percentage of times that the cumulative average produces a better MAPE than the unit cost method. In an effort to provide analysts an effective guideline to use when choosing between the unit cost and cumulative average methods, these charts are shaded. Dark gray indicates bins where the cumulative average percentage equals a value below 45%. White backgrounds indicate bins where the cumulative average rises above 55%. Light gray indicates bins where the cumulative average method produces superior results between 45%–55% of the time. The results from the bins can be seen in Table 2.

Conclusion

Under most conditions, the unit cost theory of learning curves is found to predict results better than the cumulative average theory. The unit cost theory performs equally or better than the cumulative average theory when the number of units increases beyond 40 units. The amount of variation in the data and the slope of the learning curve affects the decision between using the unit cost and cumulative average methodologies when the number of units falls below 40.

While other research examines the superiority of various learning curve methodologies, none perform detailed simulation research on major Air Force aircraft acquisition programs. However, Moses performed two similar studies in 1990 and 1994 for the Naval Post Graduate School. In these studies, Moses compared the accuracy (1990) and bias (1994) of log-linear methodologies with models using a production rate adjustment factor. The key similarity to the present study was that Moses identified some of the same variable programmatic factors used in this research, including, number of data points, LCS, and production quantity. While these studies generated synthetic data and highlighted the fact that many variables impact learning model selection, the studies do not examine accuracy between the unit cost and cumulative average models. An analysis is provided here of the accuracy of the unit and cumulative average learning curves new to this field of research.

Overall, results allow readers to determine which theory will provide an analyst improved knowledge upon what learning curve to use for their research or program. A user-friendly table (Table 2) was created, which allows the reader to determine which theory to use dependent on program characteristics. The "overall variance" factor from Table 2

TABLE 2 Percent of time cumulative average has a lower MAPE score under different conditions as determined by the three factors

		Overall variance				
		0–0.2	0.2–0.4	0.4–0.6	0.6–0.8	0.8–1
5–10 Actuals						
LCS	80–85	40.78	44.79	47.87	48.54	49.66
	85–90	44.94	49.70	50.31	51.92	54.21
	90–95	49.78	51.75	54.71	57.70	59.56
	95–99	53.08	59.14	62.03	62.91	63.35
11–20 Actuals						
LCS	80–85	3.16	17.03	29.24	32.62	34.71
	85–90	9.43	29.43	35.12	38.46	41.78
	90–95	18.83	37.08	42.96	47.91	51.03
	95–99	33.40	50.50	55.88	58.03	58.87
21–30 Actuals						
LCS	80–85	0.00	7.62	18.78	22.88	25.52
	85–90	3.39	18.57	25.70	29.58	33.02
	90–95	10.47	27.79	34.80	40.13	43.54
	95–99	24.96	43.66	49.72	52.12	53.08
31–40 Actuals						
LCS	80–85	0.26	4.68	13.40	17.44	19.92
	85–90	1.95	13.17	20.16	24.26	27.63
	90–95	7.27	22.39	29.82	35.01	38.35
	95–99	20.33	39.30	45.41	47.77	48.51

Dark gray—bins where cumulative average percentage equals a value below 45%, white—bins where cumulative average rises above 55%, light gray—bins where cumulative average produces superior results between 45%–55% of the time.

represents the amount of error within the data, while the “actual” factor depicts the number of units used to create the learning curve. The white cells describe conditions when the cumulative average theory produces better results, the dark gray cells describe conditions in which the unit cost theory produces improved results, and the light gray cells describe conditions that both unit cost and cumulative average produce similar results. When conditions fall within the light gray ranges, the user can choose which method they prefer. The user should understand that the cumulative average theory produces better results when the number of units remains low with high variability of the costs; conversely, the unit theory outperforms the cumulative average theory under most conditions and always when the number of units exceeds 40. This article provides analysts a tool to determine when to use which learning curve for the most accurate prediction; moreover, a more accurate prediction enables better allocation of funds.

Acknowledgment

The authors would like to thank Captain Tarah D. Mitchell for her contributions to relevant background information.

References

Anderson, T. P. (2003, January 21–28). The Trouble with Learning Curves. Presentation to the *DoD Cost Analyst Symposium*, Williamsburg, VA.

Belkaoui, A. (1986). *The Learning Curve: A Management Accounting Tool*. Westport, CT: Quorum Books.

Contract Pricing Reference Guide. (2011). Wright-Patterson AFB, OH: Air University, Air Force Institute of Technology. Retrieved from https://acc.dau.mil/CommunityBrowser.aspx?id=406579&lang=en-US

Crystal Ball®,Fusion Edition (2011). Redwood Shores, CA: Oracle, Inc.

Defense Acquisition University. (2008). Learning Curves (BCF 106: Fundamentals of Cost Analysis). Fort Belvoir, VA.

Jensen, R. (1991). Reconciliation of Learning Rates Between the Cumulative Average vs Incremental Learning Rate Models. *Scandinavian Journal of Management*, 7(2), 137–142.

Malashevitz, S., Williams, B., & Kankey, R. (2004). *Handout for Three Day Learning Curve Workshop: Unit and Cumulative Average Formulations*. Fort Belvoir, VA: Defense Acquisition University.

Moses, O. D. (1990). *Learning Curve and Rate Adjustment Models: Comparative Prediction Accuracy under Varying Conditions*. Monterey, CA: Naval Postgraduate School.

Moses, O. D. (1994). *Alternative Learning Curve Models: An Analysis of Forecast Error*. Monterey, CA: Naval Postgraduate School.

Shea, C., & Thomson, K. (1994, September). *Analysis of the Effects of Fixed Costs on Learning Curve Calculations.* (MS thesis, AFIT/GCA/LAS/94S-6). Graduate School of Engineering and Management, Air Force Institute of Technology (AU), Wright-Patterson AFB, OH.

Stump, E. (2002). All About Learning Curves. Presentation to the *2002 SCEA Conference*. Retrived from www.sceaonline.org/.../PTAS11%20Stump_All%20About%20Learning%20Curves.ppt

About the Authors

Trevor P. Miller, M.S., is a cost analyst for Air Force Space Command concentrating on SBIRS and Space Fence programs. His educational background and his research interests focus on simulation and statistical modelling.

Austin W. Dowling, M.S., is a Cost Analyst for the B-2 program with the Aeronautical Systems Center. His educational background and his research interests focus on simulation and statistical modelling.

David J. Youd, M.S., is a the Cost Chief for the Space and Missiles Systems Center's developmental planning directorate (SMC/XR). His educational background is in Accounting, Management, and Governmental Cost Estimating.

Dr. Eric Unger is the Cost Estimating Branch Chief at Space and Missile Systems Center. He received a B.A. in Mathematics and Economics from Northwestern University, an M.S. in Acquisition Management from the Air Force Institute of Technology, and a Ph.D. in Policy Analysis from the Pardee RAND Graduate School. He served previously as the Director of the Cost Analysis Graduate Program at the Air Force Institute of Technology and Chief of Cost of MILSATOM at Los Angeles Air Force Base. His research focuses on the policy impact of quantitative cost analysis.

Edward D. White, III, Ph.D., is an Associate Professor of Statistics within the Department of Mathematics and Statistics at the Air Force Institute of Technology. His teaching and research interests are in design of experiments, linear and nonlinear regression, and statistical consulting.