

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: The Rate and Direction of Inventive Activity: Economic and Social Factors

Volume Author/Editor: Universities-National Bureau Committee for Economic Research, Committee on Economic Growth of the Social Science Research Council

Volume Publisher: Princeton University Press

Volume ISBN: 0-87014-304-2

Volume URL: <http://www.nber.org/books/univ62-1>

Publication Date: 1962

Chapter Title: Predictability of the Costs, Time, and Success of Development

Chapter Author: A W. Marshall, W H. Meckling

Chapter URL: <http://www.nber.org/chapters/c2138>

Chapter pages in book: (p. 461 - 476)

Predictability of the Costs, Time, and Success of Development

A. W. MARSHALL AND W. H. MECKLING

THE RAND CORPORATION

NOWADAYS, tales of how important discoveries or inventions have been made by accident or in contradiction to the best scientific evaluation are familiar to almost everyone. And the proposition that development of new products or processes is an uncertain business is generally accepted without serious dispute. But when it comes to translating this generality into uncertainties associated with particular current development proposals, or when it comes to discussing the implications of uncertainty for the way development projects should be managed, the apparent agreement in point of view is often quickly dissipated. To a large extent the differences that arise do so over the question of the extent of the uncertainty in development—over questions like, Are estimates of cost of production likely to be off by 25 per cent or by 300 per cent?

In its role as advisor to the Air Force on research and development policies RAND is continually forced to make use of predictions about particular development projects. Inevitably, we have become concerned over the confidence that can be attached to such predictions. In this paper we present the results of some recent research into the extent and nature of the uncertainty in new developments. Our general conclusions are:

1. Early estimates of important parameters are usually quite inaccurate, in two respects. First, such estimates are strongly biased toward overoptimism. Second, aside from the bias, the errors in estimates show a substantial variation. That is, even if estimates are multiplied by an appropriate standard factor to eliminate the bias, a non-negligible source of error remains.

2. The accuracy of estimates is a function of the stage of development, i.e., estimates improve as development of the item progresses. This also means that estimates for development projects representing

NOTE: The authors are greatly indebted to Eugene Brussell for his pioneering research on the accuracy of cost estimates, and to Robert Summers for expanding Brussell's work and providing penetrating comments.

only "modest advances" tend to be better than those for more ambitious projects.

We hope in the balance of this paper to make these propositions more meaningful through analysis and discussion as well as presentation of some data.

One word of warning is perhaps in order. Most of our experience has been limited to military development (the Air Force in particular) and what we have to say is largely centered about the problems of development in that area. Although we feel our experience does have relevance to the problems of development generally, we are by no means confident that it applies with equal exactness in all areas. There are some good reasons for believing that privately financed projects may show different tendencies. In our data, estimates of cost of production of the type of equipment most closely resembling regular commercial equipment, i.e., cargo and tanker aircraft, appear to be quite accurate. This is probably partly because of the modest advances usually sought in these programs, and partly it is because for some of these there are commercially available substitutes.¹ There are, of course, many reasons why military development programs are likely to differ from ordinary commercial ventures. Military developments tend to explore very near the margin of technical feasibility; commercial development goes only so far as market profitability. The policy of cost-plus contracting followed in most military projects is not common practice in private developments. Under cost-plus contracting there is very little penalty attached to understating costs, hence contractors have a decided incentive to do so in order to get contracts. The tendency to underestimate costs is, therefore, not very surprising in military development.

The term research and development covers a wide spectrum of human activity ranging from the discovery of new principles to the construction and testing of operational equipment like a commercial aircraft that itself embodies very little that is really new and untried. In this paper we shall not attempt to cover this entire spectrum. Instead we shall focus on what might be called operational development as distinguished from basic research and exploratory development. By the former we mean, roughly, the effort to take ideas or components that have been tested experimentally and embody them

¹ This latter condition leads to a selective process and since we only observe the development of military cargo aircraft that are roughly competitive with commercial carriers no large factor increases in cost estimates from early to late estimates are found.

in useful equipment, as distinguished from the generation of new ideas and their experimental testing. The development of the first jet engine and of the transistor are examples of the latter, while the development of a jet aircraft or a transistorized bombing-navigation system are examples of operational development. Although this division is crude, it defines the area of discussion sufficiently for our purposes.²

Ultimately, the measure of success of any development project is the difference between costs and the value of output. But predictions regarding particular projects are seldom framed in terms of that criterion. Instead they take the form of estimates of separate ingredients of success such as costs, performance, etc.

To discuss predictability rationally then, we need an exhaustive classification of factors that affect success. One such classification that we have found useful consists of four subclasses: (1) costs (development and production); (2) performance;³ (3) time of availability;⁴ (4) utility.

Utility

Because the first three factors are largely quantitative⁵ while the last, utility, is largely qualitative, most of our study of the problem of predictability has centered about the first three, and we shall concentrate on them here. But, in a paper purporting to discuss how well we can foresee the outcome of research and development, it would be a serious error of omission to neglect entirely uncertainties about utility.

This is particularly true of the development of military equipment. The utility of a potential weapon system usually depends on a host of circumstances about which there is great uncertainty. What weapon systems will the enemy have? What kinds of war or wars will it be necessary to be prepared to fight, and where? What will happen to political alliances in the future? What competing or complementary weapons will be developed? And more could be asked. The outcome

² This restriction on the subject matter is not imposed because basic research and exploratory development are less uncertain than operational development. In fact, an imposing list of examples where the best experts failed to foresee the potential of new ideas could be constructed. It is simply a matter of keeping the size of the problem within reasonable bounds.

³ Performance is used here very broadly to include all the qualities of any equipment or system that contribute to its utility. For military equipment the kinds of qualities included are speed, range, accuracy, pay load, maintainability, operability, etc.

⁴ Realistically, availability consists of a schedule of quantities available through time, though we shall consider here only the first operational availability.

⁵ This statement is qualified because certain aspects of performance, like ease of handling, are not quantitative.

of a particular development can, and frequently does, hinge critically on answers to questions of this sort. Thus, ten years ago we did not foresee the rapid development of nuclear warheads by the Soviet and its concomitant strides in ballistic missiles. In fact we did not foresee our own development of hydrogen warheads and the implications of this for the utility of ballistic missiles. But these developments had a tremendous impact on the utility of other weapons such as air-breathing intercontinental missiles that we were then intent on developing. In other instances we have tended to overestimate the quantity or quality, or both, of enemy weapons. During the Taiwan air battle we were pleasantly surprised by the overwhelming superiority of the weapon systems we had supplied to the Nationalists. In short, although it is difficult to quantify uncertainties about the utility of proposed developments, there is no doubt that they exist and are highly significant.

The Equivocal Nature of Predictions

One of the main difficulties encountered in assessing the accuracy of development predictions is the fact that the predictions themselves are seldom formulated in a way that permits verification. Perhaps the most obvious (and most pernicious) example of this is the oft-repeated statement that something is "technically feasible" or "within the state of the art." The "something" referred to is usually a set of performance characteristics for equipment whose development is under consideration. Depending on the interpretation one supplies, assertions like these are generally either useless or irrefutable. If they are intended to mean only that such performance is physically possible, they are quite useless in deciding whether a particular development project deserves support. Since proposed developments seldom violate physical laws, at *some* time and at *some* cost the advertised performance can almost always be achieved. On the other hand, if they are intended to mean more than that—that such performance is attainable at an "interesting" cost and in an "interesting" time period—they are irrefutable until costs and availability are explicitly quantified. It is of course the latter meaning that is ordinarily conveyed, whether it is intended or not, and that meaning provides a kind of ideal prototype illustrating the source of much of the difficulty encountered in measuring the accuracy of development predictions.

In our classification of factors affecting the success of a development program, we listed (aside from utility) three parameters or classes

of parameters—costs, performance, and time. In practice, trade-offs among these three are always possible. There is a production surface relating them. A given performance can be attained earlier if greater costs are incurred; or, for given costs, earlier availability is possible if lower performance is accepted, and so on. The relevance of these trade-offs to development prediction is fairly obvious. Predictions of any one parameter, e.g., performance, that do not specify what is assumed about the other two, are very elusive propositions indeed. For the predicted parameter may always be achieved at the sacrifice of either or both of the other two.

Cost Estimates

One variant of this kind of ambiguity that is particularly troublesome is the standard procedure for estimating costs. When early estimates are made of what it will cost to produce or develop something new, the estimator typically bases his estimate on the current design and the currently planned program for development. If he is estimating cost of production, he gets a total cost by costing the various components as presently conceived and aggregating those. If he is estimating the cost of development, he estimates the cost of test articles, engineering man-hours, etc., as presently planned and aggregates those. He does not specify what performance he is associating with the particular design nor does he indicate the date at which that performance is to be operationally available. He is simply costing a physical configuration or the physical resources contemplated in the current development plan, or both.

As development proceeds, however, these initial designs and plans are almost invariably changed, either because of unforeseen technical difficulties that forestall meeting performance requirements, or because the customer decides it is essential that the equipment be modified so as to keep pace with changing predictions of enemy capabilities, new operational concepts, and new technological possibilities. A guidance system that was supposed to weigh a few hundred pounds turns out to weigh a ton; a development program that was supposed to require fifteen test articles requires forty-five; revised estimates of enemy defense capabilities dictate additional pay load in the form of countermeasure gear; and more. This kind of thing is not the exception, but the rule. Such changes will have an important effect on costs, and the costs finally incurred will not be those of the initial design or the initial development program, but the cost of whatever

actually is produced and whatever program actually is conducted. The difficulties this poses for verifying original estimates are obvious. If costs turn out to differ materially from original predictions, the estimator absolves himself on the grounds that he costed a different program or different product; and, since the initial program or product never emerges, it is usually not possible to know exactly what it would have cost.

In fact, however, absolution is not as simple as all that. The purpose of such cost estimates is to foster a better allocation of resources to research and development projects. Our interest in estimating costs emanates from our desire to economize by making better research and development decisions. In that context we are not concerned with what it would cost to construct particular physical configurations or to carry out particular plans. Instead, we are interested in what it will cost to get some level (or levels) of performance by some date (or dates). Suppose, for example, we are trying to decide whether to initiate the development of a new missile. An analysis of the potential of the missile must run in terms of the performance characteristics of the missile—range, pay load, vulnerability, reliability, etc.—attainable in a certain time period. Accordingly, it is the cost corresponding to the assumed level of performance that is relevant to making the decision, not the cost of a particular development program and a particular physical configuration.

In practice, estimates of the cost of particular development programs and physical configurations are universally married to particular performance characteristics in analyses directed toward making development decisions. In other words, such estimates are used as if they were in fact estimates of the cost of achieving the expected level of performance.⁶ As a practical matter, therefore, it is appropriate to treat cost estimates, however they are generated, as if they were estimates of the cost of achieving the expected levels of performance by a given date. By so doing and by observing the changes in these estimates we have a measure of the uncertainty actually confronted in making research and development decisions.

In principle it would be possible to factor the total error in cost estimates as they are prepared into two parts: (1) the part due to errors in the costing of the configuration supplied to the cost estimator (i.e. the intrinsic error in cost estimating); and (2) the part due to

⁶ Moreover, this is generally done with the active support of those who generate the cost estimates!

PREDICTABILITY OF COSTS, TIME, AND SUCCESS

changes in the configuration as development progresses. In practice it has not been possible to carry out this separation. However, it is our belief that the intrinsic errors in costing a fixed configuration tend to be small relative to errors due to changes in the configuration in the costing of most major items of military equipment. This means that the main improvements in the accuracy of cost estimating, especially for items of equipment or weapon systems embodying moderate or large changes in technology, must be sought in improved forecasts of the final configurations or other statistical means of adjusting early estimates. However, large uncertainties will remain for these estimates in any case, and those uncertainties should be taken into account in their use. Perhaps estimates should be presented in such a way as to make this possible. As estimates are now produced, the results show this would involve at least adding to each estimate an estimate of its bias and mean square error. Ideally it would involve supplying the user with a complete set of sufficient statistics.

Turning now to the data, Table 1 summarizes some data on the history of estimates of cost of production⁷ for twenty-two major items of military equipment.⁸ For each item of equipment we have calculated

TABLE 1
TOTAL FACTOR INCREASES IN AVERAGE CUMULATIVE COST OF
PRODUCTION, UNADJUSTED

| Fighters | Factor | Bombers | Factor | Cargoes and Tankers | | Missiles | Factor |
|-------------------|------------------------|---------|--------|---------------------------|--------|----------|-------------------------|
| | | | | Factor | Factor | | |
| 1 | 5.6 | 1 | 8.7 | 1 | 1.7 | 1 | 57.6 |
| 2 | 3.6 | 2 | 3.5 | 2 | 1.6 | 2 | 20.7 |
| 3 | 3.1 | 3 | 1.5 | 3 | 1.0 | 3 | 11.1 |
| 4 | 2.1 | | | 4 | 1.0 | 4 | 10.3 |
| 5 | 1.9 | | | | | 5 | 1.5 |
| 6 | 1.5 | | | | | 6 | 1.3 |
| 7 | 1.4 | | | | | | |
| 8 | 1.2 | | | | | | |
| 9 | 1.2 | | | | | | |
| Mean | 2.4 | | 4.5 | | 1.3 | | 17.1 (9.0) ^a |
| Mean, all classes | 6.5 (4.1) ^a | | | | | | |

Note: Factor is the ratio of the latest available estimate to the earliest available estimate.

^a Excluding missile case No. 1.

⁷ Not including the cost of development.

⁸ These studies of the accuracy of estimates of cost of production of military equipment are mainly the work of Eugene R. Brussell. We are particularly indebted to him since, so far as we know, his work is the only systematic attempt to study this problem.

the ratio of the latest available estimate of the cumulative average cost of production to the earliest such estimate available.⁹ That ratio is labeled factor in Table 1. Thus if the latest cost estimate were three times the earliest available estimate, the factor would be three. Table 1 shows these factors for four classes of equipment—fighters, bombers, cargoes and tankers, and missiles. (This breakdown simply follows conventional military lines, but it is useful because, as we shall see later, the technical advance incorporated in new equipment correlates fairly well with this classification system.)

The ratio between the most recent estimate and the earliest estimate of the cost of production for these items ranges from a minimum of 1 to a maximum of 57. If we include the extreme missile case, number 1, the over-all mean factor increase is about 6.5; if we exclude it, the mean increase is roughly 4. It is worth noting that there exist substantial differences in the means among the four classes of equipment. For cargoes and tankers the mean factor was only 1.3, while for missiles it was 17.¹⁰

The factors presented in Table 1 are unadjusted. In particular, no adjustment has been made for changes in price levels, and no adjustment has been made for disparities between actual output and the output contemplated at the time the early estimates were made. Both of these adjustments tend to reduce the factors, i.e., reduce the size of the estimating error. Since changes in price levels have generally been upward, the factors are reduced when the latest estimates are deflated. At the same time experience indicates (at least for the kinds of military equipment considered here) that cumulative average cost of production is a decreasing function of total output. In practice, total output

⁹ The data Brussell was able to get are particularly messy. Therefore, a good deal of judgment has had to go into the construction of these estimates of factor increases. But even after the most prudent treatment, the data from which the factors were generated leave much to be desired and a good deal of caution is needed in interpreting the results. The earliest estimate available for some systems was made later in development than for other systems. In some cases the latest available data were actual data on cost of production. In other cases, the latest available data were still estimates either because a program was cancelled or because development was incomplete. Sometimes it was not possible to determine just what the original estimator had included in his estimate in the way of subsystems, spares, etc. If the error in the cost estimates was generally small, e.g. 25 per cent, these deficiencies might be important. In fact, however, the error tends to be so large that the defects in the data are not too troublesome. In addition, on balance, the data were dealt with in such a way that these estimates are, if off in one direction or another, likely to understate the factor increases.

¹⁰ The vehicles included in the six missile cases represent a much more heterogeneous assortment than for the other three classes of equipment. At least one of each of the following is included among the six: air-to-air missiles, surface-to-air missiles, air-to-surface missiles, and surface-to-surface missiles.

PREDICTABILITY OF COSTS, TIME, AND SUCCESS

usually turns out to be much smaller than is originally anticipated to a large extent because costs turn out to be greater than anticipated. Therefore, when the original estimate is adjusted to take account of the reduced output, the effect is almost always to reduce the factor.

Strictly speaking, the unadjusted factors in Table 1 overstate the magnitude of the error in cost estimates. Why do we then bother to include Table 1? First, because a comparison of Table 1 and Table 2 provides an indication of the impact of the output and quantity adjustments. But second, and more importantly, because decisions are sometimes made with little or no appreciation of the functional relationship between costs and outputs. For example, system A, that is now under development, is estimated to cost \$X million per copy based on some output, Q_1 . Thereafter when system A comes up for consideration its cost is taken to be \$X million without regard to output. Thus, even though the cost estimator is thinking in terms of a schedule of average costs associated with appropriate outputs, a single point on that schedule is sometimes the cornerstone for a decision.

Table 2 presents adjusted factors for the same twenty-two items of

TABLE 2
TOTAL FACTOR INCREASES IN AVERAGE CUMULATIVE COST OF
PRODUCTION, ADJUSTED

| Fighters | Factors | | Bombers | Factors | | Cargoes and Tankers | | Factors | | Missiles | Factors | |
|-------------------|---------|-----|---------|---------|-----|---------------------|-----|---------|---|----------|---------|--|
| | A | B | | A | B | A | B | A | B | | | |
| 1 | 3.9 | 4.0 | 1 | 6.2 | 4.0 | 1 | 1.4 | 1.6 | 1 | 14.7 | 6.4 | |
| 2 | 2.6 | 2.5 | 2 | 2.8 | 2.8 | 2 | 1.5 | 1.5 | 2 | 9.4 | 6.0 | |
| 3 | 2.0 | 2.0 | 3 | 1.1 | 1.2 | 3 | 1.0 | 0.9 | 3 | 4.4 | 2.7 | |
| 4 | 1.5 | 1.5 | | | | 4 | 1.0 | 0.8 | 4 | 7.2 | 7.1 | |
| 5 | 1.7 | 2.1 | | | | | | | 5 | 1.5 | 1.3 | |
| 6 | 1.2 | 1.2 | | | | | | | 6 | 1.1 | 0.8 | |
| 7 | 1.0 | 0.8 | | | | | | | | | | |
| 8 | 1.0 | 1.0 | | | | | | | | | | |
| 9 | 1.1 | 0.6 | | | | | | | | | | |
| Mean | 1.8 | 1.7 | | 3.4 | 2.7 | | 1.2 | 1.2 | | 6.4 | 4.1 | |
| Mean, all classes | A B | | 3.2 | 2.4 | | | | | | | | |

equipment. The adjustment is twofold to take account of the two problems discussed above. First, an aircraft composite price index was used to deflate the latest cost estimates. Second, in each case the factor

was adjusted to take account of any difference between the output on which the early estimate was based and the output on which the most recent estimate is based.

You will note that Table 2 contains not one but two sets of adjusted factors. We have elected to show two sets because of the tricky nature of the adjustment process. Discretion and judgment are called for in deciding what adjustments are most sensible. The two sets of factors in Table 2 were prepared by two different individuals,¹¹ and represent somewhat independent judgments as to what is appropriate. Both proceed essentially from the same unadjusted data, however.

As is evident from Table 2, the effect of the adjustments is to reduce the size of the factors. The mean factor for all twenty-two items is about cut in half, to 3.2 for the A set of factors and to 2.4 for the B set of factors. For the missile class as a whole, the mean factor is cut by about 2/3 or 3/4, depending on which set of factors is used. Most of the reduction incidentally is a result of the output adjustment rather than of the price level adjustment. For the twenty-two systems the price level increase from the date of the earliest estimate to the date of the most recent estimate ranged from 3 per cent to 52 per cent and averaged only 28 per cent.

The data in Table 1 and in Table 2 are consistent with the two main propositions made at the beginning of this paper. The cost estimates for these twenty-two items of equipment were decidedly biased toward overoptimism. Even after the adjustments, the most recent estimates on the average exceeded the earliest available estimates by some 240 to 300 per cent. Moreover, among the various items of equipment there were substantial differences in the size of the increase, ranging from slight decreases in some instances to factors of 6 or 7, using even the most modest set of factors.

The difference in the mean size of the factor among the different classes of equipment is also revealing. The smallest increase on the average occurred in the cargo and tanker class, while the largest increase occurred in the missile class. The explanation for this lies in our second main proposition to the effect that the size of the error in estimates is a function of the stage of development or the magnitude of the advance being sought. The performance demanded of new cargo and tanker aircraft, such as range, speed, and even sometimes payload, is usually less than what has already been achieved in other air-

¹¹ One set of factors was computed by Eugene R. Brussell, the other set by Robert Summers.

craft, particularly bombers. Moreover, the subsystems that together make up the air vehicle as a rule are “off-the-shelf” items. The engines, for example, are likely to have already been in operation for some time in other aircraft. Even the airframe is often only a slight modification of some design already in use. In brief, in terms of both performance and physical characteristics of the equipment, only very modest innovations are embodied in cargo and tanker developments. Very likely, the costing errors in these cases are almost totally due to the first of the two sources of error mentioned earlier. The factors for cargo and tanker aircraft probably represent an upper bound on the currently attainable level of accuracy in cost estimating.

At the other extreme, during the period covered by most of these estimates, missile development encompassed what was in many respects a new and radically different technology. Computers and other electronic equipment were required capable of solving complex guidance and control problems without the aid of man in the vehicle, and many of the missiles incorporated propulsion systems with which there had been little previous experience. Meanwhile performance was demanded that meant an order of magnitude improvement over anything we had achieved before.

In other words, the technology and the performance which characterize the missile programs of Table 1 and Table 2 called not for the kind of modest advances sought in cargo and tanker programs but for quite ambitious advances. This, we think, is the reason for the much larger bias and the larger variance exhibited by the missile estimates. Factor increases for cost estimates associated with a second generation of missile systems are likely to be nearer those shown for fighter and bomber aircraft.¹²

Further evidence on the same point is presented in Table 3. Technical experts were asked to classify our twenty-two development programs according to the technical advance sought in each—small, medium, or large. The consensus of that poll was then used to classify the factors as shown in Table 3. The results evidence a correlation between factor size and the size of the advance being sought—the mean factor being an increasing function of the size of advance.¹³

¹² Though the subject is not covered here, we have also compared predicted and actual development costs and found much the same story as for production costs. Actual costs generally prove to be much higher than originally anticipated, with a large variance in the error from one project to another.

¹³ The relationship is even stronger if the effect of a third variable, namely the date of the estimates, is removed. Among the systems that represented large advances, the first

EFFICIENCY IN R AND D

TABLE 3
FACTORS CLASSIFIED ACCORDING TO TECHNOLOGICAL ADVANCE

| | <i>Small</i> | | <i>Medium</i> | | <i>Large</i> | |
|------|--------------|----------|---------------|----------|--------------|----------|
| | Factor A | Factor B | Factor A | Factor B | Factor A | Factor B |
| | 1.5 | 1.5 | 2.8 | 2.8 | 1.1 | 1.2 |
| | 1.7 | 2.0 | 2.6 | 2.5 | 1.0 | 1.0 |
| | 1.0 | 0.8 | 2.0 | 2.0 | 1.0 | 0.8 |
| | 1.4 | 1.6 | 1.2 | 1.2 | 6.2 | 4.0 |
| | 1.0 | 0.9 | 1.1 | 0.6 | 1.1 | 0.8 |
| | 1.5 | 1.5 | 1.5 | 1.1 | 14.7 | 6.4 |
| | | | | | 3.9 | 4.0 |
| | | | | | 4.4 | 2.7 |
| | | | | | 7.2 | 7.0 |
| | | | | | 9.4 | 6.0 |
| Mean | 1.3 | 1.4 | 1.8 | 1.7 | 5.0 | 3.4 |

Availability Predictions

As we have indicated above, time, or how long a development will take, is another important parameter in predicting the outcome of development programs. This is particularly true of military developments where each side struggles frantically to achieve technological superiority in its weapons.

Availability predictions, like cost predictions, exhibit both a decided bias toward overoptimism and substantial variation in the extent of the optimism. Table 4 gives the "slippage" in years for ten of the twenty-two systems covered in Tables 1 and 2, along with slippage factors.¹⁴ Slippage is here defined as the difference between early estimates of first operational dates and actual first operational dates. The slippage factors are the counterpart of our cost factors. They are the ratio of actual time taken to get systems into operation to early predictions of the time it would take to make them operational. Data on availability is also often treacherous, requiring careful auditing and interpretation. Determining when a system actually became operational is a case in point. New equipment in the hands of the operating commands does not mean that an operational capability has been achieved. Which of several early estimates should be used is also frequently a problem. In general, for the systems in Table 4 we have

three were cases in which the earliest cost estimates available were, nevertheless, made fairly late in the program. They are thus based on much better information than was generally the case. For that reason the resulting factors were relatively small.

¹⁴ Data on all twenty-two systems are not available.

PREDICTABILITY OF COSTS, TIME, AND SUCCESS

TABLE 4
SLIPPAGE IN AVAILABILITY

| System | Slippage (years) | Slippage Factors |
|--------|------------------|------------------|
| 1 | 5.0 | 2.5 |
| 2 | 3.0 | 1.6 |
| 3 | 3.0 | 1.5 |
| 4 | 3.0 | 1.5 |
| 5 | 2.0 | 2.0 |
| 6 | 2.0 | 1.5 |
| 7 | 1.3 | 1.3 |
| 8 | 0.7 | 1.2 |
| 9 | 0.5 | 1.2 |
| 10 | 0.3 | 1.1 |
| Mean | 2.0 | 1.5 |

used conservative dates—dates that make the error in estimated availability as small as is reasonably possible.

In spite of this, the errors in estimates of availability dates for those ten systems average 2 years and range as high as 5 years, while the factors range as high as 2.5 and average 1.5. Other studies of slippages, covering a much larger number of systems, tend to confirm the results of Table 4. Estimates of availability are almost invariably optimistic, with the error in estimates varying widely from one weapon system to another.¹⁵

Finally, it is worth noting that, like cost estimates, availability estimates tend to be more accurate the less ambitious the particular project and, of course, tend to be more accurate in the later stages of development than they are in the early stages.

Performance Predictions

The last important measure of success to be considered is performance. Availability and costs each could be expressed in terms of a single vector, time and dollars respectively. Performance, however, cannot be reduced to a single vector. Performance consists of a multitude of

¹⁵ In connection with slippages a word is in order on so-called program stretch-outs. Sometimes program plans are revised and operational dates postponed because of insufficient development funds. But there are two variants of this. There is the case where development expenditures are held to levels below what was anticipated when initial plans were laid, and there is the case where development expenditures, although constrained, are nonetheless as large or larger than was initially planned. To call the latter a stretch-out is, at the very least, misleading. And as a matter of fact the former occurs only infrequently. None of the ten systems here were stretched-out in that sense.

attributes—range, speed, altitude, reliability, maintainability, vulnerability, operability, accuracy, etc. Most of these can be quantified, but others are largely qualitative, e.g., ease of handling. It will be necessary, therefore, to treat estimates of performance much more broadly than estimates of cost and availability.

Estimates of performance also differ from estimates of cost and availability in another important respect. Typically, in weapons development great emphasis is placed on performance. Most new weapons are developed around specific detailed performance requirements laid down by the military—requirements that are taken very seriously. The penalties incurred by contractors for not meeting performance requirements are more severe than those for failure to meet availability schedules or failure to live within original cost estimates. As a result, whenever circumstances dictate a retreat from early plans, it is usually the costs or availability, or both, that give ground. If, for example, technological difficulties with electronic equipment prohibit the attainment of the required performance on the original schedule at original costs, usually either the availability date is postponed, or costs are increased, or both. Degradations in performance are seldom tolerated.

This means that estimates of performance tend to be realized to a greater degree than estimates of either costs or availability. However, among the twenty-two systems in Table 1, most fell short of performance expectations in one or more respects. Range or accuracy was sometimes less than expected. More often than not, reliability was poorer than expected. But the amount by which performance fell short was usually small in comparison to the extension of time or the increase in costs that occurred. Moreover, in some instances performance actually turned out to be somewhat better than expected. In short, though estimates of performance tend to be slightly higher than what is finally achieved, they are much more nearly fulfilled than are predictions of cost and availability.

Summary

The data presented above are neither as comprehensive nor as unambiguous as one would like. Nevertheless, they do give a reasonably accurate picture of recent experience with predicting the outcome of development projects. The truth is that estimates have been quite inaccurate. Cost increases on the order of 200 to 300 per cent and extensions of development time by 1/3 to 1/2 are not the exception,

but the rule. In addition, the size of the error in estimates has varied widely from one weapon to another.

The optimistic bias is not hard to understand. Contractors are anxious to have their proposals accepted by the military, and the military itself is anxious to have development proposals supported by the Department of Defense and Congress. The incentive to make optimistic estimates is thus very strong. On the other hand, the contractual penalties for having been overoptimistic are generally small. But apart from any bias, the military have been living for the last decade or so, and will continue to live for the foreseeable future, in a period of unprecedented technological revolution. Therefore there have been great uncertainties involved in all of their planning and perhaps especially in the R and D area. Errors were bound to be large on the average. The real problem is to explain why so large a portion of the error shows up statistically as bias rather than variance.

The variability in size of the errors observed in individual cases stems from two sources. One is just the basic uncertainty that characterizes all development work. The other is the difference in technological advance sought in different systems. For systems incorporating many new ideas and major improvements in performance the error tends to be larger than for less ambitious projects.

Steps can and should be taken to improve development cost estimates, especially to remove the bias. But even if this is done, significant uncertainties remain to harass the analyst and the decision maker. A more detailed analysis of the nature of these uncertainties and their implications for research and development decision making is the subject of the succeeding paper by Burton H. Klein.

