

ADA033667

J 12

A METHODOLOGY FOR ESTIMATING JET ENGINE COSTS
EARLY IN WEAPON SYSTEM ACQUISITION

1 AUGUST 1976

CAPT MICHAEL A. YANKE

DDC
RECEIVED
DEC 17 1976
RECEIVED
D

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A METHODOLOGY FOR ESTIMATING JET ENGINE COSTS EARLY IN WEAPON SYSTEM ACQUISITION		5. TYPE OF REPORT & PERIOD COVERED INTERNAL
7. AUTHOR(s) MICHAEL A. YANKE, CAPTAIN, USAF		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS AIR FORCE CONTRACT MANAGEMENT DIVISION ✓ STUDIES AND APPLICATIONS OFFICE KIRTLAND AIR FORCE BASE, NEW MEXICO 87117		8. CONTRACT OR GRANT NUMBER(s)
11. CONTROLLING OFFICE NAME AND ADDRESS AIR FORCE CONTRACT MANAGEMENT DIVISION STUDIES AND APPLICATIONS OFFICE KIRTLAND AIR FORCE BASE, NEW MEXICO 87117		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE 1 AUGUST 1976
		13. NUMBER OF PAGES 36 (12) 38 pgs
		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) STATEMENT A: APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) STATEMENT A: APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) JET ENGINES COST ESTIMATION CONFIDENCE INTERVALS MODELS ACQUISITION		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The Department of Defense (DOD) is deeply concerned about developing accurate initial estimates for weapon system production costs. An area of particular interest is providing estimates of future production costs for jet engines. Current parametric models used by the Air Force identify engine cost as a function of output variables. Other DOD agencies consider relating input variables as well as output variables to production costs. This study was designed to find a better way to estimate engine production costs. The → <i>over</i>		

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

results of this research include the following findings:

- (a) current Air Force cost-estimating models are operationally ineffective;
- (b) raw materials-related variables are highly correlated with cost and should be considered in developing future cost-estimation models;
- (c) statistical validation of cost models should incorporate confidence interval testing at a specified alpha level for each prediction; and
- (d) the use of confidence intervals is the correct statistical approach for developing cost estimates which may be used in decision making.

ACKNOWLEDGEMENT

This paper is derived from the unpublished masters thesis entitled, A Proposed Methodology for the Estimation of Jet Engine Costs in the Early Phases of the Weapons System Acquisition Process by Capt Rodney W. J. Mullineaux and Capt Michael A. Yanke.

RESUBMIT FOR		
WRITE	Write Section	<input checked="" type="checkbox"/>
TYPE	Staff Section	<input type="checkbox"/>
REVISIONS		<input type="checkbox"/>
BY		
AUTHORITY OR AVAILABILITY CODES		
CLASS. AND/OR SPECIAL		
A		

DDC
RECEIVED
DEC 17 1976
D

Background

Department of Defense (DOD) budgetary reviews have placed increased emphasis on obtaining the most effective weapon system per dollar invested. The anticipation of budget cuts heightens the importance of the DOD's efforts to establish and use the most accurate forecasting techniques. The rapid advancement of technology and higher labor costs have added to the difficulty of making the correct decisions about future military procurements. Over the past ten years various techniques have been developed to aid the government decision maker in the estimation of weapon system costs. The potential for sound economic decisions based upon proper cost-estimating techniques has been stressed for many years by DOD officials.

Problem Statement

The Air Force needs to develop an improved technique for estimating jet engine production costs in the early stages of a weapon system program development. Valid cost-estimating procedures early in the conceptual and validation phases of any weapon system provide the basis for an improved cost estimate of any major weapon system procurement for its total useful life to the Air Force.

The lack of proper planning and forecasting can create unexpected fiscal problems. The inability to consistently provide accurate cost-estimating information continues to strain DOD resources. The problem of inaccuracy in these estimates is not entirely the fault of any one person or agency. Voluminous cost,

design, and manufacturing data is created during the conceptual and validation phase of a jet engine. Present cost-estimating techniques involve detailed studies which require several thousand hours of effort. In spite of these efforts, the uncertainty involved in the use and precision of cost estimates remains high.

Sensitivity analysis for engine design changes and their resultant effort on preliminary engine costs estimates cannot be obtained using current Air Force methods. If a cost estimate were desired for a new developmental engine that was similar to a current production engine except for a change in materials, this estimate could not be obtained using existing Air Force estimating models.

Cost-Estimating Models

A cost estimate may be defined as a judgment regarding the future cost of an object, commodity, or some service available for use. This judgment may be arrived at formally or informally by a variety of methods. There are five generally accepted cost-estimating techniques: engineering; rates, factors and catalog prices; cost estimating relationships (CERs); specific analogies; and expert opinion. Expert opinion and specific analogy concepts are used in most cost-estimating situations. All five methods can be classified into two categories - accounting and statistical.

The primary assumption justifying the use of an engineering-accounting method is that detailed information, time, and money are available. The engineering method is a step-by-step process of compiling cost equations which reflect known relationships.

It is an intensive look at each cause and effect relationship in the total production of, for example, a jet engine. Therefore, when detailed information is available, this technique is a reasonable method for assessing total production costs.

The engineering method has several advantages:

1. It is more accurate than the CER method because the engineering method incorporates known inputs at a very detailed level. As more information is accumulated, the uncertainty of the estimate is reduced.

2. Application to separate parts of a system can be accomplished independently since the detailed data is available.

3. The engineering method facilitates tradeoff decisions among competing alternatives due to the availability of detailed cost breakdowns.

4. Since individual items can be studied independently, the engineering method allows for greater sensitivity analysis.

There are, however, several drawbacks:

1. The engineering method relies heavily on a micro or building block approach, thereby limiting its use as a preliminary costing method. Decision making in the conceptual and validation phases is predominately made on macro concepts.

2. The engineering method takes more time and is more costly than a CER method. With technological changes taking place continuously in jet engines, it could be a monumental task to create and maintain a data base which could be used by the engineering method.

3. Too many details could pose difficulty for decision makers in review and evaluation of alternative engine proposals. A capability to summarize and present meaningful information would have to be established.

4. Any subjective input information regarding unknown parameters might have a tendency to bias the reliability of the overall method.

In many cases prior technological improvements and historical experience allow previously developed weapon systems to be compared with a new system proposal. If data on performance and physical characteristics, costs, and metalurgical composition are available on older engines and the new derivative engine is not significantly different from the old, it is valid to project costs of the new engines based on the data of previous engines. Costs of engine systems may be related to physical performance or metalurgical composition through cost-estimating relationships. Parametric cost-estimating techniques do not rely on a detailed description of the inputs of the system. Historical cost experience is used to develop relationships between system characteristics and cost. Once sufficient information is available, a statistical analysis can then be used to verify the parameters as suitable CERs.

This method has several advantages:

1. CERs can be used early in the decision making process if they are developed using sound statistical methods.

2. CER methods can be readily developed by use of computer-processed data and computer models. The computer allows for a great deal of analysis of parameter inputs to be accomplished in a relatively short period of time. This facilitates the tradeoff-analysis process.

3. CERs provide a basis for constructing confidence intervals about the predicted cost.

Some disadvantages are also associated with the CER method:

1. The technique is not usually applicable to radically new systems. CERs rely on past data to allow the proposed system to fit into a class of systems preceding it, or to be similar enough to justify the CER approach with some adjustment. It is statistically unreliable to go beyond the limits of the data base used to establish the CER.

2. Since the CER method provides estimates based on specific knowledge of past data, a continuing need for updating the forecast exists. The cost of maintaining accuracy of model parameters in the CER method should never exceed the benefits realized from the use of the model.

3. The scope of the CER method is general. As more details about the system or component parts are desired, the CER method loses strength, and the more detailed knowledge required would be better obtained through the engineering method. As a weapon system progresses from the conceptual through the validation to the beginning of the full scale development phase, sufficient information has usually been accumulated which could be

used by the engineering method. If the decision making processes have not generated enough specific details about the weapon system procurement, then the CER method may still be the appropriate way to estimate costs.

Literature Review

Two Air Force agencies are currently capable of providing jet engine cost estimates as part of overall weapons system cost analysis. The Office of the Secretary of Defense Cost Analysis Improvement Group (CAIG) is the overall DOD monitor for cost estimating for weapon systems. Input to the CAIG from within the Air Force come from two primary sources: (1) the System Program Office (SPO) and (2) the Air Force Independent Cost Analysis (ICA) program. An ICA group performs cost analysis independent of the SPO for the same weapon system. Inputs are processed through the chain of command to support cost comparison and decision making at the Defense Systems Acquisition Review Council (DSARC) level. Procedures for the ICA are set forth in Air Force Regulation 173-11.

The RAND Corporation has developed the cost estimating models currently being used by the Air Force to predict jet engine costs. Efforts to improve the RAND data base by collecting information on military and commercial engines have resulted in valuable information previously unavailable. Nelson and Timson of RAND used CERs to develop an Air Force model. Through the application of the Nelson and Timson model, engine costs for the 1950's and 1960's production technology were substantiated.

A study by the Navy to develop aircraft engine costing techniques was authorized by the Naval Air Systems Command (NAVAIRSYSCOM) in late 1968. A method was used to establish CERs between price and material composition of turbofan engines. The experimental efforts of Brennan and Taylor drew considerable attention to the materials used in engine production. Another Navy cost-estimating model developed by Brennan, et al, showed high correlations of physical as well as metallurgical variables with engine costs. This approach to cost estimation was rather unique in that a concerted effort was devoted to approximating overall costs as they apply to materials and technology considerations.

Birkler, in 1975, confirmed the significance of using a model capable of considering changing technologies. The basis for this consideration in the Navy model is keyed to the Maurer Factor (MF). Initial efforts at defining material cost parameters were done by R. J. Maurer of NAVAIRSYSCOM in 1966 and 1967. The MF is defined to be:

$$MF = \Sigma(\text{Relative Weighting Factor} \times \text{Material Gross Weight})$$

The materials approach has several merits. First, it takes into account the advent of super alloys. The cost and development of these alloys have driven engine procurement costs up. Second, these alloys have forced the development of new, more sophisticated and expensive manufacturing techniques. With these alloys, engines can be made lighter, more durable to stress, and adaptable to higher temperatures. As a result, engine performance

is improved by lower fuel consumption, a wider performance envelope, and fewer maintenance repair actions. If these advantages are to be realized, the cost of the engine will have to increase accordingly. Theoretically, the materials selected for engine manufacture during the conceptual and validation phases are important cost-driving factors.

With the two approaches as a basis, a model can be developed which will improve initial cost estimates for jet engines. The model developed should be capable of being used in the conceptual and validation phases of engine development. Specific hypotheses tested are:

1. A combination of engine parameters exists which can be used to develop a statistically valid model or models that predict jet engine costs.

2. A functional relationship exists between the material composition of a jet engine and the engine cost.

Research Methodology

The research methodology was designed to develop a model which can be used to predict the cost of a military jet engine. Actual cost of the jet engine was selected as the dependent variable. Cost was adjusted to FY76 dollars using the indices in AFR 173-10. The independent variables were identified as engine performance parameters and materials used in the production of the engine.

Several independent studies have been performed that have identified major cost-driving variables. Based on studies and

interviews with Air Force and Navy industrial engineers and cost analysts, a derived table of variables was constructed (see Table 1).

A review of official government documents, technical publications and independent government-contracted research papers produced most of the information on the variables. Personal interviews with jet-engine, and cost-estimating experts at the Naval Air Development Center (NADC), the Air Force Aeronautical Systems Division (ASD), and contractor engineering personnel provided information that could not be obtained from the primary source documents. The major source of data for the dependent variable, cost and the independent variables selected was the USAF Propulsion Characteristics Summary (Airbreathing). In the Air Force, it is commonly referred to as the "Gray Book". This source was selected since it is an official Air Force document and contains most of the data needed for the study. Other sources of data include:

1. Aviation Week & Space Technology, Vol 102, No 11, March 17, 1975.
2. US Department of the Navy, Propulsion Characteristics document.
3. RAND Corporation Report, R-1288-PR, March 1974.
4. DD Form 346, Abbreviated Summary Bill of Materials (as modified by the contractor).
5. Aviation Week & Space Technology, Vol 104, No 11, March 15, 1976.

Table 1

List of Independent Variables for the Expanded Data Base

x ₁	=	Maurer Factor (MF)	x ₁₁	=	Thrust to Weight Ratio (TWR)
x ₂	=	Design Pressure Ratio/SLS (PR)	x ₁₂	=	Diameter (inches) (DIA)
x ₃	=	Max Rated Air Flow/SLS (AF)	x ₁₃	=	Thrust (maximum) (THRMAX)
x ₄	=	Afterburner (AB)	x ₁₄	=	Specific Fuel Consumption (maximum) (SFCMAX)
x ₅	=	Length (inches) (LENGTH)	x ₁₅	=	Length to Diameter Ratio (LDR)
x ₆	=	Dry Weight (WEIGHT)	x ₁₆	=	C Material (CMATRL)
x ₇	=	Thrust (cruise) (THRCR)	x ₁₇	=	Number of Turbine Stages (TS)
x ₈	=	Revolutions per Minute (cruise) (RPMCR)	x ₁₈	=	D Material (DMATRL)
x ₉	=	Specific Fuel Consumption (cruise) (SFCCR)	x ₁₉	=	Number of Compressor Stages (CS)
x ₁₀	=	Turbine Inlet Temperature/SLS (TIT)	x ₂₀	=	Turbofan (TF)
			x ₂₁	=	Revolutions per Minute (maximum) (RPMMAX)

6. Production Unit Inventory - Monetary Summary (R2800),
DO24BDT10, 31 Dec 75.

During a preliminary trial data collection effort, a major obstacle was encountered. DD Form 346 data was not available from Air Force sources. The requirement to provide this data to the Air Force is an optional determination of the contracting officer. This obstacle was partially overcome through visits to contractor and Navy sources.

The primary area of concern was to construct a model which would:

1. Not be complicated; i.e., which can be used easily by any engineer (or other cost estimator) to arrive at a cost estimate in a matter of minutes.
2. Be sensitive enough that parameter changes would detect cost variations to the satisfaction of the engineer or other cost estimator.
3. Account for technology change.

Table 2 identifies the models that were developed. Each data base was randomly divided into two groups by the computer. Eighty percent of the engines used to formulate a model were identified as the model building group. The remaining twenty percent of the engines were identified as "test engines" which were reserved to analyze the predictive capability of the model; how well the actual cost of the test engines was predicted by the model. The eighty-twenty percent split in the data base is a generally accepted data division so that sufficient data points are

Data Base	Regression Form Used	
	Linear	Exponential
RAND (all data)	Model 1	Model 2
	Model 1A	Model 2A
Expanded Data:	R2800	Gray Book
	Model 3	Model 4
Engine Materials Included	Model 7	Model 8
Engine Materials Not Included		Model 9
		Model 10

Table 2
Model Identification

selected to construct a model and check the predictive capability of the model.

The regression models identified in Table 2 were analyzed using standard statistical tests. Coefficient of determination (R^2), F-tests on the overall regression and individual independent variables and confidence intervals about values on the regression surface were used to evaluate the statistical significance and predictive capability of the models. As used in this study, model validation includes not only the normally employed F-tests, but also analysis of predictive capability using confidence intervals associated with the predicted costs of test engines. The results of an analysis of predictive capability constitute a measure of the model's utility of application.

To confirm the utility of application of a model as a cost estimator, confidence intervals about the values on the regression surface were constructed. The regression surface represents the locus of point estimates of the conditional mean values of the dependent values. The actual values of the dependent variable will not normally lie on the regression surface. Since there is a scattering of points above and below the regression surface, it is of interest to be able to ascertain whether or not this surface is indeed representative of the process underlying the data, within some specified degree of confidence. By calculating a confidence interval about the value on the regression surface for a fixed set of independent variables, the following statement can be made: There is a preset percent confidence that a definable

range about the predicted value on the regression surface will contain the value of the actual observation of the dependent variable for a set of given values of the independent variables.

Predictive capability is defined as a subjective determination by the decision maker as to how many actual costs he will allow to fall outside confidence intervals and still be willing to accept the model as a predictor. To improve the predictive capability of the model, it is highly desirable to reduce confidence interval half-widths. Generally, reductions in half-widths are achieved when the sample size is increased. Small samples of test data produce confidence intervals that increase rapidly as the number of test cases decreases.

For the performance of all testing which was completed, the following test criteria were selected:

1. To accept a multiple materials-related model, the first variable selected by the stepwise regression procedure must have been a materials-related variable and have an r of .80 or greater.

2. Both the model and each independent variable included must be statistically significant at the $\alpha = .05$ level. Caution must be exercised in concluding statistical significance. The following more sensitive test should be conducted: To maintain a confidence level of $(1-\alpha)$ on the overall system of simultaneous tests on the individual x_i , an equivalent value α' , must be determined for the individual tests. A conservative determination α' is obtainable from the Bonferonni Technique: $\alpha' = \alpha/(p-1)$. Clearly, as p , the number of independent variables, increases,

α' decreases. Since this is true, it would be possible to conclude an independent variable statistically significant at $1-\alpha$ when in fact it was not within the content of simultaneous tests on the x_i 's.

Data Analysis and Findings

Three models were selected for detailed analysis. All of the models which were developed were constructed in pairs: one linear and one exponential. Each pair used the same data points for model development. The same withheld engines were used to test the predictive capability and evaluate the utility of application for the models. A summary of all the models developed is given in Table 3.

Models 2A, 4 and 8 were selected for the detailed analysis. These models were selected for specific reasons. The RAND Standard Model (Model 2A) was chosen to determine its statistical validity. Models 4 and 8 which contained costs listed in the Gray Book were also chosen. Within the time span of the research effort, the R2800 cost data would not be verified as quickly as the Gray Book cost data. Linear models were chosen over power models because of consistency of variable order selection, slightly higher R^2 values, and more stability when small changes in independent variable values were introduced. The fallacy of accepting cost estimates based on a percent of actual cost rather than on a percent of predicted cost was substantiated in every model. This fallacy can easily lead to a misinterpretation of the predictive capability of a model.

Table 3
Model Summary

Model Number	Variables Included in the Model	R ²	Model F Value
1	THRMAX, PRESTM, VOL, MACH, MQT, TOA, TIT, WEIGHT, SFC, QUNTY, PSPAN	.99	40.2
1A	THRMAX, MACH	.77	24.5
2	THRMAX, PRESTM, PSPAN, VOL, MACH, TOA, MQT, WEIGHT, SFC, TIT, QUNTY	.99	39.7
2A	THRMAX, MACH	.91	71.6
3	DMATRL, THRCR, CS, THRMAX	.99	2383.4
4	CMATRL, AF, TWR, DMATRL	.99	388.4
5	(no model reported)		
6	MF, DMATRL, TWR, PR	.99	150.4
7	TIT, THRMAX, TS, SFCCR, TF, DIA, RPMCR, RPMMAX	.96	121.1
8	TIT, AF, SFCMAX, CS, RPMMAX	.91	62.3
9	PR, TIT, CS	.89	126.1
10	AF, CS, TIT, SFCMAX, RPMMAX	.90	56.6

RAND Models

Four models were constructed from the RAND data base. An attempt was made to (a) duplicate the RAND Standard Model, (b) determine if a linear model developed using the same data would produce statistically better results than the RAND Standard Model, (c) check the method used by Nelson and Timson in the development of their model to verify whether or not the models developed by their method could be statistically validated, (d) construct confidence intervals about values on the Standard Model regression surface to investigate whether or not the RAND model had any significant predictive capability, and (e) to inspect the plots of sample data residuals to investigate the validity of the assumption of homoscedasticity of error variances.

The RAND models which were developed are summarized in Table 4. In Model 1, three of the eleven variables (PRESTM, VOL, MACH) were statistically significant at the $\alpha = .05$ level. In Model 2, only one variable (TOA) was statistically significant and only at $\alpha = .10$. In Model 1A, only one of the two variables (THRMAX) was statistically significant at the $\alpha = .05$ level. In Model 2A, the RAND Standard Model, only one of the two variables (THRMAX) was statistically significant at the $\alpha = .05$ level.

The Nelson-Timson Standard Model was constructed using all 18 data points. These data points were the result of a combination of all engine dash numbers for a given engine. For example, even though there are more than 10 dash numbers for the J57

Table 4
Summary of RAND Models

Model Number	Cost Equation
1	$\begin{aligned} \text{Cost} = & -331642.9 + 10.9 (\text{THRMAX}) + 12.6 (\text{PRESTM}) + 250076.6 (\text{VOL}) \\ & + 161489.8 (\text{MACH}) - 6164.0 (\text{MOT}) + 24.2 (\text{TOA}) + 431.7 (\text{TIT}) \\ & - 68.1 (\text{WEIGHT}) - 410704.6 (\text{SFC}) + 6.8 (\text{QUNTY}) + 2672.2 (\text{PSPAN}). \end{aligned}$
1A	$\text{Cost} = -183271.9 + 33.5 (\text{THRMAX}) + 202709.2 (\text{MACH}).$
2	$\begin{aligned} \ln (\text{Cost}) = & -5.46043 - .00416 \ln (\text{THRMAX}) + .16766 \ln (\text{PRESTM}) \\ & - .25024 \ln (\text{PSPAN}) + .2114 \ln (\text{VOL}) + .42899 \ln (\text{MACH}) \\ & + 1.84114 \ln (\text{TOA}) - .78069 \ln (\text{MOT}) + .42157 \ln (\text{WEIGHT}) \\ & + .26349 \ln (\text{SFC}) - .50861 \ln (\text{TIT}) - .02135 \ln (\text{QUNTY}). \end{aligned}$
2A	$\ln (\text{Cost}) = -9.21127 + .85158 \ln (\text{THRMAX}) + .35924 \ln (\text{MACH}).$

engine, and each represents changes in cost, material composition, and technology, all are considered as one engine. None of these data points were reserved for testing the predictive capability of the model. The inclusion of all data points in the model development is not generally considered statistically acceptable. RAND used the cost of the one-thousandth engine as the dependent variable in their regression equations.

In duplicating the RAND Standard Model, it was discovered that the variables THRMAX and MACH had, for no stated reason, been chosen from among the original 11 independent variables. Model 2A above represents a duplication of the regression of cost against these two variables. Even though the overall model passes the F-test, only one of the two variables passes its individual F-test at the .05 level of significance. This implies that, although two independent variables were used in the model, only one makes a statistically significant contribution to the model's ability to account for the total variation at the $\alpha = .05$ level of significance.

The predictive capability of this model was analyzed with confidence intervals to determine whether the model possessed any utility of application. No confidence intervals about the values on the regression surface were developed in the RAND study for the Standard Model. A severe deficiency in the Standard Model is evidenced by the magnitudes of its confidence interval half-widths: the half-widths of the confidence intervals ranged from a minimum of \$1,800,000 to a maximum of \$2,000,000. For

example, the JXX engine, which was used in the construction of the RAND Standard Model, actually cost \$XX,XXX (1973 dollars). The half-width of the confidence interval for its cost estimate is \$1,800,000. This finding means that, using the Standard Model, an analyst could have been ninety-five percent confident of correctly estimating the actual cost of the JXX engine with a predicted value as large as \$1,980,000.

The RAND study did not indicate whether or not a test for homoscedasticity was performed after the variables were logarithmically transformed. The constant variance property is a requirement in the development of a model through multiple linear regression analysis. If the dependent variable distribution variances are not constant after transformation, heteroscedasticity prevails. When heteroscedasticity occurs, but all other assumptions of multiple linear regression are met, the estimators obtained by the least squares procedure are still unbiased and consistent, but they are no longer of minimum variance. A visual inspection of the transformed RAND data residuals plots suggested normal distributions with constant error variances.

Expanded Data Base Models

All other models were developed from the expanded data base as defined in Table 1. Residual plots were constructed for Models 4 and 8. A visual inspection of these plots suggested normal distributions with constant error variances. The equations for the expanded data base models are summarized in Table 5.

Table 5

Summary of Expanded Data Base Models

Model Number	Cost Equation
3	$\text{Cost} = -70075.5 + 828.0 (\text{DMATRL}) - 20.5 (\text{THRCR}) + 33652.9 (\text{CS}) - 14.1 (\text{THRMAX}).$
4	$\text{Cost} = 840698.8 + 394.5 (\text{CMATRL}) - 937.0 (\text{AF}) - 102112.0 (\text{TWR}) - 75.0 (\text{DMATRL}).$
5	(no model reported).
6	$\ln (\text{Cost}) = -.6682 + 2.28528 \ln (\text{MF}) - .44925 \ln (\text{DMATRL}) - 1.91150 \ln (\text{TWR}) - 2.44975 \ln (\text{PR}).$
7	$\text{Cost} = 101727.9 + 1004.2 (\text{TIT}) + 7.3 (\text{THRMAX}) - 295363.1 (\text{TS}) - 1727715.8 (\text{SFCCR}) - 350817.3 (\text{TF}) + 10667.4 (\text{DIA}) + 338.5 (\text{RPMCR}) - 307.6 (\text{RPMMAX}).$
8	$\text{Cost} = -1773919.3 + 690.1 (\text{TIT}) + 1052.9 (\text{AF}) + 97323.8 (\text{SFCMAX}) + 25083.6 (\text{CS}) + 27.4 (\text{RPMMAX}).$
9	$\ln (\text{Cost}) = -8.97794 + 1.16755 \ln (\text{PR}) + 2.29394 \ln (\text{TIT}) + .46865 \ln (\text{CS}).$
10	$\ln (\text{Cost}) = -13.39932 + .71896 \ln (\text{AF}) + .67362 \ln (\text{CS}) + 1.89741 \ln (\text{TIT}) + .26927 \ln (\text{SFCMAX}) + .7037 \ln (\text{RPMMAX}).$

Development of the Model 4 was based on those engines for which materials data and Gray Book cost data were simultaneously available. The entire data base comprised a sample of only 12 engines. The sample was much smaller than had been desired because the optional Summary Bill of Materials (DD Form 346) is the only source for the materials-related variables' data. In spite of the small data base, emphasis was directed toward the identification of trends in the order of variable selection into the models.

The resultant equation for Model 4 can be found in Table 5. The coefficient of determination was .99679 and the Model F value 388.4. All independent variables were found to be statistically significant at the $\alpha = .05$ level. When 10 engines were used to construct the model, the costs for the remaining two engines were predicted within 25 percent of actual cost; one engine predicted within 25 percent of predicted cost. Model 4 was one of three that were selected for an in-depth analysis for the following reasons:

- a. Cost data was considered reliable.
- b. The model was considered representative of those models which contained material factors.
- c. The linear form was considered more accurate in predicting costs when small changes in independent variables were introduced.
- d. Homoscedastic conditions were confirmed by the inspection of plotted residuals.

The predictive capability of Model 4 met the criteria set forth, namely that the model be capable of providing cost estimates within 25 percent of actual cost. On the surface the model appeared to satisfy all hypotheses requirements. The high R^2 indicates that most of the total variation is accounted for, and the F ratios of the model and of each independent variable were significant at the $\alpha = .05$ level. The model also demonstrated the importance of materials. Two of the four independent variables are materials-related. For example, "C" material alone accounted for over 75 percent of the explained variation.

A 95 percent confidence interval was constructed about the values on the regression surface for each of the two test engines. The half-widths of the confidence intervals associated with Model 4 exceeded \$5,000,000. The widths of these confidence intervals parallel those obtained from Model 2A. The small data base and large confidence intervals of Model 4 suggest that caution should be exercised when using any model constructed under these conditions.

Model 8 was developed from a data base consisting of 49 engines. The model was then tested with 12 engines. Materials-related independent variable data were not available for these models.

The resultant equation for Model 8 can be found in Table 5. The coefficient of determination for Model 8 was .90955 and the Model F value, 62.3. All the independent variables selected into the model were significant at the $\alpha = .05$ level. Twelve engine

costs were predicted. Not all the predicted costs met the 25 percent criteria. Model 8 was chosen for detailed analysis for the following reasons:

- a. Costs in the Gray Book were considered the most reliable data available.
- b. The model was considered to be representative of these models formulated which did not consider materials-related variables.
- c. The 49 data points which were contained in the sample were considered representative of the total population.
- d. A visual inspection of the residuals plots showed that the homoscedastic conditions were met.
- e. Twelve test engines were considered an adequate number to test the predictive capability of the model.

A 95 percent confidence interval about the values of the regression surface for all 12 test engines considered simultaneously was constructed and analyzed. Then a 95 percent confidence interval was constructed about the values of the regression surface for each of the 12 test engines individually. The differences in the half-widths of these two kinds of confidence intervals is due to a change in the α level of the "t" statistic. When the test engines are considered simultaneously, the α level for each individual test engine is α/m where "m" is the number of engines in the simultaneous system. This approach ensures that the overall system has at least an α level of significance. Thus, when the test engines were taken as a whole, the half-widths of their confidence intervals are expanded.

The Model 8 confidence interval half-widths associated with the individual test engines were found to range from \$306,000 to \$375,000. When considering the test engines as a group, the half-widths of the associated confidence intervals ranged from \$479,000 to \$587,000. A 95 percent confidence interval about a cost estimate is an interval centered at the point estimate in which an analyst is 95 percent confident the actual value is included. When the engines were considered collectively, all 12 engines' actual costs fell within the 95 percent confidence intervals about the predicted costs.

Conclusions and Recommendations

The results of the analysis show a significant finding: models can be constructed with high R^2 values, which statistically satisfy an F-test and provide tolerable estimates, but fall short of producing usable estimates, as evidenced by confidence intervals constructed about the values on the regression surfaces. Depending on the degree of certainty the model is required to maintain, the acceptance of a model should be based on the width of the confidence interval at a prespecified α level.

The RAND Standard Model is a tool to provide a cost estimate in the early stages of weapon acquisition. The model can create and nurture a false sense of security for the user who is unaware of its possible shortcomings. The findings lead to the suspicion that the Standard Model was never tested to ascertain its predictive capability. The model was built with all

the available data and as such no predictive test could be made. As indicated in Table 3, a relatively high R^2 for the Standard Model was identified. The RAND Report offers a model as an optional aid to the cost analyst. Statistical acceptance and usefulness of the Standard Model is left up to the user. Using the prespecified level of significance in this study of $\alpha = .05$, one variable, THRMAX, is easily accepted, while the other, MACH, is easily rejected as statistically significant. The concepts of individual variable significance and confidence intervals constitute important elements in model acceptance determination. Confidence intervals ranged from four to forty times the cost of an engine.

The Air Force should adopt a cost-estimating methodology for model development which includes confidence intervals and validation at a prespecified level of significance. The following reasons are given in support of this statement:

1. Current techniques are inadequate.
2. Cost-estimation methodologies have not reached the state of the art where USAF management personnel have been able to place confidence in current models.
3. The predictive capability of a model should be tested by data not used to build the model.
4. The optional approach to cost models leaves the decision maker vulnerable to costly mistakes if an individual level of confidence was not built into the model based on the user's stated criteria.

5. Confidence intervals of the current RAND Standard Model leave some doubt as to its utility of application.

6. Confidence intervals and prespecified levels of significance force the decision maker to provide guidelines to the researcher when the model is being developed.

7. The R^2 term does nothing but define the amount of variation explained by the model for the data which was used to build the model. R^2 implies nothing about the predictive capability of the model.

8. An increase in R^2 can be obtained by adding independent variables to the model. The additional variables alter the degrees of freedom associated with the overall model F value. Even though the R^2 may be high, the overall model may fail the F-test as a result of the changes in degrees of freedom. Basing model acceptance on R^2 alone may lead to a model which is not useful.

9. Each term in the model should be evaluated for statistical significance. The findings indicate that terms which do not contribute significantly to the overall explanatory power of the model can produce models that are not capable of providing reliable estimates. A term in Model 2A, MACH, is one variable which was not statistically significant at $\alpha = .05$ but was included in the model.

In the RAND Report engines were grouped by type. Even though there are more than ten different dash numbers for the J57 engine which represent changes in material composition,

technology and cost, all the dash numbers were considered together as a whole. The same type groupings were made with the J85, J79 and TF30 engines even though the cost differences within each engine class were more than ten percent, and turbine inlet temperature differences were as much as 100°F.

Lumping engine data in this manner can have an overall effect of "diluting" information about relevant variables. Aggregation of data can hide significant changes to engine parameters over its production life. Due to the number of purchased engines in various years, engine costs have been known to vary from high initially to low for some derivations to high again later on. This variation in cost is factual even when engine costs are standardized in common year dollars. Examples of this phenomena can be found in the Gray Book. Prices for jet engines are renegotiated every year. Engines undergo component improvement and modification as a result of technological changes. As such, each engine dash number is unique in itself and should be treated as such in cost modeling rather than consolidated into a family of engines. A review of engine dash number parameter data in the Gray Book substantiates this finding.

Constructing cost-estimating models must be accomplished in such a way that some data points are withheld. The withheld data is then used to construct confidence intervals around the regression surface and to test the predictive capability of the model. Confidence intervals were constructed using the RAND data. The confidence interval half-widths were unrealistically

large. The magnitude of the confidence intervals, expressed in dollars, indicates that the model's utility of application in engine cost estimation is limited.

The data base must be sufficiently large enough not only to build a model, but also to allow the withholding of enough data points to test the predictive capability of the model. All the materials models and RAND models were constructed from a small data base. Each of these models resulted in high R^2 values and was statistically significant using the F-test. However, when the models were validated to include testing of each variable in the model and predictive capability, the results were unacceptable from an operational standpoint.

Confidence intervals should be included in model validation procedures. Models 4 and 8 were both statistically acceptable using R^2 values and F-test criterion. However, when confidence intervals were constructed about the predicted costs, the degree of uncertainty attributed to a cost estimate using Model 8 was substantially smaller than that of an estimate using Model 4. This comparison identifies two requirements which should be met when building cost-estimating models. First, a large number of engines should be considered. Models 3, 4, and 6 were constructed from a data base of 12 engines or fewer, and in each model statistical significance was attained but utility of application could not be substantiated from a larger data base demonstrated substantial improvement in the predictive capability and confidence in the cost estimate. Second, materials and materials

related variables are important considerations in developing cost-estimating models for jet engines. Other studies have emphasized the importance of input characteristics, such as materials and materials-related variables, and technology. Models 3, 4 and 6 confirm this finding. The MF rationale is a potentially powerful cost-estimating parameter. Model 6 selected the MF as the most powerful variable in explaining variation. Even with the limited amount of information available on the MF and "C" and "D" materials, these variables were selected into models over conventional engine parameters such as turbine inlet temperature and maximum thrust. These materials-related models developed are the first known attempts by Air Force personnel to consider inputs rather than only outputs of an engine as cost drivers.

Models 3, 4 and 6 containing material variables are not unlike the RAND Standard Model. The similarities are in high R^2 values, a satisfactory F ratio, a small data base, and extremely large confidence intervals. There are two exceptions to this statement. First, all of the data was not used to build the expanded data base models. Second, the withhold data was used to test the predictive capability of each model. The high correlations between the materials-related variables and cost are consistent throughout Models 3, 4 and 6. Based on these results, materials-related variables should be considered as prime candidates for independent variables in future cost-estimating models for jet engines.

Of the models listed in Table 3, Model 8 is the cost-estimating model recommended for use. The findings for Model 8 indicate cost estimates can be made with a 95 percent certainty that the actual cost will be within \$375,000 of the estimate. The independent variable data used to develop this model is reliable and verifiable through more than one source such as the Air Force Gray Book and its Navy equivalent. The variables included in Model 8 allow it to be sensitive to engine design changes. Gray Book figures provided the best available cost information. Costs from this book can be verified through comparison with the most recent price negotiating memoranda.

Parametric cost-estimating models are not difficult to construct. Such models can be readily developed using computers. Specific information on desired parameters can be accumulated and exercised by means of a model. However, statistically verifiable and reliable cost-estimating models are not easy to develop. Model validation should include a way to test the predictive capability of the model.

Acceptance of cost-estimating models should not be based on predictions within a given percent of actual cost. This type of thinking presupposes actual costs are available for comparison immediately following a cost estimate. Care must be taken when mixing actual and predicted costs for the establishment of acceptance criteria.

The inconsistency of the R2800 report to reflect engine cost figures in adjusted dollars may be a problem. The effect

of current AFLC policy results in an engine inventory value that is much less than the true value of this inventory. No indices are used to update out-of-production engine costs. Costs for engines in production reflect the latest price negotiating memorandum values. This practice is very misleading, as each year the quantity, as well as the price for engines, changes.

Organization and control of aircraft jet engines in the Air Force is fragmented. Each organization involved from engine acquisition through operational support had inadequate power or stature to manage its engine program effectively. The gap in engine management appears to have begun in 1961, when the Air Material Command was split into the Air Force Systems Command and the Air Force Logistics Command. The gap continues to increase with the emergency of the System Program Office in AFSC. As a result of overlapping responsibilities and duties among Air Force agencies regarding jet engines, cost-estimation techniques have not been adequately developed or employed in the conceptual or validation phase of propulsion system acquisition.

The SPO generally controls only one engine procurement program. Decisions are made from a parochial view and dominate over any long-run life cycle cost consideration. Some recent engine management studies reviewed report this same perception. Advanced development and testing is not adequate prior to production commitment and, as a result, more costly component improvement programs must be established and maintained across command

lines. The cost of these component improvement programs could be reduced by spending more money in the research and development phases of acquisition.

The Air Force Wright Aeronautical Laboratories, an organization involved in engine acquisition from an engineering design point of view, has not been chartered nor is manned, in our judgment, to provide cost estimates. The Propulsion Branch, Naval Air Development Center, maintains a centralized cadre of personnel with engineering skills which provide cost analysis for Navy engine procurements. In the USAF, the ASD Comptroller personnel, in cooperation with the independent cost analysis groups, provide accounting-type cost estimates. Air Force parametric costing for engines is based on RAND models. By the time ASD Comptroller personnel get involved in cost analysis, it is too late for these groups to have any significant impact on procurement cost. The design has already been determined. The independent cost groups price the design specifications. The engineers, who are involved prior to design commitment, can have a significant impact on cost if they are involved in parametric cost estimation early in the conceptual and validation phases of engine acquisition.

Recommended Methodology

The model-building technique embodied in this research is a new methodology for engine cost estimating in the Air Force.

The following is the recommended algorithm for model development:

1. Using a computer program, randomly divide the data for the engines into two groups. Eighty percent of the engines are used to build the stepwise regression model. The remainder are used to test the model.

2. Using a step-wise regression program such as the one available in the SPSS package, build the model. Ensure the model meets the statistical significance criteria for each independent variable and the overall model.

3. Using the engines in the test group, test the predictive capability of the model.

4. Set a level of significance for the confidence intervals.

5. Review utility of application; accept or reject the model based on an evaluation of predictive capability and utility of application.

Bibliography

1. Mullineaux, Captain Rodney W. J., USAF, and Captain Michael A. Yanke, "A Proposed Methodology for the Estimation of Jet Engine Costs in the Early Phases of the Weapon System Acquisition Process". Unpublished masters thesis, SLSR 02-76A, School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB, Ohio, June 1976.