AN INVESTIGATION OF THE RELATIVE ACCURACY OF THREE-TIME VERSUS SINGLE-TIME ESTIMATING OF ACTIVITY DURATION IN PROJECT MANAGEMENT

By

FREDRICK W. SWIFT

Bachelor of Science
University of Notre Dame
Notre Dame, Indiana
1959

Master of Science
University of Alabama
Tuscaloosa, Alabama
1966

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
May, 1971

AN INVESTIGATION OF THE RELATIVE ACCURACY OF
THREE-TIME VERSUS SINGLE-TIME ESTIMATING
OF ACTIVITY DURATION IN
PROJECT MANAGEMENT

Thesis Approved:

Earl J Ferguson
Thesis Adviser

James E Shamblin

M. Palmer Terrell

J. Leroy Folks

D. D. Durham
Dean of the Graduate College

PREFACE

The objective of this research is to investigate the relative accuracy of different estimating procedures in project management. The desirability of getting accurate, dependable estimates early in projects is vitally important to reflect actual, ultimate time and cost of the program. Since many crucial decisions are made during the initial phase, it is greatly advantageous to the project manager to know activity times as accurate as possible in the early planning stages.

It is the thesis of this study that extraneous psychological factors contaminate an honest estimate of work to be completed at some future date. Various uncertainties, due to unknowable facts such as a change in manufacturing techniques, new materials that might be required or even unknowns in the labor conditions that might prevail will always exist and cannot be eliminated. Improvements can be achieved, however, by circumventing factors involving certain aspects of human nature.

A procedure is developed whereby it is possible to determine if advantages exist with the three-time estimating procedure over the single-time estimate. Comparison of the data is made and an estimate of the difference in relative accuracy is established.

I would like to take this opportunity to express my sincere appreciation to the members of my graduate committee for their advice and guidance: Dr. Earl J. Ferguson, Thesis Adviser, who gave generously of his time and suggestions; Dr. James E. Shamblin, Committee Chairman, for his assistance and guidance; and Dr. M. P. Terrell, all of the School of Industrial Engineering and Management; and Dr. J. L. Folks, Chairman of the Statistics Unit.

I would like to recognize the Marshall Space Flight Center (NASA) for sponsoring my academic program and for permitting the use of supporting data for this study. Also, Mr. Roy Ferguson, of the LTV Corporation, gave valuable assistance and made many helpful suggestions.

I want to thank Velda Davis for the excellence in preparing my manuscript and for editorial suggestions as well as handling the details of submission, etc.

I wish to express a special thanks to my supervisor, Mario H. Rheinfurth, who not only permitted me time, but also provided invaluable suggestions and advice throughout my graduate pursuits.

Finally, I am deeply grateful for the patience and understanding from my wife, Pat, who typed my many rough drafts; and my children, Bill, Melinda, Mike, and Brian, who granted me time alone. It is my sincere desire that they share in the rewards of this experience as they have had to bear the burdens of its trials.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

PROBLEM DEFINITION

Introduction

The widespread adoption of critical path methods (CPM) and the Program Evaluation and Review Techniques (PERT) during the past decade stands as proof of the real value of these management tools in industrial and governmental environment. Many variations of the critical path methods such as line-of-balance (LOB), activity balance line evaluation (ABLE), and continuous milestone, are extensively used in management control today. Conceptually these methods all have a common goal; that is, to control a specific project so that time and cost schedules may be met. The theoretical foundation of these methods is sound and the underlying assumptions and mathematical concepts are well known and understood. A brief summary of the important features of CPM and PERT is given in the next chapter of this thesis, but for a rigorous derivation, Kelley (16) or Moder and Phillips (22) may be consulted.

The basic features of CPM and PERT have much in common, with the differences in each system directly related to the fundamental aim or nature of the project being managed. For those projects which have relatively well-defined objectives and which use technologies of relatively well-known predictability, it is common practice to use CPM exclusively. There also exists a reasonably well-established practice of using PERT for planning and control of projects having a relatively high degree of uncertainty associated with the activities comprising a network. Closely related to these techniques are the companion concepts of CPM/cost and PERT/cost.

Computer Utilization for Network Calculations

When the number of activities is large, the computer is normally utilized to facilitate routine calculations, and the cost of running the computer must be added to the total project cost. This factor suggests that the input data, in the form of activity time estimates, need to be as accurate as possible in the early stages of the project. Whenever activity time estimates are modified, a new computer run must be made at additional expense and in the event of close profit margin, this can be critical in terms of profitability or even company survival.

In the earliest phase of the program, the manager must decide on the type of analysis he will make. The computer programs which are available will have differences in input and output data; each program may have a different cost to install and debug on the computer and, of course, computer time for each run may be different. For the network programs currently available, a general rule is that the CPM (time only) is the least expensive followed by PERT (time only), CPM/cost and the most expensive is PERT/cost.

Another factor when choosing between PERT and CPM is the form of the results. Many people have difficulty interpreting the probability statements after a PERT analysis has been made and, therefore, probability calculations are extra cost without adding any value. The decision to use PERT or CPM, therefore, may become one of weighing the extra cost of running a PERT program against the risk of modifying the single-time estimate permitted in the CPM analysis.

Individual Company Modifications

Many companies were interviewed which have adopted their own variation of CPM and use it merely as a guide; others have been frustrated with the results in their model and discontinued its use altogether. One of the more frequent complaints among users is that the individual activity estimates are not accurate and must be changed after the project is under way. The cost can become prohibitive, and as one manager pointed out, once changes are permitted, more time is spent controlling the program than on controlling the project.

When schedules do slip and program milestones are not met, each company has its own system of "fixes", but most frequently the solution is to spend more money. Since money can usually "buy" time, many projects are found to use an excessive amount of overtime or to be burdened with an excessive number of people. Normal crew sizes grow to inefficient sizes, and costs often sky rocket. As shown in a report by Croft (7), most complex, large scale defense and space programs often have associated costs that were grossly underestimated.

Specific Area of Emphasis

One of the basic ingredients of PERT or CPM is the estimated activity times. In the PERT analysis, three estimated times are required while CPM analysis requires only one estimate. This is the most critical aspect of the entire analysis; yet, in developing this management technique, it has not received adequate study. Much of the discussion on time estimates up to the present has put the emphasis on who should make the estimates, how they should be made and level of detail in the estimates (13, 22). This thesis discusses some of the reasons a person might intentionally or unconsciously give inaccurate estimates. Some estimators may want to make an impression on the boss by making pessimistic estimates and then finishing the job early; others may intentionally underestimate so they can later work overtime. A study is made using both types of estimates which demonstrates that more reliable results can be obtained by using the three-time estimating procedure.

Research Hypothesis

It is hypothesized that an estimate of an activity duration time made by a single-time estimate procedure is a statistically significant variation from a mean-time estimate for that same activity by the same person using a three-time estimate procedure. If this hypothesis is correct, a practical advantage can be gained by using the three-time estimate procedure combined with the less costly and more easily understood CPM analysis. The advantages to be gained are more than economic. The attendant psychological advantage of knowing the system is working becomes an incentive for the project manager and intangible benefits are derived from increased morale.

In addition to a discussion of the general factors that are involved in making decisions relating to activity estimates, this thesis will choose a model whereby a manager can determine if a significant difference exists in the two methods for his particular company or division. A comprehensive research study was performed during which many activities from a variety of industries were reviewed. Six were then selected from a particular industry for the purpose of making a detail test. Using these selected activities, estimates were obtained using both the single-time estimate procedure and the three-time estimate procedure and it was established that a significant difference (in both a practical and a statistical sense) exists between the two methods. The null hypothesis can be stated as follows: H_o: the mean of the single-time estimates is equal to the mean of the three-time estimates. The alternate hypothesis is that the means are not equal. In Chapter III Mahalanobis' significance test is applied and a F-statistic is calculated.

The results in Chapter III show that the null hypothesis must be rejected. This means that the mean of the single-time estimates is not the same as the mean of the three-time estimates.

Psychological Factors

One reason for the difficulty in obtaining accurate activity estimates is the various motives of the people involved in the subsystems. In complex systems today, there are many people involved at all levels of the project. Not only are planners from different departments trying to make estimates that interact with each other but oftentimes different and diversified contractors and sub-contractors must provide crucial inputs that will impact the entire program. At each level, the human element, both individually and in complex combinations, must be skillfully treated if the over-all project is going to be successfully planned. If it could be determined which human factors cause a bias, and how much bias each causes, a more accurate estimate could be determined for each activity and ultimately for the entire program. A good manager must, therefore, be able to provide for the human (or psychological) needs of people just as he must provide wages and acceptable working conditions. Towards this end, it is suggested that the three-time estimate technique be used in the CPM method of management control. This will provide a psychological margin of safety while maintaining a margin of latitude in either direction.

From the investigation, it became apparent that if the concept of program control is not accepted at some level of management, the project will not be done as estimated regardless of the technique. One high level manager made the comment, "milestones exist only in the minds of the planners." Clearly when this feeling exists in top management, accurate estimates would not be expected at lower levels. Likewise, if one contractor does not believe in using modern management techniques, he may give hastily prepared estimates not realizing that success for the entire project depends on the accuracy of his portion of the network.

An interesting observation was made by an executive of a large construction company. He indicated that his experience in the construction industry revealed that subcontractors who were not familiar with quantitative management techniques would give ambitious estimates while those knowledgeable in the area of PERT and CPM tended to give pessimistic estimates. This study would substantiate that idea. If estimators are aware that their portion of the project might be on the critical path, which means that their performance will be closely monitored, they tend to be conservative. If a penalty charge will be imposed for late finishing dates, a pessimistic estimate might be given.

Interdependence of Time, Cost and Performance

The concepts of CPM and PERT were primarily developed in the 1950's. These concepts were refined and further developed during the 1960's to include the interdependence of performance, time and cost. These three parameters are the most important from the managers viewpoint, but many other psychological factors become involved when planners of different organizations must be coordinated to produce a good CPM chart. Each organization, and each planner in each organization has a different set of goals. Factors such as the planners relationship to the job (does he do the work or supervise others doing the work), how many people can work efficiently on the activity, and is the activity long or short term, all enter, either consciously or unconsciously, into planning a project. These factors should not influence estimates but since people are involved, they usually do. Some estimates may be intentionally biased for various motives, while others may not give proper attention to detail which produces inaccurate results. It is an extremely difficult task to determine which psychological factors are operating and virtually impossible to measure the thought process involved when an individual plans a project. His own motives are complex and these are compounded with the organization which has its own set of ultimate goals. As one cynic has observed, "Its always easier to arrive at a firm conviction about a problem after you know what the boss thinks."

Responsibility for Estimates

In today's complex society, the fear of making mistakes (over estimating a project, resulting in loss of contract, or underestimating, resulting in penalty payments) coupled with the problem of responding appropriately to management, results in one finding himself in an undesirable situation. Writing on organizational stress, Kahn (15) states:

> We are perhaps more familiar than any other people with the twin dilemmas of responding appropriately to opposing and incompatible forces, or to forces for which both source and target remain obscure. Conflict and ambiguity are among the major characteristics of our society, and we are marked by them.

It is felt, therefore, that the estimates that people must make and ultimately assume responsibility for, are often biased by factors which should not be relevant to the analysis. To compensate for this dilemma, it is suggested that planners be provided with a "psychological cover" via a three-time estimate procedure while retaining all the advantages of the lower cost and more simply understood CPM network analysis. It will be argued by some that a planner will merely produce his estimate with equal bounds on either side. Even if this should occur, a better estimate should result since he knows a certain latitude is available and his estimate will still be "right".

Experience and Judgment Needed for Estimates

When a person is asked for an estimate, he is in a sense being asked to predict an event that will take place at some future time. To do this he must rely on his past experiences and try to relate what previously occurred to what he anticipates will happen in the future. The current and future environment may be changing and some difficult judgments might be necessary to successfully complete a project. Kahn, Wolfe et al. (15) distinguish between the objective environment and the psychological environment. They state the objective environment of a person consists of "real" objects and events, verifiable outside his consciousness and experience.

In a practical sense, the objective environment here would involve the physical operations necessary to complete a given task. In the construction of a missile, for instance, each component would have to be manufactured entailing a series of drilling, boring, and machining operations. The objective environment would include the time required to obtain tools and jigs and to complete the particular part.

The psychological environment of a person would be the conscious and unconscious representation of the objective environment. The individual could view the objective environment of producing a missile part and intentionally underestimate the time required in the expectation of working overtime at a later date. Conversely, he could unconsciously overestimate the time required if he views the situation as being different from what it really is. The way a person views a situation, thus, becomes more important than the true situation itself. The person, in effect, "creates" his own reality.

The objective environment manifests sufficient uncertainties because of the "level of the problem and the nature of information available" without the imposition of attendant psychological factors. When projects are well defined and the operator has repeated the task many times before, the decisions which are necessary in making the CPM charts are not generally subject to large errors. On the other hand, the real necessity for CPM and PERT becomes apparent when new projects are attempted. When one attempts to estimate "how long or "how much," he must rely on intuition and judgment since he often does not have detailed information on the activity. In discussing this problem with managers in various industries, they all contend that each project has unique features that did not have to be considered in similar projects.

Since it is apparent that one must apply the judgement and experience factors with their attendant uncertainties to most large projects, it is most advantageous to minimize all extraneous factors. Regardless of how good the system is or how good the procedure works, people must invariably be in positions of control, they must rely on past experience, exercise some judgment, make decisions, and ultimately be responsible for these decisions.

Quantitative management has been practiced for a long time but probably has not yet reached its full potential. There are many refinements to be made but often these refinements are made in the mathematical model instead of in the application of the model. It appears that some poor decisions have been made using highly refined models; whereas, much improvement could be attained by better use of existing models. It is proposed that this objective be accomplished by using the three-time estimate technique as a management control discipline.

CHAPTER II

HISTORICAL REVIEW OF SOME MANAGEMENT CONTROL TECHNIQUES

Critical Path Method

The beginning of a new era in management control resulted from the work of Morgan R. Walker and James E. Kelley in 1959. The work on this new concept began in 1956 and culminated with the publishing of their work in a paper entitled "Critical Path Planning and Scheduling." In this paper, they developed the essentials of a technique that is presently known as critical path methods or more commonly CPM (3). Until the development of CPM, Gantt charts were probably the most universally used management tools. Gantt charts which were invented by Henry Gantt around 1900 were used extensively by contemporary managers with modifications and variations to suit their individual problems. Even today, Gantt charts are used by managers with their own innovations.

The construction and use of CPM charts are relatively straight forward and easy to master. Both CPM and PERT (discussed in the next section) are based on higher mathematics, but only simple arithmetic is needed to construct and understand the charts. To fully understand the charts, one must work through the entire system, and experience is essential to gain facility in the the technique.

The CPM network is activity oriented (job oriented) and network diagramming is necessary before one can proceed with the project. When the user begins the project, he needs to show the dependency of one job on another and in effect plan the entire project. The emphasis in this paper is given to the next phase of the project which some writers feel is the most critical. Many feel that knowledge gained in the network phase alone is worth the effort, time, and money expended in planning the project (3). They feel that this phase causes detailed planning which helps identify bottlenecks and visually present the "big picture."

Level of Detail

Another feature that should be kept in mind is that there are different levels of networks depending on the level of management. Top management's concerns are different from that of foremen. Top management is mainly interested in major milestones while a foreman needs to know all the operations that might make up one milestone. The networks, therefore, might look different but they may be for the same project. A series of related activities can become a single activity for a network at a higher level of management. Regardless of the appearance of the network, once the network is sketched, the next step is estimating the activity times. It is primarily at this stage that subjectivity becomes a factor in the analysis. In the planning phase, errors in logic can be detected but estimating cost and time requires experience and judgment. Any errors introduced during this phase may not be detected until late in the program.

In a CPM network, each activity duration is given a single-time estimate using convenient work units for each activity. The normal unit is the work day or week. Any unit is permissible but it is mandatory that the unit be consistent for the whole project. A normal level of manpower, equipment and any other standard conditions should be assumed when making the activity estimates. The estimate of an activity time should be an independent entity and should not anticipate activities that precede or follow it.

By whatever means used, the project manager is expected to finally determine a single time, which is the mean activity duration time. These times are then combined and later monitored so that the manager can actively control the whole program. The importance of good time estimates cannot be overemphasized. In addition to the normal benefits that are derived from good project management there is a valuable intangible. If the project proceeds on schedule a psychological advantage is obtained because the different organizations retain faith in the program. A frequent lament from managers indicate that confidence is lost in the total project if one area (or series of activities) fall significantly behind. In fact, one manager flatly stated that these management systems "won't work" for his type of industry. "More time was spent trying to modify, correct, or otherwise update the estimates than was spent doing the work."

Most planners agreed that if CPM or PERT could be prepared accurately it would be used more frequently than it is now. It might be mentioned here that many contracts require a PERT or COM chart and often times a consultant is needed to prepare a final analysis. Unexpected contingencies which are beyond the control of the managers are not normally included but some managers include normal weather conditions in their estimate. Most, however, prefer to simply increase the total project by a safety factor after the total project has been planned. This avoids any duplication of "padding" and the planner can assume more control of his project.

Program and Evaluation Review Technique

Program and evaluation review technique, more commonly called PERT, was developed in 1958. In its search for a modern management technique the Navy and a team from industry combined to develop a system that would meet their needs for the Polaris Missile System. They originally described PERT as "a fast, flexible management control tool for coordinating complex research and development programs." PERT was successfully applied to the Polaris Missile Project and was credited with trimming two years off the time originally thought required to design and implement a project of this size and complexity (11).

Some of the early applications of PERT include the C-141 Jet Transport at the Lockheed Georgia Company, the B-70 at North American Aviation and the Communication Satellite at the Defense Communications Agency.

The basic theoretical concepts of PERT are well known and understood by individuals who use the technique. It is not necessary to develop these concepts in this thesis but the fundamental aspects of the time estimates and what the estimates represent will be given here.

PERT Time Estimates

Under the PERT system, three possible activity duration times are estimated:

a = optimistic time - an unlikely but possible activity time duration if all goes well.

m = most likely time - a modal value of the time distribution or the activity time which would likely occur more often than another value.

b = pessimistic time - unlikely but possible activity time duration if all goes badly.

From these data it is possible to calculate the weighted average of the three-time estimates to estimate the expected time for the activity. Then the expected time, t_e, is found:

$$t_e = \frac{a + 4m + b}{6} .$$

The developmental logic underlying this weighted average formula need not be discussed here since many references to it are available in the relevant literature (21, 22).

From the three-time estimates, PERT systems use commonly known statistical procedures to calculate standard deviation, variance, and subsequently, probabilities of completion of activities to predetermined completion dates. Theoretically this calculated time should represent the mean of the beta distribution. This time represents the time which allows a 50 per cent chance of finishing the activity as scheduled. Figure 1 demonstrates the relationship of the estimated times.

In theory, it should make no difference which estimates are obtained first; in practice the extremes should be obtained so the situation can be bracketed (29). It might be noted here that many distributions can give the same mean. For instance, the following distributions 13 - 15 - 17; 3 - 16 - 23; 5 - 17 - 17 all have a mean of 15 but each is a different representation of the anticipated activity time. One should get single-time estimate of 15 for all three cases,

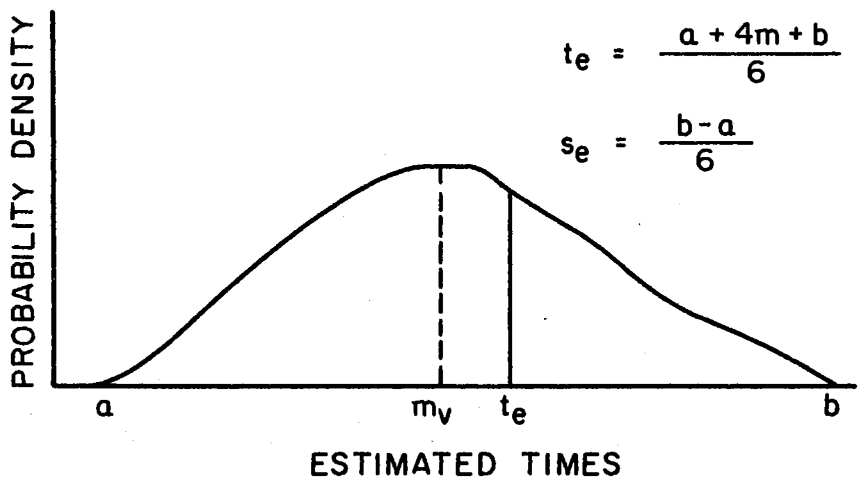


Figure 1. PERT-Time Relationship

but in actual practice one would probably get something entirely different. G. N. Stilian (29), in discussing the guidelines for PERT applications, states,

> ... an estimate is the means by which subordinates inform us of their best technical judgment of a situation. It is interesting to note in this connection that once people have become convinced that we are actually living by this rule, we can get three time estimates much more rapidly and accurately than we could previously get one commitment.

Other Techniques

Other management techniques in use today include various modifications to network scheduling but none have been as universally favored as PERT and CPM. Similarities exist between all the network theories and techniques such as line of balance (LOB) have been in use since the 1940's. The LOB technique is useful when the project manager is not in close contact with the project. It operates on the principle of exception or deviations from planned performance and not on a continuous examination of the plan as a whole. Even though this technique may be applied to a variety of programs, it finds its most frequent application in manufacturing and production operations.

Continuous Milestone

Another associated technique is called continuous milestone. The project planner does not monitor activities but instead looks at a group of activities which culminate in an easily identifiable milestone. The project manager then tries to keep all milestones on schedule and again operates on the management by exception principle. The inherent disadvantage is that as portions of the project fall behind, it is nearly impossible to find the exact cause of the delay. Often times money is inefficiently spent trying to catch up.

Many other types of control are used for special purposes and in special situations but PERT and CPM continue to gain acceptance because of their flexibility.

CHAPTER III

RESEARCH PROCEDURE

Review of Historical Data

Has PERT been more successful than CPM? Or asked another way, can it be said that PERT has been more accurate in past programs than other methods of management control techniques? An attempt was made to resolve this question, but the results are inconclusive if taken from historical data alone. The managers who use PERT exclusively are convinced that it is the best method, while those who use another method can show that their methods are more accurate.

In a recent study conducted at MIT, an investigation was made to determine the effects of PERT on R + D organizations. Based on a study of 31 projects, it was concluded that PERT leads to improvement of schedule performance without any noticeable effect on technical performance. The test was made between what was termed PERT versus NON-PERT, but the most interesting observation was the speculation on how the improvement was achieved. Seelig and Rubin (27) attribute the improvement to a large degree, to "improvement in communication." This kind of communication is not that received through the formal authority structure, but rather the informal kind that is so essential to the successful implementation of any program. This indicates that model refinement may not add any improvement, but use of a particular model may affect other factors which become the vehicle of improvement.

An independent study conducted for this paper reviews the results of 100 activities that were taken from successful application of the CPM and PERT techniques. Data was obtained from companies which have used both techniques and they were asked only for projects where they felt the technique had been used properly and successfully. The results of the study are shown graphically in the histograms of Figures 2 and 3. The sample means and variances were calculated from the relationship on page 121 of Miller and Freund (20).

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{k} x_i f_i$$

$$s^2 = \frac{n \cdot \sum_{i=1}^{k} x_i^2 f_i - \left(\sum_{i=1}^{k} x_i f_i \right)^2}{n(n - 1)}$$

Using the tabulated data in Table I, the sample means were calculated to be

$$\bar{x}_{PERT} = 35.2 \qquad s^2_{PERT} = 701$$

$$\bar{x}_{CPM} = 36.9 \qquad s^2_{CPM} = 738.7 \quad .$$

To make a significance test, the t statistic was calculated by

$$t = \sqrt{\frac{n}{s^2_{CPM} + s^2_{PERT}}} \left| \bar{x}_{CPM} - \bar{x}_{PERT} \right|$$

$$t = .448$$

when this is compared to a table value of t_α it is seen that the sample means are the same. This would seem to indicate that no difference

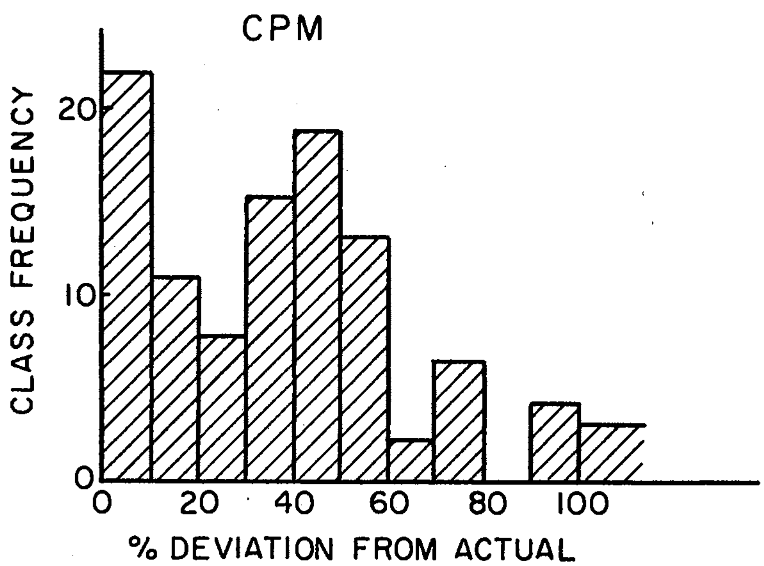


Figure 2. CPM Histogram

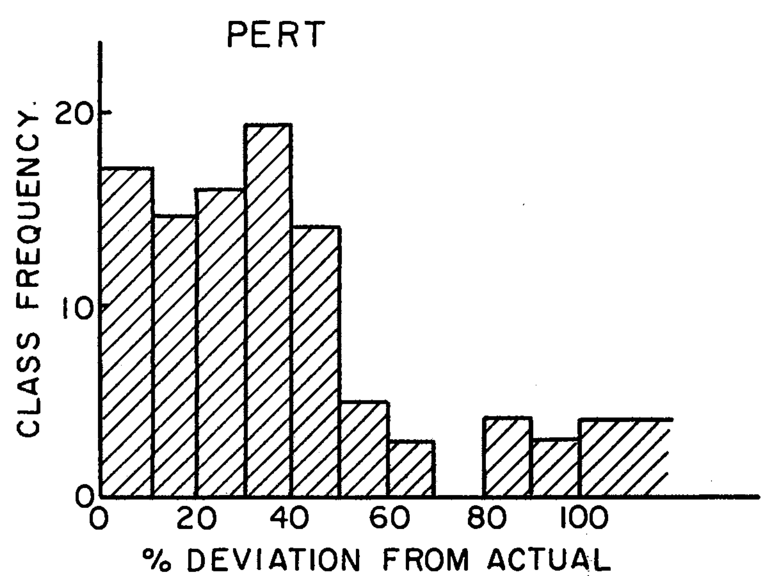


Figure 3. PERT Histogram

TABLE I

FREQUENCY DISTRIBUTION

Class Mark (x_i)	f_i (PERT)	f_i (CPM)
5	17	22
15	15	11
25	16	8
35	19	14
45	14	18
55	5	12
65	3	2
75	0	6
85	4	0
95	3	4
105	4	3

exists between the two techniques. A closer look at the histograms, however, indicates that the CPM projects exhibit a bimodal frequency distribution with one mode in the 0-10 per cent deviation range. This would substantiate the supposition that some pressure might be exerted to finish these projects on time. Some authors (3, 22) feel the advantage of a single-time estimate sets a firm date and extra effort is made to meet the schedule. However, the next mode occurs at 40-50 per cent which, in general, is greater than PERT errors. The PERT histogram does not indicate a well defined mode but shows a fairly constant frequency at the lower range of deviations.

The results presented here are based on historical data without regard to overtime, cost, or efficiency. The following study provides a better measure of accuracy between the two methods in that estimates are made for "future" projects as far as the estimators are concerned. The results are obtained using the same estimators for the same projects while using both types of techniques.

Model for Comparison

In 1936, R. A. Fisher introduced the concept of discriminant function analysis. This new line of research led to a new method of deriving test criteria suitable for multiple variates. The concept can be used here to reduce the several estimates by different people to a single variate which can then be used to determine if there is a difference between the two classes of estimates. The procedure described would try to find a critical value of an index such that any number below the index would belong to one class of estimates and values above the index would belong to the other. If overlap does occur, it is possible to test for significance to determine if there is a real difference. The discriminant function analysis finds its primary utility in classifying different species of plants or animals. If one were trying to determine which class or group a sample belonged to, he could make use of several measurements and combine them in the manner described below to obtain a single variate. This index number would serve to classify the particular sample as belonging to one or the other species.

The primary need for this study was to find a model which would test the stated hypothesis; that is, determine if there is a significant difference between the two methods of estimating. Because of its special characteristic of separating two groups of data, the discriminant function analysis is very amenable to this study. This model permits all the estimates of one type to be combined to provide a single variate for each project. Using the appropriate statistic the hypothesis is tested to determine if there is a significant (statistical) difference between the two methods.

As noted in the literature (3, 23, 25) the principal difference between a linear discrimination function and an ordinary linear regression function arises from the nature of the dependent variable. While a linear regression function uses values of the dependent variable to determine a linear function that will estimate the values of the dependent variable, the discriminant function possesses no such values or variable, but uses instead a two-way classification of the data to determine the linear function. Before determining the function which represents the single estimate for a given project the following notation is defined:

T_o = scaled value of actual activity time.

T_i = scaled value of estimated time by estimator.

$x_i = \frac{T_i - T_o}{T_o} \times 100$.

Z = transformed x variable.

i = 1, 2, ..., k, where k is the number of available estimator.

$Z = \sum_{i=1}^{k} x_i \lambda_i$ where λ_i are coefficients to be determined.

The first step in the solution will be to determine the λ's by some criteria that will allow Z to serve as an index for differentiating between members of the two groups of data. In other words, all the CPM estimates are combined in a linear combination to give a single variable for each project. Likewise, all estimates for the PERT members are combined to give a single value of Z. The development of the discriminant function analysis can be found in Hoel (12) and is summarized in Appendix A. Finally a test of significance is made using Mahalanobis' generalized distance between two samples.

Source of Data

Ideally one would want to test unbiased time estimates using both estimating methods for the same activity; that is, estimate the same activities using each of the two methods, neither of which would influence either of the two estimates. A direct experimental approach to the testing of the hypothesis does not appear feasible. It would be desirable to have one estimator make a CPM estimate and then make a PERT estimate for a series of projects. Since it would be impossible for the estimator to completely ignore one estimate while making the other, the procedure developed here avoids having the same person make both types of estimates on the same activity. Although one person will be performing both types of estimates, he will, in all cases, use only one method per activity.

A number of projects were examined from different types of industries ranging from small subcontractors to million dollar construction industries to billion dollar government projects. After talking with many of the different managers, it was concluded that the most uniform, consistent set of data could be obtained from the manufacturing industry where machined parts were the end product. Another desirable aspect of the chosen activities was the availability of historical work records. It was necessary to find projects that had enough detail in the work statement to present a uniform package to planners to make reasonable estimates and at the same time historical records had to be available so the actual time could be determined. Through the cooperation and assistance from a firm in the aerospace industry, enough background data was obtained to make a successful study. The activities used in this study were typical airplane and missile components and involved making estimates for the manufacturing time required to make a part. The time of interest was determined by the interval from the time engineering drawings were available, to the time the component was manufactured and ready for subsequent use.

Even though the data used was not classified confidential, the information was proprietory and all time estimates are coded. The actual time of the longest project was assigned the value of 100 and all other times are normalized on this actual time. Some estimates have values over 100 but this simply means the estimated time was higher than the longest actual time.

Data Reduction

After the data was obtained and coded it was tabulated for the CPM estimate and the expected time was calculated for the PERT estimate according to the accepted formula:

$$t_e = \frac{a + 4m + b}{6} \quad .$$

These data are shown in Table II. Since the estimates are to be combined to yield a single estimate for each activity, the ordering of the events is not pertinent and for clarity of presentation, the shortest in terms of actual activity time is labeled No. I up to the longest which is labeled No. VI. It would be convenient to recall the hypothesis to be tested is whether the three-time estimate varies significantly from the single-time estimate. From the data shown in Figure 4, no clear pattern can be determined and it is doubtful that inferences could be made from this figure. From the cursory examination of Figures 5 and 6 it can be seen that the widest range of estimates is found in Figure 5. This is to be expected since the PERT estimates gives estimates for least likely times in both directions; that is, pessimistic and optimistic times. The points on Figure 6 cannot be directly compared to Figure 5 because they are the values of the pessimistic, most likely, and optimistic times. As pointed out before, the expected time must be calculated before a comparison can be made. Archibald (3) warns of the balanced estimate approach when obtaining three-time estimates. This means making the estimate with an equal spread for the optimistic and pessimistic times. Experience shows that actual times, when reported against a balanced time estimate, are frequently closer to the pessimistic time than to the most-likely time;

TABLE II

CODED DATA

Job No.	T_o	CPM Estimates	PERT Estimates (T_e)
I	47.32	141.46	128.46
		94.63	83.34
		110.24	57.73
		73.17	66.66
II	77.07	141.46	143.10
		117.56	98.78
		128.78	74.80
		96.58	76.02
III	79.51	171.71	142.68
		134.63	110.56
		141.95	76.83
		104.88	84.95
IV	91.71	175.61	155.29
		148.78	129.26
		158.54	82.51
		82.93	106.10
V	92.68	166.58	162.61
		109.75	138.22
		130.73	92.68
		95.12	111.39
VI	100.00	165.12	162.61
		80.49	109.76
		152.19	104.07
		95.12	92.68

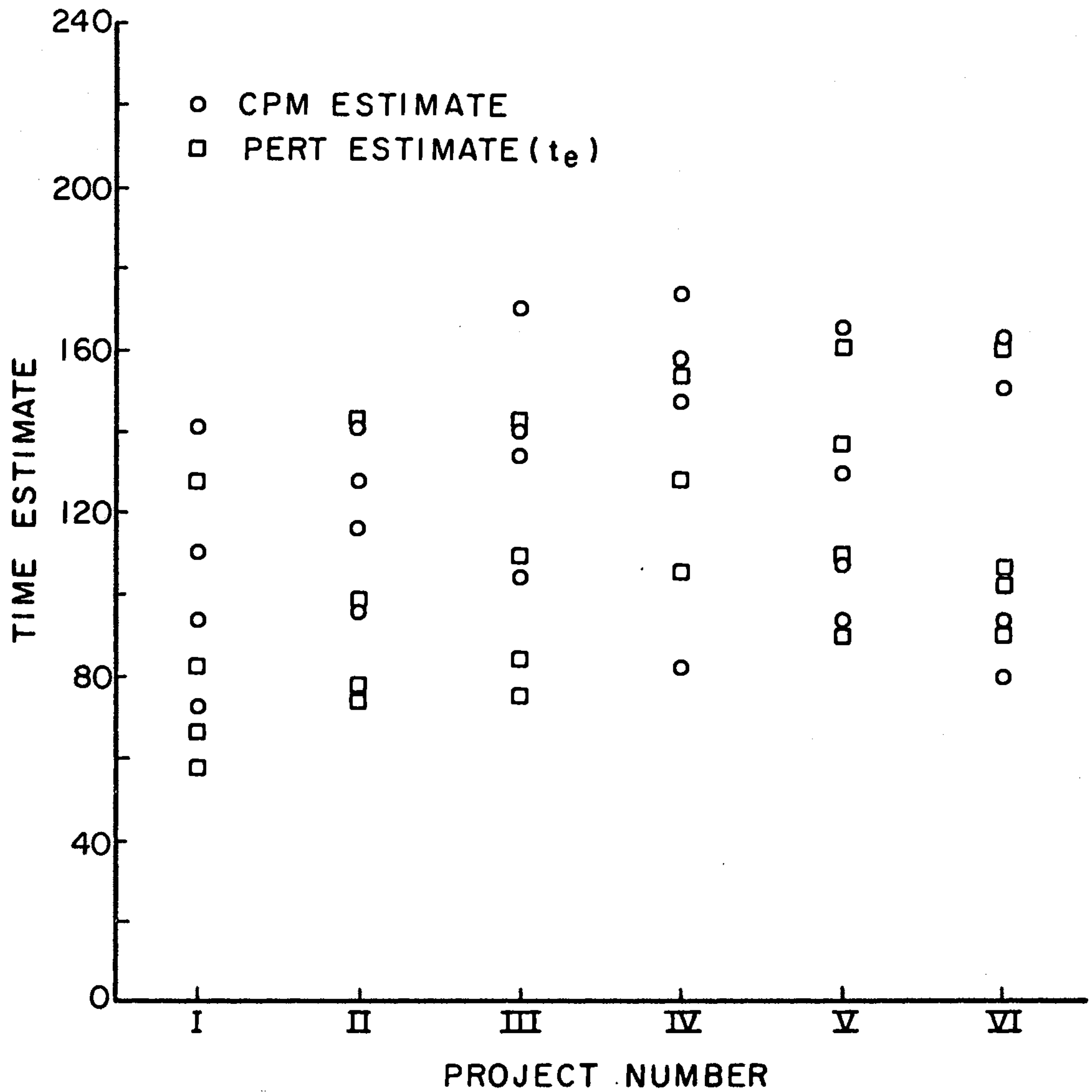


Figure 4. Comparison of PERT and CPM Time Estimates for Selected Projects

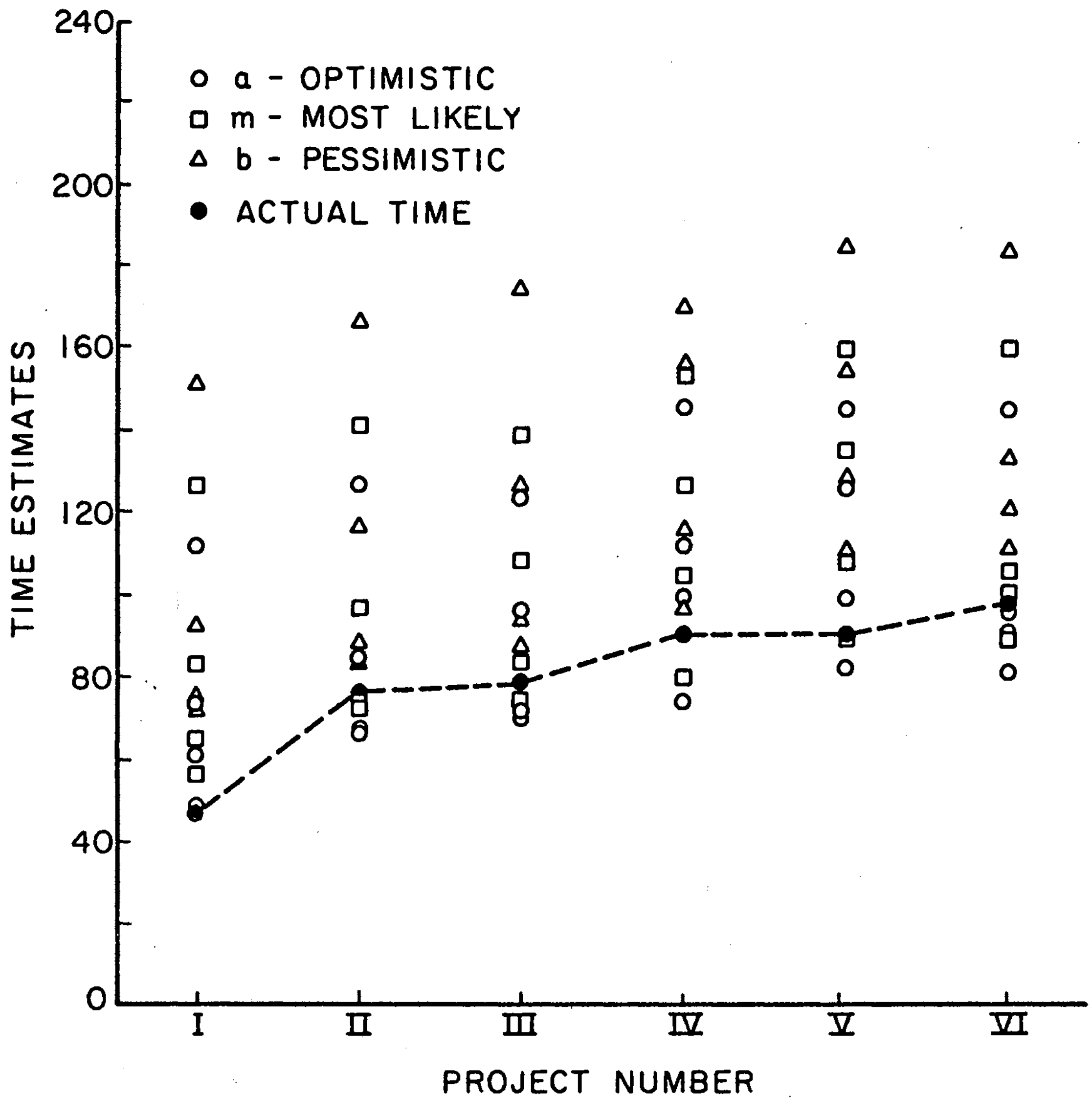


Figure 5. PERT Time Estimates for Six Different Control Projects

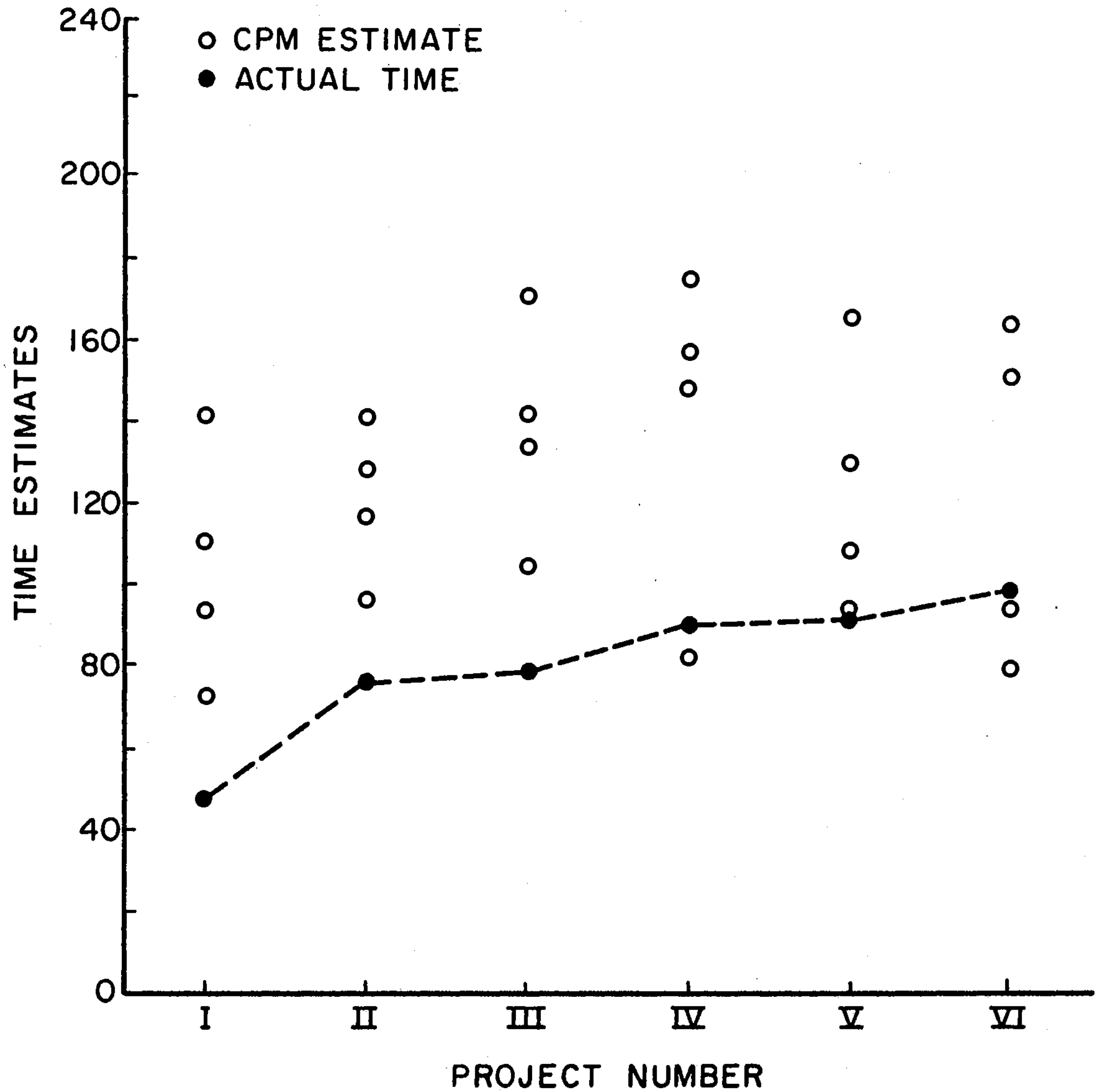


Figure 6. CPM Time Estimates for Six Different Control Projects

this indicates that the estimator did not allow for possible delays that should have been considered. Therefore, although estimates may not necessarily be skewed toward the pessimistic, the planner should review the estimating rationale if he receives too many balanced estimates, to be sure that the pessimistic time has accounted for everything. This problem did not appear in this study; however, this would certainly be a contender for a problem area and alert managers should guard against this type of impulsive estimate. Even when this does happen a better estimate is probably obtained since it still provides the estimator with a "psychological cover" in that he can give what he feels is an accurate estimate and then, by providing a and b estimates, feel confident that the actual time will be somewhere in his bounds. Figure 7 presents the expected value of the PERT estimates. The actual completion times are shown on this chart to demonstrate the relationship between actual and estimated times. The highest upper bound (which corresponds to the most pessimistic time) and the least lower bound (the most optimistic time) are also shown. These do not reflect one individual estimator, but may be from different estimators on any given project. Even though the expected times do not always accurately estimate the activities, the two bounds usually include the actual time. Also of interest is that the actual time normally lies on the optimistic side of the estimates. This would tend to substantiate the contention that those familiar with the technique tend to be conservative. This point will be explored further in the next chapter.

Comparing Figure 6 with Figure 7, a noticeable difference between the two types of estimates begins to emerge. However, only preliminary inferences can be made from examining the data and charts. The data

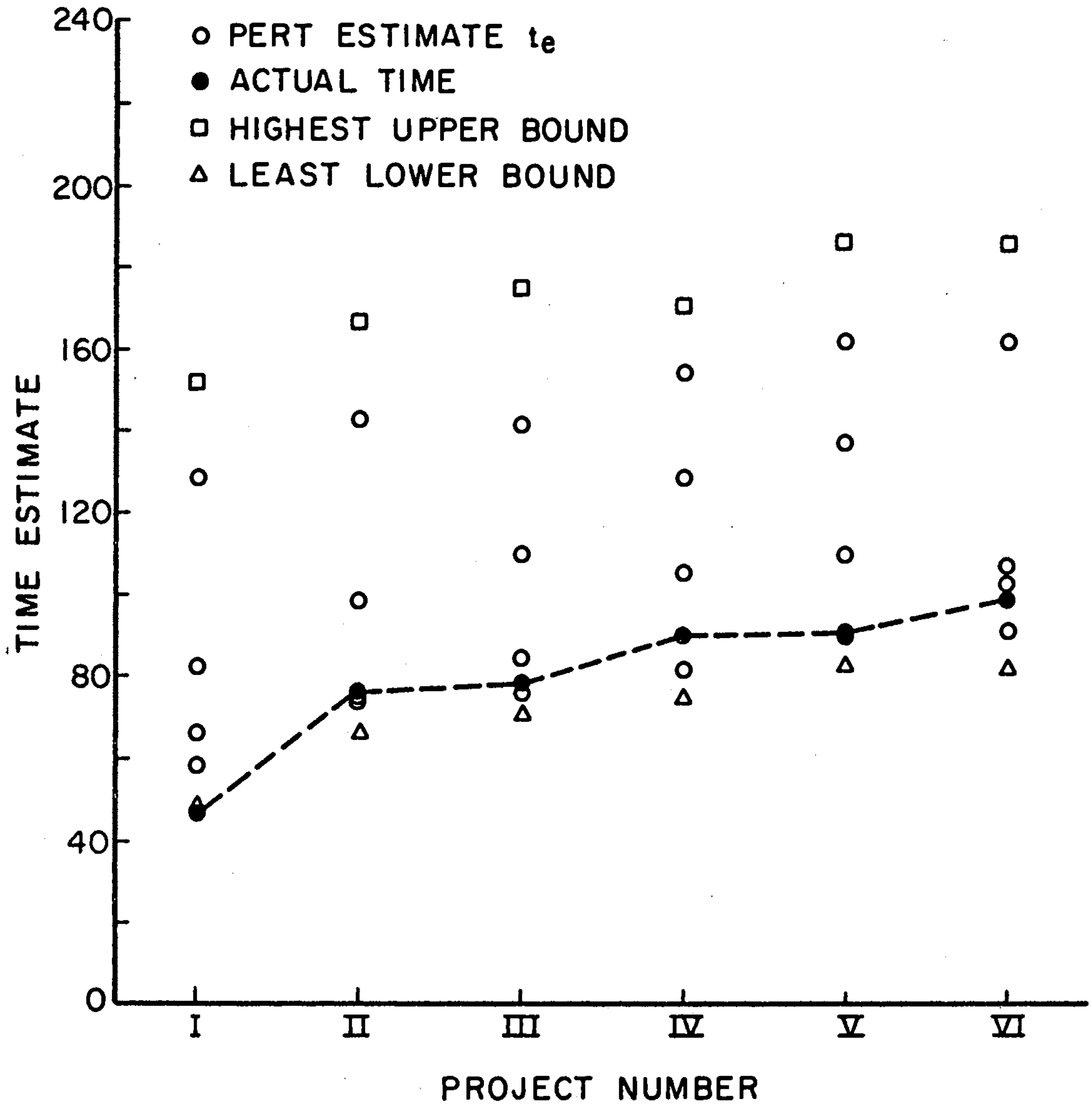


Figure 7. Calculated Expected Activity Time From PERT Estimates for Control Projects

must be reduced to fit the model described in the appendix to determine if a significant difference exists between the two types of estimates. Since it was necessary to calculate many terms for the R matrix and then compute the matrix inverse, a small computer program was written for this portion and the results are shown in tabulated form. Sample calculations are given at each step and the results are shown.

Discriminant Function Determination

The first step in determining an index is to arrange the data in standard form. Recalling the form of each independent variate:

$$x_i = \frac{T_i - T_o}{T_o} \times 100$$

Table III is determined. The following parameters are defined for subsequent use in the analysis.

$\bar{x}_{ij}$ = average of i^{th} estimator for the j^{th} type of estimate

$i = 1, 2, 3, 4$

$j = 1, 2$

x_{pij} = individual estimates where p refers to the estimator, i to the type of project, and j to the project.

$dp = \bar{x}_{p1} - \bar{x}_{p2}$.

Next, values for R_{ij} must be calculated and use is made of the computer for the calculations.

$$R_{pq} = \sum_{i=1}^{2} \sum_{j=1}^{6} (x_{pij} - \bar{x}_{pi})(x_{qij} - \bar{x}_{qi})$$

TABLE III

TABLE OF x_i VARIATES

		Project						$\bar{x}_{ij}$
		I	II	III	IV	V	VI	
CPM	x_{11}	198.95	83.55	115.96	91.48	79.74	65.12	105.80
	x_{21}	99.98	52.52	69.32	62.22	18.42	19.50	53.66
	x_{31}	132.97	67.09	78.53	72.87	41.06	52.19	74.11
	x_{41}	54.63	25.31	31.91	3.03	2.63	4.88	20.40
PERT	x_{12}	171.47	85.68	79.45	69.33	75.45	62.61	90.67
	x_{22}	76.12	28.17	39.05	40.94	49.14	9.76	40.53
	x_{32}	22.00	2.95	3.37	10.03	0.00	4.07	7.07
	x_{42}	42.87	1.36	6.84	15.69	20.19	7.32	15.37

TABLE IV

TABLE OF $\bar{x}_{ij}$ VALUES AND d_p VALUES

$\bar{x}_{11} = 105.80$	$\bar{x}_{12} = 90.67$	$d_1 = 15.13$
$\bar{x}_{21} = 53.66$	$\bar{x}_{22} = 40.53$	$d_2 = 13.13$
$\bar{x}_{31} = 74.11$	$\bar{x}_{32} = 7.07$	$d_3 = 67.04$
$\bar{x}_{41} = 20.40$	$\bar{x}_{42} = 15.37$	$d_4 = 5.03$

To calculate the values of R_{23}, the x values are obtained from Table III and the $\bar{x}$ values are obtained from Table IV and the calculations would be as follows.

$$R_{23} = \sum_{j=1}^{6} [(x_{21j} - \bar{x}_{21})(x_{31j} - \bar{x}_{31}) + (x_{22j} - \bar{x}_{22})(x_{32j} - \bar{x}_{32})]$$

$$\begin{aligned} R_{23} = {} & (99.98 - 53.66)(132.97 - 74.11) + (76.12 - 40.53)(22.00 - 7.07) \\ & + (52.54 - 53.66)(67.09 - 74.11) + (28.17 - 40.53)(2.95 - 7.07) \\ & + (69.32 - 53.66)(78.53 - 74.11) + (39.05 - 40.53)(3.37 - 7.07) \\ & + (62.22 - 53.66)(72.87 - 74.11) + (40.94 - 40.53)(10.03 - 7.07) \\ & + (18.42 - 53.66)(41.06 - 74.11) + (49.14 - 40.53)(0.00 - 7.07) \\ & + (19.50 - 53.66)(52.19 - 74.11) + (9.76 - 40.53)(4.07 - 7.07) \end{aligned}$$

$$R_{23} = 5326.74 .$$

These values are shown in the R matrix below.

$$R = \begin{bmatrix} 19968.10 & 10361.70 & 8851.69 & 6911.11 \\ 10361.70 & 7316.77 & 5326.74 & 4150.63 \\ 8851.69 & 5326.74 & 5428.96 & 3441.24 \\ 6911.11 & 4150.63 & 3441.24 & 3194.07 \end{bmatrix}$$

As mentioned in the appendix and noted here, the R matrix is symmetric; that is

$$R_{ij} = R_{ji} \quad .$$

To determine values for λ, the following set of equations must be solved.

$$\sum_{i=1}^{4} \lambda_i R_{ri} = d_r \qquad r = 1,\ 2,\ 3,\ 4$$

Individual values are obtained from the R matrix above and d_r values are from Table IV.

$$\lambda_1(19968.10) + \lambda_2(10361.70) + \lambda_3(8851.69) + \lambda_4(6911.11) = 15.13$$

$$\lambda_1(10361.70) + \lambda_2(7316.77) + \lambda_3(5326.74) + \lambda_4(4150.63) = 13.13$$

$$\lambda_1(8851.69) + \lambda_2(5326.74) + \lambda_3(5428.96) + \lambda_4(3441.24) = 67.04$$

$$\lambda_1(6911.11) + \lambda_2(4150.63) + \lambda_3(3441.24) + \lambda_4(3194.07) = 5.03$$

Computer solutions to above set of equations yield:

$$\lambda_1 = -\ .0098831$$

$$\lambda_2 = -\ .01151939$$

$$\lambda_3 = \ .0495881$$

$$\lambda_4 = -\ .0154969$$

Finally the Z transformation may be made for each activity.

$$Z = \sum_{i=1}^{4} \lambda_i x_i$$

For activity number III of the single-time estimate

$$Z_{CPM(III)} = - .0098831(115.96) - .01151939(69.32) + .0495881(78.53) - .0154969(31.91)$$

$$Z_{CPM(III)} = 2.48135 .$$

All the Z transforms are calculated in a similar manner and are shown in Table V.

TABLE V

INDEX VALUES

(Table of Z Transforms)

	I	II	III	IV	V	VI
CPM	4.3988	2.2468	2.48135	2.7954	1.7043	2.22339
PERT	- .58874	- .28397	- .46724	- .2859	- .95351	- .08592

Figure 8 depicts the Z transformation for the two different types of activities. It is noted that the transform separates the two estimates and a clear picture can now be observed. To show there is a significant difference in the two methods, a significance test is made in the next section.

Figure 8. z Transform of PERT and CPM Values

Test of Significance

The procedure used here is described in the appendix and the theoretical basis can be found in Rao (25).

Beginning with

$$d_p^2 = \sum_{i=1}^{p} \sum_{j=1}^{p} W_p^{ij} (\bar{x}_{i1} - \bar{x}_{i2})(\bar{x}_{j1} - \bar{x}_{j2})$$

the following notations are made.

p refers to the number of significant variates and is four for this study,

N_1 is the number of activities in the CPM population, and

N_2 refers to the number in the PERT population.

R_{ij} inverse is obtained from the computer print out and

$$W^{ij} = (R_{ij})^{-1}(N_1 + N_2 - 2)$$

d_p^2 is calculated from the relation

$$d_4^2 = \sum_{i=1}^{4} \sum_{j=1}^{4} (N_1 + N_2 - 2)(R_{ij})^{-1}(\bar{x}_{i1} - \bar{x}_{i2})(\bar{x}_{j1} - \bar{x}_{j2})$$

$$d_4^2 = \sum_{i=1}^{4} \sum_{j=1}^{4} (N_1 + N_2 - 2)(R_{ij})^{-1} d_i d_j$$

where $\bar{x}_{ij}$ are previously calculated constants tabulated in Table IV.

$$d_4^2 = 29.4868 \quad .$$

The statistic

$$\frac{N_1 N_2 (N_1 + N_2 - p - 1)}{p(N_1 + N_2)(N_1 + N_2 - 2)} d_p^{\,2}$$

should have the known F distribution with degree of freedom p and $(N_1 + N_2 - 1 - p)$

$$F_{CAL} = 15.48 \quad .$$

Using the table from Miller and Freund (20), the tabulated values for four and seven degrees of freedom for the F distribution are:

$F_{.05}(4, 7)$	$F_{.01}(4, 7)$	F_{CAL}
4.12	7.85	15.48

The calculated value of F is greater than the tabular value of $F_{.01}$ and significance is established at the .01 level. Since the null hypothesis says the two methods of estimating are equivalent, the null hypothesis must be rejected and it is concluded that the two methods are significantly different in the statistical sense. The relative accuracy of the two methods is discussed in Chapter IV.

CHAPTER IV

CONCLUSIONS

General Comments

It is an accepted fact that the management technique of CPM and PERT have proven their worth in many projects in various types of industries over the past twenty years. The complexity and magnitude of multi-million projects makes it mandatory to use some type of management control if the objectives are to be achieved and kept within acceptable financial limits. One manager pointed out that in the construction industry, final payment may be withheld until the completion of the job. Since this may involve millions of dollars, the interest alone for a single day may be appreciable. If the CPM analysis saves only one day it can make the entire study a profitable one.

In recent years, the feeling that PERT was too costly to operate has led to its diminished use. Some have questioned the theoretical assumptions behind the PERT formulation and others have objected to the fact that it provides a "rubber yardstick" when controlling the project (13). It can hardly be denied, however, that many useful benefits have resulted from its implementation. There are some who complain that three properly prepared estimates are not feasible; the tendency would be to make one estimate and improvise the other two. Waldron (30) states: "The use of PERT has diminished to minor importance since more valuable information concerning the establishing of activities has been realized with use."

It is a guide that is sometimes used as an indicator of uncertainty. One school of thought states that if the Probability for the Final or Objective Event is between 0.25 and 0.65, this is an acceptable status for a Cost-Plux-Fixed-Fee contract. If the Probability is more than 0.65 on a CPFF contract, then the agency has a right to suspect that too many resources are being applied to that contract (overloading). If the Probability falls below 0.25, then they may feel that the contract is understaffed, or insufficient resources are being applied.

The results of this study demonstrate that an advantage can be obtained by using three estimates, especially if many different people are depended on for inputs. Some consultants demonstrate an acceptable degree of accuracy for single time estimates where they depend only on themselves for the entire project, but a manager might have a more flexible control over individual activities if he had three estimates to work with.

Accuracy Improvement

Today's complex business makes it mandatory for managers to take full advantage of the available knowledge to properly assess the environment and act accordingly. Taking the action and making decisions based on considered judgment is so vital to the success of modern business that he should always be searching for ways to improve. Most managers are aware of the diversity of people that must be considered when planning a project and they are aware of various outlooks, work habits and various needs. This diversity demands a similar diversity in manager ingenuity. A successful manager must provide a means whereby he may adequately control a project while simultaneously satisfying the physical and psychological needs of the people relating to the project. This is a difficult and delicate task.

Those who have worked extensively with PERT or CPM know the difficulty of preparing adequate estimates for the analysis. The technique may be superior to any other, the underlying mathematical assumptions may be completely fulfilled, but project managers are cognizant of the large deviations in estimates. This study substantiates the large variations possible in this difficult aspect of estimating. In no way should this study imply that the estimators do not adequately prepare their figures. Network techniques are extremely beneficial when charting new and heretofore unfamiliar projects and this by its nature means uncertainty. In light of this, some of the estimates were remarkably accurate. The efforts of this study have been directed towards diminishing the effect of extraneous psychological factors that a person consciously or unconsciously injects in the analysis. The analysis demonstrates a wide margin for improvement is available, but unfortunately not all of it is attainable. There is a degree of uncertainty that will remain because of the changing environment.

Improvement From Psychological Factors

To assess the improvement available from psychological factors, two parameters were examined. First, the mean estimate for each category was determined and the deviation from actual was calculated. Since all variables are the same for both types of estimates the difference between the two was used as a measure of the improvement possible. The results are shown in Table VI and also depicted graphically in Figure 9.

TABLE VI

PER CENT DEVIATIONS OF AVERAGE ESTIMATES

Project	CPM	PERT	CPM-PERT
I	122.0	77.57	44.43
II	57.0	27.38	29.62
III	74.1	30.1	44.0
IV	54.1	29.0	25.1
V	35.1	36.18	- 1.08
VI	22.5	17.5	5.0

A second aspect of useful information is the dispersion or variation in the data. This factor shows consistency of estimates and, therefore, the coefficient of variation was calculated according to the formula

$$cv = \frac{s}{T_o} \times 100 \quad .$$

The results are shown in Figure 10 and tabulated in Table VII. This measure gives the standard deviation as a percentage of the mean and is independent of the scale of measurement.

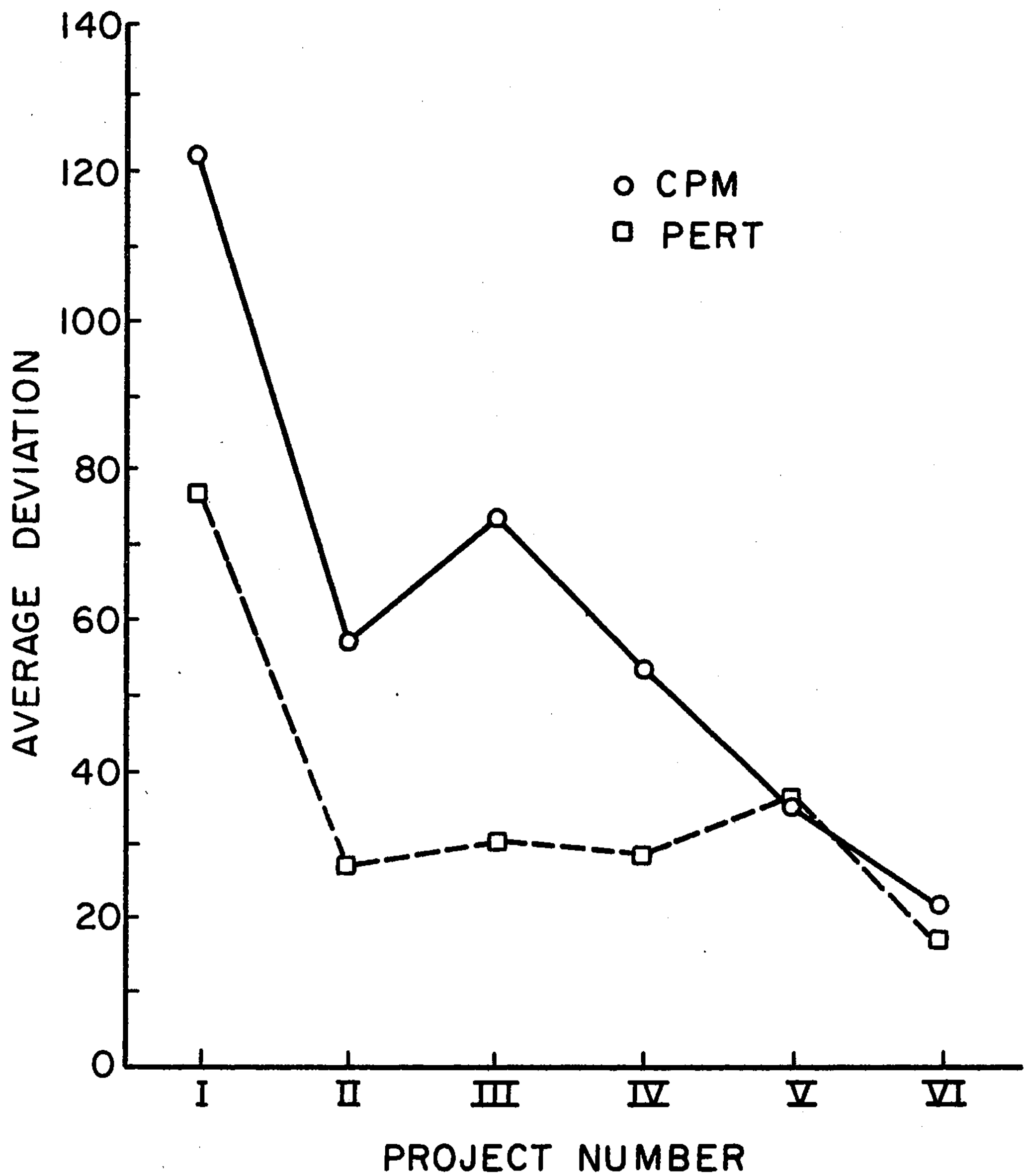


Figure 9. Deviation of the Mean Estimate for Each of the Control Projects

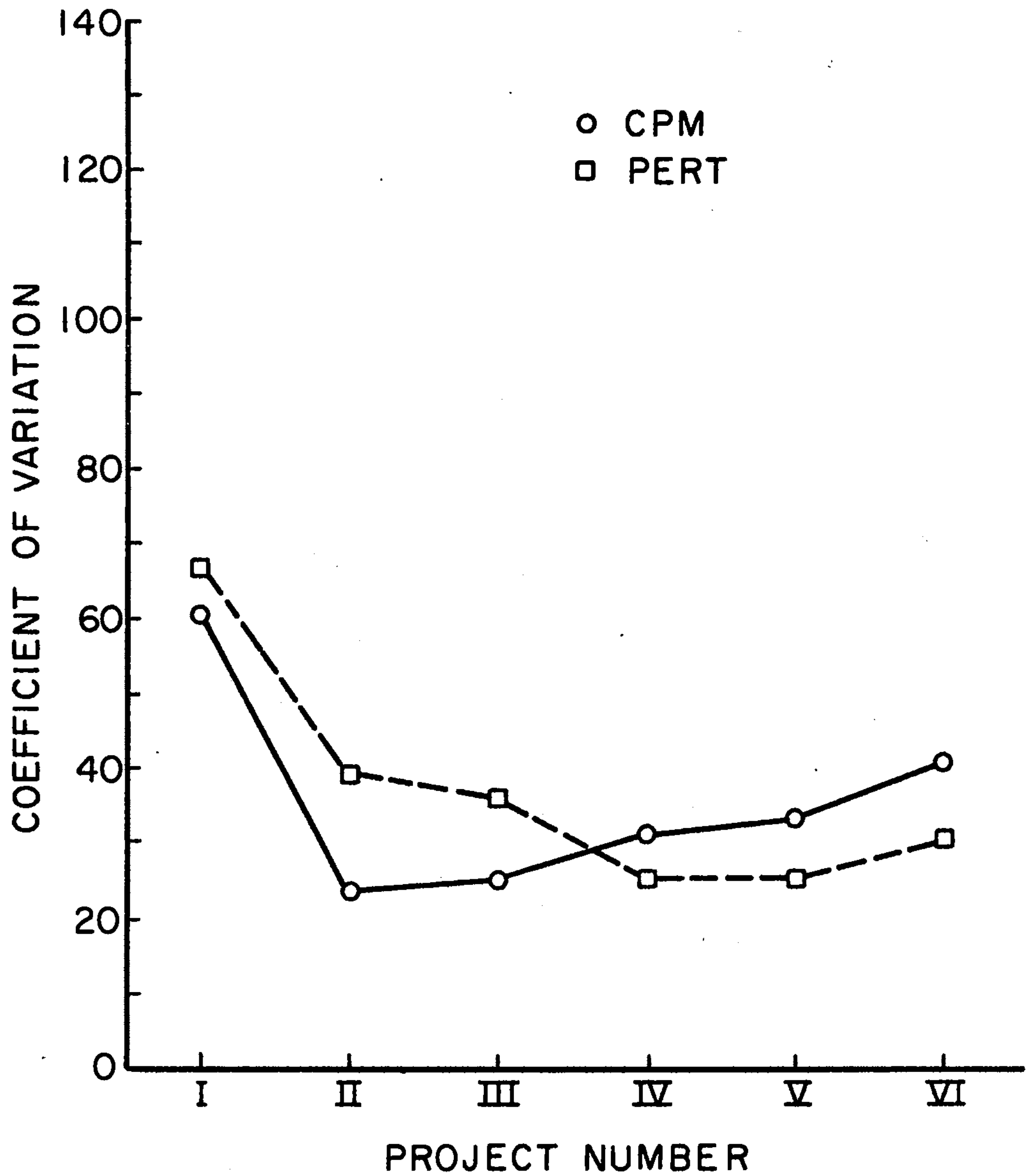


Figure 10. Coefficient of Variation

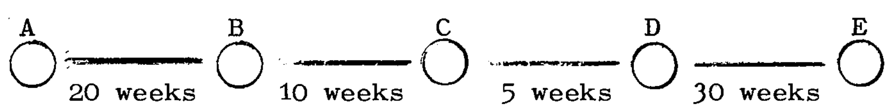
TABLE VII

COEFFICIENT OF VARIATION

Project	CPM	PERT
I	60.70	66.49
II	24.71	43.76
III	25.72	36.74
IV	31.09	25.80
V	33.47	26.42
VI	41.68	31.00

Application to Future Networks

For management control, it is desirable to have the mean estimate as close to nominal as possible and also desirable to have small variations about the nominal. For instance, a typical critical path might have the actual times which should be as shown on the diagram below.



The total time from A to E should take 65 weeks. If estimates of 10 - 20 - 15 - 20 were received from four different estimators the total estimated time of the project would also be 65 weeks, but from a manager's point of view the project would be difficult to control because the underestimated times would be impossible to meet without overtime or other complicating factors. For this situation, the average deviation for the entire critical path would be zero, but the coefficient of variation would be high.

For the other case, if estimates of 22 - 15 - 7 - 35 are received, the coefficient variation is small but the total estimated time becomes 79 weeks. Even though the coefficient of variation is smaller, the average deviation is higher and either situation complicates a manager's job.

The desirable situation would be to have a small mean deviation and also have a small coefficient of variation. The results of this study and the data presented here would indicate the three-time estimate produces more accurate results than the single-time estimate based on these factors.

From the data presented, it can be concluded that improvements up to 50 per cent can be attained if the psychological contributors to inaccuracies can be eliminated.

Just as there is no sure fire method of finding the "best" way to complete a given project, there is no guaranteed formula for producing an exact estimate. One must study the available alternatives, and in light of the current environment choose the appropriate technique for controlling the project. The technique suggested here provides a manager with a useful concept that enables him to take advantage of two techniques by extracting the favorable aspects of each and dovetailing them into a workable scheme.

CHAPTER V

FUTURE APPLICATION

The model proposed in Chapter III can be used to determine if an advantage can be obtained by making three estimates for a planners CPM analysis. Careful records would have to be maintained in the initial stage, but it is hard to imagine a situation where an inferior analysis would result from using the three-time estimate technique. The technique would have its primary use where a planner must depend on other departments, other companies, or simply other people. In large, complex projects, it is generally the case that inputs will be made from various departments and from more than one contractor or sub-contractor. Since the greatest deterrent from using PERT is the cost factor, many managers have simply overlooked the possibility of using three-time estimates. Many existing CPM computer programs can be modified to accept three-time estimates however, and the analysis can be completed in a deterministic fashion.

If it can be assumed, as suggested by some managers that estimates of sub-contractors with no knowledge of CPM tend to be optimistic, and pessimistic when obtained from planners who know of the technique, it would be very easy to compensate for this when using the three-time estimate technique. It would become much more flexible if the manager had enough control over the computer program so he could actually select which of the three times to use. An experienced manager could profit considerably if he diminished the psychological factors involved in the input data he receives. It appears to be true that one is strongly influenced if he knows how the data is to be used: A manager can circumvent this aspect if he obtains three estimates without revealing how they are to be used.

Many managers feel that a company must own its computer programs if it is to have adequate flexibility. This is to give more control over the way the data is treated. In addition, it gives more familiarity and more accessability to modify the program from time to time. If one wanted a degree of sophistication, he could keep records on various departments or sub-contractors and by use of charts on page 215 in Moder and Phillips (22) he should get excellent results. By using the technique described herein, the rising costs of PERT analysis per se could be avoided, but the many advantages of PERT analysis, both psychological and real could be realized, perhaps the familiar lament "It won't work" will not be expressed so frequently if one can depend on the accuracy of estimates he receives.

Extension of Study

The model used in this study and the subsequent experimental application demonstrates that it could be made to be a useful applied tool. The degree of improvement would vary between industries and ingenuity of managers, but it is felt that measurable improvement is available in all cases; the degree of improvement depending on many factors. The adage that "the work expands to fill the available time" is very relevant in many projects, especially where proposals are to be made, or surveys are to be conducted. A frequent question is "How much time do I have?" It is quite common that the time is completely used and the quality suffers if inadequate time is allowed. On the other hand, if excess time is available, the quality may not increase but the time is usually consumed.

There is a wide latitude for possible improvement in the area of estimates, but this thesis has been confined to the psychological factors which may result in a degradated analysis. It would be of great interest if these factors could be individually isolated and properly assessed. During the early stages of any program, human nature restrains an honest appraisal of the physical environment and it would be advantageous to measure this effect.

Another interesting possibility would be to use three estimates in a different combination. For instance, a small construction project could be simulated and estimates made by people familiar with CPM and also by those with no knowledge of it. If optimistic times of the former are used in conjunction with the pessimistic times of the latter and this produces an accurate chart, this would substantiate the experiences of the construction industry.

A SELECTED BIBLIOGRAPHY

(1) Anderson, Richard C. Management Strategies. New York: McGraw-Hill, 1965.

(2) Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: John Wiley & Sons, Inc., 1958.

(3) Archibald, Russell D., and Richard L. Villoria. Network-Based Management Systems (PERT/CPM). New York: John Wiley & Sons, 1967.

(4) Berners-Lee, C. M. Models for Decision. London: The English University Press Limited, 1965.

(5) Burgess, A. R., and J. B. Killebrew. "Variation in Activity Level on a Cyclical Arrow Diagram." Journal of Industrial Engineering, Vol. 13, No. 2 (March-April, 1962).

(6) Cochran, William G. Sampling Techniques. New York: John Wiley & Sons, 1953.

(7) Croft, F. Max. A Model for Predicting Cost Trends for R&D Launch Vehicle Contracts." (Unpublished Ph.D. dissertation, Oklahoma State University, 1970.)

(8) DOD/NASA PERT/COST Guide. Washington, D. C.: Government Printing Office, 1962.

(9) Duncan, Acheson J. Quality Control and Industrial Statistics. Homewood, Illinois: Richard D. Irwin, Inc., 1952.

(10) Haire, Mason. Psychology in Management. New York: McGraw-Hill, 1956.

(11) Hansen, B. J. Practical PERT. Washington, D. C.: America House, 1964.

(12) Hoel, Paul G. Introduction to Mathematical Statistics. New York: John Wiley & Sons, 1962.

(13) Horowitz, Joseph. Critical Path Scheduling. New York: The Ronald Press Company, 1967.

(14) Johnson, D. M. Psychology of Thought and Judgement. New York: Harper and Brothers, 1955.

(15) Kahn, Robert L. et al. Organizational Stress. New York: John Wiley and Sons, 1964.

(16) Kelley, J. E., Jr. "Critical Path Planning and Scheduling: Mathematical Basis." Operations Research, Vol. V, No. 6 (1961), pp. 296-320.

(17) "Line of Balance Technology." A Graphic Method of Industrial Programming. Department of the Navy. Navexos P1851 (Rev. April, 1962).

(18) MacCrimmon, Kenneth R., and C. A. Ryavec. "An Analytical Study of the PERT Assumptions." Operations Research, Vol. XII, No. 1 (January-February 1964), pp. 16-37.

(19) McLucas, Hon. John L. "The New Look in R & D Management." Armed Forces Management (December, 1969), pp. 44-46.

(20) Miller, Irwin, and John E. Freund. Probability and Statistics for Engineers. New Jersey: Prentice-Hall, 1965.

(21) Miller, Robert W. Schedule, Cost and Profit Control With PERT. New York: McGraw-Hill, 1963.

(22) Moder, Joseph J., and Cecil R. Phillips. Project Management with CPM and PERT. New York: Reinhold Publishing Corporation, 1964.

(23) Morrison, Donald F. Multivariate Statistical Methods. New York: McGraw-Hill Book Company, 1967.

(24) Phillips, Cecil R., and Charles R. Beek. "Computer Programs for PERT and CPM." Operations Research Incorporated. 2nd ed. (28 February 1963).

(25) Rao, C. Radhakrishna. "Tests of Significance in Multivariate Analysis." Sankhya, 7, 1947.

(26) Rao, C. Radhakrishna. "Tests with Discriminant Functions in Multivariate Analysis." Sankya, 7, 1946b.

(27) Seelig, W. D., and Irwin M. Rubin. "The Effects of PERT on R&D Organizations." Alfred P. Sloan School of Management. Massachusetts Institute of Technology (December, 1966).

(28) Siegel, Sidney. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.

(29) Stilian, Gabriel N. et al. PERT: A New Management Planning and Control Technique. New York: American Management Association, 1962.

(30) Waldron, A. James. Applied Principles of Project Planning and Control. New Jersey: A. James Waldron, 1966.

(31) Wasserman, Paul, and Fred Silander. "Decision-Making, An Annotated Bibliography." Graduate School of Business and Public Administration, Cornell University, 1966.

(32) Yamane, Taro. Statistics, An Introductory Analysis. New York: Harper and Row, 1967.

(33) Young, Robert K., and Donald J. Veldman. Introductory Statistics for the Behavioral Sciences. New York: Holt, Rinehart and Winston, 1965.

APPENDIX

LINEAR DISCRIMINANT FUNCTION

The model for the discriminant function analysis is taken from Hoel (12) and the development of the test of significance is taken from Rao (25). Let x_{pij} represent the value of x_p for the j^{th} project in the i^{th} group; $i = 1$ refers to CPM estimates, $i = 2$ refers to PERT estimates in this study. p refers to the number of estimators ($p = 1, 2, 3, 4$) and j refers to the project $j = (1, 2, 3, 4, 5, 6)$. $\bar{x}_{pi.}$ represents the mean value of $\bar{x}_p$ for all of the projects in one class. Recalling that z is to serve as the index for clasifying the two groups, it is evident that it is desirable to separate the two groups as far as possible. Unfortunately, as the two groups are separated farther apart, the variation of the values within a group becomes increasingly large for both groups. The increase in the separation of the index values is obtained, then, at the expense of an increase in the separation of the values of z within each group.

The objective is to choose a plane that separates the index values of the two groups as far as possible relative to the variation within the groups. Using

$$s = (\bar{z}_1 - \bar{z}_2)^2 \qquad (A.1)$$

as a measure of separation between the groups and

$$V = \sum_{i=1}^{2} \sum_{j=1}^{6} (z_{ij} - \bar{z}_i)^2 \qquad (A.2)$$

as a measure of the variation of the z values within the groups, it becomes necessary to determine the plane (defined by λ's, which maximizes $\frac{s}{v}$. From

$$z = \sum_{i=1}^{k} \lambda_i x_i$$

it follows that

$$\bar{z}_1 - \bar{z}_2 = \sum_{i=1}^{k} \lambda_i (\bar{x}_{i1} - \bar{x}_{i2})$$

and

$$z_{ij} - \bar{z}_i = \sum_{\alpha=1}^{k} \lambda_\alpha (x_{\alpha ij} - \bar{x}_{\alpha i}) \ .$$

To obtain an expression for s, call $d_p = \bar{x}_{p1} - \bar{x}_{p2}$

$$\bar{z}_1 - \bar{z}_2 = \sum_{i=1}^{k} \lambda_i d_i$$

$$s = (\bar{z}_1 - \bar{z}_2)^2 \ . \qquad (A.3)$$

A similar expression is obtained for v. Let

$$R_{pq} = \sum_{i=1}^{2} \sum_{j=1}^{6} (x_{pij} - \bar{x}_{pi})(x_{qij} - \bar{x}_{qi}) \ .$$

From Equation (A.2),

$$v = \sum_{i=1}^{2} \sum_{j=1}^{6} (z_{ij} - \bar{z}_i)^2 = \sum_{i=1}^{2} \sum_{j=1}^{6} \left[\lambda_1(x_{1ij} - \bar{x}_{1i}) + \ldots + \lambda_4(x_{4ij} - \bar{x}_{4i})\right]^2$$

$$v = \sum_{i=1}^{2} \sum_{j=1}^{6} \sum_{p=1}^{4} \sum_{q=1}^{4} \lambda_p \lambda_q (x_{pij} - \bar{x}_{pi})(x_{qij} - \bar{x}_{qi})$$

$$v = \sum_{p=1}^{4} \sum_{q=1}^{4} \lambda_p \lambda_q \sum_{i=1}^{2} \sum_{j=1}^{6} (x_{pij} - \bar{x}_{pi})(x_{qij} - \bar{x}_{qi})$$

$$v = \sum_{p=1}^{4} \sum_{q=1}^{4} \lambda_p \lambda_q R_{pq} \; . \tag{A.4}$$

These expressions can be combined to yield:

$$\frac{s}{v} = \frac{\sum_{p=1}^{4} \sum_{q=1}^{4} \lambda_p \lambda_q d_p d_q}{\sum_{p=1}^{4} \sum_{q=1}^{4} \lambda_p \lambda_q R_{pq}} \; . \tag{A.5}$$

It is desired to obtain values for the λ's which make $\frac{s}{v}$ a maximum. This requires

$$\frac{\partial \left(\frac{s}{v}\right)}{\partial \lambda_r} = 0$$

for r = 1, 2, 3, 4. This can be expressed in the form

$$\frac{v \dfrac{\partial s}{\partial \lambda_r} - s \dfrac{\partial v}{\partial \lambda_r}}{v^2} = 0$$

$$\frac{\partial v}{\partial \lambda_r} = \frac{v}{s} \frac{\partial s}{\partial \lambda_r} \qquad r = 1, 2, 3, 4 \tag{A.6}$$

to obtain expressions for $\frac{\partial v}{\partial \lambda_r}$ and $\frac{\partial s}{\partial \lambda_r}$, v and s are expressed in matrix form, and the symmetry of the matrix is observed. λ_r occurs in every term in the r row and r column. Also, $R_{ij} = R_{ji}$

$$\frac{\partial v}{\partial \lambda_r} = 2 \sum_{i=1}^{6} \lambda_i R_{ri}$$

$$\frac{\partial s}{\partial \lambda_r} = 2 d_r \sum_{i=1}^{6} \lambda_i d_i \qquad r = 1, 2, 3, 4 .$$

When these expressions are inserted in Equation (A.6), it will reduce to

$$\Sigma \lambda_i R_{ri} = c d_r \qquad r = 1, 2, 3, 4 \tag{A.7}$$

where

$c = \frac{v}{s} \Sigma \lambda_i d_i$ and is independent of r. The values of R and d are determined by the problem and become constants for any given problem. The equations must be solved to determine the λ's (which are not unique). Since c is a function of the λ's, it is usually chosen to be unity, and Equation (A.7) is solved for λ. Another common practice is to choose c such that λ_1 is unity and the rest of the λ's are calculated accordingly. Since either procedure (or any other) for assigning a value to c would be equivalent as far as discriminating between the two groups is concerned, c will be chosen to be unity for this analysis.

The procedure then is to obtain estimates from various planners for a series of projects; calculate the required parameters to obtain a set of linear equations in λ, by some means calculate all the λ's; using the values of the previously determined λ's, calculate a set of z's to represent a single value for each project; and finally determine if there is a significant difference between the two methods.

Test of Significance

To test the hypothesis that there is no difference in mean values of the two groups, the test statistic

$$\frac{N_1 N_2 (N_1 + N_2 - p - 1)}{p(N_1 + N_2)(N_1 + N_2 - 2)} d_p^2$$

is calculated. This statistic should follow the well-known F distribution with (p) and $(N_1 + N_2 - 1 - p)$ degrees of freedom.

The test statistic for significance is taken from Rao (25) and is the results of work by Mahalanobis in 1936. Let N_1 and N_2 be the sample sizes from two populations, π_1 and π_2, characterized by (p + q) variates. The sample means for the i^{th} character are represented by $\bar{x}_{i1}$ and $\bar{x}_{i2}$ for π_1 and π_2, respectively. The estimated value of the covariance is given by

$$(N_1 + N_2 - 2)w_{ij} = \sum_{t=1}^{N_1} (x_{i1t} - \bar{x}_{i1})(x_{j1t} - \bar{x}_{j1}) + \sum_{t=1}^{N_2} (x_{i2t} - \bar{x}_{i2})(x_{j2t} - \bar{x}_{j2}) .$$

Mahalanobis' distance between the two populations as estimated from the sample on the basis of the first p characters is

$$d_p^2 = \sum_{i=1}^{p} \sum_{j=1}^{p} w_p^{ij} (\bar{x}_{i1} - \bar{x}_{i2})(\bar{x}_{j1} - \bar{x}_{j2})$$

where (w_p^{ij}) is reciprocal to (w_{ij}), i, j = 1, 2, ..., p. To test the hypothesis specifying no difference in mean values of the p characters for π_1 and π_2, the statistic

$$F = \frac{N_1 N_2(N_1 + N_2 - p - 1)}{p(N_1 + N_2)(N_1 + N_2 - 2)} d_p^2$$

can be used as the variance ratio with (p) and $(N_1 + N_2 - 1 - p)$ degrees of freedom. To use the statistic for this analysis, the covariance matrix must be modified by the constant $(N_1 + N_2 - 2)$ before taking the inverse. Using the notation given earlier,

$$R_{ij} = (N_1 + N_2 - 2)w_{ij} .$$

Since R_{ij} must be calculated to obtain the λ's, it can also be used to calculate the generalized distance:

$$w_{ij} = \frac{1}{N_1 + N_2 - 2} R_{ij}$$

$$w^{ij} = (N_1 + N_2 - 2) R_{ij}^{-1} .$$

Then,

$$d_p^2 = \sum_{i=1}^{p} \sum_{j=1}^{p} (N_1 + N_2 - 2)R_{ij}^{-1} (\bar{x}_{i1} - \bar{x}_{i2})(\bar{x}_{j1} - \bar{x}_{j2}) .$$

When this calculated F-statistic is compared with the known F distribution, it must be smaller than the tabular value if the two groups come from the same population. If the calculated value is larger than the tabular value, then a significant difference does exist.

VITA

Fredrick W. Swift

Candidate for the Degree of

Doctor of Philosophy

Thesis: AN INVESTIGATION OF THE RELATIVE ACCURACY OF THREE-TIME VERSUS SINGLE-TIME ESTIMATING OF ACTIVITY DURATION IN PROJECT MANAGEMENT

Major Field: Engineering

Biographical:

Personal Data: Born in Springfield, Missouri, July 4, 1936, the son of William H. and Loretto E. Swift. Married Patricia L. Dark on June 8, 1957.

Education: Graduated valedictorian from St. Agnes High School, Springfield, Missouri, in May, 1954; received Bachelor of Science degree in Aeronautical Engineering from the University of Notre Dame, Notre Dame, Indiana, in February, 1959; received the Master of Science degree in Engineering Mechanics from the University of Alabama, Tuscaloosa, Alabama, in May, 1966; completed program in Management Institute and Labor Relations for supervisors sponsored by the U. S. Civil Service Commission, Atlanta, Georgia, in December, 1966; completed requirements for the Doctor of Philosophy degree, with a major in Industrial Engineering and Management, at Oklahoma State University, May, 1971.

Professional Experience: Employed by the Douglas Aircraft Company, Tulsa, Oklahoma, from June, 1958 to November, 1959 as Associate Engineer; employed by Douglas Aircraft Company, Charlotte, North Carolina, from November, 1959 to June, 1961 as a Project Engineer; employed by Marshall Space Flight Center (NASA), Huntsville, Alabama, from June, 1962 to present as an Aerospace Engineer.