ARPA ORDER NO.: 189-1

RM-6269-ARPA
AUGUST 1970

A METHODOLOGY FOR COST FACTOR COMPARISON AND PREDICTION

Alvin J. Harman
Assisted by Susan Henrichsen

prepared for
ADVANCED RESEARCH PROJECTS AGENCY

Rand
SANTA MONICA, CA. 90406

ARPA ORDER NO.: 189-1

RM-6269-ARPA
AUGUST 1970

A METHODOLOGY FOR COST FACTOR COMPARISON AND PREDICTION

Alvin J. Harman
Assisted by Susan Henrichsen

This research is supported by the Advanced Research Projects Agency under Contract No. DAHC15 67 C 0141. Views or conclusions contained in this study should not be interpreted as representing the official opinion or policy of Rand or of ARPA.

Rand
SANTA MONICA, CA. 90406

Published by The RAND Corporation

PREFACE

Two questions have dominated analysis of the weapon acquisition process in recent years and are especially relevant to an environment of constrained military budgets: What is an appropriate procurement strategy for a particular system? How can we increase the reliability of cost-effective comparisons between systems with uncertain costs? This Memorandum provides insights into each question. First, it analyzes the closeness of actual to predicted costs for weapon systems of the 1960s and compares the results with a similar analysis of earlier postwar experience. Second, it suggests a way of anticipating the cost outcomes of acquisition programs that involve various levels of technological risk.

The research reported here is an outgrowth of work first reported in preliminary form in RM-6072-PR, System Acquisition Experience.[1] The model described in Section IV of the Memorandum has been refined, with primary emphasis on describing the underlying motivation for the methodology of cost factor comparison.

This Memorandum, and particularly Sections II and III, is somewhat more technical than RM-6072-PR. Specialists who are concerned with statistical analysis of the acquisition process are the main components of the intended audience for those parts. Sections I and IV (Discussion) and the hypothetical example of a predictive application of the model (in Section V) should be of particular interest to decisionmakers in the Deaprtment of Defense and to the acquisition-policy echelons of the Army, Navy, and Air Force. For them, this study may offer insight into the probable cost# growth and range of uncertainty of future procurements -- *if* the recent past is a suitable guide to such insight.

Empirical results presented here are based on a modified version of the data used in RM-6072-PR. Two features of this data base should

---

[1] R. L. Perry, D. DiSalvo, G. R. Hall, A. J. Harman, G. S. Levenson, G. K. Smith, J. P. Stucker, System Acquisition Experience, The Rand Corporation, RM-6072-PR, November 1969.

be kept in mind in evaluating the statistical results. First, data were available for only a small number of systems.[1] The data for the systems of the 1960s were originally obtained in the spring of 1969 by the Director of Defense Research and Engineering, aided by Rand, through questionnaires on 21 systems developed by the three services.[2] That survey information has been expanded, refined, and up-dated to December 1969 for use in the final phase of estimation of this analysis.[3] Sufficient information for the model has now been obtained for 15 programs. For some of the more recent of these programs, projections of "actual" cost or time of operational status are still preliminary.

Second, the measure of technological advance used in the analysis was obtained by a rather limited survey conducted within Rand. These measures are, at best, crude approximations of the concept of the relative level of technological advance sought in the development of various weapons over two decades. Further work is in progress to refine and measure this concept.

---

[1] See also Perry _et al_., _System Acquisition Experience_, especially Sections I and II.

[2] DDR&E chose the sample of systems with the advice of the individual services.

[3] A detailed discussion of these data is presented in the Appendix.

SUMMARY

This Memorandum describes a methodology for comparing and predicting cost experiences of weapon system procurements, using cost factors (the ratios of actual to predicted costs) as a basis. Although cost factors can be greater or less than unity, past experience has shown that in the usual case actual costs exceed predicted costs. Cost factors having a value greater than unity can be caused by inaccurate cost estimates (implying that estimates are, on the average, lower than efficiently executed actual outcomes would be) or by cost growth arising from such factors as scope change (alteration of goals or specifications subsequent to the start of a development program), technological uncertainty (the general unpredictability of cost for risky new technology), or various inefficiencies. Cost factors alone do not provide sufficient information to distinguish among such contributors.

Also, cost factors are not adequate by themselves for making cost comparisons, since different types of systems may be more or less difficult to estimate accurately. For example, one weapon system may have a cost factor of 1.2 and another a cost factor of 1.4. But if the first was a brief program involving essentially off-the-shelf technology and the second required a major advance in technology and its cost had to be predicted very early in the conceptualization stage of the program, the _apparent_ 40 percent increase in costs of the second might well in retrospect be considered to be money well spent and the 20 percent increase of the first could be considered as evidence of unacceptably poor estimation or inefficient management.

An attempt to make adequate cost comparisons must take into account the influence on costs of other aspects of procurement, including program duration and the degree of technological difficulty encountered in development. This proposition leads to recognition of the need for a theoretical model within which the various influences on costs can be structured.

The model developed here considers several of the more pertinent influences on the size of a cost factor. One influence is program

length. A longer program may result in a high cost factor either because of increased time during which the management of the program is inefficient, or because the cost estimate was based on a very early (and therefore vague) configuration of the weapon. Another influence on the cost factor is the requirement for advanced performance of the system leading to a large desired technological advance and therefore a greater intensity of effort (for a given program length) in development of the system. This may result in a high cost factor because of the increased probability of development difficulties or because of estimation optimism that such difficulties will not occur.

The model is analyzed using two sets of data -- the first essentially composed of systems developed in the 1950s, and the second from the 1960s.[1] The various structures for the model all indicate that both longer programs and larger technological advances usually lead to higher cost factors. The comparison of acquisition experience between the two decades is based on the "Aircraft and Missiles" subsample of the 1960s data, which is most closely comparable to the 1950s data. If the parameter estimates for the model can be characterized as descriptive of the system acquisition process, then the statistical results show essentially no (net) change in the process over the last two decades. That is, for programs comparable in length and difficulty, 1960s procurements would have resulted in actual costs exceeding estimates by roughly the same proportion as had 1950s procurements. However, there is evidence that the "typical" program of the last decade was structured somewhat differently from the previous experience. On the average, development program length seems to have been somewhat shorter and the

---

[1]In the text, quotation marks are placed around these decades when they are used to name the samples, which cover somewhat broader ranges of years (see Section III). These samples were compiled at different times and use somewhat different bases for measurement of costs of the systems. In particular, the method of price deflation and assumptions of quantity adjustment of the actual and estimated costs are different; the refinement of the 1960s costs were chosen to be conservative — to err, if at all, in the direction of understating rather than overstating any improvement in cost accuracy between the decades (see the Appendix).

magnitude of technological advance sought has been kept somewhat lower than in the 1950s. When the (different) "typical" programs of each decade are predicted by the acquisition process (equation) characteristic of the decade, the cost factor for the 1960s is somewhat lower. There is thus some indication of less bias in cost estimation or overrun of actual costs in the 1960s than previously.[1]

Projections can also be made concerning the probable difference between actual and predicted costs of future programs as well as a range of uncertainty of estimation inaccuracy or actual cost growth -- provided that such future programs are sufficiently similar to those in the present sample.[2] However, the data that produced these results are crude. Thus, conclusions concerning the range of future uncertainty should be considered as tentative and subject to refinement by improved measurement of the appropriate variables and modeling of their interrelationships.

The results only suggest the range of possible cost uncertainty based on past experience; the policymaker must still cope with the fact that weapon systems involving requirements for high levels of technological advance will continue to be components of our force structure. Acquisition of such systems involves unavoidable cost uncertainty. If this uncertainty is to be reduced or hedged, there must first be a substantial improvement in the effectiveness of procurement strategies and processes -- which may require fundamental changes in organization and decisionmaking as well as in incentives and procedures.

---

[1] See Section IV for a full elaboration of these results; the "Discussion" subsection reviews the equation results and analyzes them graphically.

[2] And, of course, provided that the cost estimates have been obtained by a method that has not already "corrected" them for the "usual" overall growth of costs. The method of projection is discussed by example in Section V.

ACKNOWLEDGMENTS

The author would like to thank Rand colleagues R. L. Perry and G. K. Smith for their suggestions during this undertaking, as well as the other co-authors of RM-6072-PR -- D. Di Salvo, G. R. Hall, G. S. Levenson, and J. P. Stucker -- for their comments during the early stages of model design. J. A. Dei Rossi, D. E. Emerson, R. Erler, G. S. Fishman, R. B. Johnston, A. W. Marshall, P. R. McClenon, R. R. Nelson, J. P. Newhouse, R. L. Petkun, W. D. Putnam, and T. P. Schultz were also generous with their help on and suggestions concerning this study. Susan Henrichsen provided truly excellent research and editorial assistance throughout the development of this Memorandum. Any remaining errors are the responsibility of the author.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

A complete analysis of weapon system procurements should weigh the costs of various systems against their effectiveness. Effectiveness, of course, ought to be measured in terms of (1) the extent to which the weapon produced meets the threat actually present at the time when the weapon is in the inventory, and (2) the length of time during which it remains an important element of the inventory. Traditionally, however, weapon procurement studies have had a more limited goal.

When a weapon system acquisition is reviewed and evaluated, whether within the Department of Defense or by Congress, the closeness of the actual cost to earlier estimates is a common concern. A cost factor -- the ratio of the actual cost to the estimated cost -- is frequently used as a measure of the degree of closeness. We must have legitimate methods of using popular measures such as cost factors in these evaluations, but we must also appreciate the more limited role of cost factors within a complete cost-effectiveness analysis of the procurement process. As an illustration, consider the conditions necessary for cost factors to be of central importance in a cost-effectiveness analysis.

Suppose that a new weapon system is contemplated and that we know with certainty that it will perform as expected and be available to the inventory when expected. Let us also suppose that all other weapon systems being considered for procurement have this performance and schedule certainty. We would then know the effectiveness of each of the alternatives only to the extent that we could be confident of our knowledge of the presence and extent of the threat that these systems would counteract. Only in the case of certain knowledge of future threats would the cost-effectiveness comparison of the weapon systems reduce strictly to the cost components. Then the probability of selecting the most cost-effective weapon systems would be directly related to the <u>accuracy</u> with which we could predict the costs.

The accuracy of the cost estimate has several dimensions. Two of the essential features are bias and dispersion. Bias refers merely to any consistent tendency for estimates to be lower or higher than actual cost outcomes; dispersion relates to the probability that an individual estimate may differ from the actual cost, even after the usual bias of estimates is taken into account. Assume that two weapon systems, with approximately the same range of estimated costs, have been proposed to combat a particular threat. If one type of weapon system tended to have an upward bias in its cost estimates, and another type had downward-biased estimates, the cost-effective comparison between these types of weapons might imply the wrong selection. There would be a high probability of selecting the second type, when in fact the first would be the better choice. Similarly, if the dispersion of cost estimates of weapon systems is high, we should be less confident of correct cost-effective comparisons between contemplated procurements because we would be less confident that current estimates (on which a decision must be based) are sufficiently close to what the systems will actually cost.

Cost factors are mainly useful in the bias aspect of the accuracy of our knowledge about weapon system costs. For this reason, comparison of cost factors is a sufficient procedure for evaluating procurement experience only when dispersion as well as the other possible uncertainties mentioned above are not present.[1]

Aside from its limits in importance in cost-effectiveness studies, the cost factor also has limited interpretability. For example, if a weapon system is estimated to cost $4 million and actually costs $5 million, the cost factor for this system is 1.25. Many times this is described as "25 percent cost overrun." But the "blame" for the discrepancy between estimate and actual really cannot be identified simply from their ratio. Inspection of the value of a cost factor only tells the degree to which the actual cost and the estimated cost

[1]The methodology developed here allows for statistical analysis within which dispersion as well as bias can be investigated. See Sections IV (Discussion) and V below.

differ; it does not reveal the reason for the difference. The cost factor by itself does not distinguish between bias and overrun; the estimate may have been poor or the manufacturer or program management may have been inefficient.

In summary, cost factors are of limited value to the analysis of procurement effectiveness in several important respects. Only with confidence in our information on schedule, performance, and future threat, as well as with confidence of no dispersion component to our cost estimating accuracy would cost factors be the sole determinant of cost-effectiveness comparisons. Even if we know how large cost factors will be, additional information is needed (about estimating procedures, or contractor or management methods, incentives and objectives) before we can determine how they can be reduced. Still, cost accuracy is one element of the procurement effectiveness question. If cost accuracy has improved over past experience, we can be more confident than in the past of the correctness of our decisions based on cost-effectiveness comparisons.

The analysis to be presented in the following pages deals essentially with how one might compare the degree of cost accuracy between two periods of time. The specific application will be a comparison of the experience in cost accuracy in the 1960s with the experience of the 1950s. The question of why our actual experiences differ between the two decades and what these differences may be attributed to is not the direct objective of this Memorandum.[1] The empirical estimates of the model can also be used to predict probable cost growth and range of uncertainty of future procurements -- if the recent past is a suitable guide for such insights.

---

[1]See, for example, the research reported in Perry et al., System Acquisition Experience, and a proposal for future improvement, E. Dews, Buying Competitively After Designing to a Price, The Rand Corporation, forthcoming.

II. THE MODEL

PREVIOUS METHODOLOGIES

Although studies of the policy-related issues of weapon systems acquisition have varied substantially, the actual empirical underpinnings of these analyses have been quite similar.[1] Table 1 displays succinctly and somewhat simplistically the evolution of studies using cost factors.

Throughout the 1950s the objective was to find the earliest possible estimate for a given project and divide it into the most accurate and up-to-date appraisal of what the project was actually costing. Cost factors during this period were the ratio of unit costs of production at the time of the estimate and for the actual production item. The quantity proposed at the time of estimate was often substantially different from the quantity actually produced. Adjustment to make the quantities comparable left considerable uncertainty in what the cost factor for a particular weapon system should be. The Marshall-Meckling study used this "earliest-estimate basis" and reported averages for two different sets of cost factors for the same set of weapon systems. The averages reported under "B" in Table 1 are based on the set of factors calculated by Brussell and those under "S" are based on the set derived by Summers for the same weapon systems.[2]

---

[1] A. W. Marshall and W. H. Meckling, Predictability of the Costs, Time, and Success of Development, The Rand Corporation, P-1821, December 1959, also in The Rate and Direction of Inventive Activity, Princeton University Press, Princeton, 1962; R. Summers, Cost Estimates as Predictors of Actual Weapons Costs: A Study of Major Hardware Articles, The Rand Corporation, RM-3061-PR, March 1965, also in T. Marschak, T. K. Glennan, Jr., and R. Summers, Strategy for R&D: Studies in the Microeconomics of Development, Springer-Verlag, New York, 1967. The extensive data base compiled by E. R. Brussell of Rand during the years 1957-1961 formed the empirical basis for these two studies. See also M. J. Peck and F. M. Scherer, The Weapon Acquisition Process: An Economic Analysis, Harvard University Press, Cambridge, Massachusetts, 1962; and F. M. Scherer, The Weapons Acquisition Process: Economic Incentives, Harvard University Press, Cambridge, Massachusetts, 1964.

[2] Marshall and Meckling, Predictability of the Costs..., pp. 13-15.

Table 1

METHODOLOGIES

MARSHALL - MECKLING (1959)

Method: Means of Cost Factors (Actual/Earliest Estimate)

Result: Average for:

Fighters		Bombers		Cargo & Tankers		Missiles	
B	S	B	S	B	S	B	S
1.8	1.7	3.4	2.7	1.2	1.2	6.4	4.1

SUMMERS (1962)

Model: $F = Ke^{a_1 t + a_2 tA + a_3 A + a_4 A^2 + a_5 L + a_6 \ (T - 1940)}$

Result:

$$F = 11.9e^{\left[\begin{array}{llllll} .097t - & .032tA - & .311A + & .015A^2 + & .008L - & .075 \ (T - 1940) \\ (.47) & (-1.7) & (-1.6) & (2.1) & (4.0) & (3.8) \end{array}\right]}$$

Definitions:

F = cost factor; t = the timing of the estimate within the development program (expressed as a fraction of the program length); A = level of technological advance sought; L = length of the development program; and T = calendar year. The t statistics are presented in parentheses below coefficient estimates.

The Summers study contributed further to the methodological basis for evaluation and comparison of the cost aspects of weapon acquisitions. Summers hypothesized a causal relationship between the bias in cost estimates (as reflected in part by the extent to which the cost factor deviates from unity) and certain characteristics of the project, including the timing of the estimate within the course of the project. In other words, he acknowledged that there was often more than one estimate of what the project would cost and that these estimates would tend to be more accurate (or, at least, closer to the actual outcome) at later stages of the program. By taking the earliest estimate each time and comparing them across projects, an investigator could not be sure of having found an equally early estimate for each project or even equally early ones within types of projects. Summers therefore related the cost factor to the timing of the estimate within the development program (t), the level of technological advance sought for the new weapon (A),[1] the length of the development program (L), and the year in which the estimate was made (T). His main underlying hypothesis was that as the proportion of time increased, the estimators would have a better idea of what the program would actually cost and therefore the cost factor would be smaller.

Summers points out that at the beginning of a development period, the strategy would be to choose from among the various possible configurations of the system the one that seems to be the least costly to achieve the performance objectives. As time goes on, the trial and error method of achieving the desired performance leads to successively more expensive configurations. Thus the cost factor will approach unity from above as the estimated cost approaches the actual cost from below. He also hypothesized that, other things equal, a long development period will also contribute to more bias in the cost factor. Furthermore, the more technologically advanced the system under development, the more likely it is that those less expensive methods that have some hope of reaching the performance goals will not in fact achieve those goals.

---

[1]This survey measure is also used in the analysis reported below. See p. 25 for a description of A.

Finally, advancement in time is entered simply to represent the continuing improvements that are presumed to take place in the cost estimator's ability to predict the actual outcome with the information available to him.

The form of Summers' model is simply "the most satisfactory" of a variety of functional forms that were explored empirically. The indicated directions of bias from the features elaborated above do, in fact, find support in the empirically estimated model shown in Table 1. However, the estimated values for the part of the model quadratic in A lead to the rather peculiar result that for values of A up to about 11, increasing values of A lead to less bias, whereas for values of this technological advance measure above that point, the amount of bias increases.[1] Although the time variable (T) is suggestive of the constant progress in our ability to estimate what costs will be, it also leads to the rather unfortunate conclusion that by the 1960s (earlier or later in the decade, depending on the values of t, A, and L) we are bound to be predicting costs not merely less optimistically (F greater than but closer to unity), but increasingly pessimistically (F less than unity and declining) — that is, increasingly <u>less</u> accurately.

It is vital to our interpretation of such results that we understand the motivation underlying a model for analyzing cost factors before we can elaborate an explicit formulation amenable to empirical analysis.[2]

THEORETICAL MODEL

To understand the reasons for an overrun or a bias to be reflected in a cost factor, one must characterize the way actual and estimated costs might arise. Costs must be related to other important features

---

[1]Summers' sample values of A ranged from 5 to 16.

[2]The discussion of the theoretical model presented in this next section provides the basis for the analysis in Section IV of Perry *et al.*, *System Acquisition Experience*, as well as for the elaboration of that analysis reported in this Memorandum. See also Peck and Scherer, *The Weapons Acquisition Process*, Part III (esp. chaps. 9 and 11).

of an acquisition program — schedule, performance, intensity of effort, and technological advance sought. This characterization is not an attempt to describe mathematically the way costs actually arise from either production or management activities, nor is the estimated cost function an attempt at mathematically describing the cost estimating procedures that have been used. The mathematical relationships are essentially an attempt to capture the major influences on costs without trying to capture the mechanism by which they influence. That is, to analyze a phenomenon empirically, it is not necessary to describe in detail the mechanism by which the phenomenon occurs.[1]

The notation that will be used for this discussion is presented in Table 2. The cost aspect of the model can be characterized by two equations:

$$C_a = f_1(L_a, E_a, R_a) \tag{1}$$

$$C_e = f_2(L_e, E_e, R_e) \tag{2}$$

The main emphasis in this model is on a narrowing in the range of potential outcomes that is achieved during the development portion of the program. Thus, two of the three arguments of the f_1 actual cost function pertain to development -- program length and intensity of effort. The third argument is merely the resources required to produce the system configuration settled upon in development.[2]

It is not unreasonable to assume further that the decisions made during development determine the resources needed during production;

---

[1]On this point, see M. Friedman, Essays in Positive Economics, University of Chicago Press, Chicago, 1963, Chapter I.

[2]A production function characterization (labor and capital services as well as "learning by doing" resulting in units of the system) would be a reasonable approach to achieving a measure of "resource requirements." This will not be elaborated here. Instead, it is merely necessary to assume that decisions made as a result of the development phase -- characterized by L_a and E_a -- imply a certain commitment of resources to produce the system specified.

Table 2

NOTATION FOR CHARACTERIZING A PROGRAM

Definitions[1]

A	Technological advance sought
C^*	Total program optimal (that is, minimum) cost
C_a	Total program actual cost
C_e	Total program estimated cost
E^*	Optimal (that is, minimum-cost) level of effort (for example, engineering man-hours) per month for development
E_a	Actual level of effort per month for development
E_e	Estimated level of effort per month for development
L^*	Optimal (that is, minimum-cost) development program length, in months
L_a	Actual months for development
L_e	Estimated months for development
P	Desired performance
R_a	Actual resources devoted to production of the system
R_e	Estimated resources devoted to production

Identity

Cost Factor: $F = \frac{C_a}{C_e}$

Note:

[1]The variable "M", which is used as a surrogate for program length throughout the empirical analysis, is defined in Section III.

that is,

$$R_a = f_3(L_a, E_a), \tag{3}$$

and

$$R_e = f_4(L_e, E_e). \tag{4}$$

For example, if an aircraft is to be designed during a short development phase with minimum effort, the designers may simply choose to use a costly to produce but powerful fan jet engine; if they have more time for designing, they may be able to achieve desired performance objectives with a turboprop engine. These decisions determine the cost of the production phase of the procurement.

The usual U-shaped curves relating costs to program length are shown in Fig. 1. There is presumed to be a development strategy -- program length and development effort -- that will lead to the minimum total cost (C*) of development and production (see Fig. 1a). For shorter program lengths, lower development costs might arise, perhaps by choosing at the outset a quite costly configuration to produce but one that is reasonably certain to achieve performance goals. Thus, the total costs -- including the high production costs -- will be larger than C*. This figure can be thought of as a range of procurement strategies from concurrency of development (shorter development programs in which all components are developed concurrently with integration of the system at the last stage of development) to sequential development (longer programs in which the essential parts of the configuration are assembled and tested earlier in the development phase). The curve of Fig. 1a indicates that a very compressed schedule for development may be able to achieve the desired system (characterized, perhaps, by an elaborate set of minimum performance goals), but that one can also seek to design a minimum cost development strategy.[1] For a different weapon with more demanding performance requirements

---

[1]The minimum of the curve could actually remain flat over a range of development program lengths.

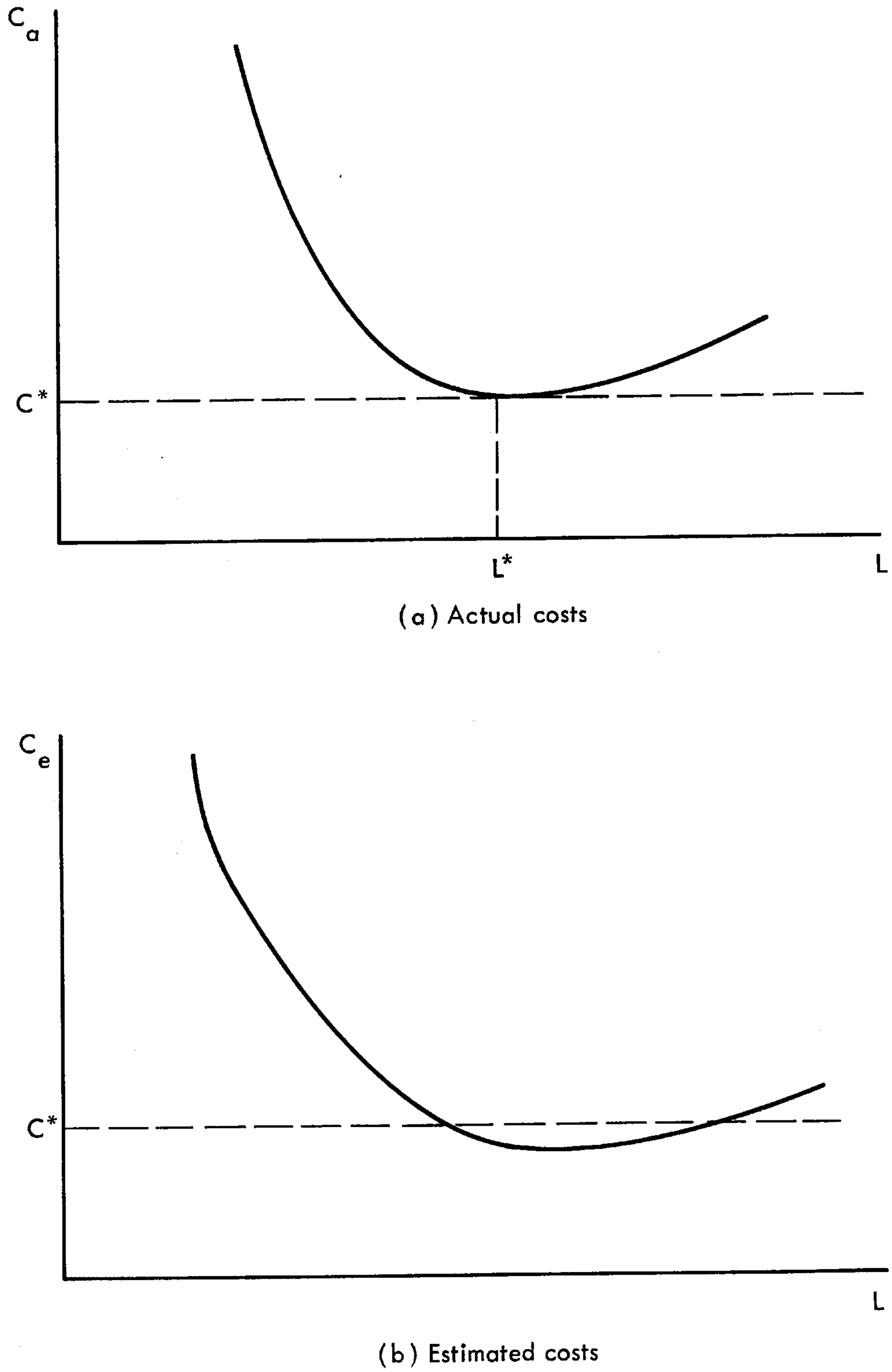


Fig. 1—Relationship of cost to program length

the requisite levels of effort, program length, or resources for production would shift upward.

In summary, equation (1) deals with features affecting the actual program costs, which are presumed to be related to the actual development schedule (L_a), level of effort (E_a), and requisite production resources (R_a). Note that the degree of cost increase from a schedule stretchout, for example, is not necessarily assumed to be equal to the cost increase associated with an equal amount of schedule compression (that is, the "U" shaped curve in Fig. 1a is probably not symmetric). In fact, the curve undoubtedly approaches an asymptote at least for very short lengths of development. A final influence on actual costs is the total accumulation of inefficient contractor performance or administrative laxity in controlling costs. This phenomenon may be captured in large part by the length of time of the development program (L_a) and the resources in production (R_a).[1]

Cost estimates characterized by equation (2) take into account the same basic influences,[2] except that there is an added interpretation that can be given to the inclusion of L_e. This term not only reflects the "U" shaped curve of higher costs for deviations from optimal program length, but may also capture a range of behavior from accurately estimating the inefficiencies in the procurement procedures to estimating costs in an increasingly optimistic or myopic manner as the distance to the end of development of the system increases.

These equations form a scheme for translating reasonable hypotheses by which costs and cost estimates arise into a scheme for analyzing the bias in cost estimates or the extent of cost overruns, using cost factors. To see this translation, consider a cost factor, which is a measure of this bias or overrun:

---

[1]In this use, L_a may be somewhat deficient indicator of inefficient use of resources in production [through equation (3)].

[2]Let me emphasize again that I am not trying to suggest a method for actually estimating costs, but only for summarizing the process sufficiently accurately that it can be analyzed empirically.

$$F = \frac{C_a}{C_e} \quad . \tag{5}$$

Substituting from equations (1) to (4) above into (5) yields

$$F = F(L_a, L_e, E_a, E_e) \tag{6}$$

F deviates from unity (perfect estimate and perfect cost control or the same extent of overestimate and cost slippage) only through the following mechanisms:[1]

- The height of the minimum cost level is underestimated.
- The effect on costs of deviations from optimal development program length is underestimated.
- The effect on costs of deviations from optimal development effort is underestimated.
- Management or production inefficiency is not accounted for in the estimate, or is misjudged.

The first point is simply the possibility of direct bias due to the cost estimator's understanding of the height of the usual schedule-to-cost curve (Fig. 1b). It may be, for example, that he thinks the curve (and therefore the minimum point) is lower than it actually is, and therefore the line at the level of actual minimum costs (C*) intersects the estimated curve, rather than being tangent to the curve as it is in fact.

Bias may also be due to the cost estimator's optimism in the degree to which schedule compression or stretchout leads to larger costs. This part of the bias, in other words, comes from an optimistic view of the steepness of the sides of the schedule-to-cost curve (Fig. 1a). Another possible source of bias comes from a similar optimism with respect to the effect of effort on cost. The final possible source of bias, as well as the cost overrun possibility,

---

[1]Bias in costs has almost always appeared as overly optimistic estimates or actual overruns — that is, cost factors greater than unity. Thus, all bias influences discussed here are stated in those terms.

is suggested by the last point of the list. There may be inefficiency of the program manager or contractor in conducting the program, or uncertainty within which the cost estimator must make his projection. These possible sources of bias or overrun in costs from length of development or level of effort are summarized graphically in Figs. 2 and 3 (in each figure, the other variable is held constant).

To see the underlying influences behind these sources of bias in costs, we must look further into the other two features of a program that are co-determined with the cost of the acquisition; namely, the length of the program and the amount of effort.[1]

The measure of effort per unit of time is presumed to be related to schedule and performance features of the desired system according to the following relationship:

$$E_a = f_5(L_a, A(P)), \tag{7}$$

$$E_e = f_6(L_e, A(P)). \tag{8}$$

There is an important assumption embodied in equation (7), as illustrated by Fig. 4. The parable for following Fig. 4 may be described as follows: The conceptualization of a new weapon system, deemed essential to future defense against a perceived threat, takes the form of specification of the increased performance characteristics needed from the new weapon (P). These characteristics are portrayed schematically simply as the height up the vertical scale in Fig. 4(a). Corresponding to this performance characterization, there is a set of technological advance levels which, if attained, make it likely that the performance characterizations specified will in fact be achieved. The very fact that a new weapon is deemed necessary implies that the performance characterization of this new weapon will be somewhat higher than those in existing stock, and therefore that the level of P will

---

[1]Throughout this discussion, effort is thought to be measured in terms of some work load; for example, it might be measured as engineering man-hours per month for the development phase of the acquisition.

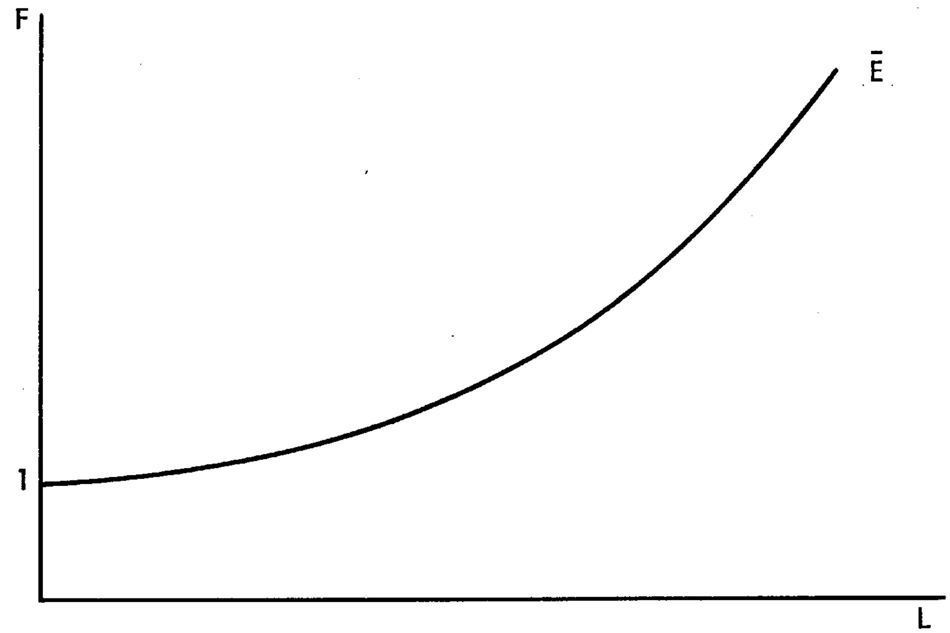


Fig. 2—Relationship of cost factor to program length

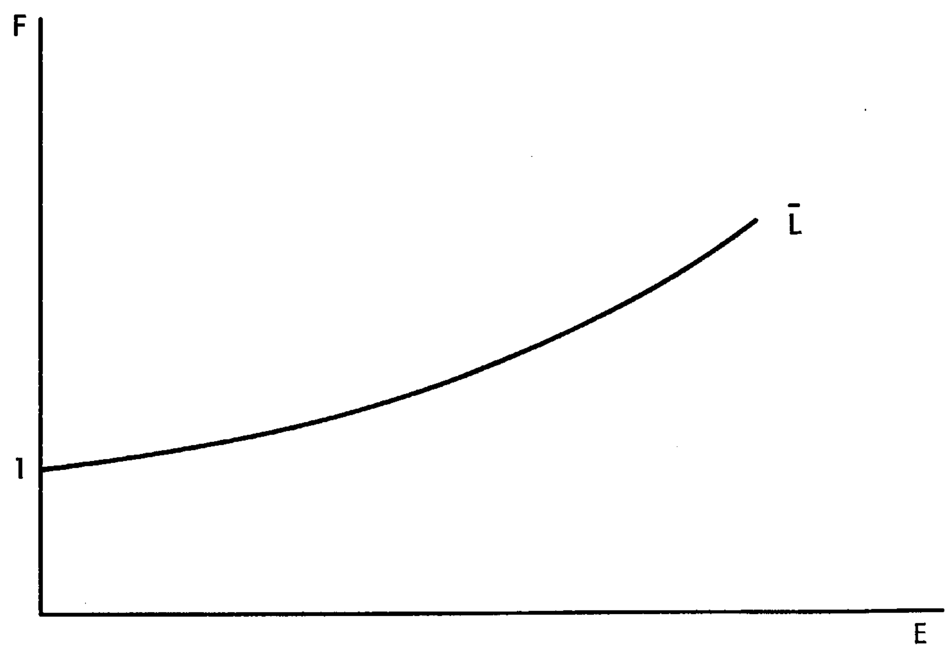
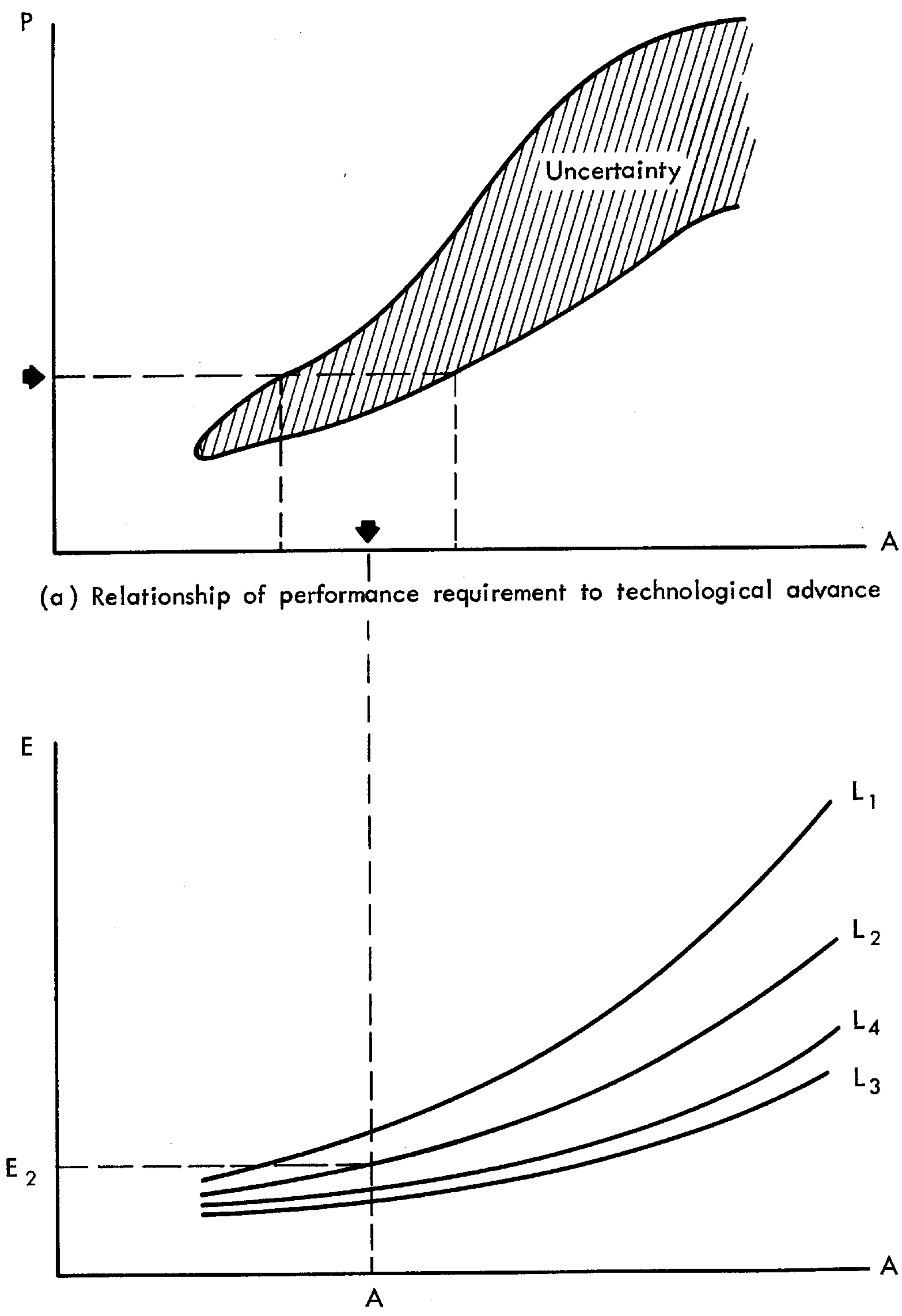


Fig. 3—Relationship of cost factor to effort



(a) Relationship of performance requirement to technological advance

(b) Relationship of effort to technological advance sought for a given program schedule

Fig. 4—Influence of performance specification

never be zero.[1] Correspondingly, the fact that this weapon is something new implies that there will be a certain minimum technological advance (possibly measured as the amount of skill necessary to put together a new configuration based on off-the-shelf components).

The parable continues to the assumption that within this band of technological advance, a particular level of technological advance is sought. Corresponding to this level and corresponding to a prescribed schedule for achievement of the new weapon system (given technical efficiency), a level of effort is uniquely determined [see part (b), Fig. 4]. This is, of course, a highly simplified characterization of the way in which weapon procurements are actually structured. It is, however, a useful simplification for the purpose of concentrating on cost aspects of weapon systems, while still taking into account, at least broadly, the performance features of the systems.

If we suppose that the level of A in Fig. 4 corresponds to curve A_2 in Fig. 5, then the increasing values of L, labeled $L_1, \ldots\ L_4$ in Fig. 4, correspond to increases in L on the horizontal axis of Fig. 5. This latter figure portrays another characteristic of equation (7). In addition to showing how performance and technology influence the level of effort of the development phase, it portrays the assumption that one can reduce the level of effort by lengthening that program's development (that is, the length of time until the weapon is to become operational). It is also assumed, however, that one can reduce that effort to a minimum level, at which point any increases in the length of the program incur diseconomies. For technological advance level A, that point in length of program corresponding to minimum effort is L_3 (L_3s in Fig. 4 and Fig. 5 are the same).

---

[1] This description is possible only for performance measured in one dimension as is done in this example. The appropriate concept for a vector of performance measures, some of which may not be as advanced as for existing weapon systems, might be a measure of the length of the vector relative to an appropriately selected set of basis vectors.

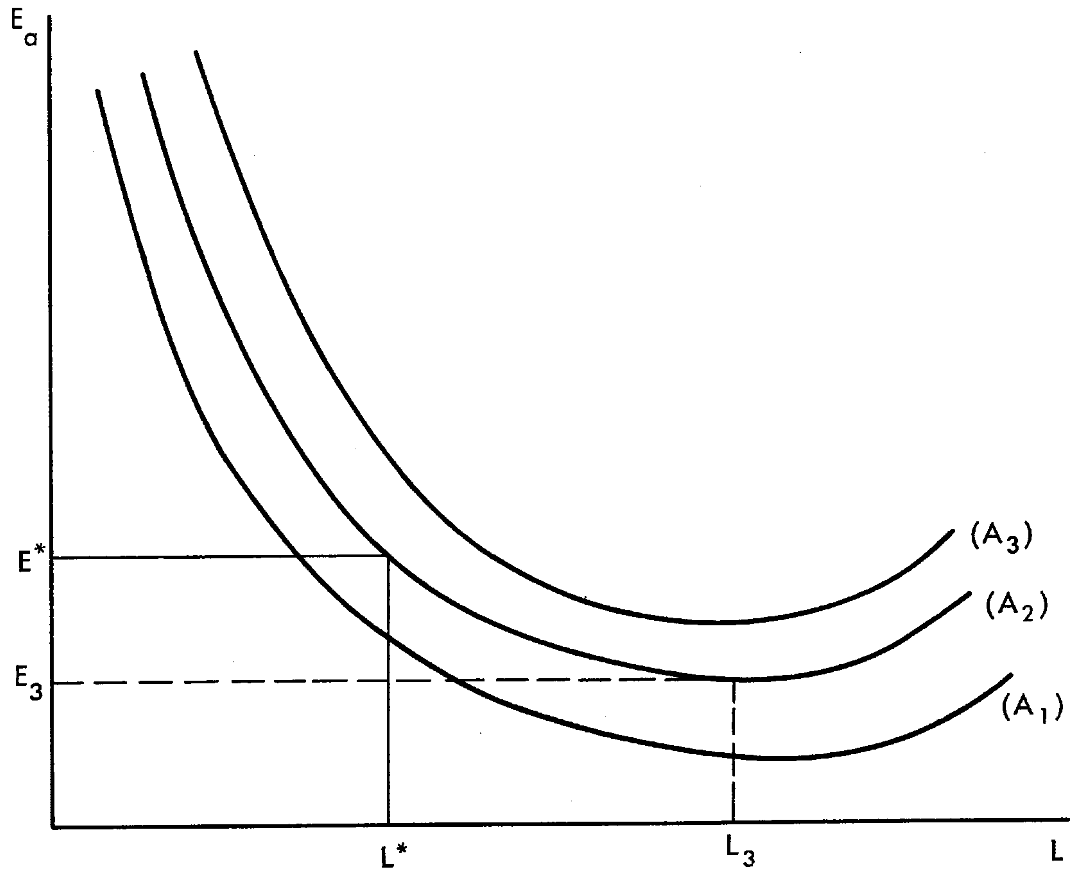


Fig. 5—Tradeoff between length of development
and intensity of effort

Equations (9) and (10), then, complete the model by specifying values for the actual program length and the estimate for that length:

$$L_a = f_7(L^*), \tag{9}$$

$$L_e = f_8(L^*). \tag{10}$$

The schedule for the program is presumed to be designed with cost in mind and is therefore related to the minimum-cost schedule; it is also influenced by purely external factors such as the desire that the system be available prior to the realization of the anticipated threat. Although the estimate of program length parallels this reasoning, it may be overly optimistic of the minimum-cost schedule for longer programs [that is, $f_7' > f_8'$].

The three main features of this model -- cost estimation bias or cost overrun, length of development, and level of effort -- can now be combined graphically as shown in Fig. 6. This 3-dimensional diagram combines the descriptions of Figs. 2, 3, and 5. The faint curved lines represent a surface of values of F for various combinations of E and L. This surface comes into the "room" delineated by the F-L plane (from Fig. 2) as the "back wall"; the F-E plane (from Fig. 3) as the "side wall"; and the E-L plane (from Fig. 5) as the "floor." On the "floor," the minimum E-L combinations for a given level of technological advance are displayed. For each technological advance level, the minimum-cost L and E, (that is, L* and E*) determine a point on the surface; these points for varying levels of technological advance form the line depicted by the broken arrow. The value of any F on the surface can be derived from equation (5).

To this point, we have not even considered the possibility that the level of technological advance sought does not lead to achievement of the performance characteristics chosen (recall that Fig. 4a depicted a band of uncertainty). If it does not, or if the performance demands are increased, there may be a decision in the course of the program to increase A. Thus, in Fig. 6, the arrow on the "floor," depicting an outward shift in the possible effort and program length combinations,

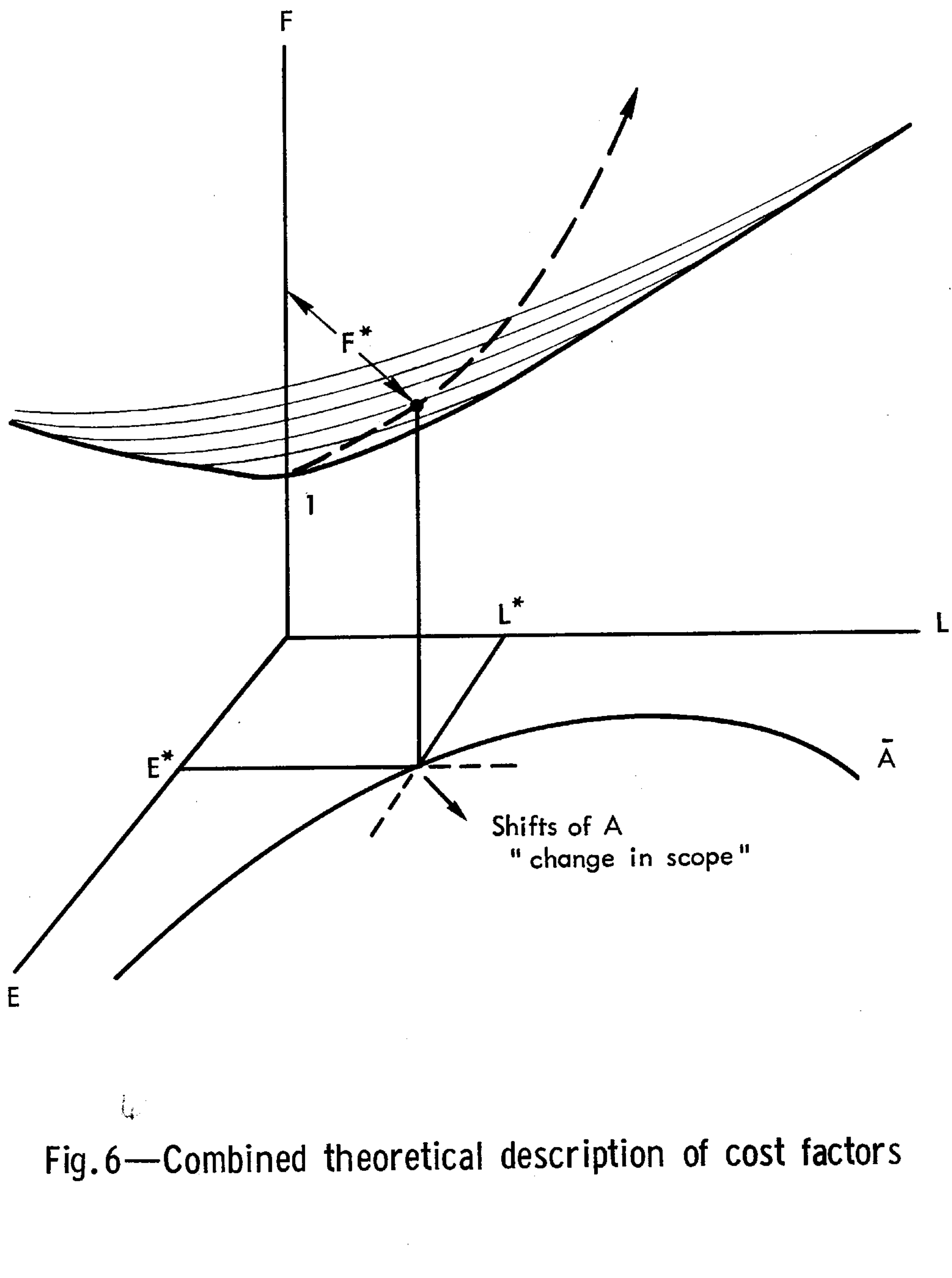


Fig. 6—Combined theoretical description of cost factors

can be considered a characterization of technological uncertainty or of what is often called "changes in scope." This latter expression is usually used to describe either the customer's or the developer's changing the objectives of a program during its development phase. The effect, whether large or small, is to alter the configuration or performance of the article being designed or tested and thus to change the optimal program length or intensity of effort. It is a reasonable assumption that the development program is even less "optimally" structured to achieve the modified system with minimum cost than it was to achieve the original configuration. Attempts are usually made to capture the tendency for scope change in the cost estimate based on past experience. Nevertheless, scope change almost always appears as an addition to estimated program costs and is an important contributor to cost growth.

In summary, this model presents a simplified representation of the influences of performance, technological advance, and scheduling on the degree of bias or overrun in costs. The scheme relates the three dimensions of cost bias, program length, and intensity of effort. The next section contains an explicit model specification that can be analyzed empirically.

III. EMPIRICAL FORMULATION

Comparison of bias in cost performance between the 1950s and the 1960s can be accomplished through the model presented graphically in Fig. 6. This comparison can be made by measuring the extent to which the curved surface characteristic of cost factors in the 1950s flattened out closer to the unity plane parallel to the "floor" for the 1960s.[1]

Equation (6) can be simplified for the empirical exploration because the acquisitions under investigation have all either been concluded or are sufficiently advanced that we can be reasonably confident of knowing the actual cost and consistent date for the "end of development" of these projects.[2] Thus the L_e argument of the function in equation (6) does not actually need to enter a function used to analyze the data available to us. Although this is true for analyzing the historical data, in using any such analysis to predict future results, we must keep in mind that we will not know program length with accuracy, just as we will not know the cost of the program with accuracy (see Section V).

---

[1] In evaluating the accuracy of estimates of weapon system costs, dispersion around the actual cost, as well as bias, must be considered. As mentioned above, we cannot be very confident in making decisions concerning choice of acquisitions if, on the average, our predictions are correct, but for any particular prediction, the actual cost could easily differ by a factor of two or three.

The model used in this analysis is directed at identifying bias; other aspects of the total question of predictive accuracy will be considered in evaluating the empirical results in Sections IV (Discussion) and V.

[2] For several of the systems included in the DDR&E survey reported in Perry <u>et al.</u>, <u>System Acquisition Experience</u>, the program was not sufficiently advanced to determine the date with the accuracy implied by this statement. These projects were therefore excluded from the sample and are not included in the list of available projects presented in the Appendix.

To simplify notation, L_a is replaced by L in this section and subsequently. Since no data were available to distinguish E_a from E_e, a general intensity of effort (E) measure will be used in the empirical model.

The natural assumption for the way development program length (L) and intensity of development effort per month (E) influence the functions specified in Section II is that their product (L•E) is the essential feature -- measuring total effort in development. The following equation, then, captures the general nature of the prediction of system cost bias described in Fig. 6:

$$F = ae^{bL \cdot E} \qquad (11)$$

As mentioned above, it is assumed in using F in this equation that the procurement is sufficiently far in the past that actual cost is known. The cost factor is presumed to be exponentially related to the total effort involved in bringing the system to fruition.[1] This formulation captures the effect on cost factors characterized in the theoretical model; namely, that for longer programs or programs requiring high levels of effort), the bias due to optimism in cost estimation or to inefficiencies in program management tends to be larger. Thus the coefficient b is hypothesized to be positive.

---

[1]Throughout this Memorandum I have emphasized that the basic motivation for this analysis is to develop a framework within which cost factors can be analyzed. The implication from such an empirical formulation (by multiplying by C_e) is that

$$C_a = aC_e e^{bLE}$$

and therefore that

$$\frac{\partial C_a}{\partial L} = abEC_e e^{bLE}$$

$$= bEC_a$$

That is, the marginal impact on actual costs of a change in development program length depends not only on the monthly effort in development but also on the magnitude of costs. Thus, implicit in the use of a cost factor in this type of model is the assumption that the impact of changes on cost growth is in percentage terms in the program's costs and not simply in additional dollars. The validity of this assumption is conjectural; further investigation is continuing to determine if another structure using cost factors or a model that abandons the cost factors of common reference would be preferable.

Along with the presumption that estimates for longer programs will tend to be more biased, there is also the likelihood that these longer programs, or ones involving larger technological advance, tend to have generally more cost uncertainty. Thus, as the total program effort (L·E) increases, both larger bias and dispersion of the estimates are presumed to occur. The exponential form of the function is chosen so that one obtains the positive association between bias and intensity of program effort. Furthermore, in the form in which the equation will be estimated, heteroscedasticity is reduced. Thus, the equation can be transformed into an expression linear in the coefficients as follows:

$$\log F = \log a + bL \cdot E + \varepsilon \qquad (12)$$

in which ε is the statistical disturbance term. The argument is that the variability of F is likely to increase as L·E increases. However, it is a more reasonable assumption that the variability of log F remains essentially constant (as total program effort increases) and thus the disturbance in equation (12) is more nearly homoscedastic.[1]

The final requirement for analyzing this model is the choice of measures for the development program characteristics, intensity of effort (E) and program length (L).

The data available for this study are presented and discussed in the first section of the Appendix. The two sets of data used are called "1950s" and "1960s" samples since the weapon systems in each are predominantly from these periods. However, the actual range of sample observations (from earliest estimate to latest operational

---

[1]Of course, heteroscedasticity could be reduced in other forms of the model besides the exponential form. And in fact, the empirical results might display an essentially linear slope over the range of estimates. The hypothesis here is that estimates deviate from actual costs at least by a linear function of total program effort, and very possibly at an increasing rate. Since the exponential form leads to the implication of very large increases in F for very long or difficult development efforts (for example, as L·E increases), one should be cautious not to draw inferences about systems in the future (as in Section V) that are outside the range of observations of L·E in the current samples.

status) of each is February 1947 to July 1961 for the "1950s," and January 1958 to December 1969 for the "1960s." Therefore, quotation marks will be used on decade references whenever they are employed as the names of the samples.[1]

No direct measure was available of the intensity of effort in development of the systems in these samples. However, a variable quantifying the magnitude of technological advance sought for each of the systems under consideration has been obtained. The values of this variable, A, for the systems of the "1950s" were compiled through a survey of Rand engineers taken by Robert Summers. The respondents were asked "to rate subjectively the magnitude of the improvement in the state of the art required for each of the development programs."[2] The results were then consolidated to yield a value of A for each system. These measures were taken as given for use in this study; no adjustments or reevaluations were attempted. Values of A for the "1960s" systems were also obtained through a survey. The respondents were asked to estimate the magnitude of technological advance sought on a scale of 0 to 20 (or more) and were instructed to make these estimates comparable to the measurements for the "1950s." The estimates were then averaged to obtain a value of A for each system. These surveys were limited in scope and resulted in only a few responses. Measurement of the technological advance sought is thus at a preliminary stage; an expanded and refined survey would be desirable. This survey measure of technological advance sought, along with development

---

[1]The data used to establish cost factors for the "1960s" were originally gathered by Rand during a study for DDR&E, but have been updated to reflect the best estimates of actual costs as of December 1969 for programs still in progress.

Factors for the "1950s" are all for Air Force projects. These data were gathered by Eugene R. Brussell in the late 1950s and early 1960s, and the factors were calculated by Robert Summers for use in Cost Estimates as Predictors.... The cost factors themselves have been adjusted for both quantity and price level changes during the course of the programs. See Perry et al., System Acquisition Experience, Sections I and II for a discussion of the scope and limitations of the original "1960s" sample. See also the discussion in the Appendix.

[2]Summers, Cost Estimates as Predictors..., p. 25.

program length, can be transformed into a surrogate for the intensity of effort variable [see equation (7)]. Several transformations of A (and L) into E will be investigated in the empirical results reported in Section IV.

To understand the measure of development program length used in the analysis, consider Fig. 7 which displays the level of program activity through time. "Program activity" might be defined as the amount of both development and production resources committed to the program; it could, perhaps, be measured by dollars spent per year or by engineering and production man-hours per year. Figure 7 is not intended to portray the course of any particular program or programs; it is, rather, a theoretical representation of the course of programs in general. Individual programs might differ in several important respects.

First, the areas beneath the development and production curves would probably vary with the magnitude of technological advance sought for a specific program. For example, the area beneath the development curve might be larger, relative to the area beneath the production curve, for higher levels of technological advance. In addition, "humps" in the development curve might arise because of unsuccessful tries, and "humps" in the early part of production might occur because of retooling needs. There might also be additional large development commitments after initial operational delivery (IOD) because of new versions or models of the weapon system. Finally, the later part of the production curve might vary with changes in the total number procured and the timing of delivery, for such purposes as unforeseen replacement needs or maintenance of the production line. This variability is depicted in the later part of production by a band rather than a line. Since the purpose of Fig. 7 is to present a simplified and generalized concept of the course of program activity through time, the qualifications listed above are not explicitly depicted.

Within the framework, the measure of L for the systems of the 1960s is reasonably clear. The date of the estimate made near the time of DoD approval of a development program provides an initial

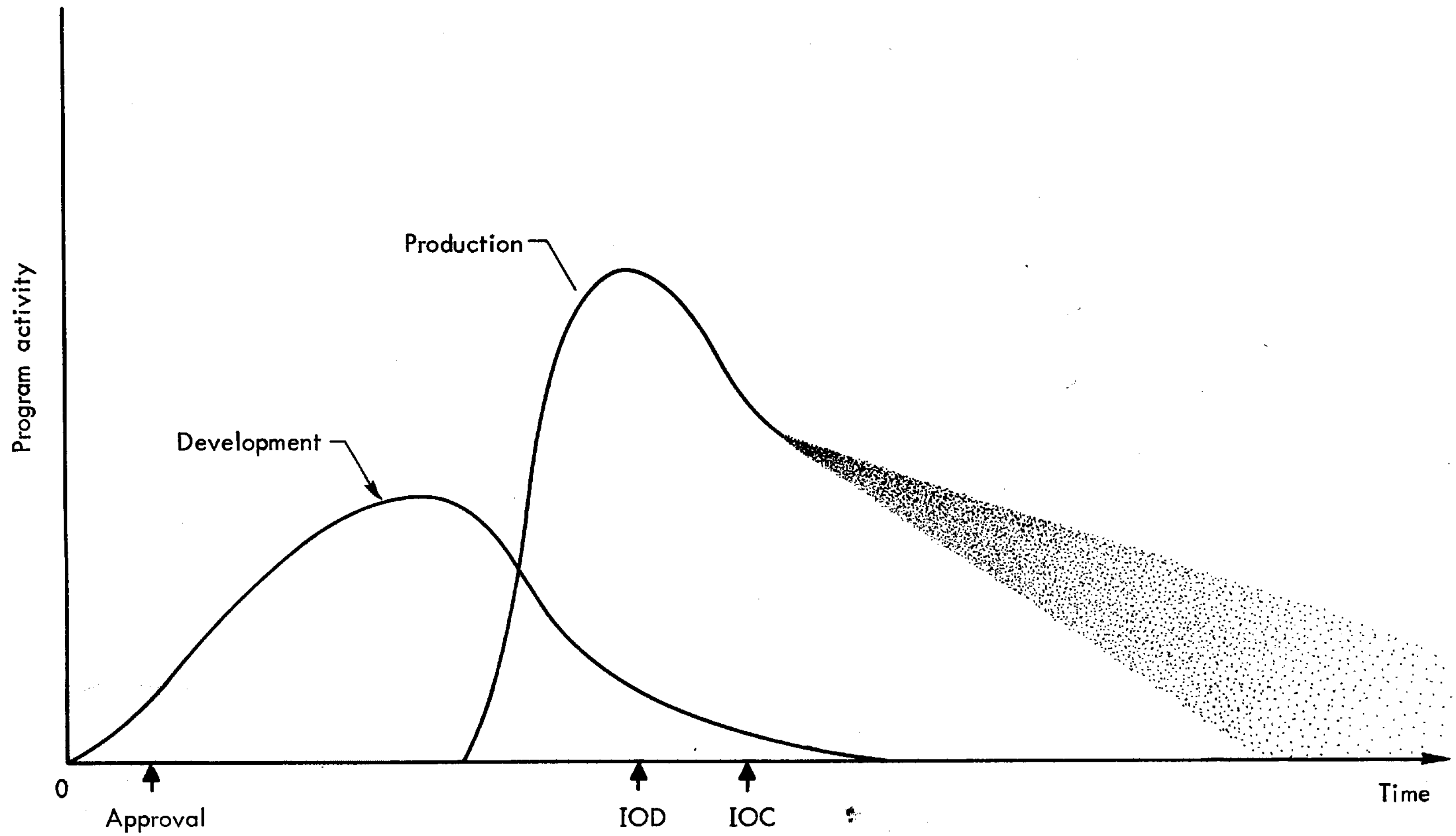


Fig. 7—Levels of program activity

base point because it marks the approval of the commencement of engineering development and provides the least ambiguous indicator of the beginning of a program. IOD was chosen as the end point because it seems reasonable to assume that most of the development uncertainties captured by the model are resolved by the time of IOD.[1] The justification of this end point follows from the conceptualization of the course of program activity as displayed in Fig. 7. IOD occurs after the development phase is largely completed. This means that the length of time from program approval to IOD is the relevant time period during which most of the technological uncertainty leading to cost uncertainty in development is resolved. Also, any initial difficulties in tooling-up will be overcome during this time so that the size of this investment would also be known. Thus, two estimation uncertainties captured by the model (namely, early stage of conceptualization and major technological advance) and a major production uncertainty (tooling-up cost) will be largely resolved by IOD.[2] If the contractor is not

---

[1]This end point was chosen mainly because such information was available for all systems in this decade. Initial operational capability (IOC) would be another possible choice for the end of the program; however, the data available were not sufficient to allow the use of that measure of L.

[2]To the extent that problems with the weapon system arise after IOD and require re-engineering or re-tooling, this end point for program length will be deficient. This may be a problem, for example, for the F-111 and the C-5A programs.

This discussion of Fig. 7 reveals another way of interpreting the predictability of production outcomes from development activity (as hypothesized in equations (3) and (4) in Section II). The total cost estimate at the beginning of the program can be broken down as the sum of estimates of development and production costs (that is, $C_{e_0} = C_{d_0} + C_{p_0}$). Similarly an estimate of the costs could be made near the end of development at IOD ($C_{e_{IOD}} = C_{d_{IOD}} + C_{p_{IOD}}$). At the end of the program, the actual costs are known ($C_a = C_{d_a} + C_{p_a}$). Since development is essentially over by IOD, the estimate $C_{d_{IOD}}$ will be very close to C_{d_a}. The additional assumption embodied in equations (3) and (4) is that since the configuration is resolved by the development program, inefficient in production, the total cost of the weapon system could be estimated with reasonable certainty by IOD. In the empirical analysis, any inefficiency in production will probably contribute to an upward influence on the coefficient of this measure of program length or on the intercept, so that although inefficiency is not measured separately, it is not ignored.[1]

In most of the empirical results presented below, however, a somewhat different measure of time in development is actually used. The new measure is required since the choice of program length for the "1950s" is otherwise rather arbitrary. The data for the "1950s" include cost factors calculated from estimates made at various points throughout the program (see Table 8), rather than one cost factor for each program calculated from an estimate made near the time of program approval. It was not possible to establish a beginning point for program length (and obtain an estimate made at that time) for the systems in the "1950s" sample comparable to that used for the programs of the "1960s" sample. IOD was still used as the end point.

For the "1960s" sample there are also estimates available at several points in the program which can be used in the analysis. A new variable "M" therefore has been defined as the number of months between the date of the estimate on which each cost factor is based and IOD. This variable (M) will probably average fewer months than would program length (L), but hopefully will provide a better comparison than would using earliest estimates or all estimates in the "1950s," and comparing those results to estimates made near the time of approval in the "1960s."

$C_{p_{IOD}}$ will be quite close to C_{p_a}. Any remaining bias in predicting production costs should be captured by the intercept "a" in equation (11).

[1]By measuring L as the time from program approval to IOD, we preclude any possibility of the model evaluating inefficiency in production separately. This task is also made difficult, however, by the use of cost factors, which lump estimation inaccuracy and inefficiency together.

IV. COMPARISON OF THE 1950S AND 1960S

From the empirical model and data measurements described in Section III, the cost experience of the last two decades can be analyzed. This will be done not only by estimating the model with the full sample for the "1950s" and "1960s," but also using two subsamples.[1]

The full coverage of the "1950s" sample includes fighter, bomber, cargo, and tanker aircraft, and ICBM, IRBM, and other missiles; the "1960s" sample includes fighter, cargo, attack, and STOL aircraft as well as a helicopter, sonars and a sonobuoy, a space propulsion vehicle, a battlefield support missile, and two programs related to an ICBM. The primary subsample, aircraft and missiles, actually includes the full sample for the "1950s," but eliminates the sonars and sonobuoy from the "1960s" sample. Therefore, this subsample provides the most comparable results for the comparison of acquisition cost experience between the last two decades. The other subsample reported here includes only aircraft.

EQUATIONS

We shall first explore the appropriate transformation of A to use as a proxy for E. To accomplish this, the variable M defined at the end of Section III will be used throughout as a proxy for L. An additive function of the proxies for development program characteristics has been chosen so that separate coefficients can be estimated for M

---

[1]As mentioned in Section III, the headings "1950s" and "1960s" are used throughout for convenience of reference. Actually the "1950s" sample spans the time from programs begun in 1947 to those operational in 1961, while the "1960s" programs were begun anywhere from 1958 to 1966 and became operational by 1969. The "1950s" data were obtained by Rand researchers in the early 1960s; the "1960s" sample was obtained early in 1969, also by Rand researchers. In the latter sample, the attempt was made to obtain a cost estimate established near the time of DoD approval. It was considered preferable to maintain the "1960s" sample to include all of the cost factors based on such estimates, rather than to combine these estimates with the others in the "1950s" and thereby create non-uniform sampling procedures within the data for one of the periods.

and various functions of A. Thus the structure for the model resulting in the estimates presented in Table 3 is linear in M and f(A):

$$\log F = \log a + bM + cf(A) \tag{13}$$

With M substituted for L in equation (12), this function is derived from the assumption that equation (7) relating program length and technological advance to effort takes the form:

$$E = 1 + \frac{cf(A)}{bL} \quad . \tag{14}$$

Four functions f(A) are investigated: $\log_e A$, A, A^2, and e^A . For ease in comparing the degree of fit of the various functions, Table 4 presents the different values of R^2 -- the proportion of the variation of the cost factors explained by the model. Each line contains a separate sample coverage, so the values of R^2 are directly comparable only across a line. In every case, an increasing function of A is preferable to simply A or $\log_e A$. Only for the full sample for the "1960s" does the R^2 loom slightly larger for the A^2 form than for the e^A form; therefore, the latter form is taken to be the best functional form for this linear structure.

Returning to Table 3, then, Part IV contains the coefficient estimates for the structure of most interest. In this part, as well as in the others, the primary point of interest is that the coefficient "b" of the program length surrogate declines between the "1950s" and "1960s," while the coefficient "c" of technical difficulty increases. In every case, "b" for the "1960s" is essentially zero (usually slightly negative) and insignificant.[1] By contrast, "c" is positive -- quite significant and usually about 4 times as large in the "1960s" as in the "1950s." Thus program length has been the main feature associated with cost optimism in the "1950s" while technical difficulty has been the major influence on cost factors in the "1960s."

---

[1]These negative values of "b" lead to the somewhat peculiar implication [from equation (14)] that months of development (L) and intensity of effort (E) are positively related rather than inversely. Since the negative values are all quite insignificant, the more appropriate conclusion may be the f(A) alone is related to E; this structure form is taken up below.

Table 3

VARIOUS TRANSFORMATIONS OF "A" IN LINEAR STRUCTURE OF THE MODEL

Coverage[a]	log a	b	c	Standard Error	R^2	Sample
			I. log F = log a + bM + c [log (A)]			
1. Aircraft & Missiles						
"1950s"	-.413	.0128	.1195	.487	.36	55
	(-0.5)	(4.6)	(0.4)			
"1960s"	-.733	-.0006	.4657	.179	.42	21
	(-2.6)	(-0.3)	(3.5)			
Combined	-.802	.0100	.3149	.439	.32	76
	(-1.9)	(4.7)	(1.8)			
2. Aircraft						
"1950s"	-.763	.0122	.3105	.436	.37	34
	(-1.0)	(3.5)	(1.0)			
"1960s"	-.626	.0025	.3945	.169	.55	13
	(-2.2)	(0.9)	(2.9)			
Combined	-.801	.0103	.3488	.384	.36	47
	(-1.9)	(3.9)	(1.9)			
3. Full sample						
"1960s"	-.753	-.0008	.4971	.219	.34	25
	(-2.3)	(-0.3)	(3.3)			
Combined	-.714	.0099	.2877	.436	.31	80
	(-1.8)	(4.7)	(1.7)			
			II. log F = log a + bM + cA			
1. Aircraft & Missiles						
"1950s"	-.294	.0126	.0148	.486	.36	55
	(-0.9)	(4.5)	(0.5)			
"1960s"	-.196	-.0006	.0516	.165	.50	21
	(-1.5)	(-0.3)	(4.2)			
Combined	-.415	.0097	.0326	.436	.33	76
	(-2.2)	(4.5)	(2.0)			
2. Aircraft						
"1950s"	-.349	.0120	.0292	.435	.38	34
	(-1.1)	(3.4)	(1.1)			
"1960s"	-.168	.0024	.0427	.157	.61	13
	(-1.2)	(1.0)	(3.4)			
Combined	-.365	.0100	.0353	.382	.37	47
	(-1.8)	(3.8)	(2.1)			
3. Full sample						
"1960s"	-.171	-.0008	.0541	.210	.39	25
	(-1.1)	(-0.4)	(3.7)			
Combined	-.361	.0096	.0298	.434	.32	80
	(-2.0)	(4.6)	(1.9)			

Table 3 (cont'd)

Coverage[a]	log a	b	c	Standard Error	R^2	Sample
			III. $\log F = \log a + bM + cA^2$			
1. Aircraft & Missiles						
"1950s"	-.234	.0123	.0008	.486	.37	55
	(-1.2)	(4.4)	(0.7)			
"1960s"	.037	-.0005	.0026	.155	.56	21
	(0.4)	(-0.3)	(4.8)			
Combined	-.254	.0095	.0015	.434	.34	76
	(-2.0)	(4.4)	(2.2)			
2. Aircraft						
"1950s"	-.192	.0119	.0013	.434	.38	34
	(-1.0)	(3.4)	(1.1)			
"1960s"	.023	.0024	.0021	.148	.65	13
	(0.2)	(1.0)	(3.7)			
Combined	-.187	.0098	.0016	.380	.37	47
	(-1.4)	(3.7)	(2.2)			
3. Full sample						
"1960s"	.077	-.0008	.0026	.205	.42	25
	(0.7)	(-0.3)	(4.0)			
Combined	-.214	.0094	.0014	.432	.32	80
	(-1.8)	(4.4)	(2.1)			
			IV. $\log F = \log a + bM + c[e^A]$			
1. Aircraft & Missiles						
"1950s"	-.153	.0112	.34E-07	.473	.40	55
	(-1.3)	(4.2)	(1.8)			
"1960s"	.182	-.0001	.16E-06	.145	.62	21
	(2.6)	(-0.1)	(5.3)			
Combined	-.096	.0088	.46E-07	.423	.37	76
	(-1.0)	(4.2)	(3.0)			
2. Aircraft						
"1950s"	-.053	.0113	.36E-07	.428	.40	34
	(-0.4)	(3.2)	(1.5)			
"1960s"	.126	.0026	.14E-06	.132	.72	13
	(1.5)	(1.2)	(4.5)			
Combined	-.027	.0093	.47E-07	.375	.39	47
	(-0.3)	(3.5)	(2.4)			
3. Full sample						
"1960s"	.249	-.0004	.15E-06	.219	.34	25
	(2.5)	(-0.2)	(3.3)			
Combined	-.069	.0088	.42E-07	.423	.35	80
	(-0.7)	(4.2)	(2.8)			

Notes:

The t-statistics are presented in parentheses below coefficient estimates.

In part IV, the "c" coefficients are very small (since e^A is usually very large); the coefficient values are listed in scientific notation. For example, the "1950s" Aircraft and Missiles "c" coefficient is $.34 \times 10^{-7} = .000000034$.

[a]The "1950s" sample includes only aircraft and missiles, so estimates for the "1950s" (1) Aircraft and Missiles sample and (3) Full sample of each part are the same.

Table 4

R^2 FOR VARIOUS TRANSFORMATIONS OF "A" IN LINEAR STRUCTURE

Coverage	log A	A	A^2	e^A	Sample Size
Aircraft and Missiles					
"1950s"	.36	.36	.37	.40	55
"1960s"	.42	.50	.56	.62	21
Combined	.32	.33	.34	.37	76
Aircraft					
"1950s"	.37	.38	.38	.40	34
"1960s"	.55	.61	.65	.72	13
Combined	.36	.37	.37	.39	47
Full Sample					
"1960s"	.34	.39	.42	.34	25
Combined	.31	.32	.32	.35	80

2

Statistical measures of the question of overall improvement between the decades can be handled in two ways: by a test of equality of the coefficients in the equations for the two decades, and by a test of the accuracy of prediction by one decade's equation of the systems represented in the other decade's sample. The former is performed by means of an F test, the latter with Theil's U statistic.[1] Using the aircraft and missiles sample, the hypothesis that there is no change in the coefficients between the two decades is not rejected by the usual statistical criteria of a 5 or 1 percent significance level. However, the hypothesis would be rejected at about the 10 percent level (the F statistic is 2.02 with 3 and 70 degrees of freedom).

The other indication of very little change in cost estimation accuracy or cost control over the last two decades comes from the measure of predictive accuracy of this structure with "1950s" parameter estimates extrapolated to the "1960s" data. The Theil U statistic (which ranges from 0 for perfect prediction of actual values to 1 for maximum inequality between predictions and actuals) is .139. While not terribly low, the predictions are not uniformly biased or asymmetrical in their prediction of the range of actual cost factors for the 1960s.[2]

A comparison of parameter estimates between the various sample ranges reveals that the values are rather insensitive to the coverage of systems. For the "1960s" equation, only the aircraft sample produces a value of "b" for the preferred structure (part IV) that is

---

[1]See the "Statistical Measures" section of the Appendix for brief descriptions.

[2]As discussed in the "Statistical Measures" section of the Appendix, the U statistic is an overall measure of predictive accuracy of an equation. The extent of inaccuracy can be broken down into the proportion of inaccuracy of prediction due to "unequal central tendency" (U^M), "unequal variation" or asymmetry of forecast over the range of actual values (U^S), and "imperfect variations" or lack of high positive correlations between actual and predicted values (U^C). If inaccuracy of prediction is present, the most desirable type is the third (that is, a large proportion of U^C). In this test, $U = .1391$, and the extent of inaccuracy is broken down as $U^M = .0212$, $U^S = .0002$, and $U^C = .9786$.

somewhat significant and positive. The values of "c" remain the same across the samples in this part of the table, but the intercept for the "1960s" increases with the scope of the sample; from the values for the intercept in log form one can transform them to average cost factors (before development program characteristics are considered) of 1.13 for aircraft, 1.20 for aircraft and missiles, and 1.28 for the full sample.

One other structure for the model [equation (12)] is also of interest. In this form, the program length surrogate and technological advance function enter multiplicatively:

$$\log F = \log a + bM \cdot f(A) \quad . \tag{15}$$

The assumed form of equation (7) for this structure is:[1]

$$E = f(A) \ , \tag{16}$$

where the f(A) to be explored will be A and e^A . The results are presented in Table 5.

As in the results for the additive structure, the most rapidly increasing function of A -- the e^A surrogate for E -- provides the least squares fits with highest "R^2s;" the fits are even slightly better than for the additive model. In every case the "b" coefficients are positive and significantly different from zero. The values are slightly lower for the "1960s" than for the "1950s" in the "M·A" formulation, but about twice as high as the latter in the "$M \cdot e^A$." Both structures reveal intercepts higher for the "1960s" than for the "1950s"; for the "$M \cdot e^A$" structure the intercepts are significantly different from zero in $\log_e$ form and therefore imply cost factors greater than unity on the average, even before taking account of the relevant development program characteristics.

The significance tests reveal even less difference between the decades for the multiplicative forms of the "aircraft and missiles"

---

[1]The assumption could be stated slightly more generally as E = cf(A). However, the effect of this would merely be that the coefficient "b" would contain the factor 1/c. If "c" remains the same within each of the two decades, then the comparison between the decades is not affected by this substitution.

Table 5

MULTIPLICATIVE STRUCTURE

Coverage[a]	I. $\log F = \log a + b\, M \cdot A$					II. $\log F = \log a + b\, M \cdot e^A$				
	log a	b	Standard Error	R^2	Sample	log a	b	Standard Error	R^2	Sample
1. Aircraft & Missiles										
"1950s"	-.087 (-0.8)	.00087 (6.2)	.460	.42	55	.159 (2.2)	.15E-08 (6.7)	.446	.46	55
"1960s"	.086 (1.1)	.00046 (2.6)	.196	.26	21	.179 (5.9)	.36E-08 (6.5)	.126	.69	21
Combined	-.058 (-0.7)	.00083 (7.3)	.405	.42	76	.180 (3.5)	.15E-08 (8.0)	.388	.46	76
2. Aircraft										
"1950s"	-.008 (-0.1)	.00096 (5.2)	.399	.46	34	.198 (2.5)	.18E-08 (5.3)	.397	.46	34
"1960s"	.087 (1.2)	.00065 (4.1)	.152	.60	13	.216 (6.3)	.33E-08 (6.8)	.105	.81	13
Combined	.006 (0.1)	.00091 (6.2)	.347	.46	47	.217 (3.8)	.18E-08 (6.4)	.344	.48	47
3. Full sample										
"1960s"	.142 (1.6)	.00044 (2.1)	.241	.16	25	.237 (5.2)	.32E-08 (3.6)	.209	.36	25
Combined	-.029 (-0.4)	.00080 (7.2)	.406	.40	80	.201 (4.0)	.14E-08 (7.8)	.391	.44	80

Notes:

The t statistics are presented in parentheses below coefficient estimates. In part II, the "b" coefficients are very small (since e^A is usually very large); the coefficient values are listed in scientific notations. For example, the "1950s" Aircraft and Missiles "b" coefficient is $.15 \times 10^{-8} = .0000000015$.

[a]The "1950s" sample includes only Aircraft and Missiles, so estimates for the "1950s" (1) Aircraft and Missiles sample and (3) Full sample of each part are the same.

equations than they did for the additive formulation. In no case is the hypothesis of no change in the coefficients between the decades rejected at even the 25 percent significance level.[1] And the Theil U statistic, measuring the predictive accuracy of the "1950s" equation in the "'1960s' aircraft and missiles" sample, has a value of only .10 in each of the two f(A) formulations; thus, the predictive accuracy is reasonably good. Only about 2 percent of the inaccuracy is due to systematic forecast bias in the "M·A" formulation, but 80 percent was attributable to systematic errors in the "$M \cdot e^A$" forecasts.[2]

The "1960s" equations provide considerably less accurate "predictions" of the "1950s" outcomes.[3] The Theil U statistics are .31 for the "M·A" formulation and .63 for the "$M \cdot e^A$."[4]

In sum, none of the structures explored indicates a significant difference between the 1950s and 1960s in the ability of the "system acquisition process" to estimate costs accurately or avoid actual cost overruns for a given development program. Although the results for each decade are insignificantly different in a statistical sense, the implication of some of the coefficient estimates in each decade is that the process itself tends to produce a higher cost factor for a given program in the 1960s than it would have in the earlier years. Despite this "deterioration" of the process, the outcomes for the "1960s" have

---

[1] The test for changes in the coefficients in part I, item 1, leads to an F statistic of .55 with 2 and 72 degrees of freedom; for part II, item 1, the F statistic is .98.

[2] The Theil U statistic for the "$M \cdot e^A$" structure is .1020 (with U^M = .1921, U^S = .6940,and U^C = .1138). For the "M·A" structure the Theil U statistic is .1002 (with U^M = .0117, U^S = .0112, and U^C = .9772).

[3] While extrapolation of an equation backward in time is of less interest in a policymaking sense, it is a useful exercise in determining whether the "system acquisition mechanism" has undergone a net improvement between decades.

[4] The complete set of statistics shows the "$M \cdot e^A$" structure for the "'1960s' aircraft and missiles" equation has a U statistic of .6349 [with U^M = .1464 (predictions <u>exceeding</u> actual cost factors on the average), U^S = .7725, and U^C = .0811]; for the "M·A" structure the U statistic is .3097 [with U^M = .1186 (predictions <u>less than</u> actuals on the average), U^S = .7339, and U^C = .1475].

been no worse than in earlier years because, in part, the level of technological advance attempted has been kept lower on the average. I will take up this aspect of the comparison between the decades in the "Discussion" subsection below.

Two further questions about the equations should be considered. First, how sensitive are the results to the measure of program length that has been used; and second, how sensitive are the results to the weighting of the programs within the sample?

As discussed in Section III, the best available measure of program length for the "1960s" is the number of months from the date of the estimate established near the time of DoD approval of initiation of development to the initial operational delivery. This amounts to using only one estimate for each system in the sample -- the earliest. In Table 6, the same selection of observations from the sample has been used for the "1950s," although these do not necessarily establish the beginning of development; in some cases they are merely preliminary guesswork at an early stage of system conceptualization. Therefore, this form of the data has not been used for comparison of system acquisition experience between the decades. Table 6 presents the estimates using this "earliest M" for each system as a surrogate for L and shows that by comparison with the results in Table 5, there are only quite minor differences: the intercepts for the "1950s" are usually somewhat higher and the "R^2s" lower in the "M·A" formulation, while the corresponding intercepts for the "1950s" are lower in the "$M \cdot e^A$" formulation. The results of the statistical tests are essentially unchanged.[1] Therefore, the results on the whole appear to be reasonably insensitive to the particular choice of a surrogate for development program length.

Table 7 displays the results of estimating the "1960s" multiplicative formulation with each system's observation(s) weighted by the

---

[1]For the "M·A" structure: the F statistic is .55 with 2 and 29 degrees of freedom, and the prediction measures for the "'1950s' aircraft and missiles" equation in the "1960s" data are U = .1094 (with U^M = .2097, U^S = .0002, and U^C = .7901). For the "$M \cdot e^A$" structure: the F statistic is .35, and U = .0814 (with U^M = .0671, U^S = .7402, and U^C = .1927).

Table 6

"M" BASED ON EARLIEST ESTIMATES USED TO APPROXIMATE "L" IN MULTIPLICATIVE STRUCTURE

Coverage[a]	I. $\log F = \log a + b\ M \cdot A$					II. $\log F = \log a + b\ M \cdot e^{A}$				
	log a	b	Standard Error	R^2	Sample	log a	b	Standard Error	R^2	Sample
1. Aircraft & Missiles										
"1950s"	.021 (0.1)	.00078 (3.4)	.521	.39	21	.268 (2.2)	.15E-08 (4.2)	.479	.48	21
"1960s"	.088 (0.8)	.00035 (1.5)	.191	.19	12	.180 (4.5)	.33E-08 (4.0)	.131	.62	12
Combined	-.007 (-0.1)	.00076 (4.5)	.430	.39	33	.240 (3.2)	.15E-08 (5.7)	.387	.51	33
2. Aircraft										
"1950s"	.078 (0.4)	.00081 (2.5)	.498	.32	15	.296 (2.0)	.15E-08 (2.4)	.504	.31	15
"1960s"	.064 (0.6)	.00057 (2.4)	.168	.53	7	.197 (4.5)	.32E-08 (4.6)	.108	.81	7
Combined	.057 (0.4)	.00078 (3.2)	.415	.34	22	.270 (2.8)	.16E-08 (3.2)	.414	.34	22
3. Full sample										
"1960s"	.230 (1.6)	.00022 (0.7)	.273	.03	15	.280 (4.2)	.29E-08 (1.9)	.246	.22	15
Combined	.049 (0.4)	.00072 (4.3)	.431	.35	36	.281 (4.0)	.15E-08 (5.4)	.391	.46	36

Notes:

[a]See notes to Table 5.

Table 7

MULTIPLICATIVE STRUCTURE FOR "1960S" WEIGHTED BY PROGRAM SIZE

Coverage[a]	log a	b	Standard Error	R^2	Sample
		I. log F = log a + bM·A			
1. Aircraft & Missiles	.100 (1.3)	.00061 (4.6)	.172	.53	21
2. Aircraft	.123 (1.5)	.00062 (4.6)	.143	.66	13
3. Full sample	.109 (1.5)	.00060 (4.8)	.177	.50	25
		II. log F = log a + b(M·e^A)			
1. Aircraft & Missiles	.226 (7.4)	.32E-08 (9.5)	.104	.83	21
2. Aircraft	.244 (6.9)	.31E-08 (8.3)	.091	.86	13
3. Full sample	.234 (7.4)	.32E-08 (8.9)	.119	.78	25

Notes:

[a]See notes to Table 5.

size of the program (measured as total dollars actually spent). These results indicate a slight upward shift of the curves from those presented in Table 5 -- the "b" coefficient is higher in the aircraft and missile and full samples and only slightly lower in the aircraft sample. The intercepts are higher for both subsamples but slightly lower for the full sample. Overall, these results are also little different from the main results on the multiplicative structure presented in Table 5.

Let us turn, then, to further interpretations of the results of Tables 3 and 5, and the implications of these results for the evaluation of system acquisition experience.

DISCUSSION

In this subsection I present the results displayed above in briefer graphic form. I shall concentrate on three structures for the model set out in Section III, equation (11); namely, the results for the aircraft and missiles samples for the two decades:

Structure	Aircraft and Missiles	
	"1950s"	"1960s"
(a)	$F = .858e^{.0112M+.034x10^{-6}[e^{A}]}$	$F = 1.20e^{-.0001M+.16x10^{-6}[e^{A}]}$
(b)	$F = .917e^{.00087\ M\cdot A}$	$F = 1.09e^{.00046\ M\cdot A}$
(c)	$F = 1.17e^{.0015x10^{-6}[M\cdot e^{A}]}$	$F = 1.20e^{.0036x10^{-6}[M\cdot e^{A}]}$

Let me first briefly review the conclusions from the statistical results presented above.

The coefficient estimates in the various structures for the model can be described as characterizations of the system acquisition <u>process</u> in each of the two decades. Average values of the development characteristics -- "M" (months between date of estimate and IOD) being used as a surrogate for development program length and "A" being a measure of technological advance sought -- characterize the way in which developments have differed between the decades. The statistical tests discussed above led to the implication that there has been no significant difference in the process between the two decades; one even finds that the process seems to have "deteriorated" for some types of development programs in the sense that for a given set of development program characteristics the cost factor for the "1960s" implied by the model would be higher than for the "1950s."[1] Despite this implication for the process, the programs' outcomes -- which can be characterized as having resulted from programs in the 1960s that were somewhat shorter and of somewhat lower technological difficulty on the average -- have shown a "typical" 1960s program to have a somewhat lower cost factor than a "typical" 1950s program.[2] That is:

---

[1]This is a somewhat different conclusion from that presented by Perry et al., Systems Acquisition Experience, in which we concluded (p. vi):

> The analysis shows that, on average, cost estimates for the 1960s were about 25 percent less optimistic than those for programs of the 1950s. Thus, if reduction in bias (or reduced optimism) is a realistic index of "better," there is evidence of improvement in the acquisition process.

The model on which that analysis and the present one were based has remained essentially the same (although the specific structural specifications for the model have been further refined). The main reason for the changes in the implications is an improved and up-dated data base for the "1960s." All the data used in that earlier analysis have been rechecked and a few errors detected and eliminated. But mainly, the estimate of actual cost of the more recent acquisitions (especially C-5A and F-111) have been updated to December 1969 estimates. The earlier analysis used estimates of about a year earlier (the earlier C-5A estimated actual cost had been made in October 1968; the F-111 had been made March 1969). See the Appendix for additional discussion.

[2]There is nothing contradictory in this statement. In fact, it has a parallel in the observation that although a system may cost more per unit than expected, the total actual expenditure may not differ greatly from the total anticipated -- only the quantity procured may have changed.

The average values of M in the table are based on the full sample of systems; they are therefore probably low estimates of development program length (see Section III). Average A is computed with each system weighted equally.

Coverage	Average Value M	Average Value A	"Typical" Cost Factor for Structure (a)	(b)	(c)
"1950s"	42.8	12.2	1.40	1.44	1.19
"1960s"	37.8	8.9	1.20	1.27	1.20

"Typical" cost factors for structures (a) and (c) could be derived equally plausibly from the average of the program effort measure (E, which is approximated by functions that include e^A). The results of this exercise show a somewhat lower "typical" 1960s cost factor for structure (c) also:

Coverage	Average Value M	Average Value e^A	"Typical" Cost Factor for Structure (a)	(c)
"1950s"	42.8	$e^{14.6}$	1.49	1.35
"1960s"	37.8	$e^{12.3}$	1.24	1.23

For consistency of presentation, the graphic comparison of the decades by each of the three structures (Fig. 9) will make use of the average "A" results.

To gain an appreciation for the implications of these results consider Fig. 8. It is most convenient, in presenting a figure in two dimensions, to consider first the equation for each of the three structures of the model that was obtained from the aircraft and missiles sample for the two decades combined. Thus, each part of Fig. 8 presents a single equation plotted for 4 different values of A. For structures (a) and (c) -- those with the transformation of technological advance sought (A) to effort per month of development (E) involving e^A -- the increases in A have little effect until A gets to the largest end of the sample range. Since these are the best structures (in the statistical sense of having the largest "R^2s"), the implication is that seeking very large technological advancement is the main reason for serious underestimation or actual cost

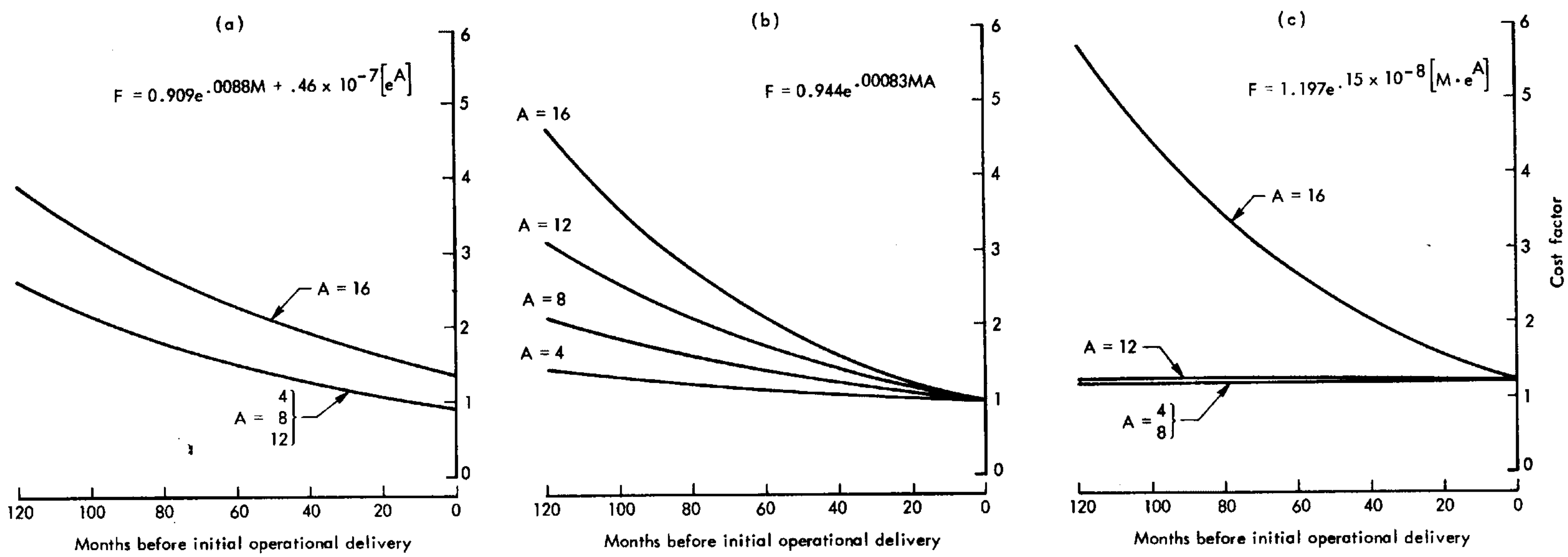


Fig. 8—Influence of technological advance on cost factors

overruns.[1] Since A enters directly in structure (b), the curves for equal changes in A rise uniformly.

Let us move on, then, to Fig. 9. In this figure as in the previous one, a horizontal line at the level of unity (cost factor = 1) would be interpreted as a perfect cost estimation or as perfect cost control at the level of the cost estimate for any length of time prior to operational status of the system. The feature that has been observed and that this analysis has been directed toward is that cost estimation or inefficiency typically leads to actual cost exceeding estimated cost. The extent of bias in each of the two decades is conveyed by the height of the curve above the line at the unity level.

The band around each regression line corresponds to one standard error of estimate.[2] The shading range along the horizontal axis indicates each sample's range of times of estimates (used in calculating M). The broken lines beyond the shading indicate extrapolations from these sample ranges. The "1950s" curves are plotted for the average value of A for that decade (12.2), while the "1960s" curves are at the level of that decade's average value of A (8.9). The graphical interpretation of the statistical insignificance of the differences between the equations and of the high predictive ability of the "1950s" equations for the "1960s" data is that the curves for the "1960s" falls almost entirely within these one standard error bands over the entire range of "1960s" values of M. The bands themselves indicate a narrowing of the dispersion aspect of accuracy in predicting costs; the "1960s" equation has a narrower one standard error band for each structure than does the "1950s" equation.

---

[1] However, the results may merely indicate that the survey measure used for A is not very sensitive to lower levels of technological advance, but can only distinguish these from the very highest. Resolution of this question of the quality of the measure versus a real implication about our past experience must await more refined measurement of A.

[2] See the "Predicting Cost Growth and Uncertainty" section of the Appendix for a discussion of the method used to plot these curves from the equations in Tables 3 and 5.

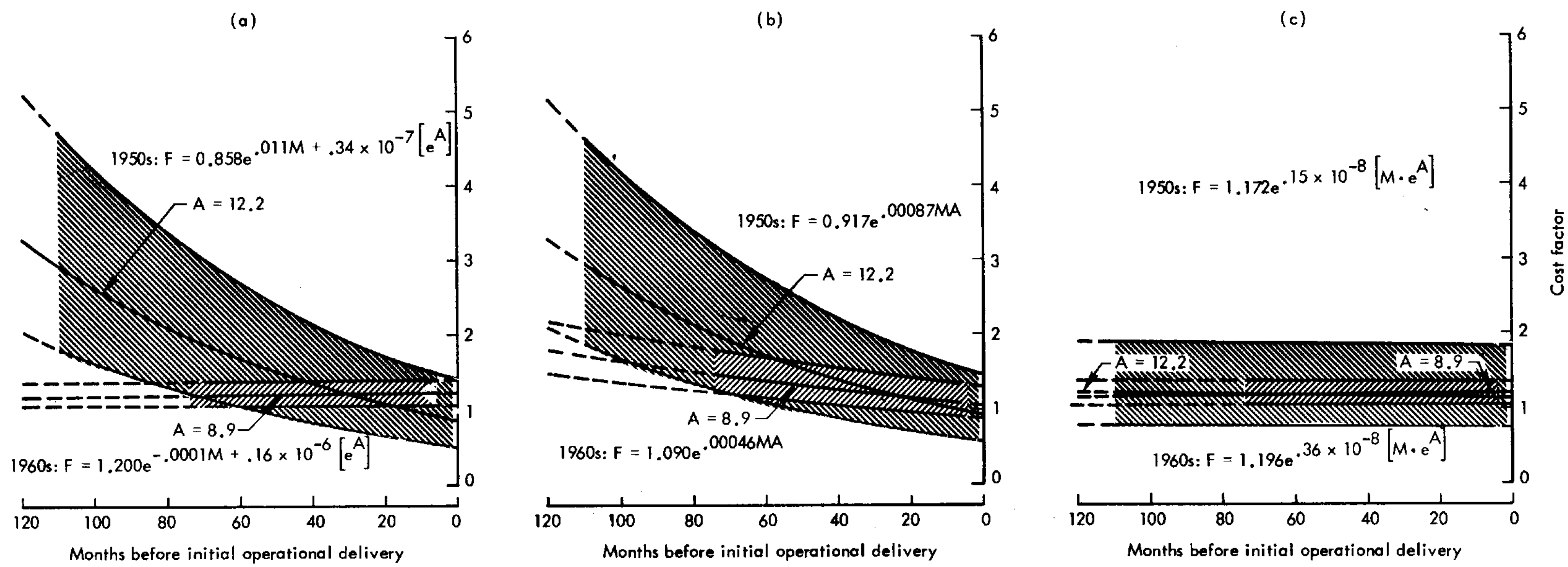


Fig. 9—System acquisition process in 1950s and 1960s

Structure (a) reveals that when the separate effects of program length and technological advance on cost factors are measured, only the "1950s" experience seems to have been sensitive to program length. In this part of Fig. 9 the height of the "1950s" curve for a 42.8 month development program is 1.40, the "typical" cost factor for structure (a) presented above. Moving along the "1950s" curve, the height at 37.8 months of the same technological advance level is 1.32. The shift to the "1960s" curve and the lower average level of A leads us to the height of the "typical" cost factor for the "1960s" from structure (a), 1.20. Thus, the statistical result that the system acquisition process is essentially unimproved over the two decades is compatible with the result that cost factors of (different) "typical" programs have declined for the last decade from the preceding.

V. HYPOTHETICAL PROCUREMENT

One of the limitations of cost factors mentioned in the introduction is that they give no indication of the magnitude of program costs. Let us make such a translation by considering the acquisition of a new weapon system -- a fighter aircraft. We can also consider the following question: Based on past experience, how can we anticipate the extent to which the actual cost of the proposed fighter will differ from the estimate? That is, how can the bias of the cost estimate or the inefficiency of the procurement process be quantified to yield a cost figure that is likely to be close to the actual cost? The results of Section IV could be used to obtain such an approximation of actual cost. The following example, however, illustrates the inherent uncertainty of such a process.

To apply the empirical results, we must assume that experience in the future will be similar to what we have observed in the 1960s; then the equations estimated for the "1960s" aircraft sample presented in Tables 3 and 5 can be used to predict the possible bias of estimate or cost overrun (see Fig. 10).[1] Suppose the new fighter is subjectively assessed to require a technological advancement of 12 by comparison with systems in the "1960s" sample.

Then, if the predicted cost of this new weapon system is one billion dollars and it is expected to be operational in 42 months,[2] the best point estimate from structure (a) is that the billion-dollar program would actually cost an additional \$290 million (F = 1.29).[3] But the range of possible cost variation is at least plus or minus

---

[1]See the "Predicting Cost Growth and Uncertainty" section of the Appendix for a detailed discussion of the use of these equations in deriving F values.

[2]This is the average development program length for systems in the "1960s" sample, based on the lengths of time from DoD approval of development of each system to operational status.

[3]In this discussion, I explicitly discuss only the results in part (a) of Fig. 10. The other two structures lead to best point estimates of additional cost of (b) 520 million and (c) 270 million.

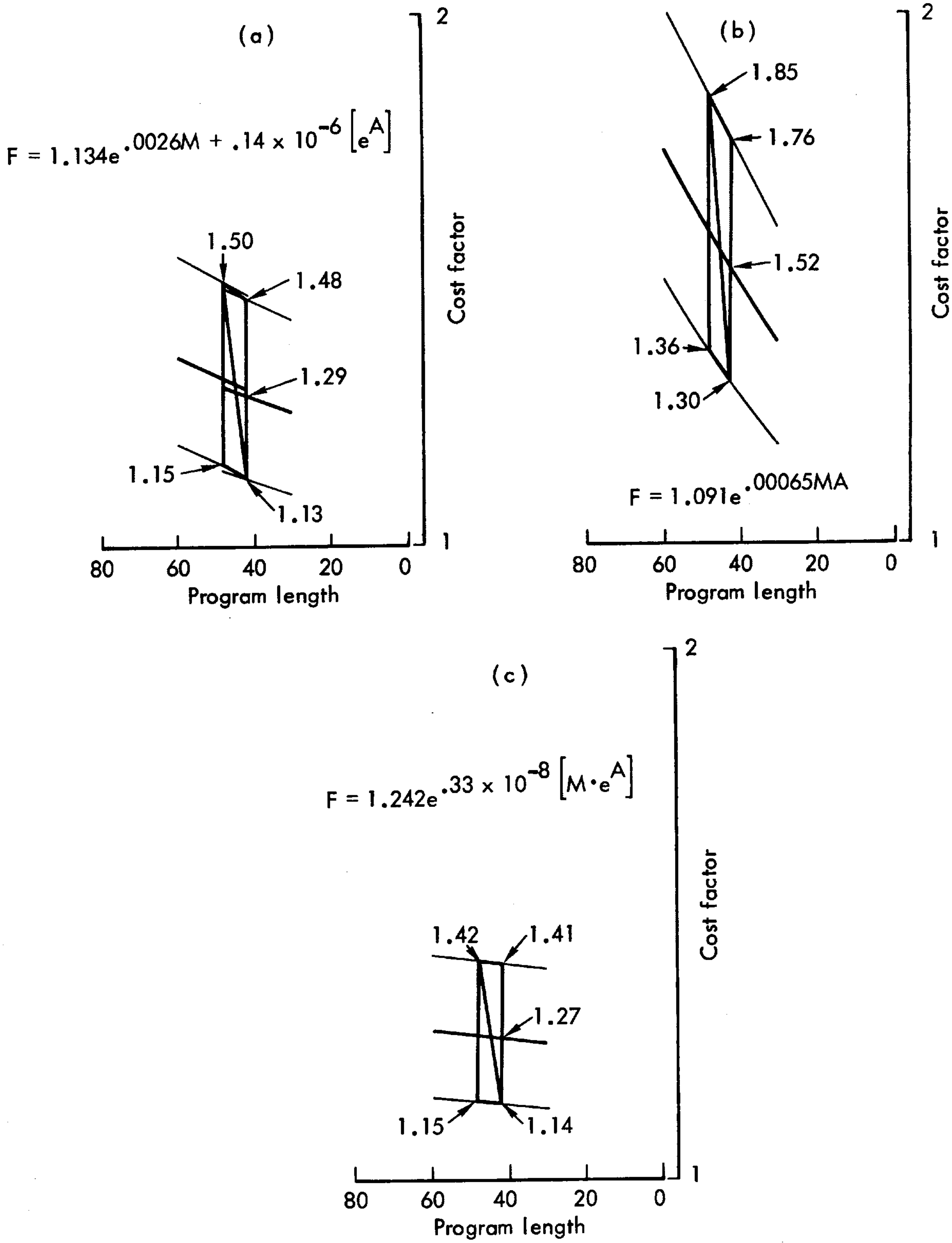


Fig. 10—Example of prediction uncertainty

one standard error of this "'1960s' aircraft" equation. That gives us a range (represented by the right-hand side of the box in Fig. 10) from \$130 million more than the predicted cost (F = 1.13) to \$480 million more than predicted (F = 1.48).

Furthermore, the fighter may not become operational in the length of time predicted; it might actually remain in development for an extra 6 months. What needs to be emphasized here is that at the time a program is under consideration, we are no more certain of the exact schedule within which it will come to fruition than we are of the exact cost. In the past, both of these have generally been different from expectations. The range of uncertainty then, assuming a schedule slippage of 6 months, is portrayed by the left-hand side of the box. At this point (M = 48) the curve itself is higher and the range of uncertainty is greater -- ranging from an additional \$150 million (F = 1.15) to an additional \$500 million (F = 1.50).[1]

Thus, when inaccuracy in the predictions or inefficiencies in the execution of both cost and schedule aspects of programs are considered, the overall range of uncertainty is represented by the diagonal of the box in part (a) -- 130 million to \$500 million more than the original

---

[1]The additive structure (a) differs in an important respect from structures (b) and (c). The model [in equation (11), Section III] postulates that cost factors are related to development program length (L) and intensity of effort (E). All structures use M (months from date of estimate to initial operational delivery) as a surrogate for L. Structures (b) and (c) use a function of A as a surrogate for E [see equation (16), Section IV]. But structure (a) uses an expression involving A <u>and</u> L as a surrogate for E [see equation (14), Section IV], Therefore, if the left-hand side of the box in part (a) of Fig. 10 were drawn on the same curves as the right-hand side, the drawing would merely depict a schedule stretch-out (lengthened L but reduced E), not a schedule slippage (lengthened L and E held the same) as in parts (b) and (c). Therefore, the appropriate left-hand side of the box is the range of a new curve with A increased sufficiently for the result of this new A and program length of 48 in equation (14) to be equal to the value of the old A of 12 and program length of 42 in that equation. The required increase in A for this adjustment was .134 which produced the slightly higher set of curves in part (a).

prediction.[1] If, in addition, the scope of the program were increased (for example, a fighter with even greater range was requested by the user's command), the A would be larger in fact and the whole box would shift upward.

Finally, the confidence that we have in this range of estimates depends on the extent to which we believe that the influences affecting the development of this new system are suitably reflected in the rather small sample of the 1960s systems analyzed above. Even if this new program is sufficiently similar that we can have some confidence in this rough approximation of the extent of our uncertainty about costs of the new system, the uncertainty is rather large -- on the order of one-third to one-half billion dollars for a billion-dollar program.

---

[1]Without carrying through the same discussion for parts (b) and (c) of Fig. 10, it should be clear that the comparable range of uncertainty for the multiplicative structure using e^A is (c): \$140 million to \$420 million more than the prediction. The multiplicative structure with simple A entered (as a surrogate for E) produces a somewhat higher range (b): \$300 million to \$850 million more than predicted.

Appendix

DATA BASE AND SUPPLEMENTARY PROCEDURES

DATA BASE

The two sets of data are labeled "1950s" and "1960s" since the weapon systems in each set are predominantly from these periods. However, the actual range of sample observations (from earliest date of estimate to latest IOD) for the "1950s" is February 1947 to July 1961, and for the "1960s" January 1958 to December 1969.

Sources of Data

The cost factors and months between the dates of estimate and initial operational delivery are presented in Table 8. The data base for the "1950s" was compiled by Eugene Brussell of Rand and has been used in several past Rand studies. This sample includes only Air Force systems. The cost factors are those calculated and analyzed by Robert Summers.[1] They are based on unit costs of production items (for the "1960s" they are based on total cost, which includes development and production). However, the development portion of the total cost of systems of the 1950s was relatively small except for the ICBMs so that the effect of this difference in the data is minimal.

The "1960s" data base originally consisted of the questionnaire returns for a survey of 21 systems selected by DDR&E with the advice of the services.[2] This sample includes Army, Navy, and Air Force programs. Originally, sufficient data to compute cost factors were available for 12 of these systems. Additional data have become

---

[1] Summers, Cost Estimates as Predictors.... Following Summers, three of the "1950s" observations were not used in the analysis because these extremely early estimates were for systems drastically different from the systems ultimately procured. These data are enclosed in brackets in the listing.

[2] The survey methodology and data base are described in Perry et al., System Acquisition Experience, Sections I and II. The survey results were compiled by Dominic DiSalvo and James Stucker, who also contributed to the additional refinement of the sample undertaken by Susan Henrichsen and reported below.

Table 8

DEFLATED COST FACTORS AND MONTHS TO FIRST DELIVERY
("1950s" and "1960s")

System	Factor	Months	System	Factor	Months	System	Factor	Months
			I. The "1950s" Air Force Systems					
F-102/106	1.30	11	B-58	1.95	27	IRBM Thor	1.33	32
	2.35	30		3.64	65			
	1.30	30		5.10	74	Snark	1.53	58
	1.49	45		4.00	84		1.17	60
	2.46	77					3.10	96
	4.06	90	B-52	.87	2		[5.60	127]
				1.46	6		[6.40	149]
F-101	1.12	24		.97	14			
	.57	54		1.44	44	ICBM Titan	1.00	46
				1.33	44			
F-100	1.20	27		2.62	64	ICBM Atlas	.88	40
							.85	47
F-94C	2.56	14	B-47	1.55	16		1.32	57
				.70	22		.82	69
F-89	1.52	6		1.25	34		.77	78
	2.04	37						
			C-133	1.55	13	Falcon	.51	27
F-86D	.78	23					2.09	45
			C-130A	1.49	21		3.80	59
F-86A	.93	20					2.52	68
			KC-135	.80	17		[2.76	95]
F-84F	2.02	13		.92	26			
				.81	37	Bomarc	1.00	26
F-84C	1.55	15					1.16	26
							1.49	30
							2.22	62
							3.48	92
							5.90	100
							7.10	110

Table 8 (continued)

System	Factor	Months	System	Factor	Months	System	Factor	Months
			II. The "1960s" Systems					
AIR FORCE			NAVY			ARMY		
F-111	2.07	62	A-7E	1.40	15	LOH (Hughes)	1.09	60
	1.95	56						
	2.02	42	OV-10A	1.10	34	Pershing I	1.12	63
	1.41	18					1.01	40
			SQS-26AX	2.34	26			
C-5A	1.36	63				Pershing IA	1.07	57
	1.38	51	SQS-26CX	1.55	39		1.03	42
	1.09	15						
			Difar	2.05	39			
C-141	1.16	29		1.04	18			
	1.41	25						
A-7D	1.23	20						
Minuteman II Airborne Command Post	1.12	15						
	1.28	6						
Minuteman II Guidance and Control	1.60	36						
Titan III-C	1.06	75						

available since the completion of the survey in the summer of 1969; the original data base has now been revised, up-dated, and expanded to include a total of 25 observations for 15 systems. The revision has included a review of the cost data available for all systems in the sample. Several sources were used to update and verify the survey data and to obtain new information from which additional cost factors could be calculated. The sample was also expanded by the inclusion of two additional systems, A-7D and A-7E. The SQS-26 program was separated into two systems, SQS-26AX and SQS-26CX.[1] "Actual" costs for six of the systems (F-111, C-5A, A-7D, A-7E, Difar, and SQS-26CX) are in reality the most recent projections of what the actual costs will be; these programs are all still in production and actual costs are therefore not yet known with great accuracy.

Of the total amount of program costs represented by this sample, 80.3 percent is for Air Force systems, 12.2 percent for Navy systems, and 7.5 percent for Army systems. Approximately half of the sample (8 systems) accounts for 93.6 percent of the total cost, while two very large programs (C-5A and F-111) account for 52.7 percent of the total.

The figures listed in the "Months" column of Table 8 give the number of months from the date of the estimate to IOD. In some cases, there were two dates associated with one estimate: the date the estimate was "established" and the date of the document in which the estimate appeared. Whenever this occurred, the "date established" was used. It was possible to determine IOD dates rather firmly for all systems but one. No IOD date was given for SQS-26AX, and it was necessary to approximate this date from other information. The figure given in Table 8 is known to be correct within four months, more or less.

---

[1] The total SQS-26 program extends over a very long period of time (more than 10 years) and has undergone substantial change. In addition, the AX version and the CX version were estimated, contracted for, and procured separately. The SQS-26CX is currently being reported by the Navy as an individual weapon system.

Some uncertainty exists for many of the cost factors used in the analysis, from difficulties in making the necessary quantity and price-level adjustments.

Quantity Adjustment

Whenever quantity changes occur during a program, adjustment must be made to reconcile the estimate with the actual quantity produced. Adjusting the estimated cost so that it is for the same quantity as the actual cost is obviously preferable to calculating a cost factor as a ratio of "total" costs for two different quantities. Such a cost factor certainly would not accurately reflect the closeness of actual to estimated cost. The accuracy of assessment of the cost-quantity relationship for a system is a part of the overall accuracy of cost estimation. Therefore, the ideal method of quantity adjustment is to base quantity-adjusted cost factors on estimates for the actual quantity delivered that were generated by the original estimator and used by decisionmakers. Such a cost factor would be a useful ratio of the actual to the estimated cost for evaluating estimation accuracy. If, however, no information is available to indicate what the estimator assumed the cost-quantity relationship to be, any adjustment introduces some uncertainty into the cost factor since the estimated cost used to calculate the factor has been adjusted by someone other than the estimator.

For several "1960s" systems (A-7D, A-7E, Pershing IA, Minuteman II - Airborne Command Post, and Difar), quantity-adjusted estimates were available; these were used to calculate the cost factors. For another system, C-141, the quantity change was small enough that the choice of learning curve makes little difference in the adjusted cost factor. Two systems, C-5A and F-111, have undergone rather substantial quantity changes, and it was not possible to determine with any certainty what the relevant learning curves are. The estimates for both systems were adjusted by use of an 85 percent learning curve.[1] Because the

---

[1] I am indebted to Gilbert Levenson for his advice and assistance on the necessary qunatity adjustments.

adjustments were rather large, considerable uncertainty exists for these cost factors.

For the "1950s" systems, a review of Summers' underlying documentation for the data reveals that quantity adjustments were made for approximately two-thirds of the cost factors. Some uncertainty therefore exists for many of these factors, because of the difficulty of determining the relevant learning curve.

Price Adjustment

Some uncertainty exists for all cost factors owing to price-level adjustments. For purposes of this study, it seemed necessary to make such adjustments for several reasons. First, it is assumed that cost estimation accuracy does not include the ability to predict price-level changes accurately. A comparison of the acquisition experience of the 1950s and the 1960s would be biased by a failure to adjust for possible differences in the rate of inflation for the two decades, over which the Defense Department has very little control.

Cost factors could be deflated accurately and with little uncertainty only if a number of conditions were met. First, we would have to know what allowance for inflation, if any, has been included in the estimate so that the actual cost could be adjusted by the difference between the allowance made for inflation and the inflation that actually occurred. We would also have to know the timing of program expenditure, and we would need a price index relevant to each program in the sample. At least this much information would be needed to adjust a cost factor by the difference between the estimated and the actual rate of price increase. Since such information was not available for any of the systems, it was necessary to devise some other method of deflating the cost factors.

Several methods were considered and regression results were obtained for each method in order to ascertain the sensitivity of the results to the method of deflation. These methods and the results will be described below. First, regression results were also obtained for the "1950s" and the "1960s" using undeflated cost factors to determine to what extent the magnitude and fit of the estimated coefficients depend on deflating cost factors. The results for the two multiplicative structures for the aircraft and missiles sample (using undeflated cost factors) are shown below. The t statistics appear in parentheses below the coefficient values.

	$\log F = \log a + bMA$		$\log F = \log a + bM \cdot e^{A}$	
Coverage	log a	b	log a	b
"1950s"	.029 (0.2)	.00095 (6.0)	.308 (3.6)	.16E-08 (6.1)
"1960s"	.110 (1.4)	.00056 (3.2)	.234 (7.6)	.38E-08 (6.7)

Comparison of these results with those shown in Table 5 (where deflated cost factors were used) reveals that the fits do not vary a great deal. This indicates that the explanatory power of the model is not due merely to the use of deflated cost factors.

The "1950s" cost factors were deflated by Summers; his method is not reported, but the resulting factors have been taken as given. For the "1960s" factors, two different methods and two different price indexes were considered. The first method (Method 1) was to multiply the undeflated factor by a figure equal to the price index for the year in which the estimate was made divided by the price index for the year in which IOD occurred. In other words, $F_{deflated} = F_{undeflated} \times (P_{year\ of\ estimate}/P_{year\ of\ IOD})$. This method would be correct only if both program expenditure and price increase were at a constant rate throughout the program and if the time from the date of the estimate to IOD constituted half the full program length. In fact, this deflation is probably conservative both because half the spending probably has not been completed by IOD and because prices have generally increased at an increasing rate in the 1960s.

The second method (Method 2) was an attempt to deflate the "1960s" cost factors by the difference between an approximation of the amount of price increase assumed by the estimator and the actual amount of price increase which occurred from the date of estimate to IOD. No information was available on the amount of price increase allowed for in the estimate. It was therefore assumed that the estimate included an allowance for a constant annual rate of price increase equal to the average annual rate of increase for the three years prior to the year in which the estimate was made. This resulting net price increase was obtained by subtracting the "estimator's assumed rate" from the actual average rate of price increase and deflating the cost factor over the period from the date of estimate to IOD.[1] In other words, $F_{deflated} = F_{undeflated} \times [1/(1 + \text{net price increase})^{(\text{year of IOD} - \text{year of estimate})}]$. Since this method assumes that some allowance for inflation was included in the estimate when in fact there may have been none, this method is probably more conservative than Method 1, described above.

Both Method 1 and Method 2 were used with two price indexes,[2] resulting in four sets of deflated cost factors. Using these four sets of data, the following results were obtained:

Deflated by		"1960s" Aircraft and Missiles Sample			
	Price	log F = log a + bMA		$\log F = \log a + bM \cdot e^A$	
Method	Index	log a	b	log a	b
1	1	.086 (1.1)	.00046 (2.6)	.179 (5.9)	.36E-08 (6.5)
2	1	.105 (1.4)	.00046 (2.7)	.198 (6.7)	.35E-08 (6.6)
1	2	.056 (0.6)	.00023 (1.1)	.068 (1.8)	.33E-08 (4.8)
2	2	.098 (1.3)	.00052 (3.1)	.209 (7.6)	.36E-08 (7.4)

[1] If this method produced a "net price increase" that was negative, no further deflation was made (that is, the "net price increase" was set equal to zero).

[2] Price index 1 is the Wholesale Price Index for Machinery and Equipment available in the 1970 Annual Report of the Council of Economic Advisors. Price Index 2 is an unpublished aircraft price index (excluding avionics and engines), developed by Harry G. Campbell.

The comparison shown above reveals that the regression results are rather insensitive both to the method of deflation and the price index used. The results are nearly the same using both methods with Price Index 1. When Method 2 is used, the results for both price indexes are similar. The only case in which the coefficient values differ very much is that of using Method 1 with Price Index 2.

All of the results reported in this study used Method 1 and Price Index 1 (the Wholesale Price Index for Machinery and Equipment), to deflate the cost factors of the "1960s." This index was chosen because the "1960s" sample represents a number of widely differing systems and subsystems, so that this rather general index might be more applicable than would a more specialized one. Also, it leads to relatively "conservative" deflation in that the average rate of price increase is lower for it than for Price Index 2. In any case, the choice of method and price index for deflation of the cost factors appears to have a minimal effect on the results.

Technological Advance Sought

Table 9 displays the measures of "technological advance sought" (A) and the contractor for each of the systems in the samples. The values for A were obtained by rather limited surveys conducted within Rand (see the discussion in Section III). These measures are, at best, crude approximations of the concept of the relative level of technological advance sought and should be considered preliminary measures for this variable.

STATISTICAL MEASURES

Standard significance tests are not really appropriate to the kinds of data samples used in this study. As discussed above, the criterion for inclusion of a system in the sample was the availability of data: through past Rand research in the case of the "1950s" sample, and through selection of systems by DDR&E with substitutions made by the individual services for some of the original requests for the "1960s" sample. Therefore, neither of them is a random sample of all

Table 9

WEAPON SYSTEMS AND "A" LISTED BY CONTRACTOR[a]

	Boeing		General Dynamics		Douglas		Lockheed		McDonnell		North American		Northrop		Republic		Martin Marietta		Hughes		Other[b]	
	System	A	System	A	System	A	System	A	System	A	System	A	System	A	System	A	System	A	System	A	System	A
"1950s"	B-47	15.9	F-102/106	15.2	C-133	7.8	F-94C	8.5	F-101	8.9	F-86A	14.3	F-89	12.1	F-84C	8.5	ICBM Titan	12.0	Falcon	13.9		
	B-52	12.2	B-58	16.0	IRBM Thor	12.9	C-130A	7.0			F-86D	14.0			F-84F	10.0						
	KC-135	8.3	ICBM Atlas	14.0							F-100	12.6										
	Bomarc	16.0									Snark	16.0										
Average A: 12.2		13.1		15.1		10.4		7.8		8.9		14.2		12.1		9.3		12.0		13.9		
"1960s"	MMII-ACP	8.0	F-111	14.8			C-141	5.5			MMII-G&C	11.5					Pershing I	10.0	LOH	6.7	SQS-26AX	13.0
							C-5A	8.3									Titan III-C	8.5			SQS-26CX	10.0
											OV-10	9.8					Pershing IA	7.0			Difar	7.0
																					A-7E	7.0
																					A-7D	7.0
Average A: 8.9		8.0		14.8				6.9				10.7						8.5		6.7		8.8

Notes:

[a]The systems listed under each contractor are in chronological order, according to the times at which they became operational.

[b]The contractors for these systems are as follows:
SQS - 26AX - General Electric
SQS - 26CX - General Electric
Difar - Magnavox
A-7E - LTV
A-7D - LTV

acquisitions undertaken during a decade. Still, significance tests are commonly used to provide "feel" for the extent to which the implications of an analysis warrant acceptance. For this purpose, the following outline of the statistical measures reported for the additive and multiplicative structures for the model may be helpful.

The first indication of significance is provided by the value of the t statistic reported beneath each coefficient in the tables. For the larger sample sizes (over 30), the coefficients are deemed different from zero at the 10 percent significance level if $t \geq 1.31$, at the 5 percent level if $t \geq 1.70$, and at the 1 percent level if $t \geq 2.46$. For the smallest sample (7 observations in Table 6 and thus 5 degrees of freedom), "significant" t values must exceed 1.48 (10 percent), 2.02 (5 percent), and 3.36 (1 percent).

The test of significantly different equations between the decades is accomplished by a test based on the F distribution.[1] The multiplicative structure for the model, for example, can be written as

$$F = a_{50s} e^{b_{50s} M \cdot f(A)} , \qquad (17)$$

$$F = a_{60s} e^{b_{60s} M \cdot f(A)} , \qquad (18)$$

and for the estimates from an equation for the two decades combined, as

$$F = a_c e^{b_c M \cdot f(A)} . \qquad (19)$$

The test for the entire equation, then, is based on the hypothesis:

$$a_{50s} = a_{60s} \quad \text{and} \quad b_{50s} = b_{60s} ,$$

[in which case equation (19) is sufficient to describe the past two decades], with formulations (17) and (18) as the alternative. For

[1]On such statistical tests, see, for example, J. Johnston, Econometric Methods, McGraw-Hill Book Company, New York, 1963, Sections 4-3 and 4-4.

the aircraft and missiles sample, the significance levels are Fs of slightly less then 2.18 [10 percent level for 3 and 60 (instead of 70) degrees of freedom], and 2.35 [10 percent level for 2 and 60 (instead of 72) degrees of freedom], which are almost never attained for the aircraft and missiles results that have been investigated.

The final statistic reported in Section IV is the Theil U statistic, which Theil calls an "inequality coefficient."[1] It is calculated from the actual values of the cost factors (F) over the relevant sample, and the predicted values (F_p) from the equation under investigation; namely:

$$U = \frac{\sqrt{\frac{1}{n} \sum (F_p - F)^2}}{\sqrt{\frac{1}{n} \sum F_p^2} + \sqrt{\frac{1}{n} \sum F^2}} \tag{20}$$

Theil describes this statistic as follows (A's for actuals, P's for predicted):[2]

> The coefficient U is -- except for the trivial case where all P's and A's are zero, when it is indeterminate -- confined to the closed interval between zero and unity.... We have U = 0 in the case of equality: $P_i = A_i$ for all i. This is clearly the case of perfect forecasts. We have U = 1 (the "maximum of inequality") if there is either a negative proportionality, or if one of the variables is identically zero: $rP_i + sA_i = 0$ for all i and for some nonnegative r and s (not both zero). In other words, U = 1 if there is a non-positive proportionality between the P's and the A's. This is indeed a case of very bad forecasting, for it means either that always zero predictions of non-zero actual values are made (s = 0), or that non-zero predictions are made of actual values which are always zero and hence easy to predict (r = 0), or that predictions are positive (negative) if actual outcomes are negative (positive) in a remarkably regular manner (r and s $\neq$ 0).

The three types of predictive inaccuracy are simply proportions of the numerator of U related to the difference between the means $\bar{F}_p$ and

---

[1]See H. Theil, Economic Forecasts and Policy, Second Revised Edition, Amsterdam, North-Holland Publishing Co., 1961 (Section 2.5).

[2]*Ibid.*, pp. 32-33.

$\bar{F}$, the difference between the standard deviations s_{F_p} and s_F, and a residual component. Thus, the "unequal central tendency" component is

$$U^M = \frac{(\bar{F}_p - \bar{F})^2}{\frac{1}{n}\sum (F_p - F)^2} \quad ; \qquad (21)$$

the "unequal variation" or asymmetry of forecast component is

$$U^S = \frac{(s_{F_p} - s_F)^2}{\frac{1}{n}\sum (F_p - F)^2} \quad ; \qquad (22)$$

and the "imperfect covariation" or residual component is

$$U^C = \frac{2(1 - r)s_{F_p} s_F}{\frac{1}{n}\sum (F_p - F)^2} \qquad (23)$$

where r is the correlation coefficient between the actual and predicted values.[1]

A METHODOLOGY FOR EVALUATING CONTRACTORS?

Table 10 displays the results for a variation of the multiplicative structure that might give us some insight into the relative "performances" of some of the contractors in the defense industry. By knowing which contractor was responsible for which weapon system, one could hope to distinguish the "b" coefficient for each corporation.[2] The individual corporations shall remain nameless since the results of the table do not necessarily indicate that a corporation is more or less efficient in developing weapon systems. Although this methodology is a step toward evaluating contractors, the data are insufficient to the task of distinguishing relative corporate performance.

---

[1] For further details, see *ibid.*, section 2.5.2.

[2] The method of determining which corporations should be included in this analysis was based solely on the availability of the data and on the premise that a corporation could be included only if it had at least one system in the sample in each of the two decades.

Table 10

ANALYSIS OF CORPORATIONS USING MULTIPLICATIVE STRUCTURE AND AIRCRAFT AND MISSILES SAMPLES

Coverage	log a	b_1	b_2	b_3	b_4	b_5	b_6	b_7	Standard Error	R^2	Sample
I. "M·A" Structure											
"1950s"	-.154 (-1.3)	.00107 (6.0)	.00088 (5.0)	.00591 (2.3)	.00066 (2.7)	.00028 (0.3)	.00124 (3.6)	.00103 (1.6)	.450	.51	55
"1960s"	.155 (2.3)	-.00006 (-0.1)	.00068 (5.6)	.00033 (1.4)	.00038 (1.4)	-.00018 (-1.0)	-.00016 (-0.5)	.00087 (1.0)	.111	.84	21
Combined	-.040 (-0.5)	.00096 (6.5)	.00079 (5.6)	.00112 (1.9)	.00056 (2.7)	.00018 (0.5)	.00104 (3.5)	.00078 (1.4)	.400	.48	76
II. "$M \cdot e^A$" Structure											
"1950s"	.027 (0.4)	.17E-8 (6.7)	.23E-8 (6.8)	.14E-4 (2.5)	.94E-9 (2.6)	-.36E-8 (-0.1)	.13E-7 (3.5)	.40E-7 (1.4)	.386	.64	55
"1960s"	.132 (3.2)	.47E-6 (0.2)	.39E-8 (7.8)	.76E-6 (2.1)	.90E-7 (2.9)	-.56E-7 (-0.8)	-.86E-6 (-0.4)	.67E-5 (1.6)	.103	.86	21
Combined	.106 (2.0)	.16E-8 (7.1)	.22E-8 (7.5)	.14E-5 (1.3)	.82E-9 (2.6)	-.15E-7 (-0.3)	.11E-7 (3.6)	.32E-7 (1.3)	.351	.60	76

Note:

The results shown here are for the equation: $\log F = \log a + b_1 M_1 \cdot f(A_1) + \ldots + b_7 M_7 \cdot f(A_7)$. In other words, using the multiplicative structure, a separate "b" coefficient was estimated for each corporation shown above, with that corporation's systems' program lengths and level of technological advance used for M_i and A_i. The first 6 numbers stand for six different corporations; the 7th includes the data for all others that could not be separately estimated. The numbers shown in parentheses below the coefficient estimates are t statistics.

The estimate of the "b" coefficient merely describes the degree to which the contractor's cost factors deviate from unity; it does not describe why the factors differ. Recall a point made at the beginning: that a cost factor cannot distinguish between contractor inefficiency, which might lead to too high an actual cost, and an inaccurate cost estimate made early in the program. The results appearing in Table 10 might occur either because cost estimates for systems developed by certain contractors are uniformly more optimistic than for others, or because of contractor inefficiency.

One would thus want to incorporate a system for evaluating cost estimates before he could distinguish contractor efficiency. But even then, the contractor could be evaluated only in terms of cost performance, given the technological advance required. One would want to include total system performance in a full evaluation of a contractor.

PREDICTING COST GROWTH AND UNCERTAINTY

In Section V, various structures of the model were used to predict the probable cost growth and range of uncertainty for a hypothetical procurement. This is done by specifying values for the independent [M and f(A)] and calculating the values of F from

$$F = e^{[\log a + bM + cf(A)]} \tag{24}$$

for the additive structure, and from

$$F = e^{[\log a + bM \cdot f(A)]} \tag{25}$$

for the multiplicative structure. The range of uncertainty considered in Section V was plus or minus one standard error (s.e.). This is calculated by solving

$$F \pm s.e. = e^{[\log a + bM \cdot f(A) \pm s.e.]} \tag{26}$$

for the multiplicative structure, and similarly for the additive structure. The coefficients and the standard error for each equation are given in the tables in Section IV.