Assessing the Impact of Confidence Levels in Funding and Budgeting NASA Science Missions

Robert Bitten
The Aerospace Corporation
2310 E. El Segundo Blvd.
El Segundo, California 90245
(310) 336-1917
robert.e.bitten@aero.org

Charles Hunt
NASA Headquarters Cost Analysis Division
300 E Street SW
Washington, DC 20456
(202) 358-0803
charles.d.hunt@nasa.gov

***Abstract*—NASA Procedural Requirement 7120.5E requires that NASA utilize joint confidence level (JCL) analysis to set budget and funding guidelines for projects within its portfolio of missions. This paper addresses analysis that was conducted to determine the effect of different confidence levels on the performance of different cases of mission portfolios. The results show that the most effective confidence level varies depending on the case but that NASA's current policy is best for a traditional mix of missions and may vary significantly for a single mission portfolio. An overview of the policy, methodology, cases and results are discussed.**

TABLE OF CONTENTS

1. INTRODUCTION

NASA formalized the practice of producing probabilistic cost estimating and confidence level budgeting in 2006. The "portfolio effect" was one of the primary drivers to NASA's original probabilistic cost policy implementation. [1] The "portfolio effect" is defined as the tendency for the risk on a well-diversified holding of investments to fall below the risk of most and sometimes all of its individual components.

Using the portfolio principles, individual project confidence levels can roll up to higher or lower confidence levels at the program level. Applied to an Agency's mission portfolio, the portfolio effect can be applied to understanding the relationship between confidence at the Agency's (or Program's) level and confidence at individual project level. For the portfolio effect to work, projects within a program (within a portfolio) that turn out not to require their entire original budget must be managed in such a way that their unused budget is available to other projects. These unneeded resources are then available to be used for projects which exceeded their budgets. This allows decision makers to fund projects at lower confidence levels while achieving higher confidence levels from an Agency or Program viewpoint. [2]

NASA's original policy was focused solely on cost confidence only and did not address schedule confidence. It also assumed that the Agency would be able to take maximum advantage of the portfolio effect. Because of these assumptions, two aspects of policy could be improved. Firstly, it was realized that a project's plan is programmatic risk posture needs to be measured with both cost and schedule confidence levels. A project's optimistic schedule can directly influence a conservative cost estimate. Secondly, original policy did not address the concept of risk moral hazard. Moral hazard, is when people tend to adjust their behavior in response to perceived level of risk. A classic example of moral hazard deals with how drivers respond to antilock brakes – they have been shown to drive faster, follow closer, and brake later. Essentially the perceived risk posture of the environment changes how the drive behaves.

How does this apply to NASA CL policy? It is reasonable to assume that projects will not follow the typical "portfolio effect" assumptions. If you give a project manager more funds, the project manager will have a tendency to find a way to spend – e.g. an additional test to ensure technical success. This assumption, that projects will not give back their unused budget/schedule to the program/Agency, counteracts the portfolio effect.

NASA policy evolved to account for the two deficiencies mentioned above. Firstly, NASA policy now includes both cost and schedule confidence level. This new policy is commonly referred to as Joint Cost and Schedule Confidence Level – or simply JCL. The JCL measures the probability that a project will be within cost AND schedule. Secondly, NASA set up a tiered reserve strategy.

NASA's current budgeting policy, as stated in NASA Procedural Requirement (NPR) 7120.5E, is to budget projects at a 70 percent joint cost and schedule confidence level. [3] This policy is in response to historical cost overruns that were prevalent in NASA's history when previous budgets were primarily based on a project's initial estimate. [4, 5, 6] Figure 1 shows this change in policy in which projects are now budgeted at the 70% confidence level and funded at the 50% confidence level as opposed to the 20% confidence level at which, historically, NASA missions were budgeted and funded. This policy results in a project being funded at the Management Agreement while being budgeted at the Agency Baseline Commitment. This allows both the project and NASA Headquarters (HQ) to hold Unallocated Future Expenses (UFE) which are costs that are expected to be incurred but cannot yet be allocated to a specific element of a program's or project's plan. The UFE managed above the project provides more flexibility to NASA HQ to fund other projects that may exceed their Management Agreement within a given portfolio.

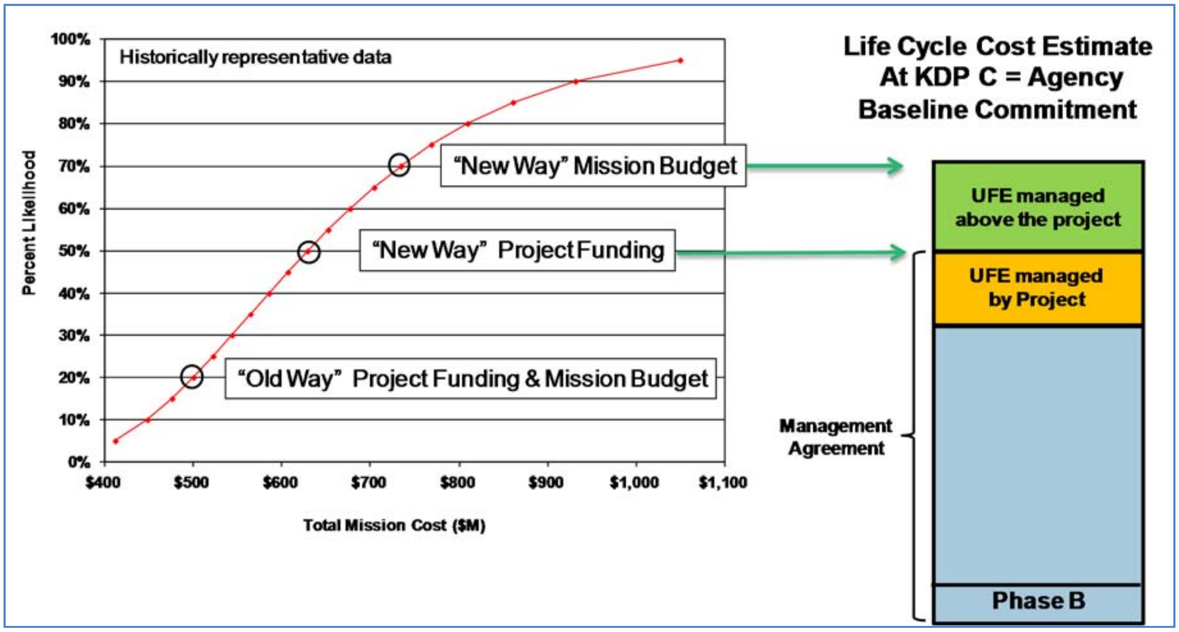


Figure 1: Simple Explanation of Current Budgeting and Funding Policy

2. Analysis Overview

Definition of Terms

This study was initiated by NASA's Cost Analysis Division (CAD) at NASA Headquarters to understand the potential impact of variations in the policy on a set of missions within a portfolio. For the study, two new terms were developed for identify the Budgeting Confidence Level (BCL), which is the Agency Baseline Commitment at which a mission is budgeted, and the Funding Confidence Level (FCL) which is the Management Agreement at which a project is funded. These terms can be shown graphically in Figure 2 and is the basis for the case matrix that was investigated for the study. The difference between BCL and FCL is the HQ Unallocated Future Expense (UFE) that is available to help fund overruns of other projects in the portfolio.

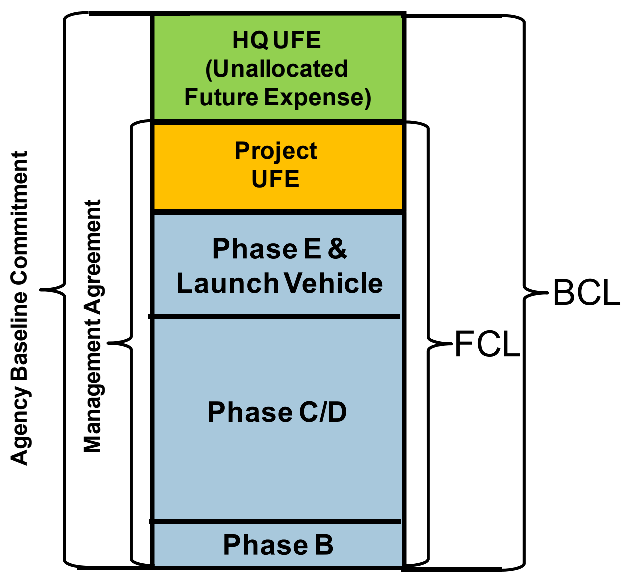


Figure 2. Definition of BCL and FCL

Case Matrix

A case matrix was defined to assess the varying levels of BCL and FCL that were to be studied. The basic case matrix, as shown in Figure 3, consists of different combinations of BCL and FCL. The current NASA policy is highlighted in the red box which the mission is funded to a 70% BCL and the project is funded to a 50% FCL.

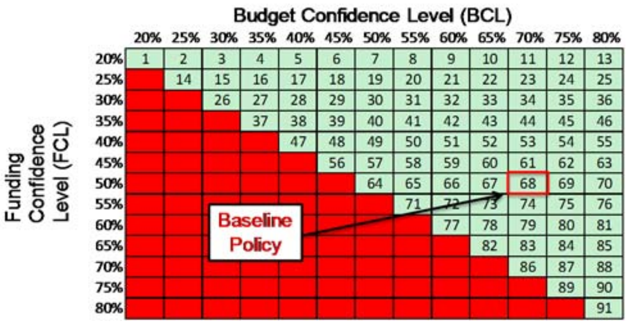


Figure 3. Study Case Matrix

Each BCL level in the case matrix requires that project planned start dates be changed to "pack in" the missions at a given portfolio budget. Each FCL case within a BCL only requires that project funding change commensurate with the confidence level of the FCL, thereby providing differing levels of HQ UFE. By definition, if BCL = FCL, then no HQ UFE is available to help cover projects that may overrun their budget. The HQ UFE can be seen graphically in Figure 4 which shows the UFE available in the annual budget, as identified by the white line, after all of the projects, as identified by the multicolored areas, are funded at their FCL.

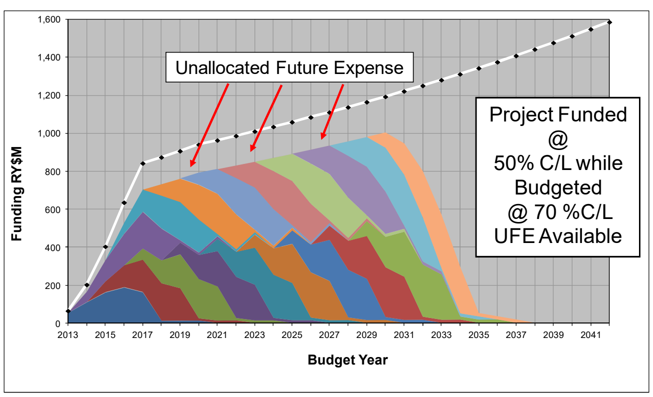


Figure 4. Graphical Depiction of UFE for a Case

Figures of Merit

To determine which case is the most effective, a series of Figures of Merit (FoM) were developed. These FoMs include:

1) Percentage portfolio cost growth over initial plan
2) Time to launch the first twenty planned missions
3) Percentage of planned missions launched in a 20 year period
4) Percentage of time that the planned missions exceed a 15% cost growth threshold breach requirement

The first FoM assesses the variable cost associated with an equal content within a given BCL. The second FoM assesses the variable time associated with that same equal content. The third FoM assesses the variable content over an equal period of time. The fourth FoM measures the overall volatility of the missions within the portfolio and is based upon the threshold breach requirements that are outlined in Section 103 of Public Law 109–155 which is requires that Congress be notified when "the development cost of the program is likely to exceed the estimate provided in the Baseline Report of the program by 15 percent or more." [7] For the purpose of the analysis, the threshold breach value is defined as the BCL plus 15% of the Phase C/D cost of the project not including the launch vehicle.

Methodology

To assess the FoMs, The Aerospace Corporation's Sand Chart Tool (SCT) was used. The Sand Chart Tool is a probabilistic simulation of budgets and costs that simulates a program's strategic response to internal or external events that cause cost and schedule to grow. It was developed with the specific intent to assess the effect of confidence levels on the interaction of multiple elements within a portfolio and have been used to assess program and portfolio performance. [8, 9] Algorithms are derived from historical data and experiences and can provide long-term program/portfolio analysis over a 10 to 20 years period. SCT is used to assess the domino effect for other projects in a program portfolio such that when one is stretch or delayed due to cost or schedule overruns, other projects in the portfolio are affected. This domino effect adds cost due to inefficiencies of starting & delaying projects. SCT applies real world penalties to projects, as shown in Figure 5, to modify other missions in the portfolio when other projects experience cost and schedule overruns. Applying these penalties within the simulation allows for a realistic assessment of interaction of multiple program elements or multiple missions within a given portfolio. Further description of the SCT can be found in reference 9.

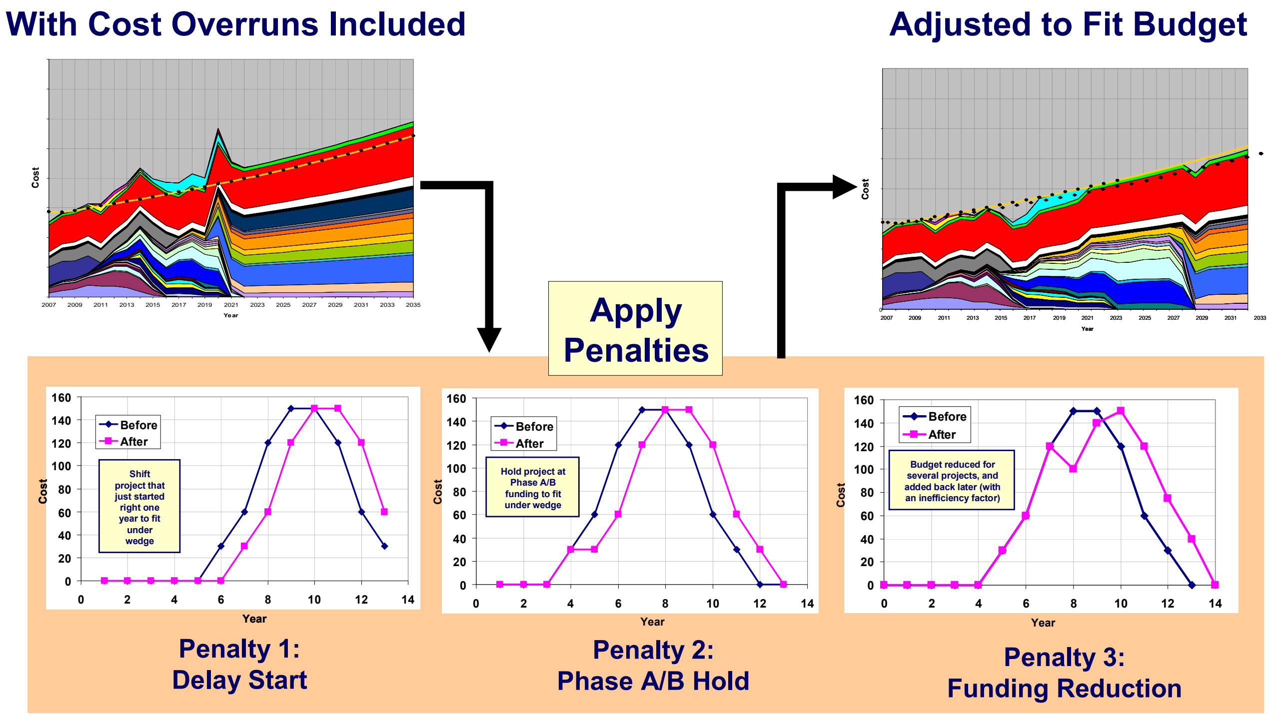


Figure 5: Penalties Included in the Sand Chart Tool for Aligning Funding with Available Budget

Scenario Overview

To ensure that the assessment was robust, multiple scenarios consisting of different portfolio were investigated. One way to measure different "types" of portfolios is by the Gini coefficient. The Gini coefficient (GC) measures the inequality among values of a frequency distribution. [10] Figure 6 shows that the GC is equal to the area marked A divided by the sum of the areas marked A and B, i.e. GC = A / (A + B). A GC of zero expresses perfect equality, where all values are the same, for example, where every project in a portfolio is exactly equal in cost/schedule programmatic posture. A GC of one expresses maximal inequality among values, for example, a single project program. This study addresses three distinct portfolios:

- Scenario 1: A representative mixed SMD portfolio (GC between 0 and 1)
- Scenario 2: A portfolio with all projects being equal (GC= 0)
- Scenario 3: A single project program (GC = 1)

Figure 6 is an example of the mixed portfolio scenario 1 where three distinct mission classes, 11 small $200M, 9 medium $700M and 5 large $1.5B, make up the total portfolio for a GC of 0.7. The goal of utilizing different scenarios is to cover the "corner solutions" of possible portfolio makeup.

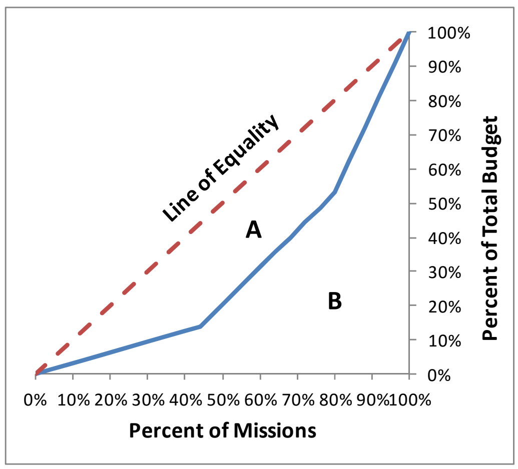


Figure 6: Graphical Representation of Gini Coefficient

3. Scenario 1 Results: Mixed Portfolio

Scenario Input

The initial scenario represents a mixed portfolio of Category (CAT) 1, 2 and 3 missions as defined in NPR 7120.5E. Figure 7 shows the S-curves used for the analysis which were created using the coefficient of variance derived from historical data. These S-curves were the basis for representing a typical mixed NASA portfolio of higher cost CAT1, medium cost CAT2 and lower cost CAT3 missions.

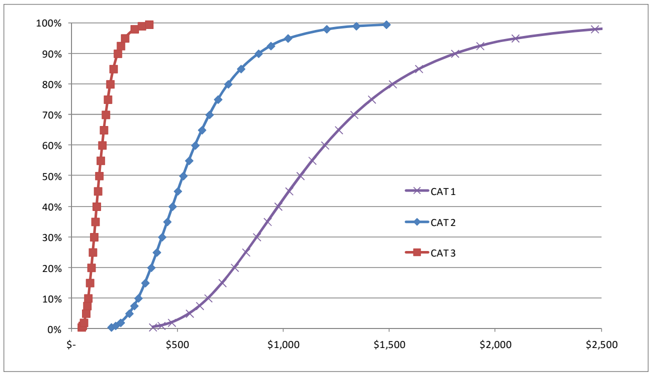


Figure 7: S-curves used in Scenario 1

Variation in BCL

To understand the impact of varying BCL, the first cases that were run along of the diagonal of the case matrix shown in Figure 3 where BCL = FCL. The result of the first FoM, portfolio cost growth over the initial baseline plan, is shown in Figure 8. As expected, the growth is lower for the higher BCL cases and higher for the lower BCL cases as fewer overruns occur at the higher BCL levels.

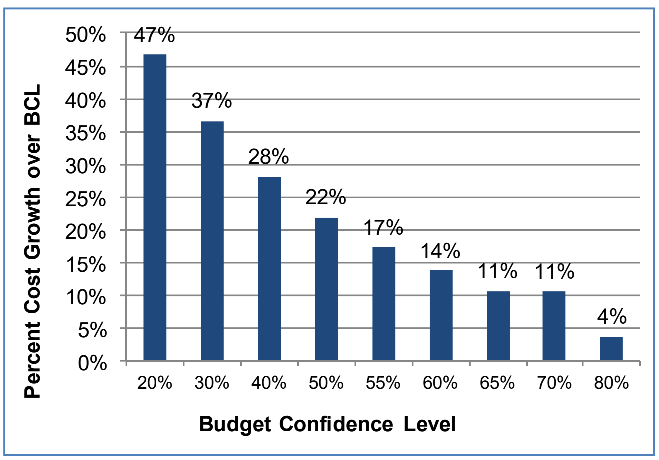


Figure 8: Cost Growth for Varying BCL

The result of the second FoM, time to launch the first 20 mission, is shown in Figure 9. The results show more variability due to the different launches of the 20 missions but the general trend is as expected with the less time needed to launch for the higher BCL cases and the greater time needed for the lower BCL cases as cost and schedule overruns occur which move follow on missions out.

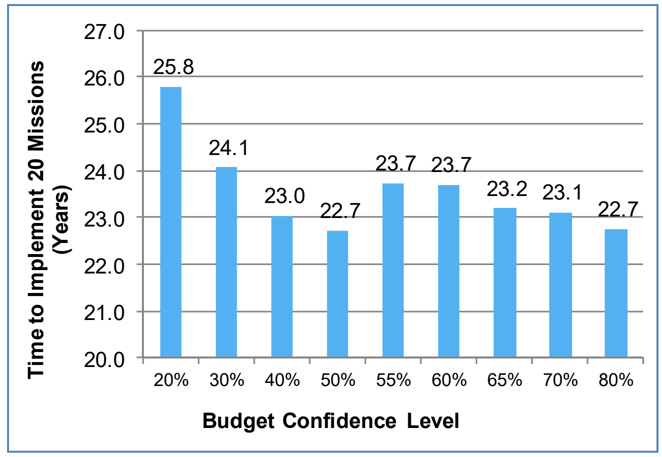


Figure 9: Time to launch 20 missions for Varying BCL

The results of the third FoM, percentage of planned missions launched in 20 years, is shown in Figure 10 with the results as expected with a greater percentage launched for the higher BCL cases.

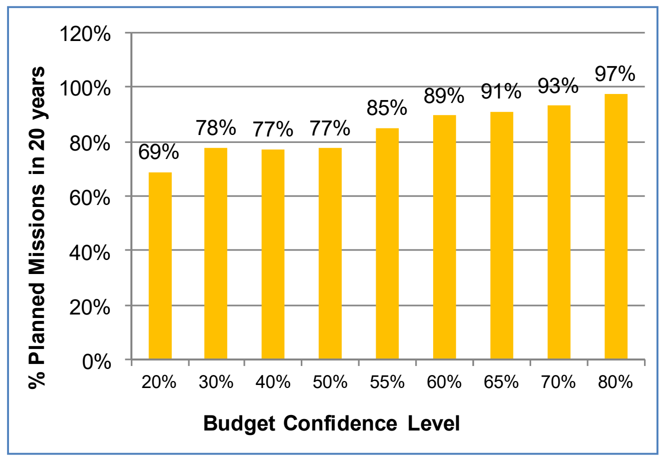


Figure 10: Percent Planned missions for Varying BCL

The results of the fourth FoM, percentage of missions exceeding threshold breach, is shown in Figure 11 and is also as expected with the lower threshold breach at the higher BCL levels.

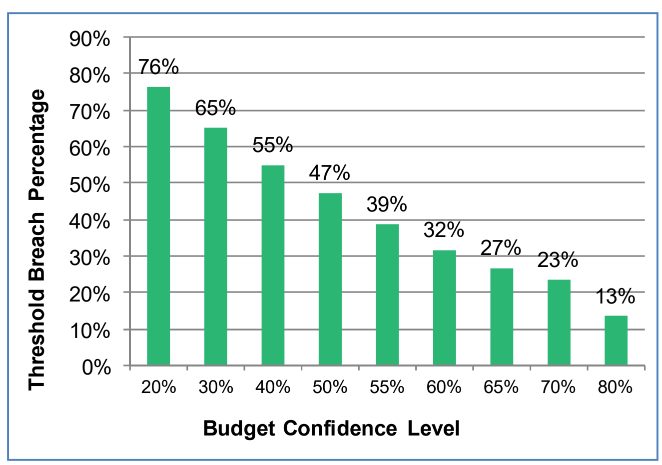


Figure 11: Percent Threshold Breach for Varying BCL

The FoMs stated address performance relative to plan. One additional consideration is the number of absolute missions that are launched in a given amount of time. Figure 12 is a variation of FoM #3 looking at the absolute number of missions implemented in 20 years as a function of the BCL and FCL. As can be seen, the number of missions implemented is similar, within one full mission, for BCLs from 55% to 70% for FCLs of 50% or less, as indicated by the red box in Figure 12.

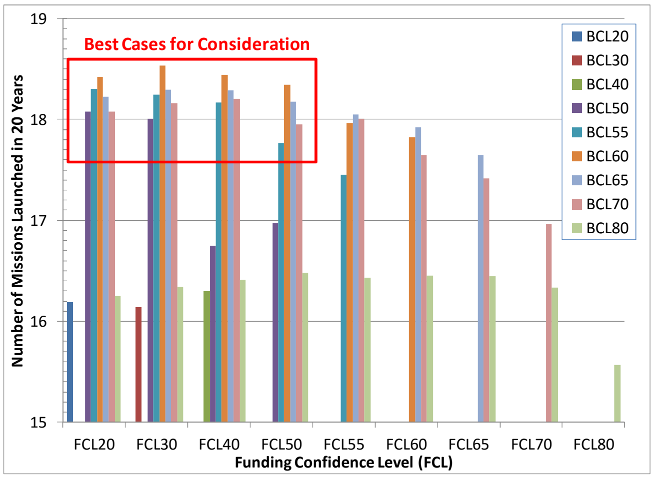


Figure 12: # Missions Implemented for Varying BCL

Choosing the "Best" BCL

Looking at the combination of FoMs shown in Figures 8, 9, 10, and 11, it is clear that a trend emerges as the performance of each FoM worsens as BCL decreases. Looking at this general trend with respect to the results of Figure 12 indicates that the highest BCL should be chosen where the number of missions implemented is similar – as indicated by the red box in Figure 12 – so that performance to plan is maximized. This occurs at the 70% BCL case where the number of missions implemented are similar to the 55%, 60% and 65% BCL cases but the performance of the 4 FoMs relative to plan, as shown in Figures 8 through 11, are better as a whole. This result is consistent with the policy stated in NPR 7120.5E.

Variation in FCL

In addition, the variation in FCL was investigated for a given BCL. For the example shown in Figure 13, variations of FCL for the BCL case of 70% were for each FoM. Figure 13 shows that each of the parameters is best at the 40% to 50% FCL except for the Threshold Breach which is always getting worse as FCL is decreasing. For this reason, the combination of 50% FCL at 70% BCL looks like a reasonable baseline policy for budgeting and funding a mission in a mixed portfolio scenario. This is also consistent with current NASA policy.

Comparison of Old Way vs. New Way

Figure 14 isolates the comparison of the new policy, with a BCL of 70% and FCL of 50%, vs. the old policy of, historically, having both a 20% BCL and 20% FCL. As can be seen, the new policy is substantially better for each of the study FoMs.

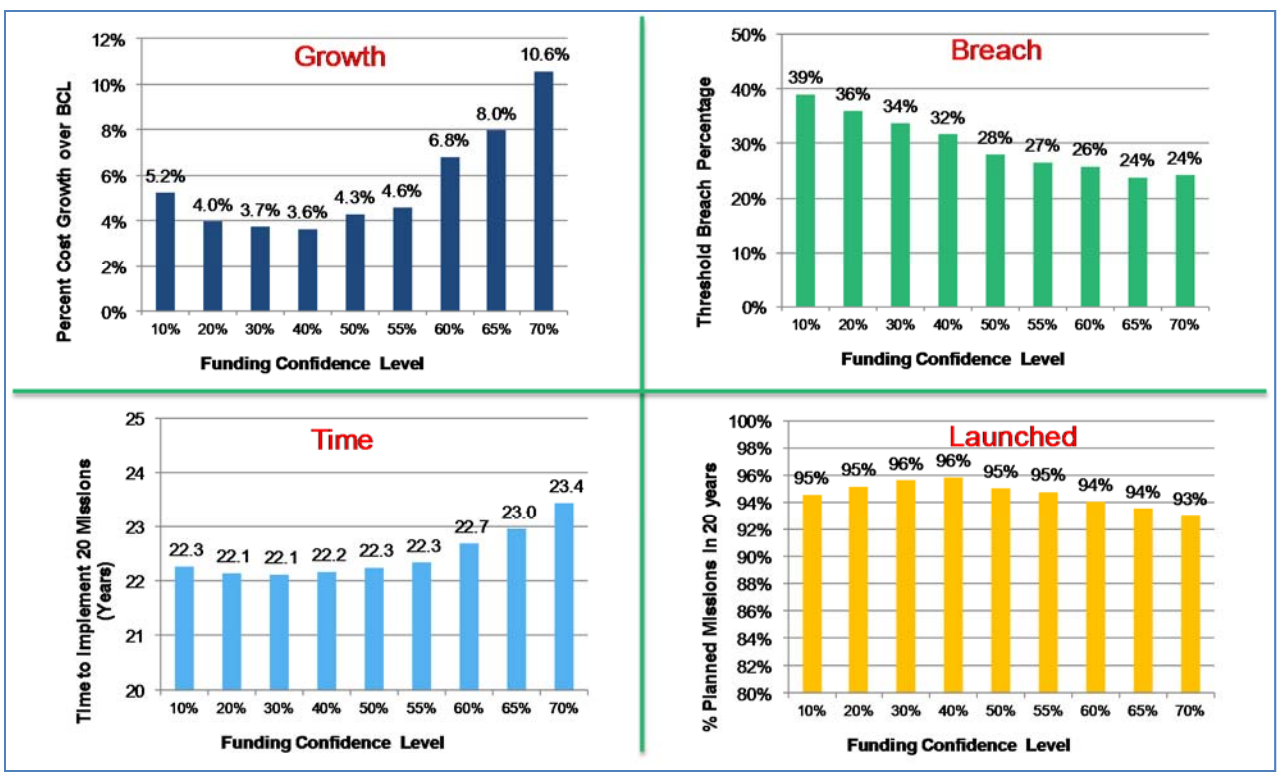


Figure 13: FoM Results for 70% BCL and Varying FCL for Scenario 1 Mixed Portfolio

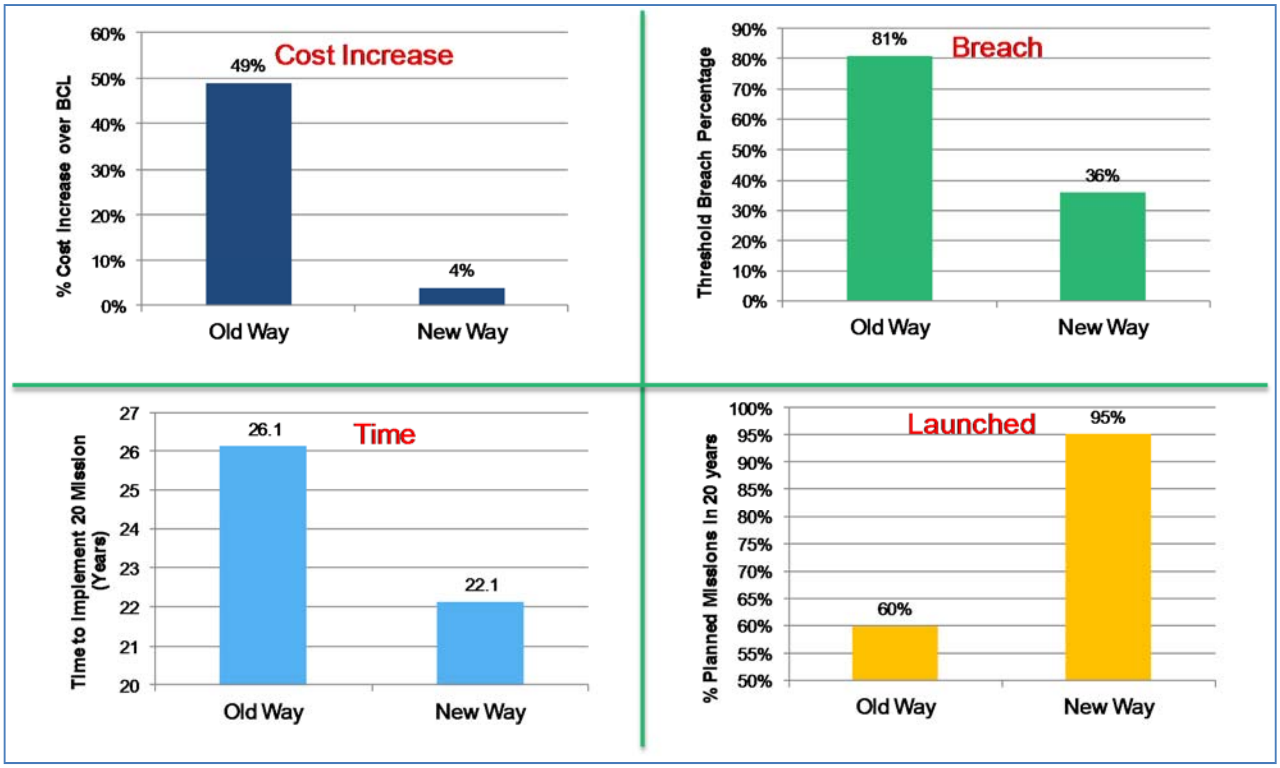


Figure 14: FoM Results Comparing "Old Way" of Budgeting & Funding vs. New Policy

4. Scenario 2 Results: Same Mission Portfolio

Scenario Input

To understand the potential sensitivity to a different portfolio mix, scenario 2 was developed which consists of all Category 2 missions with the same planned funding profile, total cost and schedule for a $700M, five year development with three years of mission operations. This scenario represents a Gini coefficient of 0, were all projects are equal, and is used primarily to determine if the overall results are affected by a different portfolio of missions.

Choosing the "Best" BCL

A similar approach was used to determine if the "best" BCL is different in scenario 2 than scenario 1. Figure 15 shows the absolute number of missions launched in 20 years for scenario 2 and is a parallel of Figure 12 for scenario 1. The results are different than in scenario 1, however, where the number of missions implemented is maximized for BCL cases of 50% and 55% for FCLs of 20%, 30% and 40%, as indicated by the red box in Figure 15. This trend is different than scenario 1 and indicates that the mixture of missions does have an effect on the results. The difference in results from scenario 1 and 2 imply that each portfolio would require its own unique analysis to determine sensitivities to the mixture of missions.

Figure 16 shows the FoMs for scenario 2 for the overall variations in BCL. As can be seen, the FoMs exhibit the same general trends as the scenario 1 BCL FoMs as shown in Figures 8, 9, 10 and 11 in that performance worsens as the BCL decreases. Similar to scenario 1, this implies that the "best" BCL should be the largest BCL in which the number of missions is maximized. Combining this information with Figure 15 implies that a BCL of 55% is the most best choice for scenario 2.

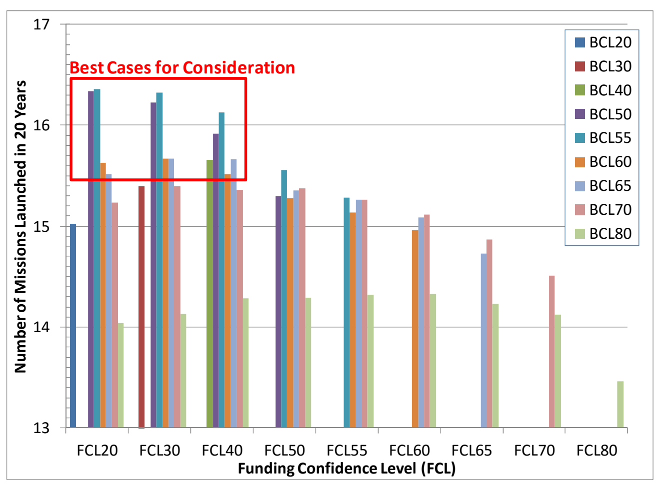


Figure 15: # Missions Implemented for Varying BCL

Variation in FCL

Figure 17 shows the results for all 4 FoMs for the scenario 2 case for 55% BCL. As shown, most metrics start to flatten at 40% FCL while the minimum threshold breach is at 40% FCL. In this case, a 55% BCL and 40% FCL looks like the best overall performance for all metrics.

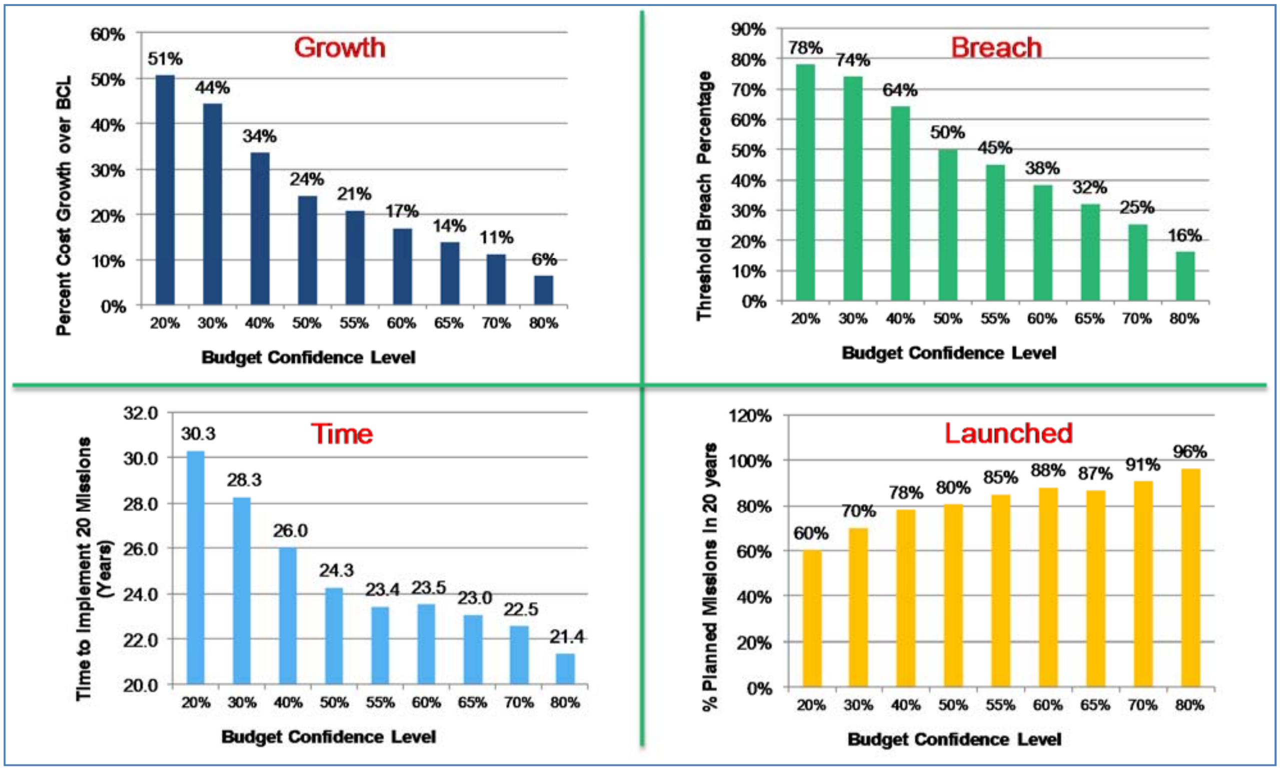


Figure 16: FoM Results for Varying BCL for Scenario 2 All the Same Mission

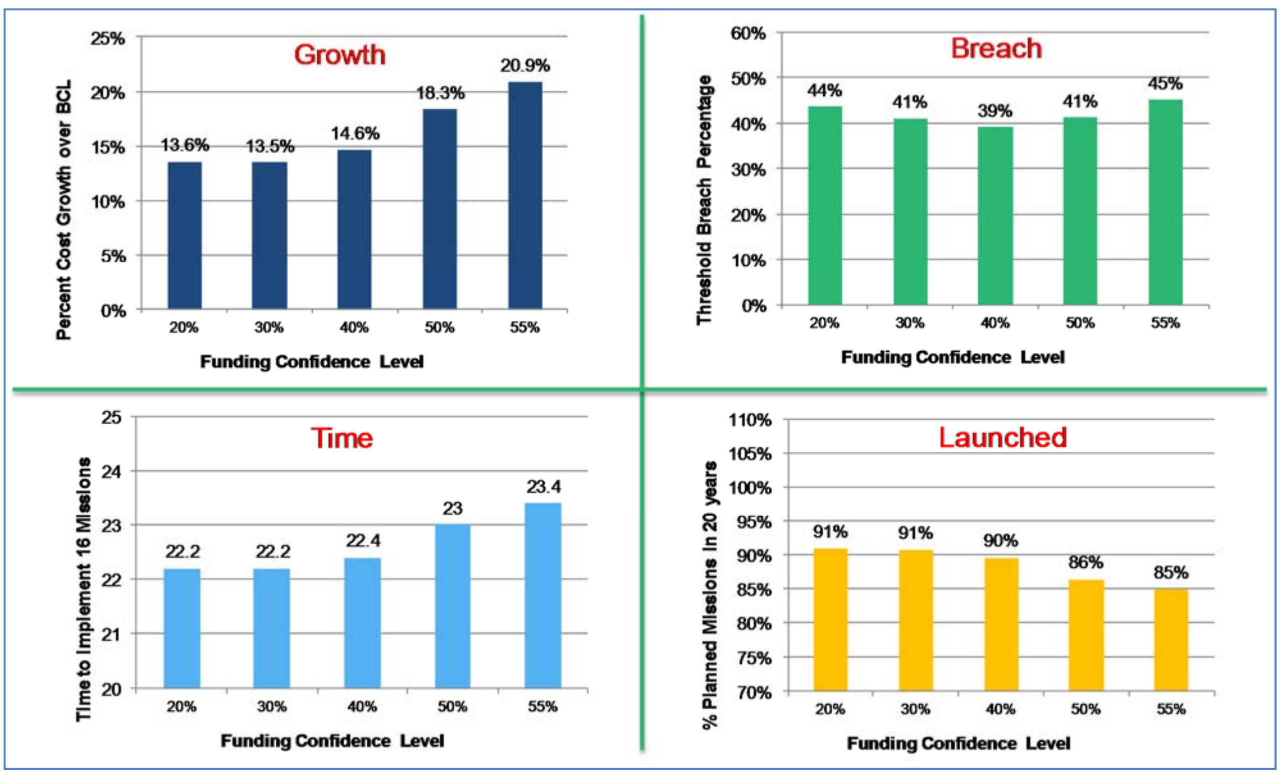


Figure 17: FoM Results for 55% BCL and Varying FCL for Scenario 2 All the Same Mission Portfolio

5. Scenario 3 Results: Single Flagship

Scenario Input

For the single Flagship mission scenario, it was asssumed that the mission had a $4B life cycle cost which includes $250M for launch and 5 years of ops at $250M total which results in a $3.5B in development cost. It is assumed that the funding ramps up and then is capped at $400M annually with the first year at $150M, then $300M and then $400M annually to launch. For the single project Flagship, it is assumed that all funding available to the program is also available for the project such that FCL = BCL for fall cases. Additionally, since the output looks at a single mission, the primary difference in the case results is time to launch. Because of the fixed annual budget, the cost draw must fit under the annual funding constraint so that the only variable in play is delay in launch date.

Choosing the "Best" BCL

To better understand the results of the scenario, it is best to look at the specific likelihood of cost for a given mission. Figure 18 shows the likelihood for BCL equals 20% and 50%. The graph shows the behavior is as expected where the 20% BCL case is a straight line at the 20% likelihood but then grows beyond as the draws above 20% cause the cost to grow due to the penalties within the model. The 50% BCL case shows similar performance with a straight line cost to 50% likelihood and then growing beyond. As can be seen, the cost for the 50% BCL case is less than the 20% case starting at the 45% likelihood at which point the 20% BCL case is more expensive than the 50% BCL case. When comparing these two cases, it is important to look at the area difference under each curve to see which is less expensive which percentage of the time.

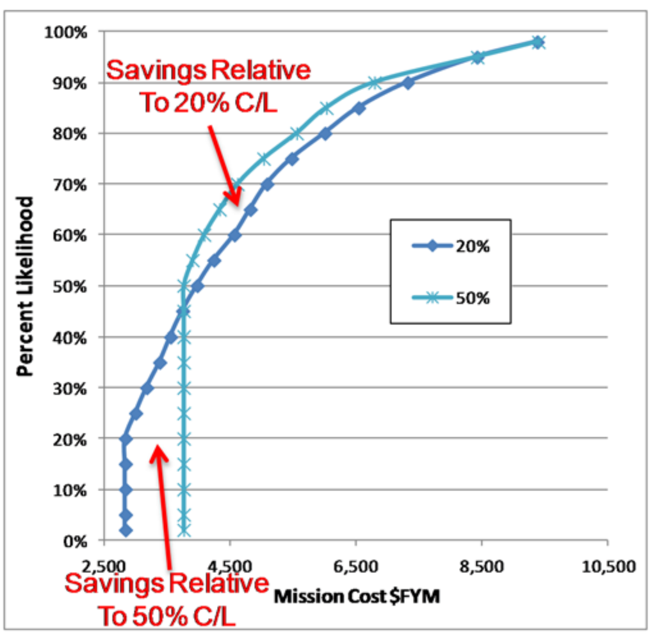


Figure 18: Cost Differences for Different Cases

Another way to see this comparison is by looking at the average mission cost, as shown in Figure 19. As can be seen, the "knee in the curve" for the average mission cost is at the 45 to 50% BCL level, which implies that, for single flagship missions, it may be better to budget at the 50% BCL level as opposed to the 70% BCL per NPR 7120.5E.

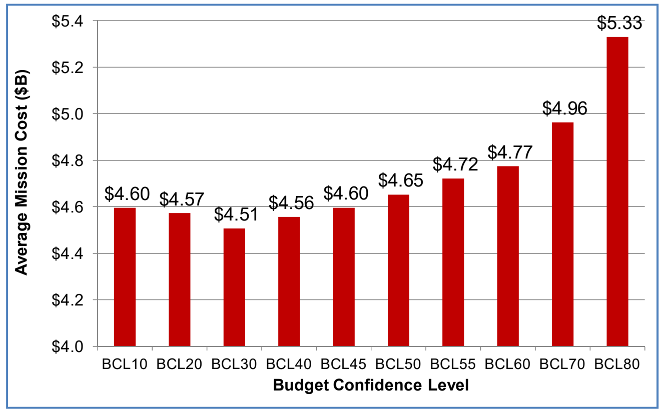


Figure 19: Average Mission Cost for Scenario 3

The lower BCL, however, comes at a price. Figure 20 shows the threshold breach percentage as a function of BCL. This demonstrates a similar trend as scenario 1 and 2 where lowering BCL results in a higher threshold breach percentage. The tradeoff then is potentially have a lower average mission cost but at the risk of a higher threshold breach percentage. This tradeoff is important as single project flagship missions are typically highly visible to the public, Congress and the Office of Management and Budget (OMB).

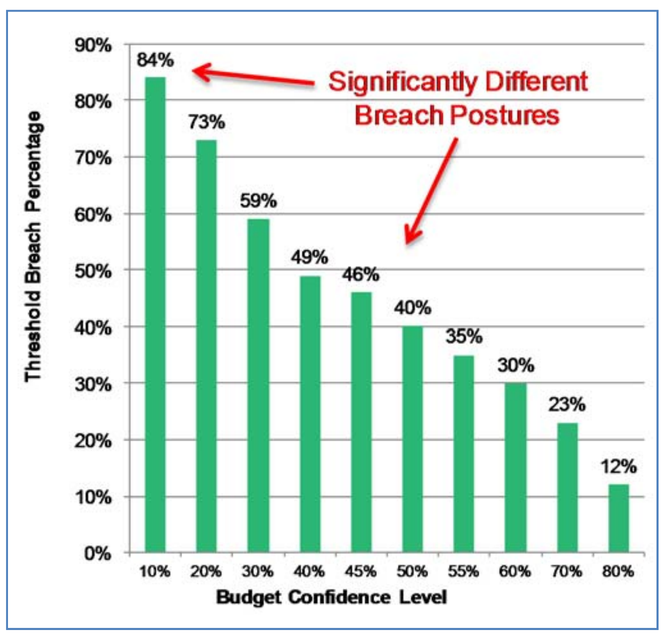


Figure 20: Threshold Breach Percent for Varying BCL

Figure 21 shows the complete set of FoMs, which shows that potentially lower BCL values may have a lower average mission cost and at an earlier launch date but that this would come at a price of having larger deviations relative to plan.

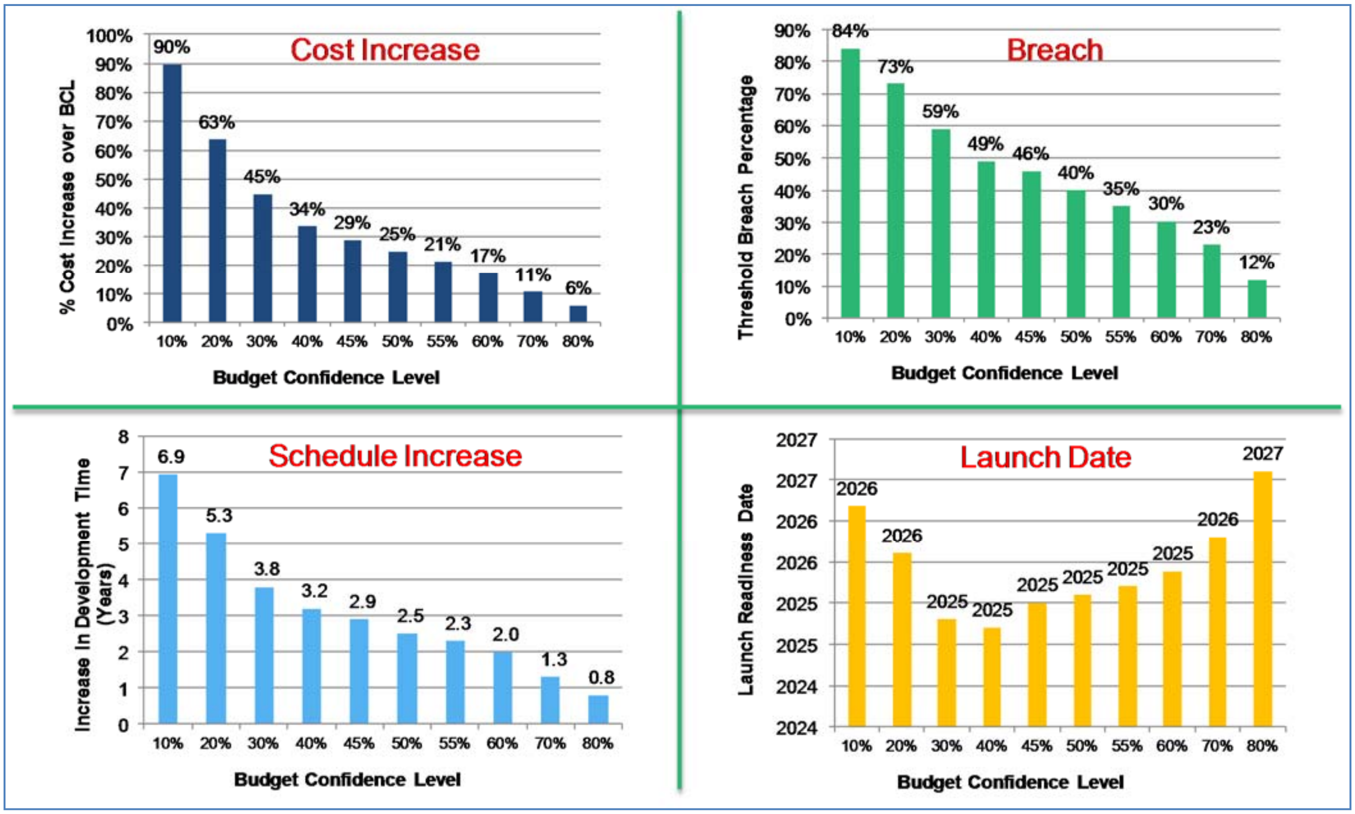


Figure 21: Varying BCL for Scenario 3 Single Flagship Mission

6. SUMMARY

The analysis shows that for a typical portfolio of multiple loosely coupled missions, NASA's baseline JCL policy of budgeting projects at the 70th percentile and funding to at least the 50th percentile seems to be a sound strategy.

However, for single-project Programs, NASA's baseline JCL policy may be less than an optimal strategy as data is indicates that a confidence level of ≤ 50th for Budgeting Confidence Level may be more cost effective depending on acceptable level of threshold breach tolerance.

Different budgeting and funding strategies could be beneficial depending on management FOM priorities and portfolio characteristics so it is recommended that an analysis like that conducted in this study be applied to the specific portfolio of missions to be examined to determine the optimal budgeting and funding confidence levels.

REFERENCES

[1] "NASA Risk Adjusted Cost Estimates"; J Hamaker NASA Headquarters Cost Analysis Division; International Society of Parametric Analysts Annual Conference, May 23-26, 2006

[2] "The Trouble With Budgeting to the 80th Percentile", T Anderson.; 72nd Military Operations Research Society Symposium; June 22 – 24, 2004.

[3] "NASA Space Flight Program and Project Management Requirements", NASA Procedural Requirements (NPR-7120.5E), 14 August 2012

[4] "Using Historical NASA Cost and Schedule Growth to Set Future Program and Project Reserve Guidelines", R. Bitten, D. Emmons, C. Freaner, IEEE Aerospace Conference, March 3-10, 2007

[5] "Optimism in Early Conceptual Designs and Its Effect on Cost and Schedule Growth: An Update", R. Bitten, C. Freaner, D. Emmons, IEEE Aerospace Conference, March 6-13, 2010.

[6] "Explanation of Change (EoC) Study: Approach and Findings", R. Bitten, D. Emmons, F. Bordi, C. Scolese, IEEE Aerospace Conference, March 2-9, 2013.

[7] Section 103, Baselines and Cost Controls, Public Law 109–155, December 30, 2005

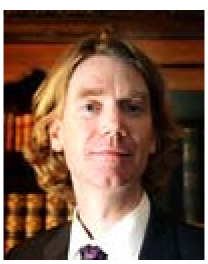
[8] "An Integrated Architecture Design/ Cost/Scheduling Approach for Future Space Exploration Program Affordability," R. Bitten, D. Emmons, I. Min, and T. Radcliffe, 1st Space Exploration Conference, January 30 - February 1, 2005

[9] "Affordability Assessments to Support Strategic Planning and Decisions at NASA", D. Emmons, M. Lobbia, and T. Radcliffe, R. Bitten, IEEE Aerospace Conference, March 6-13, 2010.

[10] Damgaard, Christian. "Gini Coefficient." From MathWorld--A Wolfram Web Resource, created by Eric W. Weisstein. http://mathworld.wolfram.com/GiniCoefficient.html

BIOGRAPHY

Robert Bitten *is a Principal Engineer at The Aerospace Corporation and has conducted independent cost estimates for NASA proposal evaluations and independent assessments for a variety of different NASA missions and organizations. He is a winner of the President's Award, one of The Aerospace Corporation's highest honors, for his effort in assessing the cost effectiveness of different alternatives in the in the Hubble Space Telescope Remote Servicing Module Analysis of Alternatives. He also won the 2007 NASA Cost Estimating Support Contractor of the Year Award that is awarded to recognize an individual who has provided "outstanding contractor support to the NASA cost estimating community and significantly contributed to the field of cost estimating." Bob has a Bachelor's of Industrial & Systems Engineering degree from the Georgia Institute of Technology and an MBA from Pepperdine University.*



Charles Hunt *is currently a senior cost analyst for the NASA Headquarters, Strategic Investments Division, with the Office of the Chief Financial Officer with the responsibility for establishing and maintaining cost and schedule estimating policy, conducting and funding cost and schedule research, providing tools and models to the broader NASA cost community, coordinating programmatic support for Agency Standing Review Boards as part of the Agency Programmatic Analysis Capability, as well as providing cost and schedule estimating support to the NASA Administrator through Agency-level studies. Mr. Hunt is involved with the communication of NASA cost policy to both internal and external stakeholders. Mr. Hunt joined NASA Headquarters in 2007. He began his career as a cost analyst for the Engineering Cost Office at Marshall Space Flight Center performing numerous cost estimates in support of projects, studies, and Agency level analysis. Charley is a graduate of Tennessee Technological University with a Bachelor of Science in Industrial and Systems Engineering and of Johns Hopkins University with a Masters of Art in Applied Economics.*