Appendix A

The Sensitivity of Operational Aerospace Plane Costs To Required Confidence and Actual Reliability

Appendix A

The Sensitivity of Operational Aerospace Plane Costs To Required Confidence and Actual Reliability

The proposed experimental National Aero-Space Plane (NASP), the X-30, is being designed with a goal of 99.999 percent reliability—i.e., to have only 1 chance in 100,000 of failing catastrophically during a flight, assuming no human error. NASP program officials have said that later first-generation aerospace planes, which they term NASP-Derived Vehicles (NDVs), would probably be designed to have a similar reliability. Recognizing that the design reliability does not account for possible human error in maintaining and flying the vehicle, NASP program officials assume the actual reliability of a first-generation NDV would be lower—nominally 99.8 percent, but 99.9 percent in the "best case, " and 99.5 percent in the "worst case."[1]

NASP program officials assume there is little risk that an NDV of lower reliability would be flown on operational missions, because they expect NDV development would be halted if the proposed X-30, or a later prototype NDV, fails to demonstrate acceptable reliability (99.5 percent) in its flight test program. Thus, in their view, the cost at risk would be not billions of dollars of greater than expected failure costs but only those funds spent to develop and build X-30s.[2]

This argument hinges on a critical assumption: that if the test vehicles turn out to be unacceptably unreliable, the test program will detect that fact with high confidence. The validity of this assumption cannot yet be decided, because the NASP program office has not yet specified what kind of confidence (statistical or subjective) they require, nor how much is enough, nor how it will be calculated from test results. These details are important, because a test program cannot determine the reliability precisely. The test flights might be a lucky streak, or an unlucky streak; the actual reliability could differ significantly from the successful percentage of test flights. However, a properly designed test program can determine the confidence level with which the required reliability has been demonstrated.

In choosing a required confidence level, NASP program officials face a dilemma—as would any manager of a launch vehicle development program: Requiring too little confidence could allow acceptance of a vehicle that is actually unacceptably unreliable; operational vehicles of similar design and reliability would probably fail often enough to incur staggering failure costs. If, on the other hand, too much confidence is required, a vehicle that is actually highly reliable might be rejected, and the savings potentially realizable by using operational vehicles of similar design and reliability would be forfeit.[3]

NASP program officials must also choose the *type of confidence* to require. They could require statistical confidence to be demonstrated, or they could calculate the confidence level by Bayesian inference, which would use the results of the flight tests to update a subjective prior probability distribution over possible values of reliability (see box A-A).[4] The former choice would require a very large number of flights to demonstrate the required reliability with high statistical confidence.[5] A problem with the latter choice is that there would be risks of optimism and pessimism. If the prior distribution is optimistic, the reliability might be low but the vehicle would be accepted and later losses incurred. If pessimistic, reliability might be high but the vehicle would be rejected and potential savings unrealized.

Because the type and level of confidence with which 99.5 percent reliability must be demonstrated has not been specified, and because the actual reliability will never be known precisely, this appendix shows how the life-cycle costs of acquiring and operating a mixed fleet of launch vehicles-including NDVs, if accepted—would depend on the type and level of confidence required in testing and on the actual reliability.

Cost Estimates, If Statistical Confidence Is Required

The type and level of confidence required and the actual reliability determine the probability that the test program will be successful; if it is, NDVs will be acquired and operated, and their actual reliability will affect the failure costs incurred.

Figure A-1 shows estimates of the probabilities with which test vehicles of various reliabilities would demonstrate 99,5 percent reliability with various levels of statistical confidence in 100 test flights. If 40 percent

[1] NASP Joint Program Office staff, personal communication, Jan. 18, 1990.

[2] Furthermore, they expect that the value of NDV technology "spun off" to other applications such as aircraft and launch vehicles would compensate for some of the cost at risk.

[3] If more test flights were conducted, a reliable Vehicle could demonstrate acceptable reliability with acceptable confidence and allow these potential savings to be realized, but against this must be weighed the expense and delay of the extra tests.

[4] National Research Council, *Post-Challenger Evaluation of Space Shuttle Risk Assessment and Management* (Washington, DC: National Academy *Press,* January, 1988), app. D.

[5] One hundred test flights, if all successful, would provide only 39.4 percent statistical confidence in a 99.5 percent lower confidence bound on reliability.

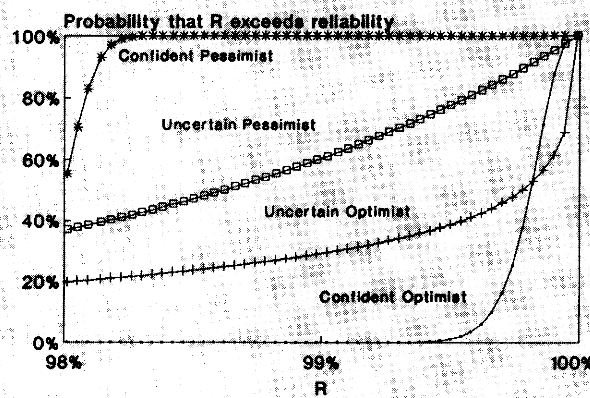
Box A-A—Estimation of Reliability by Bayesian Inference

In classical probability theory, a system is assumed to *have* a definite reliability, even though the reliability is not known precisely. The percentage of tests (or uses) that are successful is an estimate of the reliability, and statistical confidence bounds indicate the uncertainty of estimates. If no tests have been done, reliability cannot be estimated by statistical methods.

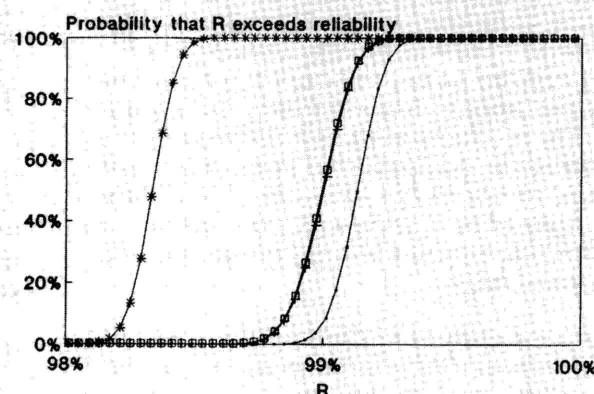
In contrast, the Bayesian view is that the reliability of a system may be 10 percent, or 90 percent, or any other possible value. A person's beliefs and uncertainty about the reliability, even before tests have been conducted, may be expressed by a set of probabilities: the probability that the reliability is less than 10 percent, the probability that it is less than 20 percent, and so forth. Collectively, such probabilities specify what is called a cumulative probability distribution over reliability. Because it represents one person's beliefs about the reliability, the distribution is called a subjective probability distribution (SPD). Another person may have different beliefs about the reliability of the same system; that person's beliefs would be represented by a different SPD. The Bayesian interpretation of probability does not require the SPDs of different subjects to agree, but it does require each subject's SPD to be self-consistent. For example, an SPD cannot specify that a subject believes the reliability to be less than 10 percent with 50 percent probability and less than 20 percent with 30 percent probability!

Four Subjects' Subjective Probability Distributions Over Reliability

A Priori
(Before testing)



A Posteriori
(After 9,900 successes in 10,000 tests)

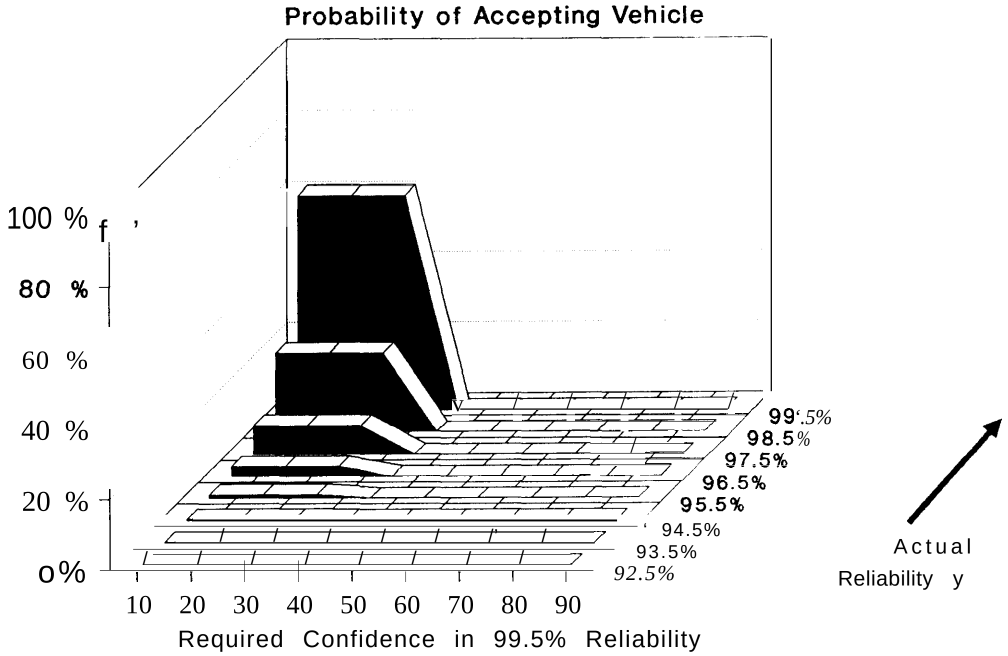


SPDs which are not based on actual tests of the system are called prior SPDs: they are SPDs estimated prior to testing. Prior SPDs may be based on expert judgement, considering tests of subsystems or analogous systems. They may also reflect guesses, hunches, mysticism, or complete ignorance. When test data become available, SPDs may be updated; however—*and herein lies the value of Bayesian inference*—each SPD must be updated in a logically consistent manner, using Bayes's theorem, which is stated and proved in many textbooks on probability. SPDs updated in this manner are called posterior SPDs: they are SPDs estimated *a posteriori* (after testing).

Differing prior SPDs become more similar after Bayesian updating, as the figures below illustrate. The figure on the left shows portions of the prior SPDs of four subjects, who are identified as "Confident Pessimists," "Uncertain Pessimist," "Uncertain Optimist," and "Confident Optimist." The prior SPD of the Confident Pessimist indicates that the Confident Pessimist is about 55 percent confident that the reliability is less than 98 percent and almost 100 percent confident that the reliability is less than 98.25 percent. The Uncertain Pessimist is less than 40 percent confident that the reliability is less than 98 percent even though his SPD implies the same expected value of reliability—98 percent-as does the SPD of the Confident Pessimist (the expected value of reliability cannot be read from a graph of the SPD; it must be calculated from the SPD). Similar readings of points on the SPDs of the Uncertain optimist and the Confident Optimist reveal why they are so named.

The figure on the right shows portions of the posterior SPDs of the same four subjects, updated after 9,900 successes have been observed in 10,000 tests-a success rate of 99 percent. The posterior SPDs of the Uncertain Pessimist and the Uncertain Optimist are almost indistinguishable; influenced by the test results in a logically consistent manner, their prior SPDs have converged, that of the Uncertain Pessimist (who expected 98 percent reliability *a priori)* becoming more optimistic, and that of the Uncertain Optimist (who expected **99.8 percent reliability** *a priori)* becoming more *pessimistic.* The SPDs of the Confident Pessimist and the Confident Optimist also became more similar but did not fully converge; the high confidence implicit in the prior SPDs (and apparent as steep slopes in the graphs) caused them to be influenced less by the test results.

Figure A-1—Probability That Test Vehicles Will Demonstrate Acceptable Reliability If Statistical Confidence Is Required



Based on 100 Monte Carlo samples.

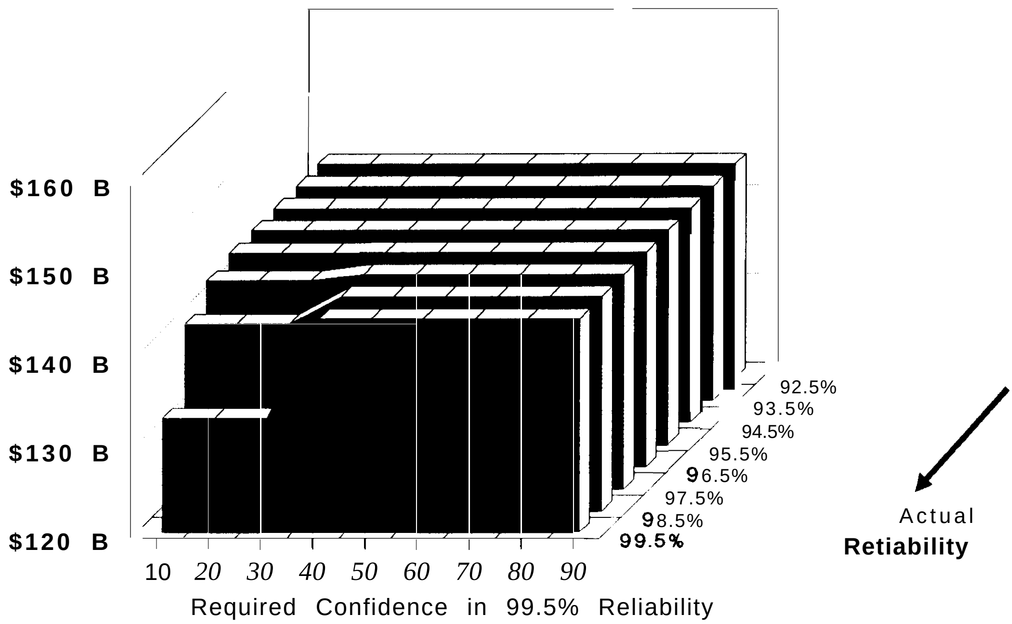
SOURCE: Office of Technology Assessment, 1990.

statistical confidence, or more, is required, a vehicle would be rejected even if all flights were successful. If, however, only 30 percent statistical confidence is required, a 99.5-percent reliable vehicle would be accepted with a probability of about 80 percent.

Figure A-2 shows how the expected present value of the life-cycle costs of flying the missions in OTA's Low-Growth mission model[6] through the year 2020 would depend on actual reliability and required confidence, if statistical confidence is required. The greater the

[6] U.S. Congress, Office of Technology Assessment, *Launch Options for the Future—A Buyer's* GM'&, OTA-ISC-383 (Washington, DC: U.S. Government Printing Office, July 1988).

Figure A-2—Expected Present Value of Mixed-Fleet Life-Cycle Cost If Statistical Confidence Is Required and NDV Costs Are As Estimated by NASP JPO



Based on 100 Monte Carlo samples.

SOURCE: Office of Technology Assessment and National Aero-Space Plane Joint Program Office.

X-30'S reliability turns out to be, the greater the savings can be.[7] If managers require at least 10 percent statistical confidence, they risk little or nothing compared to the costs if NDVs were not attempted (figure 1-2) if the X-30'S reliability turns out to be low, because there is little chance that an unreliable vehicle would be accepted, and no NDV development costs are assumed to be incurred before the decision on whether to proceed with NDV development.[8]

Methodology

OTA calculated these estimates assuming that NDV-related costs are uniformly distributed over ranges estimated by the NASP Joint Program Office (JPO), which are shown in table A-1. OTA also assumed:

1. X-30 development costs are sunk costs.
2. The test program will consist of 100 test flights *to orbit and back;* other (e.g., suborbital) test flights may be conducted but will not be used to estimate reliability on orbital flights.[9]
3. During the test program, the X-30S will not be modified in any way, or operated in different ways, that would make reliability differ from flight to flight.[10]
4. The government will decide in 2000 whether to proceed with NDV development,[11] based on whether

[7] The actual reliability of operational NDVs is assumed to be the same as that of the X-30s or prototype NDVs used to demonstrate reliability in the flight test program.

[8] These estimates exclude the costs of developing the x-30 (or other prototype); if Congress decided now to forego development of an NDV, it could save a few billion dollars by hatting the NASP program. See U.S. Congress, Congressional Budget Office, *Reducing the Deficit: Spending and Revenue Options* (Washington, DC: U.S. Government Printing Office, 1990), pp. 68-70.

[9] The JPO actually plans only about 75 to 100 test flights, most of them suborbital. Data from suborbital test flights and ground tests of vehicle systems and components could be used to estimate reliability on orbital flights, but this would require developing a component-level reliability model that describes how component failures could cause vehicle loss and how component failure probabilities depend on details of vehicle assembly, maintenance, and operation (e.g., on the speeds and altitudes at which the vehicle has flown).

[10] This assumption simplifies analysis. In fact, the X-30s could be modified—e.g., after a failure—in an attempt to increase reliability, but "wiping the slate clean" late in the test program might reduce confidence that the required reliability has been demonstrated.

[11] Some have suggested that if th, X-30 is successful, NDVs might be developed privately to service the market for space tourism. For estimates Of the demand for round trips to orbit as a function of ticket price, see DoD & NASA, *National Space Transportation and Support Study 1995-20)0, Annex* B: Civil Needs Data Base, Version 1.1, vol. I—Summary Report, Mar. 16, 1986, pp. 3-3 1—3-32, and Gordon R. Woodcock, "Economics on the Space Frontier: Can We Afford It?," SS[*Update* (Princeton, NJ: Space Studies Institute, May/June 1987).

Table A-l—Ranges of Costs Estimated by NASP Joint Program Office (index **year not specified)**

	Best Case	Nominal	Worst Case
Development	$3,000 M[a]	$4,000 M	$6,000 M
Facilities (per NDV)	$25 M	$50 M	$75 M
Production (first NDV)	$700 M	$800M	$1,100 M
(% learning[a])	85 %	90 %	95 o/o
Operations (per NDV-year)	$10 M	$15 M	$30 M
(per flight)	$0.8 M	$1.2 M	$2.2 M
Failures (per failure)	$1,500 M	$2,000 M	$2,500 M

a M = million
b I.e., the incremental unit cost of the nth NDV will be (%/1 OF)@@ times the incremental unit cost of the first NDV, where % is the percentage learning.

SOURCE: NASP Interagency Office, 1990

the required reliability (assumed to be 99.5 percent) is demonstrated with a specified *statistical* confidence.

5. If the government decides not to proceed with NDV development, the missions in the mission model will be flown by Shuttles, Titan IVs, and Medium Launch Vehicles. Construction of facilities required for launching Titan IVs at the rates required (which began earlier as a hedge) will continue.
6. If instead the government proceeds with NDV development,
   (a) Construction of only those Titan IV facilities required for launching at the rates required to complement the NDV will continue, possibly after a delay.[12]
   (b) The actual reliability of operational NDVs will be the same as that of the X-30s or prototype NDVs used to demonstrate reliability in the flight test Program.[13]
   (c) Enough NDVs will be procured to make the probability of losing them all to attrition no greater than one percent, assuming a reliability of 99.5 percent.[14]
   (d) NDVs will fly all the reamed missions in OTA's "Low-Growth" mission model on a 1:1 basis (i.e., one NDV flight substituting for one Shuttle flight) and half of Titan missions 1:1, beginning the year of initial operational capability, which OTA assumes will be 2005.
   (e) [f and when all NDVs are lost to attrition, another NDV will be procured at the same incremental unit cost as the first NDV.[15]

The probability that X-30S or prototype NDVs would demonstrate the required reliability (99.5 percent) with the required confidence during the flight test program was calculated for each combination of actual reliability and required confidence considered. This probability-the "acceptance probability ''-was used in calculating the life-cycle cost of the mixed fleet. Titan and Shuttle costs depend on the number of missions Titans and Shuttles are required to fly, which depends on whether NDVs are accepted and used to complement Titans and supersede Shuttles.

For the case in which they are, the costs of NDV development, facilities, production, operation, and failures are estimated by Monte-Carlo techniques-i. e., random-event simulation. For each of 100 scenarios, values for each of the uncertain costs in table A-1 were generated pseudorandomly[16] and used to calculate the life-cycle costs of the NDV fleet. The number of operational failures in each year was also generated pseudo-randomly, based on the actual reliability assumed, and used to calculate NDV failure costs. For each value of actual reliability y considered, the difference between the 70th percentile of NDV costs and the median value of NDV costs was used as the "STAS cost risk''--i.e., the cost risk as defined in the Space Transportation Architecture Study[17]—for the NDV fleet.

Sensitivity to Greater-than-Expected NDV Costs

To gauge the sensitivity of the estimates in figure A-2 to greater-than-expected NDV costs, OTA estimated costs by the same procedure but assumed NDV-related costs are uniformly distributed over the ranges in table A-2. The lower bounds of these ranges areas estimated by

[12] Because they may not be needed as soon, and delaying expenditures for facilities allows them to be more heavily discounted.

[13] Operational NDVs could be designed to differ from the X-30s or prototype NDVs—e.g., to have more engines-with the intent of making them more reliable. If so designed, detailed reliability models (footnote 9) of both X-30s and operational NDVs would be needed to estimate operational NDV reliability on the basis of X-30 flight tests. If operational NDVs differ significantly from X-30s, X-30 flight tests may provide little information about NDV reliability; in any case, the updating procedure would be much more complicated than updating based solely on test flights of similar vehicles under similar conditions.

[14] According to this criterion, the fleet size should be eight for the Low-Growth mission model (534 NDV flights).

[15] This is Optimistic; it neglects procurement delay and the remote possibility that one NDV may be required to fly more flights than the NASP JPO assumes it will be able to: 250 to 500, but nominally 400.

[16] The costs were assumed to be distributed uniformly between the worst-case and best-case values in table A-1.

[17] U.S. Congress, Office of Technology Assessment, op. cit.. footnote 6.

the NASP JPO (table A-l); the upper bounds of these ranges are twice the upper bounds of the ranges estimated by the NASP JPO (in the case of percentage learning, the upper bound is twice as close to 100 percent). Figure A-3 shows the resulting cost estimates.

Cost Estimates, If Subjective Confidence Is Allowed

OTA has also estimated savings and losses for cases in which the government decides whether to proceed with NDV development based on the *subjective* (rather than *statistical)* confidence with which the required reliability (assumed to be 99.5 percent) is demonstrated. The confidence level is calculated by Bayesian inference. For illustration, the estimates were calculated assuming the prior distribution of the "Confident Optimist" of box A-A, which implies an expected reliability of 99.8 percent-the same nominal reliability estimated by the NASP JPO. Figure A-4 shows the proto-NDV acceptance probabilities, and figure A-5 the life-cycle costs, estimated under these assumptions, assuming OTA's Low-Growth mission model and NDV-related costs uniformly distributed over ranges estimated by the NASP JPO (table A-l).

To gauge the sensitivity of the estimates in figure A-5 to greater-than-expected NDV costs, OTA estimated costs by the same procedure but assumed that NDV-related costs are uniformly distributed over the ranges in table A-2. Figure A-6 shows the life-cycle costs estimated under these assumptions.

Figure A-5 shows that if NDV costs areas estimated by the NASP JPO, there is little risk that an unacceptably unreliable vehicle would be accepted and, as a consequence, the mixed-fleet life-cycle cost (figure A-5) would exceed that of the current mixed fleet (the Titan-IV option in figure 1-2).

Table A-2—Ranges of Costs Assumed by OTA for Sensitivity Analysis

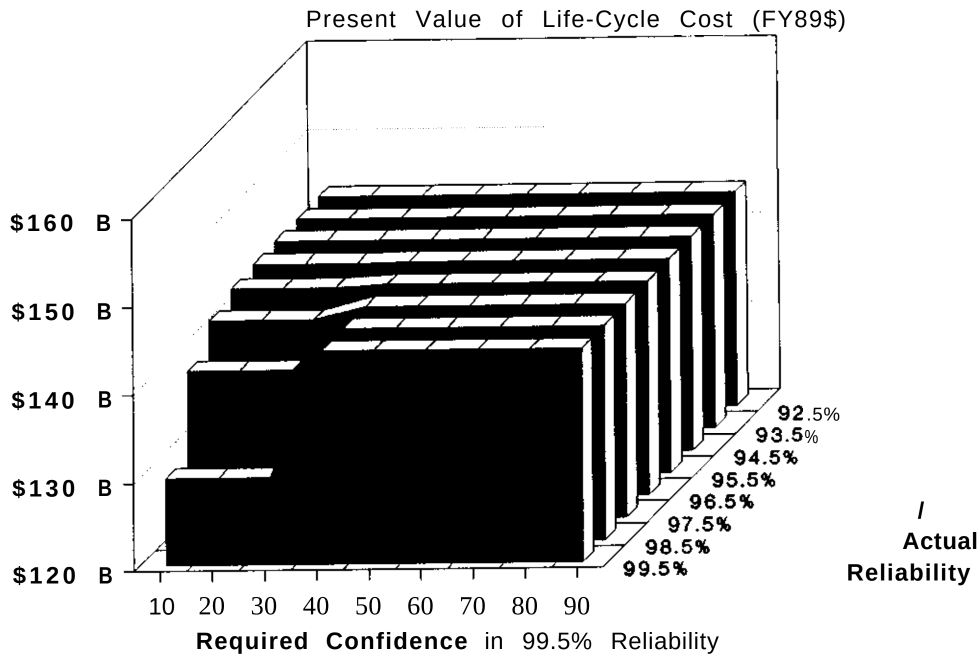
	Best Case	Worst Case
Development	$3,000 M[a]	$12,000 M
Facilities (per NDV)		$150 M
Production (first NDV)	$700 M	$2,200 M
(% learning)	857.	97.570
Operations (per NDV-year)	$10 M	$60 M
(per flight)	$0.8 M	$4.4 M
Failures (per failure).	$1,500 M	$5,000 M

[a]M = million.

SOURCE: Office of Technology Assessment, 1990.

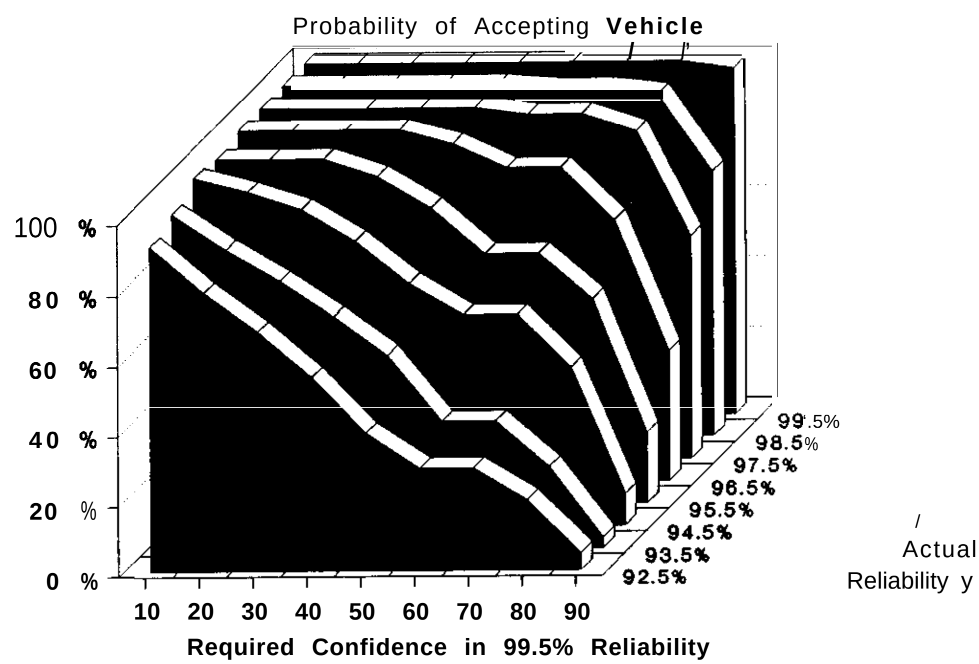
However, if NDV costs can range up to twice the upper bounds estimated by the NASP JPO, figure A-6 shows that there could be a significant risk of loss caused by accepting an unacceptably unreliable vehicle. For example, if only 10 percent confidence is required and actual reliability turns out to be 92.5 percent, the median life-cycle cost would be about $16 billion more than if the NDV were rejected, or not attempted, because the failures that would occur in the test flight program would probably not reduce the confidence in NDV reliability (over 98.9 percent, a priori) below 10 percent. If 90 percent confidence were required (see figure A-6), or 10 percent statistical confidence were required (see figure A-3), or prior confidence in NDV reliability were lower, this risk could be made negligible, but this would also reduce the probability of accepting, and benefiting from, a reliable NDV,

Figure A-3-Expected Present Value of Mixed-Fleet Life-Cycle Cost If Statistical Confidence Is Required and NDV Costs May Be 2X NASP JPO Estimates



Based on 100 Monte Carlo samples.

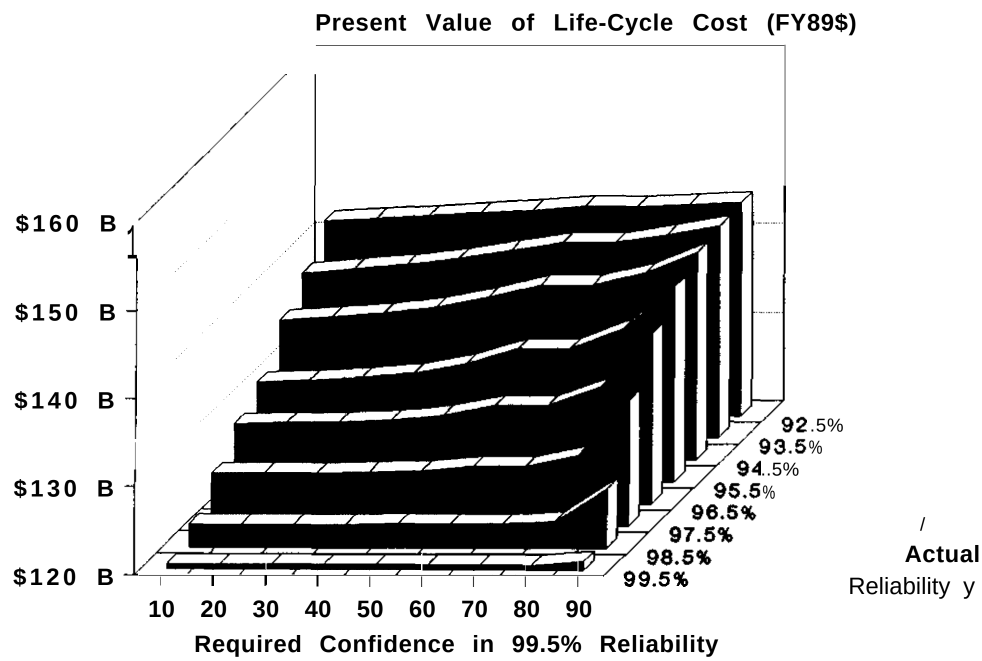
Figure A-4--Probability That Test Vehicles Will Demonstrate Acceptable Reliability If Subjective Confidence Is Allowed



Based on 100 Monte Carlo samples.

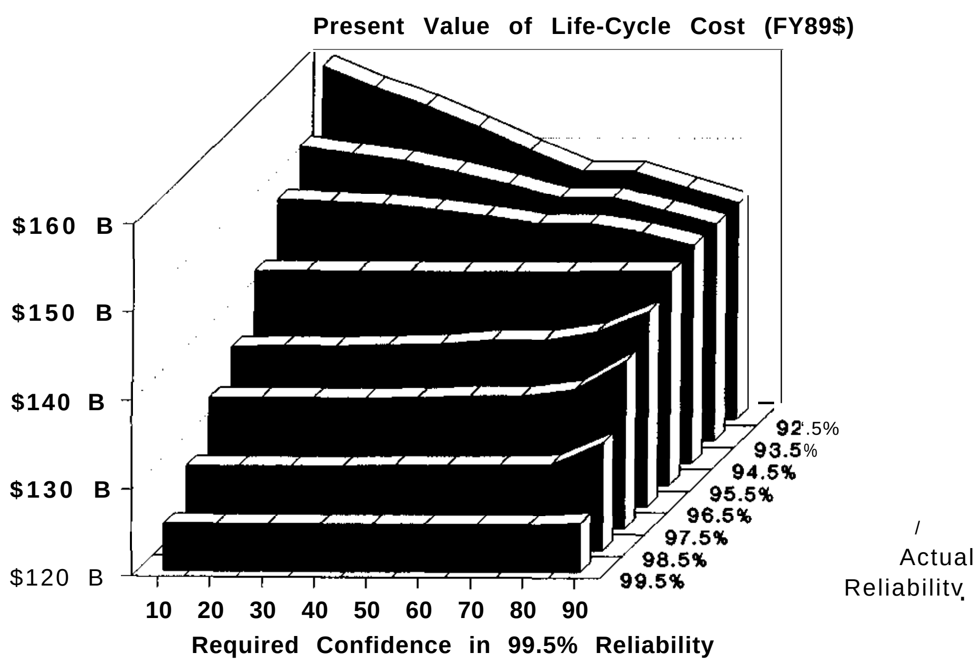
SOURCE: Office of Technology Assessment and National Aero-Space Plane Joint Program Office, 1990.

Figure A-5-Expected Present Value of Mixed-Fleet Life-Cycle Cost If Subjective Confidence Is Allowed and NDV Costs Are As Estimated by NASP JPO



Based on 100 Monte Carlo samples.

Figure A-6--Expected Present Value of Mixed-Fleet Life-Cycle Cost If Subjective Confidence Is Allowed and NDV Costs May Be 2X NASP JPO Estimates



Based on 100 Monte Carlo samples.

SOURCE: Office of Technology Assessment and National Aero-Space Plane Joint Program Office, 1990.