U.S. Coast Guard Research and Development Center
1082 Shennecossett Road, Groton, CT 06340-6096

**Report No. CG-D-19-99**

# An Analytical Cost Estimation Procedure

FINAL REPORT
JUNE 1999

This document is available to the U.S. public through the
National Technical Information Service, Springfield, VA 22161

Prepared for:

U.S. Department of Transportation
United States Coast Guard
Office of Financial Systems (G-CFS)
Washington, DC 20593-0001

# N O T I C E

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

This report does not constitute a standard, specification, or regulation.

Marc B. Mandler, Ph.D.
Technical Director
United States Coast Guard
Research & Development Center
1082 Shennecossett Road
Groton, CT 06340-6096

Technical Report Documentation Page

| | | |
|---|---|---|
| **1. Report No.** CG-D-19-99 | **2. Government Accession Number** ADA369101 | **3. Recipient's Catalog No.** |
| **4. Title and Subtitle** An Analytical Cost Estimation Procedure | | **5. Report Date** June 1999 |
| | | **6. Performing Organization Code** Project No. 9500.1 |
| **7. Author(s)** Dr. Toke Jayachandran | | **8. Performing Organization Report No.** RDC-252-99 |
| **9. Performing Organization Name and Address** Naval Postgraduate School Monterey, CA 93943-5216 | U.S. Coast Guard Research and Development Center 1082 Shennecossett Road Groton, CT 06340-6096 | **10. Work Unit No. (TRAIS)** |
| | | **11. Contract or Grant No.** MIPR DASW01-94-C-0054 |
| **12. Sponsoring Organization Name and Address** U.S. Dept of Transportation United States Coast Guard (G-CFS) Washington, DC 20593-0001 | | **13. Type of Report & Period Covered** Final, August 1998- June 1999 |
| | | **14. Sponsoring Agency Code** Commandant (G-CFS) U.S. Coast Guard Headquarters Washington, DC 20593-0001 |

**15. Supplementary Notes**

The R&D Center's technical point of contact is Mr. Clark Pritchett, 860-441-2653, email: cpritchett@rdc.uscg.mil.

**16. Abstract (MAXIMUM 200 WORDS)**

Analytical procedures that can be used to do a sensitivity analysis of a cost estimate, and to perform tradeoffs to identify input values that can reduce the total cost of a project, are described in the report. A problem that has been developed in Project Analysis and Cost Estimating PACE (the Coast Guard's cost analysis system) is used as an example for the procedures. The total project cost is approximated using a model from economics, the Cobb-Douglas function. A procedure is then developed using techniques from economics and calculus to produce the total differential, the estimated change in project cost. This approach is accurate for linear cost models and can be made accurate for non-linear models by simply adding higher order terms to the differential. A procedure is also described for estimating the parameters of the Cobb-Douglas model from data from historical cost estimates. The Cobb-Douglas model can still be used to assess the effects of small changes in parameter values on the total cost even when model parameters cannot be estimated with high precision because historical data is limited.

| | |
|---|---|
| **17. Key Words** mathematical, economics, cost, model, Cobb-Douglas, PACE | **18. Distribution Statement** This document is available to the U. S. public through the National Technical Information Springfield, VA 22161 |

| | | | |
|---|---|---|---|
| **19. Security Class (This Report)** UNCLASSIFIED | **20. Security Class (This Page)** UNCLASSIFIED | **21. No of Pages** | **22. Price** |

Form DOT F 1700.7 (8/72) Reproduction of form and completed page is authorized.

## Acknowledgment:

The author wishes to thank Mr. Jim Wilson from the Cost Analysis Research Division (CARD) of the Institute for Defense Analyses (IDA) for contributing the appendix to this report. Mr. Wilson was responsible for the design and development of PACE for the Coast Guard.

## Executive Summary:

Analytical procedures that can be used to do a sensitivity analysis of a cost estimate, and to perform tradeoffs to identify input values that can reduce the total cost of a project, are described in the report. A problem that has been developed in PACE (the Coast Guard's cost analysis system) is used as an example for the procedures. The total project cost is approximated using a model from economics, the Cobb-Douglas function. A procedure is then developed using techniques from economics and calculus to produce the total differential, the estimated change in project cost. This approach is accurate for linear cost models and can be made accurate for non-linear models by simply adding higher order terms to the differential. A procedure is also described for estimating the parameters of the Cobb-Douglas model from data from historical cost estimates. The Cobb-Douglas model can still be used to assess the effects of small changes in parameter values on the total cost even when model parameters cannot be estimated with high precision because historical data is limited.

# TABLE OF CONTENTS

INTRODUCTION ..... 1
MATHEMATICAL MODELING OF COST ..... 2
AN EXAMPLE ..... 4
STATISTICAL ESTIMATION OF PARAMETERS ..... 5
CONCLUDING REMARKS ..... 6
REFERENCES ..... 7
APPENDIX ..... A-1
Cost Factors ..... A-1
Users Project Information ..... A-2
Putting It All Together ..... A-2

[ This page intentionally left blank. ]

## INTRODUCTION

The U.S. Coast Guard has recently developed a new desktop computer-based cost analysis software system called PACE[1] (Project Analysis and Cost Estimating). It is intended for use by Coast Guard personnel to assess the cost of a new project such as the replacement of an existing hangar with a new modernized hangar, or to analyze cost issues associated with the acquisition of a new cutter. The inputs to the system are the various factors/variables that influence cost. In the case of a hangar replacement, these factors could relate to the demolition of an existing hangar, environmental cleanup, building a new hangar, including the construction costs per square foot, and the number of supervisory personnel required. It is the intent that PACE will be used as a vehicle to perform sensitivity and trade-off analyses to arrive at an "optimal minimum cost option" before requesting funding for the project. It is also expected that when PACE is fully functional, it will be used to perform many of the day-to-day budget related computations.

Coast Guard analysts will be able to use PACE to obtain preliminary information as to which of the cost factors that are within their control offer the best leverage for cost optimization. For the hangar replacement example, the Coast Guard has control over the size and arrangement of a new hangar, as well as the number and the ranks of the supervisory personnel needed. They may want to know which one of these two factors has the better potential for cost reduction. The initially assigned values for each of the controllable cost factors is decreased by a certain fixed percentage (say 10%) and the resulting change in the project cost is noted; it may be useful to store the cost outputs for future reference and analysis. Those factors that did not produce a significant cost change, i.e., the cost savings did not exceed a pre-selected minimal threshold value (say 5%) are eliminated from further consideration. The remaining variables can then be rank ordered[2] in terms of their importance for reducing cost. This is followed by an examination of the influence of certain logically selected (based on the results of the previous step or the analyst's intuition) cost factor combinations, after increasing/decreasing the assigned values by prescribed amounts and observing their impact on project cost. This step will lead to an identification of a subset of "most promising factor combinations" for a more extensive tradeoff analysis. An added advantage of this method is that the Coast Guard analysts will have at their disposal information on how the controllable variables affect cost and will be able to make cost optimization judgements on an informed basis.

Tradeoff analyses can be performed, using PACE, by changing the values assigned to the variables in the "most promising" factor combinations (selected using the process described above) and analyzing the project cost outputs to arrive at the "optimal" (minimal cost and meeting the Coast Guard needs) set of values for the cost factors. A drawback to this method is that a large number of input combinations may need to be tested in order to identify the "desired" input choices. An alternative is to develop an analytical approach for performing sensitivity and tradeoff analyses. The objective of this study is to propose a mathematical model that describes a functional relationship between the cost factors and the total cost of a project. Calculus techniques can then be applied to identify the "influence" of each of the cost factors and objectively determine the

[1] PACE is described in References 1 and 2.
[2] PACE has the capability to rank order cost elements.

"desirable input combination." A procedure for selecting and analyzing an appropriate mathematical model[3] that expresses cost as a function of the inputs is discussed below.

## MATHEMATICAL MODELING OF COST

We begin by postulating that the total cost of a project, say C, is a differentiable function of k input variables $x_1, x_2, \ldots x_k$, i.e., $C = f(x_1, x_2, \ldots x_k)$. Certain partial derivatives of this function provide important information relevant to cost analyses.

The partial derivatives (the rates of change of cost w. r. t. each $x_i$)

$$\delta C/\delta x_i, \quad i = 1, 2, \ldots k,$$

when multiplied by small changes in $x_i$, $dx_i$, are the marginal costs. They measure the effects of arbitrarily small changes in the variables $x_i$ on the total cost C.

The total differential

$$dC = (\delta C/\delta x_1)dx_1 + (\delta C/\delta x_2)dx_2 + \ldots + (\delta C/\delta x_k)dx_k$$

is a measure of the change in total cost when all of the input variables are perturbed by small amounts $dx_1, dx_2, \ldots dx_k$. Note that the differential is a linear approximation to the change in cost. In general, when the $dx_i$ are small, the differential provides a good approximation. If a higher precision is desired, terms involving second-order partial derivatives can be added to the formula; the resulting new formula is essentially a second-order Taylor polynomial approximation. Of course, the approximation can be improved even more by including third order terms.

Sometimes, one is interested in describing relative changes in cost and the inputs. This can be achieved by using what economists refer to as "elasticities" ($\varepsilon$) that bring relative changes in two quantities in relation to each other. Specifically,

$$\varepsilon_i = (dC/C) / (dx_i/x_i) = (dC/dx_i)\,(x_i/C)$$

[3] See the appendix Overview of Data Storage in PACE about the limitations of extracting a mathematical model from the databases in PACE.

is the elasticity of the $i^{th}$ variable. It represents the percent change in cost when the $i^{th}$ input variable $x_i$ is changed from its current value by a certain marginal percentage $dx_i$ , while holding the values of the other variables fixed.

Three generic type of models that are often used by cost analysts are the following.

$$\text{Linear Model:} \quad C = a_0 + a_1x_1 + a_2x_2 + \ldots + a_kx_k$$

$$\text{Exponential Model:} \quad C = e^{(a_0 + a_1x_1 + a_2x_2 + \ldots + a_kx_k)}$$

$$\text{Cobb-Douglas Power Model:} \quad C = a_0\, x_1^{a_1}\, x_2^{a_2} \ldots x_k^{a_k}, \qquad (1)$$

$$\text{where } a_1 + a_2 + \ldots + a_k = 1$$

The power model, or the Cobb-Douglas model, in particular, has been successfully used by econometricians (see reference [3]) to approximate the relationship between cost and the input variables, in several inherently different applications. The reasons for the easy adaptability of this model are that the structure of the model and the number of parameters $a_1$, $a_2$, . . . $a_k$ allow enough flexibility to fit many situations, and that it is fairly straight forward to estimate the parameters statistically, if historical data is available.

We propose the use of the Cobb Douglas equation (1) to model the total cost of a Coast Guard project as a function of the inputs. Generally, one begins by estimating the unknown parameters of the model, namely, $a_0$, $a_1$, $a_2$ . . . . $a_k$ statistically, using historical data on the input variables and the total costs from past projects (similar to the project under investigation). However, such archived data may not be readily available for many of the Coast Guard projects. Then, it is necessary to assign values to these parameters in some logical fashion. One approach could be to choose the values of $a_i$ in proportion to the relative sizes of the input variables $x_i$. As an example, if the number of input variables is three and $x_1 = 2$, $x_2 = 3$, $x_3 = 5$, the unknown parameters $a_1$, $a_2$, and $a_3$ may be assigned the values 2/10, 3/10/, 5/10, respectively. Equation (1) now becomes

$$C = a_0\, [x_1^{0.2}\, x_2^{0.3}\, x_3^{0.5}]$$

A value for the parameter $a_0$ is then chosen so as to match the expected cost C of the project. The formulas for the cost derivatives and the differential, for this particular model, work out to be:

$$\delta C/\delta x_i = a_i(C/x_i) \quad i = 1, 2,...k \qquad (2)$$

and $$dC = C^* [(a_1/x_1)dx_1 + (a_2/x_2)dx_2 +...+ (a_k/x_k)dx_k]. \qquad (3)$$

The following example adapted from the PACE User's Manual [1] demonstrates the step-by-step process for selecting a mathematical model to represent the relationship between the cost factors and the total cost for a Coast Guard project.

## AN EXAMPLE

A fictitious project to demolish an existing hangar and replace it with a new hangar is used to demonstrate the workings of PACE, in Section 8 of the PACE User's Manual [1]. It can be seen from the data in Figure 8-7 of the Manual that the major factors influencing the total cost of the project are the square footage of the existing hangar, the desired square footage for the new hangar, the number and the ranks of the Coast Guard personnel assigned to the project, and the environmental compliance requirements for the new hangar. Since the cost of the environmental cleanup is very small, this factor will not be included in this demonstration. Therefore, the three input variables and their initial values are $x_1$=5 units (50K square feet for the old hangar; a unit is taken to be 10K square feet), $x_2$=3 (30K square feet for the new hangar) and $x_3$=7 (number of personnel assigned to the project). Since the unit costs for the three variables are approximately \$100K, \$1M and \$50K respectively, $a_1$, $a_2$, $a_3$ are assigned[4] the values 2/23, 20/23 and 1/23. The appropriate Cobb-Douglas equation (1) for the above choice of input values is

$$\$4M = C = (a_0)\,(5^{2/23})\,(3^{20/23})\,(7^{1/23})\,. \qquad (4)$$

The total estimated cost for the project is shown to be \$4M (rounded off) in Figure 8.8 of the PACE Users Manual [1]. The coefficient $a_0$ is then calculated to be \$1.229M.

[4] The costs sum to \$1.15M or 23 times \$50K so the fractions are computed accordingly.

The superscripts/powers 2/23 = 0.087, 20/23 = 0.870 and 1/23 = 0.043, are now the elasticities of the three input factors. This means, for example, that a one unit (10K square feet) increase in $x_1$ (square footage of the existing hangar to demolish) will result in an increase of \$.0696M [\$4M*(2/23)*1/5] in the total cost of the project.

If the three variables are increased/decreased by, $dx_1$ = + 0.2, $dx_2$ = - 0.5 and $dx_3$ = - 0.5 units respectively, from their current values of 5, 3 and 7, the approximate reduction in cost is determined from the differential (3) to be

$$dC = \$4.0M * [(2/23)( +0.2/5) + (20/23)( -0.5/3) + (1/23)(-0.5/7)] = \sim - \$\ 0.578M.$$

For this example, the true reduction in cost can be calculated directly by plugging in the modified values of the input variables, namely, $x_1 = 5.0 + 0.2 = 5.2$, $x_2 = 3.0 – 0.5 = 2.5$ and $x_3 = 7.0 – 0.5 = 6.5$ into equation (4). The exact change in cost is \$.587M and the absolute error in the differential approximation is about 1.5% (.009/.587).

A similar calculation using percentages would show that a 1% increase in $x_1$, a 1% decrease in $x_2$ and a 10% decrease in $x_3$ would result in a change of cost of

$$(0.087)(.01) - (0.870)(.01) - (0.043)(.1) = -1.2\%$$

(a 1.2% decrease in project cost.) Note that this 1.2% decrease in project cost is independent of the values currently assigned to the three variables $x_1$, $x_2$, $x_3$. In other words, this percentage reduction in project cost is a relative change which will be the same whether the values initially assigned to $\{x_1, x_2, x_3\}$ are {5, 3, 7} or {10, 14, 25} or any other set of three numbers.

## STATISTICAL ESTIMATION OF PARAMETERS

As indicated earlier, ideally the parameters $a_1, a_2, \ldots a_k$ in the Cobb-Douglas model (1) should be estimated using historical data from projects of a similar nature. The resulting model will then be more efficacious in describing the relationship between the inputs and the cost. We describe below the procedure for estimating the parameters $a_1, a_2, \ldots a_k$ in the Cobb-Douglas model.

Suppose data is available on the values of the k input variables $x_{ij}$, i= 1, 2,...k, j = 1, 2,...N and the costs $C_j$, j = 1, 2,...N, for N previous projects. The first step is to apply a logarithmic (ln) transformation to the data to obtain

$$y_{ij} = \ln(x_{ij}) \quad \text{and} \quad Z_j = \ln(C_j).$$

The Cobb-Douglas model (1) then translates to

$$Z_j = \ln(a_0) + a_1 y_{1j} + a_2 y_{2j} + \ldots + a_k y_{kj} \quad j = 1, 2, \ldots N.$$

$$\ln(C) = \ln(a_0) + a_1 \ln(x_{1j}) + a_2 \ln(x_{2j}) + \ldots + a_k \ln(x_{kj}).$$

This modified model expresses logarithmic costs as linear functions of the input variables, also represented on a logarithmic scale. Standard "multiple regression" techniques can be applied to statistically estimate the unknown parameters $a_1$, $a_2$,...$a_k$ and $\ln(a_0)$. Some of the commonly used statistical packages such as MINITAB, SAS or EXCEL contain routines to compute these quantities. Details of the mathematical formulas and their statistical properties of the parameter estimates can be found in reference [2].

## CONCLUDING REMARKS

This report describes a procedure for selecting a mathematical model to represent the relationship between the cost of a project and the factors that affect cost. Some of the uses for such a model are: (1) to identify the most influential factors, (2) obtain numerical measures of factor influences (elasticities) and (3) to select the combination of inputs that will provide "the most beneficial project" within the budget constraints. Of course, all models are, inherently, approximations and how accurately the model represents "reality" depends on how well we are able to determine the values for the unknown parameters in the model (such as $a_0$, $a_1$, $a_2$,…$a_k$ in the Cobb-Douglas model) and also, to some extent, on the functional form of the relationship between the factors and the cost. However, we believe, the model can still be used effectively to assess the trends in the impact of small changes in the factors on the cost, even when appropriate data are not available to estimate the model parameters with high precision. The mathematical approach to cost optimization described in this report may also be used as an adjunct to PACE. First, the "best" factor combination is determined using the analytical method and then PACE may be used to adjust the factor values to improve on the approximate analytic solution.

# REFERENCES


[1] "PACE User's Manual", Version 2.0. December 1998; Institute For Defense Analyses

[2] PACE Support Documentation Release 2.0, December 1998; Institute For Defense Analyses

[3] The Practice of Econometrics - Classic and Contemporary by R. E. Berndt, Addison-Wesley,1991.

[4] An Introduction to Mathematical Economics - Methods and Applications by G.C. Archibald and R.G. Lipsey, Harper & Row, 1976.

[ This page intentionally left blank. ]

# APPENDIX

## Overview Of Data Storage in PACE

by

Jim Wilson, Institute for Defense Analyses

PACE uses cost factors and user supplied descriptive project data to estimate the costs of alternative solutions to project requirements. One needs to understand how both types of data are stored to fully understand how PACE works under the interfaces.

### Cost Factors

Cost factor data are stored as individual cost factor components such as military pay, medical support, fuel, and spare parts. The basic PACE cost factor components table has the following data for each component:

1. A component identifier (personnel pay, PCS costs, fuel)
2. user entered component values,
3. funding types for each component,
4. a cost base year for each component, and
5. a calculated component value converted to the Project's Cost Base Year.

This latter field is calculated internally by PACE using the PACE inflation data.

$$\text{InflatedFactor}_{\text{factortype, projbaseyr}} = \text{UserFactor}_{\text{factortype, fundingtype, factorbaseyr}} * \text{CompInflRate}_{\text{fundingtype, factorbaseyr, projbaseyr}}$$

This allows PACE to use cost factor component data from multiple base years yet apply the data uniformly to each project cost estimate.

Typically, there are several cost factor components that are related to a cost driver. For example, Coast Guard Standard Personnel costs are the sum of five or more individual components and facility operating costs typically have separate components for fuel, overhauls, and maintenance parts. PACE therefore has a construct

$$\text{CostFactorSet}_{\text{type}} = \Sigma\,(\text{Cost FactorComponents}_{\text{type}})$$

When cost factor components are used as a cost factor set, the value applied is the sum of the components of the set but the value can no longer be associated with a funding type and therefore cannot be inflated to express results in terms of Current or budget costs. More on this later.

## Users Project Information

Users Project information identifies and quantifies how many and what types of resources (e.g., people, facilities) are included in a project. In many cases, users are also permitted to enter a value that is used as a "percent applicable." For example, some one could enter data to represent a cutter operating a 185 days away from home port but say that they want only 25% of the costs actually included in the estimate. Project cost entries can be generally represented as:

$$\text{Cost} = \text{CostFactorSet} * \text{Quantity} * \text{Multiplier}$$

Recall that the CostFactorSet value represents the constant dollar sum of all applicable cost factor components.

Users also record whether each entry should be applied to the total estimate as:

1. A non-recurring cost (for a specified single year)
2. A recurring cost (for a specified year range)
3. A periodic cost (with a specified starting date and recurrence period)

One last level of complexity is associated with Project data when users decide to apply a percent applicability to an entire Cost Element. Each Cost Element can contain many individual cost entries. For example, a single Cost Element can contain data for several ships or many different types of personnel.

## Putting It All Together

PACE records data so that it represents cost factor components, cost factor sets, project data, and finally the overall, multi-year cost stream that represents the total estimate, year by year, at the lowest level of resolution. We discussed the two levels of storing cost data – first by user entered cost factor component and then a PACE-created table of cost factor sets. PACE also has another internal table in each project database (i.e., the Component Cost Element or CCE table) that records information exactly as users have entered data on the data entry screens. Each entry in the CCE table contains

1. The type of data entry screen PACE used to record the data.
2. The Cost Breakdown Structure line item the user assigned to the entry (if a Direct Entry cost). If the entry uses a standard or user created cost factor set, the cost breakdown structure item is recorded by cost factor component.
3. The cost factor set identifier applied (if not a cost factor set was used).
4. The <u>value of the cost factor</u> in terms of the Project Base Year price level to be used. This may be:
5. a User Entered Factor,
6. a value from a standard Cost Factor Set, or
7. a value from a User created Cost Factor Set.).
8. The <u>annual cost</u> when the model is to use a Direct Entry value rather than a calculated value.
9. The base year that applies to a Direct Entry value. The base year that applies to a standard or user created cost factor set is actually recorded by cost factor component.

10. The quantity value. This may be:
11. a User Entered value or
12. a value taken from the User Variable table.
13. The User Variable identifier when this is used as the source for the quantity.
14. The multiplier.
15. The type of recurrence (non-recurring, recurring, or periodic) that should be applied in adding annual costs in the year by year estimate table.
16. The first year that annual costs should appear in the year by year estimate table.
17. The period of recurrence if the entry has been recorded as a Periodic cost.
18. The annual cost that should appear in the year by year estimate table. This value is derived from the cost factor, quantity, and multiplier entries and the method of calculation varies based on how the user created each entry.

At any time PACE is required to display the total project cost in any form, it uses the data in the CCE table to generate a Cost By Year table. When this occurs, PACE uses the data from CCE to know what years to generate from each entry. When an entry in CCE is based on a Standard Cost Factor Set or User Created Cost Factor Set, PACE expands the cost factor value into separate values for each component of the cost factor set so that funding type information can be included in the Cost By Year data.

The initial transfer of data from the CCE table to the Cost By Year table is done in constant dollars based on the Project Base Year. After all of the entries are generated, PACE converts the data to Discounted Costs using the Project Discount Rate. Finally, PACE generates the Current or budget year costs based on the funding types associated with each entry. (Inflation data are recorded by Funding Type).

The Cost By Year table has the following fields:

1. Alternative ID
2. Cost Element ID
3. Year
4. Cost Breakdown Structure Item ID
5. Cost Factor Type
6. Component Cost Element ID
7. Cost Category (e.g., Personnel, O&M, Construction)
8. Cost SubCategory (e.g., Staffing (Personnel), Allowances (Construction))
9. Cost Frequency Type (i.e., Recurring, Non-Recurring, Periodic)
10. Funding Type (e.g., AFC 30, AFC 01)
11. Amount (Constant dollar amount)
12. Discounted Amount
13. Inflated Amount

You will note that the Cost By Year table has picked up information from the details of the cost factor components (e.g., Funding Type) but does no carry an explicit field that ties the contents of this table back to specific cost factors. The consequence is that if you want to do the sensitivity of the result on the cost of fuel, you do no t have the data in the Cost By Year table (or any other table) to do it. You would need to write a procedure that would "join" the data in Cost By Year with the data in CCE and the Cost Factor to generate a new table with all applicable information.