METHODOLOGICAL PROBLEMS IN EVALUATING THE EFFECTIVENESS
OF MILITARY AIRCRAFT DEVELOPMENT

Thomas K. Glennan, Jr.

May 1966
METHODOLOGICAL PROBLEMS IN EVALUATING THE EFFECTIVENESS
OF MILITARY AIRCRAFT DEVELOPMENT

Thomas K. Glennan, Jr.*

The RAND Corporation, Santa Monica, California

This paper is concerned with one aspect of the effectiveness of a
process, the process of translating a body of theory and experience into
a capability to produce a military aircraft. We intend to examine
several issues dealing with the efficiency of this process. To deal
completely with the problem of the effectiveness of a process we would
also have to deal with questions associated with the quality of the
product developed by the process. We will not do so here; instead, we
will assume that the system performs more or less according to the
specifications initially set down for it.**

Having made this crucial assumption, we seek to find a relation-
ship between the cost of a development and the attributes of the

*Any views expressed in this paper are those of the author. They
should not be interpreted as reflecting the views of The RAND Corpora-
ton or the official opinion or policy of any of its governmental or
private research sponsors. Papers are reproduced by The RAND Corpora-
ton as a courtesy to members of its staff.

This paper was prepared for presentation at the 29th National Meet-
ing of the Operations Research Society of America.

**In support of this assumption we may cite the observation of
Marshall and Meckling that while most of the systems they had examined
one or more fell short in performance dimensions, "...the amount by
which performance fell short was usually small in comparison to the
extension of time or the increase in costs that occurred." A. W. Marshall
and W. H. Meckling, "Predictability of the Costs, Time, and Success of
Development," in The Rate and Direction of Inventive Activity, Princeton
1962.
developed aircraft, the qualities of the process by which it is
developed, and the development time. If such a relationship can be
found, we would be in a position to say something about the efficiency
of a particular development. There are several possible types of state-
ments that could be made.

1. We might identify certain process characteristics which lead
to lower costs.

2. We might identify certain developments as being efficient or
inefficient depending on the deviation of their actual cost from their
projected cost using the relationship we have found.

3. If we find a sharp inverse relation between development cost
and development time we could question the effectiveness of rapid
developments where such rapidity seems unneeded.

Our first task then is to attempt to establish a relationship
between development cost, aircraft attributes, development process
qualities, and time. What are the determinants of development cost?
There appears to be a "conventional wisdom" which has grown up con-
cerning the determinants of development cost. Perhaps the best exposit-
ion of these is found in Peck and Scherer, The Weapons Acquisition
Process: An Economic Analysis.* The following are my interpretations
of this conventional wisdom.

1. Intrinsic size and complexity of the aircraft.

It seems reasonable to suppose that a larger aircraft of a given
complexity will require more development effort than a smaller aircraft
of the "same" complexity. There are more drawings to be made, more

---

*M. J. Peck and F. M. Scherer, The Weapons Acquisition Process:
An Economic Analysis, Boston, 1964, Chapters 9 and 11.
analyses to be done, and more tests to perform. Similarly for aircraft of comparable size, more complex aircraft cost more to develop than less complex aircraft. Costs are therefore expected to be positively related to these factors. This holds irrespective of the degree to which the designers are attempting to achieve performance levels beyond those previously achieved; that is, independent of the state-of-the-art advance being sought.

2. Technological uncertainty.

The cost of development should be expected to be related to the technological uncertainty surrounding the project. If there is great uncertainty, false design starts will occur, costs will be expended to resolve the uncertainties, and in some cases multiple or parallel approaches to particular components will be used. Frequently the technological riskiness is associated with the loose and ill-defined term, state-of-the-art advance.

3. Development time.

Generally development costs are held to vary inversely with development time. As one tries to shorten development time, tasks that would ordinarily be done sequentially must be done in parallel. This heightens the chance of design errors, increases the costs of coordination, and raises the cost of building and dispersing a design group. However, it is sometimes argued that the shortening of development time eliminates much unnecessary work and "gold plating" and that it therefore lowers costs.* In addition, some fixed or "overhead-type" costs continue irrespective of the level of effort and would tend to increase the total costs.

---

cost of developments taking longer times. No doubt there is truth in such an assertion and that over some range of development times, the factors tend to counteract the influences toward higher costs.


Just as in production processes, there are superior and inferior ways of organizing and conducting development projects. These costs would conceptually be measured as the increased costs above some minimum "efficient" cost. Conceptually, these costs can be separated into those associated with what we will call development strategy and those associated with general inefficiency of the development organization. A development strategy is the strategy used for allocating and reallocating development sources to tasks associated with resolving design uncertainties. Such a strategy is very complex in reality but strategies have been crudely defined for aircraft development. For example, there has been considerable discussion through the years of whether or not an airframe prototype should be used as a part of an aircraft development. If one could demonstrate that lower costs were associated with development using this development strategy as opposed to a strategy which attempts to develop aircraft using other means of resolving development uncertainty, this would be a significant finding concerning the efficiency of development activities. This issue is complicated by the fact that the strategy chosen is dependent in part upon the development task. We would not expect one type of strategy to be appropriate to all types of development.*

There are other process characteristics which bear upon process efficiencies. There are superior and inferior design review procedures, there are some designers who simply have better capabilities than others, and some firms are likely to be more effective than others in controlling their costs.

5. **Productivity factors.**

It is impossible to find sufficient aircraft developments being conducted at any one time to permit a "cross-section" type of analysis to be conducted. This means that the projects examined here have taken place over time and we must therefore consider possible productivity changes associated with the factors used in development. Such changes could be expected to occur. The quality of the theory used by engineers is steadily improving. The quality of the test equipment and facilities has improved. The computer has been introduced and its use expanded through time. Conceivably these changes could have reduced the cost of development.

**SOME EMPIRICAL EVIDENCE**

Cost data on thirteen military airframes have been collected and analyzed.* Examination of these data suggests that development costs are difficult to define and that it may be necessary to resort to some fairly arbitrary definitions. For example, how should test aircraft be charged to development? The number of such aircraft has varied from six or seven to as many as thirty or more. Some of these aircraft can be refurbished and put to operational use, but others simply cannot make

---

*The airframe developments analyzed include the F-84, F-86, F-86D, F-89, F-100, F-101, F-104, F-105, F-106, B-47, B-52, B-58, and F-4.
the grade. The need for the various numbers of aircraft may be related
to the complexity of the development task but in some cases it appears
to be more nearly associated with the particular development philosophy
in vogue at the program's inception. In some cases development seems
to go on almost continuously through the production of a hundred or so
aircraft while in others the development appears complete after the
production of fifteen. Most economists would choose to separate develop-
ment activities from capital investment activities and thus would attempt
to exclude those activities associated with tooling from development costs.
Yet substantial tooling is required to produce test vehicles and in many
developments substantial parts of the tooling may have to be redesigned
because of discoveries made during design and test. Such costs should
be allocated to the development effort, but this is impossible with the
data available to us.

As a result we have made use of some fairly arbitrary definitions
in our study. First, we have imputed a cost of engineering, tooling,
and production for the first twenty-five airframes, corrected for price
rises. These costs include allocations of overhead expenses as they
were made by the firms themselves. We have done the same thing for 100
airframes. The statistical results presented here include the cost of
engineering including flight test for the first 25 and 100 airframes
and the total cost of the airframe developer's activities (including
engineering, tooling, and production) through 25 aircraft. * None of
these measures of development cost corresponds exactly with a proper

---

*We have attempted to exclude costs associated with spares, train-
ing units, specialized ground support equipment and manuals.
economic definition of R&D but they do serve as convenient and useful surrogates.

To relate these costs to the factors we have suggested may determine development costs, we have computed a series of regressions using the functional form:

\[ C = A \cdot B^\beta \cdot D^\gamma \cdot E \]

where \( C \) is the particular cost surrogate, \( B \) and \( D \) are measures of the factors influencing the cost of development and \( E \) is an error term. Naturally the small size of our sample will restrict the number of variables we can consider as determinants of development cost. Certainly we cannot expect all of the factors suggested above to be included so we seek instead to identify those which are most important.

The data in our sample for development costs have an extremely large range. For example, the engineering costs at 25 airframes range from a little over five million to over two hundred million 1962 dollars. Therefore, we concentrate first on seeking measures of intrinsic size and complexity. This problem has been examined by numerous people, primarily in connection with cost analysis activities. Generally a relationship using weight and speed seems to provide a good explanation. We tried numerous other measures including, for example, various ranges, maximum engine thrust, and measures of aircraft density. None of these other measures provided as good an explanation as weight and speed or added to the explanation provided by weight and speed. Therefore we have the following regressions, using the logarithmic form.*

*For those unfamiliar with regression analysis, a short note may be in order. The figures in parentheses are the standard errors of the estimated regression coefficients above them. While there are statistical
(1) \[ \log E_{25} = .41 + .83 \log W + 1.97 \log S \]
\[ (.14) \quad (.11) \quad (.21) \]
\[ R^2 = .92 \]

(2) \[ \log E_{100} = .46 + .91 \log W + 1.84 \log S \]
\[ (.12) \quad (.09) \quad (.18) \]
\[ R^2 = .94 \]

(3) \[ \log C_{25} = .72 + 1.04 \log W + 1.42 \log S \]
\[ (.08) \quad (.06) \quad (.13) \]
\[ R^2 = .97 \]

where \( E_{25} \) = Cost of engineering including flight test through 25 aircraft (1962 dollars).
\( E_{100} \) = Cost of engineering including flight test through 100 aircraft (1962 dollars).
\( C_{25} \) = Cost of airframe program through 25 aircraft (1962 dollars).
\( W \) = AMPR weight of aircraft in 1000's of pounds.
\( S \) = Maximum speed of aircraft expressed as mach number.

We note here the generally high correlation coefficients and the significance of the regression coefficients. However, care must be used in interpreting the correlation coefficients because of the large range of the dependent variables. For example in equation 1 the average value of the absolute deviation of actual from predicted costs expressed as a percentage of predicted cost is 23 percent. Thus, a great deal is lacking in the explanation of development cost provided by equation 1.

The conventional wisdom holds that the technological uncertainty associated with the project should have a bearing upon the cost of developing a new airframe. This uncertainty is viewed as being associated tests of significance which can be used, generally if the standard error is less than one-half of the estimated regression coefficient, there is a high probability that the true coefficient is different than zero and has the sign indicated. See B. Ostle, *Statistics in Research*, Iowa State College Press, Ames, 1954.
with the state-of-the-art advance (SOAA) being sought in the development. The SOAA seems to defy measurement. For example, suppose that speed is a dominant aspect of the airframe state of the art (SOA). Then we might consider the SOAA to be measured by the amount that the speed sought in a development exceeds the previously achieved speeds. This raises a number of thorny questions. Should we consider the highest speed achieved by aircraft previously developed by the firm or by the entire industry? Should we include research aircraft or only operational aircraft? Should the difference be expressed as a ratio or an absolute term? Why should we expect the impact of speed increases in the transonic region (where historically the problems were largely aerodynamic) to be related to the impact of speed increases above mach 2.2 (where the problems are importantly related to materials)? All of these difficulties are increased when we consider other factors besides speed, such as size and maneuverability. We simply cannot see a quantitative solution to the problem of relating performance characteristics desired from a development to the state-of-the-art advance.

Instead we asked a small group of engineers at RAND to give us their opinions concerning the state-of-the-art advance associated with the airframes in our sample. Their rankings showed a high degree of consistency which could be considered as giving us greater confidence in the measure than if there had been little correlation between their responses. On the other hand, it might simply indicate that they had all formed their opinions based upon the difficulty the developers had had with the development. We computed the following regressions using this measure.
\[ \log E_{25} = 0.51 + 0.85 \log W + 1.99 \log S - 0.15 \log A \]
\[ R^2 = 0.92 \]

\[ \log E_{100} = 0.53 + 0.93 \log W + 1.85 \log S - 0.11 \log A \]
\[ R^2 = 0.94 \]

\[ \log C_{25} = 0.89 + 1.07 \log W + 1.45 \log S - 0.25 \log A \]
\[ R^2 = 0.97 \]

Where \( A \) = measure of state-of-the-art advance obtained by polling experts with a range from 0 to 12. The results show that our measure of SOAA has no value in explaining the two engineering costs we have suggested could be surrogates for development cost. However, there is apparently some relation to the total cost through 25 aircraft.* The fact that there is a significant relationship to total costs as opposed to engineering cost is not too surprising. Military developments are conducted under conditions of considerable urgency. The result is an overlapping of development and production activities. Much tooling and production work is undertaken before the design is complete. If a design error is found, frequently the engineering needed to correct it is not too extensive. The finding of the error may, indeed, point to the solution. On the other hand, if tooling and production have proceeded on an erroneous design, the tools and production may have to be scrapped or significantly altered. Thus the real impact of design uncertainty may be in tooling and production.

*Using the "t" test, this coefficient is significantly different than zero at the 10 percent level.
The more puzzling aspect of our findings is the negative coefficient and the generally small impact that SOAA has. As a first and perhaps most tenable hypothesis, the measure we have used may be wrong. But we should not dismiss the results out of hand. Examination of the development programs and discussion with government and industry personnel suggests that there are many ways in which projects are modified to minimize cost increases due to SOAA. For example, the developer may take degradations in non-mission-essential performance to make the design job simpler. Poorer reliability and less stringent maintainability requirements may be allowed. More development time may be allowed. In other words, there may be a realistic reduction of the degree of refinement which is sought in aircraft attempting to advance the state of the art. These factors may well combine to make SOAA a less important factor affecting development cost than we might expect.*

We turn now to a consideration of productivity changes. We know no measure or measures which can give us a good feeling for the improvement of the quality of the factors going into the development process. We might assume that productivity is monotonically increasing at a constant rate through time. If this were indeed the case, there should be a negative trend effect in our data. For a given level of complexity and size less manpower and/or capital would be required to develop a new device. Unfortunately however there are many other factors which vary through time and these may be acting to counteract any productivity

*One can argue, of course, that it is not our measure of SOAA which is at fault in its failure to explain development costs. If our measure of size and complexity were more comprehensive, it might be that the SOAA would be a better explanation than it appears to be here.
trends which exist. The following regressions are obtained using the months from January 1944 to the date of first flight as a trend variable.

\[
\log E_{25} = 7.64 + 0.69 \log W + 1.13 \log S + 0.67 \log T \\
\quad \quad = 0.49 (0.11) \quad 0.43 (0.31) \quad \quad \quad R^2 = .94
\]

\[
\log E_{100} = -0.19 + 0.82 \log W + 1.32 \log S + 0.41 \log T \\
\quad \quad = 0.47 (0.11) \quad 0.41 (0.30) \quad \quad \quad R^2 = .95
\]

\[
\log C_{25} = 0.12 + 0.96 \log W + 0.94 \log S + 0.39 \log T \\
\quad \quad = 0.32 (0.07) \quad 0.27 (0.19) \quad \quad \quad R^2 = .98
\]

where \( T \) = months from January 1944 to first flight of aircraft.

As can be seen, the introduction of trend tends to change both the coefficient of speed and of weight in all of our equations. This suggests the possible appearance of a problem which is to be expected in analysis of data which come from events which take place over time. Speed, which is a measure of complexity, has been generally increasing over time. This is equivalent to saying complexity is increasing through time. The result is that our term for complexity appears to have drawn some of its statistical significance from other factors which were changing over time and which tended to drive development costs upwards. Speed was serving not only as a surrogate for complexity but perhaps for other time related variables as well.

For example, during the time period under consideration, there has been increasing government involvement in the direction of development projects. This involvement has led to a proliferation of reports, briefings, and meetings during development. Some of this may reflect increasing program complexity. Some may simply reflect increasing
bureaucratization. Both would tend to increase costs. It can be argued also that we are tending to lower the risk in development programs by increasing the analytical and experimental work associated with development programs. Thus the programs we have in our sample may involve increased costs due to changes in the risk content.

These types of factors may outweigh productivity increases on the part of factors of production. In any case we are reluctant to ascribe to any single factor a major share of the responsibility for the trend relationship we observe.

Although the effect is not so pronounced, we notice that the coefficient for weight is also affected by the introduction of a trend variable, reflecting the fact that there has been a slight trend toward heavier aircraft with time.

In sum, to the degree to which our measures of intrinsic size and complexity allow us to "normalize" the jobs done, there is no obvious decrease in the cost of performing development activities. There are several explanations which can be advanced for such an observation. There may have been the introduction through time of additional tasks hidden in our development costs. Possibly too, the risk of disastrous failure has been reduced in these projects through time by the accomplishment of a more thorough engineering job. Finally, it is possible that in a process where standards are so hard to set, we have steadily become more inefficient.

In our initial discussion, we note that we expect the time spent in development to have an effect upon development cost. It might seem that development time could be easily measured. For our purposes this
is definitely not the case. For example, the compression of tasks in
the development program is probably determined by the development time
planned at the initiation of a program rather than the achieved develop-
ment times. If a developer gambles and fails, not only will he have
wasted resources in concurrent activities that do not fit together but
he will lose time too. Two programs, both intended to develop a similar
piece of hardware and both taking the same length of time could be very
different in cost if one attempted to shorten time and failed while the
other was intended to take the longer time.

Unfortunately, however, we have little firm data on planned develop-
ment times for most of the programs in our sample. In some cases, par-
ticularly in the early programs utilizing prototype aircraft, there
simply was no planned date for the completion of development. In some
of the later programs the date jumped erratically due to requirements
changes.

We have had to use actual development times in our analysis. But
even this time is fraught with difficulties. The dating of the begin-
ning of a development is difficult. A company frequently started
preliminary design prior to a contract being signed in the programs
in our sample. * We are not far wrong in taking the signing of a first
letter contract for a development as the initiation date. ** The dating

* The situation is quite different for post-1960 programs with the
introduction of CDP. The initiation of a development effort today
clearly antedates the signing of a firm contract for a development
effort by a year or more. None of the programs in our sample was
started after 1955.

** The F-100 was clearly initiated before such a contract signing
and North American Aviation was subsequently compensated for these prior
efforts. In this case we made a correction to take this into account.
of development completion is more perilous. As noted earlier, in some programs development activities seem to continue through a very large number of aircraft deliveries. In others, development seems largely completed by the time the first operational squadron is activated. This variation can be seen by an examination of our data on engineering cost. Table 1 shows the ratio of the engineering cost at 100 aircraft minus the engineering cost at 25 aircraft to the engineering cost at 25 aircraft. A ratio of one would indicate that just as much engineering effort was required to get from 25 to 100 aircraft as was required to get to 25 aircraft.

Table 1

<table>
<thead>
<tr>
<th>Program</th>
<th>( \frac{E_{100} - E_{25}}{E_{25}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.37</td>
</tr>
<tr>
<td>2</td>
<td>.08</td>
</tr>
<tr>
<td>3</td>
<td>.32</td>
</tr>
<tr>
<td>4</td>
<td>.59</td>
</tr>
<tr>
<td>5</td>
<td>.68</td>
</tr>
<tr>
<td>6</td>
<td>.18</td>
</tr>
<tr>
<td>7</td>
<td>.65</td>
</tr>
<tr>
<td>8</td>
<td>.15</td>
</tr>
<tr>
<td>9</td>
<td>.11</td>
</tr>
<tr>
<td>10</td>
<td>.94</td>
</tr>
<tr>
<td>11</td>
<td>.17</td>
</tr>
<tr>
<td>12</td>
<td>.22</td>
</tr>
<tr>
<td>13</td>
<td>.25</td>
</tr>
</tbody>
</table>

The wide variance in these figures suggests the difficulty in dating the completion of a development. Again we retreat to a fairly arbitrary dating scheme, the month in which the twenty-fifth aircraft is accepted by the service. Using this development time we obtain the following regressions:
\[ \log E_{25} = .77 + .94 \log W + 2.05 \log S - .29 \log D \]
\[ R^2 = .92 \]

\[ \log E_{100} = .50 + .93 \log W + 1.85 \log S - .04 \log D \]
\[ R^2 = .93 \]

\[ \log C_{25} = .91 + 1.10 \log W + 1.46 \log S - .15 \log D \]
\[ R^2 = .97 \]

where \( D \) = time in months from letter contract to 25 acceptances.

While the direction of the sign of the coefficient of development time is in the direction predicted, the coefficients are never significant. To the extent that the measure of development time that we have used represents the urgency of the program, this result suggests that this factor does not seem to affect costs significantly. We note however that when a program is compressed in time, the way in which a development team does its job may change. The tolerance between components may be increased, the refinements which contribute to maintainability or manufacturability may be skipped. Thus the product that emerges from a rapid development, though it could possess the same weight and speed as another aircraft which was developed more slowly, might be quite a different product. Nonetheless the development costs could be nearly the same. This argument is the same as we introduced as a possible explanation for the lack of or possibly negative effect of SOAA.

Finally let us turn to characteristics of the development process itself. While a number of hypotheses might have been examined, we look at only one factor which might affect the efficiency of the development process, the amount of subcontracting used in the initial production.
It can be hypothesized that the problems associated with coordinating activities of outside producers, including, frequently, some design work, increase the cost of development. This issue is important as we move to higher subcontracting rates as a matter of public policy. We use as our measure of subcontracting the percentage of off-site production man hours of the tenth aircraft as reported in the AMPR documentation. The results of the regression are:

$$\log E_{25} = .27 + .78 \log W + 1.90 \log S + .17 \log R$$
\((.22)\) \((.13)\) \((.23)\) \((.21)\)

$$R^2 = .92$$

$$\log E_{100} = .37 + .87 \log W + 1.80 \log S + .11 \log R$$
\((.19)\) \((.11)\) \((.20)\) \((.18)\)

$$R^2 = .94$$

$$\log C_{25} = .54 + .96 \log W + 1.33 \log S + .23 \log R$$
\((.11)\) \((.07)\) \((.12)\) \((.11)\)

$$R^2 = .98$$

Here the coefficients have the sign we predict but only in the case of the total cost at 25 airframes do we have a value which seems significantly different than zero. The evidence would seem to lead some weak credence to our hypothesis.

We can compute regressions using more independent variables but our small sample size limits the usefulness of such exercises. We have done so and the results are generally uninteresting.

CONCLUSIONS

Based upon our attempts to identify the determinants of development costs we can point out many methodological difficulties.
1. The definition of development and the measurement of its cost is conceptually difficult. We are forced to seek a series of fairly arbitrary definitions.

2. Many of the conceptually important variables such as development urgency and state-of-the-art advance defy measurement.

3. The measures we have adopted for intrinsic size and complexity, while important and useful, have clearly failed to capture all the aspects of product quality which should be considered.

4. Our limited sample size with the data points spread through time give rise to problems of multicollinearity and inadequate degrees of freedom.

5. These factors, combined with a number of lesser importance, lead us to conclude that the measurement of the efficiency of a development using a technique such as ours would be hazardous indeed.

Nonetheless, our results are of some interest. Qualified by the sorts of problems mentioned above, we consider the following to be significant:

1. Development time does not seem to have an important influence on development costs, at least within the range of development times in our sample.

2. State-of-the-Art Advance, as determined by a group of experts, may have a negative relationship to development costs. This may be due to changes in the task goals which are not reflected in our variables.

3. There is no evidence of productivity increases of the factors of development (if they have occurred) being taken out in the lowering of development costs.
4. There is some evidence that the coordination of subcontracting efforts may raise development costs, however, these costs appear to be largely associated with tooling and/or production.