IDA | RESEARCH NOTES

ACQUISITION

PART 2: EXECUTING AND MANAGING PROGRAMS

Spring 2016

IDA is the Institute for Defense Analyses, a non-profit corporation operating in the public interest.

IDA's three federally funded research and development centers provide objective analyses of national security issues and related national challenges, particularly those requiring extraordinary scientific and technical expertise.

The articles in this edition of IDA Research Notes were written by researchers within the following four IDA divisions. The directors of those divisions would be glad to respond to questions about the specific research topics or related issues.

Cost Analysis and Research Division (CARD)
Dr. David J. Nicholls, Director (703-575-4991, dnicholl@ida.org)

Information Technology and Systems Division (ITSD)
Dr. Margaret E. Myers, Director (703-578-2782, mmyers@ida.org)

Science and Technology Division (STD)
Dr. Leonard J. Buckley, Director (703-578-2800, lbuckley@ida.org)

System Evaluation Division (SED)
Dr. Steve Warner, Director (703-845-2096, swarner@ida.org)

Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, Virginia 22311
www.ida.org

This is the second of two issues of *IDA Research Notes* focused on acquisition in the Department of Defense. The first article, by **David McNicol**, examines the relationship of acquisition policy and processes to cost growth. It finds no evidence that the efforts to strengthen the acquisition process through the years have resulted in lower or higher Procurement Acquisition Unit Cost (PAUC) growth. The research did uncover, however, a strong association between funding climate at Milestone B/II and cost growth, noting that high PAUC growth is not persistent, but rather episodic, and correlated with environmental factors outside the control of the acquisition process. It concludes with the striking observation that it seems unlikely that further broad changes in the acquisition process would have a major effect on cost growth.

The next two articles relate to Nunn-McCurdy breaches, a topic of frequent concern in the acquisition community. These breaches occur when a program experiences cost growth exceeding established thresholds. The first article, by **Patricia Bronson** and **David Sparrow**, describes an algorithm developed at IDA that, using readily available program data, improves the predictive value of poor performance indicators for existing major defense acquisition programs. The second article, by Dr. Bronson and **Chris Martin**, highlights the key findings from the IDA root cause analysis of the VTUAV Fire Scout Nunn-McCurdy breach.

DoD depends on private industry to design and produce its weapon systems. Over the years, IDA has performed numerous research projects in support of DoD's efforts to maintain a robust industrial base. **Brian Gladstone, Brandon Gould,** and **Prahant Patel** describe some recent work IDA performed to help develop the sustainment plan for large solid rocket motors.

The environment in which DoD acquires, builds, and operates national security systems is increasingly dependent on global supply chains for commercial information communications technology (ICT) and ICT-enabled components—providing opportunities for adversaries to tamper with those systems. During the past decade, IDA supported the DoD Chief Information Officer's efforts to enhance supply chain risk management (SCRM) and created an education, training, and awareness module to promote awareness of the risks inherent in the global supply chain and to increase understanding of ICT SCRM. **Thomas Barth**, **Michelle Albert**, and **Elizabeth McDaniel** examine the timely issue of managing these risks.

IDA has long been involved in assessing affordability. Two articles by **David Tate** highlight prior and current work IDA has done on acquisition affordability. The first details the history—and provides lessons learned—from 15 years of the IDA-developed Acquisition Portfolio Optimization Tool, or PortOpt. The second paper addresses the current and future challenges of predicting the effect of schedule changes on acquisition costs.

We close with an article on the F-35 Joint Strike Fighter (JSF)—the largest defense acquisition program in history. Specifically, **Lisa Veitch** and **Evan Laprade** describe IDA's role in following the developmental testing of the JSF and offering objective technical assessments to support key decision milestones. The article notes that, based on historical precedence, there is a high likelihood of future failures that are not yet identified.

COST GROWTH, ACQUISITION POLICY, AND BUDGET CLIMATE

David L. McNicol

The Problem

This article asks whether, taking account of funding climate, there is a statistically significant association between changes in acquisition policy and process and cost growth on Major Defense Acquisition Programs.

Average PAUC growth was substantially higher in a relatively constrained funding climate than in a relatively accommodating climate.

Discussions of acquisition reform over the past twenty-five years have usually put the Department of Defense (DoD) Program Manager (PM) and personnel in the program office in the foreground. These people oversee the contractors and do myriad things that must be done by the government for a major acquisition program to move forward—contracting, financial management, and test planning, among others. In the background are the contractors who typically do the development and manufacturing. A good program will not occur if the government personnel and contractors do not do their jobs well. It is equally true that if these individuals and organizations do their jobs well, a good outcome for the program is more likely.

What this focus on the DoD PM, the program office personnel, and the contractors' PMs and workers leaves out are factors they must accept as "givens." These givens are subject to changes—sometimes large and fairly sudden—that presumably have substantial consequences for program outcomes. One of the givens is DoD acquisition policy and process. A second is the DoD budget, which does not determine, but generally has a marked influence on, the funding for individual programs.

ACQUISITION REGIME AND PAUC GROWTH

DoD acquisition policy and process over the period 1970–2007 can be grouped into five successive regimes:

1. The Defense Systems Acquisition Review Council (DSARC), 1970–1982
2. The Post-Carlucci Initiatives DSARC, 1983–1989
3. The Defense Acquisition Board (DAB), 1990–1993
4. Acquisition Reform (AR), 1994–2000
5. The DAB – Post-Acquisition Reform, 2001–2007

Table 1 displays the average PAUC growth for MDAPs that passed Milestone (MS) B or (pre-2001) MS II or filed a first Selected Acquisition Report (SAR) in each of these regimes. The PAUC growth figures all are measured from the MS II/B baseline and normalized to the MS II/B total inventory objective. There are a number of interesting aspects to these data: for example, the high PAUC growth during the AR period and the lower PAUC growth for 2001–2007. The single most notable feature of these data is the absence of any trend in PAUC growth. If changes in acquisition policy and process have had a sustained influence on PAUC growth, it does not show up in this table.

Broadly, there are two ways to explain the absence of sustained effects of acquisition policy and process on the PAUC growth data. First, they may in fact not have a strong or consistent effect on PAUC growth. Second, acquisition policy and process may have substantial effects that are masked by some other factor or factors.

FUNDING CLIMATE AND PAUC GROWTH

Thinking along the lines of the second of these possibilities led to consideration of whether changes in the DoD funding climate might be associated with PAUC growth. The period 1970–2007 includes two sub-periods during which the DoD budget was relatively constrained: FY 1970–FY 1980 and FY 1987–FY 2002. It also includes two sub-periods in which MDAP new starts found funding climate relatively accommodating: FY 1981–FY 1986 and FY 2003–FY 2007. Table 2 displays the average PAUC growth data for these four sub-periods.

Table 1. Average PAUC Growth in Successive Acquisition Regimes

Acquisition Regime	Time Period	Average PAUC Growth	No. of Observations
DSARC	1970–1982	32%	48
Post-Carlucci Initiatives DSARC	1983–1989	19%	40
DAB	1990–1993	36%	11
Acquisition Reform (AR)	1994–2000	66%	27
DAB Post-AR	2001–2007	19%	25

Table 2. Average PAUC Growth in Different Funding Climates

Relatively Constrained		Relatively Accommodating	
Period (FY)	**PAUC Growth**	**Period (FY)**	**PAUC Growth**
1970–1980	35% (42)	1981–1986	12% (35)
1987–2002	53% (55)	2003–2007	7% (19)

Note: Numbers in parentheses are the number of observations available.

These data make it clear that the average PAUC growth in relatively constrained funding climates was far larger than it was in periods during relatively accommodating funding climates—by a factor of three in the first comparison and by a factor of more than seven in the second.

ACQUISITION REGIME AND FUNDING CLIMATE

Table 3 expands Table 2 by replacing the funding climate sub-periods with the acquisition policy and process regimes. This table provides results for two sets of natural experiments. First, the PAUC growth columns give the effect of changes in the acquisition regime for a given funding climate. Second, the rows show the effect of funding climate for a given acquisition regime. For example, the first eleven years of the DSARC (FY 1970–FY 1980) were in a relatively constrained funding climate, while the next two (FY 1981–FY 1982) were in a period in which the DoD budget was relatively accommodating.

Statistical analysis of the data behind the averages in this table leads to two conclusions. First, there is no statistically significant improvement or worsening of PAUC growth correlated with the different acquisition policy and process regimes. This result is not surprising for the relatively accommodating climate (column on the right). In contrast, in the relatively constrained periods (column on the left), average PAUC growth for AR and DAB post-AR is noticeably higher than the averages for previous periods, but the differences proved not to be statistically significant because of the large variance among programs in each period.

Second, average PAUC growth was substantially higher in a relatively constrained funding climate than in the relatively accommodating climate. We have only three natural experiments of changes in funding climate for a given acquisition regime,

Table 3. Average PAUC Growth by Acquisition Regime and Topline Condition

	Relatively Constrained		Relatively Accommodating	
Acquisition Regime	**Period (FY)**	**PAUC Growth**	**Period (FY)**	**PAUC Growth**
DSARC	1970–1980	35% (42)	1981–1982	11% (6)
Post-Carlucci Initiatives DSARC	1987–1989	34% (11)	1983–1986	13% (29)
DAB	1990–1993	36% (11)	None	N/A
Acquisition Reform (AR)	1994–2000	61% (27)	None	N/A
DAB Post-AR	2001–2002	57% (6)	2003–2007	7% (19)

Note: Numbers in parentheses are the number of observations available.

since two of the five acquisition regimes (DAB and AR) fall entirely within one funding climate. Each of these three natural experiments on the effect of funding climate had the same outcome—MDAPs that passed MS II/B in a relatively constrained funding climate on average had a much higher PAUC growth rate than those that passed MS II/B in a relatively accommodating funding climate for a given acquisition regime.

These differences are statistically significant at the 1 percent confidence level. The outcomes of the first two experiments are virtually identical—an average PAUC growth of 35 and 34 percent, respectively, in the two periods when funding was relatively constrained and average PAUC growth of 11 percent and 13 percent, respectively, in the two periods when the funding climate was relatively accommodating. The effect is even more pronounced in the third experiment (DAB Post-AR)—57 percent in FY 2001–FY 2002 versus just 7 percent for FY 2003–FY 2007.

DOES THE RESOURCE ALLOCATION PROCESS PLAY A MAJOR ROLE IN PAUC GROWTH?

These conclusions tend to challenge a fundamental assumption implicit in most discussions of acquisition reform: that the main causes of PAUC growth are to be found in the acquisition realm—the performance of the contractors, the effectiveness of the PM, the adequacy of the developmental test plan, and the completeness of the systems engineering plan, among others. This assumption is hard to maintain when the many changes in acquisition policy and process made in the past four decades have not had statistically significant effects on PAUC growth, but there is a significant association between PAUC growth and funding climate at the point when the MS II/B baseline was set.

The association between PAUC growth and funding climate suggests that the resource allocation process, particularly at the Service level, plays an important role in cost growth. This does not mean “budget instability.” Budget instability is a term of art for changes in MDAP funding through the annual resourcing cycle and “taxes.” Budget instability is a chronic condition, present to some degree in all periods. What we observed is a recurring pattern—that MDAPs that passed MS II/B during periods of relatively constrained funding, on average, had much higher PAUC growth than those that passed MS II/B when funding was relatively accommodating.

The conjecture that the resource allocation process plays an important role in cost growth gets some support from an unexpected direction—MDAPs with negative cost growth, of which there are twenty-nine in our sample. Negative PAUC growth is recorded if the actual cost of a program proves to be less than the cost in the MS II/B baseline. Assuming the program was funded to its MS II/B baseline, this implies that, over time, funds can be taken from the program in question and reallocated to other applications, including other acquisition programs.

The program, then, effectively can be used as a “bank”—a way to hold reserves in relative safety until they are needed. A bank of this sort is more likely to be needed in a relatively accommodating funding climate, as it can then serve as a way to delay final decisions on allocation of the higher level of funding that has become available. We would therefore expect to find a higher proportion of MDAPs with negative PAUC growth in the relatively accommodating climates, and this is what we observe. About 30 percent of our MDAPS that passed MS II/B in relatively accommodating funding climates show negative PAUC growth, compared to about 10 percent across the periods of relatively constrained climate.

MDAPs with "high cost growth," which we define as quantity normalized PAUC growth of at least 50 percent, also suggest an influence from the resource allocation process. DoD resource managers, particularly at the Service level, have only a few tools for responding to a relatively constrained funding climate. One of these is to impose top-down limits on the funding for particular MDAPs as they approach MS II/B. Plausibly, the result will be particularly optimistic programmatic and costing assumptions, which lead to an expectation that MDAPs started in periods of relatively constrained funding climate will have a larger proportion with high PAUC growth. This is again what is observed. During periods of relatively constrained funding climate, about 40 percent of MDAPs had very high PAUC growth. In contrast, during periods of relatively accommodating funding climate, only about 7 percent of MDAPs experienced high PAUC growth.

Taking both funding climates together, 85 percent of MDAPs with PAUC growth of at least 50 percent passed MS II/B during a relatively constrained funding climate. These MDAPs had an average PAUC growth of 93 percent and accounted for just over three-quarters of total PAUC growth. Excluding high cost growth MDAPs and MDAPs with negative PAUC growth, average PAUC growth across the two funding climates was just 18 percent. High PAUC growth is then predominantly a feature of programs with PAUC growth of at least 50 percent, and these programs mainly passed MS II/B in periods of relatively constrained funding climates. These points are important because they suggest that reforms directed to the average or typical MDAP may miss the real source of the problem.

IMPLICATIONS FOR DISCUSSIONS OF ACQUISITION REFORM

Our research points to three implications for a discussion of acquisition reform. First, the relevant context for understanding PAUC growth is the interface between the acquisition process and the resource allocation process. The crucial evidence behind this point is the strong association between funding climate and PAUC growth. Resource managers must think in terms of a portfolio of programs at various stages of the acquisition life cycle, from efforts in the technology base through programs nearing the end of production.

When a program is completed, it opens a resource "hole" that programs emerging from Engineering and Manufacturing Development can occupy. In turn, programs earlier in the acquisition cycle can move forward as well. When funding for acquisition turns down, these holes get smaller, close entirely, or require cuts in funding for ongoing programs. The alternatives available in this circumstance are all undesirable—cancellations of programs, delays in new starts, stretches, and unrealistic pricing. The evidence summarized here suggests that it is in this context that high PAUC growth arises.

Second, it seems unlikely that further broad changes in the acquisition process would have a major effect on PAUC growth. The research found no evidence that the efforts to strengthen the acquisition process through the years have resulted in lower or higher PAUC growth. This does not mean that the DAB process does not provide a useful discipline on acquisition programs; moreover, further changes in acquisition policy or process might be warranted for reasons of good government. The evidence does, at a minimum, suggest that the effects of changes in the acquisition process since the early 1970s have not had a dominant effect on PAUC growth.

Third, it is difficult to see that the cultures of the DoD acquisition organizations are a crucial obstacle to improved performance on cost growth. The key point is that high PAUC growth is not persistent, but rather episodic, and correlated with environmental factors outside of the control of the acquisition process. There is little PAUC growth in periods when the funding climate is relatively accommodating. It seems fair to ask if it makes sense to assert that an entrenched culture sometimes results in high cost growth and other times in low cost growth. Just how is it that the A team takes the field so quickly and quietly when the budgetary sun comes out? And why even in bad budgetary weather do more than half of MDAPs exhibit comparatively modest PAUC growth?

Dr. McNicol is a Research Staff Member in IDA's Cost Analysis and Research Division. He holds a doctorate in economics/finance from the Massachusetts Institute of Technology.

Reference

McNicol, David L., and Linda Wu. *Evidence on the Effect of DoD Acquisition Policy and Process on Cost Growth of Major Defense Acquisition Programs.* IDA Paper P-5126. Alexandria, VA: Institute for Defense Analyses, September 2014.

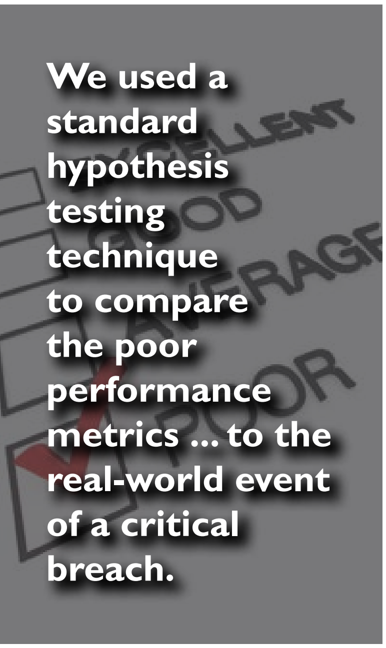
IMPROVING PREDICTIVE VALUE OF INDICATORS OF POOR PERFORMANCE

Patricia Fazio Bronson and David Sparrow

The Problem

Screening techniques need to be developed to identify Major Defense Acquisition Programs that are likely to experience a critical Nunn-McCurdy breach.

The Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, Performance Assessments and Root Cause Analyses (OUSD/AT&L/PARCA) asked IDA to develop screening techniques to identify Major Defense Acquisition Programs (MDAPs) that were likely to experience a critical Nunn-McCurdy breach. PARCA also asked IDA to develop performance assessment methods for MDAPs to support their participation in the Defense Acquisition Executive Summary (DAES) and Integrated Product Team (IPT) processes.



We used a standard hypothesis testing technique to compare the poor performance metrics ... to the real-world event of a critical breach.

This article describes and evaluates a collection of metrics of poor performance that use program data available from Selected Acquisition Reports (SARs) and DAES Web Services, and earned value management (EVM) data from the EVM Central Repository.

The metrics are based on observed events that tend to indicate poor performance:

- Instability in funding and production rate profiles
- Differences between spending forecasts and execution of those forecasts
- Differences between staffing plans and the execution of those plans
- Differences between contract forecasts and funding plans
- Cost growth on mission equipment
- Changes to the estimated costs of developing prime mission equipment
- Persistent use of the Undistributed Budget category
- Rate at which Management Reserve is spent
- Initial investment in system level tasks
- Growth in the cost of system level tasks.

Each metric is evaluated for its ability to identify programs that are likely to experience cost growth on the order of a critical Nunn-McCurdy breach. We used a standard hypothesis testing technique to compare the poor-performance metrics generated for eight programs to the real-world event of a critical breach in Program Acquisition Unit Cost (PAUC).

Figure 1 shows a sample plot of the Unit Cost Growth (UCG) for each program on the vertical axis as a function of a poor-performance index value on the horizontal axis. The poor-performance index value is a linear transformation of the poor-performance metrics for all programs, so the minimum index value is 0 and the maximum index value is 1. The horizontal blue line at 25 percent represents the real-world poor-performance event threshold value (the critical Nunn-McCurdy breach limit). The vertical blue line is the threshold value for detecting poor performance.

As with any test of this type, the sensitivity of the test is established by the placement of the vertical blue line. Move the line all the way to the right, and the results are misses and true negatives. Move the line all the way to the left, and the results are all hits and false alarms.

The results of all the tests are summarized in Table 1.

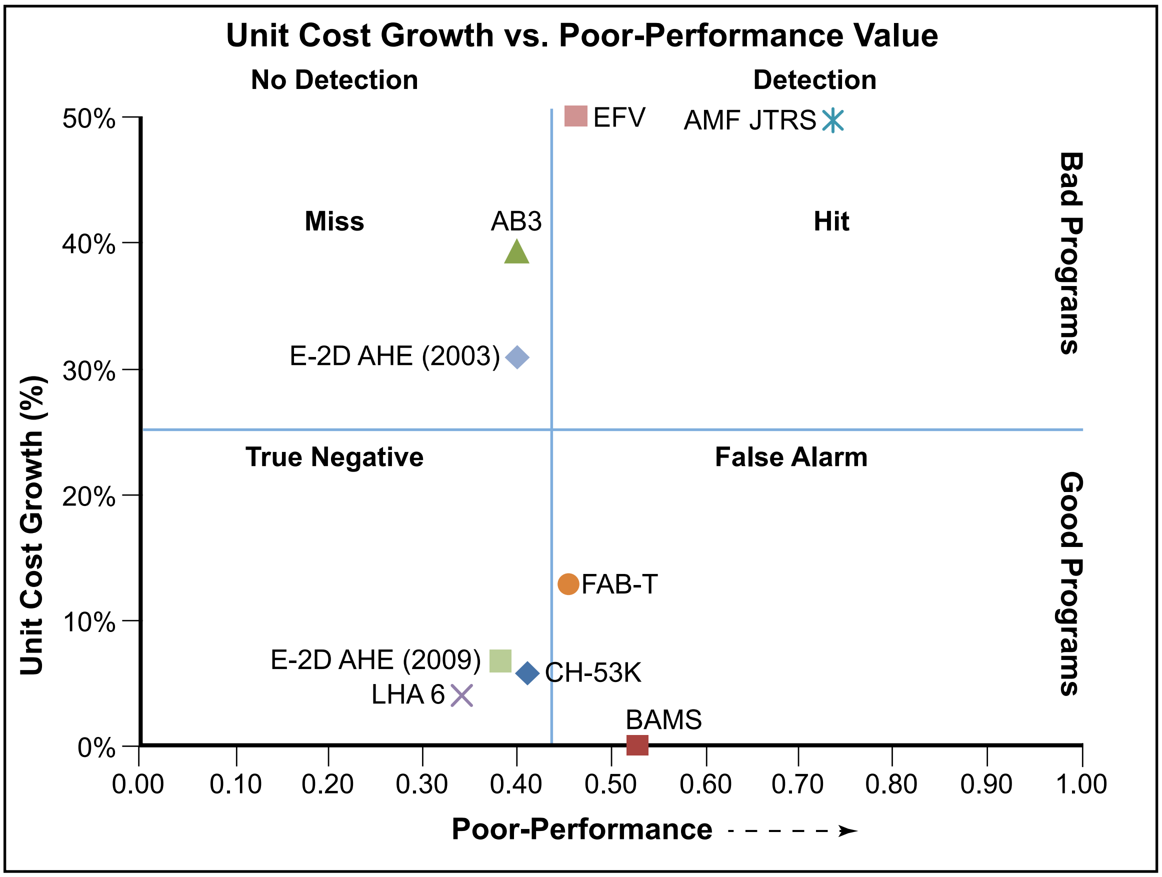


Figure 1. A Scatter Plot of Observed Unit Cost Growth and a Poor-Performance Index

Table 1. Summary of Test Results Showing Successes and Failures

	Instability in Funding and Quantity Profiles			Spending Forecasts			Execution to Staffing Plans				Volatility (#)				
	Funding	Quantity	Magnitude	Scatter	Magnitude	Scatter	Comparison of Contract Forecasts to Funding Plans	Cost Growth on Mission Equipment	Magnitude	Scatter	Undistributed Budget	Management Reserve	Initial Investment in System Level Tasks	System Level Tasks	SE SI ST Growth at 50% Complete
CH-53K								False Alarm	False Alarm						
BAMS								False Alarm					False Alarm		
AB3 Remanufacture	Hit	Miss	Hit	Miss	Miss	Miss	Miss	Miss	Hit	Hit	Miss	Miss	Hit	Miss	Miss
LHA 6													False Alarm		
AMF JTRS	Hit	Hit	Hit	Miss	Hit	Hit	Hit	Hit	Hit	Miss	Hit	Hit	Miss	Miss	Hit
FAB-T	False Alarm	False Alarm	False Alarm				False Alarm								
E-2D AHE (Jun 2003)	Miss	Miss	Miss	Miss	Miss	Miss	Miss	Hit	Miss	Miss	Hit	Hit	Hit	Miss	Miss
E-2D AHE (Jul 2009)								N/A			False Alarm	------- Not Applicable --------			
EFV	Hit	Miss	Miss	Miss	Miss	Hit	Miss	Miss	Hit	Miss	Miss	Miss	Miss	Hit	Hit
Success Rate	78%	56%	67%	56%	67%	78%	56%	50%	78%	67%	67%	78%	50%	63%	75%

The accuracy of the individual performance metrics ranges from 50 percent to 78 percent. Five of the fifteen metrics (33 percent) have a success rate between 75 and 78 percent—better than random but not excellent.

Combining the results of the fifteen metrics with a simple voting scheme yields a poor-performance metric with an accuracy of 89 percent (Table 2).

Placing the detection threshold between 3 and 4 yields an 89 percent success rate with no misses and one false alarm (Figure 2). As noted above, the highest success rate achieved by any of the individual

Table 2. Summary of Test Results Showing Indicators of Poor Performance

	Instability in Funding and Quantity Profiles			Spending Forecasts			Execution to Staffing Plans				Volatility (#)				
	Funding	Quantity	Magnitude	Scatter	Magnitude	Scatter	Comparison of Contract Forecasts to Funding Plans	Cost Growth on Mission Equipment	Magnitude	Scatter	Undistributed Budget	Management Reserve	Initial Investment in System Level Tasks	System Level Tasks	SE SI ST Growth at 50% Complete
CH-53K							1	1							2
BAMS							1	1				1			2
AB3 Remanufacture	1		1						1			1			5
LHA 6												1			1
AMF JTRS	1	1	1	1	1	1	1	1		1	1			1	11
FAB-T	1	1	1			1									4
E-2D AHE (Jun 2003)							1			1	1	1			4
E-2D AHE (Jul 2009)							N/A			1	------- Not Applicable --------				1
EFV	1				1			1						1	5

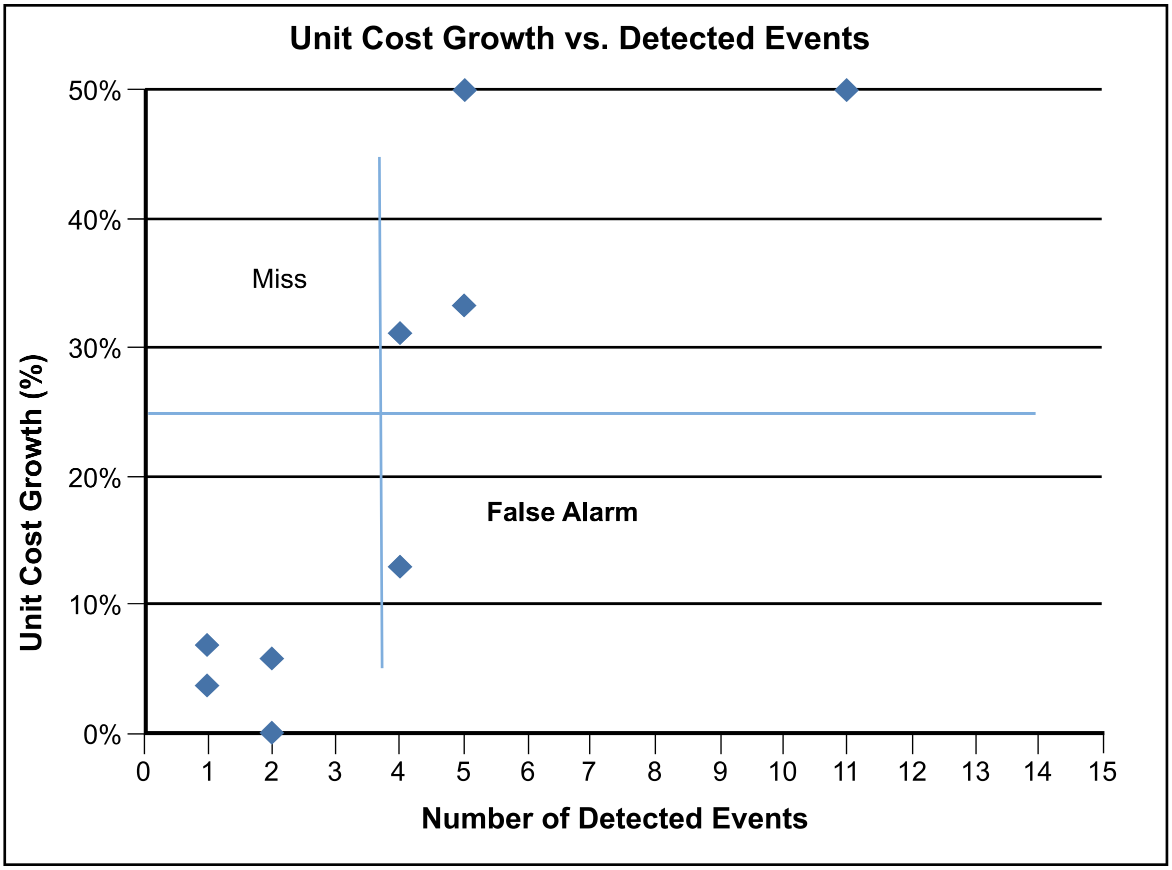


Figure 2. Unit Cost Growth vs. the Number of Events Detected

tests was 78 percent. *This result suggests that the combined test results can discriminate between those programs that are poor performers and those that are not.*

Significance tests also demonstrate that the combination algorithm was the only metric to provide sufficient evidence ($P<0.05$) to distinguish between programs that experienced UCG in excess of 25 percent and those that did not.

The concept of combining poor performance sensors to obtain improved sensor performance has a parallel in the field of radar and sensor fusion (Nicoll et al. 1991). In the early years of radar development, a graphical technique called Receiver Operating Characteristics (ROC) was used to describe how true detections and false alarms would both increase as the threshold for target declarations in a receiver was reduced. This technique recorded the sequence with which true detections or false alarms occur as the sensitivity threshold is varied from no detections to all detections.

Development of the ROC curve for the Funding Profile Instability test results is shown in Figure 3. The progression of steps can be followed by placing the poor performance threshold line (the vertical blue line in the scatter plot) to the far right at 1.0 (no detections) and moving it to

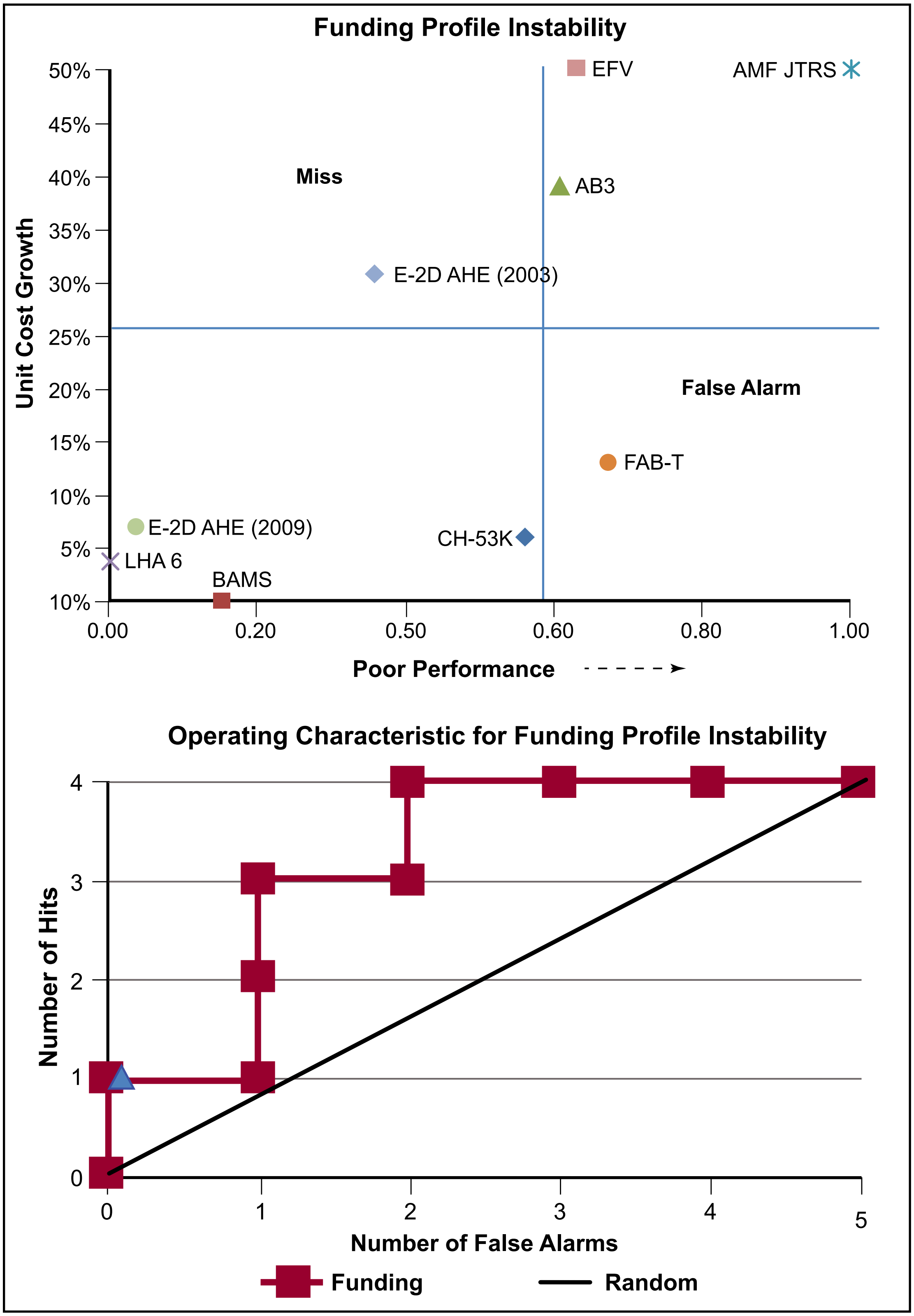


Figure 3. Operating Characteristic for the Funding Instability Metric

the left, recording +1 to the number of hits if the program marker is above the 25 percent UCG limit and +1 to the number of false alarms if the program marker is below the limit. Moving the line all the way to zero (all detections) results in four hits and five false alarms.

The “random outcome” line drawn from the origin is the expected mean for a large, normal population with a probability of detection of 50 percent (p=.5). The ROC curve for a detector with little to no value would lie close to the random outcome line. The further the ROC curve is above the random outcome line, the better the detector is at correctly identifying an event.

Development of the ROC curve for the test of Execution to Forecasts (Magnitude) metric is shown in Figure 4. This ROC curve lies closer to the random outcome line than the Funding Instability curve in Figure 3, and therefore appears to be a poorer detector than the Funding Profile Instability metric.

Figure 5 shows the ROC diagram for the combined (or fused) data in Figure 4.

This ROC diagram has two paths because the E-2D AHE (2003) and FAB-T data points have the same poor-performance index value. Either result is better than any individual test.

SUMMARY AND CONCLUSIONS

This article documents the test results for fifteen metrics of poor performance. All of the metrics use data from SARs or data from the EVM Central Repository.

These poor-performance metrics, which contribute to establishing “situational awareness” for monitoring acquisition programs, are effective tools for identifying and describing some of the problems programs encounter during the acquisition process.

The accuracy of the individual poor performance metrics in predicting UCG on the order of a critical Nunn-McCurdy breach ranges from 50 percent to 78 percent. Five of the fifteen metrics (33 percent) have a success rate between 75 and 78 percent. Combining the results of the fifteen metrics with a simple voting scheme yielded a poor-performance metric with an accuracy of 89 percent. Significance tests also demonstrate that the combination algorithm is the only metric to provide sufficient evidence ($P<0.05$) to distinguish between programs that experienced UCG in excess of 25 percent and those that did not.

The conclusion that combined results from poor detectors can exceed the detection ability of the individual detectors has a parallel in radars and sensor fusion (Nicoll et al. 1991). ROC diagrams are used to demonstrate improved performance of combined (or fused) data. This opens up the possibility that these sensor fusion techniques can be applied more broadly to MDAP-wide acquisition data in the quest for leading indicators for cost growth.

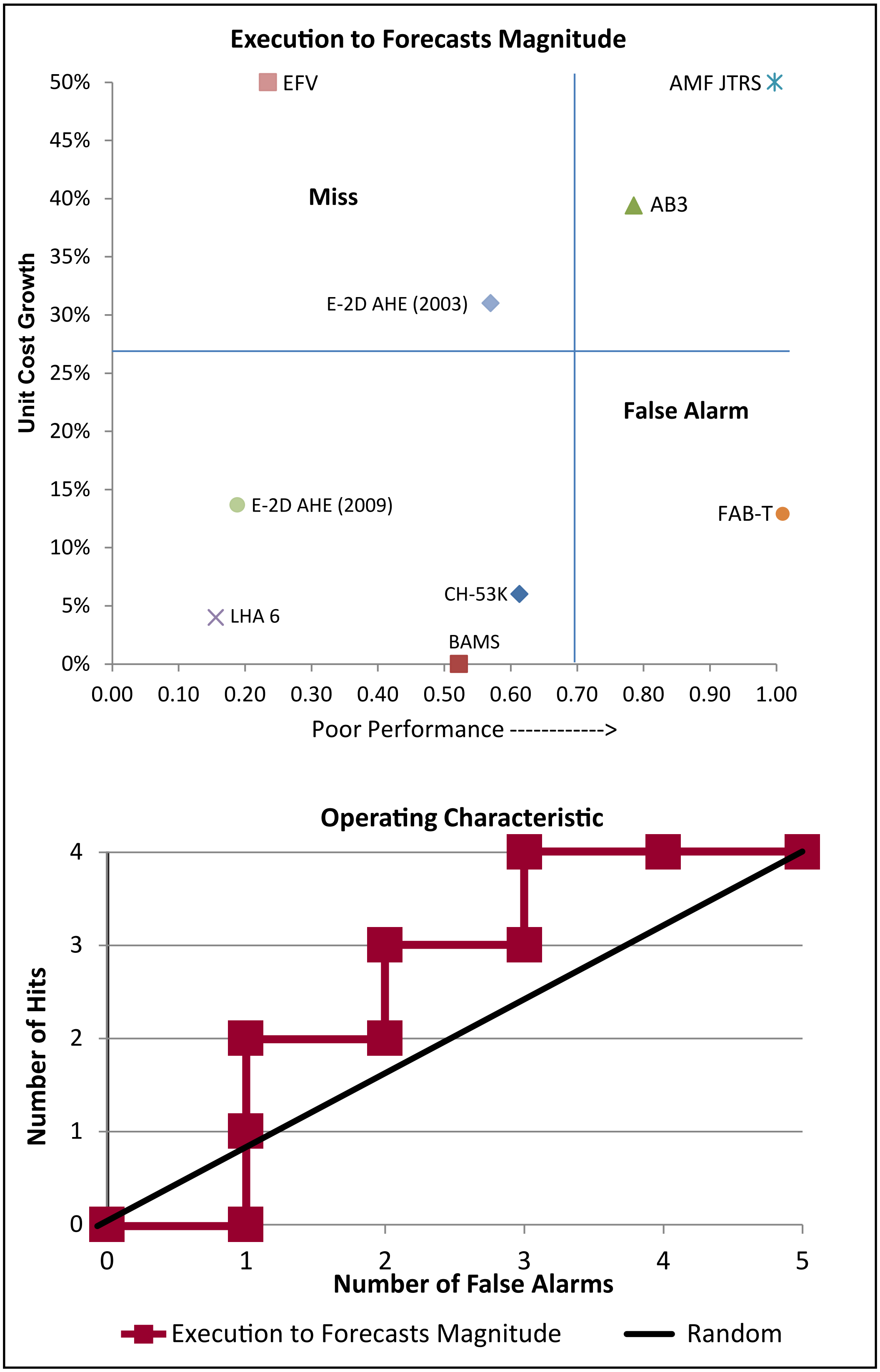


Figure 4. The Execution to Forecasts Metric

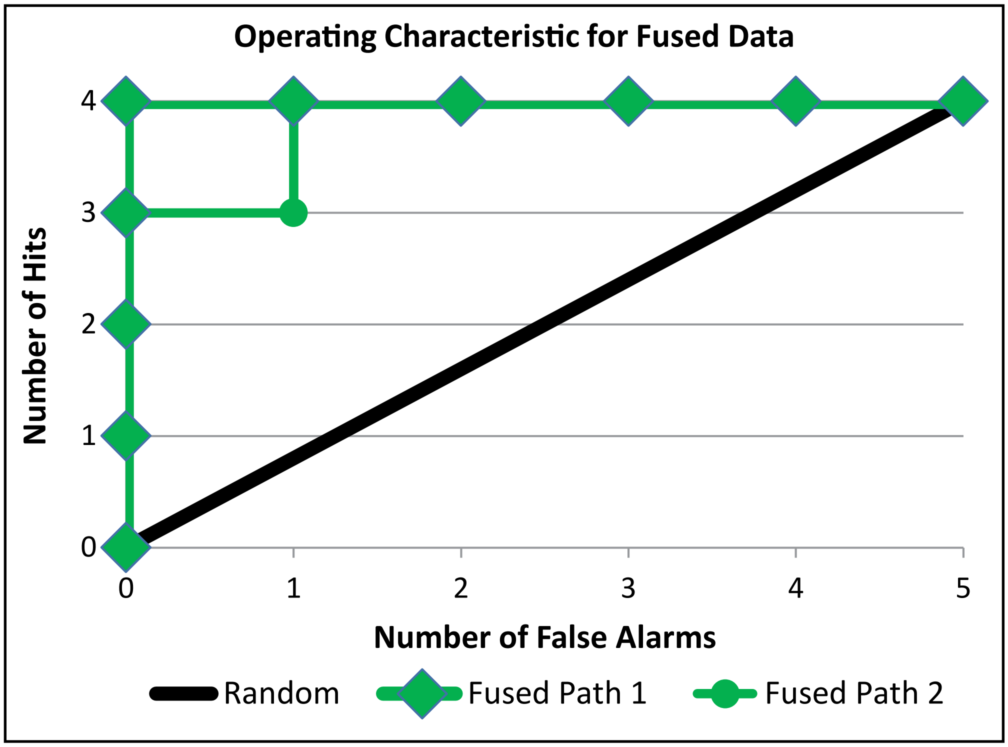


Figure 5. ROC Curve for the Fused Data

Dr. Bronson is a Research Staff Member in IDA's Cost Analysis and Research Division. She holds a doctorate in applied physics from Old Dominion University.

Dr. Sparrow is a Research Staff Member in IDA's Science and Technology Division. He holds a doctorate in physics from the Massachusetts Institute of Technology.

Reference

Nicoll, Jeffrey F., James D. Silk, and David A. Sparrow. Annual Summary Report for FY1990, Task TD-2-748, Assessment of Advanced Sensor Systems, Volume 1: Issues in Automatic Target Recognition. IDA Document D-0923. Alexandria, VA: Institute for Defense Analyses, April 1991.

ROOT CAUSE ANALYSIS OF VTUAV FIRE SCOUT'S NUNN-McCURDY BREACH

Patricia Fazio Bronson and Christopher Martin

The Problem

The root cause of the VTUAV Fire Scout Program's Nunn-McCurdy breach needs to be determined.

A Nunn-McCurdy breach occurs when a program experiences cost or schedule growth exceeding any of the established Nunn-McCurdy thresholds. The Weapon Systems Acquisition Reform Act (WSARA) of 2009 mandates that a Root Cause Analysis (RCA)—defined in WSARA as an assessment of the underlying cause or causes of growth in cost, schedule slips, or poor performance of a program—be conducted when such a breach occurs. The Director, Performance Assessments and Root Cause Analyses (PARCA) is responsible for conducting the required RCA.

At PARCA's request, IDA has conducted eleven RCAs over the past few years, in support of PARCA memoranda to the Under Secretary of Defense for Acquisition, Technology and Logistics (USD(AT&L)) describing the root causes of cost growth in programs that have experienced a critical Nunn-McCurdy breach. When the Congress requires the Department to recertify a program because it has experienced a Nunn-McCurdy breach, PARCA's memos are submitted in support of the Department's recertification decision for that program.

METHODOLOGY FOR PARCA RCAs

IDA's methodology for conducting a Root Cause Analysis of Nunn-McCurdy breaches is a four-step process centered on the Root Cause Narrative described by PARCA. This methodology repeatedly produces compelling arguments for the conclusions of the root cause:

IDA's methodology for conducting a Root Cause Analysis of Nunn-McCurdy breaches is a four-step process centered on the Root Cause Narrative described by PARCA.

1. **Official Statement of the Breach.** We take this from the Program Deviation Report to the Defense Acquisition Executive.
2. **Timeline of Events Leading up to the Breach.** We construct the initial version of the timeline from the program's historical Selected Acquisition Reports (SARs) and add to it as discovery proceeds.
3. **Root Cause Narrative.** Starting with the statement of the breach, we work backward linking the contributing factors and classifying them as symptoms; proximate causes; bad things that happened that do not have anything to do with the cost growth; and root causes. WSARA provides seven categories of root causes to consider, but permits others.

WSARA Categories of Root Causes

- Unrealistic performance expectations
- Unrealistic baseline estimates for cost or schedule
- Immature technologies or accepting excessive manufacturing or integration risk
- Unanticipated design, engineering, manufacturing, or technology integration issues arising during program performance
- Changes in procurement quantities
- Inadequate program funding or funding instability
- Poor performance by government or contractor personnel responsible for program management.

To the extent possible, we apportion the cost growth to the contributing factors, providing graphs and data as evidence without comment or conclusion.

"The purpose of the Narrative is to simply and even-handedly display the relevant facts and circumstances by which a program ended up in a ditch."(Gary Bliss)

4. **Root Cause Analysis.** We tell the story of the breach starting at the root of the problem, and allocate the contributing factors and their cost to each root cause. We discuss problems of inception and execution, and identify exogenous causes.

The RCA on the Vertical Takeoff and Landing Unmanned Aerial Vehicle (VTUAV) Fire Scout program (system shown in Figure 1) exemplifies the process.

Figure 1. MQ-8A VTUAV Fire Scout

ROOT CAUSE ANALYSIS OF VTUAV FIRE SCOUT PROGRAM

On March 10, 2014, the Program Manager for the Navy and Marine Corps Multi-Mission Tactical Unmanned Air Systems Program Office (PMA-266) submitted a Program Deviation Report that announced the VTUAV program would breach the Nunn-McCurdy critical cost thresholds of 25 percent for the Average Procurement Unit Cost (APUC) and Program Acquisition Unit Cost (PAUC) in the approved VTUAV Acquisition Program Baseline (APB).

The initial concept in 1999 for the Fire Scout was to make unmanned a Schweitzer 330 helicopter. After initial tests on the MQ-8A, the Schweitzer aircraft selected was found to have inadequate lift capacity and endurance to satisfy the Navy's desired operational needs. To address this, the Navy and Northrop Grumman further modified the aircraft to increase its performance by adding a rotor blade, extending the tail boom, and adding sponsons for additional carrying capacity. These changes to the helicopter, now designated the MQ-8B, increased the maximum gross weight and provided nearly three hours more time on station. As developmental tests proceeded, however, restrictions surfaced that limited its use.

Meanwhile, U.S. Africa Command put out an urgent request for the Rapid Deployment Capability (RDC), whose performance needs necessitated not additional purchases of MQ-8B, but the development of a third variant outside the Program of Record (POR). This new aircraft, the MQ-8C, uses a far more capable helicopter (the Bell 407), which has room and power to be a versatile weapon system. The Navy has since adopted the MQ-8C as the POR aircraft for all future procurements on the VTUAV program. According to the Navy, the more capable aircraft meant they could deploy two MQ-8Cs in place of three MQ-8Bs and reduce the total number needed to meet the Littoral Combat Ship (LCS) requirement.

Figure 2 shows major events in the evolution of the Fire Scout program from its original configuration (RQ-8A), to the configuration developed in the POR (MQ-8B), to the third configuration, developed under an RDC Joint Urgent Operational Needs (JUON) program outside the POR (MQ-8C)[1].

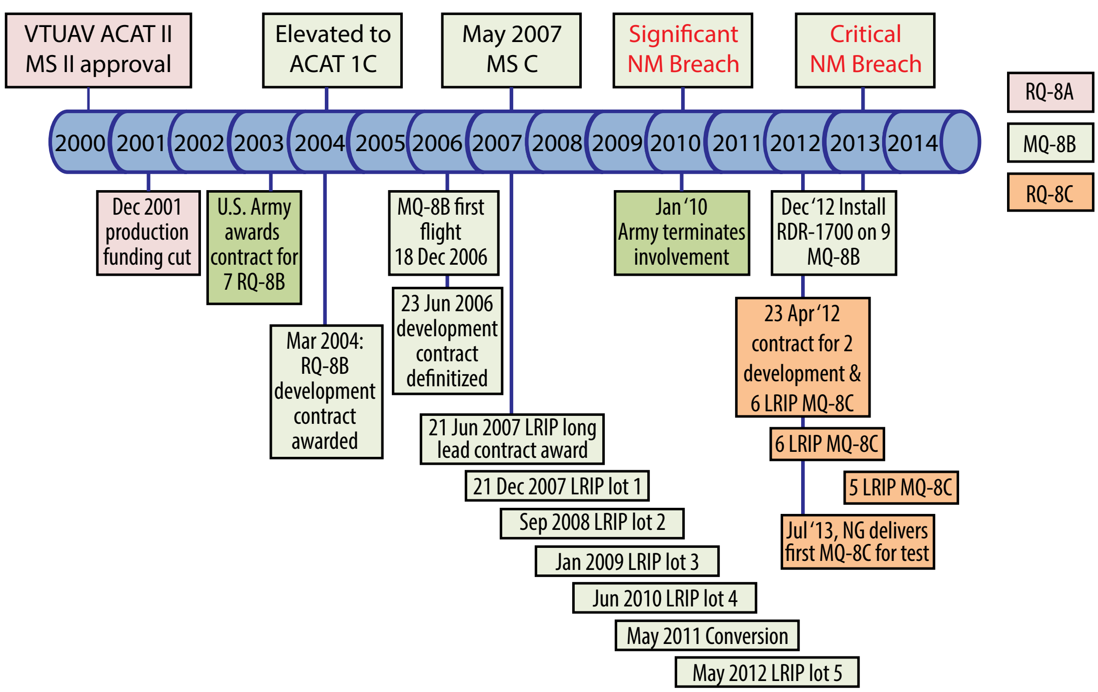


Figure 2. VTUAV Fire Scout Timeline

[1] The non-POR status means that not all costs incurred for development and procurement (of the new requirements) are included in the program baseline.

This article discusses only the critical Nunn-McCurdy breach in APUC over the original 2006 baseline, calculated to be 71.5 percent. The Navy offered that the change in aircraft purchase quantity was the sole reason for increased APUC values for VTUAV. Since quantity change rarely occurs alone, we believe that determination of other factors must precede the quantity change calculation.

The VTUAV SARs reveal that the Navy added $327 million (FY06 dollars) to the procurement program and stretched the procurement schedule by 17 years before they decided not to procure any more MQ-8Bs. Figure 3 shows the VTUAV Fire Scout planned procurement profiles from the annual SARs.

A fixed cost analysis of annual procurement cost by quantity yields an annual fixed cost estimate of approximately $20 million, which translates into a $340 million increase in the procurement cost estimate over the course of the program. The major contributors to the schedule delays were a replan of the initial, unrealistic acquisition profile; better alignment of aircraft procurement with LCS procurement; a five-year delay in the scheduled date for operational evaluation (OPEVAL); and allowance for more time for the development of the MQ-8C.

Cost experience on procurement of the VTUAV provided by the Navy revealed that recurring costs for the MQ-8B had increased by 7 percent and that the more capable MQ-8C would cost about $1 million more than the MQ-8B's current estimated cost (another 11 percent).

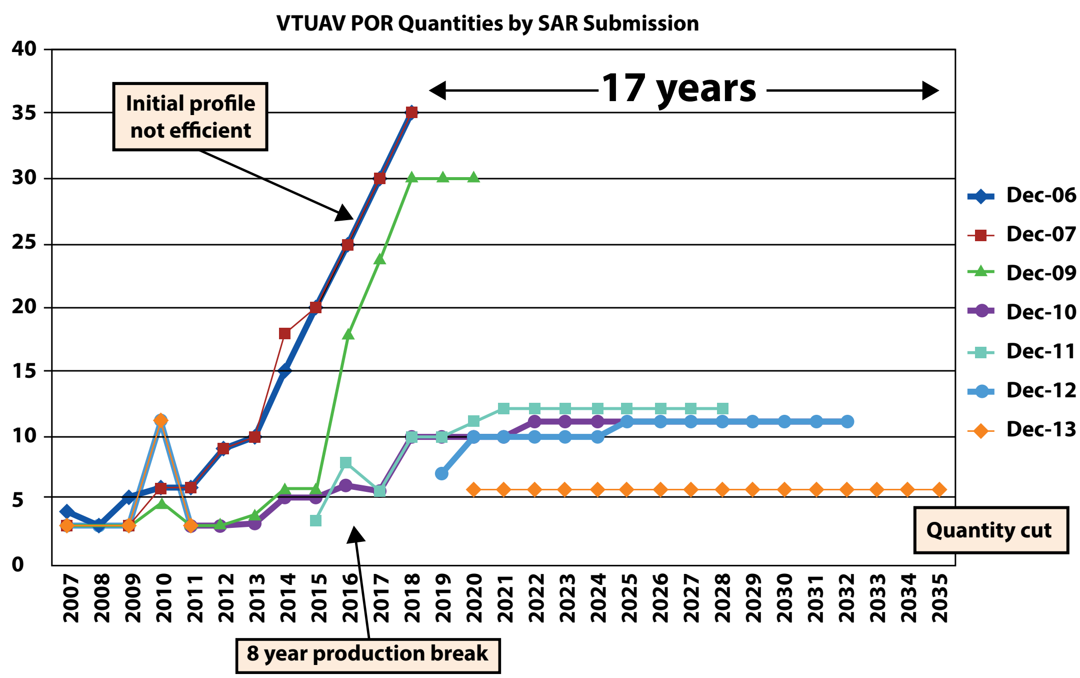


Figure 3. VTUAV POR Quantities by SAR Submission

Figure 4 shows IDA's estimates of the relative contributions for each of the reasons. There remain 32 percentage points of APUC growth for which the quantity change is accountable.

With 32 percent of the APUC increase associated with a reduced number of aircraft systems to be bought, the question is why the Navy decided to stop buying MQ-8Bs, an aircraft the Navy had reported met operational requirements. In reviewing test reports and performing analyses of demonstrated performance, we found that, during developmental testing and through experience in the field, the MQ-8B had restricted wind envelopes and reliability issues, and required greater engine maintenance costs than anticipated.

Figure 5 shows the aggregate Wind Over Deck envelopes[2] allowed for the MQ-8B on board LCS-1 superimposed on the objective and threshold requirement.[3] The white areas within the threshold "fan" are

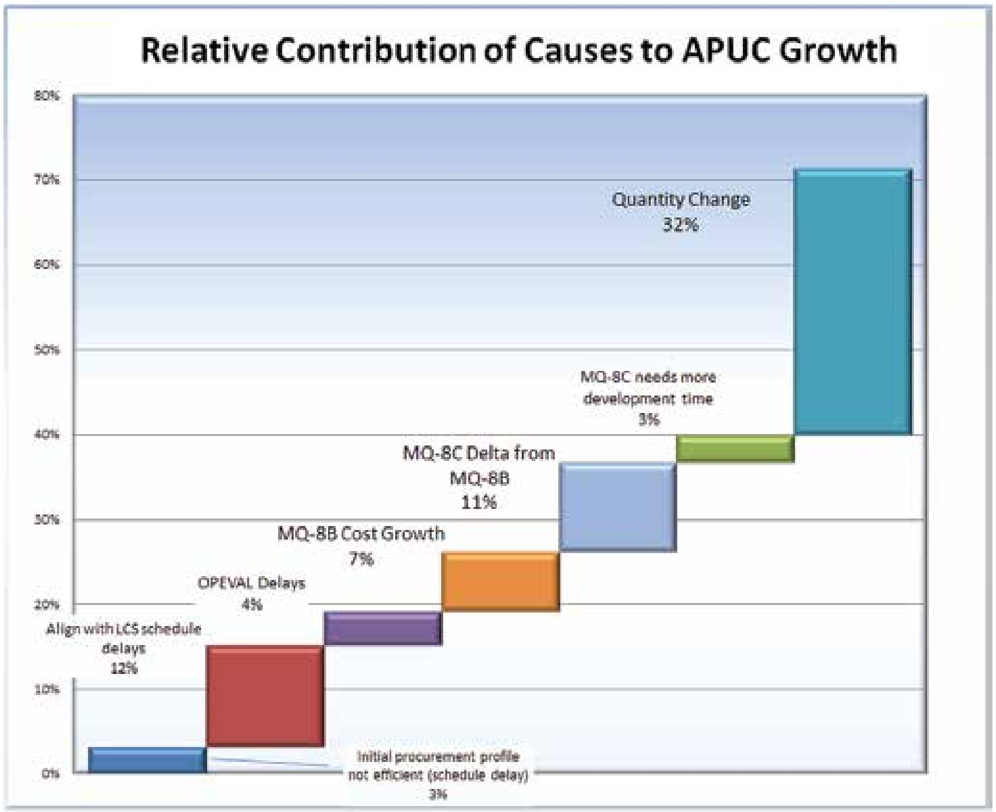


Figure 4. Relative Contributions of Causes of APUC Growth

[2] *Naval Air Training and Operating Procedures Standardization (NATOPS) Flight Manual*, Navy Model MQ—8B, Unmanned Aerial Vehicle, November 1, 2013 Change 1—March 1, 2014, Document Number A1—MQ8BA—NFM—000.

[3] Capabilities Production Document, Vertical Takeoff and Landing Tactical Unmanned Aerial Vehicle (VTUAV) System, Prepared for Milestone C Decision, Version 4.3, December 20, 2006.

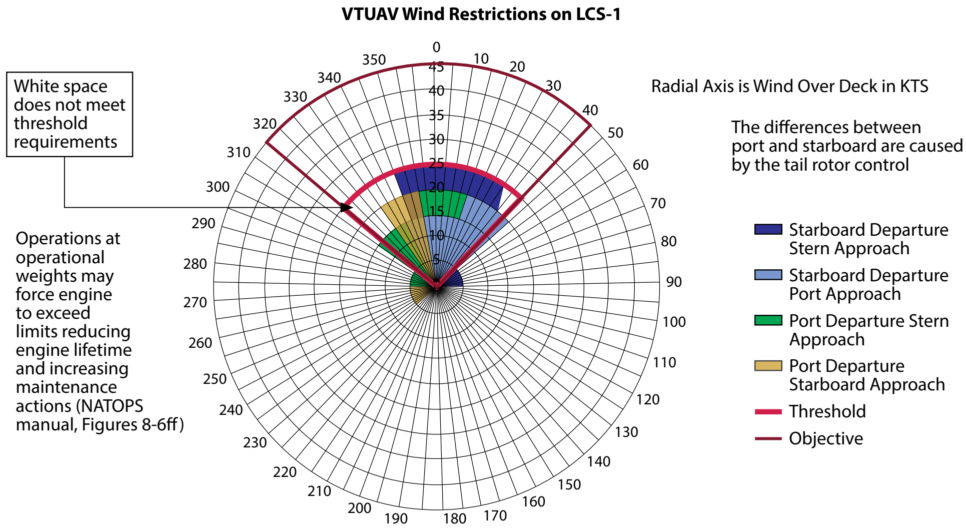


Figure 5. Aggregate Wind Over Deck Envelopes Allowed for MQ-8B On Board LCS-1

conditions in which the system does not meet the threshold requirement. The shortfalls seen in Figure 5 are due primarily to inadequacies in tail rotor authority at nominal power settings and aircraft weights.[4]

In addition, operational assessments revealed that the aircraft needed to operate at near maximum gross weight to meet time on station requirements. The engine operations at high power settings needed to meet these requirements reduced the mean time between failures, increased mean time to repair, and required additional spare parts.[5] Furthermore, the aircraft could not be operated at standard military hot temperatures due to engine overheating issues.[6] Cooling air scoops[7] are being added to the current MQ-8Bs so they can safely operate at temperatures above 30°C.

IDA assesses that the Navy abandoned the MQ-8B because it did not meet their performance expectations. Because of this, we find that the increases in recurring cost, the OPEVAL delays, the development time for the MQ-8C, and the change in quantity are all consequences of the MQ-8B design not meeting performance expectations.

Furthermore, we assess that even with the modifications made

[4] CDR Van Patrick McLawhorn, USN, and Mr. Richard Paletta, "Status of MQ-8B Developmental Testing, Inclusive of Software Increment 9.1.3.1," Report No: NAWCADPAX/ISR-2012/62, May 24, 2012.

[5] Quick Reaction Assessment of VTUAV, COMOPTEVFOR, September 12, 2012.

[6] McLawhorn and Paletta, "Status of MQ-8B Development Testing," May 24, 2012.

[7] MQ-8 AIRFRAME CHANGE NO. 16. To provide modification instructions for Engine Bay Access Panels to add cooling air scoops to reduce operating oil temperature levels.

from the MQ-8A to the MQ-8B (increasing vehicle weight, improving rotor performance, and increasing engine horsepower), the Schweitzer 330 aircraft and engine were limited in overall capability and were not going to be able to meet the payload and endurance expectations across the range of intended operational conditions. Because of this, we allocate 56 percentage points of the 71.5 percentage points in APUC growth to the WSARA Root Cause Category "poor performance by government personnel responsible for program management." The remaining APUC growth we attribute to a faulty initial procurement plan and LCS schedule delays.

Dr. Bronson is a Research Staff Member in IDA's Cost Analysis and Research Division. She holds a doctorate in applied physics from Old Dominion University.

Mr. Martin is a Research Staff Member in IDA's Science and Technology Division. He holds a Master of Science in aerospace engineering from the University of California, Los Angeles.

References

McLawhorn, Van Patrick, CDR, USN, and Mr. Richard Paletta. Status of MQ-8B Developmental Testing, Inclusive of Software Increment 9.1.3.1. Report No: NAWCADPAX/ISR-2012/62. May 24, 2012.

MQ-8 AIRFRAME CHANGE NO. 16. MQ-8B Hot Weather Delta Modification, Installation of (WUC 1000000). Cognizant Code(s): PMA-266/AIR-4.3.5.1/6.7.1.2. Issue Date: September 27, 2012.

MQ-8B Capabilities Production Document (CPD). *Vertical Takeoff and Landing Tactical Unmanned Aerial Vehicle (VTUAV) System*, Version 4.3. Prepared for Milestone C Decision, December 20, 2006.

Naval Air Training and Operating Procedures Standardization (NATOPS) Flight Manual, Navy Model MQ–8B, Unmanned Aerial Vehicle, November 1, 2013 Change 1 – March 1, 2014. Document Number A1–MQ8BA–NFM–000.

Quick Reaction Assessment of VTUAV, from Commander, Operational Test and Evaluation Force, to Chief of Naval Operations. September 12, 2012.

"VTUAV Program Deviation Report" from PMA-266 to ASN(RDA), Mar 10, 2014.

EVALUATING SOLID ROCKET MOTOR INDUSTRIAL BASE CONSOLIDATION SCENARIOS

Brian Gladstone, Brandon Gould, and Prashant Patel

The Problem

Diminishing demand for large solid rocket motors (SRMs) since 2010, coupled with plans to end NASA's programs that utilize them, has caused DoD concern regarding SRM industrial base sustainability and unit cost increases.

DoD depends on private industry to design and produce its weapon systems. A healthy industrial base for weapon systems is needed to ensure competition exists to control price and create multiple procurement options; redundancy of prime and sub-tier suppliers; and a continuous labor pipeline of scientific, engineering, and manufacturing expertise. DoD's demand for these weapons significantly affects the survivability of corporations as well as that of sub-tier suppliers. In addition, corporate choices to consolidate or leave the DoD market also have an impact on the defense supply market. DoD is increasingly finding itself with scarce suppliers for many commodities.

One recent example is the large solid rocket motor (SRM) industrial base, which has been reduced to two prime manufacturers—Aerojet and ATK—and faces extensive challenges with ever-decreasing demand from NASA and DoD. As a result of significant decreases in demand, the industrial base was oversized for expected large-SRM production, and SRM stakeholders became increasingly concerned about resulting unit cost increases and industry viability. In 2010, the Congress directed DoD and NASA to develop an industrial base sustainment plan for large SRMs.

IDA was asked to evaluate (1) whether the SRM industrial base could withstand the near-term and long-term impacts of decreased SRM demand as they were envisioned in 2010, (2) whether there are viable consolidation options, and (3) the long-term costs/savings and schedule impacts of consolidation.

STRUCTURE AND HISTORY OF THE SRM INDUSTRIAL BASE AND DEMAND FOR SRMs

Once composed of six SRM suppliers in the 1994 timeframe, the current SRM industrial base now comprises only two manufacturers—Aerojet and ATK. In addition, both companies rely on a very thin industrial base of sub-tier, often single source, suppliers. For example, AMPAC (WECCO) is the single source supplier of ammonium perchlorate, a ubiquitous and major component of propellant for SRMs.

The federal government, primarily through NASA and DoD, is largely a sole consumer, purchasing SRMs for space launch, strategic systems, missile defense (both large and small SRMs), and tactical systems. Figure 1 depicts examples of these SRMs, of which space launch at NASA consumed the most propellant by a large margin. Figure 2 depicts the historical demand for SRM propellant during 1990–2010 and the anticipated demand for propellant for 2011–2027. The significant drop in the 2010–2011 timeframe was due to the end of the space shuttle program and the cancellation of the Constellation program. However, decreases in demand for strategic systems also contributed to this decrease.

VIABILITY OF SRM PRIME CONTRACTORS AND SUB-TIER SUPPLIER RISKS

IDA interviewed representatives from GenCorp (Aerojet), ATK, and AMPAC and evaluated their companies' credit metrics against a set of benchmark companies (Pre-Castparts, Hexcel, S&P 500, DoD Primes). All three companies stated that they could withstand the decreased demand. Our evaluation of the credit metrics led to the following conclusions:

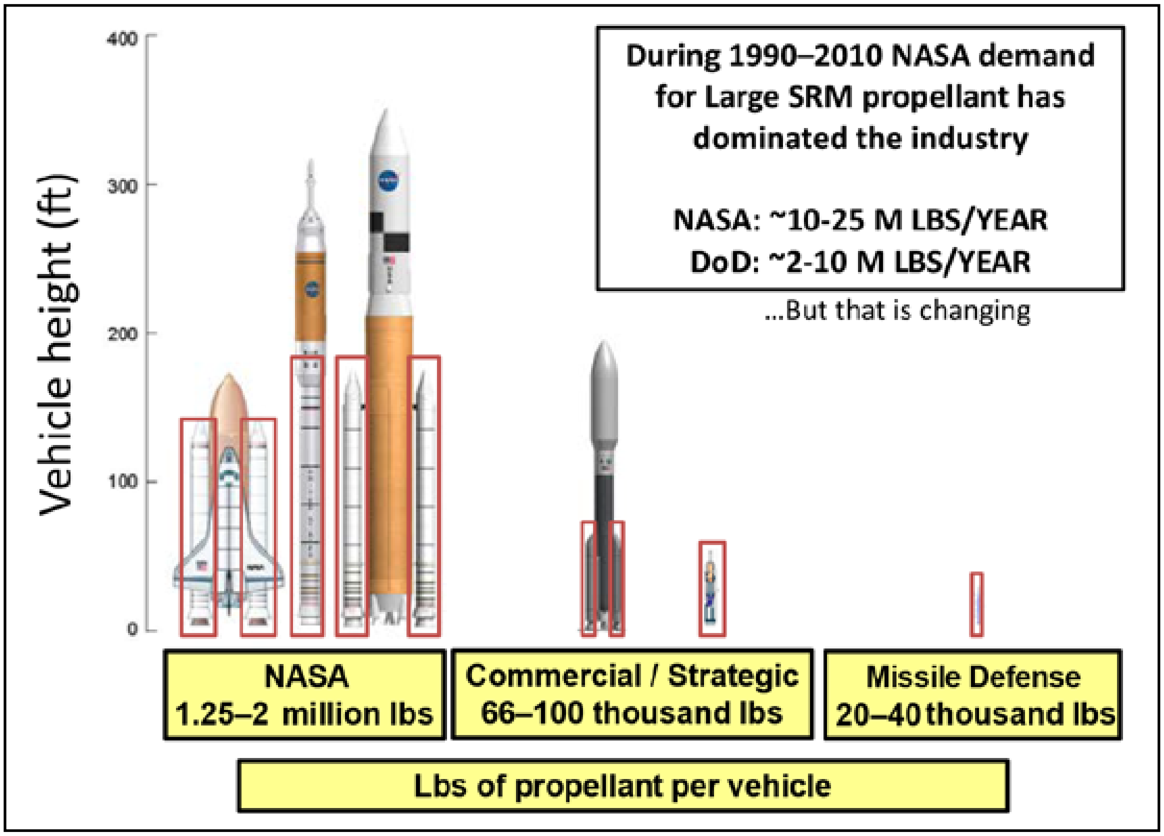


Figure 1. Examples of Current and Planned Large SRM Platforms

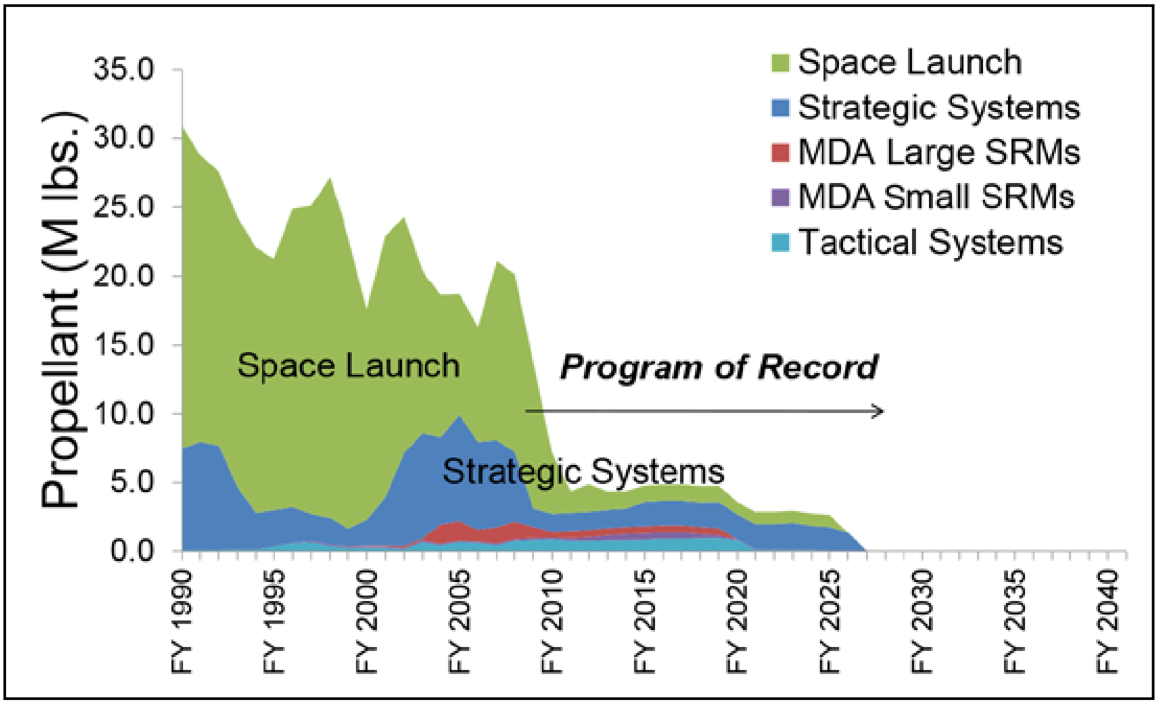


Figure 2. Historical and Future Demand for Large SRMs

1. SRM motor firms are rated high yield, also known as "junk" (below BBB-).
2. SRM firms are significantly more leveraged than the benchmarks; however, this should be manageable in the near term.
3. AMPAC's interest coverage ratio implies a D rating—cash is declining when most firms are accumulating; however, the ammonium perchlorate business is very profitable.

Thus, although there is some risk, the SRM primes will be viable in the near term.

IDA also evaluated sub-tier suppliers. Figure 3 displays actively managed suppliers grouped by our evaluation of the risk associated with each. The level of risk was assigned based on the number of programs affected, various supply issues, and whether they were a single manufacturer or sole source (or foreign supplier). An additional fourteen materials (top of Figure 4) have the potential to affect multiple programs or families of SRMs; four of these were from foreign suppliers. Seven additional materials (bottom of Figure 4) have the potential to affect a single program or family of SRMs; three of these are from foreign suppliers.

EVALUATION OF THE COST AND SCHEDULE IMPACTS OF VARIOUS CONSOLIDATION SCENARIOS

IDA evaluated the following scenarios:

1. The current industrial base: Aerojet at Sacramento and ATK at Promontory and Bacchus, Utah

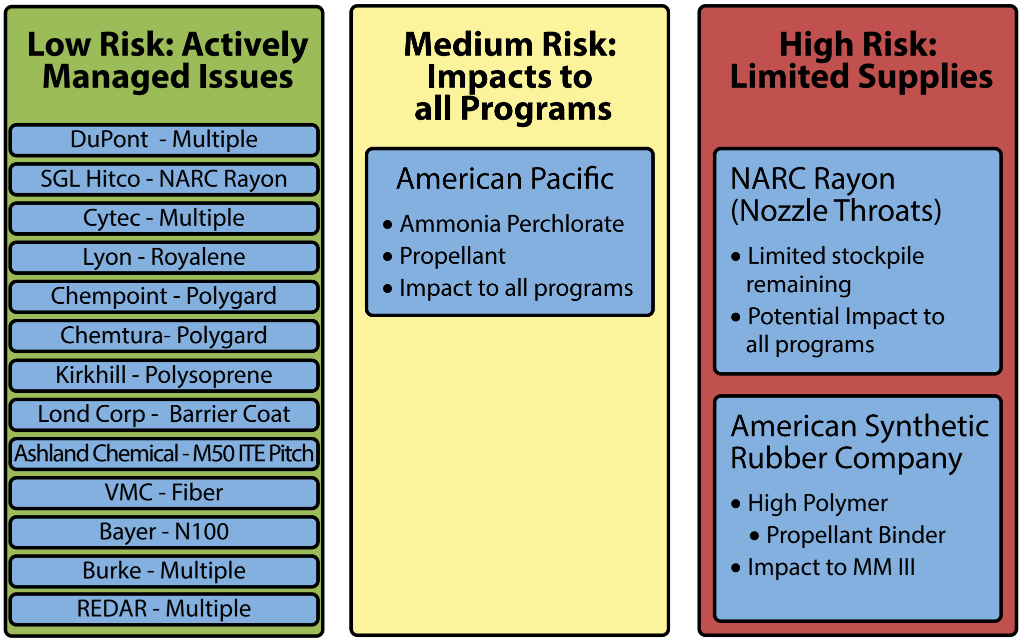


Figure 3. Actively Managed Supplier Issues

	Company	Material	Sole Source / Single Manufacturer	Domestic or Foreign
Impact to Multiple Programs / Families	Toyal America, Inc.	Spherical Alum. (X-86)	Single	US
	Sartomer	HTPB Polymer	Single	US
	Reinhold Industries	Phenolic components, molded & nozzles	Single	US
	Henkel Aerospace	Epoxy adhesives	Single	US
	Parker Hannifin CSS	Seals, gaskets	Single	US
	Arrowhead Products	Tank and Bladder	Single	US
	General Plastics	Sliver Insulation	Single	US
	Boulder Scientific Company	TPB	Sole Source	US
	Fiber Material Inc. (FMI)	ITE, C-C	Sole Source	US
	3M	HX-752	Sole Source	US
	Toral, Toho, Hexcel	Graphite Fiber		FOREIGN
	ENKA	Rayon	Sole Source	FOREIGN
	Cognis Specialty Chemical	(PEG)4500 - Motor; DDS - Igniter	Sole Source	FOREIGN
	Yong San Chemical	MA (precursor to MNA stabilizer)	Sole Source	FOREIGN
Impact to Single Program / Families	R.E. Darling	Insulation	Single	US
	Energy Research & Generation	Nozzle Baffle	Single	US
	Dongin Chemicals	2 NDPA and MNA	Sole Source	FOREIGN
	Hagedorn NC	NC	Sole Source	FOREIGN
	Heroux-Devtek	Bracket	Sole Source	FOREIGN
	Emerald Performance	HC (CTPB) Polymer	Sole Source	US
	Talley Defense System	Gas Generator	Sole Source	US

Figure 4. Potential Supplier Risk

2. Aerojet with a consolidated ATK: Aerojet at Sacramento and ATK at Bacchus with a Promontory test facility
3. ATK monopoly: ATK at Bacchus with a Promontory test facility
4. Green field government facility with contractor operators; all other sites are closed.

To evaluate the cost and schedule impacts of consolidation scenarios, the research team first had to understand the cost drivers and demand from commercial, foreign, and U.S. government entities. Next, we had to determine the total cost of the industry. After receiving data and reports from the contractors involved and various government agencies, we developed statistical relationships and found analogies or other analyses to enable estimates of total large-SRM industry costs under the scenarios listed above. Two categories of costs were considered for our estimates of these scenarios: near-term transition costs and long-term total plant costs.

The following near-term transition costs were calculated based on actual historical costs and analogies combined with IDA-derived cost models:

- Buy-out: The cost to buy out a contractor
- Close-out: The costs to close a facility or site
- Requalification: The costs to retest and requalify SRMs after a change in material, production process, etc.
- Facilities, tooling, and training.

The long-term total plant costs were:

- Direct material and direct labor
- Overhead
- Fee
- Environmental liabilities.

We first calculated the total cost of Scenario 1, the current industrial base, which was approximately $1.2 billion per year at the time of this analysis. Next, we developed cost models to evaluate SRM industry costs relative to this baseline that were consistent with the consolidation scenario. In addition, we performed risk analyses to determine upper and lower bounds of costs, savings, and schedule impacts. The results of these scenarios are illustrated in Figure 5 and Figure 6.

Of the consolidation scenarios evaluated, only Scenario 2 (Aerojet with a consolidated ATK) made fiscal sense. Internal consolidation at ATK, as defined in this analysis, had a near-term cost between $500 million and $800 million and a three-year production gap; it is likely that these costs will be recovered prior to 2035, saving the taxpayer between $0.1 billion and $1 billion. Neither the ATK monopoly scenario (Scenario 3) nor the green field government-owned, contractor-operated (GOCO) consolidation scenario (Scenario 4) made fiscal sense. The ATK monopoly and green field GOCO scenarios have considerable near-term transition costs of $1.4 billion to $6.1 billion over three to nine years as well as a three- to nine-year production gap. In addition, for both of these scenarios, it is unlikely that these costs will be recovered prior to 2035.

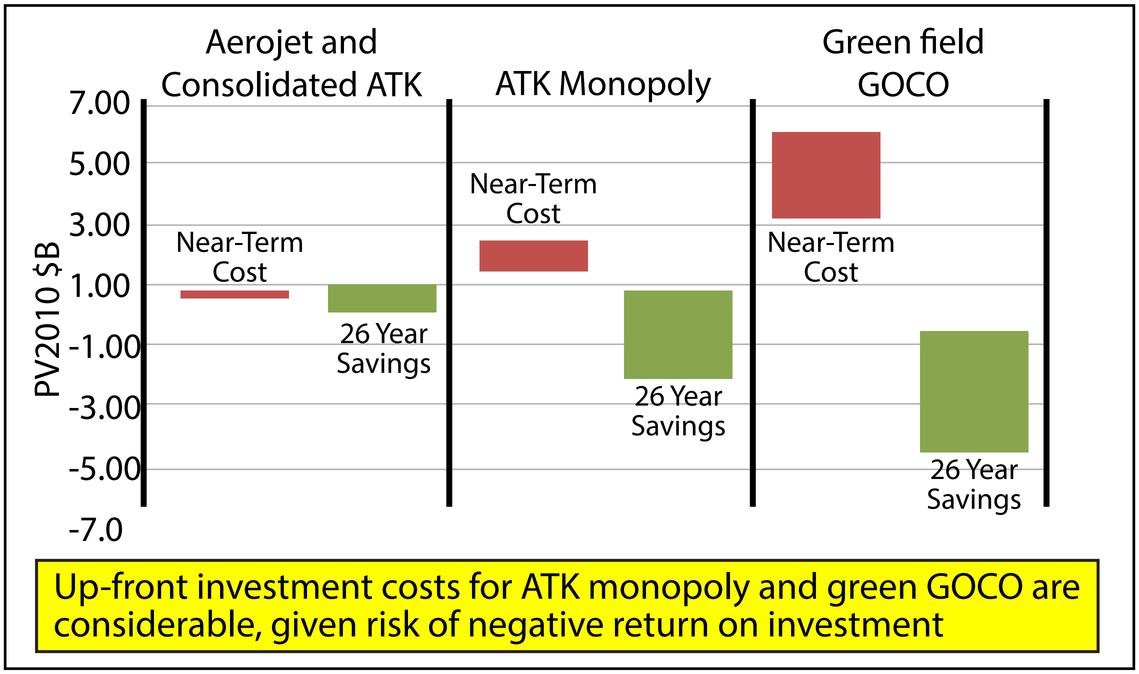


Figure 5. DoD Investment and Net Savings: 2010–2035

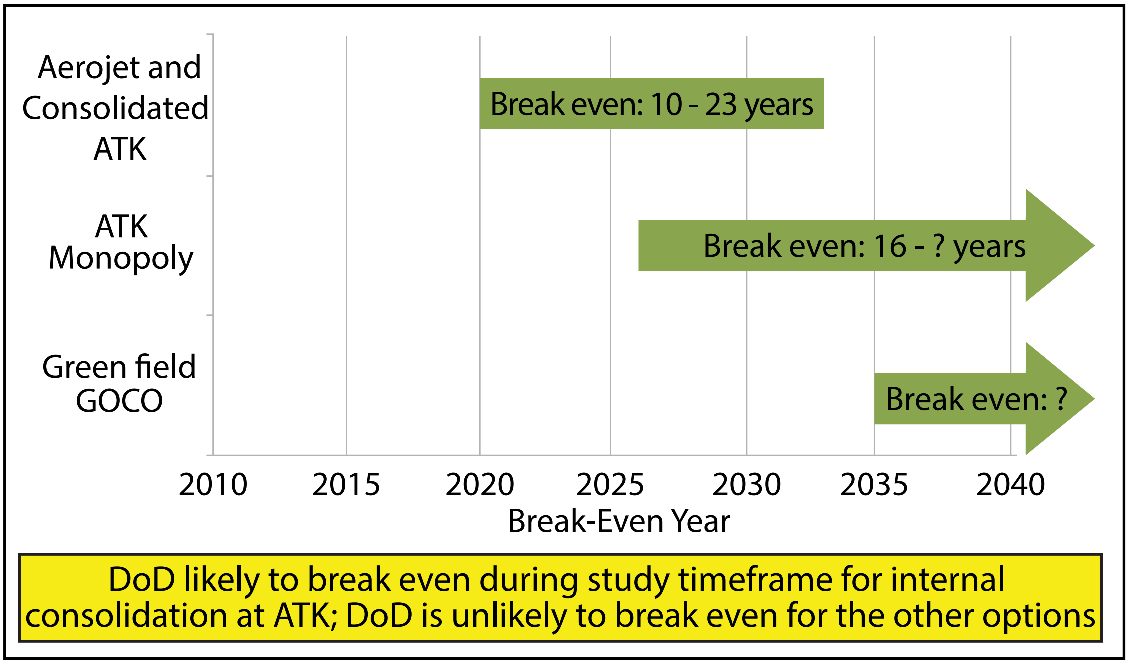


Figure 6. DoD Break-Even Time Frames

CONCLUSIONS

IDA determined that internal consolidation at ATK made the most financial sense, while the significant near-term costs, coupled with ineffective savings, made the ATK monopoly and green field GOCO options unattractive. In addition, our analysis anticipates that the large-SRM producers and subtier suppliers would likely survive. DoD proceeded with the desire to move toward Scenario 2 (Aerojet with a consolidated ATK), and encouraged ATK to consolidate its operations.

Mr. Gladstone is a Research Staff Member in IDA's Cost Analysis and Research Division. He holds a Master of Science in chemical biology/biochemistry from Harvard University.

Mr. Gould is a Research Associate in IDA's Cost Analysis and Research Division. He holds a Master of Arts in economics and foreign affairs from the University of Virginia.

Dr. Patel is a Research Staff Member in IDA's Cost Analysis and Research Division. He holds a Doctor of Philosophy in aerospace engineering from the University of Michigan.

MANAGING SUPPLY CHAIN RISKS TO DOD SYSTEMS AND NETWORKS

Thomas Barth, Michelle Albert, and Elizabeth McDaniel

The Problem

Global supply chains are vulnerable to attack or manipulation. If an adversary compromises an information communications technology (ICT) component by exploiting vulnerabilities associated with its supply chain, that component can become a system risk. ICT supply chain risk management (SCRM) relies on an in-depth understanding of the risks inherent to global supply chains and the means by which to mitigate or manage those risks.

IDA's research on ICT SCRM issues and responses contributes to the system security and performance of DoD systems and networks.

The environment in which the Department of Defense (DoD) acquires, builds, and operates national security systems is increasingly globalized. DoD weapon systems, business systems, and computer networks are dependent on commercial information communications technology (ICT) and ICT-enabled components. These components are designed, manufactured, packaged, and delivered to end users through global supply chains that create interconnected webs of people, processes, technology, information, and resources around the world (Figure 1).

Figure 1. The Supply Chain

These supply chains give DoD access to available technologies and innovations, but they also provide adversaries with opportunities to tamper with ICT components by introducing malicious code, reverse-engineering the components to access design information, finding or adding vulnerabilities they can later exploit, or inserting counterfeit parts that may increase failure rates in systems or provide additional avenues for exploitation.

The DoD Chief Information Officer (CIO), the office responsible for DoD's supply chain risk management policy, leads risk-reduction activities for DoD's acquisition programs. For over a decade, IDA has supported the DoD CIO's efforts to enhance supply chain risk management (SCRM).

Traditionally, the term SCRM refers to the logistics associated with obtaining needed components, a major concern of companies. ICT global SCRM, on the other hand, deals with targeted threats; as such, assessing and managing risk for the security of the systems requires a different set of tools. IDA's research on ICT SCRM issues and responses contributes to the system security and performance of DoD systems and networks.

ICT SCRM AWARENESS MODULE

In support of the DoD CIO, IDA published an education, training, and awareness module to promote awareness of the risks inherent to the ICT global supply chain and to increase understanding of ICT SCRM. The module captured then-current DoD policies and processes. It has a flexible format that allows for revision when policies and/or processes change and for customization for various audiences. The module, which covers SCRM throughout the acquisition lifecycle, is organized around three themes:

Theme 1. The New Insider Threat Is Not a Person – It's ICT

Theme 2. Supply Chain Risk Is a Condition To Be Managed, Not a Problem To Be Solved

Theme 3. Take Action To Manage Global Supply Chain Risk

THE NEW INSIDER THREAT IS NOT A PERSON – IT'S ICT

Theme 1 identifies the elements of SCRM and explains the national security risks associated with global supply chain exploitation. The module defines ICT as technology used for gathering, storing, retrieving, and processing information, including microelectronics, printed circuit boards, computing systems, software, signal processors, mobile devices, satellite communications, and networks.[1]

Today, most of the ICT components used in DoD systems and networks are obtained from commercial sources. These commercial products take advantage of global talent, resources, and manufacturing capabilities, and typically can be purchased at lower cost than other products. However, the globalization that lends these advantages also creates supply chains that are often opaque and difficult to trace, thereby creating security challenges. Products traverse borders and companies many times on their way to the end user and their point of integration into DoD systems or networks.

[1] IDA adapted this definition from DoD Instruction 5200.44, *Protection of Mission Critical Functions to Achieve Trusted Systems and Networks (TSN)*, and NIST Draft IR 7622, *Notional Supply Chain Risk Management for Federal Information Systems.*

Supply chain attacks can occur throughout the DoD system development lifecycle; entry points for exploitation and manipulation include component design, manufacturing, transport, delivery, installation, and repair or upgrade. Figure 2 illustrates a notional supply chain and highlights possible entry points for an adversary to manipulate or tamper with a component.[2] ICT components are often stored or transported in ways that leave them open to tampering. The notional points of manipulation in Figure 2 illustrate vulnerabilities in the supply chain environment that create opportunities for specific attacks. The impacts of such attacks can be disruption of service, insertion of malicious functionality, data exfiltration, and theft of intellectual property. The goal of supply chain risk management is to reduce a component's or system's susceptibility to supply chain threats and the potential impact of those threats.

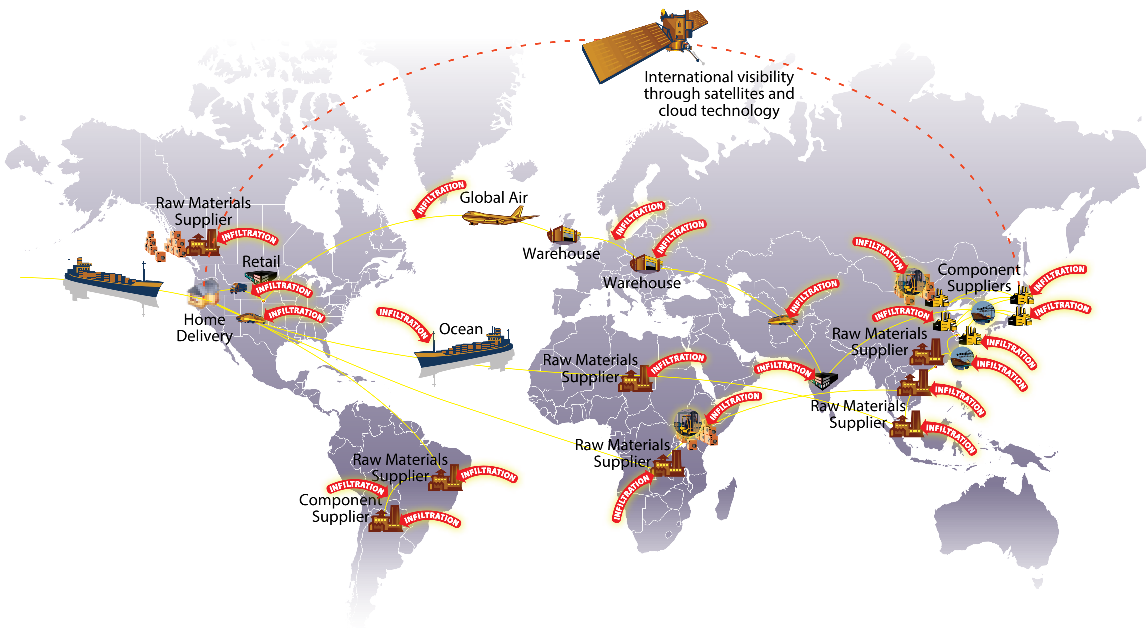


Figure 2. Manipulation along the Supply Chain

As an ICT component traverses its supply chain, it passes from country to country, company to company, and person to person. Each company has its own logistical security standards, and

[2] In the awareness module, components and systems obtained through simple procurement or as part of the Defense Acquisition Management System (DAMS) are described as having system development life cycles that span design through disposal. The module refers to the Joint Capabilities Integration Development System (JCIDS) process as the Requirements phase. The module combines some other DAMS phases for ease of understanding. The Acquisition phase refers to design, development, testing, production, and deployment; the Operation and Sustainment phase refers to operations and support; and the Disposal phase refers to system or component disposal.

SUPPLY CHAIN RISK IS A CONDITION TO BE MANAGED, NOT A PROBLEM TO BE SOLVED

Theme 2 explains why supply chain risks must be managed, discusses the key concepts of risk management in the context of ICT SCRM, and offers a range of responses to identified risks. It is not possible to anticipate or eliminate every vulnerability, so risks must be managed. And, since everything is connected today, one ICT component that has been tampered with in a DoD system or network can affect that one system or multiple systems. Risk management must be considered for every ICT component purchased or integrated into a system.

If the assessed risk is high, DoD has four basic responses: treat it, tolerate it, transfer it, or terminate it (Figure 3). Treating the risk means applying countermeasures and mitigations to lessen the consequence of a compromised component or system by incorporating risk management strategies throughout a component or system's life cycle.[3]

Transferring, tolerating, or terminating the risk should be considered if it is better to treat the risk at a later time, if there are insufficient resources to treat it now, or if available treatment options do not reduce the risk to an acceptable level. SCRM options

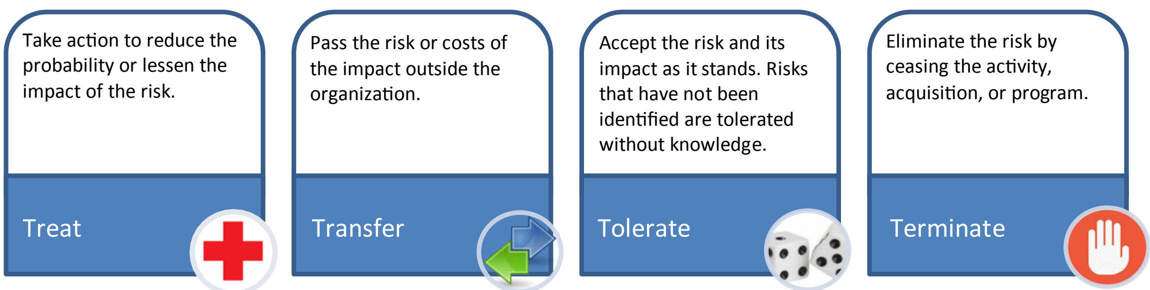


Figure 3. Four Responses to Identified Risk

range from doing nothing, which entails no effort or extra costs up front, to redesigning a system to avoid using a component with unacceptable risk mitigation options, which involves more effort and higher costs.

ICT SCRM begins with identifying critical components and functions, vulnerabilities, and threats to the supply chain, and developing strategies to respond. Limits on time and money require DoD to focus on risks to mission-critical functions, those functions that, if compromised, could degrade a system's ability to meet its core mission.

[3] According to the Joint Doctrine, countermeasures are devices or techniques applied to impair the operational effectiveness of adversary activity. In the context of ICT SCRM, countermeasures prevent adversaries from exploiting supply chain or component vulnerabilities. Mitigations are actions taken to alleviate the risks or effects resulting from vulnerabilities in critical components or systems.

TAKE ACTION TO MANAGE GLOBAL SUPPLY CHAIN RISK

Theme 3 describes the current complex, dynamic, and evolving environment of relevant government and DoD policies, standards, and strategies that guide the management of supply chain risk across the phases of the system development life cycle. DoD has articulated requirements in Acquisition policy (DoD Instruction 5000.02, *Operation of the Defense Acquisition System*, and DoD Instruction 5200.44, *Protection of Mission Critical Functions to Achieve Trusted Systems and Networks (TSN)*) and in cybersecurity policy (DoD Instruction 8500.01, *Cybersecurity*, and DoD Instruction 8510.01, *Risk Management Framework (RMF) for DoD Information Technology (IT)*). In combination, these policies provide guidance on ICT SCRM for DoD personnel.

Although much of the available policy and guidance focuses on the acquisition phase, most warfighting, intelligence, and business systems, products, and services spend the majority of their existence in the operations and sustainment phase. Risk management is essential during the design and manufacture phases of acquisition, but it is also critical during operations, routine services, maintenance, and planned upgrades or modifications.

GOING FORWARD

The ICT Global SCRM Awareness Module was designed to prompt DoD personnel to care, think, and act in response to real risks that result from the supply chains of ICT products across the life cycle. It was designed to support the efforts of Combatant Commands, Services, and agencies to understand and implement DoD's policies effectively in the face of real threats to system security and performance.

The ICT SCRM Awareness Module on DVD is available upon request; email ETASCRM@ida.org. The module's introductory video is featured on IDA's website at https://www.ida.org/SAC/SACResearchDivisions/ITSD/ITSD_Ideas_Home.aspx.

Mr. Barth is a Research Staff Member in IDA's Information Technology and Systems Division. He holds a Master of Arts in strategic studies from the U.S. Army War College and a Master of Arts in military art and science from the School of Advanced Military Studies, U.S. Command and General Staff College.

Ms. Albert is a Research Associate in IDA's Information Technology and Systems Division. She holds a Master of Arts in Journalism from the University of Missouri.

Dr. McDaniel is an Adjunct Research Staff Member in IDA's Information Technology and Systems Division. She holds a doctorate in education: supervision, curriculum, and instruction from the University of Miami.

LOOKING BACK AT PORTOPT: AN ACQUISITION PORTFOLIO OPTIMIZATION TOOL

David Tate

The Problem

Can DoD afford to buy everything it is currently planning to buy? If not, what are the alternatives?

In 1998, the Acquisition Resources and Analysis (AR&A) directorate within the Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics (OUSD(AT&L)) had reviewed the recent long-term investment plans of Major Defense Acquisition Programs (MDAPs), and had added up the proposed spending by all of the programs year by year. The result showed both sharply rising investment costs within the five-year Future Years Defense Plan (FYDP) and continued growth beyond that horizon.[1]

AR&A asked IDA to address three questions:

1. Given plausible levels of defense funding, can we afford to execute all of the current programs?
2. If not, what are the cost implications of having to rearrange and stretch programs to fit under the budget?
3. How much money are we wasting by managing each program individually, rather than trying to coordinate acquisition across the entire portfolio of MDAPs?

Thus was born PortOpt, the Acquisition Portfolio Optimization project, which has provided analytical decision support to AR&A for the past fifteen years.

PortOpt was designed to answer this optimization question: "Given this much money to work with in each year, which set of simultaneous production schedules buys everything at lowest cost?"

DESIGNING A TOOL

IDA identified three critical challenges in answering the AR&A questions. First, we would need to be able to predict how the year-by-year procurement costs of a weapon system would change if its procurement schedule were changed. Second, we would need an optimization model that could find the lowest-cost combination of procurement schedules that could fit under a given top-line budget. Third, we would need to be able to refresh these models when new programs were started and as cost estimates and requirements changed within existing programs. For the tool to be useful, this refresh would need to happen in days or weeks.

[1] Defense analysts refer to rising planned costs in the years just beyond the FYDP as the "bow wave" of acquisition.

COST AS A FUNCTION OF SCHEDULE

The unit cost of a weapon system depends in part on how many you buy and how quickly you buy them. There are several different mechanisms at work, with many underlying complexities. In the next article, “Predicting the Effect of Schedule on Cost,” we talk about these complexities, both in terms of what they are and how we can figure them out from the available data. Here, we focus only on the three cost components that were included in the original PortOpt model: direct manufacturing costs, program management costs, and overhead costs.

MANUFACTURING

The most obvious costs of making a military system are the touch labor and materials costs. It takes a certain amount of metal to make a ship or helicopter, and it takes a certain number of labor hours to turn that metal into a hull or airframe. History has shown, however, that it doesn’t take the same number of hours for every ship or helicopter of a given design. Production of complex hardware exhibits learning—the direct labor hours for the second unit you make are less than for the first unit, and the third takes fewer hours than the second, and so forth. In general, the more systems of a given design you have already built, the lower the direct labor costs of the next one will be. Perhaps surprisingly, this is still true even in these days of highly automated manufacturing. Ongoing process improvements, more efficient uses of raw materials, better subcontractor arrangements—these all contribute to a general reduction of unit manufacturing cost as the cumulative quantity increases.

Learning has consequences that are not immediately obvious. One is that if you reduce the total number of units you plan to buy, the average unit cost will go up: the units you have canceled would have been the least expensive units built, having benefitted from all of the prior learning in making the earlier units. A second consequence is that the costs of production are not evenly distributed over the buy. If you were to make thirty helicopters per year for ten years, the annual costs would be much higher in the first year than in the last year—creating a tension between the desire for stable production rates and the desire for stable annual funding levels. It also makes it more complicated to predict exactly what will happen to funding requirements if you shift a few units of production from one year to another. PortOpt uses learning curves to capture these subtle nonlinear effects of changing production plans.

PROGRAM MANAGEMENT

Manufacturing costs are not the only procurement costs associated with acquiring weapon systems. Most programs also spend significant amounts on systems engineering, program management, quality assurance, training, documentation, support, and contract management. These and similar activities are direct costs of the program, but they are not associated with specific production units, and are essentially independent of whether the program is producing ten units or 1,000 units per year.

These costs are not subject to learning, and cannot be reduced in a given year by shifting production into the future. To be able to predict how annual and total costs will change as a result of a change in production schedule, one needs to understand what portion of the original planned cost is due to program management costs.

PLANT OVERHEAD

Finally, in addition to the direct costs of production (including both manufacturing and management costs), certain indirect costs must be paid. These are the overhead costs of the project; they pay for the salaries of employees who do not charge directly to individual programs, and also create the profit for the contractors. We tend to think of these as "fixed" costs, but they actually do vary somewhat as a function of how much business the contractor is doing.

For PortOpt, we collected historical direct and indirect cost data for the largest defense contractor facilities. Using those data, we estimated the total overhead costs per year for various prime contractors as a function of the total amount of business being done by the contractor in a given year. To estimate the overhead cost impact of a schedule change, we first estimated the direct costs in every year for every program (using the learning curves and program management costs), then allocated overhead costs to all of the programs at each plant, in proportion to the direct costs incurred by those programs in each year. If a contractor also had non-defense business at a plant, some of the plant's overhead was apportioned to that as well.

Putting the pieces together, PortOpt estimates the annual costs per program under a proposed production schedule by taking the following steps:

1. Estimating direct manufacturing costs in each year, taking into account how far down the learning curve the program will be at the beginning of each year and how many units are to be produced in that year;
2. Adding program management costs to each program that is still in production in each year;
3. Combining these to get total direct costs at each contractor facility in each year;
4. Using these total direct costs to estimate total overhead at each facility in each year;
5. Apportioning these overhead costs to individual programs at each facility in each year;
6. Adding direct costs to overhead costs for each program to get total program costs in each year.

PORTFOLIO OPTIMIZATION

PortOpt was designed to answer this optimization question: "Given this much money to work with in each year, which set of simultaneous production schedules buys everything at lowest cost?" Of course, not just any production schedule is practical. Because manufacturers have limited capacity, there is an upper limit on how many units they can produce per year. Conversely, there are economic and industrial base reasons never to let production drop below a certain "minimum sustaining rate" once it has begun. Finally, there are operational concerns—we need to buy all of the units of a given system while it is still operationally useful to do so.

This leads to an optimization formulation that can be summarized as:

Objective: buy all units of all systems at minimum total cost… subject to these constraints:

- *Stay within budget every year;*
- *Deliver all units on time;*
- *Don't produce too many per year; and*
- *Once you're in production, don't produce too few per year.*

That sounds pretty straightforward, but the result is a Mixed Integer Linear Program (MILP) with thousands of variables and tens of thousands of constraints. Even using the most powerful available commercial optimization software of the day, it took some clever formulation tricks for us to be able to optimize the entire MDAP portfolio in minutes or hours on a desktop computer.

DATA REFRESH

The goal of PortOpt was to be able to continue to update the model over time without having to conduct detailed assessments of plant overhead at defense contractors, which would be time consuming and costly. For that reason, we restricted the basic inputs to the data found in the annual Selected Acquisition Report (SAR) filed by each MDAP.

The SAR gives a description of past production quantities and costs, and a forecast of future quantities and costs by year. It is often not directly comparable to the previous year's SAR, due to changes in program requirements, cost growth, changes in planned total quantities, and so on. Every year, when the new SARs were published, we imported the new forecasts into PortOpt, estimated learning curves and annual program management costs from the new data, set delivery deadlines for each system, and set minimum and maximum production rates.

USES

Over the years, PortOpt has been used for different kinds of analysis. In this section, we discuss five of these.

ANALYZING AFFORDABILITY

The first of our original motivating questions was "Can we afford to execute all of the current programs?" This became an annual exercise for PortOpt—enter a plausible future top-line procurement budget, optimize all programs within that budget, and look to see how much money would be left to use on future programs that are not yet MDAPs. This gave the most optimistic picture of *affordability*, since it assumed no cost growth, no major development delays, and a stable future budget. From that baseline, we would then run sensitivity analyses and excursions reflecting different budget levels, different amounts of cost growth, "untouchable" programs, and changes in demand for various systems. Not uncommonly, the optimization would fail to find a feasible solution, indicating that not even the most efficient possible set of production schedules could buy all of those systems under the specified restrictions.

MEASURING THE COST OF THE "BOW WAVE"

The second of our original motivating questions was "What would it cost to knock down the bow wave?"—that is, how much more will systems cost, per unit, if we have to stretch out production in order to be able to afford them? For this analysis, we would fix the procurement schedules within the FYDP to be as planned in the SARs, but allow the optimization free rein to rearrange schedules in the outyears to make them fit under a projected budget. We would then use the PortOpt costing module to compare the projected costs of the SAR schedules and the optimized schedules for all programs. The difference in cost would be directly attributable to the budget constraint.

MEASURING THE COST OF STARTING TOO EARLY

Our third motivating question was "How much money are we wasting by managing programs one by one, rather than as a single optimized portfolio?" One special case of this question arises when deciding whether to start a new program in the current year, or to wait and start it at some future time. In general, it is more efficient to have fewer programs producing at higher rates at any given time. As those programs finish, new programs can be started. However, there are many incentives in the defense world to want to start programs as soon as funding is available, and to continue them for as long as possible.

Using PortOpt, we could compare the overall and unit cost difference between the "ideal" schedule, which finishes programs as quickly as possible by delaying the start of other programs, and the "typical" schedule, which funds all programs as soon as they are ready to start, but at lower production rates than in the optimal schedule. As above, the cost difference between the optimal schedule and the "typical" schedule gives a measure of how much could be saved through more efficient scheduling, in the absense of other constraints on acquisition plans.

REPROGRAMMING AFTER UNEXPECTED EVENTS

Occasionally, unexpected events cause a sudden change in the Services' expectations about budgets, program needs, or both. In those cases, PortOpt can be used to offer suggestions about the best way to reprogram everything in response to the disruption.

One example of this occurred when the Air Force, which had been planning to lease tanker aircraft using Operations and Support funds, was required by the Congress to purchase those aircraft instead. This led to a temporary (but large) shortfall in procurement funds available for other programs in certain years. Using PortOpt, AR&A was able to understand the magnitude of the problem, and to recognize that it would be impossible to remedy without significant changes to the then "untouchable" F-22 program.

A second example occurred when the Army identified an urgent need for mine-resistant ambush-protected (MRAP) vehicles in Iraq and Afghanistan. While some of the funds used to purchase MRAPs came from supplemental budgets, others had to be offset by reductions in Army procurement spending in other areas. PortOpt offered a way to estimate the opportunity cost of buying MRAPs, in addition to the monetary cost.

FINDING EFFICIENCIES

Finally, PortOpt can be used to find short-term efficiencies in how we procure military systems. In 2007, then-Under Secretary for Acquisition, Technology, and Logistics, John Young, challenged his staff to find $10 billion in procurement savings in the FY10 through FY15 budgets, without increasing acquisition budgets in any year or delaying full fielding of any system. Using PortOpt, AR&A identified the most effective set of current programs to accelerate and complete in the near term, in order to free up funds in the designated time window. While few of these programs were in fact accelerated, the PortOpt analysis framed the discussion for OUSD(AT&L) and contributed to initiatives that nearly met the $10 billion goal.

WHAT WE'VE LEARNED

In fifteen years of working with PortOpt, we have learned quite a bit about the opportunity costs of the way we do business, some of which confirm common sense (like the first two items below) while other insights are more nuanced:

- When every program is running near its minimum sustaining rate, there is no flexibility to cope with the unexpected.
- Finishing programs saves money. Everything else increases costs.
- Starting programs as early as possible often increases unit costs and delays fielding of systems—including the systems that were started as early as possible.
- Even if our current cost estimates are accurate, they don't account for the added costs of making everything fit under the budget we will actually have.
- Optimization is nice, but its real value is in showing the mechanisms that lead to savings. In real life, no single decision maker has the authority to optimize all procurement, even within a Service.

New tools using new methods are in development to take advantage of some of these lessons and to leverage new data sources and techniques that were not available in 1998. The timing is right for a change; as budgets turn down and the emphasis on affordability increases, PortOpt-like tools can help DoD acquire as many systems as possible within available budgets.

Dr. Tate is a Research Staff Member in IDA's Cost Analysis and Research Division. He holds a doctorate in operations research from Cornell University.

PREDICTING THE EFFECT OF SCHEDULE ON COST

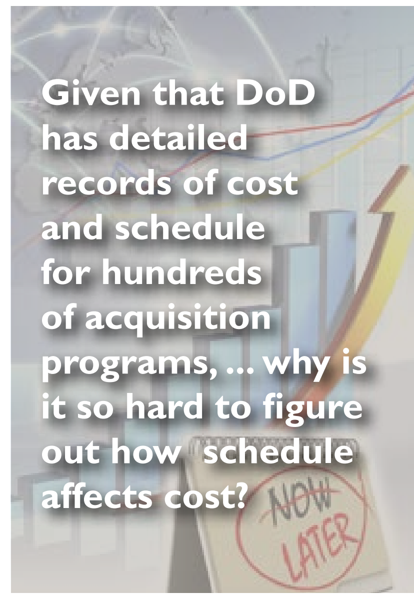
David Tate

The Problem

What is the effect of schedule on weapon system acquisition cost?

WHY DO WE NEED TO KNOW HOW SCHEDULE AFFECTS COST?

In our discussion of the PortOpt model ("Looking Back at PortOpt"), we noted that a key requirement of the PortOpt tool was to be able to predict how the lot costs of a procurement program would change if we were to change the production schedule. This is the fundamental step; if we can't estimate the cost impacts of specific schedule changes, then we can't (1) compare alternatives to the current Selected Acquisition Report (SAR) plans, (2) do sensitivity analysis, and (3) optimize.



WHY IS THIS HARD?

DoD has detailed records of cost and schedule for hundreds of acquisition programs, going back decades. Given that wealth of data, why is it so hard to figure out how schedule affects cost? We believe there are three major analytical obstacles.

OBSTACLE #1: WHICH IS THE CAUSE, AND WHICH IS THE EFFECT?

Consider three procurement programs: A, B, and C. Program A is doing fine, but due to overall budget reductions, A's production schedule is going to be stretched, which will increase unit costs. Program B has just announced significant cost growth—not caused by a schedule change—that has made its planned production schedule no longer affordable. Program C has been experiencing integration issues—its electronics are going to require a new design that uses a more expensive subcomponent, which will have to be retrofitted into existing units. This means both a cost increase (due to the new component) and a schedule slip (to accommodate the new design and the rework).

For all of these programs, unit cost went up and average production rate went down. Causally, though, we have three distinct cases:

- For Program A, schedule stretch caused cost growth.
- For Program B, cost growth caused schedule stretch.
- For Program C, technical issues caused both cost and schedule growth.

We're trying to understand only the first of those mechanisms. Unfortunately, we can't tell just by looking at historical numbers which of those cases was in effect—or whether it was some mix of all of them—for a given program. We need a way to isolate the Program A effect from the others.

OBSTACLE #2: NOT ALL COSTS REACT THE SAME WAY TO SCHEDULE CHANGES

Since 2006, SARs have broken out cost projections into subcategories: end-item recurring flyaway costs, non-end-item recurring flyaway costs, nonrecurring costs, and two categories of support costs. This is very helpful, because we don't expect all of those costs to react identically to a change in production schedule. End-item recurring costs should be most directly affected, while nonrecurring costs and non-spares support might not be affected at all. Our econometric model of how schedule affects cost will have to identify these different cost categories and treat them separately

OBSTACLE #3: LIMITED RELEVANT DATA

If I wanted to understand the relationship between the price of butter today and the price of eggs tomorrow, I could collect a great deal of historical data on butter and egg prices at various times. That would work because the natures of butter and eggs don't change much over time; all of my historical data would describe the same commodities

Acquisition programs, however, are not like that. We can look at the historical SARs for Program A, going back for as many years as Program A has been around—but do those past SARs tell us anything about the relationship today between Program A's production schedule and Program A's lot costs? Since those past SARs were published, Program A may have changed in any number of ways—new designs, revised cost estimates, requirements changes, technology insertions, planned product improvements, new contracts, new demands from the field, for instance.

Unless we could somehow correct for all of the program changes other than schedule, those past forecasts don't tell us what the estimated cost would be today for that prior planned schedule. We generally don't get to see multiple schedules (and their associated costs) for the same exact program. Since we're trying to figure out how cost varies as schedule varies, this is a major limitation.

WHAT WE'VE LEARNED

We have explored a number of competing theories about how and why unit costs change when schedules change. These theories are not necessarily mutually exclusive, which makes it even trickier to figure out how to combine them into a coherent model. Here's what we've learned so far.

FIXED COSTS AND STICKY COSTS

Some of the costs of producing a weapon system are incurred per unit time, rather than per unit produced. For example, the costs associated with running the program office do not depend much on the current production rate, or on how many units have been produced so far. Similarly, the indirect costs associated with contractor overhead are only slightly sensitive to production rates. What's more, overhead rates are "sticky"—they don't generally adjust instantaneously to changes in work level. Our model of how cost depends on schedule will need to be able to distinguish fixed costs from variable costs, and estimate how sticky the fixed costs are.

LEARNING AND FORGETTING

The previous article "Looking Back at PortOpt" also introduced the idea of learning curves. It is not uncommon to see that unit costs seem to follow a standard learning curve for most of the life of a program, but then start to climb upward again toward the end of the program. To account for this, C. Lanier Benkard suggested that producers become more efficient by gaining "experience" making units, but that this experience dissipates at a constant rate.[1] Thus, early in production (when cumulative quantity is doubling frequently), or at high production rates (when more experience is being gained per unit time), learning behavior dominates. Late in the production run, or at low production rates, the gains from learning are visibly offset by forgetting.

We investigated this model, and found that it fits many historical programs quite well. It can also be improved by combining it with a fixed cost model, so that overhead is modeled separately, while direct costs are modeled by a combination of learning and forgetting.

REGULATORY LAG

Finally, William Rogerson has proposed that the interaction between cost progress and production rate can be understood by looking at the incentives inherent in how procurement contracts are awarded.[2] In general, a new fixed-price procurement contract is awarded for each lot, with a price based on the contractor's demonstrated historical costs. If a contractor invests in management or tooling changes that reduce production costs, they will only realize extra profits from this until a new price is negotiated—typically two or three years later.

At high production rates, contractors have more incentive to reduce production costs, because they will realize extra profits on many units during the two- to three-year "regulatory lag" period before the price is renegotiated downward to reflect the lower production costs. There is also a wider range of worthwhile cost-reducing investments available, given the need to make back the initial investment costs through higher profits.

Conversely, at lower production rates the contractor has less incentive to reduce costs, as well as fewer available cost-reduction alternatives that will provide the necessary return on investment. If this theory is correct, we should expect to see less learning

[1] C. Lanier Benkard, "Learning and Forgetting: the Dynamics of Aircraft Production," *American Economic Review 90*, no. 4 (Sep 2000): 1034–1054.

[2] William P. Rogerson, "Economic Incentives and the Defense Procurement Process," *Journal of Economic Perspectives* 8, no. 4 (Fall 1994): 65–90.

at low production rates, and more learning at higher production rates. This is very different from traditional procurement models, which all assume that the learning curve slope is an intrinsic characteristic of the system being produced.

We can combine this model with a fixed-cost model, as we did with the learning-and-forgetting model. In theory, we could add forgetting to this model as well, but we will generally not have enough data to distinguish between regulatory lag effects and forgetting effects if both are present.

GOING FORWARD

IDA continues to develop and refine models of how schedule changes affect cost. Our investigations suggest that the most useful models for practical applications will combine a fixed-cost model with either a learning-and-forgetting model or a regulatory lag model, depending on the available data and type of system being procured. The model parameters will be fit at the system-type level (e.g., a single value for all helicopters, or all tactical missiles) where possible, and at the program level where necessary.

Estimating cost changes from schedule changes is complicated, and difficult—but necessary. The new models being developed will enable important new capabilities for affordability analysis, portfolio planning, analyzing proposed multi-year procurement contracts, and a variety of other activities important for DoD acquisition planning.

Dr. Tate is a Research Staff Member in IDA's Cost Analysis and Research Division. He holds a doctorate in operations research from Cornell University.

References

Benkard, C. Lanier. "Learning and Forgetting: the Dynamics of Aircraft Production." *American Economic Review* 90, no. 4 (Sep 2000): 1034–1054.

Rogerson, William P. "Economic Incentives and the Defense Procurement Process." *Journal of Economic Perspectives* 8, no. 4 (Fall 1994): 65–90.

RECENT DEVELOPMENTS IN THE JOINT STRIKE FIGHTER DURABILITY TESTING

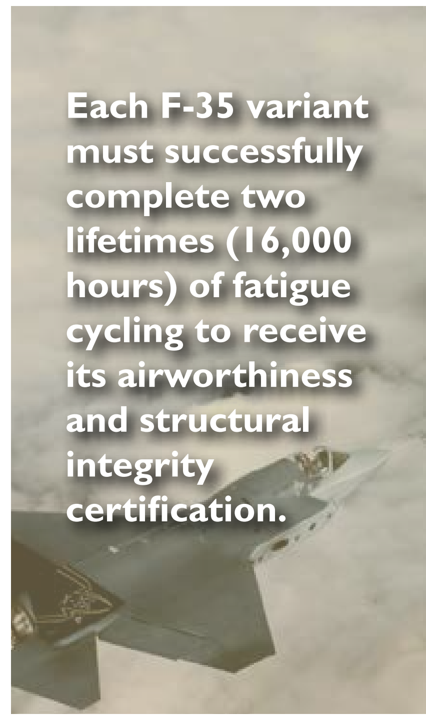
Lisa Veitch and Evan Laprade

The Problem

The Joint Strike Fighter (JSF) program—the largest defense acquisition program in history—aims to maintain U.S. air superiority by providing the next generation of Air Force, Navy, and Marine Corps fighter aircraft. It is also, arguably, the most complex defense acquisition program ever undertaken. IDA's role is to assess progress in developmental testing for USD(AT&L), exploring potential issues and testing options for consideration by the program office and by other stakeholders in the acquisition process.

F-35 DURABILITY TESTING PROGRAM OVERVIEW

The Joint Strike Fighter (JSF)/F-35 Lightning program, the Defense Department's most expensive acquisition program, is currently more than halfway through developmental test and evaluation (DT&E). Three variants of the aircraft, shown in Figure 1, are to be delivered to the Air Force, Marines, and Navy within the next five years. Each has unique capabilities. The Air Force variant is the conventional take-off and landing (CTOL) aircraft; the Marine variant is a short take-off/vertical landing (STOVL) aircraft, which can operate in austere locations; and the Navy's carrier variant (CV) is equipped for landing on carrier decks. Developmental testing (DT) of each variant requires not only flying the aircraft to analyze its flight characteristics and mission systems, but also involves full-scale ground testing to determine each design's structural integrity.

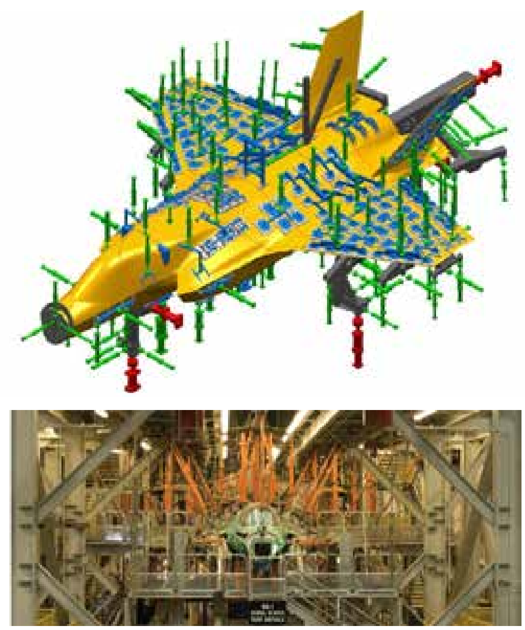


Source: Images courtesy of Lockheed Martin.

Figure 1. JSF Variants: (left to right) F-35A (CTOL), F-35B (STOVL), F-35C (CV)

Each F-35 variant must successfully complete two lifetimes (16,000 hours) of fatigue cycling to receive its airworthiness and structural integrity certification. To demonstrate this capability, a complete airframe is subjected to a spectrum of maneuver, buffet, and catapult/arrestment (carrier variant only) loads, using a specially designed rig of pneumatic jacks

(Figure 2). Each test article is heavily sensored to monitor strain and stress loadings of critical areas. Through this full-scale durability testing, the F-35 program hopes to identify points in the structure that do not meet durability requirements, incorporate fixes for these issues (when possible), and, most importantly, demonstrate that the airframe design will maintain its structural integrity as required.

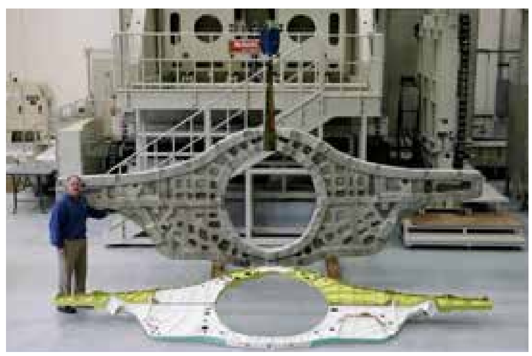


Source: Images courtesy of Lockheed Martin.

Figure 2. (Above) Schematic Showing Location of Approximately 150 Pneumatic Loading Points; (Below) STOVL Test Article Sitting in Its Test Fixture

FIXES TO ACHIEVE "FULL-LIFE" – REDESIGNS, STRAPS, AND LASERS

Since 2010, IDA has been closely monitoring the progress of the F-35 durability testing program in support of Deputy Assistant Secretary of Defense (Developmental Test and Evaluation) (DASD DT&E) assessments. During this time, an unexpected number of structural and materials issues have arisen. In particular, testing the STOVL full-scale article, referred to as BH-1, has been, and continues to be, particularly eventful. BH-1 began testing in July 2010 and shortly thereafter inspections discovered a succession of significant cracks at approximately 1,500, 7,000, and 9,000 hours of testing. These cracks occurred in major load-bearing bulkheads (like that shown in Figure 3), which hold the wings onto the aircraft and are essential to the structural integrity of the airframe.



Source: Images courtesy of Lockheed Martin.

Figure 3. STOVL Bulkhead 496, Forged from Aluminum 7085

As a result of these design shortfalls, BH-1 has required extensive repairs and experienced significant downtime, setting testing progress behind by more than a year and a half. Furthermore, due to the production and testing concurrency of the F-35 program, these failures now must be addressed as retrofit fixes for earlier low rate initial production (LRIP) aircraft, and, when possible, as bulkhead redesigns for future production aircraft, creating myriad configurations that must be certified.

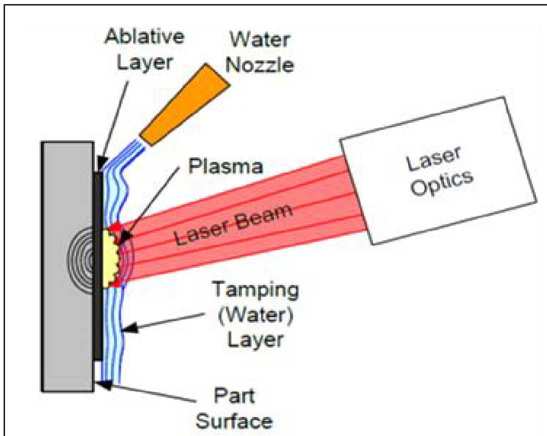
For instance, for pre-LRIP 5 aircraft, the fatigue life shortfalls identified at 1,500 hours of testing require the addition of heavy steel straps (mechanically attached reinforcement braces). For post- LRIP 5 aircraft, a bulkhead redesign, with thicker features, is being used for production. Similar modifications—straps for early production and redesigns for future production aircraft—will help address the shortfalls identified at 7,000 hours. The fatigue life shortfalls identified at 9,000 hours, however, pose a greater challenge because the necessary bulkhead redesign (thickening of features) cannot be accommodated without interfering with neighboring components. As a result, Lockheed Martin is pursuing a special method to strengthen these bulkhead areas—a unique surface treatment process referred to as Laser Shock Peening (LSP).

LASER SHOCK PEENING (LSP)

LSP is a *mechanical* process that improves a material's fatigue life by introducing a compressive residual stress at the surface. A material's surface, where there is often a combination of high strain (or stress), stress concentrations (sharp design features, machining marks), and corrosive attack, is particularly susceptible to fatigue crack initiation. A compressive surface stress layer counteracts the tensile stress environment necessary for crack nucleation and growth, thereby improving a material's crack resistance and fatigue life.

LSP uses high-energy laser pulses to create a shock wave that mechanically deforms the surface of a material (it is not used to create thermal effects). The process, shown in Figure 4, involves first coating the part surface with a sacrificial ablative layer (typically paint or tape). Water is then flowed over the part surface and a high-energy laser (1-10 GW/cm2)[1] is directed at the target region. A laser pulse vaporizes the ablative layer, creating a plasma cloud that is confined by the water layer. The rapidly expanding plasma generates a pressure shock wave (1-10 GPa)[2] that plastically compresses the metal, producing a residual stress field with a highly controllable depth and magnitude.

While LSP generates compressive stresses similar in magnitude to that of traditional shot-peening methods, the depth of this residual stress field extends far deeper below the surface (up to 2 mm for LSP versus only 0.5 mm for shot peening).[3]



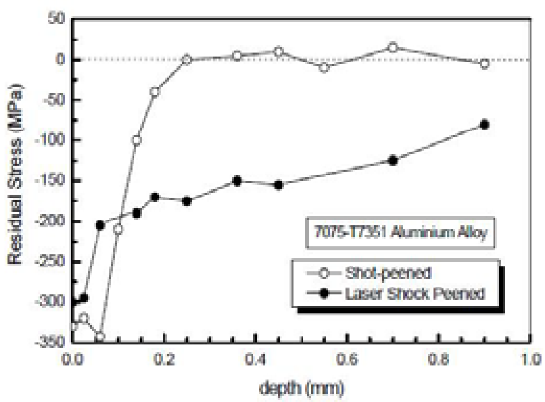
Source: Graphic reproduced from LSP Conference Presentation, April 2010, F-22 Air Vehicle Technology.

Figure 4. Laser Shock Peening Schematic

[1] Y. B. Guo, "Laser Shock Peening: Modeling, Simulations, and Applications," in *Numerical Simulations – Applications, Examples and Theory*, ed. Lutz Angermann (InTech, January 2011), 331–54.

[2] Y. Zhao, "Effects of Laser Shock Peening on Residual Stress, Texture and Deformation of Microstructure of Ti-6Al-4V Alloy" (Ph.D. diss., University of Cincinnati, 2012).

Figure 5 provides a comparison between the residual stress produced in an aluminum alloy using traditional shot peening and laser shock peening, with the latter exhibiting a significantly deeper compressive stress field.



Source: Q. Liu, "An Effective Life Extension Technology for 7xxx Series Aluminum Alloys by Laser Shock Peening," DSTO-TR-2177 (Melbourne, Australia: Defence Science and Technology Organisation, 2008), http://dspace.dsto.defence.gov.au/dspace / handle/1947/9655.

Figure 5. Comparison of Residual Stress Fields Induced by LSP and Shot Peening in Aluminum Alloy 7075-T7351

As a result, LSP provides resistance to crack growth and initiation deeper below the surface, further improving fatigue life. Additional advantages of LSP (versus shot peening) include process repeatability, less surface damage (cratering), and greater flexibility to reach hard-to-access areas.

Laser shock peening was initially demonstrated in the 1960s, but technology maturation delayed industry adoption until the mid-1990s. During the decade that followed, LSP found particular success extending the lifetime of engine fan blades, including those on the B-1B Lancer, F-16 Falcon, and Boeing's 777/787.[4] More recently, LSP has been applied to airframe structures as shown in Figure 6: F-22 Raptor titanium (Ti) wing lug,[5] T-45 steel tail hook shank,[6] and Apache/Chinook steel rotor gears.[7] LSP processing of aluminum aircraft structures has been limited to 747-8 wing panel skins[8] and T-38 aircraft main landing gear aluminum side-brace trunions (7049-T73).[9]

The F-35B bulkheads that require LSP processing are made from a relatively new 7000 series aluminum, aluminum (Al) 7085. Although a number of laboratory studies on 7000 series

[3] Ibid.

[4] Q. Liu, "An Effective Life Extension Technology for 7xxx Series Aluminum Alloys by Laser Shock Peening," DSTO-TR-2177 (Melbourne, Australia: Defence Science and Technology Organisation, 2008), http://dspace.dsto.defence.gov.au/dspace /handle/1947/9655; L. Hackel, "Corrective laser peen forming of F-18 fuselage 701 skins," CTMA Symposium, 2012.

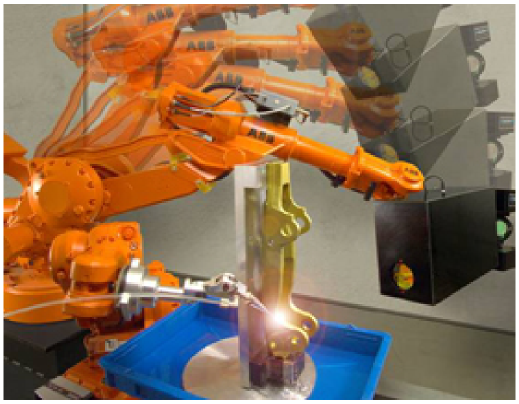
[5] D. Jensen, "Adaptation of LSP Capability for Use on F-22 Raptor Primary Structure at an Aircraft Modification Depot," 2nd International Laser Peening Conference, San Francisco, CA, April 2010; L. Polin et al., "Full Scale Component Tests to Validate the Effects of Laser Shock Peening," F-22 Program Brief, Public Release 11/18/2011, 2011.

[6] J. Rankin et al., "Effect of Laser Peening on Fatigue Life in an Arrestment Hook Shank Application for Naval Aircraft," 2nd International Laser Peening Conference, San Francisco, CA, April 19, 2010.

[7] "Laser Peening for Army Vehicle Life Extension," SBIR Award ID 62903, 2008, http://sbir.gov/ sbirsearch/detail/218723.

[8] C. Collisson, "Re-inventing the Legend: The Development of the 747-8," ICAS 2008, 2008.

[9] Liu, "An Effective Life Extension Technology for 7xxx Series Aluminum Alloys by Laser Shock Peening."



Source: Image courtesy of Metal Improvement Company.

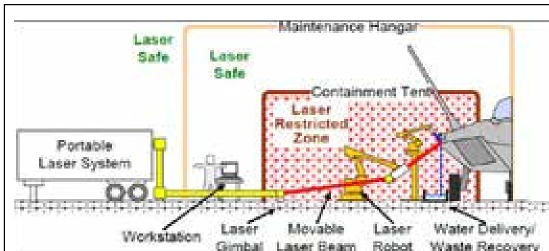
Figure 6. Time Lasped Images of an LSP Operation – LSP Robot (orange) Controls Laser (black box) Positioning and Water Injection (silver nozzle) as a Metal Part Is Processed

aluminum have demonstrated that LSP provides a significant fatigue life improvement (~3x or more) relative to the as-machined condition,[10] only a single published research effort[11] has looked specifically at LSP processing of Al 7085.

LSP DEVELOPMENT PROGRAM

An LSP development program is currently under way by Lockheed Martin to optimize process parameters for Al 7085 and to verify that LSP will provide the necessary fatigue life enhancement. The F-22 wing-lug development effort provides a useful frame of reference for assessing the road ahead for LSP qualification for F-35 bulkheads. The F-22 wing-lug LSP qualification process was completed over a four-year period, starting with coupon level testing to optimize LSP parameters—laser intensity, pulse size, duration, and number of layers (LSP parameters are highly dependent on material and geometry).

The LSP setup used for the F-22 consists of a tractor-trailer-mounted laser system and a robot that redirects the laser to the target area (see Figure 7). F-22 wing lugs are relatively easy to access after removing the aircraft's wings, providing adequate space for robot positioning and water jet placement. Access to the F-35 bulkheads will be more restricted and will likely require even greater engineering efforts in terms of aircraft disassembly, laser redirection, and water flow positioning. In addition, material and part design differentiates the F-35 bulkhead LSP development effort from that of the F-22 wing lug. The bulkhead features that have been targeted for LSP treatment are primarily webs and flanges. These features are considerably thinner and geometrically more complex than those processed on the wing lug, making residual stress control more challenging.



Source: Jensen, "Adaptation of LSP Capability for Use on F-22 Raptor Primary Structure at an Aircraft Modification Depot."

Figure 7. LSP Setup Used for the F-22

[10] Ibid.; Montross et al., "Laser shock processing and its effects on microstructure and properties of metal alloys: a review," International Journal of Fatigue 24, No.10 (2002), 1021–1036.

[11] H. Luong and M. Hill, "The effects of laser peening on high-cycle fatigue in 7085-T7651 aluminum alloy," Materials Science and Engineering: A 477, No.1-2 (2008), 208–216.

PATH TO STOVL AIRCRAFT CERTIFICATION

The airworthiness and longevity of the F-35B is critically dependent on the successful qualification of the LSP process. Currently the LSP development effort is on schedule for completion in November 2017. Assuming this timeline holds, the first production line cut-in of LSP would start with LRIP 11 (all STOVL aircraft LRIP 11 and beyond would receive LSP during production). Pre-LRIP 11 STOVL aircraft (111 in total), however, will undergo LSP processing as part of a depot modification. Given LSP's significance to the entire F-35B fleet, IDA continues to monitor the progress of the LSP development program and provide DASD DT&E with technical insights on this important piece of the durability testing program.

Dr. Veitch is a Research Staff Member in IDA's System Evaluation Division. She holds a doctorate in materials engineering from Purdue University.

Dr. Laprade is a Research Staff Member in IDA's System Evaluation Division. He holds a doctorate in materials science and engineering from Northwestern University.

References

Collisson, C. "Re-inventing the Legend: The Development of the 747-8." ICAS 2008. 2008.

Guo, Y. B. "Laser Shock Peening: Modeling, Simulations, and Applications," In *Numerical Simulations – Applications, Examples and Theory*, edited by Lutz Angermann, 331–54. *InTech*, January 2011. doi: 10.5772/13094.

Hackel, L. "Corrective laser peen forming of F-18 fuselage 701 skins." CTMA Symposium, 2012.

Jensen, D. "Adaptation of LSP Capability for Use on F-22 Raptor Primary Structure at an Aircraft Modification Depot." 2nd International Laser Peening Conference, April 2010.

"Laser Peening for Army Vehicle Life Extension." SBIR Award ID 62903. 2008. http://sbir.gov/sbirsearch/detail/218723.

Liu, Q. "An Effective Life Extension Technology for 7xxx Series Aluminum Alloys by Laser Shock Peening." DSTO-TR-2177. Melbourne, Australia: Defence Science and Technology Organisation, 2008. http://dspace.dsto.defence.gov.au/dspace /handle/1947/9655.

Luong, H., and M. Hill. "The effects of laser peening on high-cycle fatigue in 7085-T7651 aluminum alloy," *Materials Science and Engineering*: A 477, No.1-2 (2008): 208–16.

Montross, C., T. Wei, L. Ye, G. Clark, and Y. W. Mai. "Laser shock processing and its effects on microstructure and properties of metal alloys: a review," International Journal of Fatigue 24, No.10 (2002): 1021–36.

Polin, L. et al. "Full Scale Component Tests to Validate the Effects of Laser Shock Peening." F-22 Program Brief. Public Release 11/18/2011. 2011.

Rankin, J., L. Hackel, J. Harrison, J. Nemeth, and M. Leap. "Effect of Laser Peening on Fatigue Life in an Arrestment Hook Shank Application for Naval Aircraft," 2nd International Laser Peening Conference, San Francisco, CA, April 19, 2010.

Zhao, Y. "Effects of Laser Shock Peening on Residual Stress, Texture and Deformation of Microstructure of Ti-6Al-4V Alloy." Ph.D. diss, University of Cincinnati, 2012.

Past Issues

DoD Test and Evaluation: Statistical Methods for Better System Assessments

- Assessing Submarine Sonar Performance Using Statistically Designed Tests
- Applying Advanced Statistical Analyses to Helicopter Missile Targeting Systems
- Tackling Complex Problems: IDA's Analyses of the AN/TPQ-53 Counterfire Radar
- Improving Reliability Estimates with Bayesian Hierarchical Models
- Managing Risks: Statistically Principled Approaches to Combat Helmet Testing
- Validating the Probability of Raid Annihilation Test Bed Using a Statistical Approach

Technological Innovation for National Security

- Acquisition in a Global Technology Environment
- Lessons on Defense R&D Management
- Commercial Industry R&D Best Practices
- Strengthening Department of Defense Laboratories
- Policies of Federal Security Laboratories
- The Civilian Science and Engineering Workforce in Defense Laboratories
- Technology Transfer: DoD Practices

Acquisition

- Defining Acquisition Trade Space Through "DERIVE"
- Supporting Acquisition Decisions in Air Mobility
- Assessing Reliability with Limited Flight Testing
- Promise and Limitations of Software Defined Radios
- Implications of Contractor Working Capital on Contract Pricing and Financing
- The Mechanisms and Value of Competition
- Early Management of Acquisition Programs

Security in Africa

- Trends in Africa Provide Reasons for Optimism
- China's Soft Power Strategy in Africa
- Sudan on a Precipice
- A New Threat: Radicalized Somali-American Youth
- Chinese Arms Sales to Africa
- Potential of Engagement Networks in Africa
- Environmental Cooperation with South Africa

Challenges in Cyberspace

- Cyberspace – The Fifth and Dominant Operational Domain
- Transitioning to Secure Web-Based Standards
- Information Assurance Assessments for Fielded Systems During Combatant Command Exercises
- Supplier-Supply Chain Risk Management
- Internet-Derived Targeting: Trends and Technology Forecasting
- Training the DoD Cybersecurity Workforce

Today's Security Challenges

- A Framework for Irregular Warfare Capabilities
- Bridging the Interagency Gap for Stability Operations
- Developing an Adaptability Training Strategy
- Force Sizing for Stability Operations
- Detecting Improvised Explosive Devices
- Building Partner Capacity
- Combating the Trans-South Atlantic Drug Trade
- Countering Transnational Criminal Insurgents
- Understanding the Conflict in Sudan
- Planning Forces for Foreign Internal Defense and Counterinsurgency
- Test and Evaluation for Rapid Fielding Programs
- Using Economic and Financial Leverage

Resource Analyses

- Evaluating the Costs and Benefits of Competition for Joint Strike Fighter Engines
- Analysis and Forecasts of TRICARE Costs
- Effects of Reserve Mobilization on Employers
- Does DoD Profit Policy Motivate Contractors?
- Auctions in Military Compensation

Focusing on the Asia-Pacific Region

- Making Security Partners Better Resource Managers
- Collaborating with Singapore
- Inside North Korea
- Extending Trilateral Cooperation for Disasters
- Developing Human Capital in China

Homeland Security

- Port Vulnerability
- Assessing the EMP Threat
- Homeland Defense Scenarios
- Transport and Dispersion Models
- IT Security

IDA | RESEARCH NOTES

4850 Mark Center Drive • Alexandria, VA 22311-1882
www.ida.org