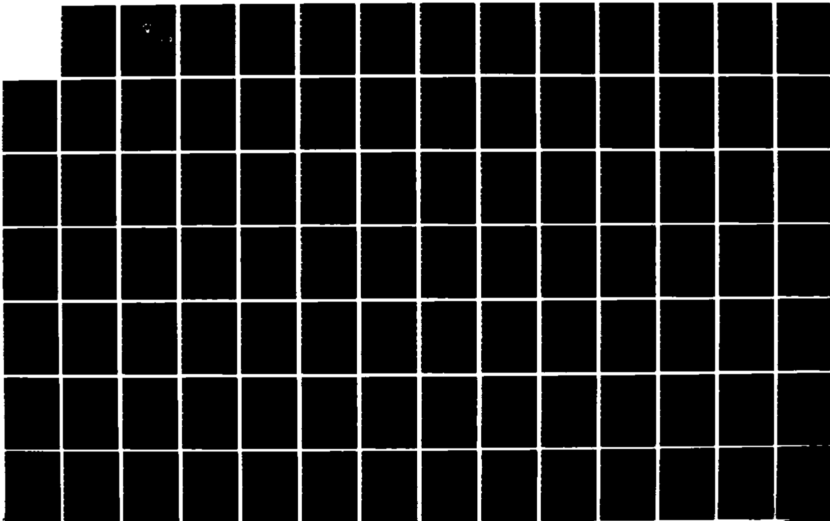
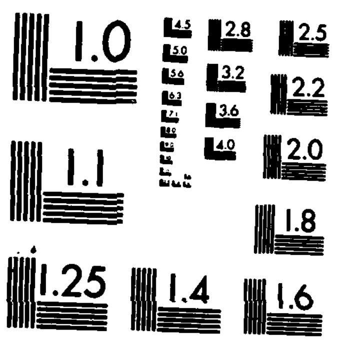
AD-A139 982 COST ESTIMATION OF NAVAL SHIP ACQUISITION(U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA I H BAIK DEC 83 1/2

UNCLASSIFIED F/G 5/3 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

2

NAVAL POSTGRADUATE SCHOOL

Monterey, California

AD A139982

THESIS

COST ESTIMATION OF
NAVAL SHIP ACQUISITION

by

Baik, In Hwa

December 1983

Thesis Advisor: Michael G. Sovereign

84 04 11 009

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A139982	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Cost Estimation of Naval Ship Acquisition		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis; December 1983
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Baik, In Hwa		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93943		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93943		12. REPORT DATE December 1983
		13. NUMBER OF PAGES 129
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Cost estimation; Acquisition; Parametric cost estimate; linear least square regression; Cost estimating relationship (CER); Correlation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The acquisition of major weapon systems is an extremely complex process involving interrelationships between a number of organizations. This thesis presents a general procedure and develops parametric cost estimate for Naval ship acquisition cost. Two different models are developed, one a 9-subsystem model, the other a single total cost model. The models were

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. (continued)

developed using the linear least squares regression technique with MINITAB statistical program on a data base of Destroyer type ships built in 1954-1966. A comparison of these two estimates with the existing RMC model's estimate was examined for Patrol Frigate construction data. The 9-subsystem estimate could be compared favorably with the RMC model cost estimate.

Approved for public release; distribution unlimited

Cost Estimation of Naval Ship Acquisition

by

Baik, In Hwa
Lieutenant Commander, R.O.K. Navy
B.S., R.O.K. Naval Academy, 1973

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
December 1983

Author: ______________________

Approved by: ______________________ Thesis Advisor

______________________ Second Reader

____
Chairman, Department of Administrative Sciences

____
Dean of Information and Policy Sciences

ABSTRACT

The acquisition of major weapon systems is an extremely complex process involving interrelationships between a number of organizations. This thesis presents a general procedure and develops parametric cost estimate for Naval ship acquisition cost. Two different models are developed, one a 9-subsystem model, the other a single total cost model. The models were developed using the linear least squares regression technique with MINITAB statistical program on a data base of Destroyer type ships built in 1954-1966. A comparison of these two estimates with the existing RMC model's estimate was examined for Patrol Frigate construction data. The 9-subsystem estimate could be compared favorably with the RMC model cost estimate.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

Cost estimation has been defined by Batchelder, [Ref. 1] as: "A judgement or opinion regarding the cost of an objective, commodity, or service." This judgement or opinion may be arrived at formally or informally by a variety of methods, all of which are reliable guides to the future. The major purpose of the cost estimation is long range planning or contract negotiation. The problem of estimating the procurement cost of major weapon systems is particularly important.

Traditionally, cost estimates for military weapons systems acquisition have been derived through Industrial Engineering (IE) techniques. These techniques are extremely time-consuming and require detailed information about the proposed equipment. In recent years estimates have been made using Cost Estimating Relationships which is defined by Baker [Ref. 2], as: "An estimate which predicts cost by means of explanatory variables such as performance characteristics, physical characteristics, and characteristics relevant to the development process, as derived from experience on logically related systems."

Cost Estimating Relationships (CERs) are mathematical equations which relate system costs to various explanatory variables. They are most generally derived through statistical regression techniques on historical cost data. There are several reasons why CERs have been and will continue to be important in the acquisition process. Early in the process when many alternative designs are contemplated, a CER based on readily available performance characteristics as explanatory variables allows the decision maker to evaluate the cost impacts of the various design and make trade-offs accordingly. Recognizing the need for and usefulness of a CER is the easy part. Developing a reliable CER is difficult at best. There are many problems the analyst must overcome in achieving this end. Identifying and collecting the data is the first and most difficult obstacle. The availability of CERs to the weapons systems acquisition process has received considerable attention in part because a reasonably large number of weapons systems have been procured since 1950 for which cost information is available. Several techniques/methods for determining an appropriate CER have been tried and are continually being improved.

This thesis's objective is to present a general procedure for development of a parametric cost estimates and to develop a model for the prediction of the total procurement cost of destroyer type naval ships that increase in precision. This thesis was limited to destroyer type ships to reduce the scope of the problem and also because of the author's experience and familiarity with this type ship.

This thesis consists of nine-chapters. A review of the general procedure for development of a parametric cost estimate is presented in Chapter III. Even though there are numerous studies about CER and several ship Cost Estimating Models, the author would like to present and contrast two sample models related to ship cost model in Chapter IV. One is the Escort Ship Cost Model (ESCOMO) developed by the Center for Naval Analysis (CNA) using performance characteristics such as maximum speed, type of weapons and sensors, endurance, range, etc. The other is RMC Cost Model developed by the Resource Management Corporation (RMC) using physical input characteristics such as weight, powerload, number of generator, etc.

Two models were to be developed and examined using the data base in Chapter VI. One model is disaggregated into nine-subsystems for estimates of total cost estimate. The other is a single total cost estimation equation. The primary criteria for comparing the predictive value of these models is the estimates of variance associated with each model. This estimates of variance is to be derived in two different ways, as discussed in Chapter VII. Finally, these two models developed in Chapter VI will be compared with the RMC MODEL to the best estimate available within the Naval Sea Systems Command.

II. APPROACHES TO COST ESTIMATION

Traditionally, weapon system cost estimates have been prepared using industrial engineering techniques. These techniques involved detailed studies of the operations and materials required to produce the new system. The cost estimate frequently required several thousand hours to produce with voluminous supporting documentation. Changes in design require extensive changes in these estimates. In spite of all the time and effort involved in preparing these estimates, there is considerable uncertainty remaining. This is evidenced by the large cost overruns cited by the annual GAO reports to Congress. Several consequences of these overruns have been:

1. A decrease in the public's confidence in the managerial ability of military leaders.

2. Acquisition of weapon systems that were not cost effective.

3. Forced reductions in the number of units purchased in order to stay under an imposed ceiling on the weapon system's acquisition cost.

4. Financial hardships experienced by military contractors in trying to meet unrealistic price estimates.

Within the last decade, a second major approach to cost estimation has come into prominence. Independent parametric cost estimation has received considerable attention in the Department of Defense as a means of increasing the accuracy of cost estimates. This procedure is based on the premise that the cost of a weapon system is related in a quantifiable way to the system's physical and performance characteristics.

Parametric cost estimates (PCE) can provide estimates during the concept formulation stage of the acquisition process before detailed engineering plans are available. These early cost estimates can be used to:

1. Identify possible cost/performance tradeoffs in the design effort.

2. Provide a base for cost/effectiveness review of performance specifications.

3. Provide information useful in the ranking of competing alternatives.

4. Suggest a need for identifying and considering new alternatives.

Historical cost data incorporate system development setbacks such as engineering and design specification changes and other items that are not identifiable at the time of design. Industrial engineering (IE) estimates tend to be optimistic in that they don't allow for unforeseen problems. Unexpected engineering or design changes usually bring about unexpected increases in system cost. Cost estimating relationships based on historical data will incorporate some of these unknowns into the cost estimate.

In the late stages of a weapon system's development, PCE's can serve as a comparison in reviewing the industrial engineer-cost estimates as they become available. Any large unexplained differences between the PCE and IE cost estimates should indicate to the decision maker that something may have been left out in either cost analysis and that further analysis may be needed. Parametric cost estimating is not intended to replace the IE estimates. It should be used along with the IE estimate to improve the accuracy of the final cost estimate.

III. METHODOLOGY FOR THE DEVELOPMENT OF A PARAMETRIC COST ESTIMATE

The methodology that follows draws heavily upon the material presented in Ref. 14 and several documents with limited distribution. Many of the ideas and techniques contained in these references are presented here in a format acceptable for unrestricted distribution.

An outline of the PCE development algorithm is presented in Figure 1. An algorithm consists of the procedural and decision steps that a cost analyst would follow in the development of the PCE. The boundaries between the individual steps are not as well defined as indicated in the figure. Several steps may be worked on simultaneously and the sequence of steps may be altered to fit the particular situation. Each step in the algorithm with its objectives, requirements, and decision criteria will be discussed separately. Steps 1 through 11 will be discussed in the development of a ship's hull subsystem CER to be used in the preparation of a PCE. Discussion of alternative methods of CER/PCE development will be presented in steps 12 through 14.

Step 1: Define the Problem and Its Objectives.

The analyst should initially strive to obtain a clear understanding of what is expected of him and the environment in which he has to work. Answer to the following questions should provide the necessary insight.

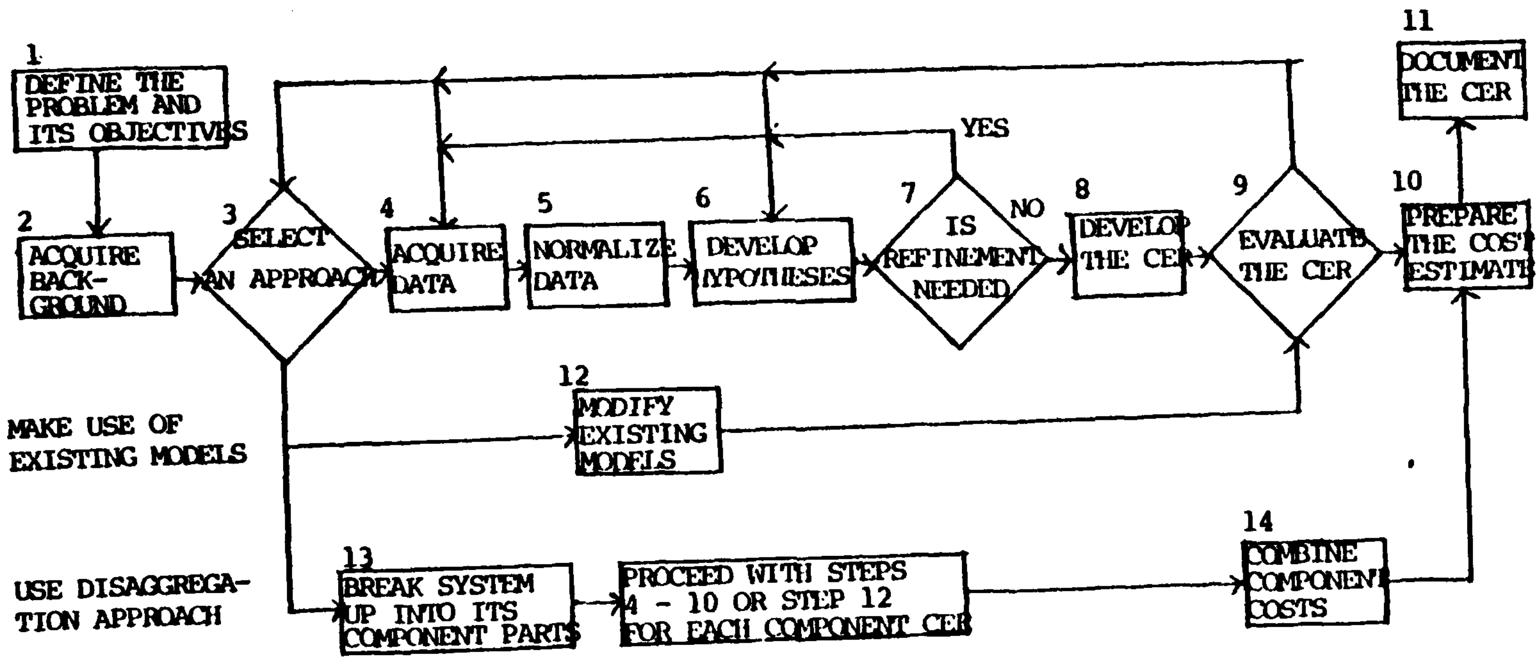


Figure 1. Parametric Cost Estimate Algorithm

1. What is the purpose of the analysis? Who is the ultimate client and what decision will be made on the basis of the analysis?

2. What is the scope of the analyst's responsibility? He should be alert for opportunities to formulate new alternatives and include these in the analysis.

3. How much time is available in which to complete the project? The amount of time available can influence both the types of data sources used and the degree of model refinement.

4. Is there any other person or agency working on the same projects? Has work been done on any similar or related projects? Often the analysts counterpart in another service or in MND will have had some experience in the project area.

5. What major sources of data and technical expertise are available? Data and associated information may be available in the project office or outside sources of information may have to be located and contacted.

6. What degree of accuracy is required in the analysis? What are the consequences of the cost estimate being too low or too high? The need for accuracy generally increases as the number of competing alternatives increases. If the analyst produces a cost estimate that errors on the high side, that particular alternative may be dropped because it seems to be too expensive. If the cost estimate errors on the low side, the prospect of a cost overrun would increase. The consequences of these possible errors were discussed in the previous chapter.

7. The analyst should be well aware of what is meant by an "independent" cost analysis. There should be open lines of communication between the independent cost analyst and the project manager. Independence does not mean that information available to the project manager should be withheld or disregarded. However, an independent attempt at evaluating such information is a necessity. Evaluation can be in the form of cross checks by using alternative techniques or information and reassessment of the unknowns and system requirements and definition.

The better the analyst prepares himself for the assigned project, the better job he will be able to do. An analyst must have not only a thorough understanding of the analysis, he must also be familiar with the system he is working on.

Step 2: Acquire Background.

> "An analyst should have a good knowledge of the kind of equipment with which he is dealing -- its characteristics, the state of its technology and the available sample." [Ref. 2]

Due to the wide diversity of weapon systems used by each of the services, an analyst will quite often not have a good technical background for the particular project he is assigned. Before the analyst is able to provide the authoritative analysis expected from him, he must have a good working knowledge of the systems involved. This is essential in order to evaluate both reference materials used and the final report produced. The two best methods of obtaining the required background are:

1. Reading texts and technical reports dealing with the subject area.
2. Consulting with technical experts in the field.

Typical questions that should be answered during this phase include:

1. What parameters are used in describing the equipment or system?

Cost estimating relationships (CER) can be divided into two major categories: input and output. Input CER's are functions of the system's input parameters such as weight, volume, density, number of component parts, operating temperatures and pressures, and in general, parameters used in the physical description of the system. Output CER's are functions of the system's output parameters. These are the parameters that are measures of the system's capabilities such as speed, operating range, payload, range of detection, etc. The input and output parameters should not be combined into one CER since problems with multicollinearity between the variables are likely to be encountered in such a model and statistical tools will become unreliable. Separate input and output CER's should be developed and this will provide the analyst with two different cost estimates for comparison.

2. What are typical values of these parameters and how and why have they changed over time?

3. What are the values of these parameters for the new system? Have any new parameters been developed or become necessary to describe it?

4. What is the current state of the art and how has this changed over the course of the equipment development?

5. What is the state of the art required for the new system? Can the new system be constructed using current technology or must new breakthroughs occur before production of the system is feasible?

6. What are the basic physical laws that determine equipment operation and what is their relationship with the system's descriptive parameters?

Once the analyst has acquired sufficient background, he will be in a position to determine:

1. Type of data to be collected.

2. Possible sources of data.

3. The kind of adjustments required to be made in the data.

Answers to several of the questions above are partially obtained through the data collection effort itself. The boundary between this step and the next is not well defined and parts of each may be done simultaneously.

Step 3: Select an Approach for the Parametric Cost Estimate Development

This step represents the first decision point that an analyst would normally encounter. The opinions available are:

1. Utilize existing system cost models.
2. Develop new cost models for the entire system.
3. Break the system up into component subsystems and use a separate model for estimating the cost of each component.

If parametric cost estimates have been developed for systems similar to the one under consideration, a search of the literature should produce the supporting documentation. Cost models have been extensively developed for ships, aircraft, and aircraft engines. Reference 11 contains a bibliography of existing cost estimateing models. Modifications of these existing models may have to be made to fit the proposed system and this process is discussed further in Step 12.

An initial data search should indicate the level to which cost data is available. If cost data is available on component parts of past systems, it may be possible to break the new system down into similar components and then estimate their cost separately. The process of breaking the systems down into its components is called disaggregation. An advantage gained by disaggregation is the likelihood of better identifying the relationships between costs and the system's parameters. Another is that some component costs are well known. The analyst should be alert for opportunities to disaggregate. Further guidance for this procedure is provided in Step 13.

Quite often the only source of cost data available is the contract price which is generally an aggregated cost of the system and related support items. This would dictate that the cost prediction models developed from it could be used only to estimate total system cost. Input and output CER's could be developed from the cost data depending on the availability of parametric data.

Step 4: Acquire Data.

> "Acquisiton of data is the process of identifying, searching out, acquiring, verifying, and recording the specific information that is of value to the analyst." [Ref. 3]

There are two basic categories of data that must be collected, each with its own unique problems.

1. Parametric data

The analyst should set down the definitions of exactly what each parameter measures. Very seldom will different data sources use identical parameter definitions. Notes should be kept of the adjustments that will have to be made so that all the data satisfies the parameter definitions. If the analyst starts with definitions and uses them as benchmarks during his collection efforts, his data collection problems will be decreased significantly.

2. Cost Data

Collection of cost data can be one of the most frustrating periods for the analyst. Chapter two of Reference 3 contains a very good summary of the complications involved in collecting cost data. The aforementioned reference, or similar Refs should be consulted prior to the initiation of cost data collection.

Data collection will constitute the largest effort in any cost analysis problem. The Cost Information Report (CIR) was established by DOD in 1966 to help alleviate the data collection problem. This reporting system was designed to collect cost and related data on major contracts for aircraft, missiles, and space programs. A newer system called Contract Cost Data Reporting (CCDR), has been instituted. In the absence of CIR type data, the analyst must resort to contract records, managerial records, or periodicals containing cost data such as the Annual Market Intelligence Reports, DPA (Defence Procurement Agency) Market Price Report.

While collecting data, the analyst should be keeping in mind the levels of accuracy and aggregation that he needs. If cost data is available down to the component level, it may be possible to proceed with a disaggregated method of cost estimating. No matter what approach is used, data collection problems can be minimized by first becoming familiar with the system's technology and second, by using consistent definitions for the cost and parametric variables.

Step 5: Normalize the Data.

Before any analysis is applied to the data, the data must be consistent and comparable. Data is normalized to decrease the effects of definitional difference, production quantity differences and yearly price changes. Price indices, learning curve factors and the definitions of the parameters are used to make the required adjustment. Listed below are several of the data adjustments often needed.

1. Cost Definition Adjustments

Different contractor accounting practices and types of contracts are the primary reasons for this type of adjustment. An analyst should state the cost definition that he wishes to use, and then adjust the data to meet his definition. It is sometimes impossible to obtain information needed for consistent adjustment. Interpretation of the final cost estimate should make allowances for this possible source of cost behavior.

2. Price Level Adjustment

It is all too apparent that inflation changes the purchasing power of the dollar dramatically. In order to compare the cost of system purchased in 1953 to the cost of a new system, the cost figures must be adjusted to "constant" dollars. The Bureau of Labor Statistics publishes many indices that can be used for this purpose. With sufficient data, specifically for the type of system being estimated this can be a very laborious process and so several general indices are available for use. The Korean Ministry of National Defense publishes a procurement index to be used for general military hardware. It is almost an impossibility to obtain an index that will remove all of the price level changes. Best results are obtained from indexes which are specialized to the type of equipment being estimated.

3. Cost Quantity Adjustments

The "Learning Curve" is a phenomenon prevalent in many industries. As the cumulative number of identical items produced doubles, the unit cost or cumulative average cost is reduced by a constant percentage. For example, in a 90% learning curve, as the cumulative output is doubled, the unit cost decreases by 90%. Here the cost of unit #5 is $5.56, the cost of unit #10 is (.9) (5.56) = $5.00. Cost curve information can be obtained from two possible sources. One source is the contractor cost records for individual units. Another source of information would be a general indstry-wide learning rate that may be published in the industry's literature.

Step 6: Develop Hypotheses.

A statistical hypothesis is an assumption about the population being sampled or the relationship between combinations of variables. Numerous statistical techniques have been developed to determine the validity of these types of hypothesis. There are two categories of hypotheses that should be developed by the analyst during his data collection effort. The first type deals with the compatibility of the different subsets and can be aggregated together into one data base. The second type of hypotheses is developed around the relationships between cost and the explanatory variables.

1. Aggregation hypotheses

Figure 2 contains a chart of ship's construction cost plotted against light ship weight of observations. It is obvious that there are five distinct subsets of data:

1) General destroyer type ship (DD)

2) Escort type ship (DE)

3) Destroyer with missile type ship (DDG)

4) Escort with missile type ship (DEG)

5) Major fleet escort with missile type ship (DLG).

The question to be considered is whether or not these five subsets of data can be aggregated into one data set to be used to construct a CER for the prediction of the cost of a new destroyer type ship. There are three possible solutions:

a. Use only the appropriate data, i.e., use only the general destroyer type ship data to predict the new system's cost.

b. Include dummy variables in the regression models to identify that subset to which the data point belonged.

c. Combine the subsets and conduct tests on the final regression model to determine if other variables in the model accounted for the difference in the weight

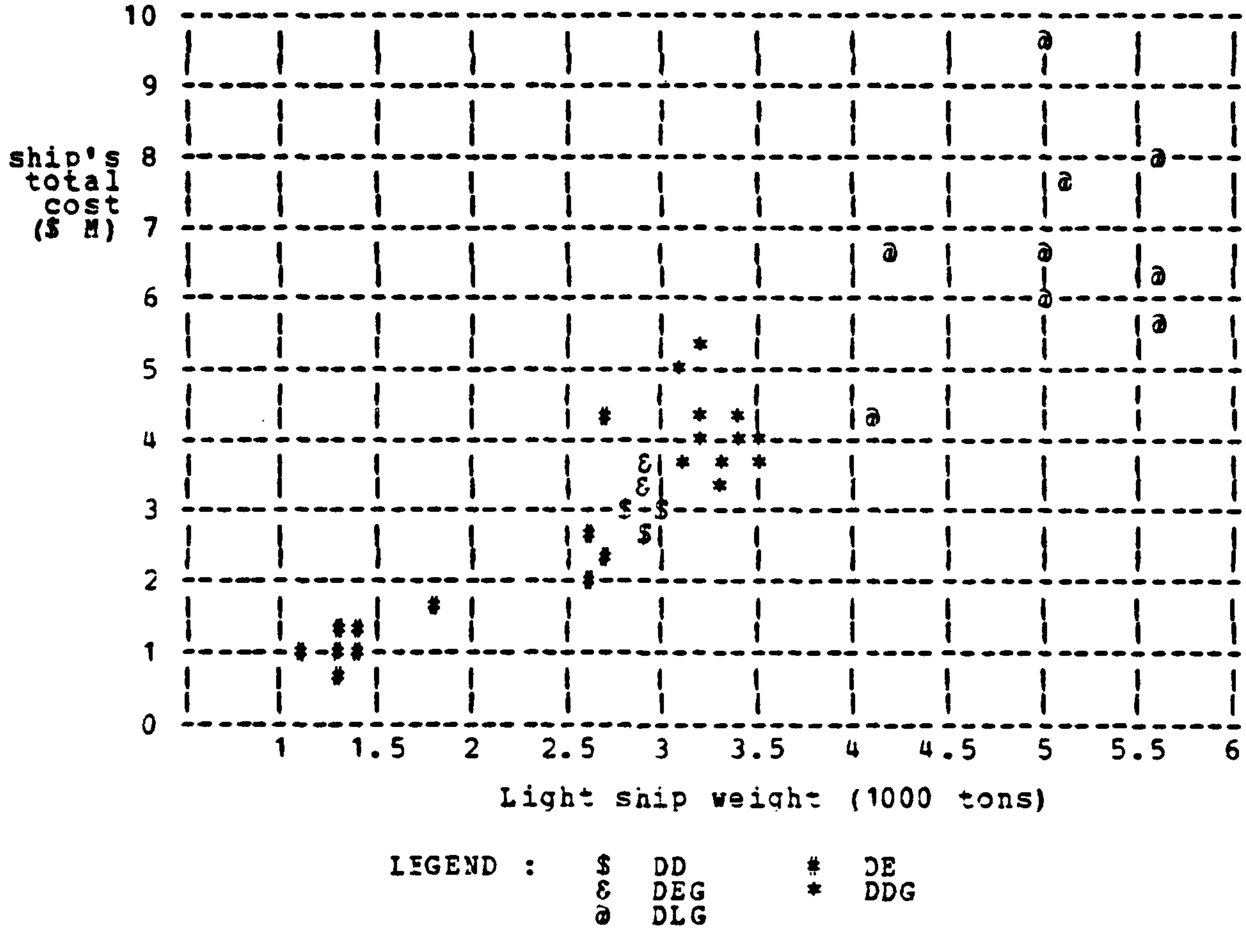


Figure 2. Ship's Total Cost vs Light Ship Weight

variable. The "Chow test" is a good method for testing the hypothesis that different subsets of data are from the same population.

2. Functional form hypotheses

The second type of hypothesis to be formed is developed around the relationships between cost and the explanatory parameters. Costs are normally expressed as a function of the independent parameters with unknown coefficients. For example, the cost of a system is normally thought to be correlated with weight. A simple hypothesis expressing this would be:

$$\text{Cost} = a + bW + e$$

a = constant term

b = unknown coefficient (cost/ton)

W = system weight

e = error term

More complex models can be developed as other relevant parameters are considered. The choice of parameters will depend on the systems' underlying technology. Parametric studies made on similar types of equipment can often suggest analogous hypotheses for consideration. Several hypothesis may be formed in this step and validity of each can be tested during subsequent steps.

Step 7: Is Refinement Needed?

After the initial data collection effort, the analyst should evaluate his data matrix. In order to use a specific parameter in a model, there should be a value of that parameter presented for each observation in the data base. The analyst may reach the conclusion that he does not have enough information in his data matrix to evaluate the hypotheses that he has formed. Possible alternatives are:

1. Collect more data to improve and enlarge the data base.

2. Limit the selection of hypotheses to make use of only those parameters for which there is a full set of observations.

3. Estimate the missing parameter values. Graybill [Ref. 10, p. 125] suggests a method of inverse estimation to provide a point estimate of the missing variable. A confidence interval for the missing variable can also be calculated. It is not known how the estimated value of the missing parameter biases the prediction abilities of the final model. Therefore, this technique should be used with care and as a means of last resort. If the analyst has the time he should strive to obtain the missing value.

After his data collection effort the analyst might find that he has enough cost data to disaggregate the system into its component parts and then estimate the cost of the individual components. If cost data is available at the component level, the analyst should proceed to steps 13 and 14, the aggregation method of system cost estimating.

Step 8: Develop the Cost Estimating Relationship

The specific analytical procedures used in the development of a CER will depend on the analyst and the computing facilities available to him. Most computer facilities will have statistical regression programs stored in the machine ready for use. The particular characteristics of a program should be studied and understood before using the program. Least squares estimation is the most commonly used method of regression analysis. References 9 and 10 are excellent sources of information on least squares procedures. Figure 3 contains the outcome of a least squares regression performed on the data in Table I. The illustrated regression line is:

$$\text{Cost} = 0.005\ 3 + 0.0013(\text{ENGPAY})$$

An analyst who obtained such a model should be concerned with the question: how well does the equation fit the data? There are several statistical measures that can give indications of the ability of the model to describe the data.

The most commonly used measure of the "Goodness of Fit" of the regression equation is the Coefficient of Determination (R^2).

$$R^2 = \frac{\text{Explained Variance}}{\text{Total Variance of the Dependent Variable}}$$

The coefficient of determination is the percentage of the variance in the data explained by the regression model. Ideally an analyst would want an R^2 to approach 1.00. An R^2 of .73 was obtained from the example regression model above. This relatively low value indicates that one independent variable, ENGPAY representing the summation of engineering and payload weight, alone does not explain all of the data. The remaining variance may be explained when other variables are considered and brought into the equation.

Figure 3 shows the relationship graphically. The solid line indicates the regression line and the dashed lines represent the standard error of estimate. The greater the dispersion of the observed values of cost about the regression line,

TABLE I

HULL SYSTEM WEIGHT AND COST DATA

Obs.	Cost	weight	Obs.	Cost	Weight
1	1.83	1902	19	1.23	769
2	2.25	1902	20	1.50	769
3	2.62	1902	21	0.78	1045
4	2.34	1946	22	1.60	1641
5	2.15	1987	23	1.30	1438
6	3.31	2027	24	1.73	1388
7	2.37	1919	25	1.55	1388
8	2.31	2027	26	1.70	1478
9	3.09	2038	27	1.65	1496
10	2.53	2027	28	3.56	2569
11	2.40	2027	29	3.06	2569
12	2.80	2027	30	4.04	2818
13	3.25	2068	31	4.21	2813
14	2.36	2035	32	5.48	2818
15	1.76	720	33	5.46	2818
16	1.20	720	34	3.47	3024
17	1.31	720	35	3.06	3039
18	1.66	720	36	4.17	3044

the less accurate the estimates that are based on that line are likely to be. If the cost data follow a Normal distribution, approximately 68 percent of the data points should fall in the area bounded by the two standard error lines. The standard error lines should not be confused with the prediction intervals constructed around point estimates. Standard error is a measure of the dispersion of the data and its relation to prediction intervals is discussed in the next step.

In comparing the standard error of one model to that of another, it is useful to compute a relative standard error of estimate. The Coefficient of variation (CV) is such a measure which relates the standard error of the mean value of the dependent variable. A value less than 20 percent for the CV is desirable.

The standard error of the model presented above is \$ 0.6045 and the coefficient of variation is:

$$CV = S.E/c = 0.6045/2.5306 = 0.238$$

$$c = \text{mean value of the dependent variable.}$$

This high value of the CV also serves as an indication that the proposed model is not well suited to the data.

In the process of constructing the CER, several models may be developed and evaluated together using the statistical measure mentioned above. The analyst should not be concerned with just maximizing R^2. There should be a logical and, if possible, a sound technologically based reason for trying a

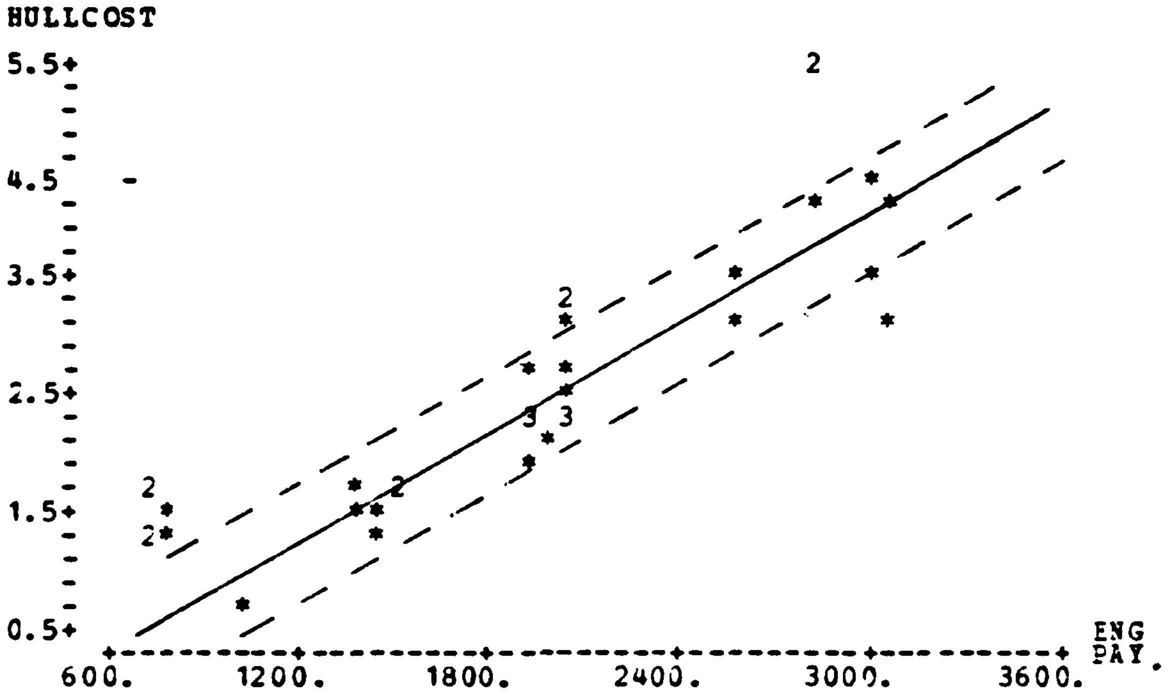


Figure 3. Hull Cost CER

particular model structure. For example, logarithmic transformation of the variable will often raise R^2, but it may also result in poorer estimates in the region of interest. The number of variables used should be restricted to those which have a logical basis and that are non-duplicating. It is best to use only input related or output related explanatory variables in one CER. There is a tendency to manipulate models to obtain a high R^2 and then later try to determine why that model produced such a high correlation. This back-order approach should be avoided since it undercuts the foundation of CER's and it can lead to serious problems when making the cost estimate for a new system.

Step 9: Evaluate the Models.

It has been previously mentioned that there are several statistical measures that can be used in evaluating a model. The coefficient of determination (R^2) and the standard error (SE) are the most commonly used measures. In addition to these, the adjusted multiple correlation coefficient should be checked. The adjusted multiple correlation coefficient is an adjustment made to the coefficient of determination for the number of degrees of freedom present in the model. If the number of degrees of freedom of the model is small, an overly optimistic picture of the performance of the explanatory variables may be obtained from R^2.

If a model contains coefficients that are not significantly different from zero, the associated variables should be dropped from the model and the model should then be determined from the significant variables. The signs and the magnitudes of the coefficients should also be studied. If cost is expected to increase with weight, then a CER containing a negative coefficient for weight would not make sense. If the CER had a large negative constant term that produced negative cost estimates for a part of the data base, then the CER would not be valid over the full range of the data base. A sensitivity analysis of the CER should be conducted to determine how the model responds to changes in the parameter values. If the model is fairly insensitive to changes in a prameter that is felt to be highly correlated with cost, then the analyst should question the suitability of the model.

There are a few hard and fast rules to be used in evaluating a model. The models' statistics must be looked at in combination since no single statistic can be a meaningful indication of the model's applicability. However, more than statistical measures are needed to analyze a CER. The analyst must satisfy himself that the model will accurately predict costs.

If the analyst is not completely confident of the model, the following may prove to be useful:

1. Recheck the definitions used for the parametric and cost data.

2. Validate any questionable data points that lie outside the expected range of values.

3. Determine if any relevant parameters have been overlooked.

4. Develop new hypotheses to be tested.

Step 10: Prepare the Cost Estimate.

The cost estimate is calculated by substituting the parameters of the proposed system into the CER. The cost figure obtained from the model is a point estimate of the actual cost and a prediction interval should be constructed around the estimate to describe the uncertainty of the estimate. When both input and output CER's are used, their point estimates and associated prediction intervals should be compared. It is very unlikely that the estimates will be the same. The interpretation and weighting of the cost estimates is up to the analyst.

In addition to evaluating the cost estimates the analyst should consider the following potential problems.

1. Estimates for systems which contain a major advance in the state of the art beyond the systems in the model's data base. The analyst should be aware that a model based on old technology may incorrectly estimate the cost of a new system containing advanced technology.

2. Very often the parametric values of the new system will lie outside the range of values contained in the data base. This requires extrapolation and faith that the model continues to be valid. If the amount of extrapolation is large, the analyst should carefully consider the possible errors inherent in the estimate.

An analyst should not blindly trust the estimate obtained from the CER. The estimate must be tempered with careful reasoning before being put into use. Cost estimating relationships are usually constructed to estimate the cost of a specified unit (first production unit, tenth production unit, etc.). The learning curve adjustments made in the normalization of the data will determine what unit cost is being estimated. Contracts normally cover the purchase of numerous identical units for a given total cost. The analyst must convert the estimated unit cost into an estimated total purchase cost. Here again learning curves play a very important role in cost estimating. The amount of learning experienced by a contractor can have a significant effect on total production cost. Total purchase cost can be easily determined using the unit cost obtained from the learning the CER and the learning curve tables contained in Reference 5. Several estimates of the total cost can be obtained by using different estimates of the learning curve slope. These estimates should be studied to observe how the total purchase cost can vary as the learning rate is changed.

Step 11: Document the Model.

The material presented in this step is taken from the documentation procedures presented in reference 14.

It is important to document a newly developed CER so that future users of the model may study it to any degree desired. Much of the material required by the guidelines given below should have been collected during the development of the model.

1. Indicate the purpose, objectives and final user of the analysis.

2. Describe the input data used and any adjustments performed on either the independent or dependent variables.

3. Identify sources and dates of the data.

4. Define each dependent and independent variable used in the analysis.

5. Provide scattergrams of the dependent variables vs. the explanatory variables used in the analysis.

6. Document the final model by including its relevant statistical information in the report.

7. Prepare a table for the final model including the observed values of the dependent variable, the estimated values and the residuals. A scattergram showing the observed costs plotted against the estimated costs should also be included.

8. List the alternatives models that were considered and the reasons why they were rejected.

9. State the major hypotheses that were formed and tested during the development of the model.

10. Provide an example to illustrate the procedure for using the final cost model.

11. Describe the limitations of the final model. Include the range of the data and any other restrictions on the population covered by the model.

The material called for by the guidelines above is a minimum of the documentation needed. An analyst should keep in mind the following principle while compiling model documentation: The model should be well enough documented so that any potential user could reconstruct the model from the information contained in the final report.

Step 12: Modify Existing Cost Estimating Models.

Considerable analysis has been performed to develop CER's for equipment such as aircraft airframes, gas turbine engines, ship hulls and related equipment. Reference 11 contains a brief summary of existing documentation available on cost estimating models. An analyst might find that he will be able to use some of the existing CER's on his current project. Use of existing models can save considerable time and effort, but they can also produce some erroneous cost estimates. The question to be considered is whether or not the existing model is completely applicable to the present project. If the existing model is not completely suitable, can the model be adjusted to reflect the changes inherent in the new system?

In order to evaluate an existing model, its documentation must be examined to determine what are the contents of the model's data base and what assumptions were made in the model's derivation. If the existing model was developed to predict the cost of airframes made out of steel and aluminum, can it be used to predict costs of an airframe that includes the use of titanium? Or, if the model is based on data for fixed wing aircraft, what adjustments would be needed to use it to predict costs for a variable geometry winged aircraft? These questions are not easily answered even if the existing model is fully documented and is understood by the analyst.

Reference 8 provides some examples of how existing CER's were modified to represent the system under consideration.

The procedures mentioned below are but a couple of the techniques that can be applied to existing CER's.

1. The control system for the new missile represented a departure from the sytems used on missiles in the data base. Weight of the control system was the explanatory cost variable in the model and it was felt that the new type of system should cost 15 percent more on a pound for pound basis than the systems in the data base. The CER was adjusted by the addition of a multiplicative term of 1.15.

   old CER: $C = a + bW$

   new CER: $C = (a + bW)\ 1.15$

   The problem with this type of approach is determining the appropriate factor to add to the CER. This factor should be based on sound opinions from experts in that particular area of technology.

2. The CER for the warhead section of the missile was also a function of weight, but the new warhead had a component in it that was not presented in the warheads in the data base. An additive term was included in the CER to reflect the use of the new component.

   old CER: $C = a + bWc$

   new CER: $C = a + bWc + d$

   It was felt that the new component's cost could be determined from other sources and its cost could be simply added to the cost obtained from the old CER. Again, competent sources of information must be utilized to provide the needed adjustment.

3. A third possible method of modifying CER's can be obtained by combining parts of existing CER's. If the old CER had the form of

   (1) $C = a + bW$

   and the CER hypothesized for the new system had another variable V in it.

   (2) $C = a + bW + cV$

the analyst could search for other CER's for similar equipment that had the variable V in it such as:

(3) $C = A + bD + cV$

If CERs 1, 2, and 3 were compatible enough in regards to their data bases and uses, the coefficient c in model 3 could be added to model 1 to produce the desired model. The problems encountered with this method deal with multicolinearity between the variables D and V, the coefficient of V will not accurately represent the actual relationship between cost and V.

The problems encountered with using existing CER's can be numerous. Using one blindly without being familiar with its development could produce cost estimates containing considerable error. The analyst must decide for himself if it would be easier to develop a new CER or spend the time and effort involved in becoming familiar with and possibly modifying an existing model. If any modifications are made, they must be based on sound technological considerations. Possibly part of the data base used in the development of the existing model could be used in the development of a new but related CER.

Step 13: Disaggregation of the System.

Quite often it is undesirable to try and estimate the total system cost with the use of just one CER. Systems may be broken down into components and then each component cost can be estimated separately. The individual component costs can then be reaggregated statistically into the system's total cost. For each component part, costs may be broken down in the following manner:

1. Initial Engineering

2. Direct Material

3. Direct Labor

4. Overhead

The exact breakdown used for a particular system would depend on the availability of the appropriate data. Some of the advantages, disadvantages and requirements of the disaggregation approach are:

1. Steps 2 through 10 or Step 12 must be used for each component cost estimate. This will require a considerable amount of time and effort on the part of the analyst to obtain his necessary data and background information.

2. Disaggregation may prove to be useful when a hypothesized CER for the total system contains many independent variables and the data base is limited in size. Each subsystem CER should require fewer and perhaps different variables than the ones required for the aggregated system.

3. The likelihood of identifying and utilizing functional forms based on technology improves as the level of disaggregation increases. Cost uncertainty of the total system can be reduced by estimating each component cost from a CER expressly.

4. Care must be taken to ensure that no parts are left out or duplicated when disaggregating the system. If disaggregation is carried too far, it will require considerable time and report and will approach the Industrial Engineering approach to cost estimation.

A cost estimating relationship will provide an analyst with a point estimate of cost. Prediction intervals developed to be placed around this point are based on the assumption that the distribution of cost follows a Normal distribution. This normality assumption leads to a distribution of costs depicted below.

When the number of data points is small, this assumption can be difficult to make. Generally, information in addition to the point estimate is available to the decision maker regarding the establishment of upper and lower bounds of the estimate. A lower bound could be the price of a similar but less sophisticated piece of equipment that is already available. An upper bound could be the maximum price that the decision maker is willing to pay for the equipment before looking for new alternatives, although such a consideration introduces additional factors. The point estimate obtained from the CER may be termed the most likely estimate and it is represented by the high point in each graph below. In addition to being able to state his high and low cost bounds, the decision maker may have some intuitive feel for the distribution of the cost estimates. Both figures in Figure 4 have the same most likely cost estimates, but the top figure displays a situation where there is a very high probability of exceeding that estimate. The lower figure illustrates just the reverse.

Quite often the Beta distribution is used to describe cost distribution because it has finite limits and an infinite variety of unimodal shapes that can be assumed. This variety of shapes can be used as figured in Figure 4 describe the particular characteristics of the estimate under consideration. The individual Beta distribution can be aggregated to provide a total cost distribution, also a Beta distribution, that incorporates the various uncertainties contained in the individual cost estimates.

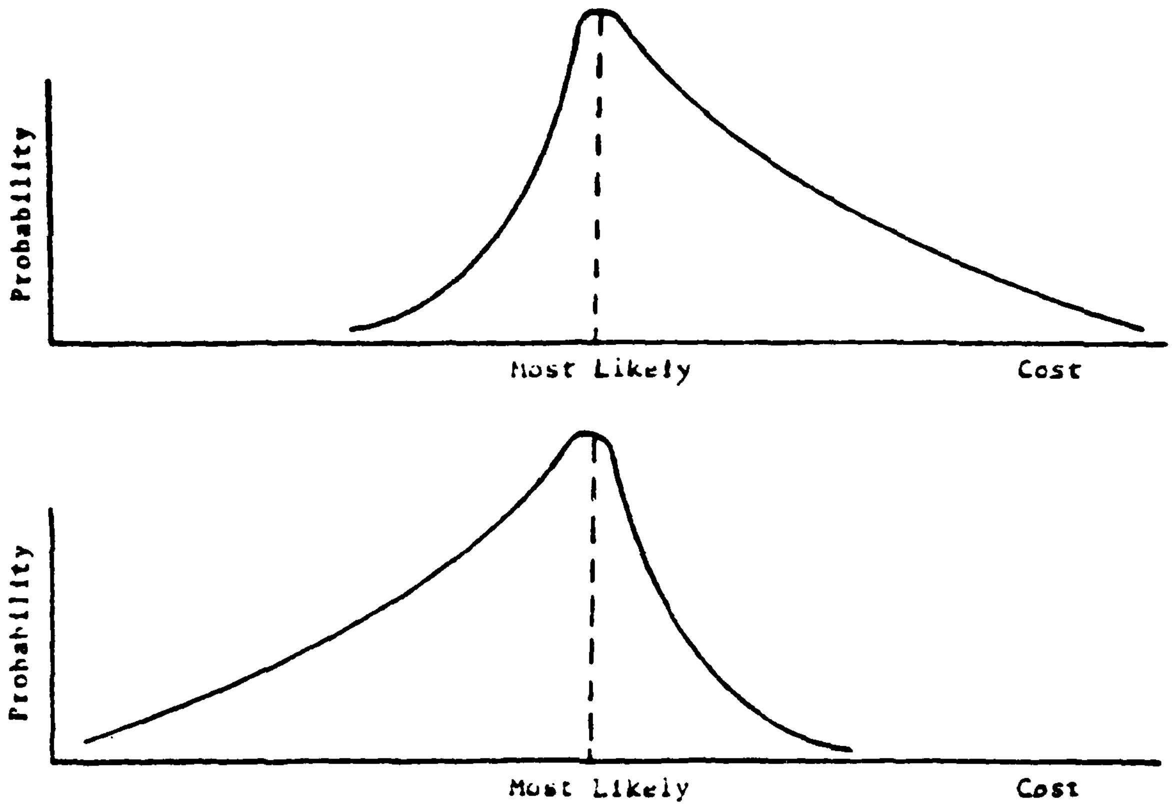


Figure 4. Beta Distribution

Reference 13 contains an excellent discussion of the procedures to follow in determining the individual Beta distributions. Reference 7 illustrates an example of how this procedure has been applied by OP-96D in preparing independent parametric cost estimates of major weapon systems.

Step 14: Aggregation of Component Costs

Step 13 described a methodology for breaking up a system into its component parts and then estimating the cost of each. In order to get a total system cost the component costs must be aggregated together. One method of doing this would be simply to add together the most likely cost for each component. This would be statistically incorrect unless the component cost distributions have certain additive properties. It would also be inappropriate because much of the information regarding the cost variances would be lost. The uncertainty in each component cost estimate has been qualified by the choice of a cost distribution. Cost distributions are likely to be quite diverse in the range of costs covered and their associated form. In order not to lose the information provided by the individual cost distributions, the cost estimates should be combined statistically using procedures such as the Summation method or the Mean square residual method which will be discussed in Chapter VII.

Summary of Methodology

The methodology for the development of a parametric cost estimate was studied from step 1 through step 14 according to the PCE development algorithm using the Ship's Hull cost data. The cost estimate of a certain system depends on analysts and given data and time, but the general method of PCE would be similar to the procedures of this paper. Several steps may be worked on simultaneously and the sequence of steps may be changed to meet the particular situation.

Step 13 suggested a method for breaking a complex system up into its components and then estimating the cost of each component in order to get a total system cost. The component cost must be aggregated together. This method was utilized in this paper.

Finally, it must be noted that the method of modifying existing models (step 12) is very dangerous. Using one blindly without being familiar with its development could result in considerable errors. Therefore, an analyst must decide himself whether to develop a new model or to modify an existing model very carefully.

IV. EXISTING SHIP COST ESTIMATING MODELS

A. ESCORT COST MODEL

1. General Description

ESCOMO is a statistically derived model produced at the Center for Naval Analysis. It is used to estimate end costs of new escort ships. The purpose of this model is to relate costs to performance characteristics of a ship substituting for the more traditional method of having costs related to physical characteristics like weight, shape, et. ... The former procedure made it quite difficult and sometimes even impossible to analyze and understand how performance affects costs and how these costs can in turn be related to desired benefits.

Research was then undertaken to derive statistical CER's between the end costs and the performance characteristics of escort ships [Ref. 6]. The definition of End Costs in this model is Total costs including Basic contract cost and Government Furnished Materials (GFM). Production characteristics were also included in the model, like quantities of ships built, the number of builders and the dates in which the ships were built.

2. Results of Analysis

In conducting an analysis, it was hypothesized that the end costs can be quantitatively related to their major performance and production characteristics. After analyzing 100 conventionally powered escorts, the natural logarithm(ln) of End Cost equation was established:

(1) LADJ$ =-0.9778 + 0.088 MAXSP + 0.57 LCRWF + 0.09 LSORNR

(12.8) (3.22) (3.84)

+ 0.0025 ORD - 0.102 LSQYD

(5.27) (-4.32)

LADJ$ = ln of end costs adjusted to 1970 Mi of dollars

MAXSP = maximum speed in knots

LCRWF = ln of crew factor (quotient of full load displacement by crew accommodations)

LSNOR = ln of sonar index

ORD = ordnance index (G&M)

LSQYD = ln of building sequence number by class within the same shipyard

The number in parentheses below the coeficients are the t-statistics of those coefficients.

F-statistics = 266.2

Multiple correlation coefficient (R^2) = 0.934

Standard error of the estimate (SE) = 0.13

Durbin Watson statistic = 2.00

Coefficient of Variation (CV) = 0.036

The values of the t-statistics, F-statistics and Durbin-Watson statistic indicate significance at the 95% level of confidence. The Durbin-Watson statistic indicates no autocorrelation in the residuals. Therefore, the hypotheses that end costs are related to certain variables that describe major performance and production characteristics in escort ships is accepted.

Even though a CER has been derived to estimate the logarithm of end cost, we are interested in predicting not the logarithm of cost, but cost itself. Several steps are required to develop an estimating relationship to predict cost. First, the antilogarithm of both sides of eq. (1) is taken to transform it to an exponential equation. The form of this equation is:

$$(2) \quad \text{expLADJ\&} = 0.37615\, e^{0.087984\ \text{MAXSP}}\ \text{CRWF}^{0.57554}\ \text{SONAR}^{0.090806}\ e^{0.0025353\ \text{ORD}}\ \text{SQYD}^{-0.10201}$$

expladJ$ = antilogarithm of LADJ$

0.37615 = antilogarithm of -0.97778

MAXSP = max speed in knots

CRWF = crew factor

SONAR = sonar index

ORD = ordnance (gun and missile) index

SQYD = building sequence by shipyard and class

As the distribution of EXP LADJ is lognormal, CER equation does not estimate the mean of the distribution of EXP LADJ$. Therefore, it must be multiplied by a corrective factor which is the antilogarithm of the quotient of the variance of LADJ$ divided by two, that is 1.1312. The equation can then be rewritten:

$$\text{CER100:} \quad ADJ\$ = 0.42550\, e^{0.087984\, MAXSP}\, CRWF^{0.57554}\, SONAR^{0.090806}\, e^{0.0025353\, ORD}\, SQYD^{-0.10201}$$

where:

ADJ$ = End cost adjusted to 1970

0.4255 = 0.376 * 1.1312

The other variables being the same as before.

CER 100 is an exponential equation that can be used to predict escort ship end costs from five explanatory variables that describe four performance characteristics and one production characteristic. This relationship was derived from statistical analysis of the costs and characteristics of 100 escort ships built for the Navy during the 1950's and 1960's. It could be used as the basis for estimating the end costs of future escorts.

B. RMC COST MODEL

1. General description

The RMC COST MODEL was developed by Resources Management Corporation (RMC) for estimating the cost of new construction in civilian shipyards (as opposed to government-owned shipyards) [Ref. 12].

The purpose of this model is to relate costs to physical characteristics like weight, shape, etc. while the ESCOMO model is to relate costs to performance characteristics of a ship. RMC approached in deriving CERs from the 13-year data base (1954-1966). The approach consisted of a stratification of the data into six groups according to ship type such as aircraft carrier, destroyer, submarine, auxiliary, amphibious, patrol/minesweeping.

RMC employed the linear least squares regression technique to develop CERs from the historical data base on ship construction costs for each ship subsystem; hull, propulsion, electrical, communication and control (C&C), auxiliary, outfitting, armament, design and engineering, construction services. The ship's characteristics as independent variables were obtained from various sources such as Ships data book, Contractor's accepted estimates, Navy contract design estimates. These variables generally consisted of characteristics that could be estimated long before ship construction began, such as subsystem weight (hull, armament, etc.) performance specifications (range, maximum speed, etc.). A complete listing of all the characteristics included in the data base by RMC may be found in Appendix B.

2. <u>Results of analysis</u>

The basic contract costs of each subsystem were then utilized as dependent variables for which CERs were developed. The established CERs are in Table II. The basic contract cost was defined as the summation of these nine cost categories plus profit and the total cost of the ship was the summation of these basic contract costs plus the cost of electronics, weapons and miscellaneous items, added after completion of basic ship construction (end-cost item).

TABLE II

9-GROUP BASIC COST CERs

(millions of dollars)

1. Hull cost

   CER Y = -0.870 + 0.00144 HULWGT + 3.794 NUC + 22.652 AR/LSW

2. Propulsion cost

   CER Y = 2.090 + 0.00640 PROWGT + 17.461 NUC - 0.0790 SERIES

3. Electrical cost

   CER Y = 0.134 + 0.283 GEN + 0.00350 ELEWGT + 2.310 NUC

4. Communication and control cost

   CER Y = 0.237 + 0.00361 C&CWGT + 1.513 NUC

5. Auxiliary cost

   CER Y = 0.09582 + 0.00176 PROWGT + 0.00295 AUXWGT

6. Outfitting cost

   CER Y = 0.150 + 0.00544 OUTWGT

7. Armament cost

   CER Y = -1.453 + 0.0068 ARMWGT + 1.151 DEDEG

8. Design and Engineering cost

   CER Y = -1.0520 + 0.00667 ARMWGT + 0.00156 PROLSW

9. Construction service cost

   CER Y = -0.01090 + 0.000241 LSW + 1.131 NUC

V. DATA BASE

One significant point should be addressed before proceeding further. It is understood that the use of contract bid data for predictive purpose is not an optimal procedure. It would be much more desirable to utilize actual ship construction costs if these costs were available. However, this is not the case. During the period from which this data was collected (1954-1966), cost accounting systems differed greatly among the various contractors, making it virtually impossible to obtain data on a uniform level of aggregation and in a manner suitable for the objectives of this thesis.

In addition, bid costs are really prices in the shipbuilding industry and are thus subject to price fluctuation. Some types of ships can only be built by certain shipyards due to the required level of expertise in electronics or weapons systems, for example. The bids on these ships could reflect a "monopoly" effect. Some shipyards are fully employed in the building of both naval and commercial ships. These shipyards might have a lower overhead, and thus produce lower bid prices, than shipyards that were not operating at full capacity. Thus, bids costs can be affected by many variables, some of which are not directly concerned with the construction costs of a specific ship.

However, contractor bid data is the most meaningful data available for the period under study (1954-1966). Using this data at least allows a preliminary effort to be made in deciding which ship's characteristics determine construction costs and within what limits of accuracy these estimates might fall.

A. DATA ADJUSTMENT

1. Bid cost data

Contract raw bid data was adjusted in three specific ways to remove cost variances due to other than ship's characteristics, as follows:

a. Cost definition adjustment.

b. Cost quantity adjustment.

c. Price level adjustment.

The price level adjustment. In this paper, the installation of Government Furnished Equipment (GFE) by the contractor required significant adjustments, especially in the propulsion category. To achieve consistency, the cost to the government of GFE was added to the appropriate cost group since the contractor's bid represented only the cost of installation and not the cost of the GFE. The cost plans supplied to the builder from an external sources was added to cost group 8, design and engineering. Again, this was done to achieve an accurate and consistent cost breakdown.

The cost quantity adjustment implies that the cost of ship construction decreases progressively with each ship in a procurement lot. The information necessary to adjust for the learning effect was derived from NAVSHIPS FORM 4282.2, UNIT PRICE ANALYSIS CONSTRUCTION, which lists contractor estimates for the nine different construction cost groups, subdivided into three categories: direct labor, direct material and overhead costs. An overall average learning curve slope was determined for all ships to apply to labor hours and material dollars for each of the 9 basic contract groups. The average learning curve slope for the data was 95.2% for 210 ships of all types (DD, AE, LSD, MSC, SSN, etc.), for which 19 bids were for 4 or more ship lots, 73 bids for 3 ship-lots, and 118 bids for 2 ship-lots [Ref. 12].

The price level adjustment refers to the variation of prices, productivity and wages over time. This data base included construction data from 1954 to 1966. To remove the temporal effects inherent in this data, 1965 was chosen as the base year, and all data from other years was adjusted to the base year by means of standard shipbuilding industry indices for price, productivity and wages.

The order in which these adjustments were made to the data was as follows:

1. application of the learning curves produced data representing one unit costs.

2. adjustments, using 1965 indices, produced data representing one unit costs in 1965 dollars.

3. addition of the cost of GFE and plans, produced data representing all basic contract costs, on a consistent level, as unit one costs in 1964 dollars.

In most cases these adjustments involved relatively small dollar differences between raw and adjusted data. For the base year, one comment is necessary. The cost quantity adjustment was carried out in terms of inflated dollars since the price level effects were treated after the cost quantity adjustments. A reversal of this order of treatment would produce different final dollar values. An explanation concerning the order of treatment would have been appropriate.

2. <u>End Cost Data</u>

End cost data was adjusted in much the same way that Contractor Bid Data was adjusted. Two specific adjustments [Ref. 12] were made to the raw End Cost Data, as follows:

a. Cost quantity adjustment utilizing a slope of 96.8 percent.

b. An adjustment for price level based on general shipbuilding, electronics, and ordnance indices with 1965 as the base year.

The two adjustments listed above provided small and consistent changes between raw and adjusted data, and were therefore accepted as reasonable.

B. THESIS DATA BASE

The adjusted values for Basic Bid and End Cost Data were accepted as a point of departure for this analysis of destroyer construction costs. However, one objective of this thesis is to examine Basic Bid and End costs simultaneously. Therefore, the End cost data were aggregated with the Basic Contract Cost data as follows:

a. Electronics End Cost was added to command and control cost,

b. Weapons End cost was added to Armament cost, and

c. Miscellaneous End cost was added to construction services cost.

The thesis data base is the adjusted construction cost data for 36 ships. These 36 ships are as noted in Table III.

TABLE III

THESIS DATA BASE SHIP TYPES

Type	Data available
DD	3
DDG	11
DE	11
DEG	2
DLG	9
TOTAL	36 ships

This breakdown of ship types in the data base cannot be considered as a representation of the proportion of each type ship in either the current or future Navy. The general purpose destroyer (DD) is obviously underpresented with only 3 ships in the data base. Thus, the data base could be considered as being biased toward guided missile ships (22 out of 36 ships). There is no way to correct a possible bias except by attempting to use weighted average values (weight the average figures for each ship type by the proportion of that type in the current or proposed Navy) or selectively dropping some of the DDGs/DLGs to gain a more correct proportional representation. However, there are relatively few ships (36) in the data base, and average figures tend to eliminate possibly important differences among ships of the same class. Also, there is no really objective way to determine the proportional breakdown of ships in a future Navy. For these reasons, it was decided to use the data base as given, while recognizing a possible bias. The data base is not truly homogenous since it contains five different types of ships. They are all bound together under the general heading of surface combatant ships. However, major differences exist, as follows:

1. DD - general purpose destroyer; good shore gunfire support capability; ASW capability; poor AAW capability.

2. DDG - general purpose destroyer with good gunfire support, ASW and AAW capability.

3. DE - ocean escort with good ASW capability only.

4. DEG - ocean escort with good ASW and close in AAW capabilities.

5. DLG - major fleet escort; extra communication and control equipment; good ASW and best AAW capabilities.

It is obvious that a DE cannot perform all of the same missions that a DLG can perform. Even so, each data point is given the same weight in the data base. In reviewing the data, it was noted some DLG type ships had significantly high costs in the following areas: hull, outfitting, construction services, weapon and cost, and electronics end cost. In addition, as described earlier, the DLG type ship can be considered to have a different operational mission than the smaller, less expensive destroyer type ships. However, a single group data will be considered since there are relatively few ship cost data in this thesis.

Using the data base, a CER will be developed using two different methods of cost disaggregation schemes. One is each 9-subsystems cost group CER. The other is summation of 9-subsystem cost groups CER. This data contains 36 physical characteristics for each ship in numerical form. This characteristics is essentially design parameters such as maximum speed, maximum draft, number of generators, hulls, armament, etc., those are listed in Appendix B.

VI. DATA ANALYSIS

All the data was fed into the "MINITAB" statistical program to use linear regression technique. In part A, a general discussion of independent variable selection criteria is followed by a detailed discussion regarding the development of each CER. In part B, CERs and a summary of statistical information relevant to each CER was listed.

A. DEVELOPMENT OF CER

1. Criteria

The choice of independent variables to be employed in each CER was based upon several criteria:

a. Each independent variable should denote a subjectively logical causal relationship with cost.

b. Each explanatory variable should exhibit a high correlated with dependent variable and a high degree of statistical independence from all other explanatory variables used in the same CER. This will be examined by correlation matrix.

c. Each variable should be input oriented, implying that its value could be obtained with a high degree of certainty before ship construction began.

d. Each CER should have higher R^2 (coefficient of determination). Addition of additional explanatory variables can never decrease R^2, the increase may be marginal and not worth the additional complexity. The point at which the increase in R^2 ceases to be meaningful can then be used to determine the best subset of independent variables.

e. A value of 20 percent or less for the CV (coefficient of Variation) is desirable. CV is a measure which relates the standard error of the model to the mean value of the dependent variable.

f. Residual plot against the computed residual value for predicted value for the dependent variable. Each residual plot was examined for points that appeared to be outliers and for any indication of a need for any type of transformation.

g. The t-statistic of the coefficient of each variable should be significant level under proper assumptions of normality.

h. Finally, the F-statistic of each CER should prove significant of the regression line at the 0.999 level.

2. Discussion of CER Development

With each CER, work was begun by an examination of variables having subjectively logical cause-effect relationship between explanatory variables and cost followed by an examination of the correlation matrix of the dependent versus each independent variable selected. Next, the variables highly correlated with the dependent variable except high intercorrelated independent variables (multicollinearity) were regressed using 'MINITAB' statistical package. During the analysis, residuals were plotted against the predicted value for the dependent variable.

Each residual plot was examined for points that appeared to be outliers and for any indication of a need for any type of transformation. Eventually a best model was selected to be the CER for each of the 9 subsystems and the single total cost equation model.

In the statistics given for each CER the F-ratio is the test value for the hypothesis that the regression is not significant. If the F-ratio is larger than the table value, the hypothesis is rejected and the R^2 is considered significant. The t-value for each independent variable is used to test the hypothesis that the coefficient of the variable in equation equals to zero, the hypothesis is rejected if the table value is less than the t-value.

a. 9-SUBSYSTEM MODEL

(1) HULL COST CER

Hull cost was available for seven variables selected by logical cause effect relationship subjectively. The correlation matrix for those variables examined is listed below:

	HULL COST	ENG PAY	LSW	ENGWGT	ARMWGT	PROWGT	PRAXWGT
ENG PAY	0.854						
LSW	0.847	0.981					
ENGWGT	0.846	0.986	0.953				
ARMWGT	0.878	0.946	0.886	0.964			
PROWGT	0.784	0.918	0.836	0.959	0.965		
PRAXWGT	0.838	0.980	0.936	0.998	0.971	0.973	
HULLWGT	0.796	0.908	0.972	0.864	0.768	0.693	0.835

All of these variables are highly correlated with hull cost. However, since the ENGPAY, the summation of engineering and payload weights, is also very highly correlated with other variables, one must insure that if ENGPAY is included in the final model, the others should not and vice versa.

Each of the variables are acceptable in the regression equation. However, ENGPAY, representing the total light ship's weight (LSW) less the weight of the hull (HULL WGT), is the most statistically satisfactory variable. The choice of ENGPAY does have a significant basis in logic since the hull cost should be directly related to the total weight of the ship's powerplant (ENGWGT) and the total weight of ship's armament, c&c equipment and outfitting (PAYLOAD). These are the weights that the hull must be designed to carry. From the regression analysis using variable ENGPAY the following models were obtained:

R^2 = 72.9	t-value for ENGPAY = 9.56
SE = 0.6045	Table t-value = 2.03
F-ratio = 91.400	(.95,34)
Table F-value = 13.1	CV = 0.238
(.999,1,34)	

This model is statistically appealing and intuitively reasonable. After residual plots were examined, many transformations were tried. However, no improvement was noted.

(2) PROPULSION COST CER

The propulsion cost is highly correlated with PROWGT, PWRLD, and ENGWGT, but there are high intercorrelation among the above variables . The correlation matrix of four possible variables are as follows:

	PROPCOST	PROWGT	PWRLD	ENGWGT
PROWGT	0.860			
PWRLD	0.671	0.699		
ENGWGT	0.825	0.959	0.519	
RANGE	0.345	0.417	-0.301	0.558

The result of the trial to get the best CER shows that two explanatory variables, PWRLD and RANGE, are selected to predict propulsion cost. The use of these variables seems extremely logical. Both represent significant characteristics of the required power-plant. PWRLD, the ratio of maximum shaft horse power to full displacement weight, is an indicator of power. RANGE, of course, is an indicator of endurance capability. Together, the required performance of a power plant is very well defined.

However, in analyzing the residual plots of this model, the evidence of increasing variance was noted. This indicates cost increase is a power function of PWRLD and RANGE. The following transformation withstands the test of logic;

LOG (cost) = a +(PWRLD) + c(RANGE)

The phenomenon of diminishing returns to scale has long been noted in the field of power-plant design. Doubling a ship's horsepower will not double its speed. From the regression analysis using two variable PWRLD and RANGE, the transformed model is as follows:

R^2 = 85.2	t-value for pwrld = 13.18
SE = 0.07837	range = 7.90
F-ratio = 32.650	Table t-value = 2.04
Table F-ratio = 8.5	(.95,33)
(.999,2,33)	CV = 0.117

This model is statistically good and intuitively reasonable.

(3) ELECTRICAL COST CER

Examining the data, ELEWGT, NO_GEN and TKWCY are logically related to electrical cost. The correlation matrix of these three variables are as follows:

	ELECCOST	ELEWGT	NO_GEN
ELEWGT	0.732		
NO_GEN	0.734	0.580	
TKWCY	0.731	0.971	0.541

All of these three variables are highly related with electrical cost. However, high intercorrelation was observed between ELEWGT and TKWCY. Thus, two variables were employed to explain the cost of the electrical power-plant and associated equipment. The use of the weight of electrical equipment, ELEWGT, has traditional justification. The inclusion of NO_GEN, an indicator variable for the number of generators, also has a logical casual relationship with cost, considering the positive coefficient of this variable. From the regression analysis, the following model was obtained:

R^2 = 68.0	t-value for ELEWGT = 3.82
SE = 0.2937	NO_GEN = 3.86
F-ratio = 35.111	Table t-value = 2.04
Table F-ratio = 8.5	(.95,33)
(.999,2,33)	CV = 0.191

This model looks good. By examining the residual plots, there is no indication of the need for transformation.

(4) COMMUNICATION AND CONTROL COST plus ELECTRONIC END COST CER

This CER demonstrated the difficulty in estimating the cost of electronic equipment. Selected by logic, three variables were examined in the below correlation matrix:

	C+E COST	C-CWGT	PROTO
C-CWGT	0.551		
PROTO	-0.042	-0.140	
MS END	-0.007	0.484	-0.166

But no variable was highly correlated with the electronics cost. However, C&CWGT served as a traditional explanatory variable, representing the amount of communication and control equipment as bulk weight.

When regression analysis was applied to the data, the first three variables were taken in, giving logical but poor statistics, i.e., R^2 is too low. The binary indicator variable for prototype ships (PROTO) and another indicator variable representing the number of missile launchers (MS-END) might be excluded because of the insignificant t-ratio. During the examination of the residual plots, an inexplicably high electronics cost was indicated for observation 24, making it an obvious outlier. Since no information could be obtained to support this high cost level, this observation was deleted from the data base while developing this CER. The model with one observation (24) removed had the following statistics:

$R^2 = 64.8$	t-value for C-CWGT = 7.79
SE = 1.856	Table t-value = 2.03
F-ratio = 60.634	(.95,33)
Table F-ratio = 13.1	CV = 0.462
(.999,1,33)	

Even these statistics are poor but acceptable in view of no alternative.

(5) AUXILIARY COST CER

Three variables are considered in this model. The correlation matrix showed as follows:

	AUX COST	PRAXWGT	AUXWGT
PRAXWGT	0.846		
AUXWGT	0.764	0.886	
PROWGT	0.816	0.973	0.756

Both the AUXWGT and PROWGT are highly correlated with the auxiliary cost. But the intercorrelation between the two was 0.756. This high correlation seems reasonable since the auxiliary and propulsion systems operate as a composite system in providing services to the ship. The auxiliary system draws steam and power from the propulsion plant and thus does not operate as a separate system. The combining of these two weights thus eliminates the problem of intercorrelation and logically explain an increase in cost as a function of both weights. PRAXWT, a single variable consisting of the weight of auxiliary equipment (AUXWGT) and the weight of the propulsion plant (PROWGT), would contribute to this CER. The model using this single variable PRAXWT had the following statistics:

R^2 = 71.6	t-value for PRAXWGT = 9.26
SE = 0.5665	Table t-value = 2.03
F-ratio = 85.724	(.95,34)
Table F-ratio = 13.1	CV = 0.251
(.999,1,34)	

These statistics are reasonable. When examining the residual plots, no particular transformation was indicated.

(6) OUTFITTING COST CER

Outfitting costs were available for two variables which display logical cost implication. The correlation matrix for those variables listed below:

	OUTFCOST	LSW
LSW	0.844	
CUTWGT	0.821	0.960

Both LSW and OUTWGT indicated a high correlation with outfitting cost. Two highly intercorrelated explanatory variables, LSW and OUTWGT, could be utilized separately to explain the cost. These variables produced approximately the same reasonable statistical results, but LSW was selected as an explanatory variable due to slightly better statistics as listed below:

$R^2 = 71.2$	t-value for LSW = 9.16
SE = 0.3198	Table t-value = 2.03
F-ratio = 83.894	(.95,34)
Table F-ratio = 13.1	CV = 0.192
(.999,1,34)	

The outfitting cost of a ship should, in essence, be directly proportional to the weight of outfitting material, including hull fittings, non-structural bulkheads, paintings, workshop equipment, and furnishings for quarters. The use of LSW as an explanatory variable is likewise logically consistent. As the LSW increases, it follows that outfitting costs would increase. When residual plots were examined, neither outliers nor indication of a need for a transformation were found.

(7) ARMAMENT COST PLUS WEAPONS END COST CER

This category of cost consists of two cost items. One is the armament cost, including guns and gun mount, ammunition handling and storage system. Another is weapons end cost, consisting of weapons cost after contractors delivery; missiles, ASROC system, etc. Armament cost is a minor portion of the total cost. The use of one CER to predict both costs as an aggregate appears more reasonable. The relationship where four variables were considered is:

	A+W COST	ARMWGT	MS_END	ELEWGT
ARMWGT	0.868			
MS_END	0.787	0.681		
ELEWGT	0.805	0.820	0.473	
OUTWGT	0.783	0.858	0.511	0.920

All of these independent variables are highly correlated with the total costs and intercorrelated among independent variables. The use of ARMWGT as an explanatory variable appears logical as does the use of the indicator variable MS-END, but MS_END was highly correlated with ARMWGT, as the coefficient of correlation is 0.681. Thus it would drop out. It is reasonable to suppose that armament costs would increase as a function of the weight of the weapons systems. When the data are processed by MINITAB, the following model was attained:

R^2 = 75.4	t-value for ARMWGT = 10.21
SE = 5.208	Table t-value = 2.03
F-ratio = 104.215	(.95,34)
Table F-ratio = 13.1	CV = 0.426
(.999,1,34)	

The negative sign for the constant is acceptable unless one is attempting to predict the armament costs of a very light, simple piece of equipment. The weight being considered in this thesis is heavy and complex. The standard error of estimate of this CER is very large, but is considered acceptable in light of the wide range of armament costs and the small size of the data. When residual plots were examined, one observation (obs. 30) is a little high, but with small sample like this thesis, it would be acceptable. There was no indication of the need for transofmration.

(8) DESIGN AND ENGINEERING COST

Six variables were chosen to be considered. The following correlation matrix shows which independent variable has close relationship with the dependent variable (design and engineering cost).

	D_E COST	ARMWGT	PROLSW	C_CWGT	LSW	HULLWGT
ARMWGT	0.310					
PROLSW	0.843	0.057				
C_CWGT	0.276	0.830				
LSW	0.367	0.886	0.170	0.971		
HULLWGT	0.382	0.768	0.214	0.944	0.972	
PROWGT	0.267	0.965	0.043	0.797	0.836	0.693

PROLSW is only variable highly correlated with design and engineering cost. The significance of PROLSW as an explanatory variable in explaining design and engineering cost appears valid since this includes the cost of drawings, technical manuals, mock-ups and models. These costs obviously would be much higher for prototype ships and more expensive for larger prototypes than for the smaller ones. The use of ARMWGT as a second explanatory variable is less obvious, unless an association is developed between armament weight and weapon system complexity, wherein increased armament weight could indicate a more complex weapon system with high design and engineering costs.

From the regression analysis using the two variables, the following models and statistics were obtained.

R^2 = 78.0	t-value for ARMWGT = 3.22
SE = 1.308	PROLSW = 10.13
F-ratio = 58.562	Table t-value = 2.04
Table F-ratio = 8.5	(.95,33)
(.999,2,33)	CV = 0.855

This model is intuitively reasonable, even though CV is a little high. It means this model is not well suited to the data. But in this thesis, the CV value is considered acceptable due to the small sample size. Several other CERs were developed but did not show a better CV. While examining the results of residual plots, four outliers (obs. 1,4,26,30) are found. However, those are not removed from the data base while developing this CER.

(9) CONSTRUCTION SERVICE COST PLUS MISCELLANEOUS END COST

This category of costs includes a potpourri of odd costs attributable to ship construction - staging and scafolding costs; hull, mechanical and electrical (HME) costs resulting from engineering changes, launching costs, trial costs, and drydocking costs. Five variables logically related to costs were chosen for analysis.

	C+M COST	LSW	PROLSW	C_CWGT	AR/LSW
LSW	0.766				
PROLSW	0.354	0.170			
C_CWGT	0.775	0.971	0.121		
AR/LSW	-0.099	0.319	-0.120	0.282	
HULLWGT	0.822	0.972	0.214	0.944	0.119

While developing the CER, LSW, C&CWGT, and HULLWGT are found to be highly correlated with these potpourri costs. However, since LSW and C&CWGT are also very highly correlated with HULLWGT, one must insure that if one of these three variables is included in the final model, the other two could not. Initially three variables: HULLWGT, PROLSW, and AR/LSW, were used. But the PROLSW and AR/LSW were dropped in the final model, since the t-ratios of those two variables are too low. The variable HULLWGT is directly related to construction costs, since a large ship needs more engineering change costs, and more drydocking costs. The statistics for the final CER are as follows:

R^2 = 67.5	5-value for HULLWGT = 8.41
SE = 1.993	Table t-value = 2.03
F-ratio = 70.761	(.95,34)
Table F-ratio = 13.1	CV = 0.313
(.999,1,34)	

R^2-value, is a little low, indicating that HULLWGT alone does not explain all of the variance in the construction costs data. The remaining variance may be explained by other variables. However, other variables did not perform any better than HULLWGT. One observation (obs. 24) is an outlier in the residual plots.

b. Single Model

The single model is an aggregation of all basic contract and end costs into a single total cost equation. Most of the variables used in 9-sub system were weights. Therefore, a logical aggregation of these weights would be in the form of LSW, the consisting of all items of outfit, equipment, and machinery. Initial attempt was made to utilize LSW and some explanatory variable used in 9-subsystem CERs in a single CER for TOTAL COSTS.

	TCOST	LSW	ENGWGT	HULLWGT	PROWGT	PROLSW	MS_END
LSW	0.913						
ENGWGT	0.866	0.953					
HULLWGT	0.886	0.972	0.864				
PROWGT	0.758	0.836	0.959	0.693			
PROLSW	0.332	0.170	0.127	0.214	0.043		
MS_END	0.812	0.781	0.775	0.735	0.716	0.068	
PWRLD	0.269	0.288	0.519	0.077	0.699	-0.161	0.312

The CER of this single total costs was attained as follows:

R^2 = 89.7	t-value for LSW = 7.08
SE = 7.054	MS-END = 3.14
F-ratio = 93.077	PROLSW = 3.48
Table F-ratio = 7.12	Table t-value = 2.05
(.999,3,32)	(.95,32)
	CV = 0.182

The resultant regression was extremely surprising in its high statistical significance. However, the variable PWRLD proved insignificant in the regression. Thus the three variables, LSW, PROLSW and MS-END were selected as explanatory variables. The logic of including LSW as an explanatory variable has its roots in its historical success in explaining costs. Larger ships cost more. The inclusion of MS-END, attributing an increase in cost to the addition of missile systems, is also logical. The cost of armament, c&c equipment, auxiliary, and electrical equipment must increase because all, to some extent, support a missile system. The relatively high cost of a missile system and associated fire-control equipment is easily realized in comparing the costs of a missile ship with the cost of a non-missile ship. The variable PROLSW accomplishes two purposes. First it demonstrates that the cost of a prototype ship is more than that of a non prototype ship. Secondly, it indicates that the cost of a larger prototype ship is more than that of a smaller prototype. Both of these concepts are logical.

The problem of outlier appeared again while this CER was being developed. There are three outliers (obs. 24, 28,30) which displayed inexplicably a little high but were not deleted from the data base because it is not significant in this small data base.

B. CERs AND STATISTICAL SUMMARIES

The Table IV present the 9-subsystem CERs and the single CER, and a summary of statistical data is listed in Table V pertinent to that set of CER equation. The statistical information provided in Table V consists of the following:

1. The computed t-value for each variable of each CER is utilized to test the statistical significance of the coefficient of that particular variable. The computed t-value should be greater than or equal to the critical t-value to demonstrate the significance of the coefficient statistically.

2. The critical t-value is taken from standard student-t tables with a significance level of .95 and a degree of freedom equal to N-K-1, where N is the number of observations and K is the number of independent variables utilized in the entire CER.

3. The computed F-ratio is utilized to test the statistical significance of the entire CER and is merely the ratio of explained variance to unexplained variance (S2). The F-ratio should be greater than or equal to the critical F-value to demonstrate the significance of the entire CER statistically.

4. The critical F-value is taken from standard F tables, using a .999 significance level and N-K-1 versus K degrees of freedom (df).

5. R^2, the coefficient of determination, is essentially a measure of 'Goodness- of Fit' of the regression equation to the data. A perfect fit with the data would be implied if R^2 equals 1.0. By definition, R^2 is the ratio of explained sums of squares to total sums of squares.

6. CV, the coefficient of variation, is a comparison between the dispersion of data points about the regression line, and the average or mean value of the dependent variable. The range of desired CV values would be 0.2 or less.

7. df, degree of freedom, represents the number of observations less the number of restrictions upon the observations. In this section df will always be computed as N-K-1.

TABLE IV

SUMMARY OF CERs

A. 9-subsystem CERs

cost sub-category	CER
1. Hull	Cost = 0.0053 + 0.0013 (ENGPAY)
2. Propulsion	Cost = EXP(-0.4330 + 0.0572(PWRLD) + 0.0001(RANGE))
3. Electrical	Cost = 0.1490 + 0.0039(ELEWGT) + 0.2620 (NO_GEN)
4. C&C + Electronics End cost	Cost =-1.5100 + 0.0328(C-CWGT)
5. Auxiliary	Cost = 0.0197 + 0.0023(PRAXWT)
6. Outfitting Cost	= 0.4360 + 0.0004(LSW)
7. Armament + Weapons End Cost	Cost =-5.7400 + 0.0825(ARMWGT)
8. Design & Engineering	Cost =-1.0100 + 0.0065(ARMWGT) + 0.0015(PROLSW)
9. Construction Service & Miscellaneous end cost	Cost = 0.0029 + 0.0048(HULLWGT)

B. Single Total Cost CER

Tcost = -4.1400 + 0.0107(LSW) + 9.15(MS-END) + 0.0029(PROLSW)

TABLE V

STATISTICAL SUMMARY

A. 9-subsystem CER

Cost -category	variable	t-value computed	t-value tabled	F-ratio computed	F-ratio tabled	R^2	CV
1. Hull	ENGPAY	9.56	2.03	91.40	13.1	72.9	0.24
2. Propulsion	PWRLD	13.18	2.04	32.65	8.5	85.2	0.12
	RANGE	7.90	2.04				
3. Electrical	ELEWGT	3.82	2.04	35.11	8.5	68.0	0.19
	NO-GEN	3.86	2.04				
4. C&C + Electronics End Cost	C-CWGT	7.79	2.03	60.64	13.1	64.8	0.46
5. Auxiliary	PRAXWGT	9.26	2.03	85.72	13.1	71.6	0.25
6. Outfitting	LSW	9.16	2.03	83.89	13.1	71.2	0.19
7. Armament + Weapons End Cost	ARMWGT	10.21	2.03	104.22	13.1	75.4	0.42
8. Design & Engineering	ARMWGT	3.22	2.04	58.56	8.5	78.0	0.86
	PROLSW	10.13	2.04				
9. Construction Service & Miscellaneous End Cost	HULLWGT	8.41	2.03	70.76	13.1	67.5	0.31

TABLE V (cont'd)

STATISTICAL SUMMARY

B. Single CER							
Tcost	LSW	7.08	2.05	93.07	7.12	89.7	0.18
	MS-END	3.14	2.05				
	PROLSW	3.48	2.05				

Note: Outlier deleted : obs.24 In CER 4

VII. ESTIMATE OF TOTAL COST VARIANCE

Throughout this paper, emphasis has been continually placed upon the development of models that determine estimates for total procurement cost. An obvious measure of each model's effectiveness consists of an estimation of the total cost variance associated with each model. Two methods are outlined in sections A and B from which two different estimates for total cost variance are obtained for the 9-subsystem model in this paper [Ref. 12]. Note that the total variance of the 9-subsystem Model included outliers so as to compare with the Total variance of the Single Model.

A. SUMMATION METHOD

Associated with each CER is an estimate of CER variance based on the summation of the squared residual values of each observation divided by the degrees of freedom of the CER. Mathematically, for each CER,

$$S_j^2 = \sum_{i=1}^{N} \frac{(\text{residual})^2}{N - K - 1}$$

where S_j^2 = estimated variance of each CER

N = number of observations

K = number of independent variables

N-k-1 = number of degrees of freedom

Within this equation, the term residual is defined as the difference between the observed cost and the cost predicted as a result of the CER. There is a residual value for each cost observation within the data base for each model of that cost.

Since the model predicts a total cost by summing the cost estimates obtained from its unique set of CERs, it would be logical to assume that an estimate of total cost variance would be the summation of the individual CER variance estimates. That is,

$$S^2 = \sum_{i=1}^{L} S_j^{\;2}$$

$$2$$

where S_2 = estimated total cost variance

S_j = estimated variance of each CER

L = number of CERs in the model.

Adoption of this technique requires the acceptance of one important assumption: That each CER produces a cost estimate totally independent of every other cost estimate within the model. This assumption is obviously difficult to accept. Nevertheless, this method is still quite useful because it allows the establishment of a lower bound on the cost variance estimate; that is, a value of the total cost variance which represents the minimum total cost variance that may be attained utilizing that particular set of model CERs.

An estimate for CER variance is automatically calculated for each CER by the MINITAB program. Table VI lists these individual CER variance estimates.

TABLE VI

ESTIMATES OF TOTAL COST VARIANCE BY SUMMATION METHOD

		S^2
A.	9-subsystem Model	
	1. Hull	0.365
	2. Propulsion	0.006
	3. Electrical	0.086
	4. C&C + Electronics End	12.909
	5. Auxiliary	0.321
	6. Outfitting	0.102
	7. Armament + Weapons End	27.123
	8. Design and engineering	1.710
	9. Construction service + Miscellaneous End	3.972
	Total :	46.594
B.	Single Model	
	Total cost :	41.091

B. MEAN SQUARE RESIDUAL (MSR) METHOD

The second method of total cost variance estimation involves the calculation of a total cost mean square residual (MSR) for each model. The following equation represents the general method utilized to calculate this value:

$$MSR = \frac{\sum_{i=1}^{N} \left(\sum_{j=1}^{L} (RESIDUAL_{ij}) \right)^2}{N - M - L}$$

where N = number of ships

M = number of variables utilized in all CERs of model

L = number of CERs utilized in model

For example, the total cost MSR for the 9-subsystem model employs the following parameters:

N = 36
L = 9
M = 12
N-L-M = 15

Note that by summing the residual values produced by each CER for a given ship, the difference between the observed and predicted total cost is obtained for that ship as an aggregate of the individual CER residual values. When these total cost residuals are squared, summed for all observations (ship), and corrected for degrees of freedom, an estimate of variance is produced for the given model.

Table VII contains a listing of the total cost residual values of each observation (ship) used in producing the 9-subsystem CERs and the Single CER.

The summation method may be thought of as a lower bound on the total cost variance. Analogously, the MSR method may be thought of as an upper bound on the total cost variance estimate: a value below which the estimate of the total cost variance is expected to lie. Implicit within the formulation of this method is the assumption that each cost group observation is dependent upon every other cost group

TABLE VII

TOTAL COST RESIDUALS (millions, 1965 dollars)

Observation	9-subsystem Model	Single Model
1	19.3734	8.3833
2	13.5494	-1.2505
3	14.6241	-0.7305
4	-3.3459	-1.9367
5	2.9428	-0.4150
6	3.5566	0.5294
7	-0.3072	-3.2439
8	4.1966	1.1694
9	8.7115	5.5606
10	5.8366	2.8095
11	-8.2934	-11.3206
12	4.9566	1.9294
13	4.8993	2.3562
14	7.4902	5.0138
15	1.7488	2.1792
16	-0.2190	-1.6032
17	-0.9190	-2.3032
18	-1.9590	-3.3432
19	0.3121	5.5671
20	-5.9266	-2.4947
21	-3.5297	2.4147
22	-0.8604	5.3718

TABLE VII (cont'd)

TOTAL COST RESIDUALS (millions, 1965 dollars)

23	-1.9775	0.3164
24	-21.9065	-20.1688
25	1.2950	3.5991
26	-10.6723	3.0112
27	-3.5207	6.2142
28	-9.9193	-13.6494
29	7.5744	3.8206
30	-19.9575	-10.1192
31	-7.2639	-6.1311
32	5.6624	8.8352
33	0.1824	3.3552
34	-10.9856	-0.8345
35	2.7496	5.2540
36	-0.6837	1.8475

observation and the degree of this dependence (covariance) is assumed to be 1.0. It is highly unlikely that this degree of interdependence will exist between all cost group observations. However, this assumption allows the creation of an expected upper bound and is therefore useful. Its usefulness is further strengthened by the enormous difficulties involved in obtaining an actual estimate for the degree of dependence that exists between the various sub category costs in a given model. If the covariance matrix were easily attainable and accurate, the need for upper and lower bounds in the development of total cost variance estimates would be eliminated.

Intuitively, the summation method rests upon the assumption that each observation of sub-category cost (hull cost, propulsion cost, etc.) is independent of all other sub category costs within the particular model.

The other extreme, the MSR method, requires that only N observations be independent, assuming that each sub-division of total cost is entirely dependent on all other sub-division of cost. Neither of these methods allows an exact determination of total cost variance, since neither of the underlying assumptions is totally correct. Thus, the best estimate of total cost variance should lie between these two extremes.

Table XIII summarizes the data used in calculating the MSR value for two models under discussion. The bottom line of the table is the computed MSR value and represents the

TABLE XIII

ESTIMATES OF TOTAL COST VARIANCE BY MEAN SQUARE RESIDUAL METHOD

	9-subsystem Model	Single Model
$SSR = \sum_{L=1}^{N} \sum_{j=1}^{L} (RESIDUAL_{ij})^2$	2542	1315
N = No. observations	36	36
L = No. CERs in models	9	1
M = No. variables in model	12	3
N-L-M = No. degrees of freedom	15	32
MSR = SSR / (N-L-M)	169.47	41.09

TABLE IX

SUMMARY OF TOTAL COST VARIANCE

	summation method			MSR method		
	S2	S	CV	S2	S	CV
9-subsystem Model	46.59	6.83	0.18	169.47	13.02	0.34
Single Model	41.09	6.41	0.17	41.09	6.41	0.17

Note: $CV = \frac{S}{TC_{AVG}}$ $\qquad TC_{AVG} = 37.620$

the expected upper bound of total cost variance as discussed earlier.

C. ANALYSIS OF RESULT

Table IX is a summary of the estimates for total cost variance calculated using both the summation method and MSR method. The left side of Table IX contains data pertinent to the lower bound (LB) on total cost variances, while the right contains data on the upper bound (UB). The square root of these variance estimates (S2), is the standard error of estimate (S) and is analogous to the standard deviation. It represents a measure of the dispersion or spread that results from a less than perfect fit of the regression line (CER) with the data. A common use of this measure of dispersion is in the computation of the coefficient of variation (CV). The CV is a ratio comparison of the standard error of estimate (S) to the mean of the dependent variable, in this case total cost. Thus, the CV represents a comparison between the expected dispersion of total cost and the average total cost of the ship in the applicable data base.

For instance, the estimate for variance (S2) of the 9-subsystem model is 46.594, using the summation method. The expected dispersion or standard error of estimate for the 9-subsystem model is just the square root of S2 or 6.83. The coefficient of variation (CV) for the 9 subsystem model is the standard error of estimate divided by the average total cost of all ships within the data base or 18%. Thus the standard error of the 9-subsystem model (using the summation) is only 0.18 of the average cost of ships in the data base.

The summation method produces a LB on total cost variance and the MSR method produces an UB. Therefore, the CV calculated by using the total cost variance obtained through the summation method produces a LB on the CV for the model. The CV calculated by using the MSR method of variance estimation produces a UB on the CV for the model. The ideal situation occurs when the UB and LB for the CV are relatively small (.2 or less) and extremely close to one another. This would imply that the standard error of estimate was small when compared to the average total cost regardless of the method used to determine the estimate of variance.

On the basis of the criteria presented in the previous section, the 9-subsystem model produces discouraging CVs. The lower bounds on the model are acceptable values; however, the upper bounds imply that the ratio of standard error of estimate to average cost could get as high as 34%. The major source of difficulty in this model is the fact that the data base contained some non-homogeneous observations.

Note that for the single model, the upper and lower bounds on CV are the same. This is because the single model utilizes only one total cost CER. Thus, the MSR method when applied to a model with only one CER produces the same variance estimate as the summation method.

VIII. COMPARISON OF THE MODELS TO RMC MODEL

The objective of this thesis is to present ship acquisition cost estimating models that provide relatively precise total cost estimates. The Patrol Frigate (PF) is designed as an escort vessel, thus the construction data is applicable to estimate total cost by using each of the models discussed earlier. A comparison of these model estimates with the existing RMC model estimates will provide a degree of validity to the approach adopted.

A. MODEL ESTIMATES

The necessary input data concerning weight allocation and ships characteristics for the PF ship are listed in Table X. These parametric input data were substituted into the CERs of both models and aggregated according to the model structure. Both models' total cost figures in 1965 constant dollar base were produced and presented in Table XI. Note that a contract profit figure of 10% of total cost has been added to each model result. This was done to make the model estimates comparable to the estimate calculated using RMC model where a 10% profit has been figured.

B. RMC MODEL ESTIMATES

1. Basic contract cost

The parametric input data of PF were put into the CER of RMC model based on basic contract costs, and

TABLE X

PATROL FRIGATE INPUT DATA

CHARACTERISTIC	VALUE	UNITS
Hullweight	1235	Long Tons
Propulsion weight	251	"
Electrical weight	160	"
C&C weight	87	"
Auxiliary weight	358	"
Outfitting weight	264	"
Armament weight	96	"
LSW	2451	"
Full Displacement	3400	"
ENGPAY	1216	"
PRAXWT	609	"
ENGWGT	769	"
AR/LSW	0.0392	-
Endurance Range	4500	Nautical Miles
MS-END	1	-
NO-GEN	3	-
PROTO	1	-
PWRLD	11.75	shp/long ton
NUC	0	-
DE-DEG	1	-
PROLSW	2451	-
SERIES	1	-

TABLE XI

COST ESTIMATES OF PATROL FRIGATE
($Millions, 1965 base)

1. 9-subsystem CER

Cost-category	Thesis Model (Basic + end cost)	RMC Model (basic cost)
Hull	1.586	1.796
Propulsion	4.888	3.617
Electrical	1.559	1.543
C&C+Electronics End	1.344	0.551
Auxiliary	1.420	1.594
Outfitting	1.416	1.586
Armament+Weapons End	2.180	0.351
Design and Engineering	3.291	3.412
Construction service +Miscellaneous End	5.931	0.580
sub total	23.615	15.03
10% profit	2.362	1.50
Total	25.977	16.53

2. Single CER

Tcost	38.344
10% profit	3.834
Total	42.178

TABLE XII

PATROL FRIGATE ESTIMATED END COST ITEMS
($ Millions, 1973 base)

COST ITEM	COST BASE	COST
Design changes	% of Lead Ship construction cost	2.000
Construction changes	% of Lead Ship construction cost	2.500
Government Engineering Support	DD 963 Estimate	1.500
NAVSEC Electronics	Sec 6271 Estimate	3.188
NAVSHIPS Sonar	PMS 378 Estimate	2.526
H/M/E Equipment	Preliminary Equipment List	1.400
NAVORD Cost	NAVORD Estimate	8.852
NAVELEX Cost	NAVSEC 6179 Info	0.695
Total End cost		22.661

aggregated by the model structure. The total cost figures in 1965 constant dollar base are presented in Table XI.

2. End Cost

The data, supplied by NAVSHIPS as end cost estimates in 1973 dollars were presented on Table XII.

3. Total Cost Estimate

According to the RMC method for computation of cost elements, the total procurement cost of the lead PF is merely the summation of basic contract cost including the 10% profit and End cost. Since the End cost estimate was given in 1973 dollars and the Basic contract cost was calculated in 1965 dollars, the End cost data were transposed back to 1965 dollars by .55 to account for inflation. The value 0.55 represents the average inflation rate for the year 1965-1963. The deflator is applicable to government purchases of durable goods. Thus the total cost of PF is:

Basic Contract Cost	:	16.53
End Cost	:	12.46
Total Cost		28.99 ($ millions, 1965)

C. DISCUSSION OF COMPARISON

The total cost estimates of the PF based on the two models presented in this paper and the estimate by RMC Model were listed on Table XIV. Note that the estimates are based on 1965 dollars in order to facilitate comparison among the Models. The deflator utilized in transposing was 0.55, which was based on information from the Department of Defense.

Table XIII

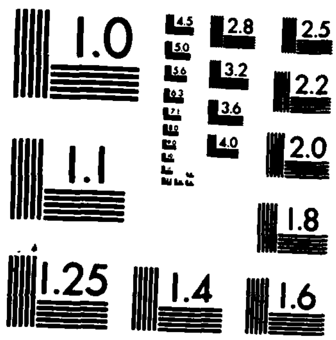
Patrol Frigate Total Cost Estimates
($ Millions, 1965)

A. Thesis Model	
1. 9-subsystem model	25.98
2. Single Model	42.18
B. RMC Model	28.99

It must be noted that the profit assumed in this thesis was 10% of the Total Cost, while the profit assumed in the RMC model is based on basic contract cost without government furnished equipment and the so-called End Cost. The estimates produced by the 9-subsystem model are 3.01 million lower in 1965 dollars than estimates by RMC model, whereas the single total cost model estimates are away from that of RMC model by 13.19 million dollars. Therefore, the 9-subsystem model can be considered comparable to the existing RMC model, but the single model is not in any way comparable and requires further study.

AD-A139 982   COST ESTIMATION OF NAVAL SHIP ACQUISITION(U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA   I H BAIK DEC 83   2/2

UNCLASSIFIED   F/G 5/3   NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

IX. CONCLUSION

This thesis has presented a general procedure for development of a parametric cost estimates in order to familiarize the reader with the approach of this thesis. According to the procedures of PCE, two different models were developed using the data based on destroyer type ships built in 1954-1966. One utilizes 9-subsystem cost group CERs and the other uses only one CER to estimate total ship acquisition cost. There were some difficulties in using the statistical method for this study. The data was too old for current cost estimating, and there were not enough data observations to make the results statistically sound, as was shown in Table V in Chapter VI. An effort, with a larger, updated data base may produce better results.

A MINITAB computer program has been provided which is readily usable in conducting linear least squares analysis and the additional time required would be minimal. The measure of effectiveness utilized was the coefficient of variation, the ratio of the standard error of estimate (S) to the average total cost of the ships. The 9-subsystem model produced a CV that could range from 18% to 34% of average total cost (37.62 million dollars). As observed in Chapter VII, this problem arose because an attempt was made to develop nine CERs with 12 independent variables from a data base that was too small to retain sufficient degrees of freedom to make the results sound. The CV for the single model is 17% of average total cost. This value is considered acceptable.

On this basis, both models might be used to estimate future ship acquisition cost as tools for rudimentary budgetary processes wherein rough ballpark estimates are all that are available.

When these two models were compared with the existing RMC model, the estimate of the 9-subsystem model was 3.01 million in 1965 dollars lower than that of RMC model while single model estimate was 13.19 million dollars higher than that of RMC model, as shown in Table XIII in Chapter VIII. These facts alone prove only that the 9-subsystem estimate is at least comparable to the RMC model. In conclusion, two rudimentary models have been developed in this thesis. Much careful consideration would be required in the use of this model even for ballpark figure budgeting, of course.

APPENDIX A

BASIC CONTRACT AND END COST CATEGORIES

A. Basic Contract Cost Categories

Symbol		Category Name	Includes
1.	Hull	Hull structure	Shell plating, planking, longitudinal, transverse frames, decks, superstructure, armor, etc.
2.	Prop	Propulsion	Boiler and energy converter, propulsion units, upstakes, propulsion control equipment, feedwater and condensate systems, etc.
3.	Elec	Electric Plant	Electric power generators, power distribution switchboards and cables, lighting systems, etc.
4.	C&C	Communication and Control	Navigation equipment, interior communication equipment, fire control systems, radar systems, radio communication systems, sonar systems, etc.
5.	Aux	Auxiliary systems	Heating, ventilating, air conditioning, plumbing, elevators, arresting gears, rudders, etc.
6.	Outf	Outfit and Furnishing	Hull fittings, nonstructural bulkheads, painting, equipment for work shops, furnishings for quarters, etc.

7.	Arm	Armament	Guns and gun mount, ammunition handling, storage systems, other weapon systems handling and storage systems, etc.
8.	D&E	Design and Engineering Services	Contract drawings, working drawings, technical manuals, lofting, mock-up and models, etc.
9.	C/S	Construction Services	Staging, scaffolding and cribbing, launching, trials, cleaning ship, drydocking, etc.

B. End Cost Categories

1.	Weapons End cost	Weapons costs after contractor delivery: missile, ASROC systems, etc.
2.	Electronics End cost	Electronics costs after contractor delivery; radar, NTDS. fire control systems, etc.
3.	Miscellaneous End cost	Disaster costs; cost of hull, mechanical and electrical changes; post delivery cost, etc.

* TCOST Total End Cost = Basic Contract cost + Profit + Miscellaneous End cost + Weapons End cost + Electronics End cost.

APPENDIX B

DESCRIPTION OF SHIPS CHARACTERISTICS

	SYMBOL	CHARACTERISTIC	UNITS	DEFINITION
1.	HULLWGT	Hull weight	long tons	These are the weight of the seven groups as described in the Bureau of Ships Consolidated Index of drawings, Materials, and Services related to Construction Conversions.
2.	PROPWGT	Propulsion weight	"	
3.	ELEWGT	Electrical weight	"	
4.	C&CWGT	Communication and Control weight	"	
5.	AUXWGT	Auxiliary weight	"	
6.	OUTFIT	Outfitting weight	"	
7.	ARMWGT	Armament weight	"	
8.	LSW	Light ship weight	"	Weight of ship complete with all items of outfit equipment and machinery but excluding cargo, stores, crew, etc.
9.	AR/LSW	Armament weight to light ship weight	-	Ratio of armament weight to light ship weight.

10.	PROLSW	Prototype times Light ship weight	-	For prototypes, this takes on the value of LSW. For non prototype, it has the value zero.
11.	NO-GEN	Number of generator	integer	Number of ship service generator.
12.	MS-END	Missile End Dummy	"	0, 1, 2 = number of launcher.
13.	PROTO	Prototype Dummy	"	1 = protype 0 = others.
14.	RANGE	Range	miles	Endurance range
15.	TKWCPY	Generator capacity	KW	Total output of all Generators.
16.	PWRLD	Power loading factor	-	Ratio of maximum shaft horsepower to full Displacement.
17.	ENGPAY	Engpay weight	long tons	Summation of engineering and payload weight

18.	PRAXWGT	Engine weight	"	Weight of Auxiliary plus weight of propulsion.
19.	ENGWGT	Engineering Weight	"	Total weight of all propulsion, Electrical and Auxiliary equipment.
20.	PAYLOAD	Payload weight	"	Total weight of all c+c, armament and outfitting equipment.
21.	DEDEG	DD or DEG Dummy	Integer	1 = *DE* or *DEG*, *0* = *others*
22.	SERIES	Series variable	"	A number that represents the position of a ship in a series of similar ships.

APPENDIX C

DATA BASE MATRIX

info

COLUMN	NAME	COUNT
C1	LSW	36
C2	HULLWGT	36
C3	PROWGT	36
C4	AR/LSW	36
C5	PWRLD	36
C6	RANGE	36
C7	ELEWGT	36
C8	NO-GEN	36
C9	C-CWGT	36
C10	PROTO	36
C11	MS-END	36
C12	AUXWGT	36
C13	OUTWGT	36
C14	ARMWGT	36
C15	PROLSW	36
C16	TKWCY	36
C17	EE COST	36
C18	WE COST	36
C19	ME COST	36
C20	HULLCOST	36
C21	PROPCCST	36
C22	ELECCCST	36
C23	C-C CCST	36
C24	AUX CCST	36
C25	OUTFCCST	36
C26	ARM CCST	36
C27	D-E CCST	36
C28	C/S CCST	36
C29	TCOST	36
C30	ENG PAY	36
C31	PRAXWGT	36
C32	C+E CCST	36
C33	A+W CCST	36
C34	C+M CCST	36
C35	ENGWGT	36
C36	PAYLOAD	36

* SHIP TYPE

DD	:	1 - 3	(3)
DDG	:	4 - 14	(11)
DE	:	15 - 25	(11)
DEG	:	26 - 27	(2)
DIG	:	28 - 36	(9)

--
note: DATA BASE MATRIX(1)
--

print c1 -c36

COLUMN	LSW	HULLWGT	PROWGT	AR/LSW	PWBLD	RANGE
COUNT	36	36	36	36	36	36
ROW						
1	2863.00	961.00	838.	0.0980	16.6700	4500.00
2	2863.00	961.00	838.	0.0980	16.6700	4500.00
3	2863.00	961.00	819.	0.0980	17.2900	4500.00
4	3154.00	1208.00	758.	0.0799	15.5600	4500.00
5	3218.00	1231.00	823.	0.0799	15.5600	4500.00
6	3210.00	1183.00	819.	0.0810	15.5600	4500.00
7	3133.00	1214.00	731.	0.0804	15.5600	4500.00
8	3210.00	1183.00	819.	0.0810	15.5600	4500.00
9	3255.00	1217.00	831.	0.0793	15.5600	4500.00
10	3210.00	1183.00	819.	0.0810	15.5600	4500.00
11	3210.00	1183.00	819.	0.0810	15.5600	4500.00
12	3210.00	1183.00	819.	0.0810	15.5600	4500.00
13	3292.00	1224.00	830.	0.0775	15.5600	4500.00
14	3274.00	1239.00	807.	0.0770	15.5600	4500.00
15	1300.00	580.00	234.	0.0623	10.2600	4500.00
16	1300.00	580.00	234.	0.0623	10.4700	4500.00
17	1300.00	580.00	234.	0.0623	10.4700	4500.00
18	1300.00	580.00	234.	0.0623	10.4700	4500.00
19	1370.00	601.00	224.	0.0292	5.2600	7000.00
20	1370.00	601.00	224.	0.0292	5.2600	7000.00
21	1841.00	796.00	301.	0.0326	7.5500	4000.00
22	2679.00	1218.00	342.	0.0392	10.2900	4000.00
23	2642.00	1204.00	342.	0.0401	10.2900	4000.00
24	2620.00	1232.00	338.	0.0416	10.2900	4000.00
25	2617.00	1229.00	353.	0.0371	10.2900	4000.00
26	2710.00	1232.00	348.	0.0410	10.2200	4000.00
27	2755.00	1259.00	360.	0.0425	10.2200	4000.00
28	4150.00	1581.00	808.	0.0778	14.8600	5000.00
29	4150.00	1581.00	808.	0.0778	14.8900	5000.00
30	5009.00	2191.00	924.	0.0750	12.1400	8000.00
31	5101.00	2288.00	922.	0.0757	12.1400	8000.00
32	5009.00	2191.00	924.	0.0750	12.1400	8000.00
33	5009.00	2191.00	924.	0.0750	12.1400	8000.00
34	5592.00	2568.00	956.	0.0581	10.7600	8000.00
35	5613.00	2574.00	951.	0.0567	10.7200	8000.00
36	5618.00	2574.00	945.	0.0566	10.7200	8000.00

--
note: DATA BASE MATRIX (2)
--

COLUMN	ELEWGT	NO-GEN	C-CWGT	PROTO	MS-END	AUXWGT
COUNT	36	36	36	36	36	36
ROW						
1	125.	4.	163.	1.	0.00	307.
2	125.	4.	163.	0.	0.00	307.
3	125.	4.	163.	0.	0.00	307.
4	115.	4.	163.	1.	1.00	333.
5	126.	4.	175.	0.	1.00	356.
6	125.	4.	163.	0.	1.00	307.
7	115.	4.	163.	0.	1.00	333.
8	125.	4.	163.	0.	1.00	307.
9	123.	4.	176.	0.	1.00	307.
10	125.	4.	163.	0.	1.00	307.
11	125.	4.	163.	0.	1.00	307.
12	125.	4.	163.	0.	1.00	307.
13	127.	4.	175.	0.	1.00	370.
14	117.	4.	177.	0.	1.00	376.
15	56.	2.	45.	1.	0.00	142.
16	56.	2.	45.	0.	0.00	142.
17	56.	2.	45.	0.	0.00	142.
18	56.	2.	45.	0.	0.00	142.
19	40.	3.	43.	1.	0.00	144.
20	40.	3.	43.	0.	0.00	144.
21	132.	3.	86.	1.	0.00	249.
22	109.	2.	153.	1.	0.00	315.
23	109.	2.	153.	0.	0.00	316.
24	109.	2.	153.	0.	0.00	317.
25	105.	2.	167.	0.	0.00	329.
26	105.	2.	176.	1.	1.00	332.
27	106.	2.	178.	0.	1.00	342.
28	148.	4.	222.	0.	1.00	417.
29	148.	4.	222.	0.	1.00	417.
30	230.	4.	240.	1.	2.00	557.
31	225.	4.	292.	0.	2.00	566.
32	230.	4.	240.	0.	2.00	557.
33	230.	4.	240.	0.	2.00	557.
34	275.	4.	280.	1.	1.00	599.
35	241.	4.	317.	0.	1.00	601.
36	241.	4.	318.	0.	1.00	601.

--
note: DATA BASE MATRIX (3)
--

COLUM COUNT ROW	OUTWGT 36	ARMWGT 36	PROLSW 36	TKWCY 36	EECOST 36	WECOST 36
1	302.	260.	2863.0	2200.00	0.3600	5.6500
2	302.	260.	0.0	2200.00	0.3900	5.6600
3	302.	260.	0.0	2200.00	0.3600	5.6600
4	275.	252.	3154.0	2200.00	3.0900	14.6800
5	250.	257.	0.0	2200.00	2.9300	14.7700
6	302.	260.	0.0	2200.00	2.9800	14.6800
7	275.	252.	0.0	2200.00	2.9900	14.7200
8	302.	260.	0.0	2200.00	2.7000	12.7200
9	272.	258.	0.0	2200.00	2.5700	11.2700
10	302.	260.	0.0	2200.00	2.6600	11.3100
11	302.	260.	0.0	2200.00	3.2400	24.0600
12	302.	260.	0.0	2200.00	2.8900	11.4800
13	264.	255.	0.0	2200.00	2.5700	13.9100
14	256.	252.	0.0	2200.00	2.8100	11.4500
15	147.	81.	1300.0	600.00	0.2100	1.1300
16	147.	81.	0.0	600.00	0.2700	0.7700
17	147.	81.	0.0	600.00	0.2600	0.7700
18	147.	81.	0.0	600.00	0.2600	0.7600
19	136.	40.	1370.0	600.00	0.2700	0.1800
20	136.	40.	0.0	600.00	0.3200	1.5200
21	153.	60.	1841.0	2000.00	3.1300	1.0700
22	273.	105.	2679.0	2000.00	3.2200	1.1400
23	288.	106.	0.0	2000.00	3.4400	1.6800
24	258.	109.	0.0	2000.00	21.3600	1.3900
25	237.	97.	0.0	2000.00	2.6700	1.3900
26	288.	111.	2710.0	2000.00	5.8000	11.3900
27	302.	117.	0.0	2000.00	4.0600	6.4500
28	330.	323.	0.0	3600.00	3.0800	23.9600
29	330.	323.	0.0	3600.00	2.8100	15.3900
30	406.	376.	5009.0	4600.00	4.4900	39.9200
31	381.	386.	0.0	4600.00	3.2800	33.8200
32	406.	376.	0.0	4600.00	2.6100	24.0300
33	406.	376.	0.0	4600.00	2.9200	24.0300
34	427.	325.	5592.0	6600.00	11.6300	23.8600
35	437.	318.	0.0	6500.00	11.3700	18.7300
36	438.	318.	0.0	6500.00	11.5000	18.7300

--
note: DATA BASE MATRIX (4)
--

COLUMN	ME COST	HULLCOST	PROPCOST	ELECCOST	C-C COST	AUX COST
COUNT	36	36	36	36	36	36
ROW						
1	2.5800	1.83000	7.15000	1.81000	0.46000	1.08000
2	2.9700	2.25000	6.66000	1.79000	1.03000	2.38000
3	2.6300	2.62000	5.83000	2.15000	0.61000	3.37000
4	6.0900	2.34000	6.60000	1.59000	0.65000	1.86000
5	5.7300	2.15000	4.68000	1.56000	1.25000	2.37000
6	3.2500	3.31000	4.86000	1.56000	0.83000	3.41000
7	4.5800	2.37000	5.99000	1.86000	1.47000	3.29000
8	4.3200	2.31000	7.09000	1.73000	0.70000	1.90000
9	2.9400	3.09000	6.36000	1.56000	0.71000	2.20000
10	3.3300	2.53000	6.26000	2.05000	1.67000	2.84000
11	5.4100	2.40000	5.50000	1.69000	1.35000	2.53000
12	4.3100	2.80000	4.90000	1.99000	0.38000	4.02000
13	4.1500	3.25000	5.02000	1.42000	0.47000	2.86000
14	4.1900	2.36000	5.09000	1.74000	0.87000	2.27000
15	1.3100	1.76000	2.61000	0.79000	0.45000	0.65000
16	1.1900	1.20000	3.27000	0.93000	0.45000	0.90000
17	1.5200	1.31000	2.56000	1.00000	0.55000	0.84000
18	1.0900	1.66000	3.29000	1.18000	0.78000	0.79000
19	1.5700	1.23000	1.65000	0.58000	0.25000	0.53000
20	3.4800	1.50000	2.02000	0.76000	0.33000	0.68000
21	4.5800	0.78000	1.95000	1.15000	0.34000	1.14000
22	8.4500	1.60000	2.98000	1.02000	0.90000	1.81000
23	7.5600	1.30000	2.78000	1.24000	0.42000	1.78000
24	10.4200	1.73000	3.72000	1.22000	0.46000	1.16000
25	4.5000	1.55000	3.35000	1.02000	0.50000	1.75000
26	9.0300	1.70000	3.16000	1.07000	0.53000	1.87000
27	4.7200	1.66000	3.10000	1.16000	0.70000	1.97000
28	10.0700	3.56000	8.97000	2.21000	1.42000	2.51000
29	5.2900	3.06000	8.28000	1.27000	0.78000	2.30000
30	9.8100	4.04000	8.99000	1.73000	1.09000	3.35000
31	10.9200	4.21000	6.48000	1.50000	2.27000	3.54000
32	6.1500	5.48000	7.53000	1.48000	0.78000	3.48000
33	8.2800	5.46000	7.55000	2.39000	2.33000	4.06000
34	13.8100	3.47000	6.21000	2.39000	0.93000	2.96000
35	8.7600	3.06000	5.58000	2.27000	0.95000	2.71000
36	9.3300	4.17000	5.61000	2.56000	1.00000	4.11000

--
note: DATA BASE MATRIX (5)
--

COLUMN COUNT ROW	OUTFCOST 36	ARM COST 36	D-E COST 36	C/s COST 36	TCOST 36	ENG PAY 36
1	2.10000	0.54000	1.5100	0.56000	27.2700	1902.00
2	1.74000	0.30000	0.4200	1.53000	28.3900	1902.00
3	1.65000	0.36000	0.8700	0.49000	28.4000	1902.00
4	1.66000	0.22000	8.5500	1.02000	51.6100	1946.00
5	1.47000	0.25000	0.7700	0.60000	40.0400	1987.00
6	1.30000	0.36000	0.6800	0.28000	39.1500	2027.00
7	1.31000	0.20000	0.8500	0.82000	41.0200	1919.00
8	1.29000	0.23000	0.4600	1.41000	38.2000	2027.00
9	1.42000	0.27000	0.1400	0.42000	34.5700	2038.00
10	1.46000	0.26000	0.2000	0.65000	35.9100	2027.00
11	1.63000	0.27000	0.6200	0.65000	50.9800	2027.00
12	1.62000	0.12000	0.5700	1.02000	37.3400	2027.00
13	1.47000	0.54000	0.5100	0.38000	37.3200	2068.00
14	1.54000	0.34000	0.5100	0.53000	33.6900	2035.00
15	1.05000	0.22000	0.3300	0.17000	11.3200	720.00
16	1.17000	0.11000	0.3400	0.16000	11.5100	720.00
17	1.31000	0.31000	0.7100	0.32000	12.0800	720.00
18	1.19000	0.13000	0.8500	0.52000	12.4900	720.00
19	0.51000	0.08000	1.0600	0.33000	8.7000	769.00
20	0.68000	0.18000	0.4800	0.45000	12.9300	769.00
21	1.59000	0.17000	1.4200	0.45000	18.4900	1045.00
22	1.55000	0.32000	1.8900	1.28000	25.0900	1461.00
23	1.55000	0.31000	0.3600	0.77000	23.9900	1438.00
24	0.90000	0.47000	0.3300	0.28000	44.7000	1388.00
25	1.42000	0.47000	0.6100	0.41000	20.2500	1388.00
26	1.42000	0.45000	0.5100	0.46000	38.4900	1478.00
27	1.53000	0.43000	0.3600	0.81000	27.2700	1496.00
28	2.67000	0.74000	1.7800	0.76000	67.5900	2569.00
29	2.06000	0.40000	1.5500	1.07000	46.2300	2569.00
30	2.93000	0.69000	11.5700	1.50000	93.9000	2818.00
31	2.54000	2.06000	1.0000	1.21000	75.3100	2813.00
32	2.73000	1.02000	0.9300	0.66000	59.3000	2818.00
33	2.38000	1.15000	0.6000	1.21000	62.9000	2818.00
34	2.72000	0.48000	10.4600	1.34000	80.0400	3024.00
35	2.61000	0.45000	0.4900	1.49000	59.5000	3039.00
36	1.89000	0.65000	0.7500	1.63000	62.5700	3044.00

--
note: DATA BASE MATRIX (6)
--

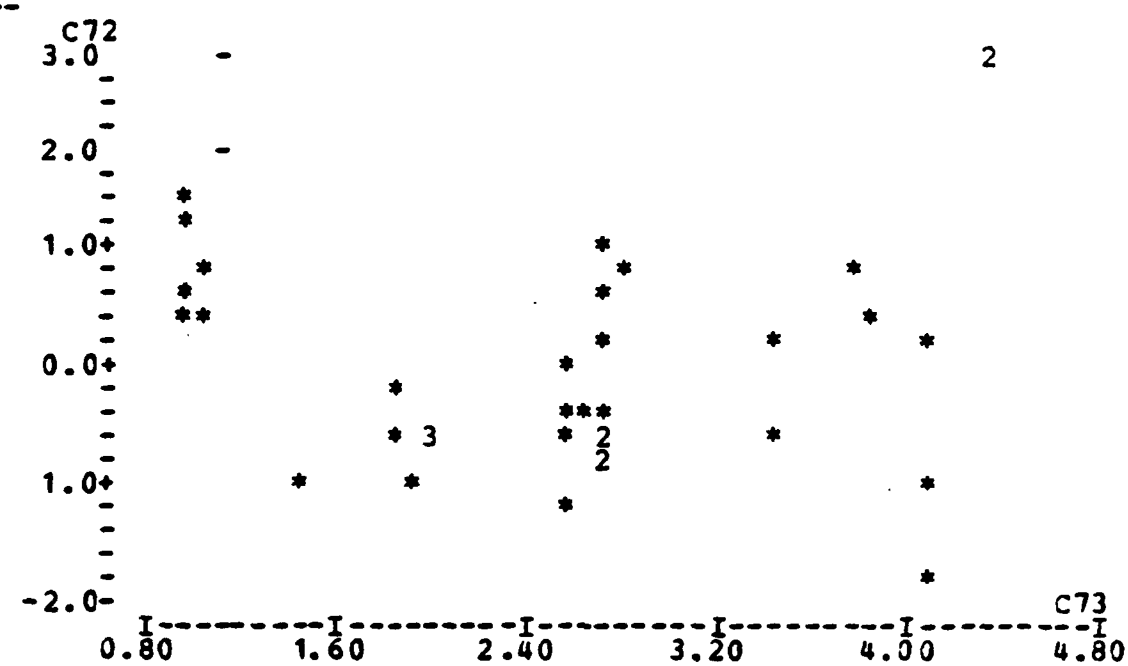
COLUMN	FRAMWGT	C+E COST	A+W COST	C+M COST	ENGWGT	PAYLOAD
COUNT	36	36	36	36	36	36
ROW						
1	1145.00	0.8200	6.1900	3.0400	1270.00	725.00
2	1145.00	1.4200	5.9600	4.0000	1270.00	725.00
3	1126.00	0.9700	6.0200	3.2400	1251.00	725.00
4	1091.00	3.7400	14.9000	6.7400	1206.00	690.00
5	1179.00	4.1800	15.0200	6.9800	1305.00	682.00
6	1126.00	3.8100	15.0400	4.0800	1251.00	725.00
7	1064.00	4.4600	14.9200	6.0500	1179.00	690.00
8	1126.00	3.4000	12.9500	5.0200	1251.00	725.00
9	1138.00	3.2800	11.5400	3.6500	1261.00	706.00
10	1126.00	4.3300	11.5700	5.0000	1251.00	725.00
11	1126.00	4.5900	24.3300	6.7600	1251.00	725.00
12	1126.00	3.2700	11.6000	4.6900	1251.00	725.00
13	1200.00	3.0400	14.4500	4.6200	1327.00	694.00
14	1183.00	3.6800	11.7900	5.0600	1300.00	685.00
15	376.00	0.6600	1.3500	1.7600	432.00	273.00
16	376.00	0.7200	0.8800	1.6400	432.00	273.00
17	376.00	0.8100	1.0800	2.0700	432.00	273.00
18	376.00	1.0400	0.8900	1.8700	432.00	273.00
19	368.00	0.5200	0.2600	1.8200	408.00	219.00
20	368.00	0.6500	1.7000	3.8100	408.00	219.00
21	550.00	3.4700	1.2400	4.9200	682.00	299.00
22	657.00	4.1200	1.4600	9.3500	766.00	531.00
23	658.00	3.8600	1.9900	7.9800	767.00	547.00
24	655.00	21.8200	1.8600	10.8800	764.00	520.00
25	682.00	3.1700	1.8600	5.0000	787.00	501.00
26	680.00	6.3300	11.8400	9.5600	785.00	575.00
27	702.00	4.7600	6.8800	5.4200	808.00	597.00
28	1225.00	4.5000	24.7000	11.4900	1373.00	875.00
29	1225.00	3.5900	15.7900	6.0700	1373.00	875.00
30	1481.00	5.5800	40.6100	10.9000	1711.00	1022.00
31	1488.00	5.5500	35.8800	13.1900	1713.00	1059.00
32	1481.00	3.3900	25.0500	6.9300	1711.00	1022.00
33	1481.00	5.2500	25.1800	10.6100	1711.00	1022.00
34	1555.00	12.5600	24.3400	14.7400	1830.00	1032.00
35	1552.00	12.3200	19.1800	9.7100	1793.00	1072.00
36	1546.00	12.5000	19.3800	10.3300	1787.00	1074.00

APPENDIX D

COMPUTER OUTPUT (ANALYSIS RESULTS)

```
--
NOTE: 1.HULL COST CER
--

--
corr c20 c30 c1 c35 c14 c3 c31 c2
--

     HULL COST ENG PAY   LSW ENGWGT ARMWGT PROWGT PRAXWGT
ENGPAY    0.854
LSW       0.847   0.981
ENGWGT    0.846   C.986  0.953
ARMWGT    0.878   0.946  0.886  0.964
PROWGT    0.784   C.918  0.836  0.959  0.965
PRAXWGT   0.838   0.980  0.936  0.998  0.971  0.973
HULLWGT   0.796   0.908  0.972  0.864  0.768  0.693   0.835

--
regr c20 1 c30, st. res c72 pred. y c73
--

THE REGRESSION EQUATION IS

Y =  0.0053 +0.0013 X1

                                      ST. DEV.      T-RATIC =
      CCLUMN      COEFFICIENT         OF COEF.      COEF/S.D.
      --              0.0053           0.2827           0.02
X1    ENG PAY      0.0013477        0.0J01410           9.56

THE ST. DEV. OF Y ABCUT REGRESSION LINE IS
S = 0.6045
WITH (  36- 2) =  34 DEGREES OF FREEDOM

R-SQUARED = 72.9 PERCENT
R-SQUARED = 72.1 PERCENT, ADJUSTED FOR D.F.

ANALYSIS CF VARIANCE

 DUE TO        DF           SS         MS=SS/DF
REGRESSICN      1       33.3966        33.3966
RESIDUAL       34       12.4239         0.3654
TOTAL          35       45.8206

F-ratio = 91.400

CV = 0.238

          X1        Y    PRED. Y    ST.DEV.
ROW  ENG PAY HULLCOST      VALUE   PRED. Y  RESIDUAL    ST.RES.
 32     2818    5.480      3.803     0.167     1.677      2.89R
 33     2818    5.460      3.803     0.167     1.657      2.85R

R DENOTES AN OBS. WITH A LARGE ST. RES.

DURBIN-WATSCN STATISTIC = 1.38

--
aver c20
  AVERAGE =      2.5306
--
```

```
--
plot c72 vs c73
--
```




```
--
note: 2.PROPULSION COST CER
--


--
corr c21 c3 c5 c35 c6
--
```

	PROPCOST	PRCWGT	PWRLD	ENGWGT
PROWGT	0.860			
PWRLD	0.671	0.699		
ENGWGT	0.825	0.959	0.519	
RANGE	0.345	0.417	-0.301	0.558

```
--
1).regr c21 2 c5 c6, st. res c74 pred. y c75
--
```

THE REGRESSION EQUATION IS

Y = - 6.03 + 0.545 X1 +0.0008 X2

	COLUMN	COEFFICIENT	ST. DEV. OF COEF.	T-RATIO = COEF/S.D.
	--	-6.027	1.062	-5.67
X1	PWRLD	0.54485	0.05487	9.93
X2	RANGE	0.0008119	0.0001159	7.00

THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
S = 0.9904
WITH (36- 3) = 33 DEGREES OF FREEDOM

R-SQUARED = 77.9 PERCENT
R-SQUARED = 76.6 PERCENT, ADJUSTED FOR D.F.

ANALYSIS OF VARIANCE

DUE TO	DF	SS	MS=SS/DF
REGRESSION	2	114.1273	57.0636
RESIDUAL	33	32.3684	0.9809
TOTAL	35	146.4957	

F-ratio = 58.175

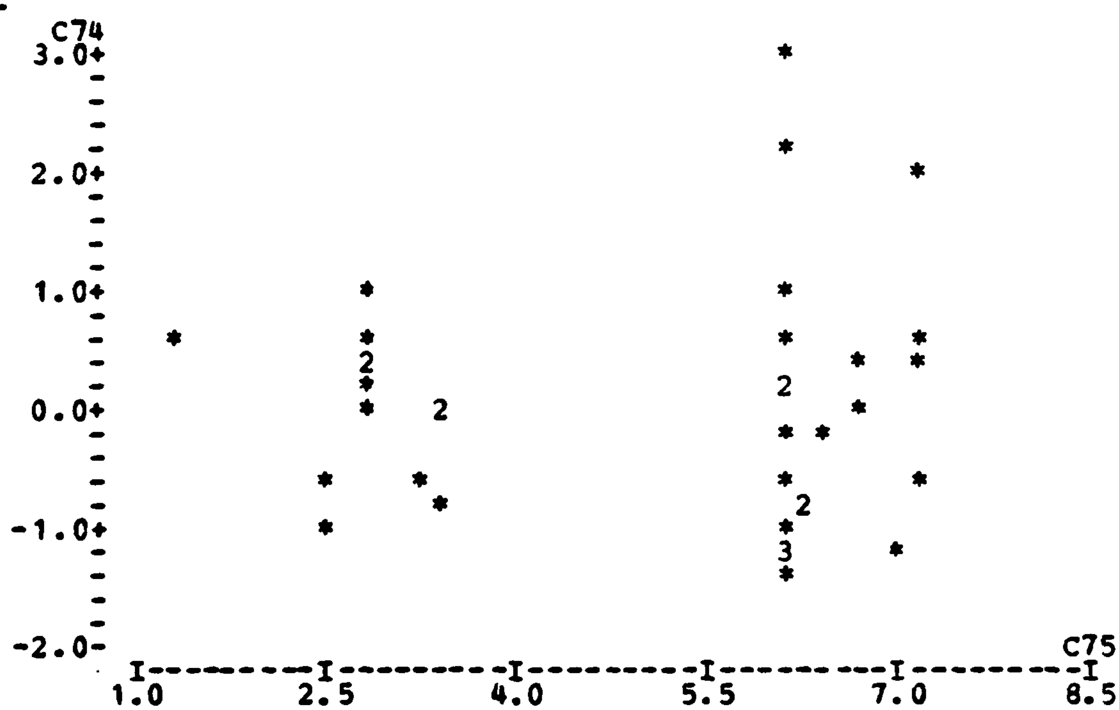
CV = 0.194

ROW	X1 PWRID	Y PROPCOST	PRED. Y VALUE	ST.DEV. PRED. Y	RESIDUAL	ST.RES.
19	5.3	1.650	2.522	0.426	-0.872	-0.98 X
20	5.3	2.020	2.522	0.426	-0.502	-0.56 X
28	14.9	8.970	6.129	0.203	2.841	2.93R
29	14.9	8.280	6.145	0.204	2.135	2.20R
30	12.1	8.990	7.082	0.353	1.908	2.06R

R DENOTES AN OBS. WITH A LARGE ST. RES.
X DENOTES AN OBS. WHOSE X VALUE GIVES IT LARGE INFLUENCE.

DURBIN-WATSON STATISTIC = 0.99

--
aver c21
   AVERAGE =   5.1008
--
plot c74 vs c75
--



--
note: Transformation (c56 = loge c21)
--
--
2).regr c56 2 c5 c6, st. res c76 pred. y c77
--

THE REGRESSION EQUATION IS
Y = - 0.433 +0.0572 X1 +0.0001 X2

```
                                               ST. DEV.       T-RATIO =
          COLUMN        COEFFICIENT            OF COEF.       COEF/S.D.
            --            -0.43341             0.08407          -5.16
X1        PWRLD           0.057215             0.004342         13.18
X2        RANGE           0.000072423          0.000009172       7.90

THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
S = 0.07837
WITH (  36-  3) =  33 DEGREES OF FREEDOM

R-SQUARED = 85.2 PERCENT
R-SQUARED = 84.4 PERCENT, ADJUSTED FOR D.F.

ANALYSIS OF VARIANCE

 DUE TO       DF          SS          MS=SS/DF
REGRESSION     2       1.170980       0.585490
RESIDUAL      33       0.202675       0.006142
TOTAL         35       1.373655

F-ratio = 32.650

CV = 0.117

         X1           Y    PRED. Y   ST.DEV.
ROW     PWRLD   T-PROCO     VALUE    PRED. Y  RESIDUAL     ST.RES.
 19       5.3    0.2175    0.3745    0.0337   -0.1570     -2.22RX
 20       5.3    0.3054    0.3745    0.0337   -0.0691     -0.98 X
 28      14.9    0.9528    0.7789    0.0161    0.1739      2.27R
R DENOTES AN OBS. WITH A LARGE ST. RES.
X DENOTES AN OBS. WHOSE X VALUE GIVES IT LARGE INFLUENCE.

DURBIN-WATSON STATISTIC = 1.08
--
aver c56
   AVERAGE =       0.66778
--
plot c76 vs c77
--
    C76
     2.5     -                                    *
       -
       -
       -              *                     *
    1.5+                                        *
       -
       -              *
       -                                     *
       -              2
    0.5+                                   * *  2
       -              *  2                   *
       -    *                                *   *
       -              *                    2 *
       -                                        **
    0.5+                                     *
       -
       -         *      *                    *
       -                 *                   *
       -                                     2
   -1.5+                                     *     *
       -
       -
       -
       -         *
   -2.5-                                            C77
        I---------I---------I---------I---------I---------I
       0.25      0.40      0.55      0.70      0.85      1.00
```

```
--
note: 3.ELECTRICAL CCST CER
--

--
corr c22 c7 c8 c16
--

           ELECCOST  ELEWGT  NO-GEN
ELEWGT      0.732
NO-GEN      0.734     0.580
TKWCY       0.731     0.971   0.541

--
regr c22 2 c7 c8, st. res c78 pred. y c79
--

THE REGRESSION ECUATICN IS
Y =    0.149 +0.0039 X1 + 0.262 X2

                                           ST. DEV.     T-RATIO =
       COLUMN      COEFFICIENT             OF COEF.     COEF/S.D.
       --              0.1486               0.1919          0.77
X1     ELEWGT        0.003862             0.001011          3.82
X2     NC-GEN         0.26155              0.06780          3.86

THE ST. DEV. OF Y ABCUT REGRESSION LINE IS
S = 0.2937
WITH (  36-  3) =  33 DEGREES OF FREEDOM

R-SQUAREC = 68.0 PERCENT
R-SQUARED = 66.1 PERCENT, ADJUSTED FOR D.F.

ANALYSIS CF VARIANCE

 DUE TO      DF           SS        MS=SS/DF
REGRESSICN    2      6.05944         3.02972
RESIDUAL     33      2.84749         0.08629
TOTAL        35      8.90694

F-ratio = 35.111

CV = 0.191

          X1        Y    PRED. Y  ST.DEV.
ROW   ELEWGT ELECCOST    VALUE  PRED. Y  RESIDUAL    ST.RES.
 31      225   1.5000   2.0637   0.0912  -0.5637    -2.02R
 32      230   1.4800   2.0831   0.0951  -0.6031    -2.17R
 34      275   2.3900   2.2568   0.1334   0.1332     0.51 X

R DENCTES AN OBS. WITH A LARGE ST. RES.
X DENOTES AN OBS. WHOSE X VALUE GIVES IT LARGE INFLUENCE.

DURBIN-WATSCN STATISTIC = 1.55

--
aver c22
   AVERAGE =       1.5394
--
```

--
plot c78 vs c79
--

```
   C78
    2.0      -
       -                           *  *           *
       -                           *
       -                                        *
    1.0+ *                         *
       -                          *
       -       *                                  *   *
       - *     *                  *2
       - *     *                   *
    0.0+       *                   *
       -       2                  *
       - *                         3
       -
       -
   -1.0+                   *       *
       -       *                                *
       -
       -
       -       *                      *
   -2.0+                                        *
       -                                        *
       -
       -
       -
   -3.0-                                                C79
        I---------I---------I---------I---------I---------I
       0.90      1.20      1.50      1.80      2.10      2.40
```

--
note: 4.C&C + ELECTRONIC END COST CER
--

--
corr c32 c9 c10 c11
--

	C+E COST	C-CWGT	PROTO
C-CWGT	0.551		
PROTO	-0.042	-0.140	
MS-END	-0.007	0.484	-0.166

--
1) regr c32 1 c9,st. res c80 pred. y c81
--

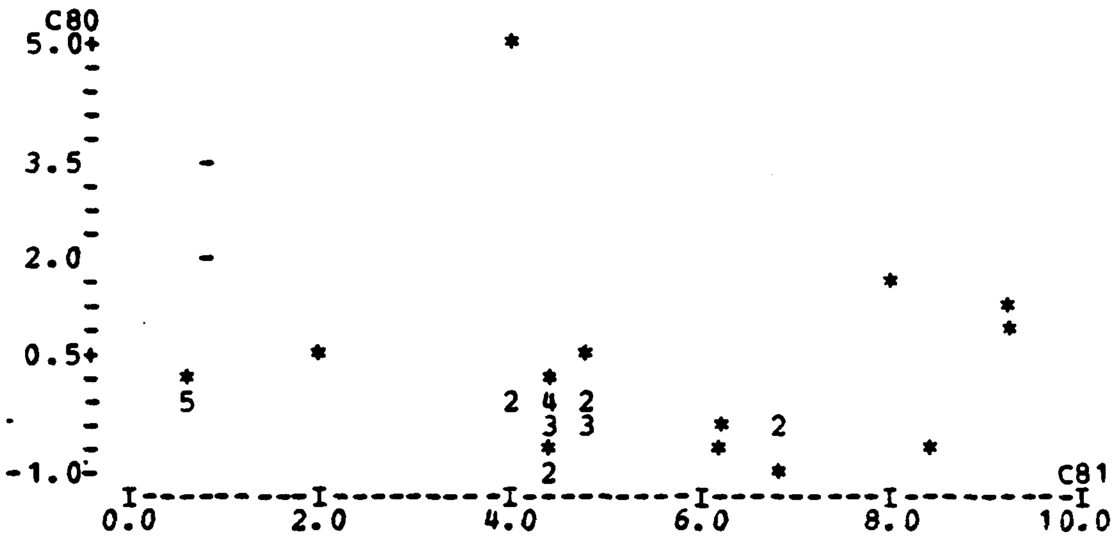
THE REGRESSION EQUATION IS
Y = - 0.769 +0.0315 X1

	COLUMN	COEFFICIENT	ST. DEV. OF COEF.	T-RATIO = COEF/S.D.
	--	-0.769	1.494	-0.51
X1	C-CWGT	0.031452	0.008162	3.85

THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
S = 3.593
WITH (36- 2) = 34 DEGREES OF FREEDOM

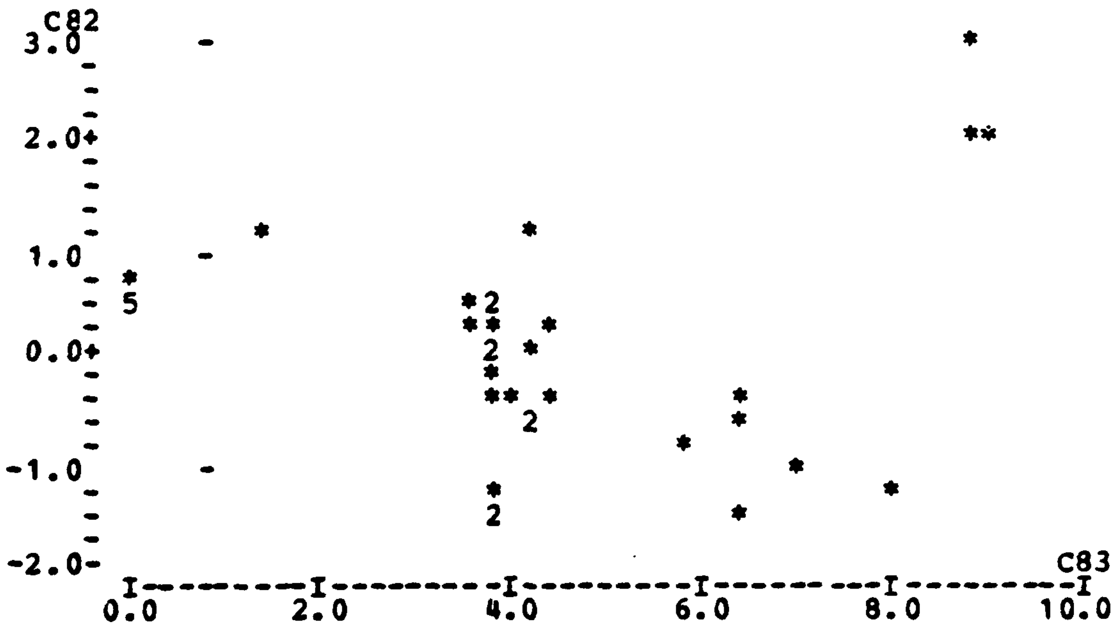
R-SQUARED = 30.4 PERCENT
R-SQUARED = 28.4 PERCENT, ADJUSTED FOR D.F.

```
ANALYSIS OF VARIANCE

 DUE TO        DF            SS        MS=SS/DF
REGRESSION      1        191.74          191.74
RESIDUAL       34        439.03           12.91
TOTAL          35        630.77

         X1         Y    PRED. Y   ST.DEV.
ROW   C-CWGT  C+E COST     VALUE   PRED. Y  RESIDUAL    ST.RES.
 24      153    21.820     4.043     0.611    17.777     5.02 R
 35      317    12.320     9.201     1.358     3.119     0.94 X
 36      318    12.500     9.233     1.365     3.267     0.98 X
R DENOTES AN OBS. WITH A LARGE ST. RES.
X DENOTES AN OBS. WHOSE X VALUE GIVES IT LARGE INFLUENCE.
DURBIN-WATSON STATISTIC = 1.75
--
plot c80 vs c81
--
```




```
--
note: remove outlier obs. 24 of c32, c9
--

--
2) regr c90 1 c91,st. res c82 pred. y c83
--

THE REGRESSION EQUATION IS
Y = - 1.51 +0.0328 X1

                                        ST. DEV.     T-RATIO =
       COLUMN      COEFFICIENT          OF COEF.     COEF/S.D.
       --             -1.5098            0.7751        -1.95
X1     C91           0.032837          0.004217         7.79

THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
S = 1.856
WITH ( 35- 2) = 33 DEGREES OF FREEDOM

R-SQUARED = 64.8 PERCENT
R-SQUARED = 63.7 PERCENT, ADJUSTED FOR D.F.
```

```
ANALYSIS OF VARIANCE

 DUE TO        DF          SS          MS=SS/DF
REGRESSION      1       208.762        208.762
RESIDUAL       33       113.618          3.443
TOTAL          34       322.379

F-ratio = 60.634

CV = 0.462

             X1            Y     PRED. Y    ST.DEV.
ROW         C91          C90      VALUE    PRED. Y  RESIDUAL    ST.RES.
 33         280       12.560      7.685      0.567     4.875      2.76R
 34         317       12.320      8.900      0.702     3.420      1.99 X
 35         318       12.500      8.933      0.706     3.567      2.08RX

R DENOTES AN OBS. WITH A LARGE ST. RES.
X DENOTES AN OBS. WHOSE X VALUE GIVES IT LARGE INFLUENCE.

DURBIN-WATSON STATISTIC = 0.74

--
aver c90
   AVERAGE =       4.0097
--


--
plot c82 vs c83
--
```



```
--
note: 5.AUXILIARY COST CER
--


--
corr c24 c31 c12 c3
--

            AUX COST PRAXWGT  AUXWGT
PRAXWGT      0.846
AUXWGT       0.764    0.886
PROWGT       0.816    0.973   0.756

--
regr c24 1 c31, st. res c84 pred. y c85
--

THE REGRESSION EQUATICN IS
Y =  0.0197 +0.0023 X1

                                           ST. DEV.      T-RATIO =
      COLUMN      COEFFICIENT             OF COEF.      COEF/S.D.
      --              0.0197                0.2595           0.08
X1    PRAXWGT      0.0022529            0.0002433           9.26

THE ST. DEV. OF Y ABCUT REGRESSION LINE IS
S = 0.5665
WITH (  36- 2) =  34 DEGREES OF FREEDOM

R-SQUARED = 71.6 PERCENT
R-SQUARED = 70.8 PERCENT, ADJUSTED FOR D.F.

ANALYSIS OF VARIANCE

 DUE TO        DF          SS          MS=SS/DF
REGRESSICN      1       27.5174         27.5174
RESIDUAL       34       10.9128          0.3210
TOTAL          35       38.4303

F-ratio = 85.724

CV = 0.251

            X1         Y   PRED. Y   ST.DEV.
ROW  PRAXWGT  AUX COST    VALUE   PRED. Y  RESIDUAL      ST.RES.
  1     1145    1.0800    2.5992   0.1014   -1.5192      -2.73R
 12     1126    4.0200    2.5564   0.0998    1.4636       2.62R

R DENCTES AN OBS. WITH A LARGE ST. RES.

DURBIN-WATSON STATISTIC = 1.76

--
aver c24
   AVERAGE =       2.2575
--
```

```
--
plot c84 vs c85
--

  C84
  3.0    -                                       *
     -
     -
     -
  1.5+                                   *  2
     -                                                   *  *
     -
     -            2**                       *
     -             *                           *          2
  0.0+   3                                  *             *
     -   2          *                       *
     -   *             *                    *2  *
     -                                             *         *
     -                                   *  *
 -1.5+                                                      *
     -
     -
     -
     -                                      *
 -3.0-                                                           C85
       I---------I---------I---------I---------I---------I
     0.70      1.40      2.10      2.80      3.50      4.20

--
note: 6.OUTFITTING COST CER
--

--
CORR C25 c1 c13
--

        OUTFCOST      LSW
LSW        0.844
OUTWGT     0.821    0.960

--
regr c25 1 c1, st. res c86 pred. y c87
--

THE REGRESSION EQUATION IS
Y =    0.436 +0.0004 X1

                                           ST. DEV.      T-RATIO =
        COLUMN       COEFFICIENT           OF COEF.      COEF/S.D.
         --             0.4358              0.1447          3.01
X1      LSW          0.00038576         0.00004212          9.16

THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
S = 0.3198
WITH (  36- 2) =  34 DEGREES OF FREEDOM

R-SQUARED = 71.2 PERCENT
R-SQUARED = 70.3 PERCENT, ADJUSTED FOR D.F.
```

ANALYSIS OF VARIANCE

DUE TO	DF	SS	MS=SS/DF
REGRESSION	1	8.5824	8.5824
RESIDUAL	34	3.4780	0.1023
TOTAL	35	12.0604	

F-ratio = 83.894

CV = 0.192

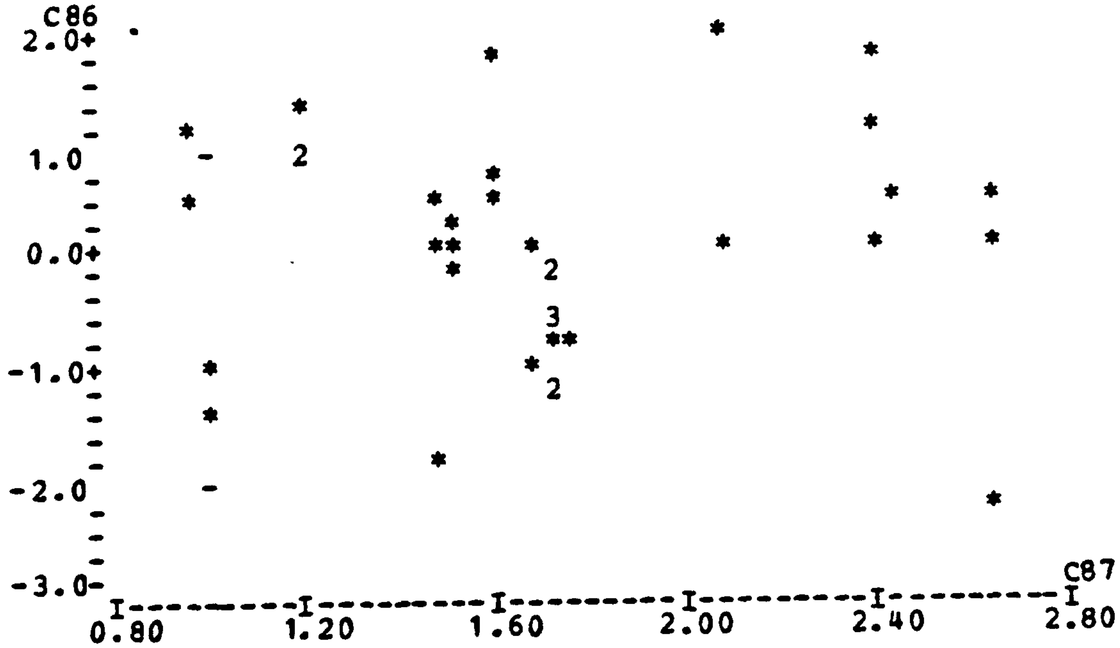
ROW	X1 ISW	Y OUTFCOST	PRED. Y VALUE	ST.DEV. PRED. Y	RESIDUAL	ST.RES.
28	4150	2.6700	2.0367	0.0668	0.6333	2.02R
34	5592	2.7200	2.5930	0.1142	0.1270	0.43 X
35	5613	2.6100	2.6011	0.1149	0.0089	0.03 X
36	5618	1.8900	2.6030	0.1151	-0.7130	-2.39R X

R DENOTES AN OBS. WITH A LARGE ST. RES.
X DENOTES AN OBS. WHOSE X VALUE GIVES IT LARGE INFLUENCE.

DURBIN-WATSON STATISTIC = 1.22

--
aver c25
   AVERAGE =   1.6683
--

--
plot c86 vs c87
--



```
--
note: 7.ARMANENT + WEAPONS END COST CER
--


--
corr c33 c14 c11 c7 c13
--

          A+W COST   ARMWGT   MS-END   ELEWGT
ARMWGT       0.868
MS-END       0.787    0.681
ELEWGT       0.805    0.820    0.473
CUTWGT       0.783    0.857    0.511    0.920


--
regr c33 1 c14, st. res c88 pred. y c89
--

THE REGRESSION EQUATICN IS
Y = -  5.74 +0.0825 X1

                                             ST. DEV.      T-RATIO =
       COLUMN      COEFFICIENT              OF COEF.      COEF/S.D.
        --            -5.737                  1.961          -2.93
X1     ARMWGT        0.082469              0.008079          10.21

THE ST. DEV. OF Y ABCUT REGRESSION LINE IS
S = 5.208
WITH (  36- 2)  =  34 DEGREES OF FREEDOM

R-SQUARED = 75.4 PERCENT
R-SQUARED = 74.7 PERCENT, ADJUSTED FOR D.F.

ANALYSIS OF VARIANCE

 DUE TO      DF         SS         MS=SS/DF
REGRESSICN    1      2826.33        2826.33
RESIDUAL     34       922.20          27.12
TOTAL        35      3748.53

F-ratio = 104.215

CV = 0.426

         X1         Y   PRED. Y  ST.DEV.
ROW   ARMWGT A+W COST    VALUE  PRED. Y RESIDUAL    ST.RES.
 30      376   40.610   25.271    1.546   15.339      3.08R

R DENCTES AN OBS. WITH A LARGE ST. RES.

DURBIN-WATSON STATISTIC = 1.36


--
aver c33
   AVERAGE =       12.213
--
```

```
--
plot c88 vs c89
--

    C88
    3.6      -                                                *
         -
         -
         -
    2.2      -                                                      *
         -           *              *
         -
         -
    0.8+        *                                    *
         -      * *    *                             *
         -
         -          4  *            4*                    2
         -              3                    2
    -0.6+                            **
         -                           *2              *
         -
         -
         -                            *
    -2.0-                             2                        C89
          I---------I---------I---------I---------I---------I
       -8.0       0.0       8.0      16.0      24.0      32.0

--
note: 8.DESIGN & ENGINEERING COST CER
--


--
corr c27 c14 c15 c9 c1 c2 c3
--
```

	D-E COST	ARMWGT	PROLSW	C-CWGT	LSW	HULLWGT
ARMWGT	0.310					
PROLSW	0.843	0.057				
C-CWGT	0.276	0.830	0.121			
LSW	0.367	0.886	0.170	0.971		
HULLWGT	0.382	0.768	0.214	0.944	0.972	
PROWGT	0.267	0.965	0.043	0.797	0.836	0.693

```
--
regr c27 2 c14 c15, st. res c94 pred. y c95
--

THE REGRESSION EQUATION IS
Y = - 1.01 +0.0065 X1 +0.0015 X2
```

	COLUMN	COEFFICIENT	ST. DEV. OF COEF.	T-RATIO = COEF/S.D.
	--	-1.0134	0.4997	-2.03
X1	ARMWGT	0.006543	0.002032	3.22
X2	PROLSW	0.0015178	0.0001498	10.13

```
THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
S = 1.308
WITH (  36- 3) =  33 DEGREES OF FREEDOM

R-SQUARED = 78.0 PERCENT
R-SQUARED = 76.7 PERCENT, ADJUSTED FOR D.F.
```

ANALYSIS OF VARIANCE

DUE TO	DF	SS	MS=SS/DF
REGRESSICN	2	200.283	100.141
RESIDUAL	33	56.430	1.710
TOTAL	35	256.712	

F-ratio = 58.562

CV = 0.855

ROW	X1 ARMWGT	Y D-E COST	PRED. Y VALUE	ST.DEV. PRED. Y	RESIDUAL	ST.RES.
1	260	1.510	5.033	0.391	-3.523	-2.82R
4	252	8.550	5.423	0.425	3.127	2.53R
26	111	0.510	3.826	0.435	-3.316	-2.69R
30	376	11.570	9.050	0.733	2.520	2.33RX
34	325	10.460	9.601	0.778	0.859	0.82 X

R DENOTES AN OBS. WITH A LARGE ST. RES.
X DENOTES AN OBS. WHOSE X VALUE GIVES IT LARGE INFLUENCE.

DURBIN-WATSCN STATISTIC = 1.83

```
--
aver c27
   AVERAGE =       1.5289
--

--
plot c94 vs c95
--
     C94
     3.0-        -
        -                                  *                          *
        -
        -
     1.5-        -    *
        -  *2                                                           *
        -   *3       *
        -           **
     0.0+           7
        -           4**2
        -            *  *   *
        -               *
        -
    -1.5+                          *
        -
        -
        -
        -                          *      *
    -3.0-                                                                C95
         I---------I---------I---------I---------I---------I
      -1.0       1.5       4.0       6.5       9.0      11.5
```

--
note: 9.CCNSTRUCTION SERVICES + MISCELLANEOUS END COST CER
--

--
corr c34 c1 c15 c9 c4 c2
--

	C+M COST	LSW	PROLSW	C-CWGT	AR/LSW
LSW	0.766				
PROLSW	0.354	0.170			
C-CWGT	0.775	0.971	0.121		
AR/LSW	-0.099	0.319	-0.120	0.282	
HULLWGT	0.822	0.972	0.214	0.944	0.119

--
1) regr c34 3 c2 c15 c4, st.res c96 pred. y c97
--

THE REGRESSION ECUATICN IS
Y = 1.84 +0.0047 X1 +0.0004 X2
- 30.4 X3

	CCLUMN	COEFFICIENT	ST. DEV. OF COEF.	T-RATIC = COEF/S.D.
	--	1.839	1.234	1.49
X1	HULLWGT	0.0047349	0.0005480	8.64
X2	PRCLSW	0.0003732	0.0002184	1.71
X3	AR/LSW	-30.35	15.91	-1.91

THE ST. DEV. OF Y ABCUT REGRESSION LINE IS
S = 1.844
WITH (36- 4) = 32 DEGREES OF FREEDOM

R-SQUARED = 73.9 PERCENT
R-SQUARED = 71.4 PERCENT, ADJUSTED FOR D.F.

ANALYSIS CF VARIANCE

DUE TO	DF	SS	MS=SS/DF
REGRESSICN	3	307.381	102.460
RESIDUAL	32	108.837	3.401
TOTAL	35	416.217	

F-ratio = 30.126

CV = 0.290

--
2) regr c34 1 c2,st. res c71 pred. y c72
--
THE REGRESSICN ECUATICN IS
Y = 0.0029 +0.0048 X1

	COLUMN	COEFFICIENT	ST. DEV. OF COEF.	T-RATIO = COEF/S.D.
	--	0.0029	0.8256	0.00
X1	HULLWGT	0.0048121	0.0005721	8.41

THE ST. DEV. OF Y ABCUT REGRESSION LINE IS
S = 1.993
WITH (36- 2) = 34 DEGREES OF FREEDOM

R-SQUARED = 67.5 PERCENT
R-SQUARED = 66.6 PERCENT, ADJUSTED FOR D.F.

ANALYSIS OF VARIANCE

DUE TO	DF	SS	MS=SS/DF
REGRESSION	1	281.133	281.133
RESIDUAL	34	135.085	3.973
TOTAL	35	416.218	

F-ratio = 70.761

CV = 0.313

ROW	X1 HULLWGT	Y C+M COST	PRED. Y VALUE	ST.DEV. PRED. Y	RESIDUAL	ST.RES.
24	1232	10.880	5.931	0.336	4.949	2.52R
34	2568	14.740	12.360	0.787	2.380	1.30 X
35	2574	9.710	12.389	0.790	-2.679	-1.46 X
36	2574	10.330	12.389	0.790	-2.059	-1.13 X

R DENOTES AN OBS. WITH A LARGE ST. RES.
X DENOTES AN OBS. WHOSE X VALUE GIVES IT LARGE INFLUENCE.

DURBIN-WATSON STATISTIC = 1.91

```
--
aver c34
   AVERAGE =      6.3606
--

--
plot c71 vs c72
--

  C71
 3.0      -
    -               *
    -
    -
 2.0+                     *
    -              **
    -
    -
    -              *                    *    *
 1.0      -
    -      *       **
    -   *          *
    -              *                  *
 0.0+                                 *
    -
    -  2       *   23
    -  2*          **
    -          2   *      *
-1.0      -              *                           *
    -                                         *
    -
    -
-2.0-                                 *                C72
      I---------I---------I---------I---------I----------I
     2.5       5.0       7.5      10.0      12.5      15.0
```

--
note: 10.TOTAL COST CER
--

--
corr c29 c1 c35 c2 c3 c15 c11 c94
--

	TCOST	LSW	ENGWGT	HULLWGT	PROWGT	PROLSW	MS_END
LSW	0.913						
ENGWGT	0.866	0.953					
HULLWGT	0.886	0.972	0.864				
PROWGT	0.758	0.836	0.959	0.693			
PROLSW	0.332	0.170	0.127	0.214	0.043		
MS_END	0.812	0.781	0.775	0.735	0.716	0.068	
PWELD	0.269	0.288	0.519	0.077	0.699	-0.161	0.312

--
1) regr c29 4 c1 c11 c15 c5, st. res c98 pred. y c99
--

THE REGRESSION EQUATION IS
Y = - 7.33 +0.0123 X1 + 7.19 X2
 +0.0030 X3 +0.0022 X4

	COLUMN	COEFFICIENT	ST. DEV. OF COEF.	T-RATIO = COEF/S.D.
	--	-7.329	5.430	-1.35
X1	LSW	0.012301	0.001141	10.78
X2	MS-END	7.193	2.182	3.30
X3	PROLSW	0.0029769	0.0008358	3.56
X4	PWELD	0.0022	0.4084	0.01

THE ST. DEV. OF Y ABOUT REGRESSION LINE IS
S = 6.997
WITH (36- 5) = 31 DEGREES OF FREEDOM

R-SQUARED = 90.2 PERCENT
R-SQUARED = 88.9 PERCENT, ADJUSTED FOR D.F.

ANALYSIS OF VARIANCE

DUE TO	DF	SS	MS=SS/DF
REGRESSION	4	13969.46	3492.36
RESIDUAL	31	1517.49	48.95
TOTAL	35	15486.95	

F-ratio = 72.345

CV = 0.180

--
2) regr c29 3 c1 c11 c15,st.res c73 pred. y c74
--

THE REGRESSION EQUATION IS
Y = - 4.14 +0.0107 X1 + 9.15 X2
 +0.0029 X3

	COLUMN	COEFFICIENT	ST. DEV. OF COEF.	T-RATIO = COEF/S.D.
	--	-4.138	3.582	-1.16
X1	LSW	0.010725	0.001515	7.08
X2	MS END	9.155	2.913	3.14
X3	PROLSW	0.0028622	0.0008232	3.48

THE ST. DEV. OF Y ABCUT REGRESSION LINE IS
S = 7.054
WITH (36- 4) = 32 DEGREES OF FREEDOM

R-SQUARED = 89.7 PERCENT
R-SQUARED = 88.8 PERCENT, ADJUSTED FOR D.F.

ANALYSIS CF VARIANCE

DUE TO	DF	SS	MS=SS/DF
REGRESSICN	3	13894.53	4631.51
RESIDUAL	32	1592.40	49.76
TOTAL	35	15486.94	

F-ratio = 93.077

CV = 0.182

ROW	X1 ISW	Y TCOST	PRED. Y VALUE	ST.DEV. PRED. Y	RESIDUAL	ST.RES.
24	2620	44.70	23.96	2.03	20.74	3.07R
28	4150	67.59	49.53	1.70	18.06	2.64R
30	5009	93.90	82.23	4.27	11.67	2.08RX
34	5592	80.04	81.00	4.67	-0.96	-0.18 X
35	5613	59.50	65.22	3.45	-5.72	-0.93 X
36	5618	62.57	65.27	3.46	-2.70	-0.44 X

DURBIN-WATSON STATISTIC = 2.21
--
aver c29
   AVERAGE =   38.848
--
plot c73 vs c74
--

```
     C73
    3.5-
       -              *
       -
       -                           *
    2.5-
       -                                           *
       -
       -                      *
    1.5-
       -
       -                                    *
       -
    0.5+                          *
       -       4       2     *
       -                      *
       -              *       2                   *
       -         *  *         *
   -0.5+              *       2*   *       *
       -                      *             *
       -         *            *            *
       -                  *2
       -                                    *
   -1.5-                                                   C74
          I---------I---------I---------I---------I---------I
          0.       20.       40.       60.       80.      100.
```

LIST OF REFERENCES

1. Batchelder, C. A., An Introduction to Equipment Cost Estimating, RAND Corp., RM-6103-sa, 1969.

2. Baker, B. N., Improving Cost Estimating and Analysis in DOD and NASA, Ph.D. Thesis, George Washington University, Washington, D.C., 1972.

3. Booz-Allen Applied Research, Inc., Cost Management Methods, 1971.

4. Bruce, M. M. and Sovereign, M. G., Parametric Cost Estimating with Application to Sonar Technology, 1973.

5. Bull, N. S., and Sandrini, L. M., Parametric Cost Estimating, paper prepared for Defence Systems Analysis Program, 1972.

6. Center for Naval Analysis Report, Escort Ship Cost Model (ESCOMO), by R. V. Wilson.

7. Chief of Naval Operations, OP-96D, Independent Cost Review, S-3A Program, 1972.

8. Chief of Naval Personnel, Advanced Submarine Sonar Technology, Volume 2, NAVPERS 93595, 1965.

9. Frank, C. R., Statistics and Econometrics, Hort, Rinehart and Winston, 1971.

10. Graybill, F. A., An Instruction to Linear Statistical Models, McGraw-Hill, 1961.

11. Helmer, F. T., Bibliography on Pricing Methodology and Cost Estimating, Dept. of Economics and Management, USAFA, Colorado, 1972.

12. Hernon, D. M. and McCumber, R. R., Estimation of Destroyer Type Naval Ship Procurement Costs, Master's Thesis, Naval Postgraduate School, 1973.

13. Klein, M. R., Treating Uncertainty in Cost Analysis: The Beta Distribution, (Preliminary version), Resource Analysis Group, OP-96D, The Pentagon, 1973.

14. Military Equipment Cost Analysis, RAND, prepared for the Office of the Secretary of Defense (Systems Analysis), 1971.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93943	2
3.	Professor Michael G. Sovereign, Code 55Zo Department of Operations Research Naval Postgraduate School Monterey, California 93943	8
4.	Associate Professor Shu S. Liao, Code 54Lc Department of Administrative Sciences Naval Postgraduate School Monterey, California 93943	1
5.	LCDR Baik, In Hwa Yimho Jutaek B Dong 101 4-100 SUBINGKODONG YONGSANKU Seoul, Korea	7

END

FILMED

5-84

DTIC