WHEN SHOULD WE START HIGH-RATE PRODUCTION OF THE B-2?

An Analysis Based on Flight Test Results

Michael D. Rich

P-7757

Published 1991 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

When Should We Start High-Rate Production of the B-2?

An Analysis Based on Flight Test Results

STATEMENT OF

MICHAEL D. RICH
Vice President, National Security Research
Director, National Defense Research Institute

RAND
Santa Monica, California

Prepared for
United States Senate
July 26, 1991

PREFACE

The research that I will describe here has a long lineage. For thirty years RAND has studied many topics associated with the development and procurement of major systems, including weapon systems and large-scale civil projects.

In 1987, the Under Secretary of Defense for Acquisition requested that RAND conduct a study of the acquisition strategy of what was then known as the Advanced Technology Bomber. Our initial results, which I briefed to Secretary Weinberger, Secretary Aldridge, and several congressional audiences in the fall of 1987, gave high marks to the risk-reduction measures taken in the early phases of the program. However, we expressed considerable concern about the planned pace of the flight test program, which had not yet begun, and especially about the scheduled timing of key production go-aheads. We made several specific recommendations about both the test program and its relationship to the production program. The subsequent evolution of those programs reflected our recommendations.

Shortly after our study, the Congress directed OSD to establish a "Cost, Performance, and Management Initiative" for the B-2. In connection with that initiative, OSD asked RAND to continue its analysis of the program's acquisition strategy. One of the questions we tackled was how to tie production commitments to progress achieved in the test program, which was one of the recommendations we made in our 1987 study. That's the part of the study I will describe today.[1]

This research was performed for OSD and conducted within the National Defense Research Institute, the federally funded research and development center sponsored at RAND by OSD and the Joint Staff. The views expressed, however, are those of the research team and do not necessarily represent those of RAND or any of its research sponsors.

[1]This part of the research was performed by Giles K. Smith, under the direction of John Birkler.

ACKNOWLEDGMENTS

This Senate testimony summarizes a portion of the research undertaken in a recent study of the B-2 acquisition strategy sponsored by the Director of Defense Research and Engineering. John Birkler co-led the overall study and Giles Smith directed the portion covered by this statement.

INTRODUCTION

Over the past decade a series of initiatives from government commissions and from the Congress have urged the Department of Defense to devote more attention to testing activities during the weapons acquisition process. One special thrust has been to encourage the Services to not only do more, and better, operational testing, but also to defer high-rate production of a new system until its operational suitability and effectiveness have been demonstrated.

Given the mandate to demonstrate operational suitability and effectiveness of a new weapon system before authorizing high-rate production, the obvious question is, *how much testing and demonstration are enough?* The complete flight test program of a modern aircraft system typically extends four to five years after first flight. The industrial lead time from funding authorization to delivery of such a system is at least three years. Thus, to wait until the end of all testing before funding the first high-rate production lot would mean delivery of the first inventory aircraft at least seven years after first flight. Such a delay would be expensive and would diminish the combat advantage provided by the technology advances incorporated in the design.

We know that system maturity (the absence of flaws and performance shortfalls) of a new airplane typically improves during the flight test program. Thus, we need some systematic basis for selecting a point in the flight test program that probably will fall short of full completion but that will provide sufficient confidence in the basic design to justify production funding. The objective of this study is to develop the necessary analysis model and to apply it to the B-2 bomber development and flight test program.

Decision Environment

The decision to authorize funding for high-rate production of a new weapon system obviously depends on many factors. The present analysis assumes that additional quantities of the B-2 bomber will be procured, and that at some point in time we will seek the economies possible through production rates somewhat higher than the Low Rate Initial Production (LRIP) of about two units per year that has characterized recent appropriations. The present study addresses certain technical and economic issues that affect the *timing* of the decision to move to higher production rates.

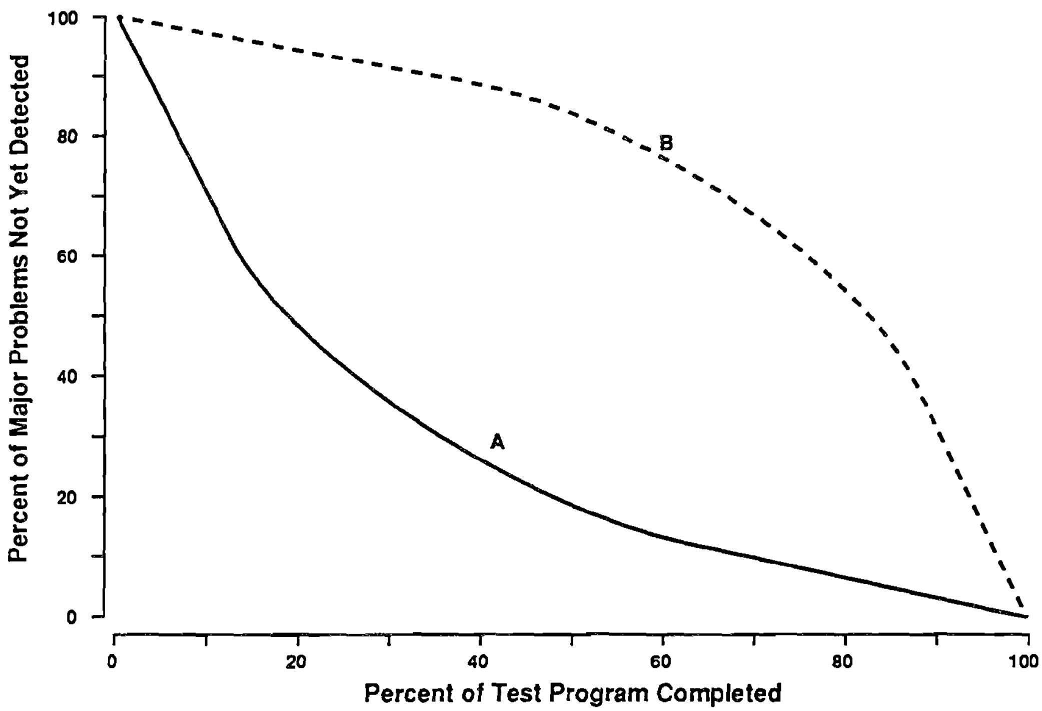
Research Approach

This research is based on the hypothesis that the fiscal and operational benefits of an early production authorization warrant accepting some risk that all problems have not been identified and corrected. The goal of the research was to identify and quantify those elements of risk that could be reduced through flight test, and to provide a model so that those risks could be systematically and quantitatively projected and then balanced against the costs of delaying production.

This basic approach can be illustrated by the sketch in Fig. 1. At the beginning of a flight test program, the probability is very high that the design contains important flaws. As we progress through the flight test program we discover those flaws, so that by the end of the combined development test and initial operational test program we expect to have found all, or nearly all, of the important problems. The flight test program can therefore be viewed as a process for reducing the risk that undetected flaws remain in the design. To help the acquisition executive who is contemplating authorization of high-rate production before the flight test program is complete, we need to know something about the *rate* of risk reduction. Two alternative conceptual models are depicted in Fig. 1. Curve A reflects the notion that important flaws tend to be revealed quite early, while Curve B suggests that it is not until near the end of flight testing that we begin to rapidly accumulate

Fig. 1

When Are Major Problems Detected?



confidence that the important flaws have been revealed. An acquisition executive who believes in curve A will be inclined to authorize high-rate production after only a modest amount of flight testing has been completed. Conversely, belief in curve B will lead to a later production start.

One of the major steps in the present study was to assemble data from the flight test phase of previous combat aircraft development programs in order to get some idea of which of the alternative models is most accurate (Fig. 2). The following programs were examined in some detail:[1]

[1] Results were compared in a cursory manner with those from two other major combat aircraft developed in recent times, the F-15 and F-16, and no inconsistencies were found.

Fig. 2
Research Approach

- Review history of recent aircraft flight test programs
  - B-1A, B-1B, F-117, F/A-18, C-5A
  - Time distribution of key events
- Apply historical patterns to estimate future progress of B-2 flight test program

- The B-1A/B-1B bomber, the most recent bomber program and possibly the one most directly analogous to the B-2;
- The F-117, our only operational stealth airplane;
- The F/A-18 fighter, another recent major combat aircraft; and
- The C-5 cargo airplane, the other large aircraft that has completed development and test during the past couple of decades.

For each program the history of the flight test phase was reviewed in detail to determine what problems were discovered, when those problems were discovered, and how long it took to make the necessary design change and to incorporate the change in the production line.

When we examined the past programs it became apparent that we needed to distinguish between some different kinds of "problems" (Fig. 3):

1. Those that seriously diminish the mission capability of the system, and that are so difficult or expensive to fix that they threaten the very life of the program;

Fig. 3

Some Problems Are Worse Than Others

Type 1: Potential program killers

-- Seriously diminish mission capability

-- Very difficult or expensive to correct

Type 2: Troublesome but manageable

-- Seriously diminish mission capability

-- Can be corrected, but require time or funding beyond original program scope

Type 3: Routine

-- Can be corrected within scope of original program

2. Other important problems that also affect mission capability but are capable of being corrected to an acceptable degree, although that corrective action causes important delays or increases in program cost.
3. Routine problems that can be corrected within the original program budget and schedule.

There are, of course, many problems discovered during flight test. Most are Type 3 (corrected within the scope of the original program) and therefore have little or no effect on major program decisions. By focusing on the two kinds of major problems that might be uncovered during flight test, we concluded that we needed a two-step decision process when addressing the high-rate production decision:

- First, it seems appropriate to demonstrate enough of the critical system capabilities so as to achieve high confidence that there are no flaws in the system concept so serious as to justify program cancellation. We refer to this as "proof of concept."
- Given proof of concept, there is a second, cost-related criterion that can be applied to the high-rate production. There is some point in the test program where the expected cost of correcting flaws not yet discovered is balanced by the expected costs of further delay in high-rate production. Starting high-rate production at that point should minimize the expected total production cost, including the cost of correcting flaws and retrofitting already-produced units.

By applying this analysis approach to the B-2 program we reached two conclusions (Fig. 4):

Fig. 4

Conclusions

- High confidence in system concept by middle of CY 1992
  - Verification of detection/survivability
  - Flight vehicle well demonstrated
  - First looks at weapon release and mission equipment
- Delay beyond FY92 authorization for starting high-rate production likely to increase total production cost

1. By the middle of 1992 we should have completed enough flight tests to provide high confidence that there are no design or performance problems so serious that they might justify termination of the program.
2. The costs of correcting flaws not yet uncovered are likely to be less than the costs of delaying onset of more efficient, higher-rate production. Any further delay in authorizing higher-rate production will almost certainly increase the total production cost of the B-2 fleet.

The remainder of this paper will describe the reasoning and analysis that supports those two conclusions.

PROOF OF CONCEPT

Military weapon systems are typically designed against very demanding performance goals, and rarely does a new system fully achieve all such goals early in its operational life. But each new system is typically configured around a few critical design concepts and performance goals, such that failure to meet those goals jeopardizes the entire program. Those goals should be demonstrated to a rather high level of confidence. The question is, how much testing is needed in order to have high confidence that there are no remaining "show stopper" problems in the design?

Historical Guidance

We approached this issue in two ways. First, we sought guidance from past experience. We reviewed a wide range of earlier aircraft weapon system programs, together with other systems that are technologically complex and challenging, to see when in the test programs project-threatening (Type 1) problems were revealed, and how often that occurred. Surprisingly, *we found very few instances where major weapon system programs were cancelled because of problems revealed or confirmed during full-scale tests.*

During the past couple of decades, which we believe reflect modern technologies and modern management practices, we found only two programs that were cancelled after full scale testing had commenced and before a substantial amount of serial production: the Army's DIVAD gun (the Sergeant York) and the Tacit Rainbow missile. There were no aircraft programs cancelled on the basis of major design or performance problems revealed during full scale testing.

Other programs have been cancelled during the full scale development and test phase. The B-1A and the T-46A are recent examples. While both suffered some performance problems, it seems clear that budget and broad political issues dominated the fate of those projects, rather than the existence of basic design problems as revealed during full scale tests.

System-Specific Criteria

One thing apparent from our review of past programs was that the few Type 1 problems actually encountered in full scale system tests had usually been identified as critical issues during the engineering development phase. That gave us confidence that we should be able to create, a priori, a list of such critical issues for the B-2 and then see when in the flight test program we are likely to have accumulated some verification of performance in those areas.

We prepared a list of those aspects of the B-2 design where performance at least close to program goals is absolutely necessary for mission accomplishment, and *where initial performance shortfalls might be quite difficult or expensive to correct.* We believe four areas satisfy those criteria (see Fig. 5):

1. **Detection/Survivability.** The basic system performance characteristics necessary to degrade enemy detection and to enhance survivability should be validated during full scale flight tests of a fully-configured vehicle. This includes demonstration of all relevant signature-reduction techniques to the extent necessary to achieve signature goals throughout the

Fig. 5

Critical Aspects of B-2 Design

	When Demonstrated
• Detection/Survivability	End 1991
• Vehicle Performance	
– Flying qualities throughout critical portion of envelope	
– Propulsion system performance	
– Range-payload capability	Mid-1991
– Airfield performance	
• Structure adequacy	
– Air loads in selected maneuvers	
– Limit load demo	End 1991
– One fatigue lifetime	
• Weapons separation	Mid-1992

anticipated spectrum of operational missions. *These performance characteristics should be adequately demonstrated by the end of this year.*

2. **Vehicle Performance.** The novel vehicle configuration makes verification of basic flight vehicle performance especially important. This includes demonstration of acceptable flying qualities and performance levels throughout the critical portions of the speed-altitude envelope, demonstration of propulsion system performance, and demonstration of acceptable handling qualities and performance during take-off, landing, and ground operation. Adequate cruise efficiency should be demonstrated at gross weights representative of useful loads for typical missions. *These characteristics have been adequately demonstrated through tests completed to date.*

3. **Structure Strength and Durability.** The unusual configuration and the extensive use of composite materials suggest that validation of structure adequacy should be the next criterion for proof of concept. That should include verification of predicted air loads in critical maneuvers, together with the common practice of taking the static test specimen to at least 80 percent of ultimate load and the fatigue specimen to at least the equivalent of one operational lifetime. *These performance characteristics should be adequately demonstrated by the end of this year.*
4. **Weapon Release.** The unusual vehicle configuration suggests that weapon carriage and clean separation should be demonstrated. *The necessary tests are expected to be performed during the first half of 1992.*

Some readers might be startled by the absence of entries in this list dealing with electronic system performance. While it is unreasonable to expect the offensive and defensive electronic suites to be trouble free, the testing of those systems already accomplished on flying test beds seems adequate to assure they can be made to perform at an acceptable level.[2]

Provided the present test schedule is met, by some time during the first half of calendar year 1992 there should be considerable confidence that the basic system concept of the B-2 is sound.

---

[2] This conclusion is based on results of a still-classified antecedent RAND study performed in 1987, and is specific to the B-2 program at this point late in the overall development process. It should not be interpreted as a general conclusion; in fact, at the beginning of any modern combat aircraft program the electronic system must be considered among the elements deserving special attention during both development and test.

MINIMIZATION OF TOTAL PROCUREMENT COST

Presuming that proof of concept has been demonstrated, should more testing be required before authorizing high-rate production? The answer to that question is based on the observation that most design flaws discovered during test can be fixed, but that occasionally (Type 2 flaws) the fix is expensive. If production is authorized early, then several units might be produced with the flawed component, requiring expensive retrofit after the flaw is discovered and corrected. That argues for extensive testing before production. However, delaying production is itself expensive. That leads to the second decision criterion:

- To minimize total procurement cost, the expected cost of correcting flaws not yet discovered should not exceed the savings achieved through prompt start of high-rate production.

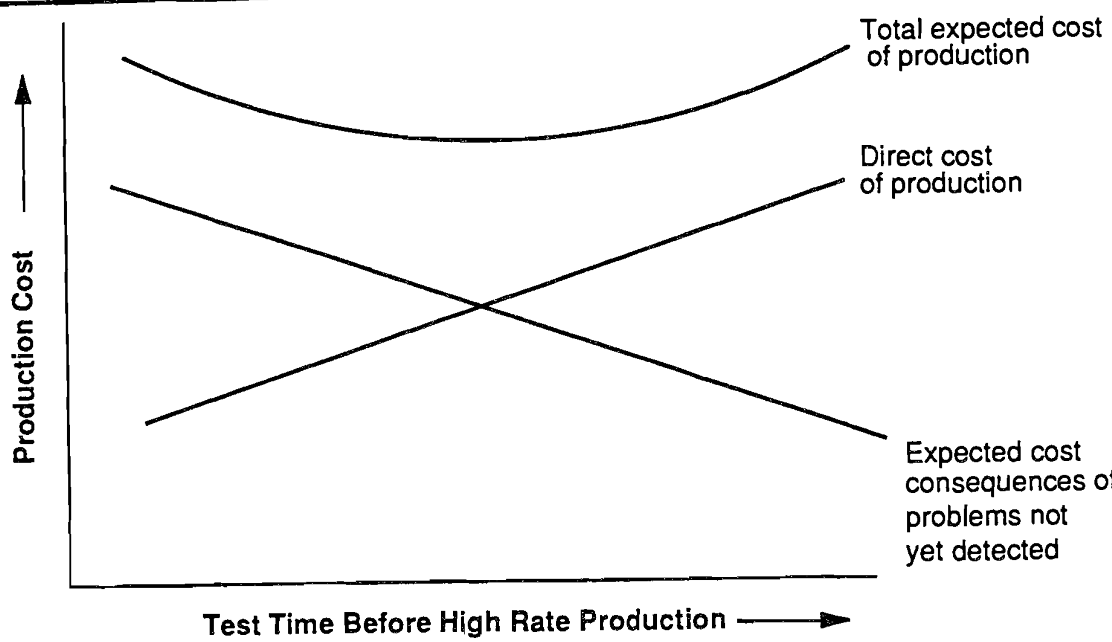
This notion can be illustrated by the simple diagram in Fig. 6. As high-rate production is delayed in order to perform more testing, the direct cost of production increases, simply because more years of overhead and other fixed costs are incurred. However, the additional testing should reduce the risk of unknown problems, and thus reduce the cost consequences of those problems. The sum of those two cost elements, which we define here as the *total* production cost, should have a minimum value at some point in time. Here we describe and apply a method for estimating the location of that minimum cost point.

Direct Cost of Production Delays

We start the process of estimating the cost consequences of delaying high-rate production by establishing a series of optional production rate profiles that, while not identical to any actual B-2 program schedules, are close enough for our purposes. We assume a Low Rate Initial Production (LRIP) rate of two units per year, which is characteristic of actual appropriations in recent years. Following a decision to produce at higher rates, a build-up sequence of four and six

Fig 6

Two Sources of Production Cost



units would occur in successive years, leading to a sustained maximum rate of 12 per year. We examined several such profiles, with each successive one containing an additional year of low-rate production.

We estimated the total flyaway cost for each of the alternative production profiles, using standard parametric cost estimation procedures.[3] *Each year of delay in moving to rates beyond LRIP costs at least $400 million dollars.*[4]

---

[3]Note that these are generalized estimates based on our somewhat arbitrary authorization profiles. Therefore, the cost values do not correspond to specific budget authorizations or requests for future authorizations, but the magnitudes and trends are consistent with official values.

[4]This estimate is based on broad military aircraft industry averages for overhead rates and fixed costs. We recognize that the B-2 program is likely to experience higher levels of such fixed costs, which would strengthen our conclusions regarding the cost benefits of early production go-ahead.

Cost Consequences of Risk

Given the estimated cost of delaying production in order to gain more test information, the next step in estimating the total expected production cost is to determine the likely cost consequences of the risk remaining at each point in time during the test program. The "risk cost" is defined (Fig. 7) as the probability of some remaining undetected flaws, times the expected cost of correcting those flaws and retrofitting any previously-completed production items. Each of those elements can be estimated on the basis of previous experience.

Probability of Remaining Flaws. In each of the five previous airplane development programs that we examined, we tried to identify the problems satisfying the "Type 2" criteria:

1. The design discrepancy had to be identified during the flight test phase. This eliminated any problems that had been identified during design or component test prior to fight test.

Fig. 7

Cost Consequences of Risks

Total Cost = Production Cost + Risk Cost

Risk Cost =	**Risk** Probability that additional tests or operational service will reveal new flaws that require corrective action	X	**Corrective costs** Cost of correcting deficiencies in units already funded when flaws were revealed

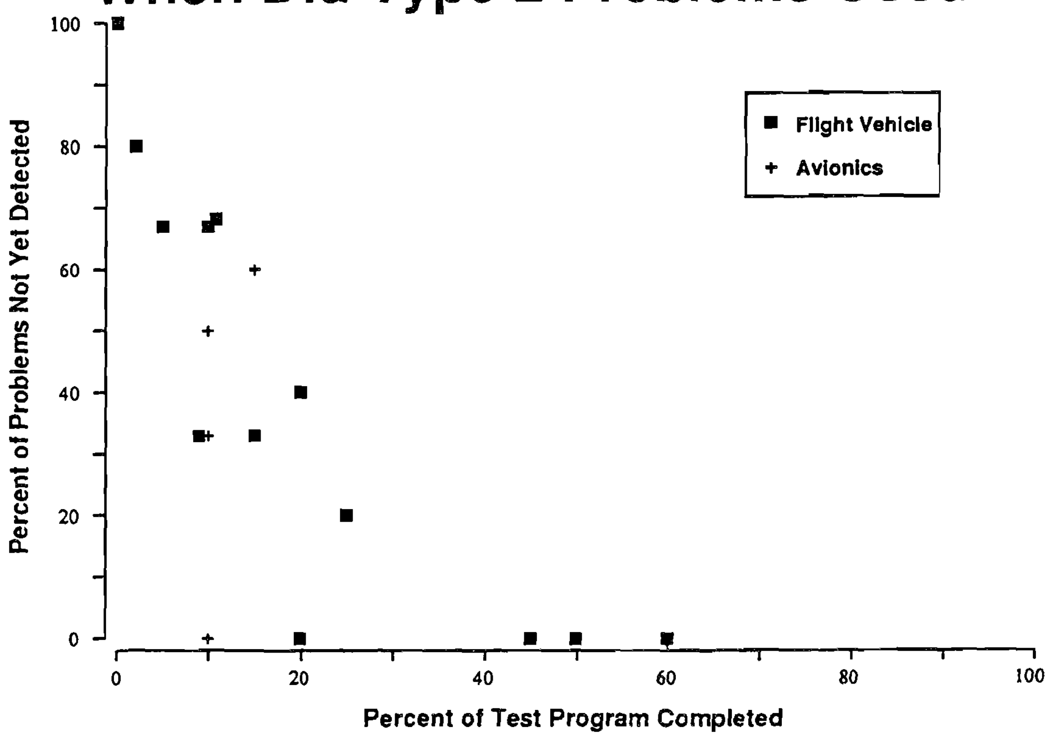
2. The discrepancy had a significant effect on the basic mission capability of the system, and therefore had to be fixed.
3. The discrepancy had to be either expensive or time consuming to fix, so that it required changing the schedule or the budget of the project to a significant degree.

We found only a few such problems in each of the programs studied.

The data are plotted in Fig. 8. It is apparent that the shape of the risk reduction trend in this data set more closely compares with Curve A in Fig. 1 than Curve B. By roughly mid-way in the test program there is a very low probability that an important flaw remains undetected. Even without further quantitative analysis, we believe this provides justificiation for rejecting the argument that production decisions should be delayed until "all testing has been completed."

Fig. 8

When Did Type 2 Problems Occur?



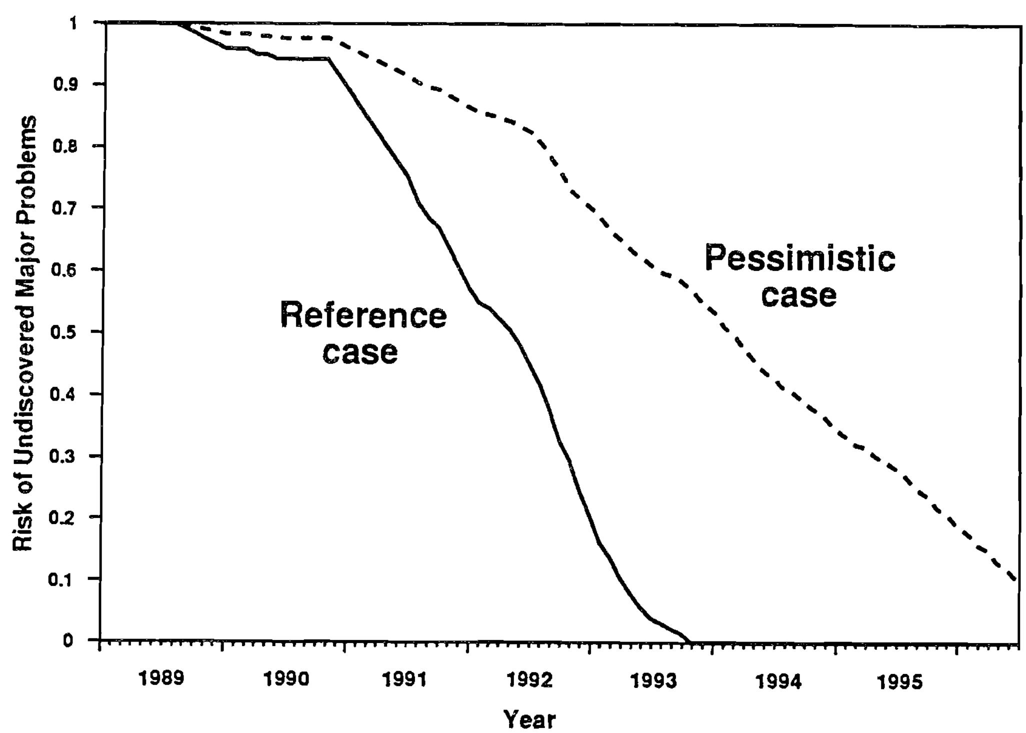
These data are plotted against percent of test program completed. By mapping the data onto the B-2 flight test plan, we can create a plot of expected-risk profile vs. calendar date. That yields the "Reference Case" program risk profile shown in Fig. 9.

That process was repeated, using the arbitrary assumption that the B-2 program might take twice as much testing as historical evidence would indicate is needed to reveal major problems. The resulting "Pessimistic Case" risk profile is also shown in Fig. 9.

Cost of Correcting Flaws. The second part of the "risk cost" model shown in Fig. 7 requires us to estimate the cost of correcting flaws that might be identified in the future. A search of the literature and available cost data bases revealed only a very few instances where such costs could even be roughly inferred. The worst case we found was the cost of replacing the entire wing on the C-5A, which cost about 16% of

Fig. 9

Predicted B-2 Risk Profiles



the original total airplane flyaway cost (using constant-year dollars). Other data points all were in the range of a few percent of original production cost. We elected to use in our calculations a very conservative and pessimistic value of twenty percent of original cost.

We would expect the cost of correcting flaws to vary with the time required to perform corrective engineering and to introduce the new design into the production line. That averages about three years, with a range of one to five years encompassing almost all data points. We used a reference value of three years, and a pessimistic upper limit of five years.

ESTIMATION OF TOTAL PRODUCTION COST

Adding the direct cost of production and the cost consequences of risk yields the estimate of total production cost for each production profile. The results are shown in Fig. 10 where we plot the projected production cost for the assumed 70-airplane buy vs. the first year for which production beyond LRIP was authorized. The lower curve is the estimate of direct production cost, whereas the upper curve includes the cost consequences of risk that we estimate are associated with each of the different years in which higher-rate production is first authorized.

We included in our calculations a series of cases in which we assumed that authorization for high-rate production had occurred in prior years. We did that to show a more complete picture of cost trends. In this chart only the last two points on the right, for FY92 and FY93, represent real opportunities at this time. We estimate that the minimum-cost point actually occurred several years ago. This is because the "risk cost" component of the total production cost is small compared to the direct production cost, so the trend of total expected cost is dominated by the cost of delaying production. Thus, *any delay (at least since FY90) in authorizing production beyond LRIP appears to cause an increase in the expected total production cost.*

We repeated the calculations using the arbitrary, very pessimistic estimate of the program risk profile. That yielded the cost estimates shown in Fig. 11. While the curve for total cost moves up and the

Fig. 10

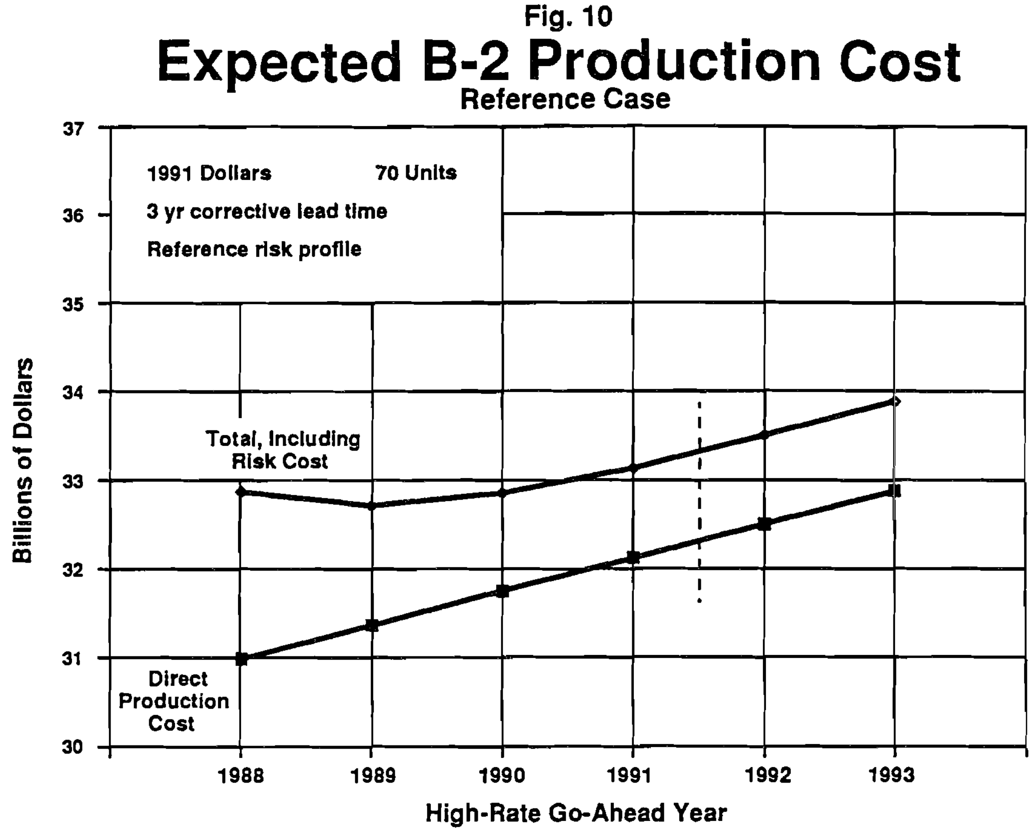
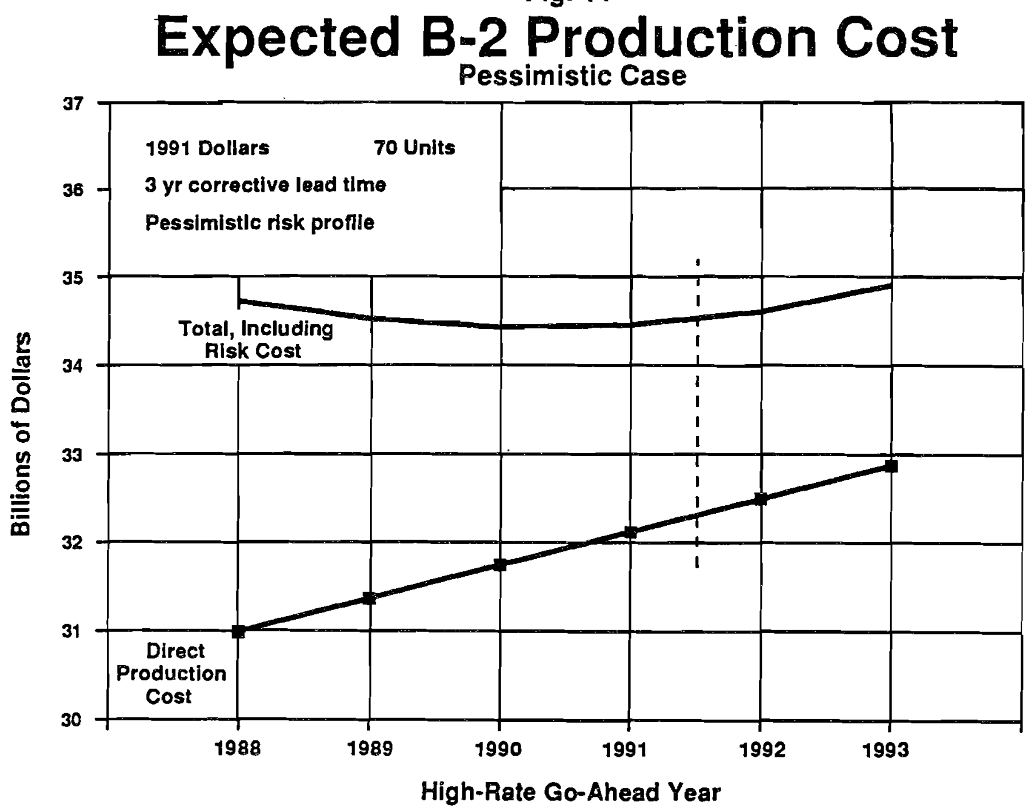

Expected B-2 Production Cost
Reference Case
1991 Dollars
70 Units
3 yr corrective lead time
Reference risk profile
Total, Including Risk Cost
Direct Production Cost
Billions of Dollars
37
36
35
34
33
32
31
30
1988
1989
1990
1991
1992
1993
High-Rate Go-Ahead Year

Fig. 11



minimum-cost point moves to the right, Fig. 11 still shows a steady trend of increasing total cost with each year of delay beyond FY92 in authorizing high-rate production.

All of the results shown above used constant FY91 dollars. To test the sensitivity of the conclusions to the escalation treatment, we repeated the analysis using then-year dollars, together with the most pessimistic combination of assumptions for the other variables (five-year lead time to correct problems, and the pessimistic risk profile). Results are similar to those for constant dollars, but with steeper slopes (i.e., larger cost penalties for each year of delay in authorizing high-rate production).

SUMMARY

This analysis supports two conclusions (Fig. 12), both predicated on the assumptions that the present test program will proceed approximately on schedule:

1. Before the middle of 1992 enough testing should have been accomplished to yield high confidence in the validity of the basic design concept for the B-2. Critical measurements will have been made on the general detection and survivability characteristics of the vehicle. A substantial amount of information should be available on the flight characteristics, structural integrity, and mission performance of the basic flight vehicle. At least initial data should be available on critical elements of the offensive systems. That information should provide substantial confidence that no subsequent problems of a program-threatening nature will be discovered.

Fig. 12

Conclusions

- High confidence in system concept by middle of CY 1992
  - Verification of detection/survivability
  - Flight vehicle well demonstrated
  - First looks at weapon release and mission equipment
- Delay beyond FY92 authorization for starting high-rate production likely to increase total production cost
- Above presumes that:
  - Test schedule is maintained
  - Problems are corrected in timely and effective manner

2. Under even the most pessimistic combinations of assumptions regarding the effect of problems not yet discovered, any delay beyond FY92 in authorizing high-rate production will likely cause an increase in total procurement cost of the system. This conclusion is valid even if the total production quantity is less than the current projection of 70 units.

It is interesting to note that in this case the goal of demonstrating confidence in system concept appears to the pacing item, because we almost certainly have passed the time when we could have achieved a minimum estimated production cost.

Every program is different, and there is no promise that the conclusions drawn regarding the B-2 program will apply to the next program. However, it does seem that a policy of delaying production until the design is highly refined and demonstrated through exhaustive flight testing is not necessarily the best course in every program.

RAND/P-7757 WHEN SHOULD WE START HIGH-RATE PRODUCTION OF THE B-2? Rich