PERFORMANCE OF THE DEFENSE ACQUISITION SYSTEM

2015 ANNUAL REPORT

SEPTEMBER 16, 2015

Approved for public release; distribution unlimited.

Report Documentation Page						*Form Approved* *OMB No. 0704-0188*
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302 Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number						
1. REPORT DATE **16 SEP 2015**		2. REPORT TYPE				3. DATES COVERED **00-00-2015 to 00-00-2015**
4. TITLE AND SUBTITLE **Performance Of The Defense Acquisition System**						5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)						5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Under Secretary of Defense, Acquisition, Technology, ,and Logistics (USD[AT&L]),3000 Defense Pentagon,,Washington,,DC, 20301**						8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)						10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**						
13. SUPPLEMENTARY NOTES						
14. ABSTRACT **This third annual report on the Performance of the Defense Acquisition System continues the effort to provide data and analysis so the Department of Defense (DoD) can see how it is doing, measure the effectiveness of ongoing efforts to improve acquisition, and learn from past experience. This annual report series includes regular updates of past analysis with more current data so we can monitor performance over time, detect trends, assess differences among our institutions, and determine relationships between policy and results. It also includes analysis on new topics to provide deeper insights and help address specific areas of concern.**						
15. SUBJECT TERMS						
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON	
a REPORT **unclassified**	b ABSTRACT **unclassified**	c THIS PAGE **unclassified**	**Same as Report (SAR)**	**211**		

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Performance of the Defense Acquisition System, 2015 Annual Report. Washington, DC: Under Secretary of Defense, Acquisition, Technology, and Logistics (USD[AT&L]), September, 2015.

3000 Defense Pentagon
Washington, DC 20301-3000
www.acq.osd.mil

FOREWORD

The plural of anecdote is not data.

—author unknown

This third annual report on the *Performance of the Defense Acquisition System* continues the effort to provide data and analysis so the Department of Defense (DoD) can see how it is doing, measure the effectiveness of ongoing efforts to improve acquisition, and learn from past experience. This annual report series includes regular updates of past analysis with more current data so we can monitor performance over time, detect trends, assess differences among our institutions, and determine relationships between policy and results. It also includes analysis on new topics to provide deeper insights and help address specific areas of concern.

Incentives are motivating better performance, but we must use them appropriately and carefully to avoid unintended consequences. The linkage between prime contract profit margin and performance is being strengthened. New data confirm that first-tier subcontract margins are generally higher than those on our prime contracts. This factual observation needs further analysis from a policy perspective. Subcontract margins should be reasonable and also tied to performance. The data suggest that we have more work to do in this area.

Cost growth on our major programs is generally at or better than historical levels, but outliers remain a problem. Some specific types of cost growth are getting better. Median biennial change in total needed program funding has been near zero since 2009 (although past growth over baselines remains). Contractors on Major Defense Acquisition Program (MDAP) contracts are doing a better job of meeting cost targets. The number of MDAP contracts started since 2009 with price reductions has increased significantly compared to earlier contracts. These results may be evidence of early success from Better Buying Power (BBP) initiatives. The BBP "should cost" initiative, which has been in place for almost 5 years, requires our managers to actively seek ways to save money and to set targets for doing so, not just to stay within their budgets. This is a major cultural change that seems to be taking hold. Another BBP initiative, supported by the data in last year's report, is the increasing the use of formulaic incentive-type contracts (both cost-plus-incentive-fee and fixed-price-incentive) that tie contractor financial results explicitly to cost performance and indirectly (because time is money) to schedule performance.

In this report, we examine some measures of government costs of doing business. Our early analysis indicates that these costs are not excessive compared to overhead costs in the defense industry itself. We have found that it is difficult to compare these costs to commercial industry measures since government regulations impose functions beyond those employed in the commercial sector. Also, it is challenging to obtain data that separate productive direct labor from overhead. Still, our data show that the percentages of government execution costs (direct and overhead) are at or below defense industry overhead levels, and they are not unreasonable in absolute percentages. While we must continually seek efficiencies whenever possible, care must be taken not to cut program management and oversight so much that fundamental acquisition functions of the government (such as requirements tradeoffs, cost estimating, market research, negotiation, contracting, systems engineering, testing, auditing, and others) are crippled. Cutting these capabilities too much is a false efficiency and a mistake we have made in the past.

Beyond these operational acquisition questions, there is evidence that we have been pursuing less complex systems with about the same or less risk since 2009. This aligns with my concern that in some areas we may not be pushing the state-of-the-art enough in terms of technical performance. This endangers our military technical superiority. In my view, our new product pipeline is not as robust as it should be at a time when our technological superiority is being seriously challenged by potential adversaries. Not all cost growth is bad; we need to respond to changing and emerging threats. It is hard with existing records to distinguish cost growth that is due to overly optimistic planning or poor execution from cost growth that is due to needed design changes that address evolving threats, technological opportunities, or other prudent factors. The broader challenge of adapting to threats while fielding warfighting capabilities is intertwined and concurrent with defense acquisition. Simply delivering what was initially required on cost and schedule can lead to failure in achieving our evolving national security mission—the reason defense acquisition exists in the first place.

This report also summarizes some important, recently completed work performed by the Institute for Defense Analyses. That work shows a very strong correlation between high acquisition cost growth for programs and tight budgetary environments (like the one at present) during program baselining. While we need to do more work to understand the underlying causes of this result, it should give us all pause that acquisition unit cost growth on such programs is consistently 3-times higher (about 30 percent versus 10 percent) than that for programs started when budgets are not as constrained.

This report series is a central part of BBP. It continues to reflect results in defense acquisition performance from ongoing DoD compliance with the Improve Acquisition Act of 2010 and the earlier Weapon Systems Acquisition Reform Act. Although similarly motivated, our efforts go beyond the specifics of those laws to seek additional insights for improving the defense acquisition system's performance. This study also fulfills ongoing Office of Management and Budget requests for evidence-based analytic studies on acquisition performance. It is encouraging to see evidence of performance improvement over the last few years. However, these results are not a reason to pause in our efforts. They should motivate us to press ahead even more vigorously.

Finally, with all the current focus on acquisition reform, I encourage the stakeholders of defense acquisition to examine this report, prior reports, and other data-driven analyses to help guide ongoing discussions and policymaking. While it is important to continue improving our policies and practices, change for change's sake isn't the answer. We should use experience supported by data-driven analysis to help ensure we don't embrace policy reforms that carry unintended adverse consequences.

The Honorable Frank Kendall
Under Secretary of Defense for
Acquisition, Technology, and Logistics

(This page was left blank intentionally)

ORGANIZATION OF THE REPORT

Highlights. The following section briefly discusses key findings from the report, providing page references to detailed discussions in the main body. It also provides some insights to illustrate how this kind of analysis is informing actions within the DoD. Though not comprehensive, it provides perspectives and insights gleaned from the entire report.

Chapter 1 provides background material on acquisition, spending levels and trends, and general perspectives on measuring institutional performance to set the stage for the analysis presented in subsequent chapters.

Chapter 2 provides analysis on outputs and outcomes of the defense acquisition system. It analyzes the performance outcomes of our acquisition institutions from a variety of perspectives: DoD-wide, by commodity type, contract- and program-level, military department, and contractors. It includes a new overview of development contracts that are statistical outliers in total cost growth (details are in Appendix A). This chapter builds on the results from prior annual reports, updating some analysis and providing new results using different datasets. To a large extent, this chapter presents an ongoing view of performance and trends.

Chapter 3 discusses new analysis on broader factors that influence acquisition outcomes. We start with how broad acquisition reforms and the funding climate appear to affect program cost growth. We follow with progress report on implementing affordability analysis and constraints. We then present new synthetic variables that appear to behave as measurements of complexity and risk in contracts, followed by measurements of them on MDAP contracts and how they correlate with contract price. Finally, the chapter closes with analysis of margins on subcontracts.

Chapter 4 provides selected input and process measures, including the acquisition workforce improvements and incentives, bid-protest rates, usage of Lowest-Price, Technically Acceptable (LPTA) for selecting contracted services, performance relative to competition and small-business goals, and preliminary results on government execution costs.

Appendix A provides detailed attributes of 27 development contracts that are cost-growth outliers.

Appendix B provides details on the statistical methods employed in the reported analyses.

(This page was left blank intentionally)

Contents

(This page was left blank intentionally)

HIGHLIGHTS OF THE 2015 REPORT ON THE PERFORMANCE OF THE DEFENSE ACQUISITION SYSTEM

A key part of improving a system is objectively measuring its performance and the effects of policies, processes, and inputs on the outcomes of the system. Without this, we cannot tell where we have problems, what is working (or not), and whether management changes are making things better (or worse). In the case of defense acquisition, the primary outcome is the value of operational capabilities delivered in time for our warfighters to address threats. Unfortunately, objectively measuring the final operational performance and value of our systems across systems and commodities is very difficult. What we can objectively measure and thus focus on in our reports is the cost, schedule, and technical performance of our acquisitions—aggregated to look for statistically significant trends together with correlates, institutional differences, and theory to inform ways to improve future outcomes. Each performance measure has its strengths and weaknesses, so we use multiple measures (e.g., at both the program and contract level) and subsequent analysis to see if the answers point in the same direction while adding experience and theoretical insights to guide our conclusions.

This is the third annual report on the performance of the defense acquisition system, using quantitative analysis of broad data to measure institutional performance. The first report (AT&L, 2013b) analyzed recent and historical data to establish performance references and begin looking for evidence of what factors affect cost, schedule, and technical performance. For example, we found that undefinitized contract actions (UCAs) usually can be employed in early procurement without incurring cost growth, but UCAs correlate historically with cost and schedule growth on development contracts. That first report also established measures of cost growth at both the program and contract level that avoid confounding issues such as quantity effects, reporting these measures across our major programs and their contracts. We realized that since our major programs last many years and that new policies affect programs incrementally, time will be needed see the results of our improvement initiatives. The journey toward data-driven policies started with the original report and continues here.

The second report (AT&L, 2014) built on the first, adding another year of data to the series of cost, schedule, and technical performance measures while reporting insights from new policy analysis. Some signs of improvements were seen—but not everywhere, and we recognized that the extremely poor performance of outliers required further analysis given their distorting effect on the portfolio and the reputation of the system. Expanded analysis of the correlation of contract type on cost and schedule outcomes found that the prevalent debate on whether "cost-reimbursement" or "fixed-price" contracts are best at controlling prices is a red herring. The real issue is how effective the incentives are for each contract type based on the situation at hand. Also, firm-fixed-price contracting may not result in fixed prices in the end because those contracts can be modified to change content as needed. We found that incentive contracts (cost-plus-incentive-fee and fixed-price-incentive) control cost, price, and schedule as well as, or better than, other types—and with generally lower yet fair margins. Each situation depends on risk, cost knowledge, uncertainty, and a number of other factors, so we need to avoid dictating a single approach.

This third annual report builds on and extends the series of data from the past two reports. We see more statistically significant trends and differences, so we have more confidence that the changes are real. Some trends show improvements while others remain stagnant. New analyses have been added to begin examining further inputs (e.g., workforce improvements and measuring government execution costs), factors that affect outcomes (e.g., budgetary climate effects on cost growth), and trends in inferred measures of complexity and risk.

Below are highlights of this third report (along with page references to more detailed discussion later in the report), but there are other relevant results in the report. As with the prior annual reports, many analyses are beginnings and indicate areas for further work, but in others we now see similar indicators in multiple measures, increasing our confidence in the results.

Many of these results parallel the recent Government Accountability Office (GAO) reports that also use Selected Acquisition Reports (SARs) but calculate performance in somewhat different ways (see GAO, 2014a, 2015b). While these reports use SARs that are 1 year behind those used in our annual performance reports,[1] the GAO also identifies improvements in some measures along with areas in which progress is still needed.

Funding Growth and Defense Acquisition Executives

Policy and execution decisions by DoD executives should bear on the effectiveness of the overall acquisition system during and after their tenures. This is particularly true for the program structure and associated baselines set at Milestone (MS) B against which future cost performance is measured. Therefore, in our annual reports we track the performance of programs started under different acquisition executives to help reinforce accountability and provides an initial look for possible trends for further analysis.

Figure H-1 and Figure H-2 shows growth in MDAP[2] Planned Total Funding in development and procurement (respectively) for active[3] and completed MDAPs against original baselines as

[1] For example, GAO's 2015 report uses the December 2013 SARs while our 2015 report uses the December 2014 SARs.

[2] MDAPs are DoD acquisition programs that are not highly sensitive classified program and are either: (a) designated as such by the Secretary of Defense, or (b) estimated to require an eventual total expenditure for RDT&E of more than $300 million (in FY 1990 constant dollars) or an eventual total expenditure for procurement (including all planned increments or spirals) of more than $1.8 billion (in FY 1990 constant dollars)—see 10 U.S.C. section 2430(a).

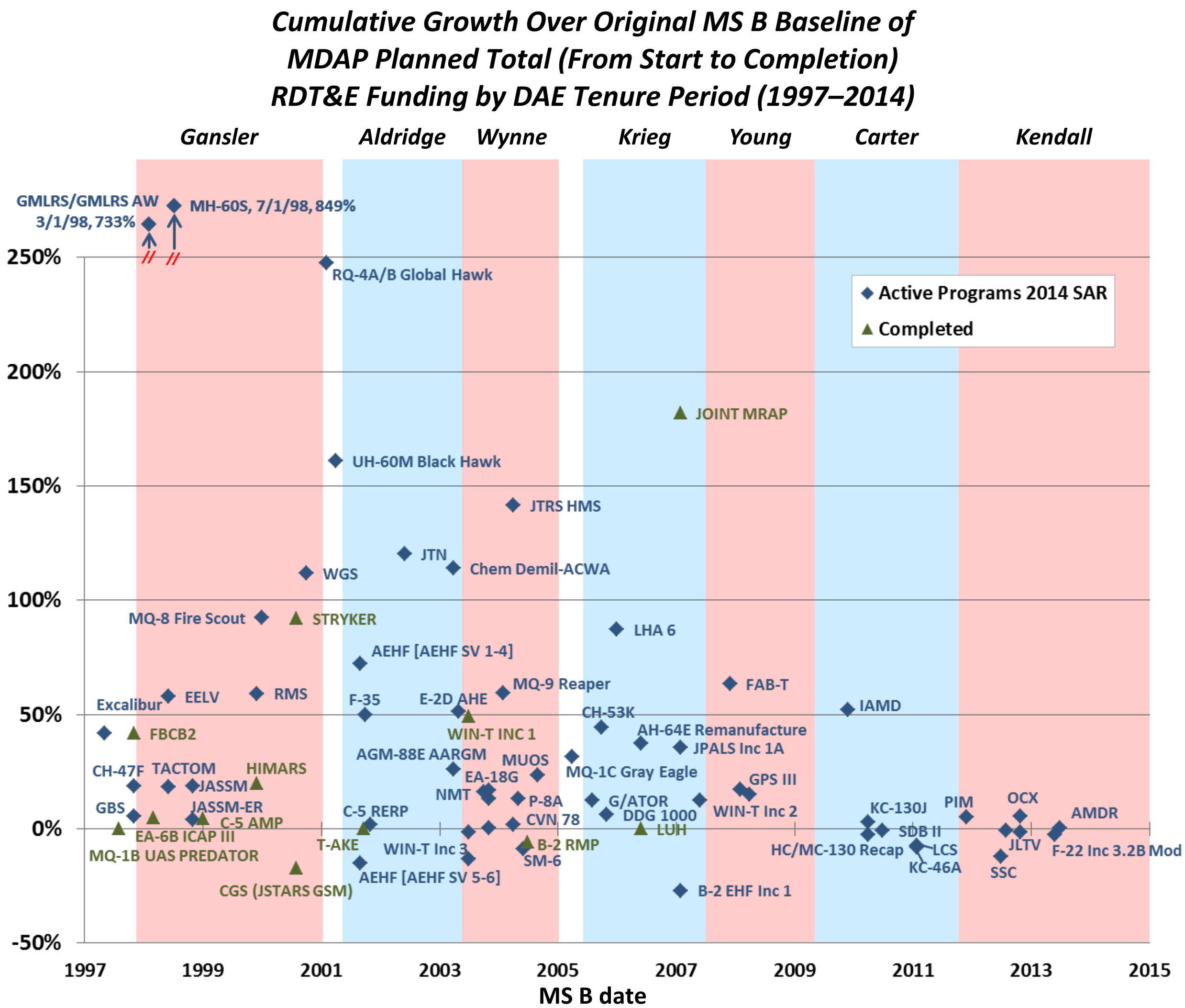
[3] In this report, *active* MDAPs are those that provide SARs and have passed MS B. MDAPs cease providing SARs and are deemed *completed* after they deliver 90 percent of total U.S. quantity or spend 90 percent of planned expenditures. See 10 U.S.C. section 2432(g).

reported to Congress in the SARs. Note that SAR funding data reflects what the program manager (PM) currently estimates will be needed in total by the end of the program for the current program configuration, including past actual funding, the current budget request, planned funding in the Future-Years Defense Program [FYDP], and planned funding beyond the FYDP to the end of the program. Growth is measured against the baseline set at the original MS B and can be positive or negative.

These figures also show the person who was the Defense Acquisition Executive (DAE) at the time of the MDAP's MS B approval. Later in the report we show similar charts for the programs started under different Service Acquisition Executives (SAEs) in the three military departments. This year we used total needed program funding instead of contract cost growth for these charts since needed funding is measured directly against the MS B baseline set by the DAE. (*See discussion starting on p. 85*).

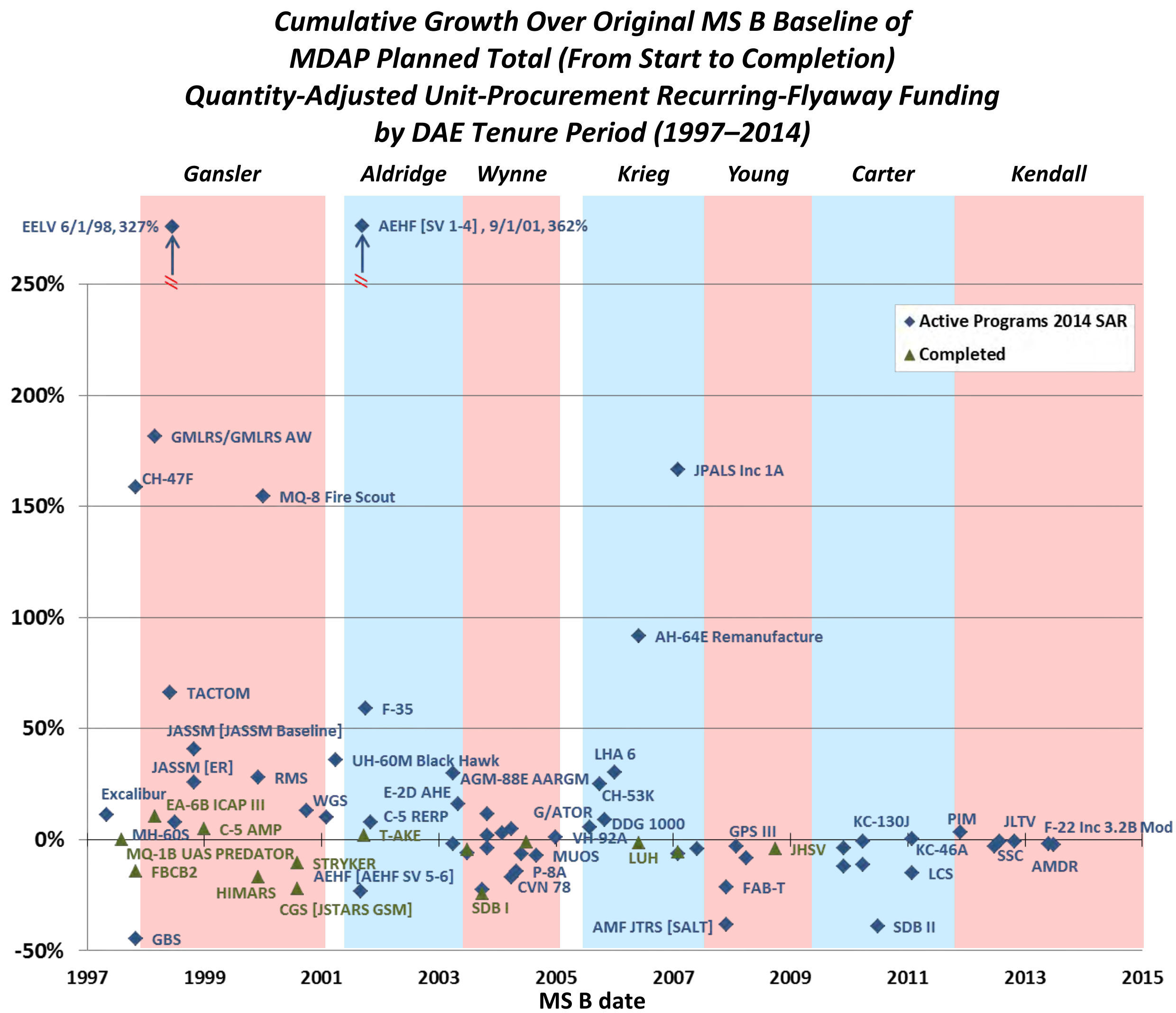
Caution is warranted, however. These charts do not reflect the effectiveness of subsequent oversight or major program changes by later DAEs during execution oversight, and they do not reflect statistical analysis to control for other internal and external variables that could have led to program success or problems. Defense acquisition is complex, and each measure has its strengths and weaknesses, so attributing performance to a single measure is subject to the limitations of that measure. For example, some programs may appear to be performing well in terms of total needed Research, Development, Test and Evaluation (RDT&E) funding but may be having problems reflected in other measures (e.g., total needed procurement funding; estimated operational costs; or cost growth on one of the program's major contracts). Thus, a combined examination of available data is important before reaching conclusions. Nevertheless, they are a crude indicator of the effectiveness of these officials' decision making. *(See detailed discussion on p. 12)*.

Figure H-1. Program Cost-Related Development Performance Baselined in DAE Periods



NOTE: This shows total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. These are percentage changes after adjusting for inflation from original MS B baseline of actual past and estimated future funding as reported in each program's latest SAR. Total RDT&E is an insightful measure because it is necessary regardless of quantity. White bars between DAE shaded regions represent periods with no confirmed executive. Relatively new programs that have not spent at least 30 percent of their original Engineering, Manufacturing and Development (EMD) schedule are not shown.

Figure H-2. Program Cost-Related Procurement Performance Baselined in DAE Periods



NOTE: This shows growth in total unit recurring flyaway needed funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. White bars between DAE shaded regions represent periods with no confirmed executive. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown.

Cost and Schedule Improvements

Recent data on MDAPs at the program- and contract-level have shown some statistically significant trends in funding, price, cost and schedule control. Many indicate improvements although complicating factors raise caveats and potential concerns.

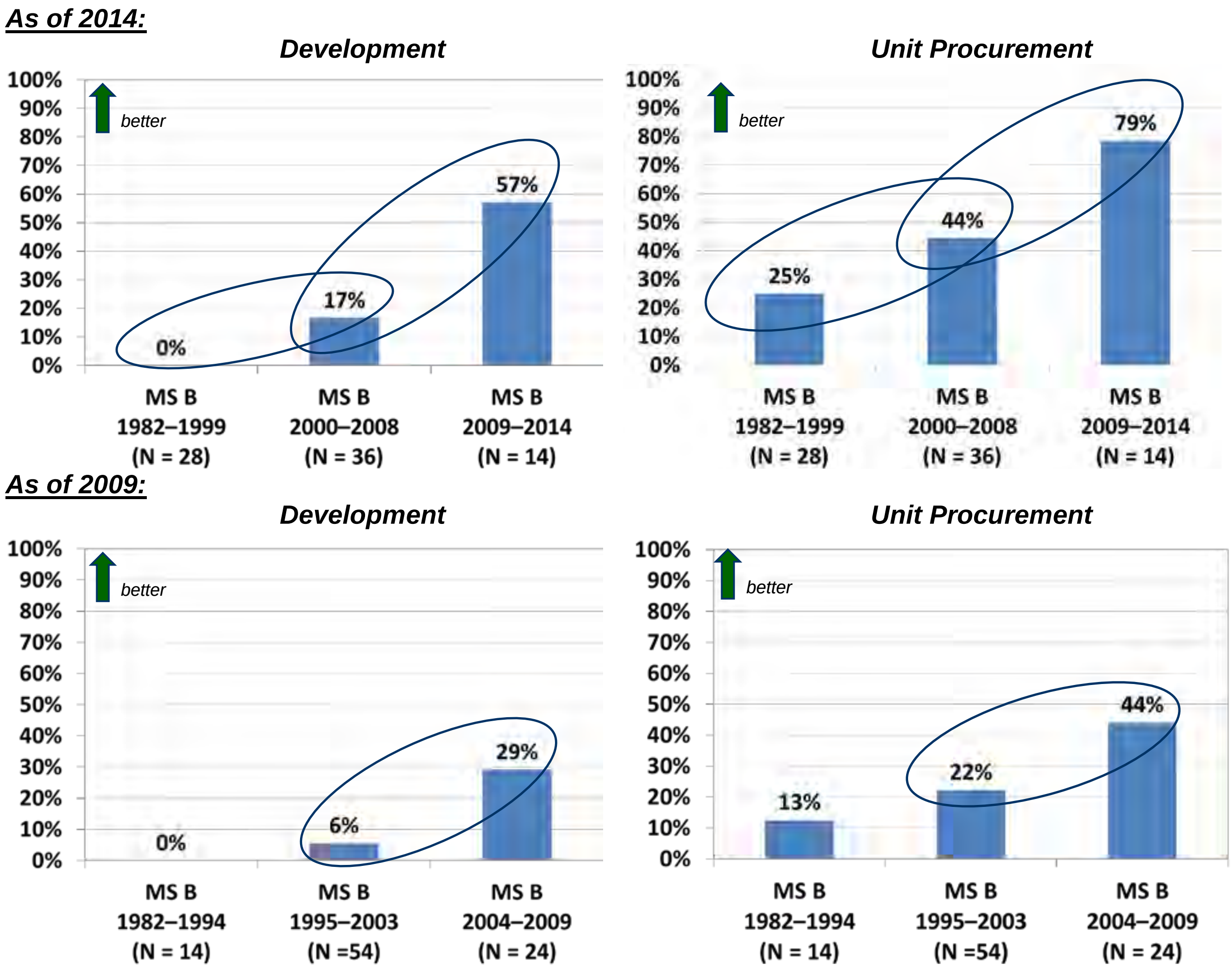
Program-Level Measures

First we highlight trends in performance at the program level.

More MDAPs are showing program funding reductions in both development and production. Relative to their original MS B baselines, more active MDAPs by proportion are estimated in the SARs to have total RDT&E funding reductions as of 2014 than as of 2009—even after we remove relatively new programs that would be unlikely to yet show growth (see Figure H-3). We will conduct further analysis to better understand these differences. *(See detailed discussion starting on pp. 40 and 64).*

Figure H-3. Planned Reductions in Program Funding

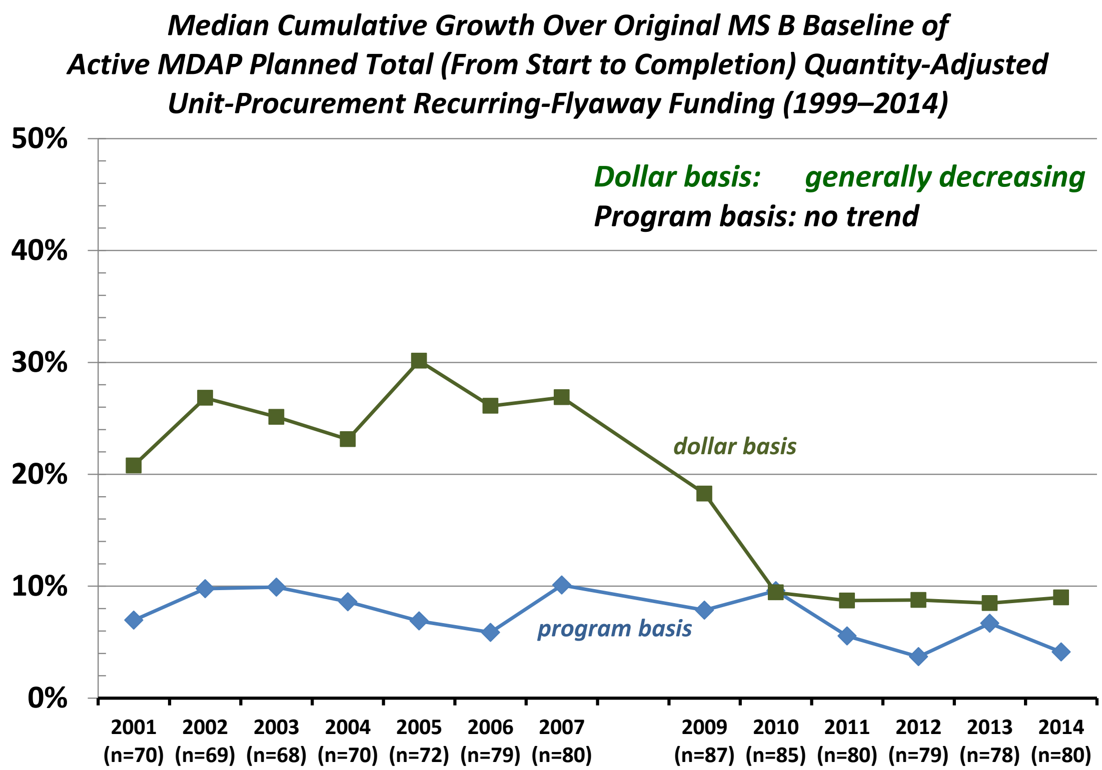
Proportions of Active MDAPs With Reductions Since Original MS B Baseline in Cumulative Planned Total (From Start to Completion) Funding (2009 and 2014 SARs)



NOTE: Development funding is total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. Procurement funding is growth in unit recurring-flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. Statistically significant differences between adjacent periods are marked with an oval. A program shows a reduction if current total is below the original MS B baseline. Relatively new programs that have not been through at least 30 percent of their original EMD schedule are not included to reduce bias from newer programs.

Lower total MDAP program funding growth since original baselines in production. Adjusting for quantity changes and the dollar size of programs, the median quantity-adjusted unit funding growth since original MS B baseline has been statistically lower since 2010 (see the dollar-basis line in Figure H-4). That figure also shows that on a program basis, the recent total unit funding appears somewhat lower but the differences are not statistically significant. In other words, larger active MDAPs (by dollar) generally have brought their growth in total unit procurement funding needs to levels close to the median for all MDAPs regardless of size. This is not the case in development, where increases are seen by program and dollar—see Figure H-18 below. *(See detailed discussion starting on p. 32)*

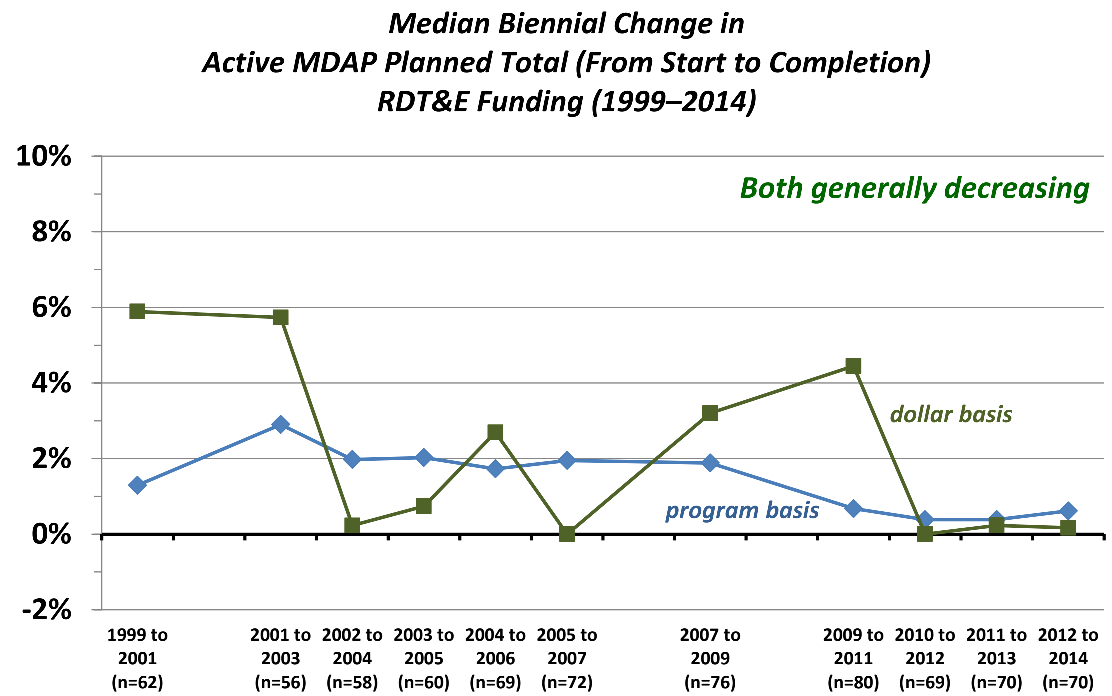
Figure H-4. Program Cost-Related Performance: Procurement



NOTE: This shows growth in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. The trend on a dollar basis (weighting by program size) is statistically significant, but there is no trend on a program basis (unweighted by dollar size). These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not included.

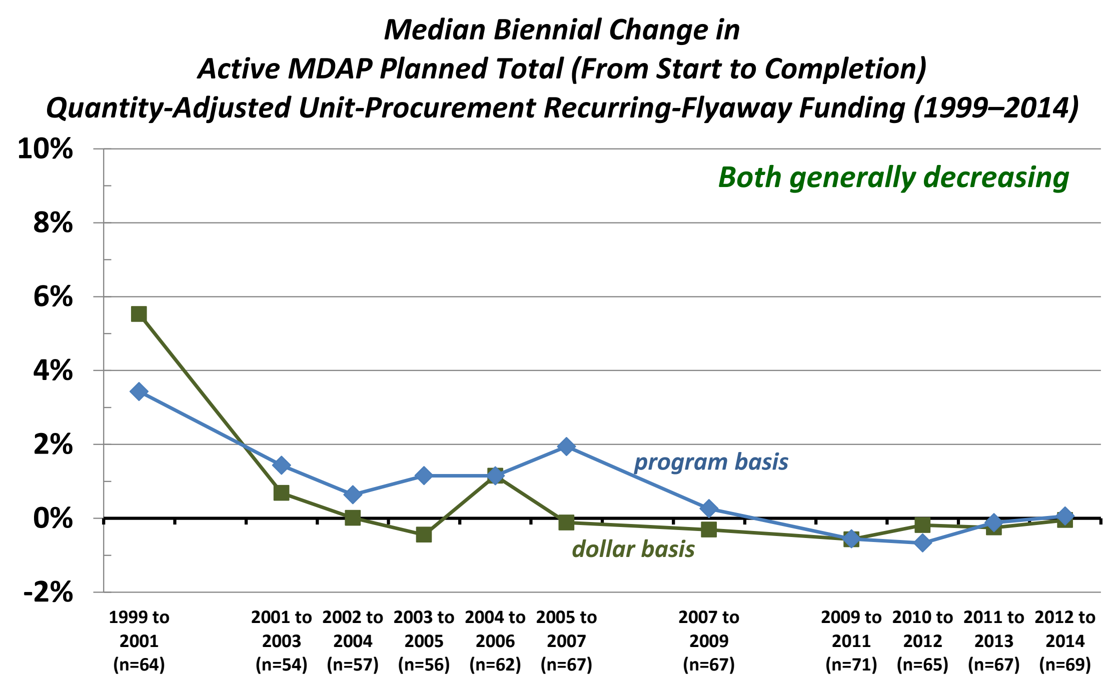
Lower biennial change in MDAP program funding for both development and production. Median biennial change in funding growth has decreased statistically in recent years both on a program basis and when adjusting for the size of programs (i.e., on a dollar basis)—see Figure H-5 and Figure H-6. Since 2010, biennial changes ranged from −1 to 1 percent. These are measured using total program RDT&E funding and quantity-adjusted unit procurement (recurring unit flyaway funding), including past and needed future funding. *(See detailed discussion starting on pp. 37 and 62.)*

Figure H-5. Program Cost-Related Biennial Performance: Development



NOTE: This measures biennial changes in total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. Both trends are statistically significant. Total RDT&E is an insightful measure because it is necessary regardless of quantity. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not included.

Figure H-6. Program Cost-Related Biennial Performance: Procurement

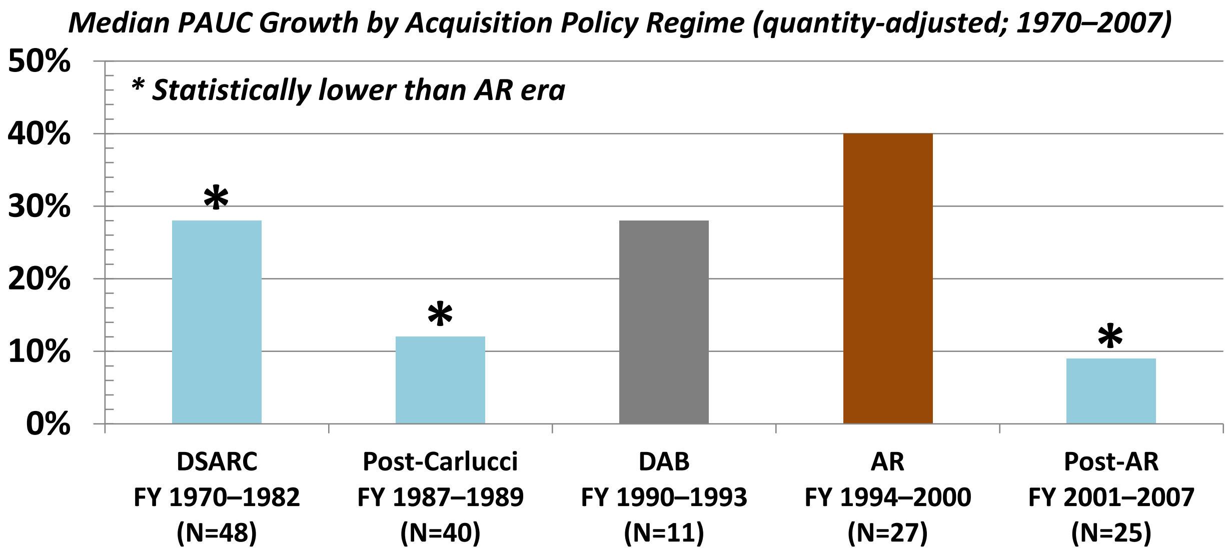


NOTE: This measures biennial changes in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. Indicated trends are statistically significant. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not included.

MDAP Program Acquisition Unit Cost (PAUC) Growth Dropped in the Early 2000s. PAUC is a different funding measure define by statute (10 United States Code [U.S.C.], sections 2430a and 2432) consisting of the total acquisition funding[4] divided by the acquisition quantity as reported in the SARs. For PAUC, "cost" is synonymous with the total amount of funding because it reflects the prices paid on RDT&E and production contracts as well as program execution costs. Later, when we discuss contracts, we will distinguish contract prices from their underlying contractor costs and margins (profits and fees).

McNicol and Wu (2014) found that MDAPs started since 2001 had significantly lower quantity-adjusted PAUC growth than those begun during the so-called Acquisition Reform (AR) era of the mid-1990s (see Figure H-7). The post-AR values appear lower than those from 1970–1993, but those differences were not reported to be statistically significant. Note that this analysis was focused on final results from these policy eras and thus did not use newer programs with MS B dates after 2007. Further analysis is warranted to see what has happened more recently as measured by PAUC, but this indicates that cost growth was lower before and since the AR period. *(See further data and discussion starting on p. 105 and in McNicol and Wu [2014])*.

Figure H-7. Program Cost-Related Performance in Policy Regimes

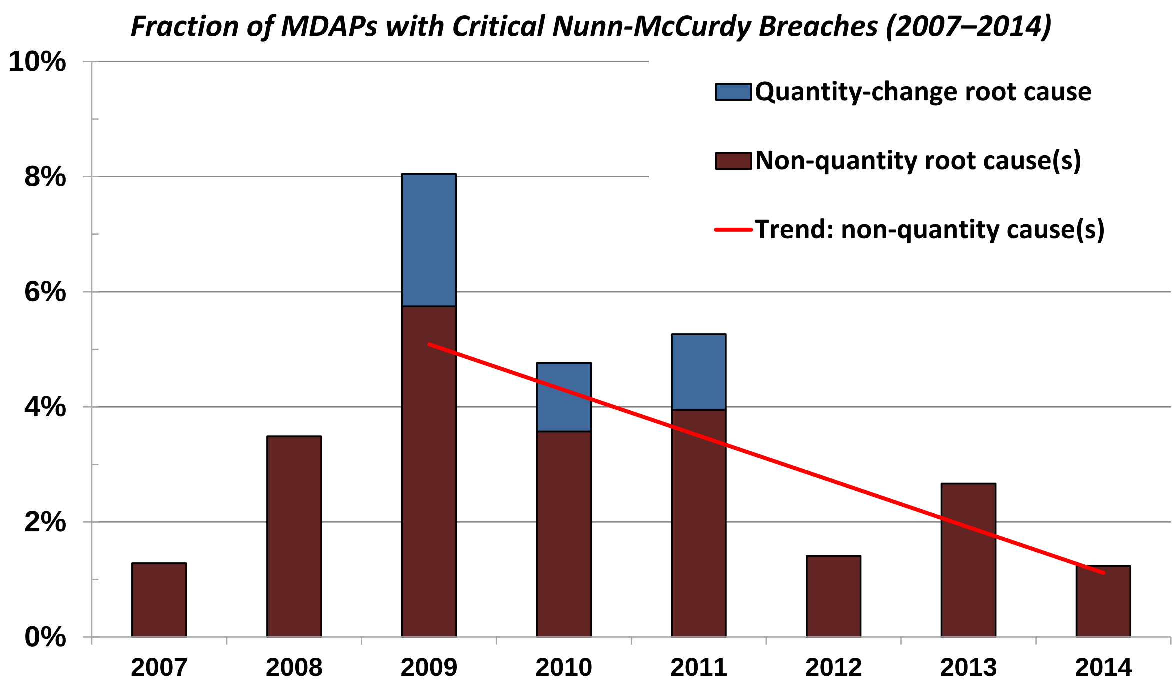


* Underlying distribution is statistically lower than that of the AR regime of 1994–2000. However, the Defense Acquisition Board (DAB) era's distribution was not different enough from the AR era to be statistically significant. Source: McNicol and Wu (2014). NOTES: The differences between the three starred era were not statistically significant. There may be further cost growth in these eras, particularly the post-AR regime. McNicol and Wu adjusted PAUC growth for any quantity changes. DSARC is the Defense Systems Acquisition Review Council (DSARC) era.

[4] *Acquisition* funding is equal to the sum of: (1) the development funding for prime mission equipment and support items, (2) the procurement funding for prime mission equipment, support items, and initial spares, and (3) the system-specific facilities funding.

Lower recent rates of Nunn-McCurdy breaches. As shown in Figure H-8, there are statistically significant downward trends since 2009 of both nonquantity-related critical breaches (shown) and all critical Nunn-McCurdy cost-growth breaches.[5] For the trend analysis, we used the breach rates instead of counts to control for changes in portfolio size between years, although the patterns are very similar because the size of the MDAP portfolio is relatively stable. *(See detailed discussion and data starting on p. 19.)*

Figure H-8. Program Cost-Related Performance: Nunn-McCurdy Breaches



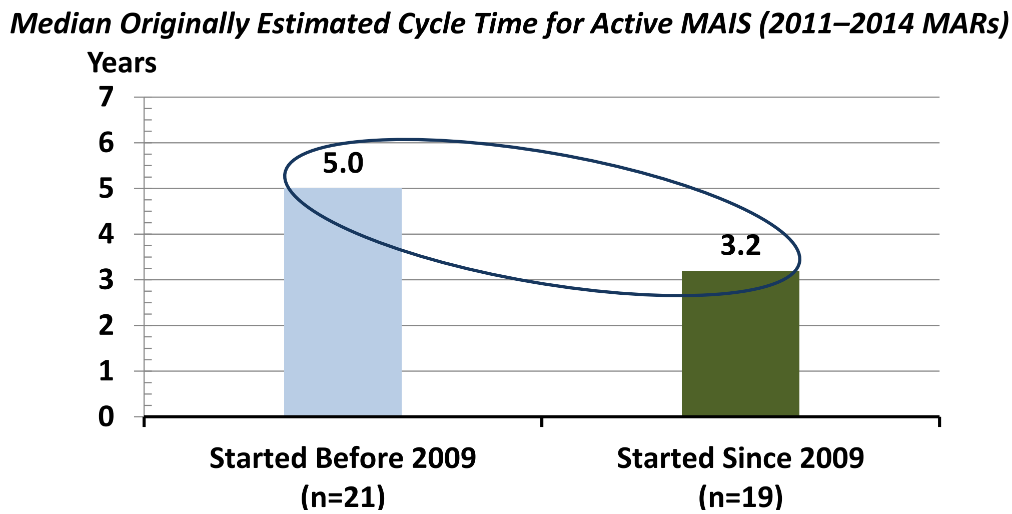
NOTE: Breaches due to quantity changes are based on root-cause analyses performed by the Office of Performance Assessments and Root Cause Analysis (PARCA) in AT&L. Since PARCA was not established until the Weapon System Acquisition Reform Act (WSARA) of 2009, it is unknown whether quantity changes were a root cause of breaches before 2009. There is a statistically significant downward trend in both total critical breaches and nonquantity-related critical breaches since 2009. Breaches are after adjusting for inflation. Since it usually takes a few years before a program might breach again, we removed programs from the portfolio count that have breached recently to avoid the potential bias towards an artificially low breach rate. Also, relatively new programs that have not spent at least 30 percent of their original EMD schedule are not included because they are unlikely to breach, and thus would bias the results downward.

[5] Nunn-McCurdy "cost" growth thresholds are established by law and trigger reporting to Congress and other specific actions by the DoD (see discussion starting on p. 19). As discussed earlier with respect to PAUC, these "cost" measures reflect funding and include the underlying contractor and government execution costs plus contractor margins (profits and fees).

Lower median Major Automated Information System (MAIS) cycle time. In addition to tracking cost and funding performance, the DoD is concerned with how quickly new capabilities can be delivered to our warfighters. Timely delivery affects operational performance against evolving threats. One important measure is cycle time or length between certain events of interest.

MAIS cycle time is particularly important given the fast pace of information technology advancement. Here cycle time is measured from either MS B or Funds-First-Obligated (FFO) to the Full-Deployment Decision (FDD). As shown in Figure H-9, the median cycle time as reported in the MAIS Annual Reports (MARs) has dropped from 5 years before 2009 to 3.2 years since 2009. This result may not necessarily mean we are faster at acquiring systems—all other things equal. For example, these results reflect any work-content changes. *(See detailed Figure 2-28 on p. 57 and associated discussion.)*

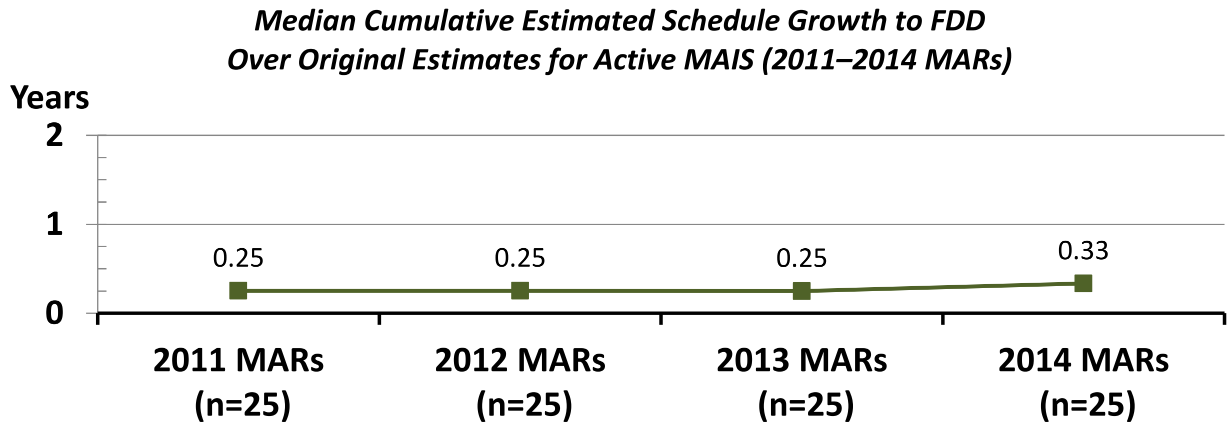
Figure H-9. Program Length: Information Systems



NOTE: These changes may reflect systemic reductions in how much work is included in a MAIS. Original estimates are those in the MAIS' first MAR. Included are the latest data on programs that appeared in at least one MAR from 2011 through 2014. Newer programs that have not completed at least 30 percent of their original EMD schedule time were excluded to avoid artificially low values.

MAIS program schedule growth. In addition to MDAPs, we track MAIS schedule growth to understand how well the DoD is keeping up with information technology and the needs for information system capabilities. As shown in Figure H-10, the 2011–2014 MARs exhibited a median growth of about three to four months. *(See detailed discussion on p. 28.)*

Figure H-10. Program Schedule Growth: Information Systems



NOTE: These measures do not control for any changes in work content or specifications. Original estimates are those reported in the first MAR for each MAIS. Schedule period is from MS B or FFO to FDD. Newer programs that have not completed at least 30 percent of their original EMD schedule time were excluded to avoid artificially low values.

Contract-Level Measures

Second, we highlight trends in performance at the contract level.

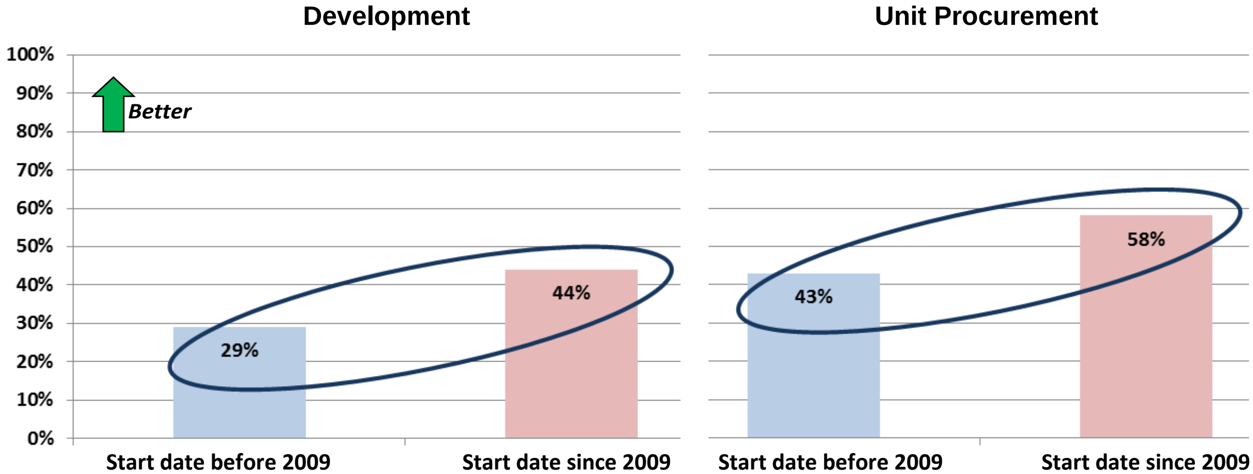
Higher proportion of contracts with negative price growth since 2009. To help get an earlier indicator of potential program performance, we also track performance on major MDAP contracts. We generally have earned-value (EV) data for major MDAP contracts, and EV data forms the basis for much of our contract analysis. Major contracts include the six largest contracts (prime, associated, or for government-furnished equipment) for each MDAP valued at more than $40 million. EV data also are available for other MDAP contracts of at least $60 million in RDT&E or $250 million in procurement or ship construction (in fiscal year [FY] 1990 constant dollars). Note that EV data are usually not provided for firm-fixed price (FFP) contracts, so this MDAP contract dataset has very few FFPs (zero of 120 in development and 6 of 160 early production contracts).

One useful measure of cost-related performance is the *price* that the DoD pays on a contract. Price differs from costs paid by the contractor in that price reflects the margin (profit or fee) on top of contractor costs. Thus, price growth reflects the net changes in work content, cost-over-target, and final margins on these contracts.

A statistically significant greater proportion of these major MDAP contracts started since 2009 generated price reductions (up from 29 percent to 44 percent in development, and up from 43 percent to 58 percent in early production—see Figure H-11). This may be due in part from the downward trend in cost-over-target. No contracts in this dataset were FFP, and only one was a hybrid with FFP. *(See detailed discussion starting on p. 70.)*

Figure H-11. Reductions in Program Contract Prices

Proportions of (Non-Firm-Fixed Price [FFP]) Contracts with Price Reductions (Negative Price Growth) From Original Contract Award for Currently Active MDAPs (controlling for inflation and maturity; 2000–2015)



NOTE: Improvements are statistically significant. Changes in price reflect the net changes in work content, cost-over-target, and final margins. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency. We controlled for maturity by ensuring the two subsamples had statistically similar work completion percentages. There were no FFP development contracts and only 6 FFP out of 160 early production contracts in this dataset.

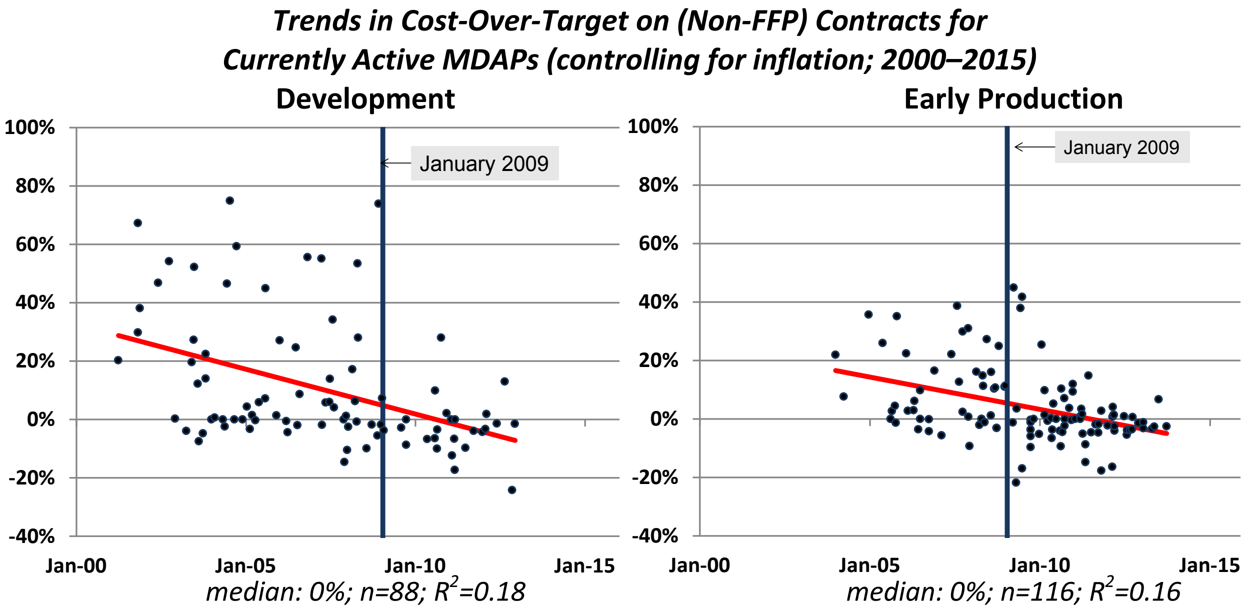
Downward trend in contract cost-over-target.[6] We also track contractor cost growth over the latest contract cost target to understand how well it is being controlled independent of work-content changes. Again, no contracts in this dataset were FFP, and only one was a hybrid with FFP.

On both development and early production major MDAP contracts starting within the years 2000–2015, contractors are doing better statistically at meeting or beating contract cost targets as shown by the latest or final PM's estimate at completion (see Figure H-12). Medians have dropped from 0 to −3 percent in development and from 11 to 0 percent in early production.

Also, the proportion of MDAP contracts started since 2009 that are at or below their cost targets (after adjusting for work-content changes) has increased significantly compared to those starting before 2009 (see Figure H-13). This improvement is indicative of the successful implementation of "should cost" and contract incentives in BBP. *(See detailed discussion starting on p. 70.)*

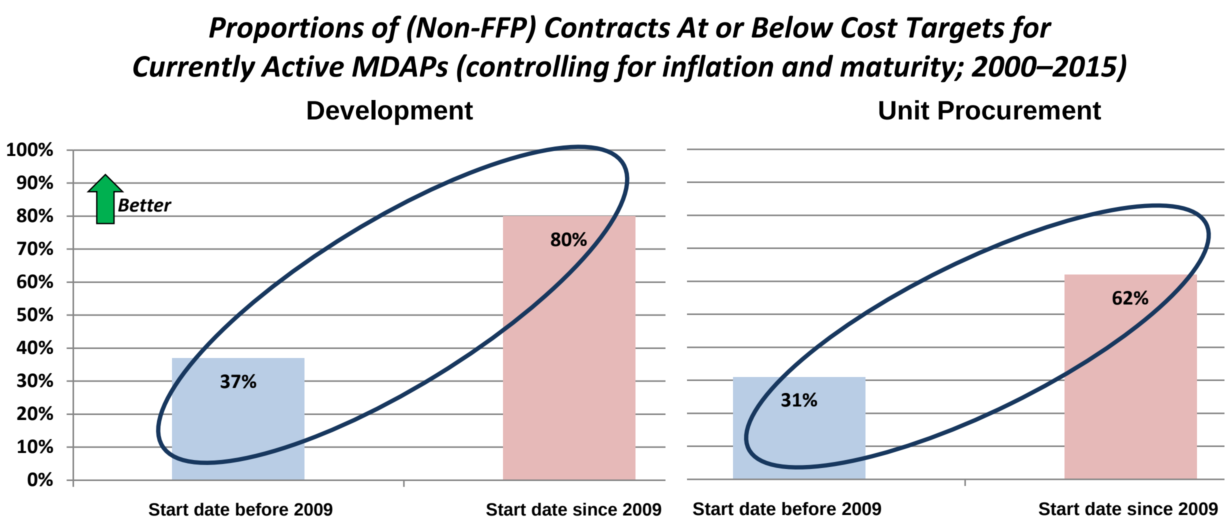
[6] Cost-over-target is the cost to the contractor over the latest target cost of the contract. It is defined as the latest PM's estimate-at-completion (EAC) minus the latest contract budget base (CBB), all divided by the initial CBB. The CBB reflects the contract target cost since contract initiation. Note that contract targets themselves are not pure, unbiased estimates of project cost and schedule. In addition to expected project scale (size), they also reflect other contract terms (e.g., share lines and incentives as well as market contestability and the general negotiating environment) and are, in part, the result of bidding strategies. Nevertheless, we use these data to tease out constructive insights, and plan to provide further analysis in a future report.

Figure H-12. Program Contract Costs Related to Targets



NOTE: Cost-over-target is cost growth relative to the latest cost target and is separate from work-content changes. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency. The differences in work-completion percentages before and since 2009 were statistically insignificant, so these results are not due to more recent contracts being immature relative to older contracts. There were no FFP development contracts and only 6 FFP out of 160 early production contracts in this dataset.

Figure H-13. Reductions in Program Contract Costs Related to Targets

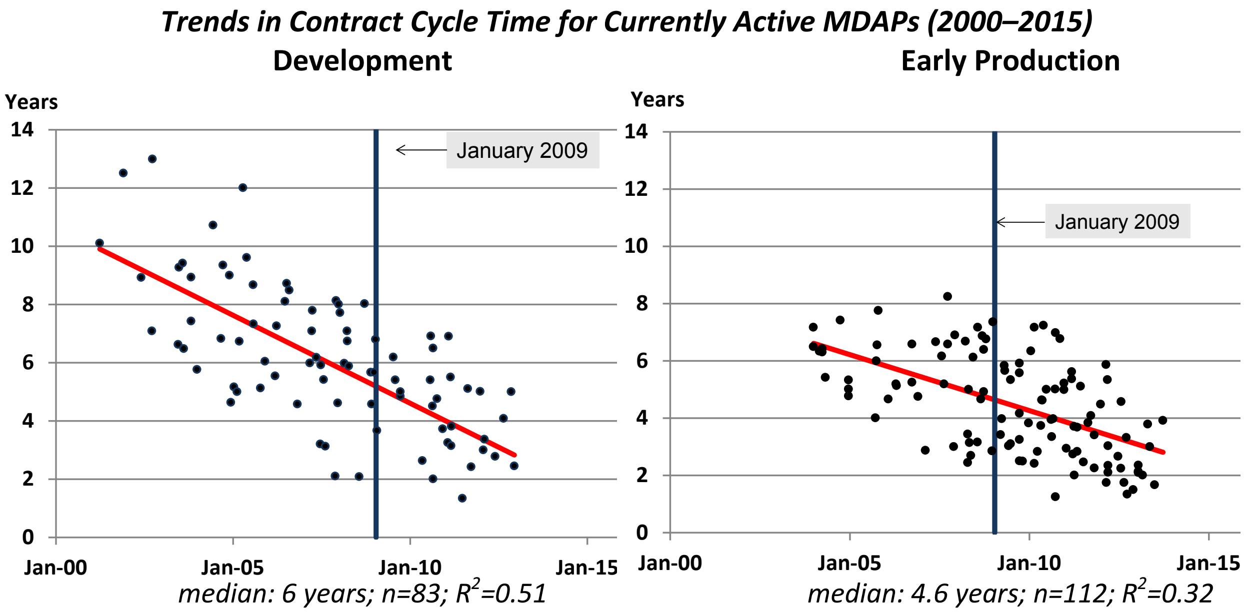


NOTE: Cost-over-target is cost growth relative to the latest cost target and is separate from work-content changes. Improvements are statistically significant. We controlled for maturity by ensuring the two subsamples had statistically similar work completion percentages. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency. There were no FFP development contracts and only 6 FFP out of 160 early production contracts in this dataset.

Downward trend in contract cycle time. For contracts, we measure cycle time as the latest contract length (in years, including any schedule growth). Contract cycle time gives an early indication of potential schedule-related effects at the program level.

Major development and early production MDAP contracts have been getting shorter over the years 2000–2015 (see Figure H-14). This is independent of changes in work content growth, which has been generally flat since 2000. This result may not necessarily mean we are faster at acquiring systems—all other things being equal—and this trend started long before BBP. This trend is likely due at least in part to fewer contracts being for new, cutting-edge systems (and thus more for system modifications). There are much fewer new-system development contracts since 2009 than before (*see Figure 2-40 on p. 79 later in the report*). New-system development contracts in this dataset were about 2 years longer at the median (7 years in length) than system modification contracts (about 5 years in length). Evidence and theory indicate this may be due to reduced complexity,[7] and there may be other underlying factors as well. Further analysis is needed. *(See detailed discussions starting on pp. 70 and 109.)*

Figure H-14. Program Contract Lengths

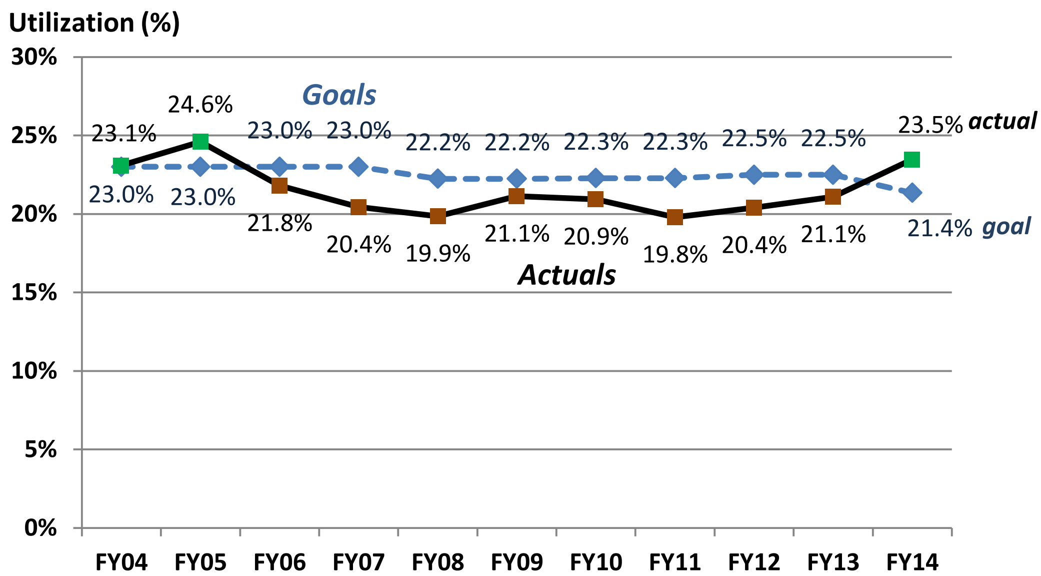


NOTE: Cycle time is the latest contract schedule length (not schedule growth). Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency, but the net results were small (e.g., in development, with the outliers the median remained at 6 years). The differences in work-completion percentages before and since 2009 were statistically insignificant, so these results are not due to more recent contracts being immature relative to older contracts.

[7] In this report, we use the term "complexity" to refer to the sophistication and complications associated with the contract work or program in question, not necessarily in the sense of non-linear or chaotic nature. It could include any of the four categories discussed by Remington and Pollack (2007)—structural, technical, directional, or temporal complexity. However, we do not explicitly assess specific aspects but rather attempt to use readily available cost and schedule data as surrogates to measure the net effect of complexity in contract work.

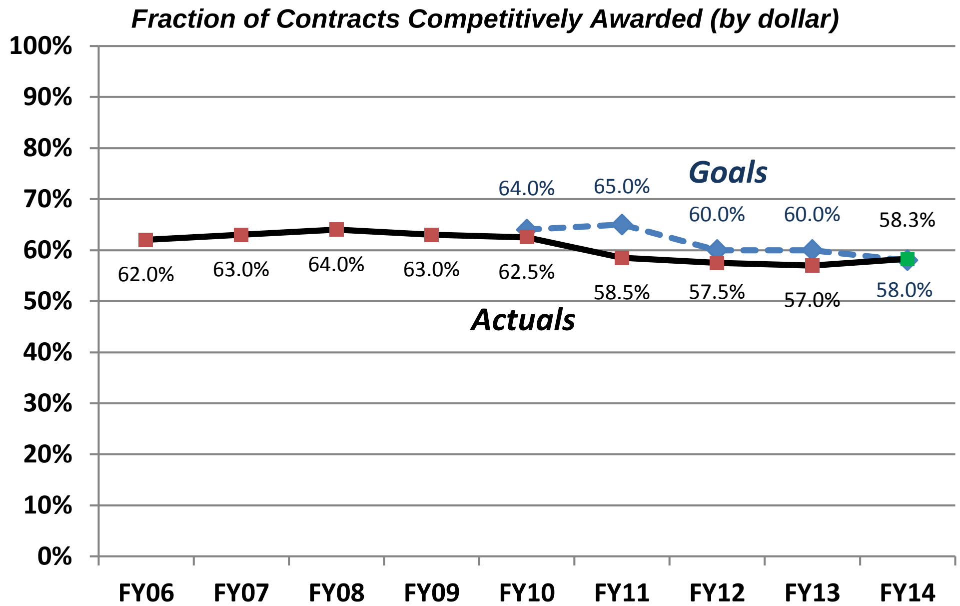
Increased small-business utilization. Figure H-15 shows actual DoD-wide small-business utilization (measured by obligations) relative to yearly goals. Recent trends since FY 2011 have been improving; we have exceeded our FY 2014 goal by 2.12 percentage points and surpassed prior years except FY 2005.

Figure H-15. Small-Business Prime-Contracting Utilization Trends: Goals and Actuals (FY 2001–FY 2014)



Competition has started to rise. Figure H-16 plots the percentage of all DoD contract dollars that were competitively awarded from FY 2006 to FY 2014. Goals were established starting in FY 2010. Since goals were established in FY 2010, we had declining actuals until we made progress in FY 2014 at reversing the trend and meeting our goal.

Figure H-16. Competition Trends: Goals and Actuals (FY 2006–FY 2014)



NOTE: We did not establish goals until FY 2010. Fraction of contracts competitively awarded is measured on a dollar basis.

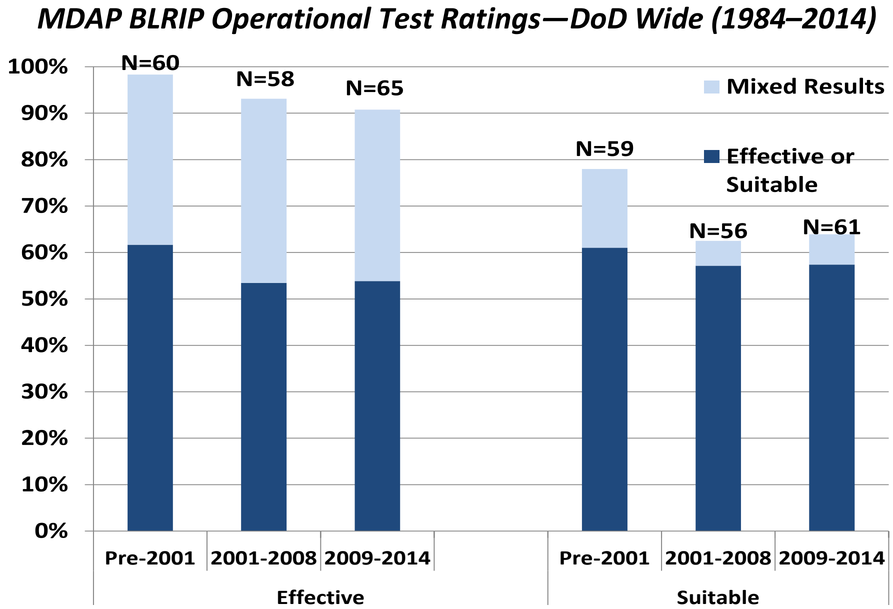
Measures Where Improvement Is Needed

Measures of performance show areas where improvement continues to be needed.

Initial operational test ratings remain about the same. The whole reason we have defense acquisition is to provide operational capabilities to our warfighters against current and evolving threats. Cost and schedule control are important, but more important is the relative value of operational benefits given costs. Operational performance goes beyond merely meeting technical requirements established before program inception. Threats can change and those initial requirements may lag operational aspects important to performance in the field. One measure of performance is the operational test results reported by the DoD Director of Operational Test and Evaluation (DOT&E) at the end of low-rate initial production [LRIP]. These Beyond-LRIP (BLRIP) results provide independent data on the operational effectiveness and suitability of the system at this point.

Figure H-17 summarizes the results of these BLRIP operational tests. The differences between the time periods are not statistically significant, and we are not able to distinguish statistically significant differences based on the incumbent DOT&E. *(See detailed discussion starting on p. 16 and 81.)*

Figure H-17. Program Technical Performance

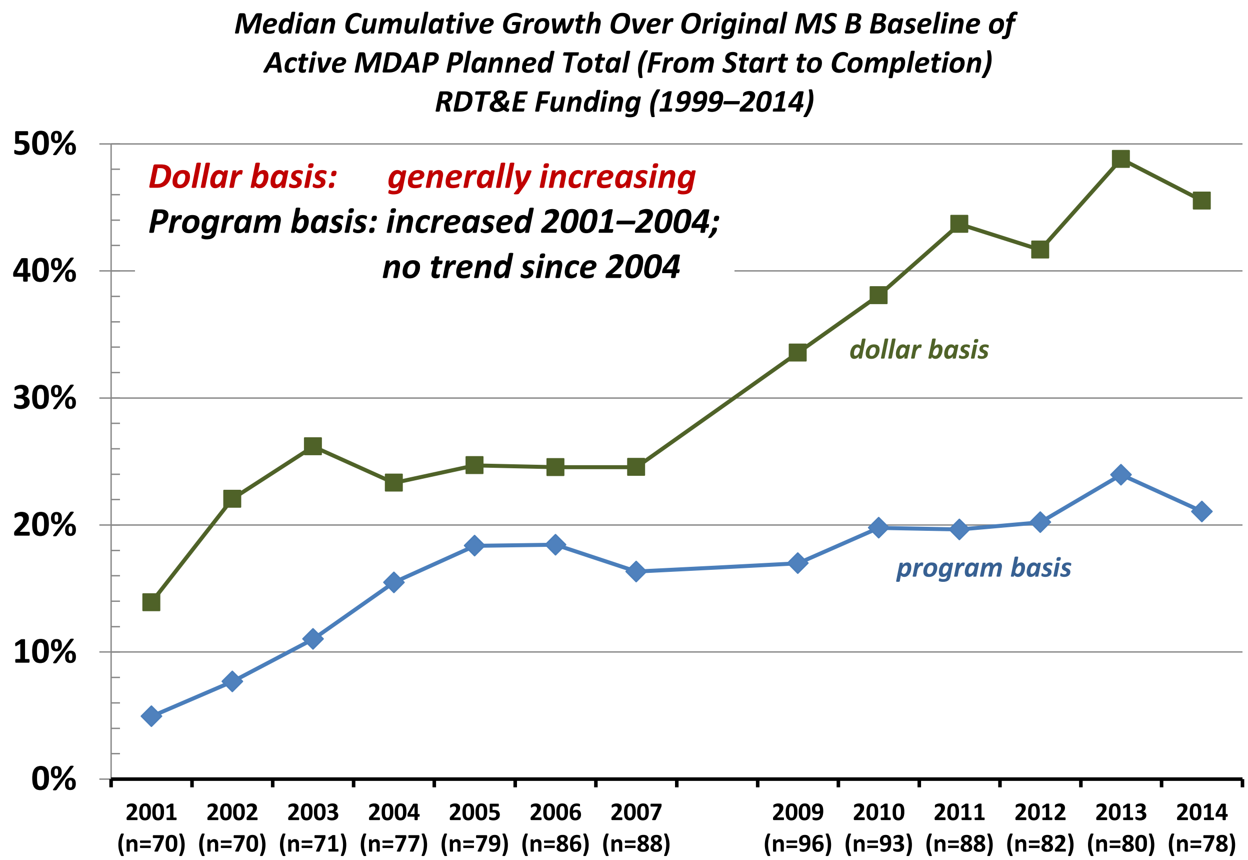


Source: DOT&E BLRIP reports.

NOTE: Differences are not statistically significant. Sample sizes differ between Effective and Suitable for some Components because effectiveness and suitability could not be determined in all cases.

Higher total MDAP RDT&E funding growth since original baselines. While biennial changes in total planned and actual RDT&E funding growth has been decreasing recently, RDT&E funding over original MS B baselines has been increasing since 2001 on a dollar basis but statistically flat since 2004 on a program basis (see Figure H-18). Since recent biennial changes in planned and actual total funding have been near zero instead of strongly negative at the median, this metric is unlikely to reverse (even if no more RDT&E growth occurs) until programs with earlier RDT&E growth (e.g., the F-35, which had significant historical development cost growth but has been stable since the Nunn McCurdy breach in 2010) exit the MDAP portfolio. *(See detailed discussion starting on p. 32.)*

Figure H-18. Program Cost-Related Performance: Development



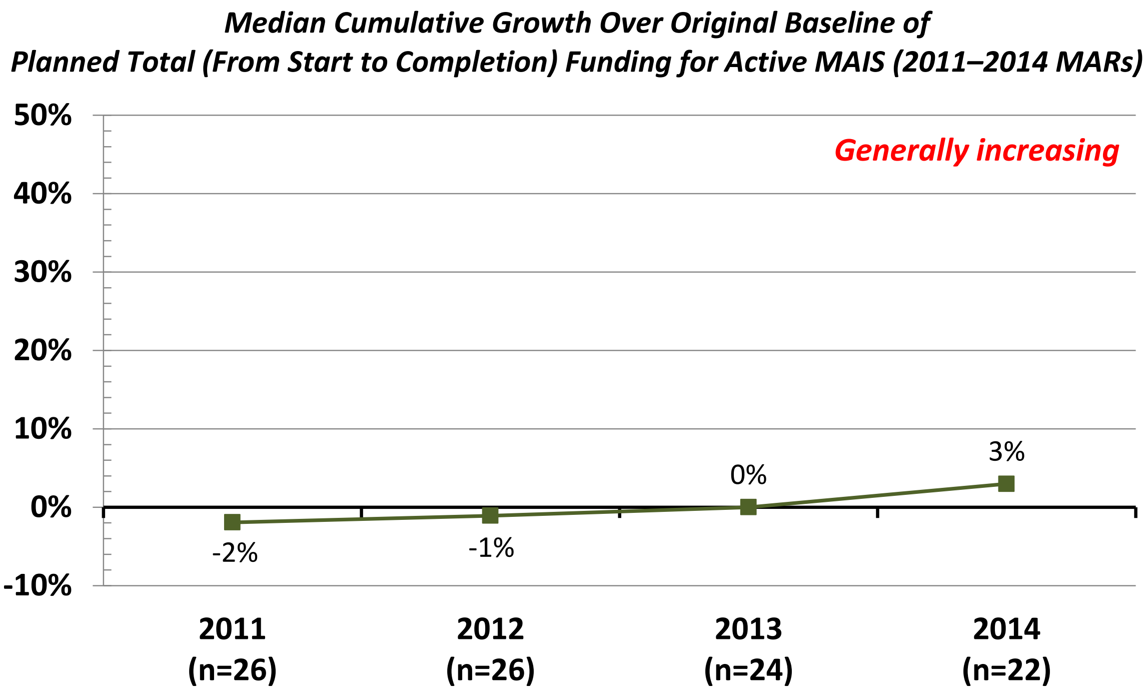
NOTE: This shows total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. Both trends are statistically significant. These are percentage changes from original MS B baseline after adjusting for inflation of actual past and estimated future funding as reported in each program's latest SAR. Total RDT&E is an insightful measure because it is necessary regardless of quantity. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not included.

Limited evidence of improved MDAP contract price performance. While proportion comparisons offer some indications that price-growth performance has improved since 2009, it does not indicate whether this is due to policy or other factors. In the first place, cost-over-target has shown signs of improvement (see above), but it is generally a small share of total cost and price growth (i.e., they also consist of work-content growth in contract modifications, which tends to dominate—see discussions throughout this and prior reports). The growth attributable to increased engineering and system requirements (as seen in contract modifications) has shown no downward trends in development and mixed changes in early production.

Detailed analysis of development contract outliers reveals some cautionary lessons. Our analysis of outlier development contracts shows that a spiral-like development strategy can be associated with large cost-growth outliers. In such cases, it is particularly hard to distinguish between expected work-content growth due to the nature of incremental development from problematic, unplanned growth, so deeper analysis is needed to try and identify specific causes. *(See detailed discussion starting on p. 45.)*

Rising median MAIS funding growth. As shown in Figure H-19, the median funding growth from original baselines as reported in the 2011–2014 MARs had been rising slowly and is now greater than zero at 3 percent in the 2014 MAR. These differences are statistically significant. *(See detailed Figure 2-8 on p. 30 and associated discussion.)*

Figure H-19. Program Cost-Related Performance: Information Systems



NOTE: Trends is statistically significant. Growth may reflect content changes. Newer programs that have not completed at least 30 percent of their original EMD schedule time were excluded to avoid artificially low values.

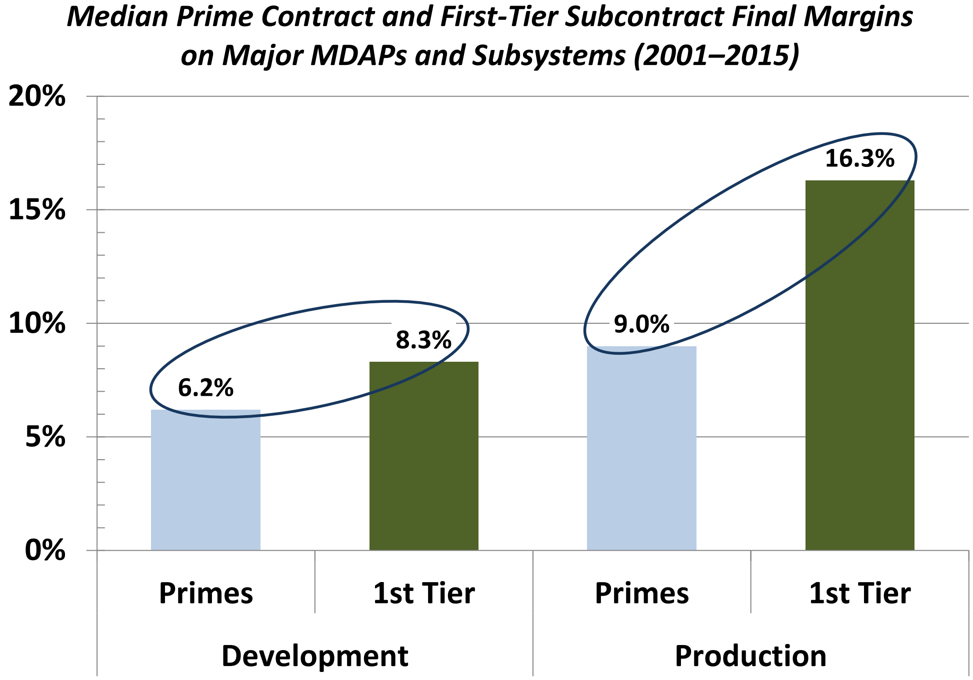
Other Observations

First-tier subcontractor final margins have been generally higher than for prime contractors. Since 2001, first-tier subcontractors earned higher margins than their associated prime contractors on the same program (at the median, about 2 percentage points higher in development and about 7 percentage points higher in production—see Figure H-20).[8] These are the large-scale subcontractors for which we have detailed cost data—not those providing commercial items. While not uniform, the differences are statistically significant and are

[8] Note that we calculated these as margins (fraction of price going to fee or profit), not markups (fraction of cost going to fee or profit). Markups at these percentages are slightly higher.

particularly large in production. Further analysis is needed to understand the reasons for these differences, but these illustrate why the DoD has been working over the past few years to motivate prime contractors to control subcontractor prices and ensure that profitability is aligned with performance—especially in production where the difference in margins is large. We are also concerned that higher subcontract margins may be motivating companies to bid on fewer prime contracts and thus reduce competition at that level. This information will be used to inform contract negotiations. *(See detailed discussion starting on p. 118.)*

Figure H-20. Program Contract Final Margins: Prime and First Tier



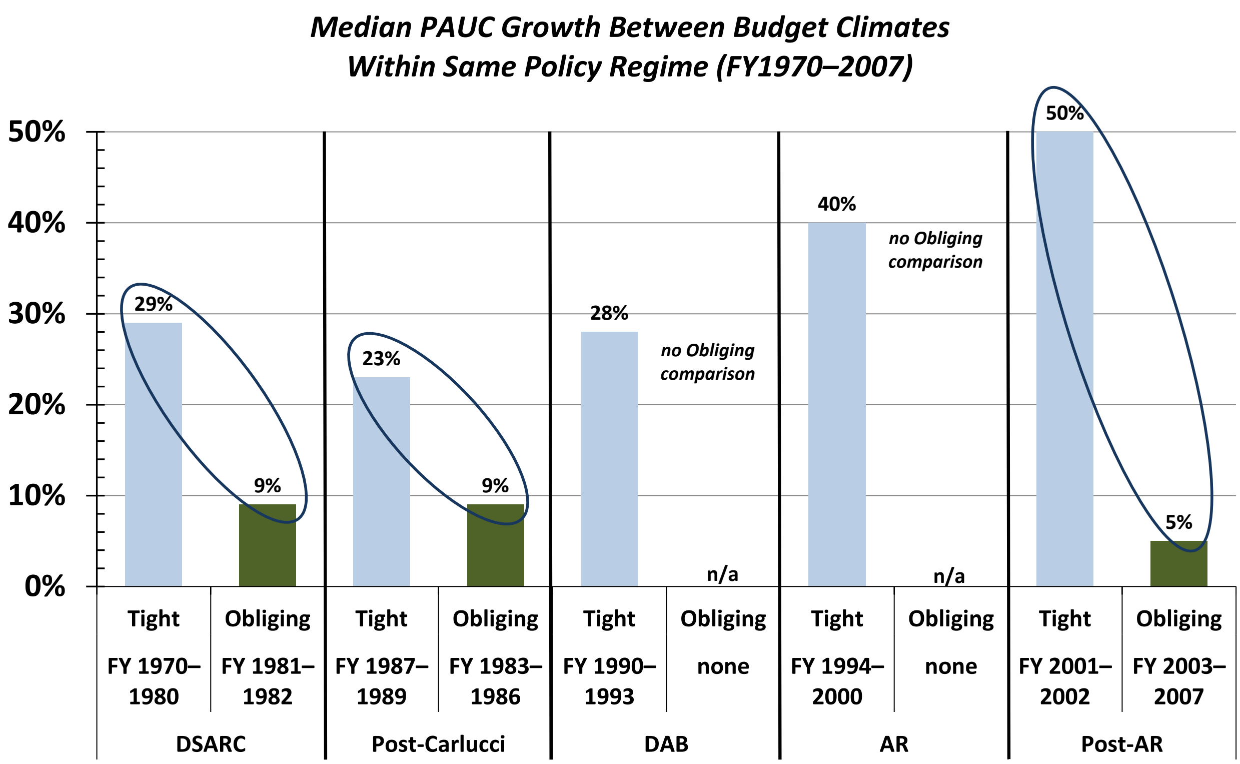
NOTE: Differences between prime and first-tier subcontractor margins are statistically significant. Values were weighted by the relative dollar size of the contracts. Margins are those for which DoD has cost data on similar activities for MDAPs and major subsystems.

Technical superiority concerns. The confluence of a number of results appears to support concern that we may be slowing our investment in technical advanced systems. Macro RDT&E budget accounts are emphasizing upgrades (budget activity [BA] 6.7) over new systems (BA 6.5; *see Figure 1-3 on p. 3*). Fewer MDAPs are for truly new systems (as opposed to modifications of existing systems) than before *(see Figure 2-40 on p. 79)*. Contract cycle time has dropped significantly, so the degree to which cycle time reflects complexity also raises concerns of reducing technical superiority *(see detailed discussion starting on p. 70).*

Not all acquisition reform attempts have been beneficial. McNicol and Wu (2014) found that MDAPs started during the reforms of the mid-1990s—which encouraged a more “hands-off” and “let industry do its job” approach and included a significant downsizing of the DoD acquisition workforce that was only partially reversed prior to 2012—produced significantly higher funding cost growth than other regimes (see Figure H-7 above). Other than the reforms of the 1990s, there were no statistically significant differences in program funding cost growth between four other major policy regimes before the current BBP regime, which was not evaluated by McNicol and Wu. *(See data and discussion starting on p. 105)*.

Tight budgets may motivate overly optimistic baselines and higher cost-related growth. Constrained budgetary climates during program baselining correlate with significantly higher PAUC growth—see Figure H-21 and McNicol and Wu (2014). Quantity-adjusted PAUC growth is much more pronounced if the program was initiated during tight budget periods (such as we currently experience). Historically out of 151 MDAPs since FY 1970, 40 showed PAUC growth of at least 50 percent. Thirty-six of those 40 programs (90 percent) were established (i.e., passed MS B or the equivalent) under tight DoD budgets. Additional preliminary analysis indicates that neither acquisition policy regime nor budgetary climates have a strong effect on the eventual cancellation rates of programs started in such periods—see McNicol, Burns and Wu (2015) and Figure 3-3 later in the report. We are in a tight budget period, so I am particularly alert to the higher likelihood of cost growth for programs that the DoD starts in this climate. This is not a time for reduced oversight of programs. *(See data and discussion starting on p. 105)*.

Figure H-21. Effect of Budget Climates on Program Cost-Related Performance

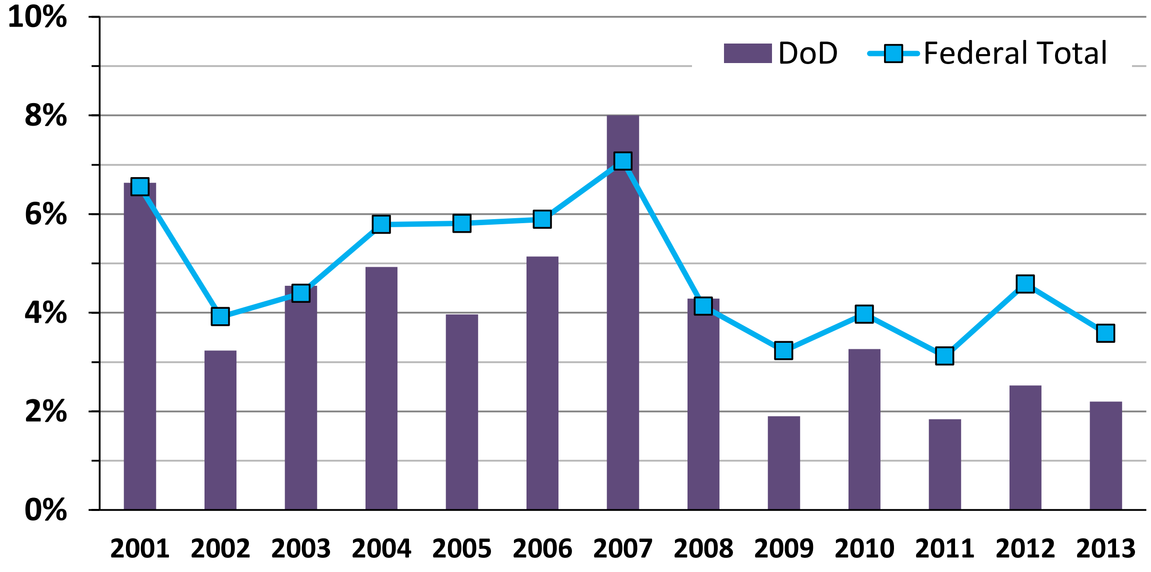


Source: McNicol and Wu (2014).

NOTES: Underlying distributions marked with ovals are statistically lower for the tight budget climate of the same acquisition regime. There were no obliging budget climates for the DAB and AR eras, so comparisons are not available in those two cases. McNicol and Wu adjusted PAUC for any changes in quantity. For consistency with the rest of our report and to better illustrate the central tendency, we report the median values instead of means from the distributions.

Bid-Protest Sustainment Rates. There has been an increase in the total number of bid protests filed with the GAO against the DoD from 2001–2013, ranging from 603 in 2001 to 1,365 in 2013. Despite this increase, the sustainment rate has dropped recently for the DoD and is below the federal total, which includes the DoD (see Figure H-22). Thus, the increased number of protests appears to reflect external Industry strategies or competitive pressures from the declining DoD budgets rather than poor DoD source-selection performance. These results are commensurate with the Congressional Research Service's recent analysis of bid-protest rates *(see Schwartz and Manual, 2015, and additional discussion starting on p. 128)*.

Figure H-22. GAO Protest Sustainment Rates (2001–2013)



Source: GAO

Actionable Insights

Our first three annual reports provide detailed analytic results broken out in different ways to provide each DoD organization and the broader community as complete a view as possible into overall and component performance. While each organization in the defense acquisition system should examine its results to see how it can improve, the following broad insights are provided to help identify specific actions that can ensue from these analyses.

Component-Level Insights

First, let us discuss insights that affect both Component leadership and PMs.

Be Particularly Careful to Ensure Realistic Program Baselines When Budgets Are Tight

A 2014 study by McNicol and Wu suggests that excessive optimism on program baselines initiated during tight budget periods (such as at present) will lead to long-term problems for the Component. Components need to explicitly recognize this and avoid setting up our successors for large overruns. For example, acquisition and Component leadership should ensure adequate risk reduction before MS B and a healthy skepticism about novel approaches that are marketed as offering substantial cost reductions (i.e., if it sounds too good to be true, it probably is). In a tight-budget climate, industry is motivated to be optimistic and take greater risk in order to win new business. DoD programmers are also motivated to put pressure on acquisition professionals to lower cost estimates and funding requirements. Because of these tendencies, I am focusing particularly on cost and schedule realism for Acquisition Category (ACAT) I and MAIS programs at milestone decision reviews. *(See detailed discussion starting on p. 105.)*

Explicitly Discuss and Track Key Framing Assumptions at Major Program Reviews

Our root-cause analyses have found that a quarter of MDAP Nunn-McCurdy breaches are due to faulty key framing assumptions[9] *(see detailed discussion starting on p. 25)*. The new DoD Instruction (DoDI) 5000.02 (AT&L, 2015a) recognizes this and requires explicit discussion of framing assumptions at major reviews so that executives can make informed risk decisions. Moreover, metrics that indicate failing assumptions will be established and tracked so that failures can be recognized and needed adjustments made as soon as possible.

Focus on Streamlined Processes and Setting the Right Schedule—Not on Arbitrary Cycle Time Targets

While reduced cycle time could enable faster response to evolving threats, it could reflect a downward trend in pursuing new, more sophisticated systems. Arbitrary targets for cycle time can do more harm than good (the poster child for this statement is the Army's cancelled Future Combat System [FCS] program). The DoD is working through BBP to streamline our processes so we can be more responsive, using collaborative, operational, acquisition, and intelligence requirements reviews and technical-maturity assessments with appropriate risk-mitigation measures to set achievable schedules. At the same time, we have to be realistic about achievable schedules or else we will set ourselves up for inefficiency and failure.

[9] A *framing assumption* is any supposition (explicit or implicit) that is central in shaping cost, schedule, or performance expectations of an acquisition program. Thus, a framing assumption reflects some risks deemed worth taking. Key framing assumptions can have a major effect on the success of the program as structured and baselined.

Program-Level Insights

Continue To Apply Formulaic Contract Types When Appropriate

Data from last year's (AT&L, 2014, pp. 87–93) indicate that when used appropriately, formulaic incentive contract types (cost-plus-incentive-fee and fixed-price-incentive) motivate cost control about as well as FFP while ensuring that the government ends up getting a good price. There is no singular "formula" that works best for all cases. In light of this, I am releasing updated guidance on contract incentives.

Apply Best Practices in Systems Engineering

Problems in systems engineering during MDAP execution have been one dominant type to date of root-cause of Nunn-McCurdy breaches (see data starting on p. 25). Take advantage of the improved quantitative metrics, approaches, and support available to PMs to monitor system development and proactively address problems as they arise.

Avoid Undefinitized Contract Actions in Development

Analysis in our first report found that from a cost perspective, UCAs can be employed successfully in early production but are concerns for developmental work. Historically, UCAs did not correlate with total cost growth on early production contracts, but they did so on development contracts (especially for ships). Thus, the data indicate that focusing attention on minimizing the use of UCAs in development should be more productive than on early production contracts.

Use Framing Assumptions and Metrics to Identify Problems and Take Prompt Action During Program Execution

Program management and supervision is an *observe, orient, decide, and act* (OODA) loop. About a quarter of MDAP Nunn-McCurdy breaches were caused by a failure to "orient" or understand and accept what the program metrics were indicating and a failure to act on the available information in a timely manner (see the discussion starting on p. 25). Managers need to analyze the available data and execute their OODA loop while sharing data and the implications up the chain-of-command to facilitate early problem identification and mitigation. The good manager raises problems early.

Identify and Manage Risk Drivers

Our annual reports show cost and schedule growth differences between the types of systems acquired. Some of those differences reflect inherent risk differences between systems. For example, major contracts on Command, Control, Communications, and Intelligence (C3I) and unmanned aerial vehicle (UAV) systems for active MDAPs show statistically high price growth, probably due to being software intensive and immature, respectively (see Table 2-7 on p. 52). Thus, these results help to inform areas where risks are prevalent and need particular attention and management.

(This page was left blank intentionally)

ACKNOWLEDGMENTS

The analysis in this report was conducted by Dan Davis, Ken Munson, Susan L. Rose, and Philip S. Antón, with much-appreciated assistance from Larry Axtell, Gary R. Bliss, Karen Cook, D. Mark Husband, V. Bram Lillard (Institute for Defense Analyses), Jeanette S. McKinney (GAO), David L. McNicol (Institute for Defense Analyses), John O'Brien (540 CO, LLC), Chad Ohlandt (RAND), Philip D. Rodgers, Moshe Schwartz (Congressional Research Service), Garry Shafovaloff, Nancy Spruill, René Thomas-Rizzo, Catherine Warner (DOT&E), Ralph O. White (GAO). Philip S. Antón, Dan Davis, Ken Munson, and Susan L. Rose were the primary writers, with very helpful reviews and comments provided by Philip D. Rodgers, Gary R. Bliss, Mildred Bonilla-Lucia, Michael E. Lebrun, and the departments of the Army, Navy, and Air Force. Benjamin Tyree and Michael Shoemaker cheerfully provided valuable editing on a very compressed schedule.

(This page was left blank intentionally)

1. THE ACQUISITION LANDSCAPE AND PERFORMANCE MEASURES

Our acquisition system—its institutions, offices, laboratories, workforce, managers, executives, and industrial partners—conducts research, provides a wide range of services, develops and produces new goods and weapon systems, and sustains these capabilities for warfighters and other operators. The performance of that system is measured relative to its outputs and outcomes of interest. Identifying internal policies, processes, workforce, and management capabilities that bear positively or negatively on those measures requires data and analysis to avoid speculative or cyclical policy choices based on current conventional wisdom and untested hypotheses.

FOLLOW THE MONEY: ANALYSIS OF SPENDING

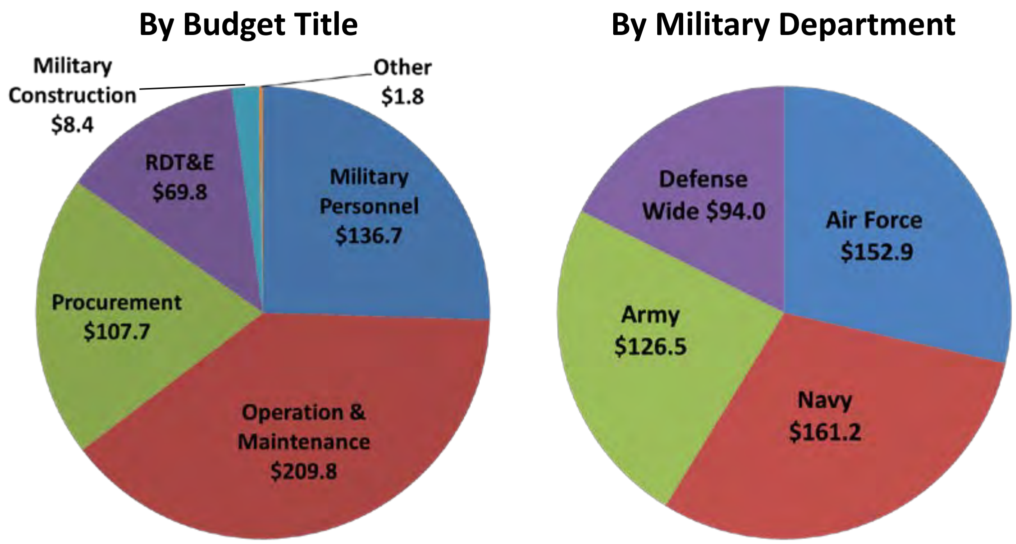
The defense acquisition system acquires goods and services to support our military forces—both now and in the future—while fulfilling our responsibility to prudently use taxpayer dollars. The DoD budgets and accounts for expenditures in various ways, each providing useful perspective on the purpose of the largest expenditures.

Spending by Comptroller Budget Accounts

Broken down by the Comptroller's budget accounts, the base President's budget (PB) for FY 2016 requests $107.7 billion for Procurement and $69.8 billion for RDT&E—see Figure 1-1. Of this $177.5 billion, 43 percent ($77.2 billion) is for programs designated as MDAPs, which provide the bulk of the readily available program data for analysis in this year's report. In addition, the PB 2016 also requests $209.8 billion for Operations and Maintenance (O&M) and $136.7 billion for Military Personnel. A sizable portion of O&M also is spent on contracts for goods and services; thus, this portion is also part of the defense acquisition system. Supplemental funding for Overseas Contingency Operations (OCO) is not included in PB 2016 figures, but a large portion of these funds is used for contracted products and services also.

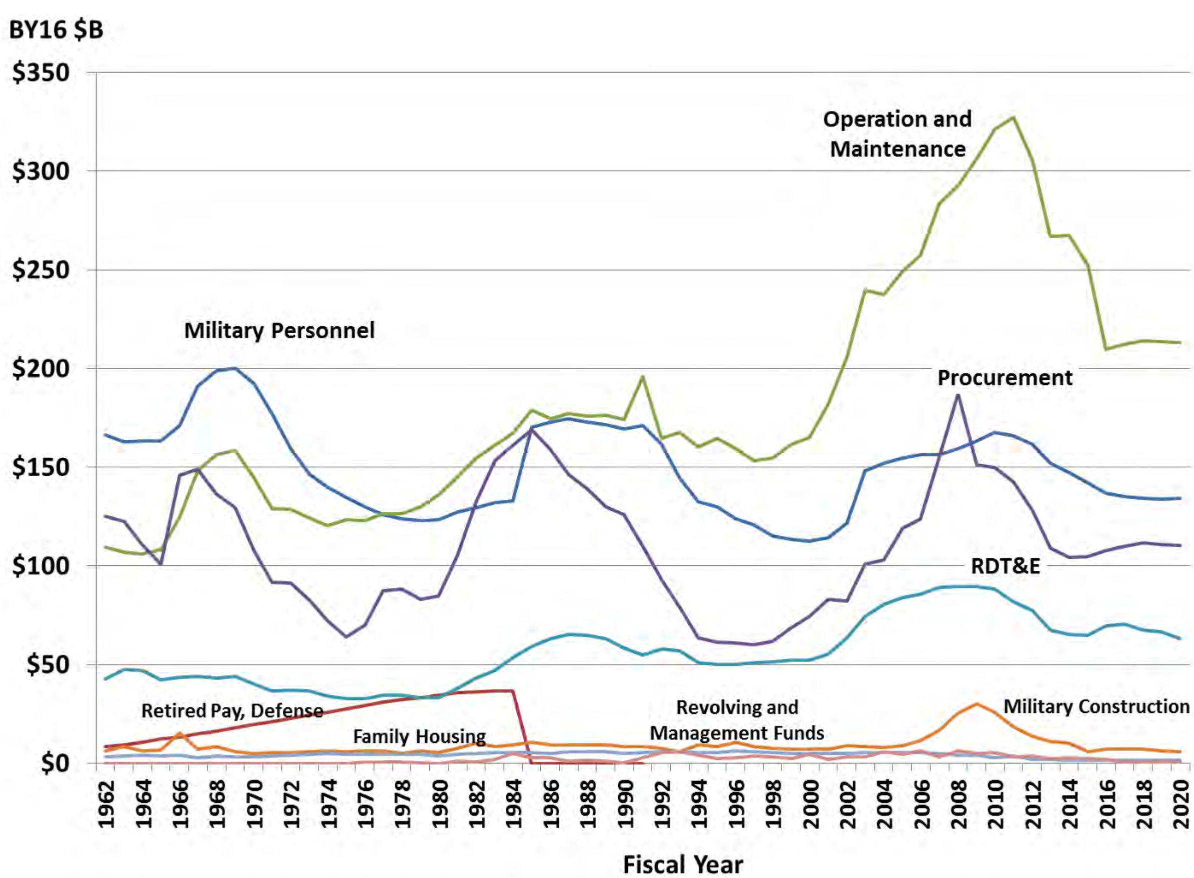
Figure 1-2 shows how defense budget accounts have changed over time and compares these to PB 2016. As reported last year, DoD budgets oscillate in a pattern that repeats about every 24 years (plus inflationary changes and noise). The current budget is on the second half of the falling portion of the general pattern. Future budgets, of course, are hard to predict, but these patterns show some structure in recent budgetary ups and downs.

Figure 1-1. Defense Budget Breakouts PB 2016 (billions of 2016 dollars)



NOTE: OCO dollars are not included.

Figure 1-2. DoD Funding by Budget Accounts: Historical and PB 2016 (FY 1962–FY 2020)

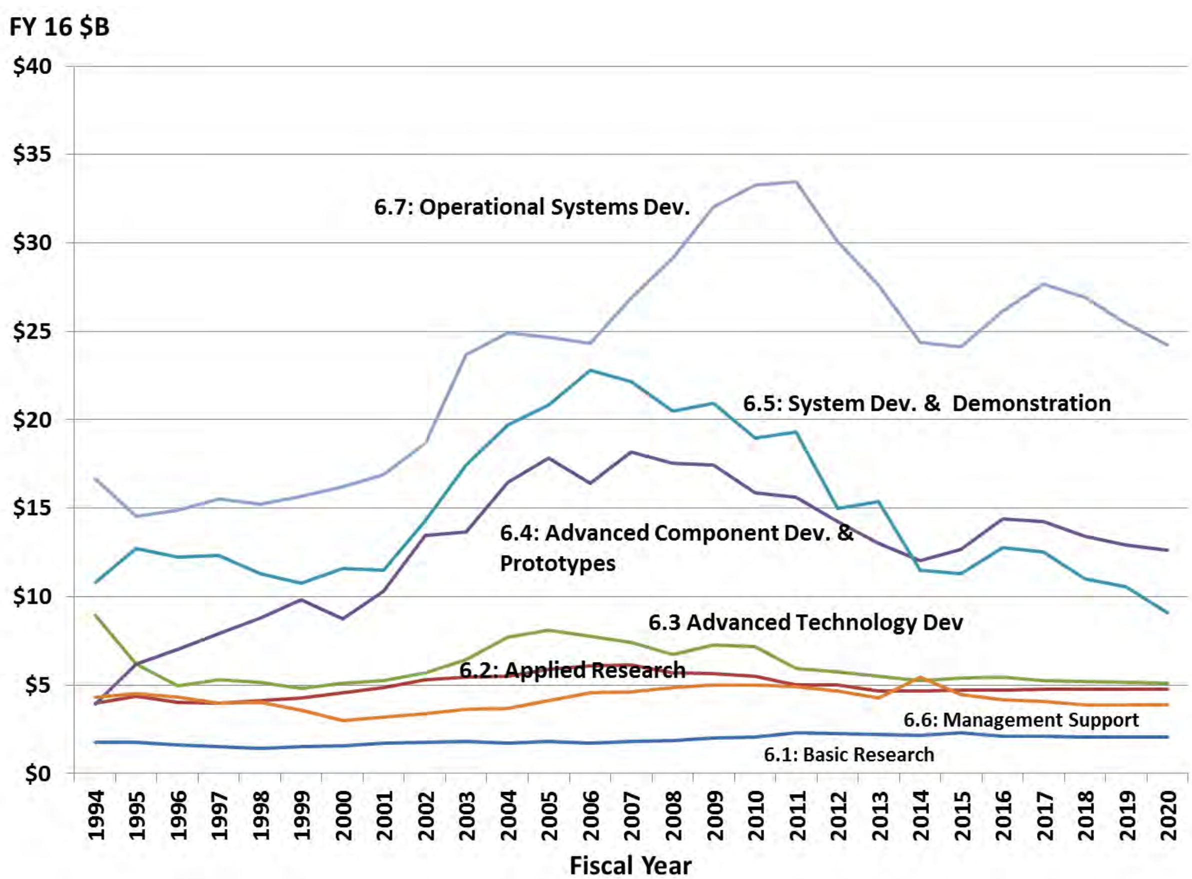


NOTE: OCO is shown in FY actual budgets up to FY 2014 but not in budgets since FY 2015. Budget amounts are adjusted for inflation and reported in billions of calendar year 2016 dollars (CY16$B).

RDT&E Budgets

Figure 1-3 shows the breakdown of DoD RDT&E funding and budgets by budget activity, going further back than in last year's report. This provides an embellished picture of how these accounts have fared historically—and especially during the recent budgetary surge and subsequent decline. Here the science and technology (S&T) accounts (6.1–6.3) are relatively flat or returned to their pre-2001 levels. Accounts for Advanced Component Development and Prototypes (6.4) and Operational Systems Development (6.7, for existing systems) are projected to come down from their peak but remain higher than the levels in the 1990s. The System Development and Demonstration (6.5) budget for new systems in the DoD's product "pipeline" is projected to decline to its lowest level in this time period. While 6.4 and 6.7 levels are not coming down as far, the low levels of 6.5 funding reinforce the DoD's concerns that we risk losing technological superiority in multiple operational domains.

Figure 1-3. Recent and Projected DoD RDT&E Funding as of PB 2016 (FY 1994–FY 2020)



NOTE: OCO is shown in FY budgets up to FY 2014 but not in the current FY 2015 or in PB 2016 figures (FY 2016–2002). Budget amounts are adjusted for inflation and reported in billions of FY 2016 dollars (FY16$B).

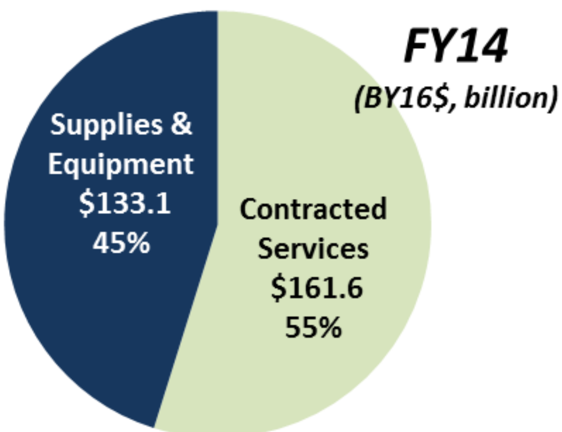
Contractual Spending by Product Service Code Portfolios

Much of what we acquire comes through contracts with industry. Thus, a different way to understand what the DoD acquires is to examine contract obligations by type rather than budget account.

The contracting community uses a categorization called *product service codes* (PSCs) to track what is procured under federal contracts.[10] The Federal Procurement Data System—Next Generation (FPDS-NG) records PSCs for every contract obligation worth at least $3,000, so this taxonomy affords us a way to quickly look across all DoD external (contracted) spending.

At the top level, spending (obligations in this case) is split between *products* (also referred to as *supplies and equipment*) and *contracted services*.[11] Figure 1-4 shows that in FY 2014, just over half (55 percent) of contract obligations were for contracted services. Some caution is warranted, however. While the acquisition community generally considers RDT&E as part of developing a physical system, contract PSCs identify research and development (R&D) as a service (i.e., it is dominated by tasks that do not produce physical end items of supply). Also contract obligations often include multiple types of work, but only one PSC is reported per obligation.

Figure 1-4. Total DoD Contract Obligations Split Between Goods and Services (FY 2014)

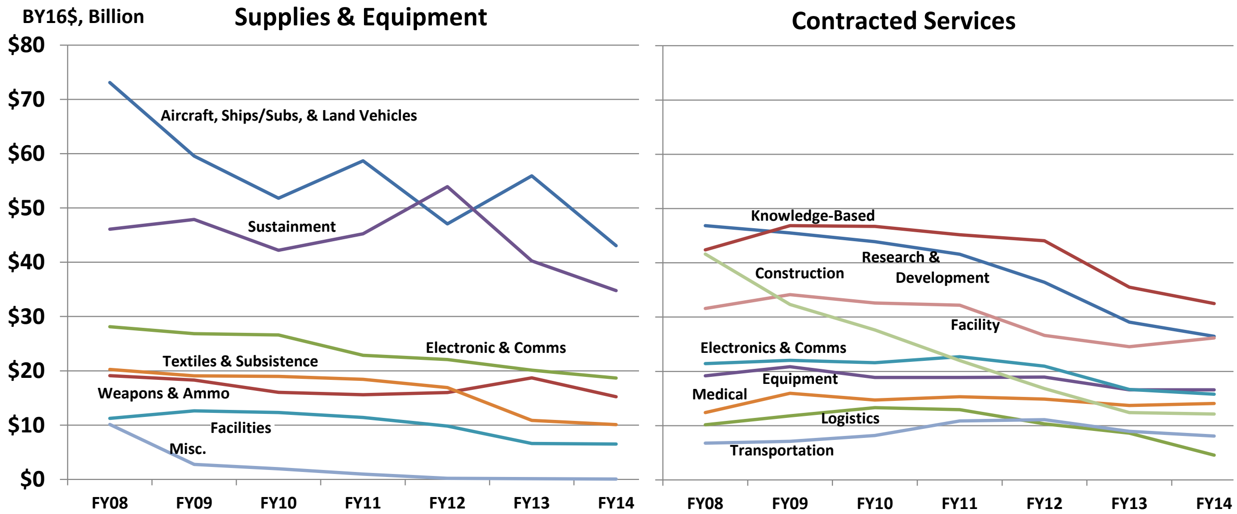


[10]See the *Product and Service Codes Manual* published by the U.S. General Services Administration (2011). PSCs are recorded in the FPDS-NG to categorize what each federal contract acquires.

[11]The Federal Acquisition Regulation defines a *service contract* as "a contract that directly engages the time and effort of a contractor whose primary purpose is to perform an identifiable task rather than to furnish an end item of supply" (see Federal Acquisition Regulation [FAR], Section 37.101). Because the DoD often refers to the military departments (i.e., Army, Navy, and Air Force) as "Services," this report capitalizes "Services" when referring to military departments but uses lower-case "services" when referring to contracted services.

Figure 1-5 shows a further breakdown of all DoD contract obligations by groupings developed to aggregate PSCs into meaningful major portfolios. Here we see some contracting portfolios have remained relatively flat over the years while others are declining with the recent budget cutbacks.

Figure 1-5. Total DoD Contract Obligations by Portfolio Group (FY 2008–FY 2014)



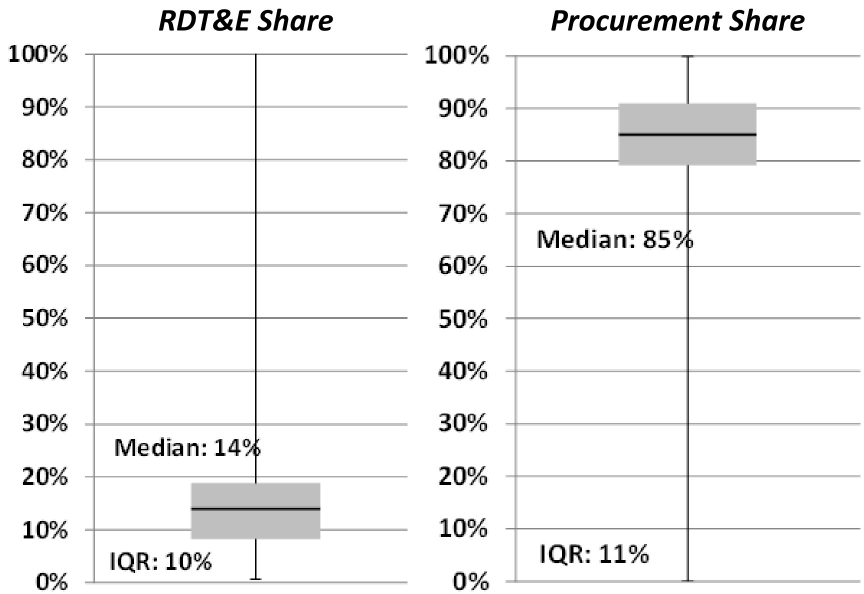
NOTE: FPDS-NG data on all DoD obligations (including OCO). Obligations for contracted services in this period ranged from 57 percent in FY 2010 to 55 percent in FY 2014. All numbers are in billions of adjusted base-year (BY) 2016 dollars.

Proportion of Major Program Spending Between Development and Production

It is important to note that relatively speaking, expenditures on development are usually much smaller than those on production for MDAPs. Thus, although attention on cost growth often focuses on the development phase of a program, the procurement phase is where the big money is. We, of course, have to get the system right in development, and that may mean spending more than planned in development to reap the benefits in production, operational support, and ultimately on the battlefield.

To gain some perspective on exactly how big the difference is, analysis of 76 active MDAPs is shown in Figure 1-6. Here we control for both program maturity and program spending (i.e., shares by phase are weighted by program funding). At the median, the procurement share is more than six times larger than the RDT&E share. Thus, while it is important to monitor performance in RDT&E, the larger cost effects are usually in procurement and sustainment. Some cost growth in RDT&E may result in better operational performance as we adjust to changing threats, and some RDT&E costs may be investments that will lower total life-cycle costs. Separating these beneficial causes of RDT&E cost growth, however, is very difficult when looking across the entire DoD acquisition portfolio of MDAPs.

Figure 1-6. Relative Shares of RDT&E and Procurement in MDAP Funding (2000–2015)



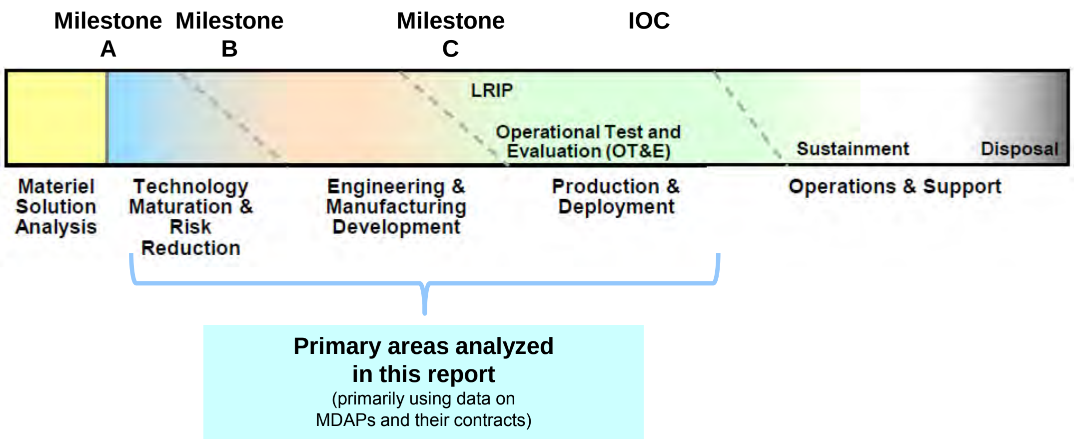
NOTE: N=76 MDAPs. Weighted by program funding.

Phases of Acquisition Assessed

This report assesses how our institutions perform—primarily using existing oversight data aggregated to look for broader performance and trends. Because of their size and the risks involved, most readily available data are on MDAPs and their measurable outcomes rather than on smaller programs and the full breadth of contracted services. Still, these data provide partial insights on the acquisition of both goods (i.e., production of the weapon systems themselves) and services (i.e., development and testing of those weapon systems)—albeit primarily on major weapon systems.

Figure 1-7 depicts a simplified program life cycle and the portion where we currently have the best data for analysis—namely, for development and production up to full operational capability (FOC). While we have some data that reflect partially on the performance in other phases (e.g., early research, analysis of alternatives [AoAs], early risk reduction, and sustainment), operation and support are reflected at best by early estimates. These other phases and uses of funds will be expanded in subsequent versions of this report as we improve data access, quality, and availability.

Figure 1-7. Program Phases Reflected in Report Data

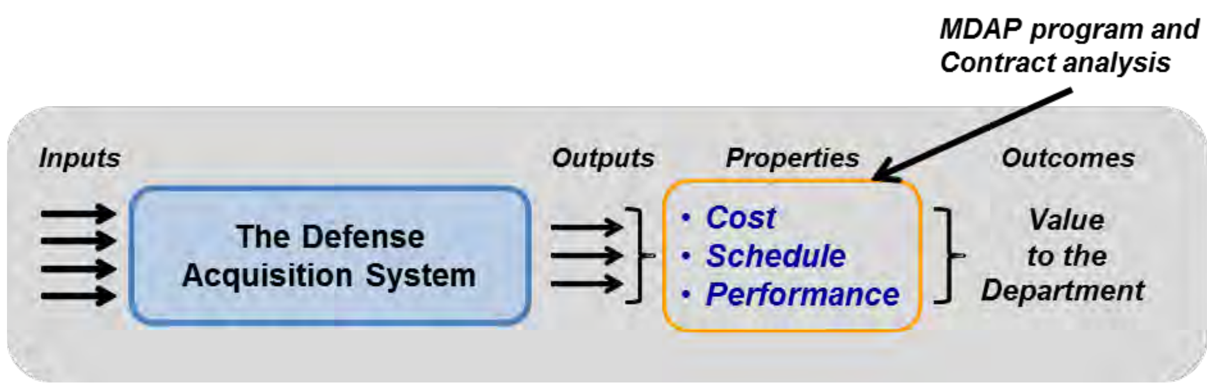


IOC = Initial Operational Capability
NOTE: This figure abstracts key elements from the example program models in the DoD Instruction 5000.02 (AT&L, 2015a).

On Measuring Performance

Institutional performance is all about acquiring value as efficiently (cheaply) as possible. Value to the DoD stems from the immediate benefits (i.e., technical performance) of the goods and services acquired in a responsive time (schedule) compared to the costs to the taxpayer. Hence, measures of cost, schedule, and performance serve as the basis for measuring the effectiveness of the acquisition system in converting inputs to outputs (see Figure 1-8). The subsequent operational benefits of the acquired capabilities to address threats and fill capability gaps is more difficult to measure objectively with available data, and we will continue to explore ways to efficiently obtain such data on outcomes.

Figure 1-8. Output Measures for the Defense Acquisition System



Understanding How Internal Functions and Processes Affect Performance

The acquisition system can be measured at two fundamental levels: (1) the major outputs and outcomes of the system, and (2) the key functions, responsible entities, and institutions accountable within the system to achieve those outputs and outcomes. The most readily available and measurable outcomes assessed throughout the report are cost and schedule growth, but some readily available information on technical performance is analyzed also.

Decomposing the acquisition system into major functional responsibilities enables analysis of how elements of the system affect the ultimate outcomes. Intermediate outputs and outcomes of key institutional functions may correlate with cost, schedule, and performance outcomes, but others may be too small or difficult to discern from available data. Nevertheless, a functional decomposition facilitates an understanding of how well the defense acquisition system performs, based on management principles and intermediary outputs and outcomes. As this work moves forward, our greatest challenge remains identifying the relationships between and among factors the DoD can affect (policies, contract terms, incentives, workforce skills, etc.) and the outcomes sought. This report is a continuing step in that process.

Much of our analysis is statistical, focusing on institutional outcomes and their trends, rather than on single acquisitions and outliers (see Appendix A for a detailed discussion of the statistical methodologies). The objective is to see how well we are doing, learn from these generalities, and change our policies and tradecraft as we seek to improve outcomes. Many of the results continue leveraging readily available data on collections of programs and contracts and examining them from different groupings and perspectives. We continue looking for statistically significant differences on samples large enough to avoid an overgeneralization from case studies.

Scope of Outcomes: Programs or Their Constituent Contracts

Our analyses often examine two main types of performance data:

- **Program-level Data**—describing measurements across the entire program (e.g., growth in Planned Total Funding from MS B baseline as reported in the SARs and MARs, including past actual funding, current funding requests, planned funding in the Future-Years Defense Program [FYDP], and estimated needed funding beyond the FYDP to the end of the program). Data sources include the SARs, MARs, DOT&E's BLRIP reports, oversight data, the Contractor Performance Assessment Reporting System (CPARS), and review documentation.
- **Contract-level Data**—describing measurements on one of the many contracts that constitute a program (e.g., the total cost growth from original negotiated contract target cost for an early lot of units procured). Data sources include EV Central Repository, FPDS-NG, GAO bid-protest data, and cost data reports.

Program-level measures show how well the acquisition system developed the ability to produce the overall program against original baselines despite quantity changes, while providing insight into whether cost growth may have been a factor in quantity changes.

Contract-level measures provide early indicators of potential program-level issues by examining performance when the DoD contracts for specific work from industry. Nearly all the actual research, development, and production on weapon systems are performed by industry partners through contracts with the DoD. Thus, examining performance at the contract level provides detailed and potentially useful indicators of performance that eventually will be seen at the more aggregate program level.

This report often switches between these types of data as we examine different types of institutions (e.g., DoD-wide or military departments) and different phases of acquisition (e.g., development or early production).

While contracts are the key execution elements of a program (i.e., most goods and even services are provided by contractors), they have different baselines (e.g., contract cost targets) set at different times than the program's MS B baseline. Performance on individual contracts can be measured earlier than their effects might show up in program-level measures. However, because there are often numerous contracts within a program, and program baselines are not equivalent to contract cost targets, an individual contract performance may not necessarily reflect the performance revealed in program-level measurements. ***Thus, it is important to recognize what type of data is discussed at each point in the report.***

Also, care must be taken to note whether cost data have been adjusted for inflation. Often the available program-level budget data we used have been adjusted for inflation (i.e., reported in "base-year" [BY] dollars), but some contract-level cost-growth data have not been adjusted (i.e., are reported only in "then-year" [TY] dollars, and insufficient temporal information was available for us to adjust the reported figures for inflation). Thus, partly because of inflation, the program-level cost-growth figures in this report may be lower than those for some contract-level analyses.

Avoiding Maturity Bias in Testing for Recent Improvements

A primary reason for systematically measuring our performance is to determine objectively whether we are improving (i.e., whether our efforts are helped by recent policy and processes changes such as WSARA of 2009, three iterations of BBP, major efficiency drives, and continued investments in the acquisition workforce training and hiring). By their nature, recent programs and contracts have less cost and schedule growth because they are newer and have not had time to realize any growth. Unfortunately, waiting until they are complete will take many years—sometimes decades. This is but one challenge in comparing performance (other variables in the portfolio include shifting commodity mixes, budgetary changes, multiple parallel policy and process changes, etc.) and is a key analytic concern.

Rather than wait for the completion of programs and contracts before measuring their performance, we take the middle ground of controlling for immature programs in many of our analyses. The cost community generally has found that programs and contracts with large cost or schedule growth will begin seeing and reflecting it in their estimates by the time they have executed about 30 percent of their originally planned schedule. Thus, analyses in this report that control for maturity exclude newer programs and contracts that have not yet reached this point. This, of course, is not the final word, but it does allow us to reflect much of the anticipated performance problems and get a reasonable sense of recent performance.

Additional methods include examining incremental (marginal) growth rather than just total growth since inception. Our program-level analyses, for example, examine biennial change as a way of seeing if growth is added (or removed) on top of original estimates. If recent programs or contracts are worsening, we should be able to see that in the marginal change data.

Measuring Performance on Contracts

Price, schedule, and technical performance are key contract outcomes of interest. Ultimately, the *cost* to the contractor of providing a good or service relates in various ways to the *price* paid by the government. Thus, we often examine cost, price, or both (when possible).

Some datasets in this report contain cost, price, and schedule data along with profit or fee (expressed as margin or markup), which allows us to analyze incentive effectiveness. Generally, we were able to adjust these cost data for inflation and thus present cost and price growth in real terms.

In most cases and where noted, contract price and cost data are adjusted for inflation (except, for example, those plots where we show total contract cost growth and outlier analyses in Chapter 2). This allows us to distinguish real price and cost growth performance independent of inflationary effects.

Analysis of Work Content Growth and Cost-Over-Target

In other datasets, we do not have profit or fee data but can break down total cost growth into two broad elements of *work-content growth* and *cost-over-target*. Work-content growth is simply the change in the contract budget base (CBB), which reflects the contract target cost since contract initiation. Cost-over-target is the latest PM estimate-at-completion (EAC) minus the latest CBB, all divided by the original CBB. Unless otherwise indicated, all these contract cost data are reported in TY dollars and are thus *not* adjusted for inflation.

Note that contract targets themselves are not pure, unbiased estimates of project cost. In addition to expected project scale (size) they also reflect other contract terms, such as the share lines, incentive terms, as well as market contestability and the general negotiating environment. Thus, in part they are the result of bidding strategies.

2. Acquisition System Outcomes and Trends

A key to improving acquisition is learning from our successes and failures. Without looking at results of past actions, we have no feedback to inform whether our hypotheses and beliefs pan out in the complicated world of defense acquisition. Objectively examining the relative effectiveness of acquisition Components and institutions while attempting to distinguish which factors and variables affect outcomes not only allows us to identify successes and failures, but also begins to lead us to specific lessons we can try to replicate—and control points we can exploit.

The following analyses examine key outcomes of cost, schedule, and technical performance of MDAPs across the DoD and by Components, commodities, and prime contractors—measured at program and contract levels. Combined, these analyses provide insight into potential cause-and-effect relationships, focusing attention on problems as early as possible, clarifying misunderstandings, and informing assessments and learning.

For our analyses of program data, note that the MDAPs examined are in a varying state of maturity—from early programs that may or may not develop future problems, to mature programs adding new capabilities to existing systems, to completed programs.

For our analyses of contract data, note that each MDAP may have more than one major contract in our datasets. Major MDAP contracts are generally those for which we have earned value (EV) data. This includes the six largest MDAP contracts (prime, associated, or for government-furnished equipment) valued at over $40 million and which are usually not firm-fixed price (FFP) contracts. This also may include other MDAP contracts of at least $60 million in RDT&E or $250 million in procurement or ship construction (in FY 1990 constant dollars). These contracts may be for development or production. Unless noted, all datasets consist of all readily available cases rather than a statistical sampling from a larger set. In some analyses when we are trying to ascertain the general tendency of the population, we remove statistical outliers using standard tests; these instances are noted. Otherwise, we include all data, including outliers. Also, we often report medians because this is a better measure of central tendency for skewed distributions than arithmetic means, which exaggerate the effect of outliers.[12]

[12] Part of the skewing in the distribution of cost change is the mathematical boundary of cost change because cost cannot decrease more than 100 percent, but it can increase more than 100 percent.

COST-RELATED FUNDING GROWTH AND DEFENSE ACQUISITION EXECUTIVES

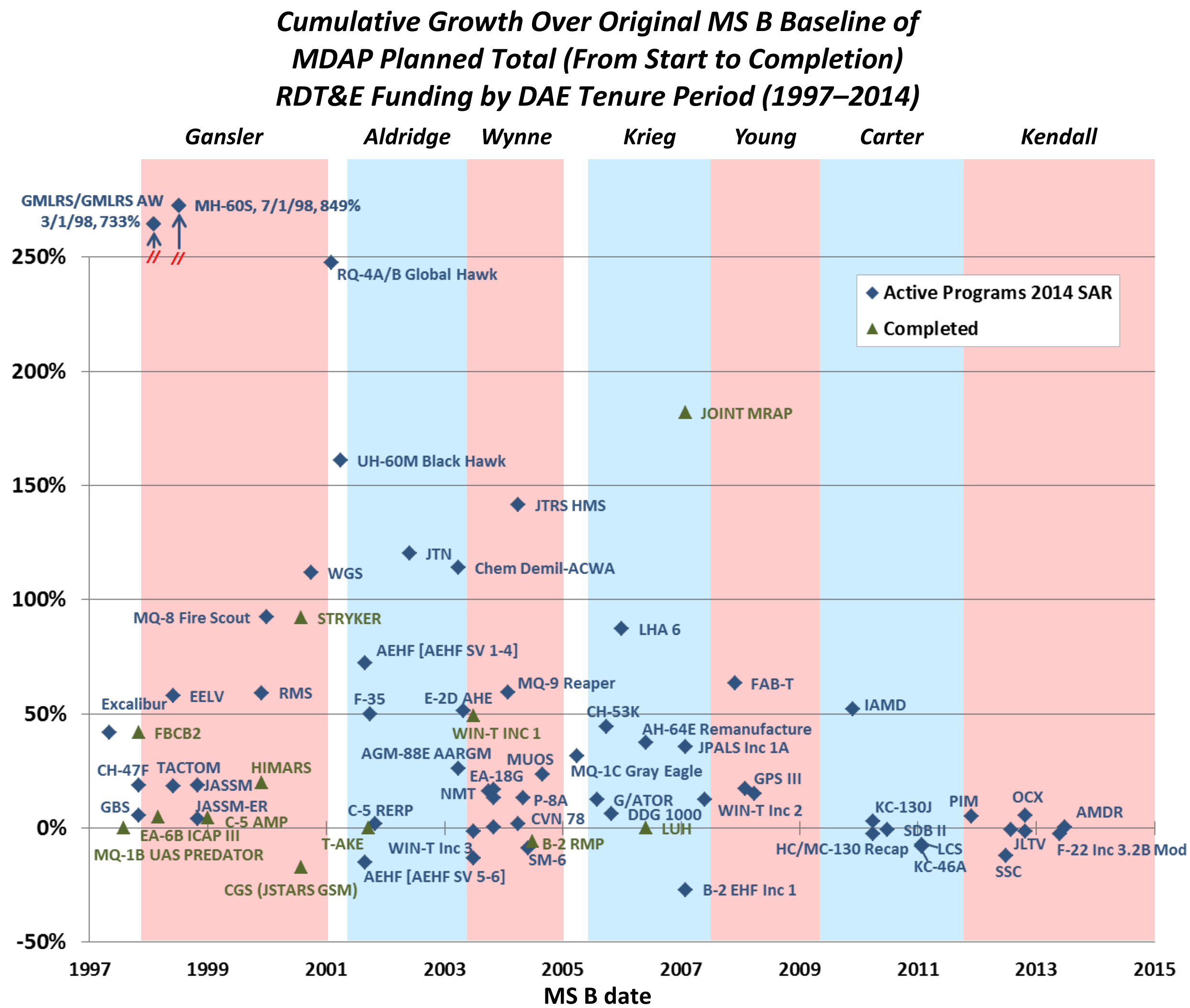
Policy and execution decisions by DoD executives should bear (in part) on the effectiveness of the overall acquisition system during their tenures. This is particularly true for the program structure and associated baselines set at Milestone (MS) B, against which future cost performance is measured. Such decisions include changes to the defense acquisition system policies and procedures (e.g., through changes in departmental regulations); approvals, certifications, and exemptions within that system; institutional organization, policies, and processes; incentives; personnel selection, training, and mentoring; guidance and execution on larger programs, including acquisition strategies and choices; and myriad other effects. More specifically, the acquisition executives chair the boards that review programs at major milestones, guiding both program directions and specific approaches to contracting. Therefore, in our annual reports we track the performance of programs started under different acquisition executives to help reinforce accountability and provide an initial look for possible trends for further analysis.

Figure 2-1 and Figure 2-2 show growth in planned total funding against original baselines for development and procurement as reported to Congress in the SARs on active and completed MDAPs. Note that SAR funding data reflect what the PM currently estimates will be needed in total by the end of the program for the current program configuration, including past actual funding, the current budget request, planned funding in the Future-Years Defense Program [FYDP], and planned funding beyond the FYDP to the end of the program. Growth is measured against the baseline set at the original MS B and can be positive or negative.

These figures also show the Defense Acquisition Executive (DAE) at the time of the MDAP's MS B approval. Later in the report we show similar charts for the programs started under different Service Acquisition Executives (SAEs) in the three military departments. This year we used total needed program funding instead of contract cost growth for these charts since needed funding is measured directly against the MS B baseline set by the DAE (see the discussion starting on p. 85).

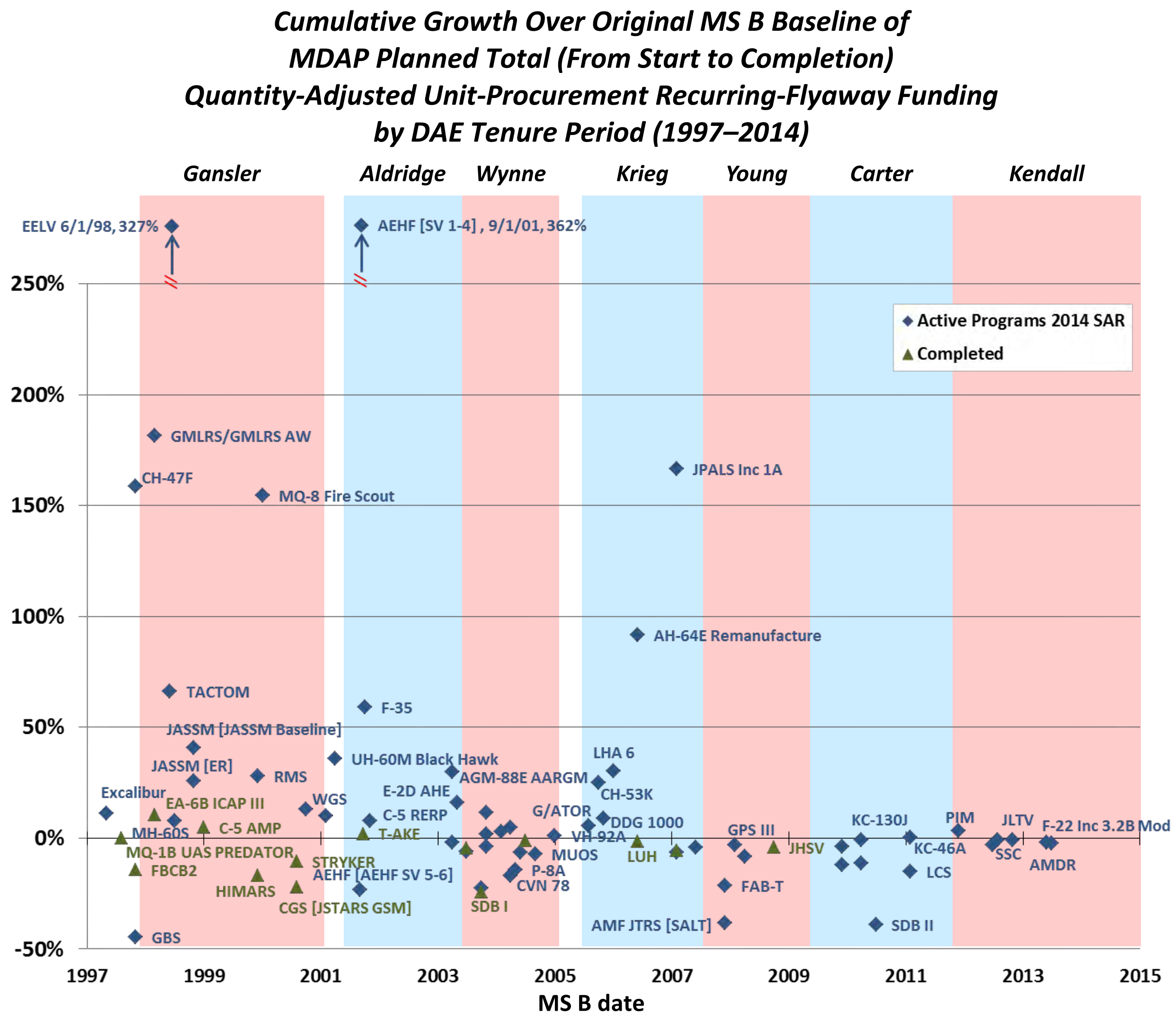
Caution is warranted, however. These charts neither reflect the effectiveness of subsequent oversight or major program changes by later DAEs during execution oversight, nor do they report statistical analysis that controls for other internal and external variables that could have led to program success or problems. Also, as we mention above, each measure has its strengths and weaknesses, so attributing performance to a single measure is subject to the limitations of that measure. For example, some programs may appear to be performing well in terms of total needed RDT&E funding but may be having problems reflected in other measures (e.g., total needed procurement funding, estimated operational costs, and cost growth on one of the program's major contracts). Thus, a combined examination of available data is important before reaching conclusions.

Figure 2-1. Program Cost-Related Development Performance Baselined in DAE Periods



NOTE: This shows total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. These are percentage changes after adjusting for inflation from original MS B baseline of actual past and estimated future funding as reported in each program's latest SAR. Total RDT&E is an insightful measure because it is necessary regardless of quantity. White bars between DAE shaded regions represent periods with no confirmed executive. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown.

Figure 2-2. Program Cost-Related Procurement Performance Baselined in DAE Periods



NOTE: This shows growth in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. White bars between DAE shaded regions represent periods with no confirmed executive. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown.

TECHNICAL PERFORMANCE OF MDAPS

While most of this report discusses outcome measures of cost and schedule, this section summarizes some readily available independent assessments of technical performance of weapon systems.

One measure of technical performance of acquisition programs is how they rate, as a group, in operational effectiveness and suitability as assessed by DOT&E.[13] Operational *effectiveness* is defined in the *Joint Capabilities Integration and Development System* (*JCIDS) Manual* as: "Measure of the overall ability of a system to accomplish a mission when used by representative personnel in the environment planned or expected for operational employment of the system considering organization, doctrine, tactics, supportability, survivability, vulnerability, and threat." Operational *suitability* is a composite evaluation that considers a system's safety, interoperability, availability, maintainability, and reliability. Operational effectiveness and suitability are not measured solely on the basis of system technical performance parameters (e.g., Key Performance Parameters [KPPs] and Key System Attributes [KSAs]). Rather, measurements are accomplished through an evaluation that includes the system under test and all interrelated systems (including weapons, sensors, command and control, and platforms) needed to accomplish a combat mission in expected environments.

Robust developmental testing occurs throughout the earlier phases of a program's life cycle, intended to provide feedback to designers to verify performance and to discover and correct issues so that, by the time operational testing is done on production representative test articles, discovery of major performance issues should be rare.

New "Mixed Results" ratings. DOT&E has been working to improve this summary data by adding a "mixed results" rating in addition to "yes" and "no" ratings. This provides an enhanced view of the technical performance of MDAPs, so this year's report now reflects these three ratings but precludes direct comparisons to the ratings in last year's report. In some cases results were not determined; those cases were excluded from each sample. In addition to any new program evaluations, the number of programs in each sample also varies from last year because for consistency we are now restricting our samples to the BLRIP summary from DOT&E.

Causes. Also new to this year, DOT&E has begun citing the fundamental causes for programs failing to be rated as fully suitable operationally. Some causes are classified, but Table 2-1 shows the categories and unclassified number of occurrences so far. The sample is too small for a meaningful breakout by Component, but this table gives some sense of the more dominant causes cited to date.

---

[13] DOT&E is independent statutorily from the acquisition organizations and is responsible for, among other things, reporting the operational test results for all MDAPs to the Secretary of Defense, USD(AT&L), Service Secretaries, and Congress.

Table 2-1. Causes of Failures to Pass DOT&E Testing (1984–2014)

Cause	Program Count	Description
Part Quality (random failures)	2	Part failing to perform its intended function before its expected "end-of-life" limit is reached (random failures)
Inadequate Design Margins *	23	Failures from engineering requirements, inadequate design (e.g., tolerance stack-up), unanticipated logic conditions (sneak paths), inadequate design margins for the environment, etc.
Manufacturing Anomalies *	1	Failures not related to inherent part reliability but which result from anomalies in the manufacturing process
System Management	26	*Requirements management*: - incorrect or insufficient implementation or interpretation of requirements, processes or procedures; - imposition of "bad" requirements (e.g., missing, inadequate, ambiguous or conflicting); OR *Interface and environment management*: - failure to provide the resources required to design and build a robust system
Wear out	0	Wear-out-related failure mechanisms due to basic device physics
No defect	0	Reported failures that cannot be reproduced upon further testing. These may or may not be an actual failure; however, they are removals and, therefore, count toward the logistic failure rate
Induced Human Factors *	3	Resulting from an externally applied stress. Examples are electrical overstress and maintenance-induced failures (i.e., dropping, bending pins, Human Factors, etc.). Can be design engineering requirements and design related.
Software Fault *	13	Failures of a system to perform its intended function due to the manifestation of a software fault caused by inadequate engineering requirements or design.

* Requirements-related cause

Source: DOT&E BLRIP reports.

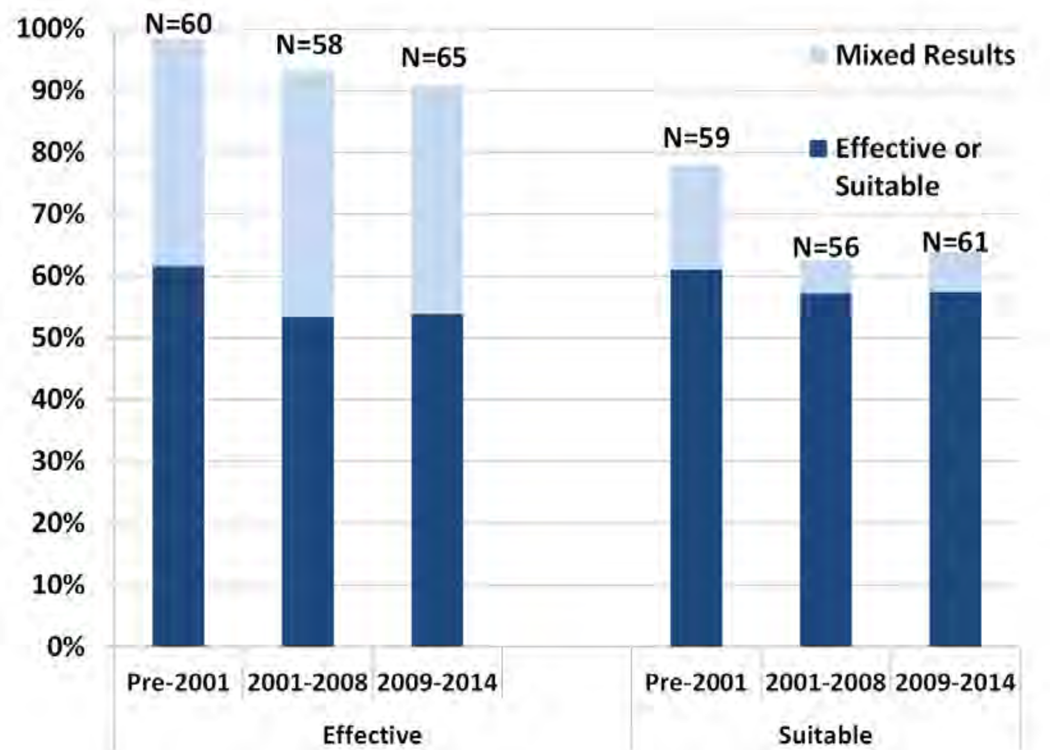
NOTE: These are unclassified results so far but do not include determinations for all programs since 1984. Not all types of causes in each category are exhibited by the programs.

Apparent Trends in Operational Testing of MDAP Effectiveness and Suitability

The following figure shows apparent trends in DOT&E's assessments of technical performance of weapon systems across the DoD. Given the relative infrequency of MDAPs coming up for BLRIP operational testing, we grouped ratings into three sets of periods: 1984–2001, 2001–2008, and 2009–2014. Generally, it appears that performance dropped since 2001 compared to prior years, especially in the number of systems rated fully effective and the number of partially suitable systems, but these changes are not statistically significant.

Figure 2-3. Program Technical Performance

MDAP BLRIP Operational Test Ratings—DoD Wide (1984–2014)



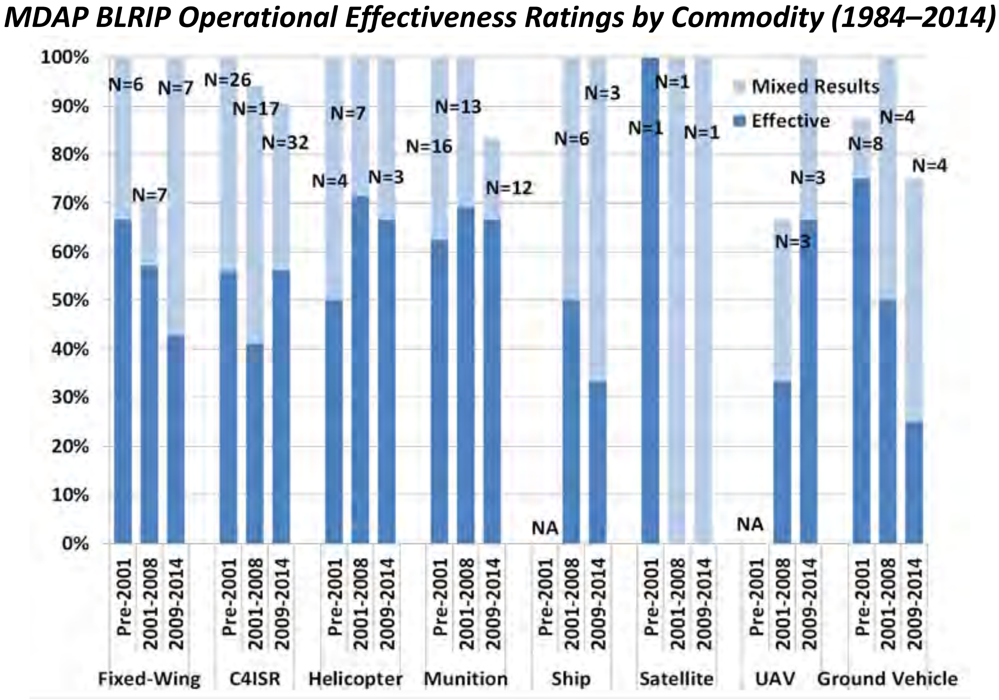
Source: DOT&E BLRIP reports.
NOTE: Differences are only apparent and may not be significant due to the low sample sizes (infrequent evaluations). Sample sizes differ between Effective and Suitable for some Components because effectiveness and suitability could not be determined in all cases.

Operational Testing of MDAP Effectiveness and Suitability by Commodity

The following figures show DOT&E's assessments of technical performance of weapon systems by commodity types. Figure 2-4 and Figure 2-5 show effectiveness and suitability ratings, respectively, for BLRIP operational tests by commodity type. While there are some differences, they are not statistically significant given the very low sample sizes.

Over all time periods, while satellite systems are at least all rated as partially effective and partially suitable, they have the lowest ratings for being completely effective and suitable (albeit with a very low sample size). Most other components (except UAVs) tend to have similar non-mixed ratings.

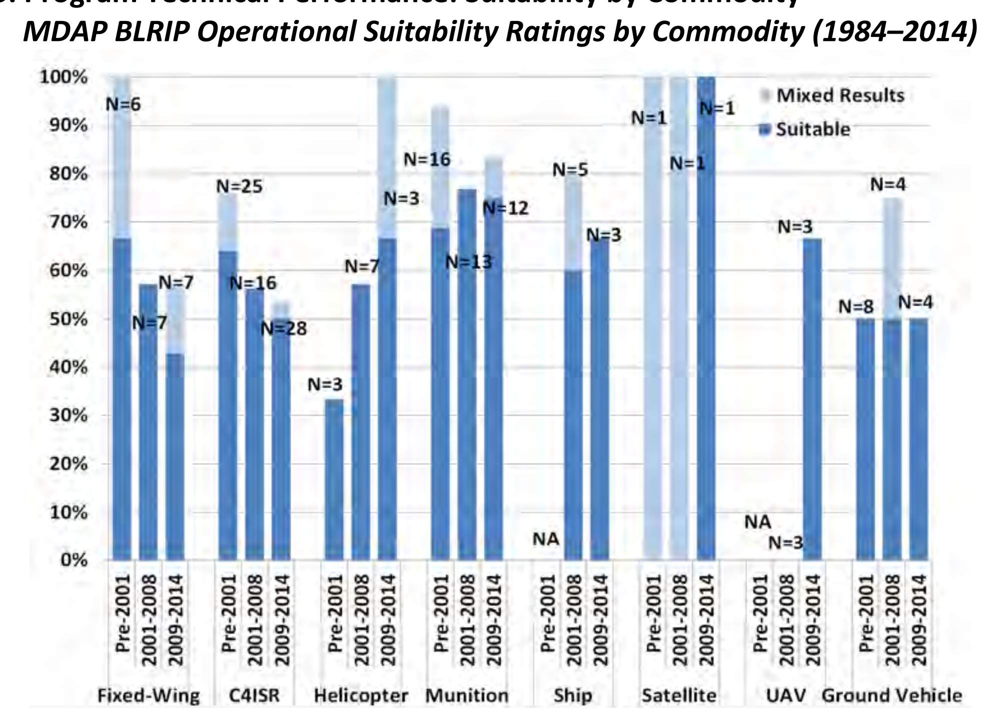
Figure 2-4. Program Technical Performance: Effectiveness by Commodity



Source: DOT&E BLRIP reports.

NOTE: Differences are only apparent and not statistically significant. Sample sizes differ between Effective and Suitable for some Components because effectiveness and suitability could not be determined in all cases.

Figure 2-5. Program Technical Performance: Suitability by Commodity



Source: DOT&E BLRIP reports.

NOTE: Differences are only apparent and not statistically significant. Sample sizes differ between Effective and Suitable for some Components because effectiveness and suitability could not be determined in all cases.

COST AND SCHEDULE PERFORMANCE: OVERALL

Nunn-McCurdy Program Breaches

Each MDAP is required by law to submit a SAR to Congress 45 days after the President's annual budget submission and under various other circumstances (see 10 U.S.C., section 2432). A SAR reflects what is included in the PB as well as a comprehensive summary of MDAP cost, schedule, and technical performance measures. Historical SAR data serve as the primary sources for much of the program-level analysis in the report due to their relative availability and comprehensiveness.

Common program cost measures[14] such as PAUC[15], which includes both RDT&E and procurement, and Average Procurement Unit Cost[16] (APUC), which includes only procurement, are codified in statute. The statute also requires that programs exceeding certain thresholds (measured by PAUC or APUC changes relative to their original and current program baselines) must go through a rigorous reexamination and certification to Congress along a variety of specified criteria. This process commonly is referred to as the "Nunn-McCurdy" process, named for the original sponsors of the legislation dating back to 1982 (see *Schwartz*, 2010, for an extensive overview of the process).

Two types of breaches are called out in the Nunn-McCurdy process: *significant* and *critical*. A significant breach is the lower threshold and is intended to warn Congress that a program is experiencing significant unit-cost growth relative to its baseline. A critical breach signifies the cost growth is even higher, triggering the formal reexamination and certification process mentioned above. The criteria for a significant breach are 15 percent from the current baseline reported in the previous SAR, or 30 percent cost growth in APUC or PAUC from the original baseline. A critical breach occurs when the program experiences 25 percent cost growth from the current baseline, or 50 percent cost growth from the original baseline. Normally, the event of breaching (crossing a threshold) is counted once for each threshold as a means for measuring program performance, even though technically we track a program that is in breach throughout the remainder of its life.

[14] Here, "cost" is synonymous with the total amount of funding because it reflects the prices paid on contracts as well as program execution costs. Later, when we discuss contracts, we will distinguish contract prices from their underlying contractor costs and margins (profits and fees).

[15] Section 2432(a)(1), Title 10, U.S.C. defines PAUC as "the amount equal to (A) the total cost for development and procurement of, and system-specific military construction for, the acquisition program, divided by (B) the number of fully configured end items to be produced for the acquisition program."

[16] Section 2432(a)(2), Title 10, U.S.C. defines procurement unit cost as "the amount equal to (A) the total of all funds programmed to be available for obligation for procurement for the program, divided by (B) the number of fully configured end items to be procured."

This year we updated this summary based on our new official list of breaches since 1997 (see Table 2-2) that resolves some historical data questions by examining any post-SAR updates of cost growth and the official breach notifications to Congress as the deciding factor. Thus, the numbers of breaches are slightly different than in last year's report (although these adjustments were relatively modest). As discussed in previous reports, it is important to note that the National Defense Authorization Act (NDAA) for FY 2006 made changes to the Nunn-McCurdy statute by adding the requirement to report unit-cost growth from the original baseline rather than the current (possibly revised) baseline. This additional requirement caused a large spike in 2005 when 11 programs had to report preexisting significant breaches. Thus, for historical comparisons, we need to compare current performance against those since 2006 because 2005 is a boundary condition and the years before 2005 were operating under different rules.

Figure 2-6 shows the critical Nunn-McCurdy yearly breach rates and counts. There are statistically significant downward trends since 2009 of nonquantity-related critical breaches (shown) and for all critical breaches. For the trend analysis, we used the breach rates instead of counts to control for changes in portfolio size between years. We also controlled for program maturity and the recency of past breaches. Immature MDAPs that had not yet executed 30 percent of their originally scheduled time between MS B and MS C were removed from the data set since they would likely bias the breach numbers to be artificially low. Also, any MDAPs that had a critical breach within the past three years were removed since they are immature relative to their new baselines, and we have yet to see a program breach a second time before this period. Thus, these adjustments provide a more conservative (harder) test and higher percentages by removing the bias from programs that are naturally less likely to breach by nature of their newness.

Table 2-2. Official DoD List of Nunn-McCurdy Breaches (1997–2015)

SAR Year	Critical	Significant[#]
1997		• Chem Demil (Legacy/ NSCMD
1998		• FMTV • Javelin • Longbow Apache
1999	• ATIRCM/CMWS • B-1B CMUP	• NAVSTAR GPS/ Satellite
2000		
2001	• CH-47F • Chem Demil-CMA/ CSD • F-22 • GMLRS • H-1 Upgrades (4BW/4BN) • LPD 17 • Navy Area TBMD[a] • SBIRS High	• B-1B CMUP • MH-60R • V-22
2002	• ATACMS-BAT: BAT P3I[b]	• Comanche • SSN 774
2003	• EELV	• F-35
2004	• Chem Demil-CMA • Chem Demil-CMA Newport	• AEHF • RQ-4A/B UAS Global Hawk • SBIRS High
2005*	• NPOESS • RQ-4A/B UAS Global Hawk • SBIRS High	• ATIRMC/CMWS* • C-130 AMP* • Chem Demil-CMA* • Chem Demil- CMA Newport* • EFV* • F/A-18E/F* • F-35* • JASSM* • JPATS* • MH-60S* • SSN 774* • ASDS[b] • GMLRS
2006	• C-130 AMP • Chem Demil-ACWA • EFV • GMLRS • JASSM • JPATS • Land Warrior[b] • WIN-T	• FBCB2
2007	• C-5 RERP	• AEHF • ARH • JAVELIN • JTRS GMR
2008	• AEHF • ARH[a] • VH-71[a,d]	• H-1 Upgrades (4BW/4BN)
2009	• Apache Block III (AB3) • ATIRCM/CMWS • DDG 1000 • E-2D AHE • F-35 • RMS • WGS	• C-130 AMP
2010	• Chem Demil-ACWA • EFV[b] • Excalibur • RQ-4A/B UAS Global Hawk	• C-27J • Inc1 E-IBCT[b] • JLENS • NPOESS
2011	• AIM-9X Block I[b] • C-130 AMP[b] • JLENS[c] • JTRS GMR[a]	
2012	• EELV	
2013	• JPALS Inc 1A • VTUAV	• AWACS Block 40/45 Upgrade • JTRS HMS
2014	• JSOW[b]	• WIN-T (Inc 2)
2015		

\# Programs that declared a significant breach and subsequently a critical breach in the same SAR year are listed only as critical breaches. Programs that declared multiple significant breaches in the same SAR year are listed only once.

* Programs in purple shading (2006–2015 for critical; 2005–2015 for significant) breached against the original baseline as per the FY 2006 NDAA. Programs in blue shading (1997–2005 for critical; 1997–2004 for significant) breached according to prior criteria, which allowed rebaselining. Eleven programs that did not have a breach prior to the new FY 2006 criteria had significant breaches as a result of this legislative change. The FY 2006 NDAA also permitted the following 25 programs to revise their original baselines to equal their current baseline estimates as of January 6, 2006, without declaring a critical breach: AEHF; AMRAAM; ASDS; Black Hawk Upgrade; Bradley Upgrade; C-17A; CH-47F; EELV; F-22A; FCS; FMTV; Global Hawk; GMLRS; Javelin; JSOW; H-1 Upgrades; Longbow Apache; LPD-17; MH-60R; Minuteman III Guidance Replacement Program; NPOESS; SBIRS High; T-45TS; Trident II Missile; V-22.

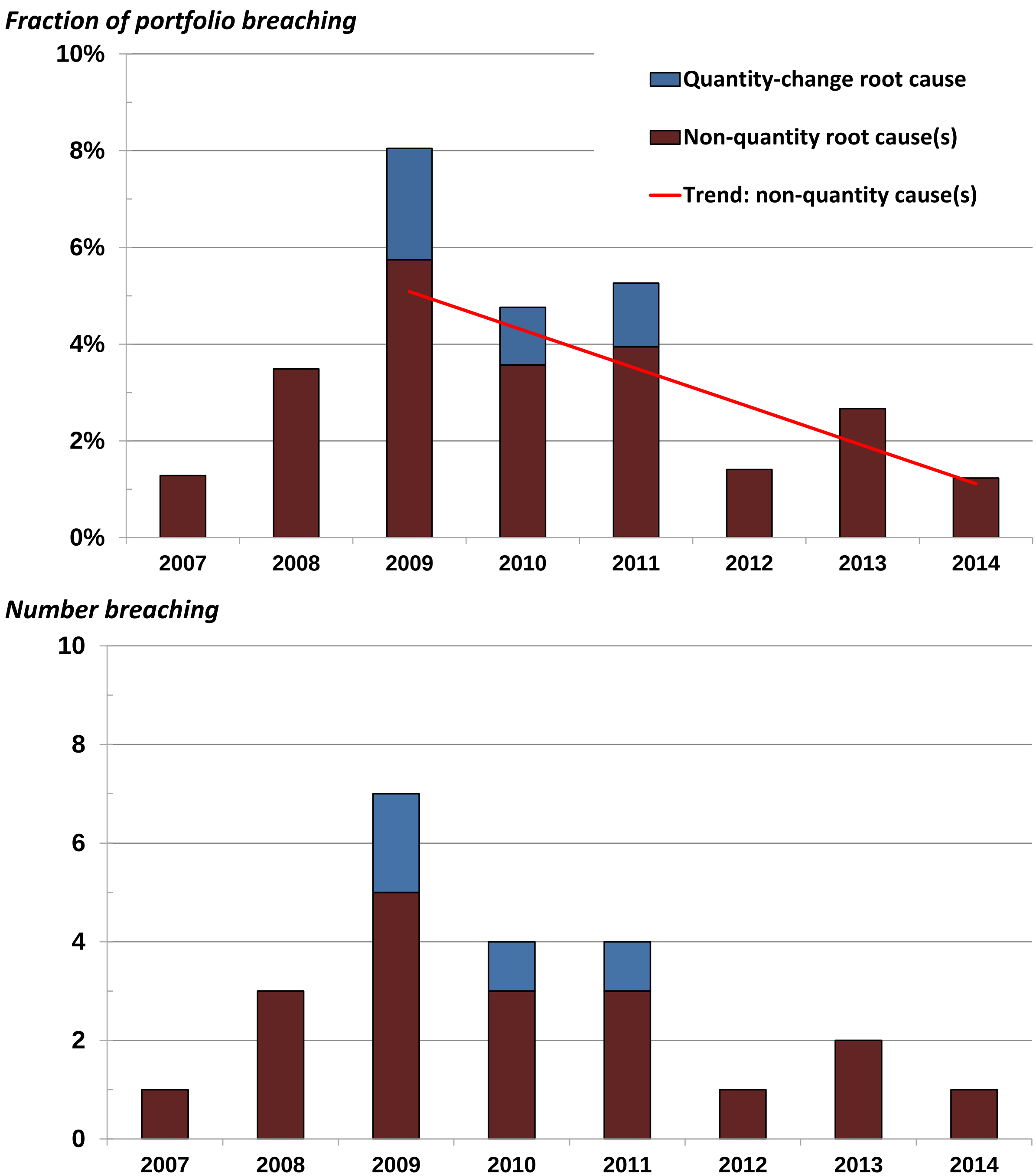
a Following a declared breach, the program was terminated rather than certified.

b Breach resulted from a decision to terminate the program.

c Breach resulted from a decision to terminate procurement phase; EMD units were completed.

d The DoD did not submit a December 2008 SAR to Congress due to a change in Administration. The VH-71 breach was reported in the March 2009 SAR, but the breach occurred in the 2008 reporting period.

Figure 2-6. Critical Nunn-McCurdy MDAP Breaches (2007–2014)



NOTE: Breaches due to quantity changes are based on PARCA root-cause analysis. Since PARCA was not established until WSARA of 2009, it is unknown whether quantity changes were a root cause of breaches before 2009. There is a statistically significant downward trend in both total critical breaches and nonquantity-related critical breaches since 2009. Breaches are after adjusting for inflation. Since it usually takes a few years before a program might breach again, we removed programs from the portfolio count that have breached recently to avoid the potential bias towards an artificially low breach rate. Also, relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown.

Causes of Recent Breaches

Here we discuss the causes of recent Nunn-McCurdy breaches and whether they relate to BBP initiatives and principles.

There were four breaches in 2013 (two critical and two significant) and two breaches in 2014 (one each) (see Table 2-3). The primary reason for the Vertical Takeoff and Landing Tactical Unmanned Aerial Vehicle Fire Scout (VTUAV) breach (about six-tenths) was a Navy capability and quantity decision, reducing total quantity from 168 to 119 and switching future quantities to the more-capable and more-expensive MQ-8C airframes. The significant breach in the AWACS Block 40/45 Upgrade was due to fleet reductions. The Joint Tactical Radio System (JTRS) Handheld, Manpack, and Small Form-Fit (HMS) radios' significant breach was due to the introduction of multiple vendors, integration costs not estimated by the program, and reduced quantity. The Warfighter Information Network–Tactical (WIN-T) breach was due to production extensions from requirements changes and realization of long-term affordability limitations. While a breach factor, the imposition of long-term affordability corrections for WIN-T, Inc. 2, is a positive sign that the DoD is working to take long-term affordability analysis seriously, which should result in early decisions and reduced sunk costs. The Joint Standoff Weapon (JSOW) breach was from quantity reduction and termination 19 years after the baseline was set. Some of the root causes were outside the direct control of the DoD. The JPALS breach was due to an external event (a Federal Aviation Administration decision). The remaining critical-breach causes were execution issues.

Our improvement initiatives such as BBP should continue to raise the cost consciousness of the DoD and sustain the downward trend in recent breach rates, but there is more to do. We have been working to improve the quality, training, and staffing levels in the acquisition workforce, especially of PMs and program executive officers (PEOs). We have seen anecdotal progress in improved execution, but it will take continued investment, focus, and time for these effects to be realized in Nunn-McCurdy breach data. Perhaps more importantly, the incentive structures for PMs, PEOs, SAEs, and Service Chiefs do not necessarily align to produce decisions that lead to better execution results. Some of this is natural given that the overarching mission of the DoD is national security and operational effectiveness, not simply meeting original cost and schedule targets regardless of changing threats. However, the discipline to make prudent risk decisions requires further work to attend to acquisition fundamentals when considering these larger operational drivers.

Table 2-3. Causes of Recent Nunn-McCurdy Breaches (2013–2014)

Year	Critical	Significant
2013	**JPALS.** The PARCA root-cause analysis on Joint Precision Approach and Landing System (JPALS), Increment 1A (JPALS, Inc. 1A) found the primary trigger was the Federal Aviation Administration's decision to indefinitely delay their previously planned phase out of Inertial-based Landing Systems. As a result, the three Services all terminated efforts to integrate global-positioning system (GPS)-based JPALS into their legacy aircraft, leading to a restructuring of JPALS and 90 percent of the cost growth.	**AWACS Block 40/45 Upgrade.** The Air Force reported that the Airborne Warning and Control System (AWACS) Block 40/45 Upgrade breach is the result of quantity reduction from FY 2015 President's Budget decision to divest and thus not upgrade seven E-3B AWACS aircraft (23% of the fleet) and Congressional reductions in FY 2013 and FY 2014. Quantity-adjusted unit cost growth in the 2014 SAR was -6%.
	VTUAV MQ-8 Fire Scout. The PARCA root-cause analysis on the VTUAV MQ-8 Fire Scout found that about 60% of the cost growth was due to the Navy's decision to reduce total quantity from 168 to 119 and switch future quantities to the more capable and expensive MQ-8C airframes. The rest of the cost growth was due to two factors: 30 percent for stretching the production schedule by 17 years (resulting in a break followed by production at rates below economic order quantities), and another 10 percent due to development and testing challenges encountered with the MQ-8B.	**JTRS HMS.** The Army reported that the primary factors that led to the significant breach on the JTRS HMS program were: change in acquisition strategy from a single vendor (per radio) to multiple vendors; vehicle integration cost requirements not previously identified as a funding responsibility of the program; and changes in the Army Fielding Strategy (less radios fielded per year).
2014	**JSOW.** The Navy reported that the critical breach on the JSOW Unitary variant AGM-154C is the result of the termination of procurement at 3,185 units instead of the original quantity of 7,800 units baselined 19 years earlier in 1995 now that more effective weapons are available against current and evolving threats.	**WIN-T, Inc. 2.** The Army reported that WIN-T Inc. 2 breached due to cost increases from an eight-year extension of the procurement schedule due to Configuration Steering Board direction to transfer total Army requirements from WIN-T Inc. 3 to WIN-T Inc. 2 in conjunction with a reduction of less expensive nodes attributed to the revised Army modernization strategy. Additionally, there was an increase in costs caused by a two-year extension of the procurement schedule due to reduced funding in FY 2021–FY 2028 identified by the Army's Long-Range Investment Requirements Analysis.

Root Causes of Critical Breaches Since 2010

Table 2-4 provides additional details on all statutory and discretionary root-cause analyses to date by the Director of PARCA (2010a-g, 2011a-e, 2012a-c, 2013, 2014a-b), including a new detailed breakout of the management performance issues. Generally, the common root causes remain the same as those discussed in our first report (AT&L, 2013b).

Poor management performance is still a root cause in about half of the MDAPs examined. However, management performance is a statutorily required category that is very broad. To provide a better understanding of specific problems that programs experienced, PARCA identified four sub-categories of poor management performance based on the analyses performed to date: Systems Engineering; Inadequate Incentives; Limited Situational Awareness; and Failure to Act on Information. Because the sub-category Systems Engineering is also broad, it was further broken down into four additional sub-categories, as shown in Table 2-4, and one of those sub-categories—Requirements Management—was also broken down into even more specific sub-categories. These breakouts are intended to provide consumers of root cause analyses products with more specific information and enable actionable changes to improve management performance. The Root Cause Analysis reports for all the programs listed in Table 2-4 are publically accessible on PARCA's Web site.[17]

Unrealistic baseline cost and schedule estimates (e.g., from framing assumptions[18] that proved invalid during program execution) were a root cause in about a quarter of the MDAPs examined.

Quantity changes were a root cause in about a fifth of the MDAPs. While a larger fraction of the programs examined experienced quantity change to some degree, PARCA determined that those quantity changes, except for the four programs indicated in Table 2-4, were not a root cause of program problems, but instead a consequence of other issues.

Actions Based on Root Cause Analyses

While the root causes of Nunn-McCurdy breaches tend to be unique in detail to each program and can involve multiple reasons, we have taken specific actions based on the overarching observations visible in Table 2-4.

Framing Assumptions. A quarter of MDAP Nunn-McCurdy breaches to date are due to faulty key framing assumptions[19] *(see detailed discussion starting on p. 25)*. The new DoD Instruction (DoDI) 5000.02 (AT&L, 2015a) recognizes this and requires explicit discussion of framing

[17] http://www.acq.osd.mil/parca/

[18] See earlier definition of framing assumptions on p. xxxvi.

[19] A *framing assumption* is any supposition (explicit or implicit) that is central in shaping cost, schedule, or performance expectations of an acquisition program. Thus, they reflect some risks deemed worth taking. Key framing assumptions can have a major effect on the success of the program as structured and baselined.

assumptions at major reviews so that executives can make informed risk decisions. Moreover, metrics that indicate failing assumptions will be established and tracked so that failures can be recognized and needed adjustments made as soon as possible.

Improved Systems Engineering. About 40 percent of MDAP Nunn-McCurdy breaches to date were caused by one or more systems engineering problems. The DoD has made significant progress in improving our systems engineering practices (e.g., in DoDI 5000.02), workforce training, and metrics to identify problems early and facilitate solutions.

Contractor Incentives. About 35 percent of MDAP Nunn-McCurdy breaches to date were caused by inadequate contract incentives. We have bolstered contract officer training and support while reinforcing the basic principles in contract incentives. For example, data from last year's (AT&L, 2014, pp. 87–93) indicate that when used appropriately, formulaic incentive contract types (cost-plus-incentive-fee and fixed-price-incentive) motivate cost control about as well as FFP while ensuring that the government gets a good price in the end. There is no single "formula" that works best for all cases. In light of this, I am releasing updated guidance on contract incentives.

Improved Execution Monitoring. About a third of MDAP Nunn-McCurdy breaches to date were caused by limited situational awareness and some instances of a failure of managers to act on available information. Better data availability and analysis (such as EV data quality improvements and linkages to Integrated Master Schedules) are improving our ability to identify and mitigate execution problems earlier.

Table 2-4. Root Causes of Critical Nunn-McCurdy Breaches and Selected Programs

	Root Causes	Fraction	Count	WGS	ATIRCM	CMWS	RMS	AB3	DDG-1000	F-35 JSF	FAB-T	Excalibur	ACWA Chem-Demil	RQ-4A/B GH	Navy ERP	GCSS-MC	JTRS GMR	JLENS	P-8A	EELV	ECSS	JPALS Inc. 1A	MQ-8 VTUAV
	Discretionary analysis	25%	5								*				*	*			*		*		
	Department		20	AF	Army	Army	Navy	Army	Navy	DoD	AF	Army	DoD	AF	Navy	Navy	Army	Army	Navy	AF	AF	Navy	Navy
	Milestone B Year (original)			2000	1995	1995	1999	2006	2005	2001	2007	1997	2003	2001	2004	2007	2002	2005	2004	1998	none	2007	2000
	Critical Breach CY			2009	2009	2009	2009	2009	2009	2009	*n/a*	2010	2010	2010	*n/a*	*n/a*	2011	2011	*n/a*	2012	*n/a*	2013	2013
Inception	**Unrealistic performance expectations**	**10%**	2		x																x		
Inception	**Unrealistic baseline estimates for cost or schedule**	**25%**	5							x			x	x	x	x							
Inception	**Immature technologies or excessive manufacturing or integration risk**	**10%**	2		x											x							
Inception	**Other**	**10%**	2												x						x		
Execution	**Unanticipated design, engineering, manufacturing or technology integration issues arising during program performance**	**15%**	3	x		x																	x
Execution	**Changes in procurement quantity**	**20%**	4					x	x			x						x					
Execution	**Inadequate program funding or funding instability**	**0%**	0																				
Execution	**Poor performance by government or contractor personnel responsible for program management**	**45%**	9				x			x	x			x	x	x	x			x	x		
Execution	*• Systems engineering*	*40%*	*8*				*x*			*x*	*x*			*x*	*x*	*x*	*x*				*x*		
Execution	*– Requirements management*	*30%*	*6*								*x*			*x*	*x*	*x*	*x*				*x*		
Execution	*> Ambiguities in combining requirements documents*	*5%*	*1*														*x*						
Execution	*> Development, translation and allocation of requirements*	*15%*	*3*								*x*						*x*				*x*		
Execution	*> Inadequately funding the program for all requirements*	*5%*	*1*											*x*									
Execution	*– Interface and environmental management*	*10%*	*2*							*x*	*x*												
Execution	*– Holistic performance attributes (e.g., reliability, weight)*	*10%*	*2*				*x*				*x*												
Execution	*– Risk assessments*	*5%*	*1*							*x*													
Execution	*• Inadequate contract incentives*	*35%*	*7*							*x*	*x*			*x*		*x*	*x*			*x*	*x*		
Execution	*• Limited situational awareness*	*30%*	*6*				*x*			*x*	*x*					*x*	*x*				*x*		
Execution	*• Failure to act on information*	*25%*	*5*							*x*				*x*	*x*	*x*					*x*		
Execution	**Other**	**20%**	4	x																x		x	x
			20																				
	PAUC Growth (%) **			18.4	290.6	25.0	79.5	25.5	86.5	57.2	25.1	198.7	39.2	14.0	—	—	92.1	215.7	-1.6	58.0	—	104.3	55.2
	APUC Growth (%) **			27.2	281.5	32.4	54.6	31.2	24.9	57.2	—	129.6	—	22.9	—	—	23.2	—	-1.1	58.0	—	129.0	71.5
	RCA Memo Year (CY)			**2010**									**2011**					**2012**			**2013**	**2014**	

* Discretionary root-cause analysis

** As reported at the time of the breach in the SAR

NOTES: Grayed rows are breakout elaborations of poor performance. ECSS did not reach milestone B.

Table 2-5 below summarizes a different analysis of Nunn-McCurdy breaches by commodity. In this case, we do not "double count" programs that have breached multiple times. This allows us to compare the types of programs that have poor cost performance (as evidenced by crossing any Nunn-McCurdy threshold) to those that have never breached during this period.

As in prior years, helicopter programs showed an abnormally high breach rate, as do the four chemical demilitarization programs. The 100 percent rate for space-launch is not statistically meaningful given only one program is in the dataset.

Table 2-5. Nunn-McCurdy Breach Rate by Commodity

Fraction of MDAPs by Commodity Type With Any Nunn-McCurdy Breach (1997–2014)

Commodity Type	Total # of Programs	# of Programs that Ever Breached	Breach Rate	# of Programs with at Most a Significant Breach	# of Programs with at least one Critical Breach
Chem Demil	4	4	100%	1	3
Space launch	1	1	100%	-	1
Helicopter	15	9	60%	5	4
Satellite	13	5	38%	1	4
Fixed-Wing Aircraft	27	10	37%	3	7
UAV	6	2	33%	-	2
Ground Vehicle	11	3	27%	2	1
Munition/Missile	31	7	23%	1	6
C4ISR	52	11	21%	3	8
Ship/Submarine	19	4	21%	2	2
Missile Defense	8	1	13%	-	1
Total	**187**	**57**	**30%**	**18**	**39**

NOTE: Compares number of programs that have crossed any Nunn-McCurdy threshold to those that have never crossed a threshold. Breaches are determined using "base-year" dollars (i.e., adjusted for inflation). These commodity types are slightly different from those reported last year. For example, sensors logically belong in the C4ISR category, and UAVs are broken out from aircraft to help reveal how they have fared.

MAIS and Business Systems

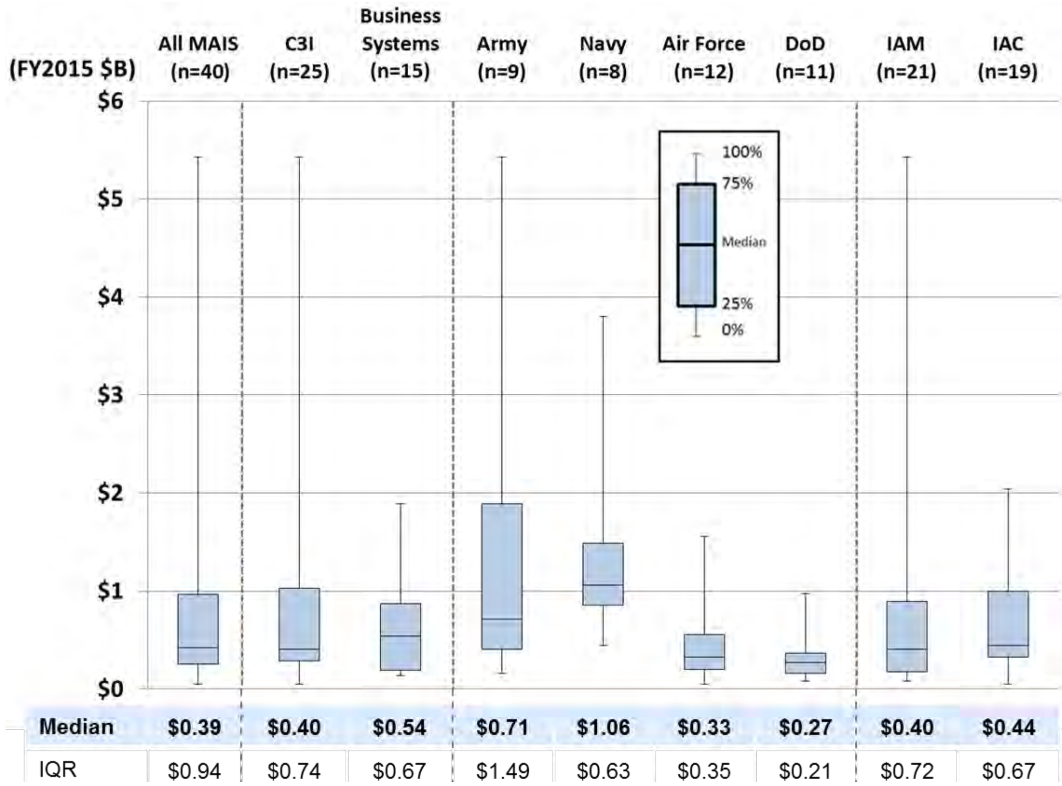
This section evaluates the performance of MAIS, including business systems. MAIS are ACAT IA programs that either meet a specified budgetary threshold[20] or are designated by the Secretary

[20] For example, one MAIS threshold is $378 million in FY 2000 dollars for total life-cycle costs (see 10 U.S.C. section. 2445 for details and other thresholds).

of Defense (or designee) as a MAIS (e.g., due to risk or other concerns) (see 10 U.S.C., section 2445a). Title 10 mandates various reports and baselining mechanisms for MAIS. Figure 2-7 shows the current dollar size of the MAIS programs reported in at least one MAR from 2011–2014.[21] At least half of the MAIS programs have original total baselines below about $500 million (except in the Navy) while others can cost billions of dollars.

Figure 2-7. Major Information System Sizes

Size Distributions (by Dollars) of Active MAIS Programs (2011–2014 MARs)



NOTE: Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The interquartile range (IQR) is the difference between the 75th and 25th percentile; IQR is analogous to variance by conveying a sense of the variability in the distribution. Immature programs that have not completed at least 30 percent of their original EMD schedule time were excluded to help control for maturity.

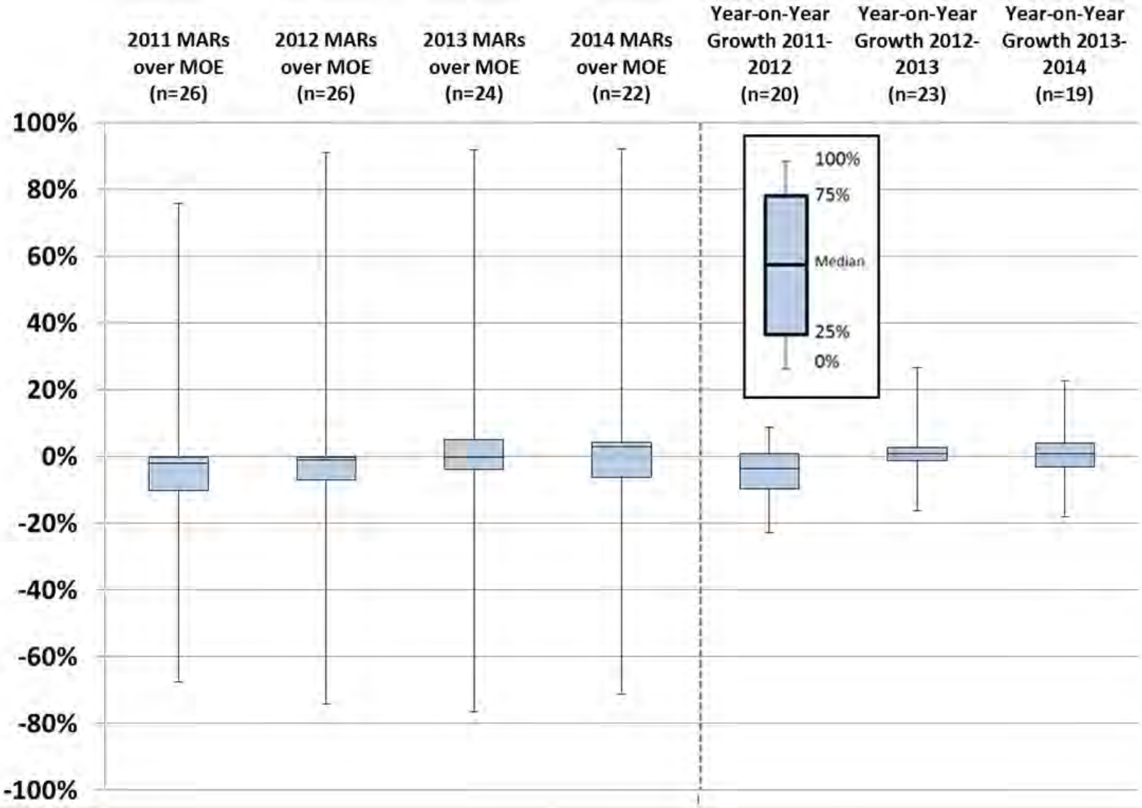
Figure 2-8 shows the funding growth from original baselines as reported in the 2011–2014 MARs. For this year's report, we control for program maturity (we deem a program as relatively mature when it passes 30 percent of the planned or actual time from Funds First Obligated [FFO] to FDD) and weight by program spend (dollars). Total funding growth from original baseline at the medians continues to be near zero with a very small increase in 2014. Looking at year-on-year funding growth, it too is near zero at the median.

---

[21]MARs are the MAIS equivalent of SARs and are provided to Congress to satisfy the requirement in 10 U.S.C., section 2445b.

Figure 2-8. Program Cost-Related Performance: Information Systems

Cumulative Growth from Original Baseline and Annual Changes of Planned Total (From Start to Completion) Funding for Active MAIS (weighted by program spend; 2011–2014)



Median	**-2%**	**-1%**	**0%**	**3%**	**-4%**	**1%**	**1%**
IQR	10%	7%	9%	11%	11%	4%	7%
Trends	***Statistically significant increased cost growth across the years, but the difference is small.***				***Statistically significant increased cost growth across the years, but the difference is small.***		

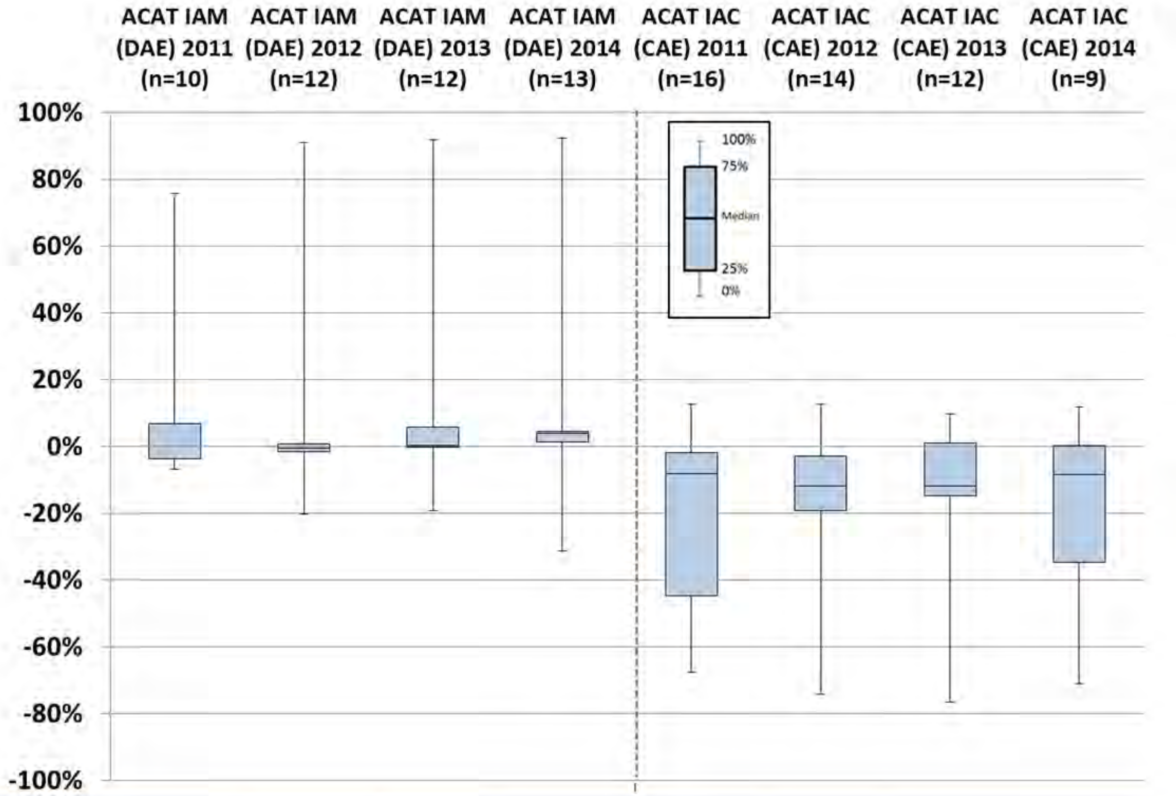
NOTE: The IQR is the difference between the 75th and 25th percentiles. Immature programs that have not completed at least 30 percent of their original EMD schedule time were excluded to avoid artificially low values.

Figure 2-9 compares MAIS funding growth from original baselines based on whether USD(AT&L) or designee is the Milestone Decision Authority (i.e., "IAM" programs) or oversight has been delegated to the Component Head or the Component Acquisition Executive[22] (i.e., for "IAC" programs). Again, we control for program maturity and weight by program spend (dollars). Here we find as in last year's report that most IAC programs have negative funding growth while IAM programs have medians at or just above zero.

[22] In the military departments, the SAE is the Component Acquisition Executive.

Figure 2-9. Program Cost-Related Performance: Information Systems by Decision Authority

Cumulative Growth from Original Estimate of Planned Total (From Start to Completion) Funding for Active IAM and IAC MAIS (weighted for program spend; 2011–2014)



Median	**0**	**-1%**	**0%**	**4%**	**-8%**	**-12%**	**-12%**	**-8%**
IQR	10%	2%	6%	3%	43%	16%	16%	35%

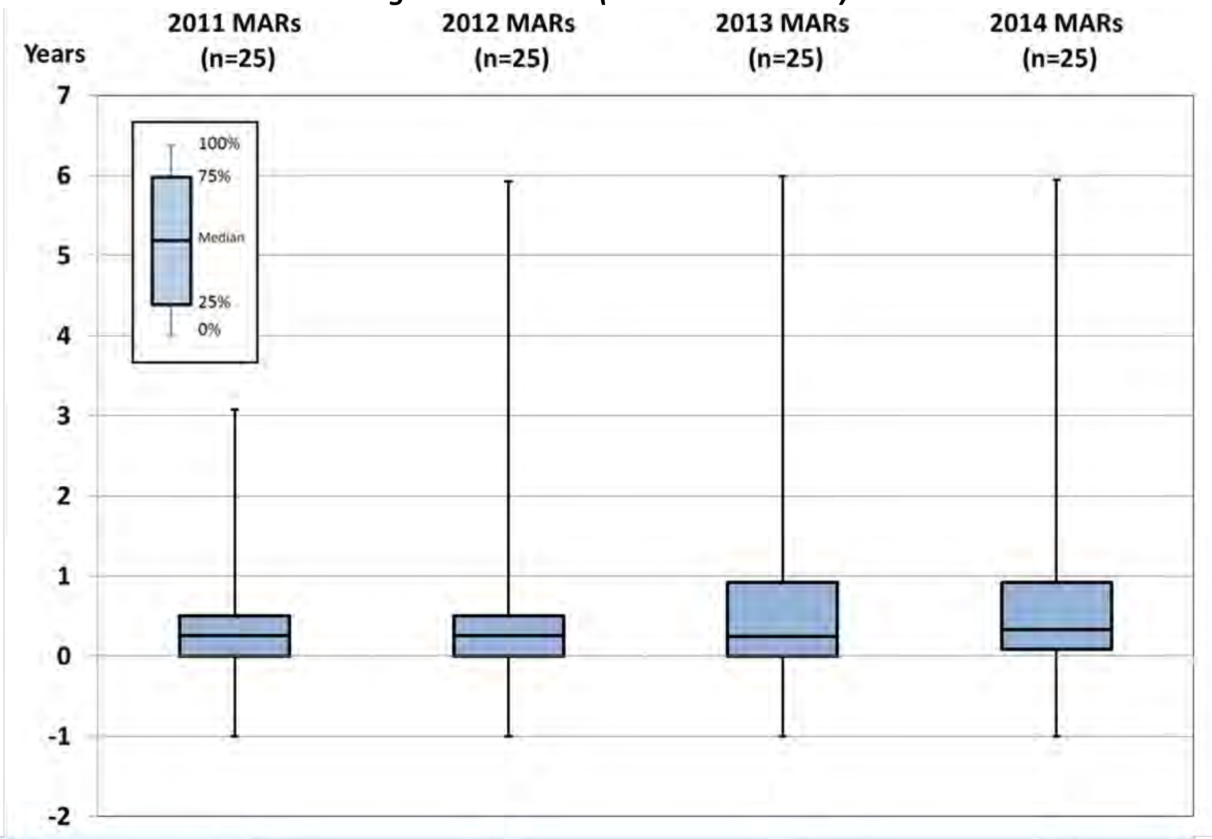
NOTE: Total funding includes O&M and Working Capital Fund. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The IQR is the difference between the 75th and 25th percentiles. Newer programs that have not completed at least 30 percent of their original EMD schedule time were excluded to avoid artificially low values.

This preliminary analysis supports anecdotal assertions that, unlike MDAPs, MAIS may be changing their scope to match available funds and schedule thresholds—especially at the Component level. We currently do not have centralized data to measure requirement changes across MAIS to test this hypothesis. If true, however, examinations of performance relative to baselines would not be as meaningful as they are on MDAPs (where general capabilities are assumed to be relatively stable).

With respect to schedule growth, the MAIS reporting in the 2011–2014 MARs exhibited a median growth of about 4 months with three-quarters of the MAIS at or below 12 months (see Figure 2-10). But there are examples of schedule growth that are considerably larger.

Figure 2-10. Program Schedule Growth: Information Systems

Active MAIS Cumulative Planned Schedule Growth to FDD from Original Estimates (2011–2014 MARs)



	2011 MARs	2012 MARs	2013 MARs	2014 MARs
Median	**0.25**	**0.25**	**0.25**	**0.33**
IQR	0.50	0.50	0.92	0.84

NOTE: Original estimates are those reported in the first MAR for each MAIS. Schedule period is from MS B or FFO to FDD. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The IQR is the difference between the 75th and 25th percentiles. Newer programs that have not completed at least 30 percent of their original EMD schedule time were excluded to avoid artificially low values.

COST AND SCHEDULE PERFORMANCE: DEVELOPMENT

First, we examine cost growth for MDAPs in development. For this year’s report we updated the performance of all contracts and added more recent MDAP contracts.

Planned Program Funding Growth: Development

We now examine MDAP development cost-related performance at the program level, using total RDT&E funding growth as the metric. Program “cost” (e.g., as defined for PAUC and APUC) is synonymous with the total amount of funding because it reflects the prices paid on contracts as well as program execution costs. Later, when we discuss contracts, we will distinguish contract prices from their underlying contractor costs and margins (profits and fees).

Generally, RDT&E must be funded regardless of how many units are produced. In that sense, they are a fixed cost for the DoD to arrive at the point where we can procure and field a capability. Thus, for RDT&E, we track total funding growth rather than by units produced (e.g., as for PAUC and APUC) to avoid confusing the effects of quantity changes with growth in RDT&E. Since we measure growth compared to initial baselines, this measure can show significant increases when a program originally was planned to involve little RDT&E but received even modest additions to address changing threats or operational needs. Still, this approach provides a means for measuring total RDT&E funding control relative to original plans.

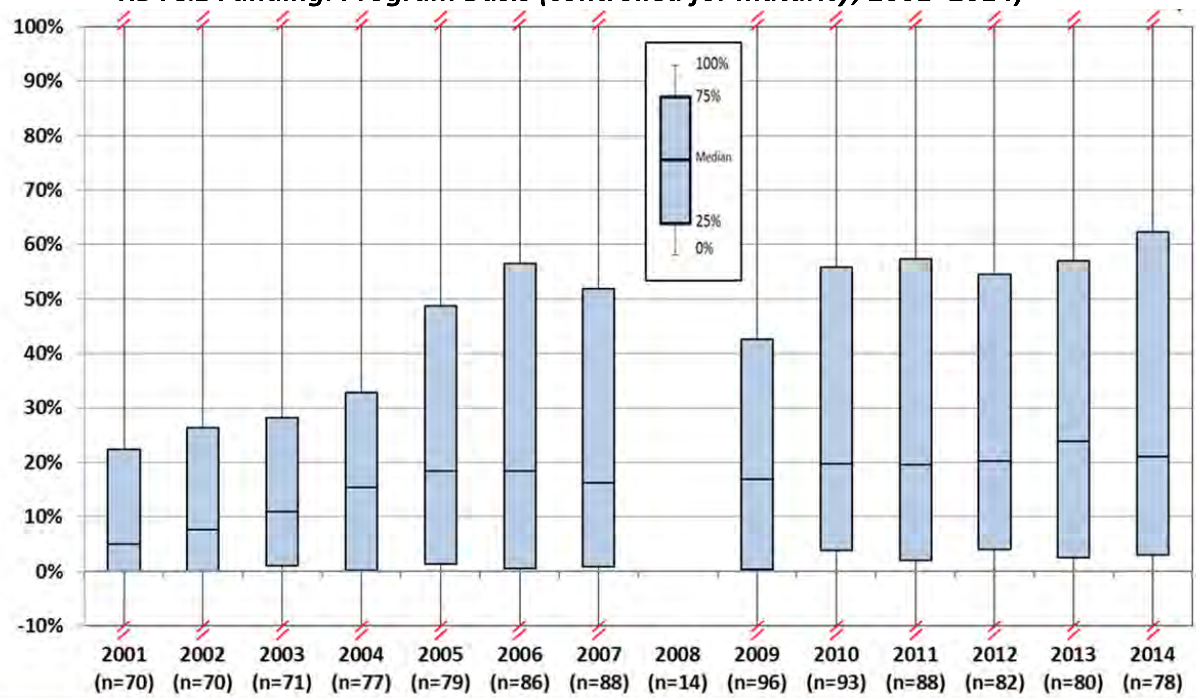
While examining total RDT&E funding from each program's original baseline estimate is important to capture the overall growth since inception, it may not be the best choice for gaining insight into recent cost-growth management. When we analyze a program from inception, we are forced to carry all growth until the program or phase of the program ceases to be active. Programs that are currently executing well but had a one-time increase in the distant past can appear to be poor performers in the long term. Therefore, we also measure biennial changes in total planned and actual RDT&E funding.

Figure 2-11 shows total cumulative RDT&E funding growth over original MS B baseline for each year's MDAP portfolio. For each analysis, we first show the main portion of the distribution (between −10 percent and 100 percent growth) followed by a separate figure showing all outliers (especially those with growth greater than 100 percent). Medians are the lines within each box. Gray-shaded columns in the table beneath each chart were periods with very low sample counts because full SAR reporting was not made in those years due to new Presidential administrations. The "x" markers above the box mark the five largest instances of program funding growth (although outliers above 100 percent only appear on the outlier charts). These outlier charts are controlled for program maturity only. Notably, the data show considerable (and sometimes seemingly conflicting) differences between the medians and the arithmetic means. This is because the data are highly skewed, and a single but very large outlier can have a large effect on the mean while not affecting the median.[23] In these cases, the best measure of central tendency is the mean.

[23] Part of the skewing in the distribution of cost change is the mathematical boundary of cost change because cost cannot decrease more than 100 percent but can increase more than 100 percent.

Figure 2-11. Program Cost-Related Performance: Development

Cumulative Growth Over Original MS B Baseline of Active MDAP Planned Total (From Start to Completion) RDT&E Funding: Program Basis (controlled for maturity; 2001–2014)



	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Median	**5%**	**8%**	**11%**	**15%**	**18%**	**18%**	**16%**		**17%**	**20%**	**20%**	**20%**	**24%**	**21%**
IQR	22%	26%	27%	33%	47%	56%	51%		42%	52%	55%	51%	55%	59%
Trends	***Generally lower than 2009–2014.***			***No trend across these years.***										

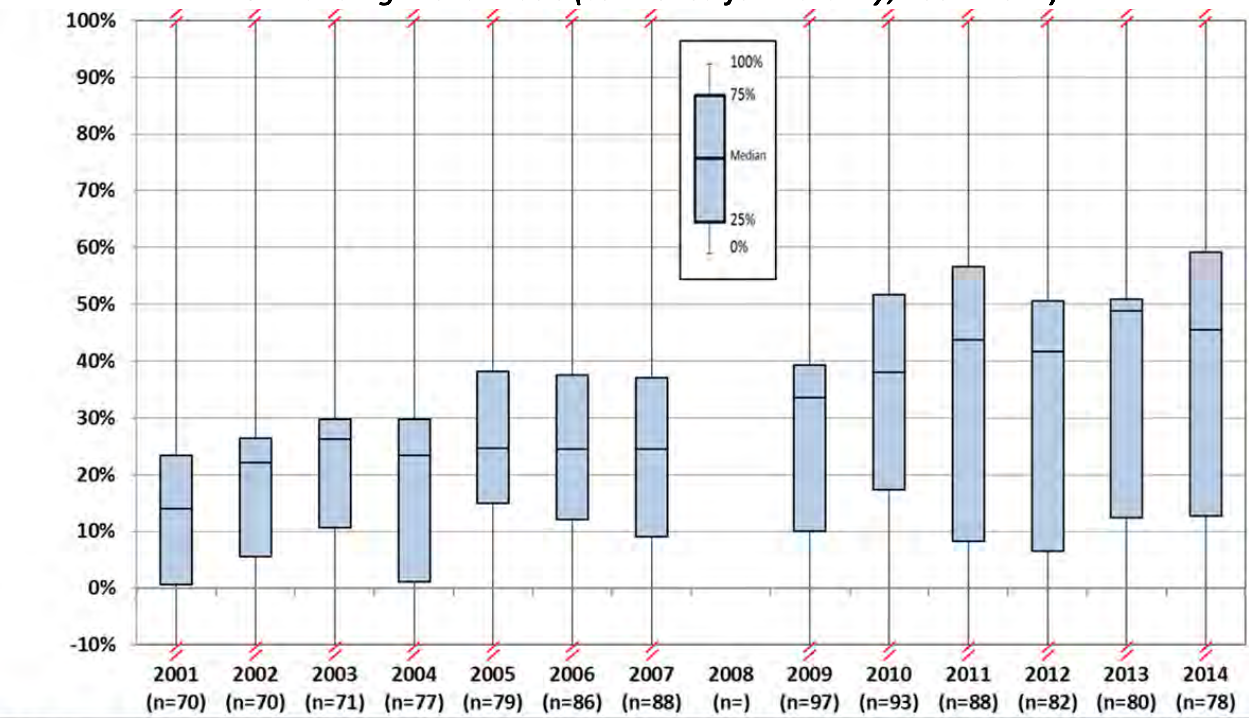
NOTES: This shows total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. These are percentage changes after adjusting for inflation from original MS B baseline of actual past and estimated future funding as reported in each program's latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The IQR is the difference between the 75th and 25th percentiles.

As with last year's results, growth has been statistically flat since 2001–2004, when the set of MDAPs active at that time had lower total RDT&E at the median.

In contrast to the results on a program basis, Figure 2-12 summarizes new analysis showing these results on a dollar basis (i.e., weighted by spending) and controlling for maturity (i.e., removing programs that have not executed at least 30 percent of their original EMD schedule from MS B to MS C). Here growth has been statistically increasing since 2001. In other words, larger programs (in terms of spending) have systematically larger total RDT&E funding growth, and that growth has been increasing. Thus, our outlier programs also are our largest by spending, and they are increasing. F-35, for example, constitutes about 20 percent of the dollars in the current MDAP portfolio and thus has a large effect when weighted by program size (dollar basis). However, the median total funding growth for F-35 is very close to the median of the rest of the portfolio. Thus, it is not driving the value upward but will tend to affect whether the median changes in the future.

Figure 2-12. Program Cost-Related Performance: Development (Weighed by Size)

Cumulative Growth Over Original MS B Baseline of Active MDAP Planned Total (From Start to Completion) RDT&E Funding: Dollar Basis (controlled for maturity; 2001–2014)



	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Median	**14%**	**22%**	**26%**	**23%**	**25%**	**25%**	**25%**		**34%**	**38%**	**44%**	**42%**	**49%**	**46%**
IQR	23%	21%	19%	29%	23%	25%	28%		29%	34%	48%	44%	39%	46%
Trend	***Increasing trend across all years.***													

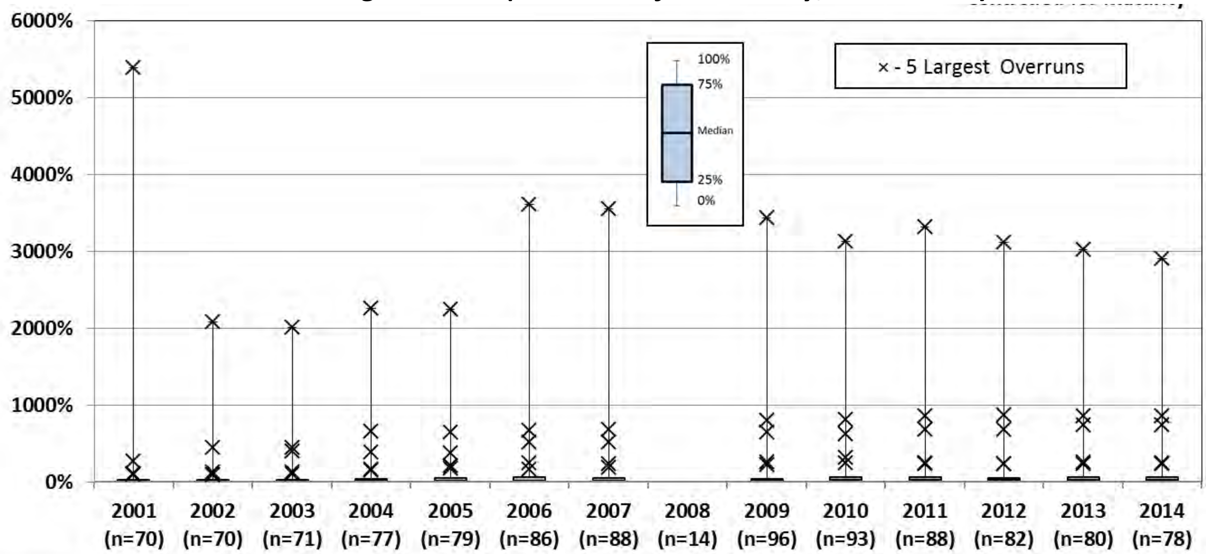
NOTES: This shows total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. These are percentage changes after adjusting for inflation from original MS B baseline of actual past and estimated future funding as reported in each program's latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The IQR is the difference between the 75th and 25th percentiles.

Figure 2-13 shows the outliers that are off the chart in Figure 2-11 and Figure 2-12. The five largest funding growth programs remain unchanged from last year's report. These outliers have very large growth percentages but are not representative of the overall MDAP portfolio. These extreme growths are not due to measurement error and so were not excluded from the analysis. Still, they do skew the aggregate data, which is an important fact for determining how to measure and discuss funding growth across a program population. Similar skewing is observed in various complex commercial projects (see, for example, Flyvbjerg et al., 2002).

Much of the funding growth from original MS B baselines are from prior years since recent marginal funding growth in RDT&E has moderated significantly at the median.

Figure 2-13. Program Cost-Related Performance Outliers: Development

Cumulative Growth Over Original MS B Baseline of Active MDAP Planned Total (From Start to Completion) RDT&E Funding: Outliers (controlled for maturity; 2001–2014)



<table>
<tr><th></th><th>2001</th><th>2002</th><th>2003</th><th>2004</th><th>2005</th><th>2006</th><th>2007</th><th>2008</th><th>2009</th><th>2010</th><th>2011</th><th>2012</th><th>2013</th><th>2014</th></tr>
<tr><td>Largest Outlier</td><td>GMLRS/GMLRS AW [Launcher]</td><td colspan="6">C-130J</td><td rowspan="5"></td><td colspan="6">C-130J</td></tr>
<tr><td>2nd Largest</td><td colspan="7">MH-60S</td><td colspan="6">MH-60S</td></tr>
<tr><td>3rd Largest</td><td>CVN 68 [CVN-77]</td><td>SBIRS High [Baseline (GEO 1-4, HEO 1-2, and Ground)]</td><td colspan="5">GMLRS/GMLRS AW</td><td colspan="6">GMLRS/GMLRS AW</td></tr>
<tr><td>4th Largest</td><td>SBIRS High [Baseline (GEO 1-4, HEO 1-2, and Ground)]</td><td>H-1 Upgrades</td><td colspan="2">SBIRS High [Baseline (GEO 1-4, HEO 1-2, and Ground)]</td><td colspan="3">RQ-4A/B Global Hawk</td><td colspan="6">RQ-4A/B Global Hawk</td></tr>
<tr><td>5th Largest</td><td colspan="2">GMLRS/GMLRS AW</td><td>H-1 Upgrades</td><td>UH-60M Black Hawk</td><td colspan="3">SBIRS High [Baseline (GEO 1-4, HEO 1-2, and Ground)]</td><td colspan="6">SBIRS High [Baseline (GEO 1-4, HEO 1-2, and Ground)]</td></tr>
</table>

NOTE: This shows total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. These are percentage changes after adjusting for inflation from original MS B baseline of actual past and estimated future funding as reported in each program's latest SAR.

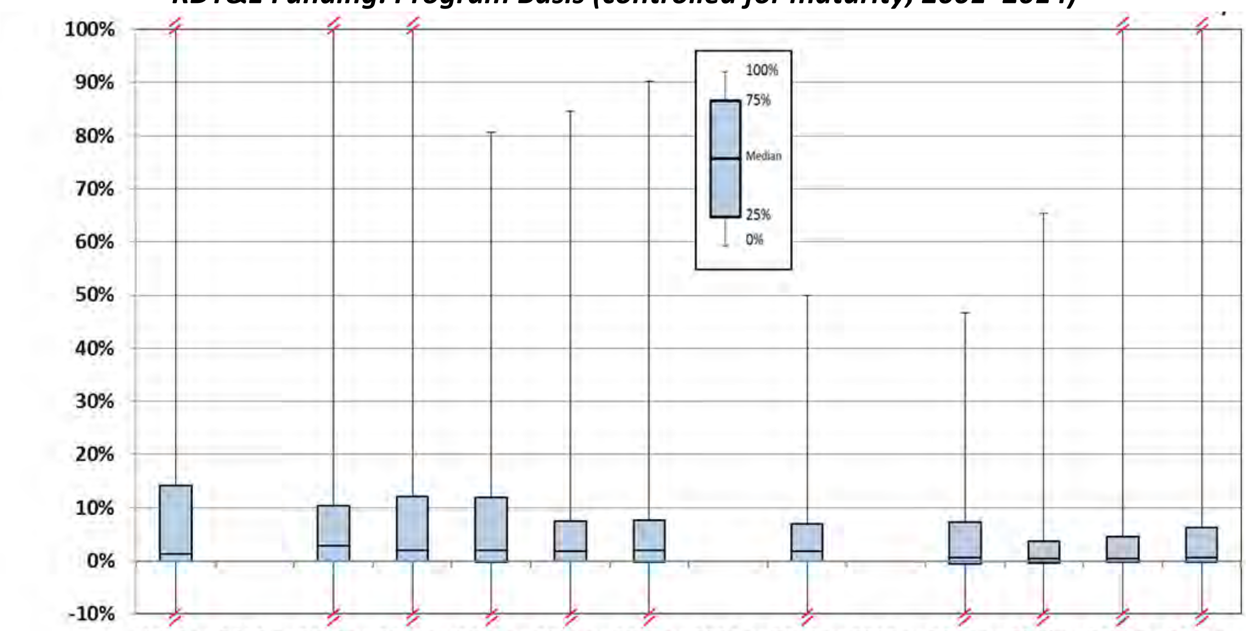
Understanding *why* a program may exhibit such a large percentage increase in RDT&E funding requires an individual examination of each case. For example, in Figure 2-13, the C-130J remains the highest outlier since 2002. This program originally was envisioned as a nondevelopmental aircraft acquisition with a negligible RDT&E effort planned. Several years into the program, a decision was made to install the Global Air Traffic Management system, adding several hundred million dollars to development and causing the total development funding growth recently to climb upward of 3,000 percent as of late. This is an example of a major change in the program rather than poor execution, although significant program changes like this are not necessarily the reason for all extreme cases of funding growth. Later on p. 44, we the attributes of funding growth outliers for MDAP contracts since FY 2000 on currently active programs, including contracts for two of the five outliers from this figure: Global Hawk System Development and Demonstration (SDD) and Space-Based Infrared System (SBIRS) High SDD.

Biennial Planned Program Funding Changes: Development

Figure 2-14 and Figure 2-15 show a continuing downward trend when examining biennial changes in total (past plus planned) program RDT&E funding growth—both on program and dollar bases (weighted by spend). The last four periods are each lower than almost all of the prior periods.

Figure 2-14. Program Cost-Related Biennial Performance: Development

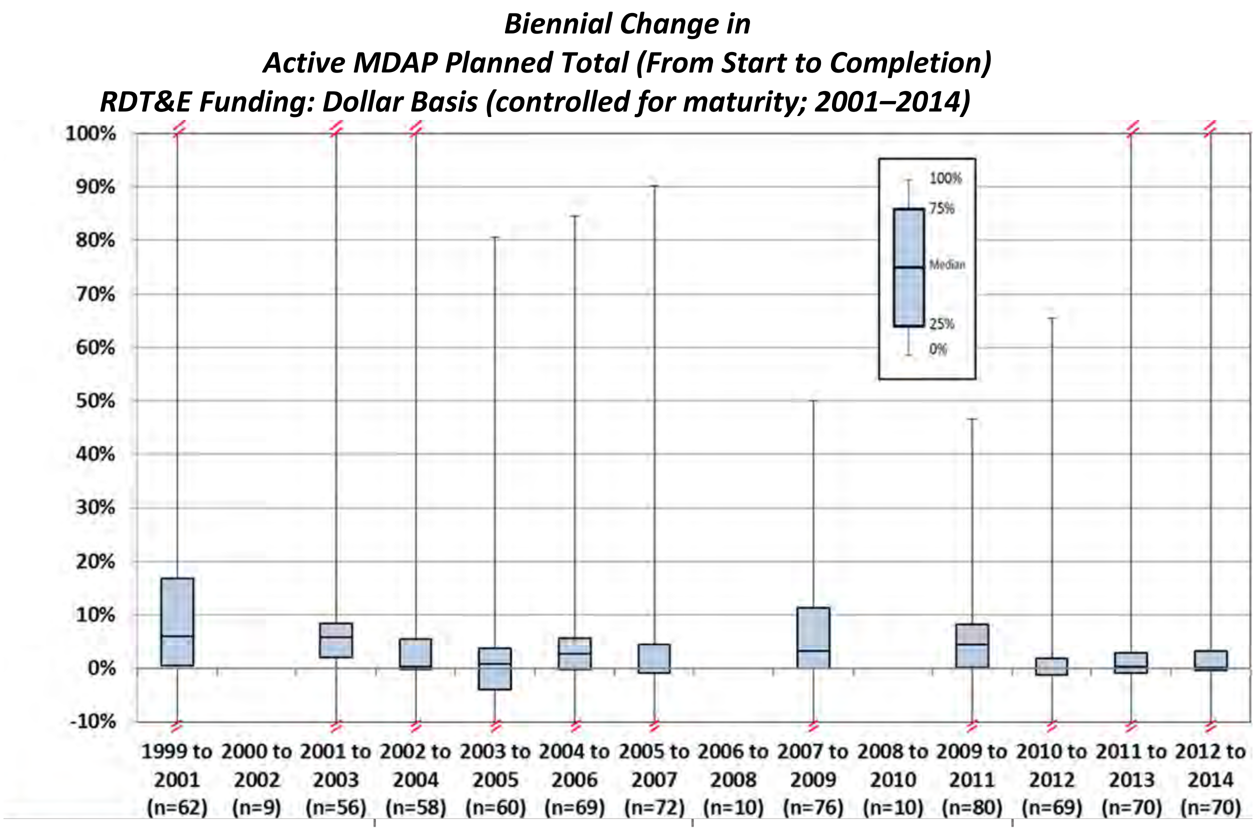
Biennial Change in
Active MDAP Planned Total (From Start to Completion)
RDT&E Funding: Program Basis (controlled for maturity; 2001–2014)



	1999 to 2001 (n=62)	2000 to 2002 (n=9)	2001 to 2003 (n=56)	2002 to 2004 (n=58)	2003 to 2005 (n=60)	2004 to 2006 (n=69)	2005 to 2007 (n=72)	2006 to 2008 (n=10)	2007 to 2009 (n=76)	2008 to 2010 (n=10)	2009 to 2011 (n=80)	2010 to 2012 (n=69)	2011 to 2013 (n=70)	2012 to 2014 (n=70)
Median	**1%**		**3%**	**2%**	**2%**	**2%**	**2%**		**2%**		**1%**	**0%**	**0%**	**1%**
IQR	14%		10%	12%	12%	7%	8%		7%		8%	4%	5%	6%
Trends	***No trend across these years.***										***Generally lower than periods ending 2001–2005***			

NOTE: This measures biennial changes in total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. These are percentage changes after adjusting for inflation from original MS B baseline of actual past and estimated future funding as reported in each program's latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The IQR is the difference between the 75th and 25th percentiles.

Figure 2-15. Program Cost-Related Biennial Performance: Development (weighted by size)



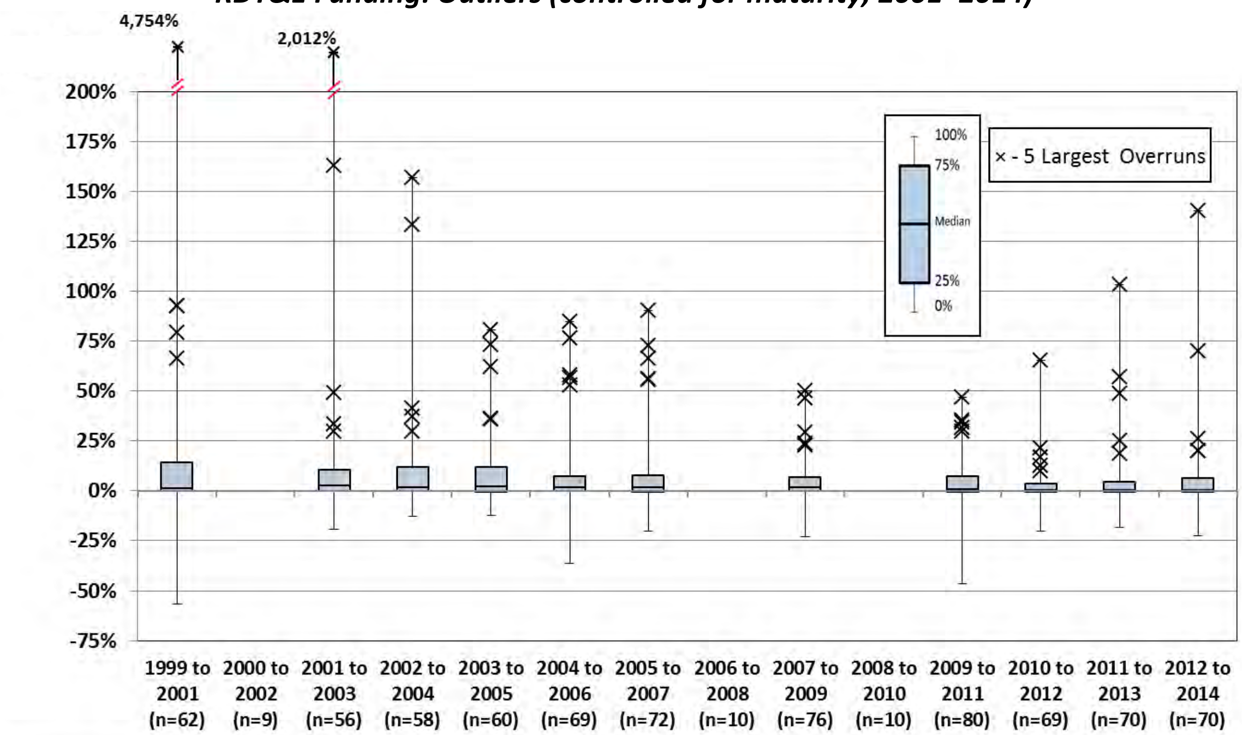
	1999 to 2001	2000 to 2002	2001 to 2003	2002 to 2004	2003 to 2005	2004 to 2006	2005 to 2007	2006 to 2008	2007 to 2009	2008 to 2010	2009 to 2011	2010 to 2012	2011 to 2013	2012 to 2014
Median	**6%**		**6%**	**0%**	**1%**	**3%**	**0%**		**3%**		**4%**	**0%**	**0%**	**0%**
IQR	16%		6%	6%	8%	6%	5%		11%		8%	3%	4%	4%
Trends				***Generally lower than prior two and next two periods.***								***Generally lower than before 2012.***		

NOTE: This measures biennial changes in total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. These are percentage changes after adjusting for inflation from original MS B baseline of actual past and estimated future funding as reported in each program's latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The IQR is the difference between the 75th and 25th percentiles.

Figure 2-16 shows the five largest programs with biennial changes in planned and actual RDT&E funding, controlling for program maturity. This includes outliers that are off the charts in Figure 2-14 and Figure 2-15.

Figure 2-16. Program Cost-Related Biennial Performance Outliers: Development

Biennial Change in
Active MDAP Planned Total (From Start to Completion)
RDT&E Funding: Outliers (controlled for maturity; 2001–2014)



Largest Outlier	GMLRS Launcher		C-130J	GMLRS	JTN	WIN -T	ARH		MRAP		AH-64E Reman	Reaper	AIM-9X Blk II	
2nd Largest	CVN-77		GMLRS	UH-60M	UH-60M	SSDS MK 2 P3I	Chem Demil ACWA		WIN-T Inc 1		IAMD	MIDS	JPALS	MQ-8 Fire Scout
3rd Largest	GMLRS		UH-60M	JTN	JTRS GMR	C-130J	SSDS MK 2 P3I		Patriot MSE		STRYKER	E-2D AHE	Reaper	JPALS Inc 1A
4th Largest	Chem Demil CMA CSD		MH-60S	WGS	NPOESS	JTN	NAVSTAR Equipment		CH-53K		JTRS HMS	IAMD	VTUAV	NMT
5th Largest	SBIRS High		NAVSTAR Equipment	MH-60S	WIN -T	Chem Demil ACWA	C-130J		LHA 6		Gray Eagle	FAB-T	E-2D AHE	EELV

NOTE: This measures biennial changes in total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. These are percentage changes after adjusting for inflation from original MS B baseline of actual past and estimated future funding as reported in each program's latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

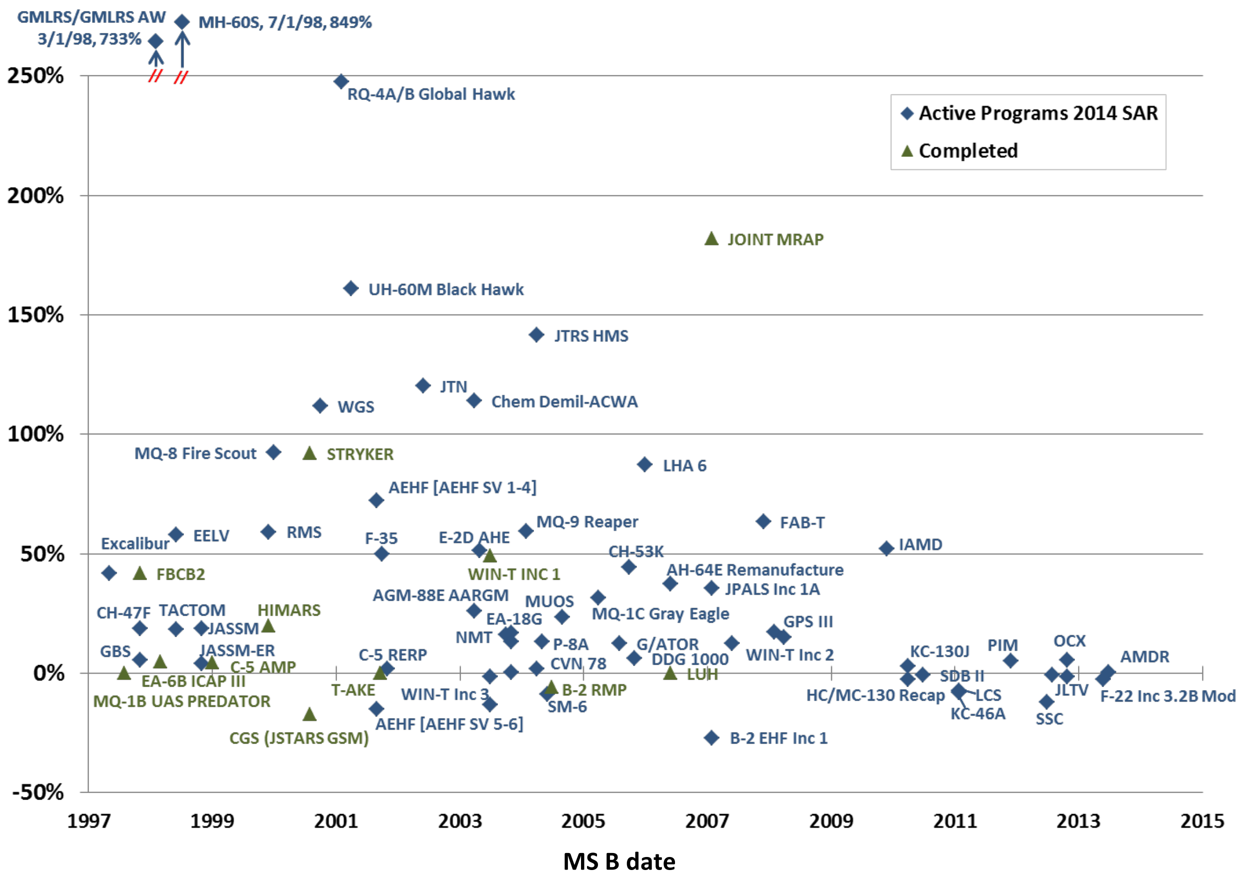
Program Funding Growth by Start Date: Development

A different way to examine program performance is to ask whether recently started MDAPs are controlling RDT&E better or worse than completed and active older MDAPs. This is important since we found in our 2013 report that MDAP contracts that start off well continue to do better, and this is commonly believed to hold at the program level as well.

Figure 2-17 plots total RDT&E funding growth (past and planned) of all currently active and completed MDAPs by their original MS B date. As before, we controlled for maturity by removing newer programs.

Figure 2-17. Program Cost-Related Performance: Development

Cumulative Growth Over Original MS B Baseline of MDAP Planned Total (From Start to Completion) RDT&E Funding (controlled for maturity; 1997–2015)



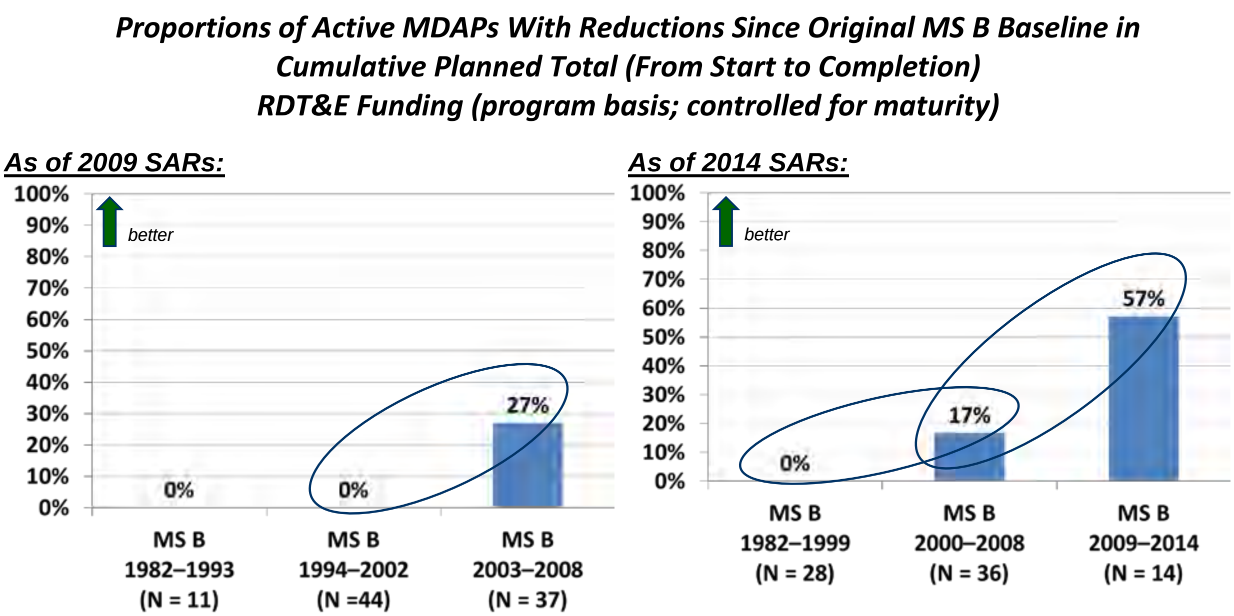
NOTE: This measures total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. These are percentage changes after adjusting for inflation from original MS B baseline of actual past and estimated future funding as reported in each program's latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown.

There is an apparent recent improvement in Figure 2-17, but it is not statistically significant when testing back to 1997; the variation explained by any potential trend (i.e., the R^2) was too low.

However, we did find that recent programs are, in fact, improving when using a proportion test comparing programs started since 2009 to earlier programs. Figure 2-18 shows the result of separating the active MDAPs based on program start date (i.e., their original MS B date) as of the 2014 SARs. After deleting immature programs (i.e., those that are too new to know the magnitude of potential problems), we find that the recent proportion of active MDAPs showing reductions (negative funding growth from original MS B baseline) in total RDT&E is significantly higher statistically at 57 percent for the period since 2009 than at 17 percent for the prior period of 2000–2008.

To help test whether this is just a result of any remaining maturity bias, we conducted the same proportion test using earlier 2009 SARs and the same time windows shifted earlier. While the later proportion (2003–2008) was significantly higher at 27 percent than the then-prior nine years (1994–2002) at zero percent, the reduction magnitudes were lower than what we see now in the 2014 SARs. Thus, while the proportions show some maturity bias remains, there is evidence that the overall levels in 2014 are higher than they were in 2009 and that recently started MDAPs are controlling development funding better.

Figure 2-18. Planned Reductions in Program Funding: Development



NOTE: This reflects total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. Statistically significant differences between adjacent periods are marked with an oval. A program shows a reduction if current total RDT&E funding is under original MS B baseline. Relatively new programs that have not been through at least 30 percent of their original EMD schedule are not included to help control for low maturity.

Conceptually, two behaviors may be at work here. First, recent initiatives (e.g., WSARA 2009; BBP initiatives, including Should-Cost and improved contractor cost control) may indeed be leading to better programs. These factors should lead to reductions in RDT&E (whether motivated initially by budgetary pressures or as a result of savings or avoidance realized), and data at the contract level shown later in this chapter show we indeed are significantly reducing RDT&E growth. Second, we might be seeing reductions due to budget cuts with content adjustments. For example, engineering and design requirements may be moderating, enabled by actions from the PM and Configuration Steering Boards and reinforced by affordability analysis. We cannot tell from these data, but the signs are encouraging.

Contract Cost Growth: Development

Contract-level cost growth data help provide early indicators of potential cost growth at the program level. This is not a perfect indicator since we measure contract cost growth from the contract budget base of the original award. The program may have planned to exercise options or use the same contract vehicle to receive additional work, which would show up here as cost growth. Still, particularly large cost growth can indicate cases that warrant further investigation to determine whether problems exist. By monitoring contract performance, we may see these problems well before they reveal themselves in the program-level SARs.

We generally have EV data for major MDAP contracts, and EV data forms the basis for much of our contract analysis. Major contracts include the six largest contracts (prime, associated, or for government-furnished equipment) for each MDAP valued at over $40 million, although EV data are usually not provided for firm-fixed price (FFP) contracts. EV data also be available for other MDAP contracts of at least $60 million in RDT&E or $250 million in procurement or ship construction (in FY 1990 constant dollars).

Figure 2-19 shows more than 20 years of contract data on total cost growth relative to initial contract cost targets for major MDAP development contracts.[24] Later, Figure 2-37 shows the same for early production contracts. These plots have been extensively updated since last year's report, including the addition of recent contracts that are relatively mature (spending at least 30 percent of their original budget; see p. 9 for a discussion on our techniques for avoiding maturity bias). Outliers and selected contracts for well-known programs are identified by program name (including some duplicates for multiple large contracts for the same MDAP).

Note that Figure A-1 through Figure A-27 later in Appendix A provide a detailed examination of the 27 development contracts in Figure 2-19 that are statistical cost outliers. These examinations provide insight into the kind of issues seen in extreme cost growth contracts and the benefits of using contract cost growth as an indicator of potential problems. For example, the Navy attributes most of the growth on the CVN 21 construction preparation (CP) contract to

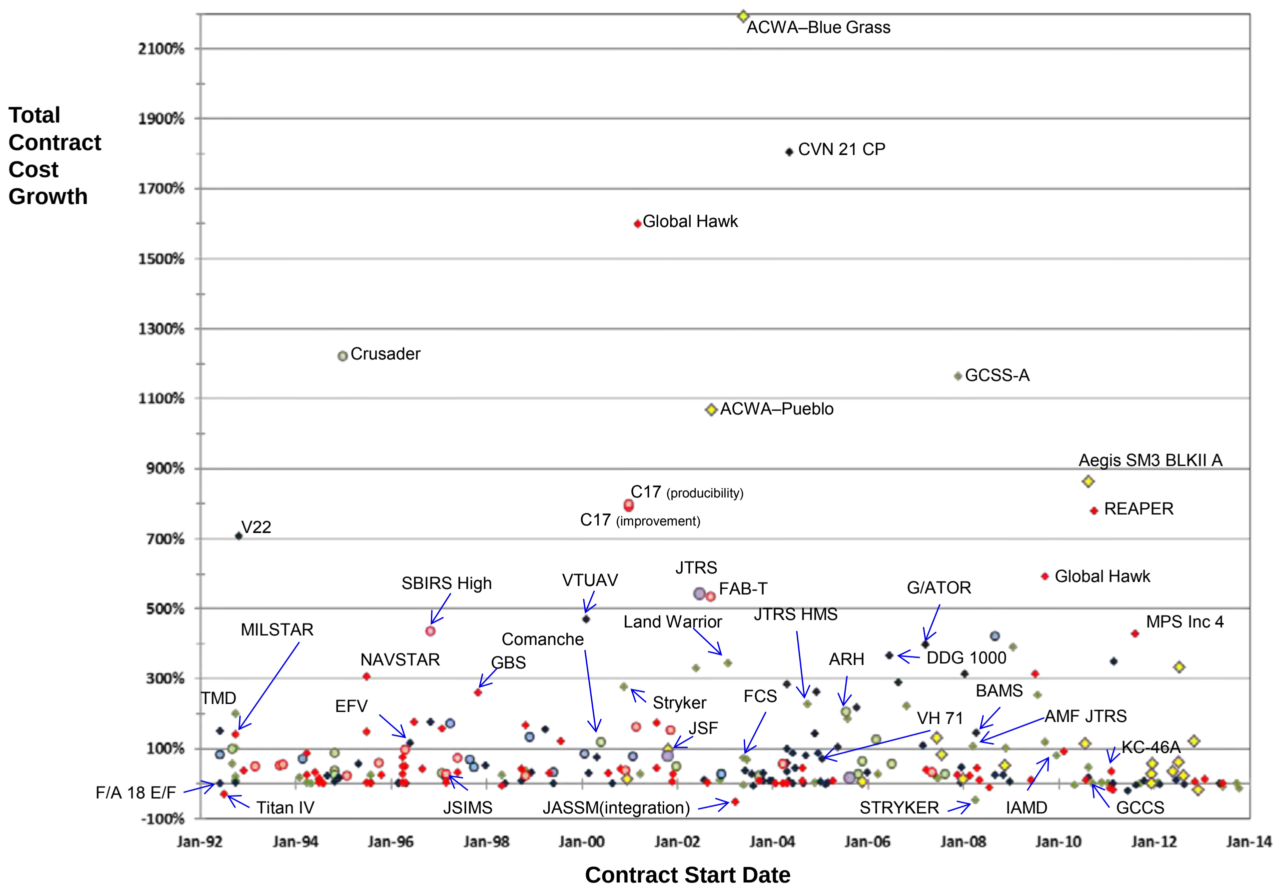
[24] "Major MDAP contracts" are defined in on p. 11.

two revisions of the lead-ship delivery date and associated effects on the contract period-of-performance (see Figure A-11 on p. 143).

These data convey the magnitude and distributions of cost growths in defense acquisition and the challenges of trying to identify causes and motivate performance ownership. Note that the scatter plots reveals significant variation and skewing in total contract cost growth measured from original contract cost target. Caution also is warranted in that these are *contract* data. Performance of the larger program against baselines may be different depending on how large the contract is and the degree to which the baseline can absorb any growth in the contract observed.

Figure 2-19. Program Contract Cost-Related Performance: Development

DoD-Wide Development MDAP Major-Contract Total Cost Growth (1992–2015)



NOTE: This shows total cost growth, including any work-content changes and cost-over-target. Army programs are shown in green; Navy in blue, Air Force in red, DoD-wide in purple and yellow. Diamonds on the charts indicate growths mostly attributable to the start of the contract; circles represent significant work added later. Source cost data were reported in TY dollars (unadjusted for inflation). Immature contracts that have not spent at least 30 percent of their original target cost were excluded to help control for maturity. See Appendix A for attributed explanations of cost growth for statistical outliers.

These contracts completed at least 30 percent of original contract dollar amount. Programs that will exhibit performance problems generally exhibit significant cost growth by this point. However, many recent contracts are not completed and may (or may not) end with cost growth rates higher than currently estimated. For example, total cost growth increased significantly since our prior update of these data on contracts for Global Hawk, Global Combat Support System—Army (GCSS-A), and guided-missile destroyer (DDG) 1000 (among others) in development; and Littoral Combat Ship (LCS), initial F-35 LRIPs, and E-2D Hawkeye in early production.

Contract Cost-Growth Outliers: Development

While outliers do not reflect the overall performance of the defense acquisition system, the detailed examination of the performance of MDAP contracts with especially high cost growth is important. These are the most significant cost-growth programs, and understanding how such cost increases occurred can help us prevent similar outliers in the future. These outliers can be analyzed individually to ascertain the attributed explanations of their high cost growth. These outliers can then be compared to more typically executing contracts to determine the primary attributes that systematically and significantly determine their relatively poor cost performance. Note that we have not yet analyzed early production contract cost-growth outliers.

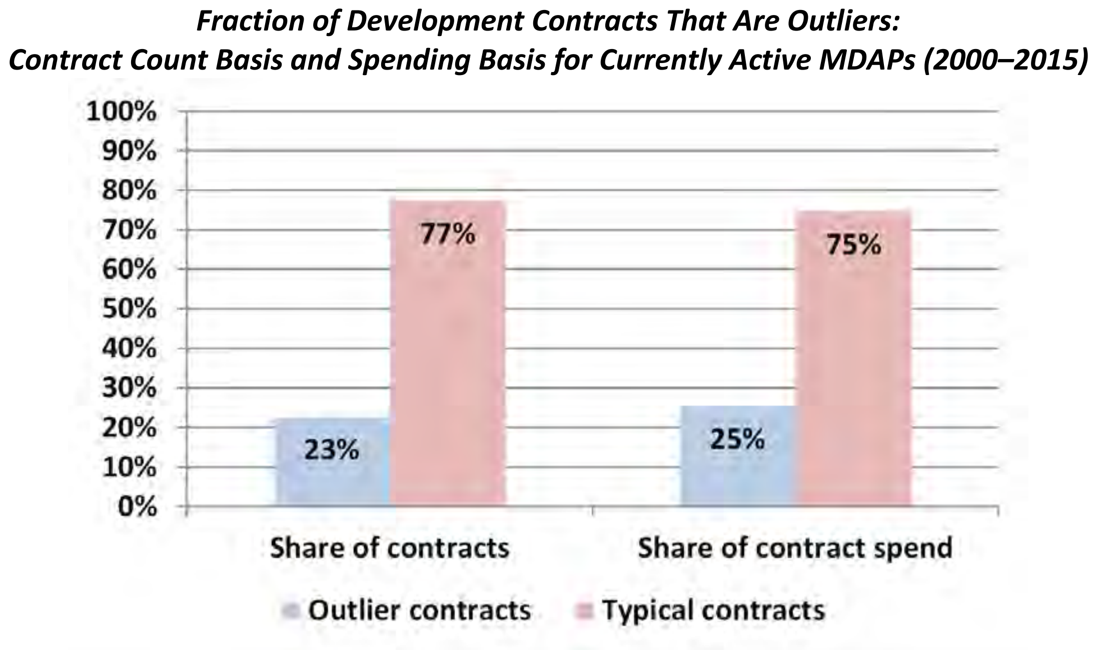
For this initial analysis, we looked only at development contracts. We first used statistical tests and analysis to select the outlying contracts.[25] After careful analysis, we found 27 outlier contracts out of 119 major development contracts from MDAP programs that were active (providing SARs) at the beginning of FY 2015, were relatively mature, and had started since FY 2000 (with the exception of SBIRS High, which started in FY 1997).

Fraction of Contracts That Are Cost Outliers

Analysis of the share of contracts in our dataset that were outliers versus the share that were typically performing contracts is shown below in Figure 2-20. Shares also are compared by the proportion of contract spending in our contract dataset taken up by outliers and typically performing contracts. About 23 percent of contracts by contract count and 25 percent in terms of spending (dollars) are cost-growth outliers. Interesting, the previously discussed Nunn-McCurdy breach rate for programs executing between 1997 and 2015 (see Table 2-5) was similar at 30 percent. Here our contract analysis is very comparable to program analysis results.

[25] The four standard statistical tests were performed on a trend regression. These were the studentized residual test, the leverage test, Cook's Distance test, and the DFFITS test.

Figure 2-20. Program Contract Cost Outliers: Development



Attributes of Cost Outliers

We then analyzed EV data and program office reports in the SARs to identify the reasons cited for cost growth seen over time in the data. These explanations fell into the following 11 categories.

- Attributes primarily resulting from contractor performance:
    - **Systems engineering** issues were most typically referred to in the SAR text as added costs arising from integration problems, or from problems with the sequencing and synchronizing of developmental activities.
    - **Contract management** problems were sometimes determined by looking at the EV data for evidence that costs-over-target may have been rolled into the contract baseline (cost target). Also, reports sometimes referred to difficulties getting EV reporting started properly, gaps in reporting, or late submission of work breakdown structure. All these deficiencies make managing the effort difficult and ultimately costly.
    - **Underestimating cost** (for a known amount of work) was often specifically mentioned in the reports. Here the established target costs turned out to be too low. These are often due to *faulty framing assumptions*.[26]
    - **Rework** often was attributed as an explanation for cost growth. This occurred when work had to be redone because the initial quality or performance was insufficient.

[26] See earlier definition of framing assumptions on p. xxxvi.

- Other attributes:
  - **Spiral-like acquisition strategies** that made cost control and management very difficult and in these cases were associated with high cost growth.
  - **Unstable engineering or system requirements** meant that during contract execution new work was added by contract modification beyond what was initially envisioned when the contracted effort began. In many cases the SARS specifically cited this attribute in addition to evidence in the EV data. These are usually dominated by detailed systems engineering requirement and specification changes to meet usually stable KPPs and KSAs, but there are few instances of KPP and KSA changes.[27] There are also cases of ambiguous KPPs and KSAs (e.g., "Net Centric") that therefore become moving targets during program execution and thus lead to work content growth.
  - **Design changes, restructuring, or replanning** sometimes were cited as reasons for excessive cost growth.
  - **Concurrency**, meaning the overlap of development and early production activities, was sometimes cited as a cause of cost growth.
  - **Acquisition/operations overlap** occurred during development as prototypes or early versions of the developing item were used in an operational or wartime environment. This sometimes contributed to program flux and cost growth.
  - **Beginning as an ACAT II program** meant the program started with less-intensive oversight early in its design and execution. This was sometimes cited as an explanation for cost growth.
  - **UCA** meant the contract was started before the final work-content plan was resolved. This was sometimes correlated with high cost growth in contract performance.

Table 2-6 below summarizes the attributed explanations of extreme cost growth for each of the outlier contracts examined, with further details for each in Appendix A. Contracts are identified by the program they support and often have additional qualifiers to help identify which contract was examined in the case of multiple major development contracts for the same MDAP. Among the outlier development contracts, there was no statistical difference between the performance of smaller and larger contracts (in dollar size terms). The total cost growth from the initially negotiated contract cost is noted for each contract in the last column. All cost growth calculations are in then-year dollars (i.e., unadjusted for inflation).

Note that this analysis is not equivalent to the PARCA root-cause analysis process whose results are discussed starting on p. 23. That process is much more intensive and, in particular, takes great care to distinguish proximal causes from root causes (i.e., it examines which causes lead to other causes). Thus, these attributes may be a mix of proximal and root causes.

[27] See, for example, GAO (2015c).

Table 2-6. Attributes of Program Contract Cost-Growth Outliers: Development

Attributed Explanations of Extreme Total Cost-Growth Outliers in Major MDAP Development Contracts (2000–2015)

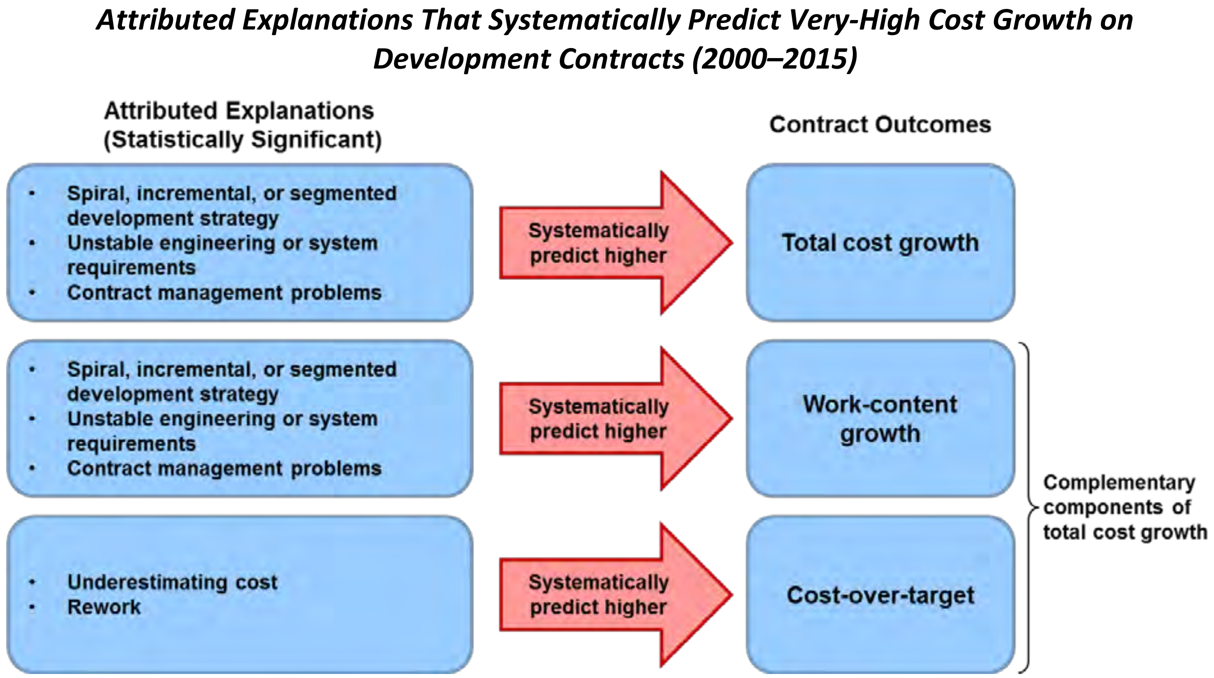
	Contractor Performance												
Contract:	Systems Engineering	Contract Management Problems	Underestimated cost (for given work)	Rework	Spiral, incremental, or segmented development strategy	Unstable engineering or system requirements	Design change, replan, or restructure	Concurrency	Acquisition/Operations overlap	Began as ACAT II	UCA	Spend ($M TY)	Contract Total Cost Growth
PIM	x	x	x			x	x			x		219	**253%**
JTN (JTRS Cluster 1)	x			x		x	x					599	**329%**
MQ-1 (SDD Ext)			x			x	x	x	x	x	x	534	**185%**
JTN (SRW)			x	x		x		x				134	**231%**
TMC (CPOF)					x	x		x	x	x		136	**736%**
GCSS-A		x	x		x	x						969	**1346%**
VTUAV MQ-8 Fire Scout						x	x			x	x	503	**466%**
LCS-1 (DD&C)			x	x			x					538	**228%**
LCS-2 (DD&C)	x		x	x			x					645	**216%**
H-1 Upgrade (ZBN)			x	x	x	x						85	**317%**
CVN 21 CP		x				x	x					2,633	**1805%**
DDG 1000 (BIW)			x			x	x					2,575	**283%**
CVN 79 CP						x						3,306	**1078%**
DDG 1000 (NG)	x		x	x		x	x					1,644	**364%**
G/ATOR	x	x	x			x	x			x		449	**400%**
AIM 9X SIP		x				x						89	**365%**
Global Hawk SDD		x	x	x	x	x	x	x	x		x	1,742	**1599%**
SBIRS High SDD	x		x			x	x					9,412	**434%**
AEHF SDD	x		x			x	x					6,965	**153%**
MQ-9 (Tgt'ing Imp)	x	x				x	x					99	**797%**
MQ-9 (SDD bridge)			x	x		x	x	x				155	**327%**
RQ-4A/B GSR	x		x		x	x						76	**682%**
MPS Inc 4					x	x						15	**493%**
FAB-T	x	x	x	x		x	x					1,585	**533%**
Pueblo CAPP		x	x	x	x	x	x				x	2,133	**1116%**
Blue Grass CAPP		x	x	x	x	x	x				x	2,880	**2197%**
Aegis SM-3 Block IIA			x			x						1,450	**918%**

NOTE: Note that this analysis is not equivalent to the PARCA root-cause analysis process, so these attributes may be a mix of proximal and root causes. Total cost growth includes any work-content changes. See Appendix A for further details on the attributed explanations of cost growth for these outliers. PIM = Paladin Integrated Management. JTN = Joint Tactical Networks. SRW = Soldier Radio Waveform. TMC = Tactical Mission Command. CPoF = Command Post of the Future. DD&C = Detail Design and Construction. ZBN = Z-variant, build new. The CVN 21 was later designated CVN-78. BIW = Bath Iron Works. NG = Nothrop Grumman. G/ATOR = Ground/Air Task Oriented Radar. AIM = Air Intercept Missile. AEHF = Advanced Extremely High Frequency. GSR = Ground-Station Re-architecture. MPS = Mission Planning System. FAB-T = Family of Advanced Beyond-line-of-sight Terminals. CAPP = Chemical Agent-Destruction Pilot Plant. SM = Standard Missile.

Predictors of Outlier Cost Growth

Figure 2-21 shows the statistical analysis of the attributed explanations of extreme cost growth in MDAP development contracts from Table 2-6. These results specify the statistically significant predictors of extreme cost growth in these outliers. The same three attributes (spiral-like acquisition strategies, unstable engineering or system requirements, and contract management problems) were statistically significant in predicting higher total cost growth and work-content growth on these outliers. This is not very surprising since the latter is usually the dominant component of the former. Two different attributes (rework and underestimating cost) were statistically significant in predicting higher cost-over-target on these outliers.

Figure 2-21. Systematic Attributes of Program Contract Cost Outliers: Development



Analysis of Individual Cost-Growth Outlier Development Contracts

Appendix A starting on p. 137 provides detailed analysis of each of these 27 cost outlier MDAP development contracts. These explanations are largely based on SARs and the evidence in the EV reports.

Cost Growth Elements in Outliers: Development

We also compared the percentile-by-percentile weighted distributions of outlier and typical development contracts for the two measures that sum to total contract cost growth: work-content growth and cost-over-target.

Figure 2-22 shows outlier contracts had much higher cost growth as a result of work-content growth (i.e., adding work to the contracts through contract modifications). The differences were statistically significant and substantial.

Figure 2-22. Program Contract Work-Content Changes: Development

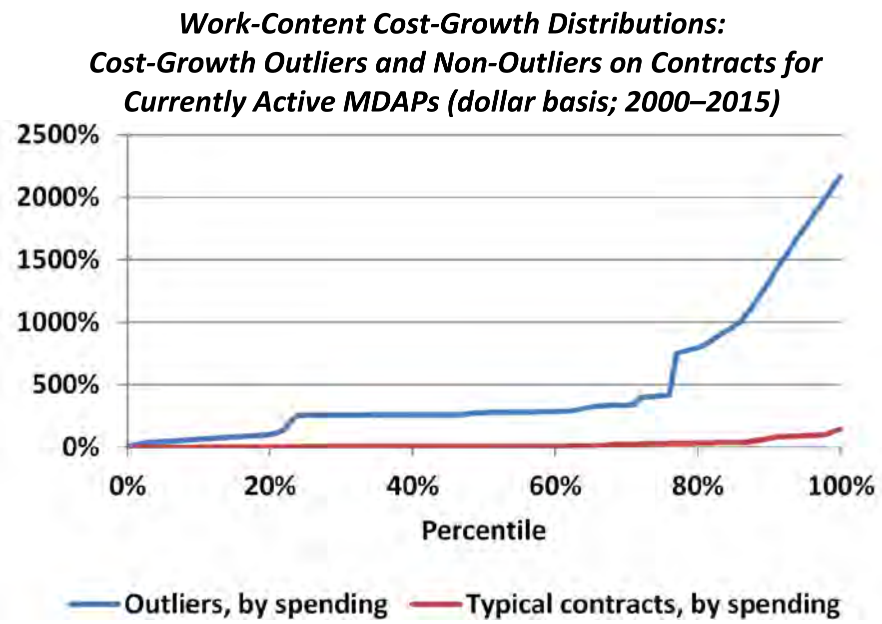
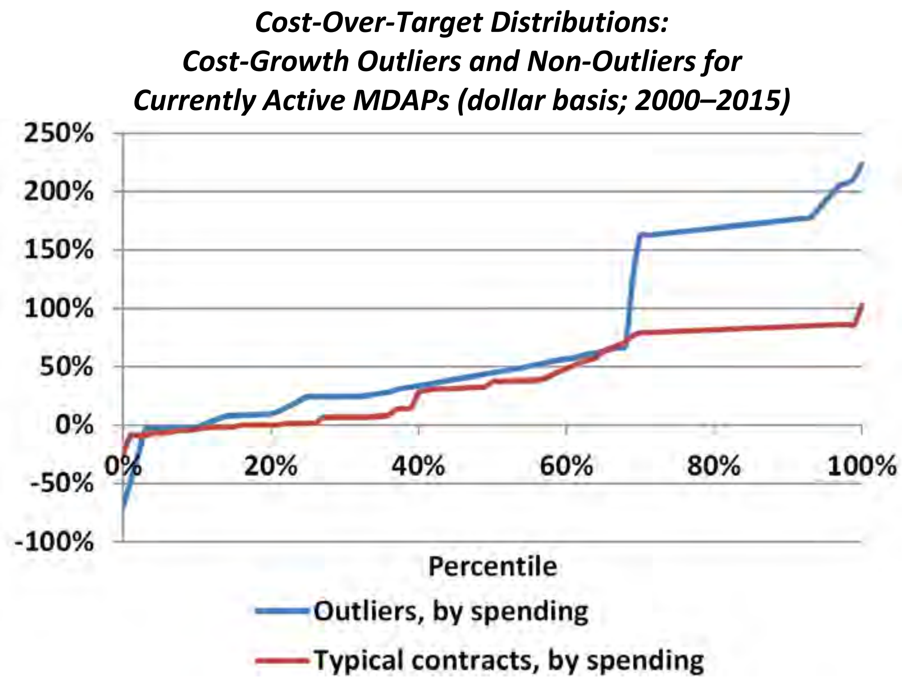


Figure 2-23 compares the percentile-by-percentile weighted distributions of cost-over-target (i.e., cost growth over cumulative target costs) for outlier development contracts and typical development contracts. The outlier contracts had much higher cost growth over cumulative target cost, especially in the last one-third of the distribution. The differences were statistically significant and substantial. Thus, outliers have high cost growth due to both work-content growth and cost-over-target.

Figure 2-23. Program Contract Costs Related to Targets: Development



Cost Outliers by Commodity: Development

We now examine whether major MDAP contracts for certain commodities are more prone to extremely high cost growth.

Figure 2-24 displays the raw number of outliers in each commodity category. In our dataset, lead-ship, C3I, and UAV development contracts were most numerous in the sample of outliers.

Figure 2-24. Program Contract Cost Outlier Count by Commodity: Development

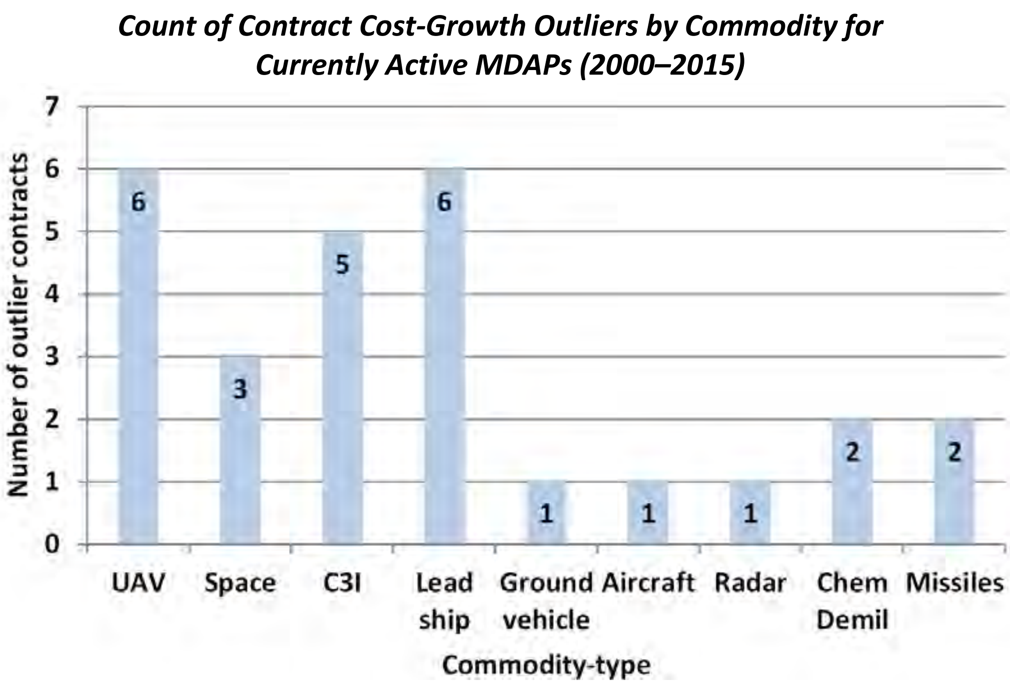
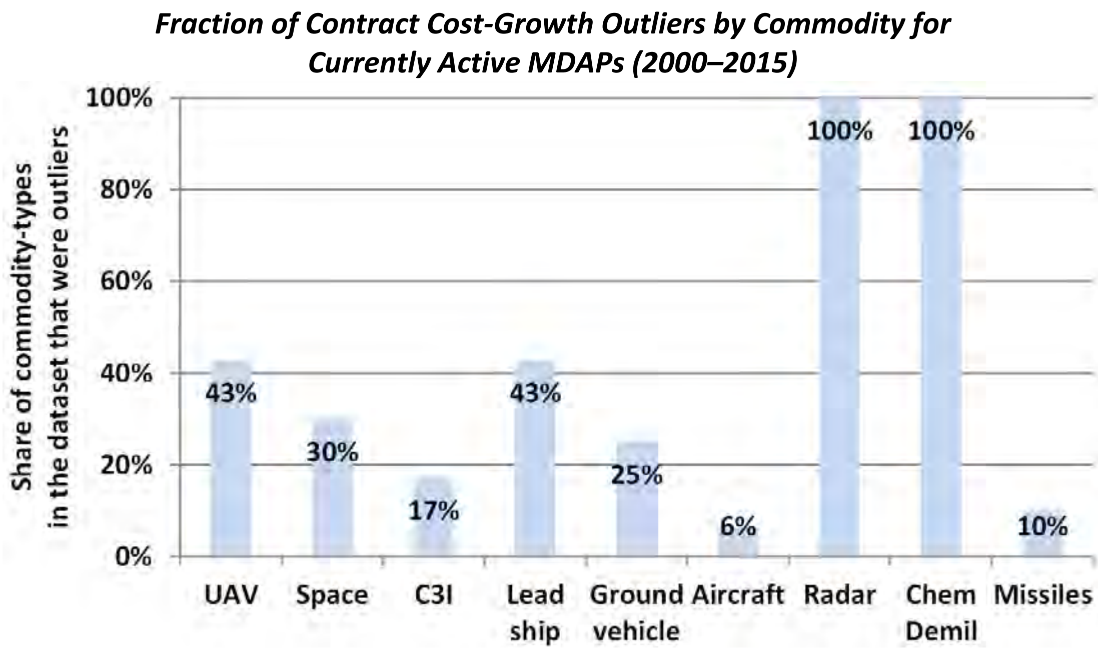


Figure 2-25 displays the fraction of statistical outliers in each commodity type. The largest shares of commodity-types in the dataset that were outliers were: UAVs, Space, lead ships, and ground vehicles. The 100 percent shares for chemical demilitarization (Chem-Demil) and radar were due to their very low number in the overall dataset (only one radar contract and two Chem-Demil contracts).

Caution is warranted in comparing these data to program-level outlier data such as the fraction of MDAP commodities that breach (recall Table 2-5 on p. 28). The number of major contracts per MDAP can differ widely between programs, and the samples include different periods. Thus, for example, while some commodities show higher outlier frequency in both (e.g., Chem-Demil), others do not (e.g., helicopters). Also, the Nunn-McCurdy breaches go back to 1997 and thus involve programs (and underlying contracts) much further back in time while this outlier dataset only examined contracts of currently active MDAPs (as of early FY 2015) with start dates back only to 2000.

Figure 2-25. Program Contract Cost Outlier Rate by Commodity: Development



Contract Price and Schedule Growth by Commodity: Development

Table 2-7 compares the price growth between commodities for an updated set of major MDAP contracts since FY 2000 (plus a SBIRS High contract, which started in FY 1997). For this analysis we controlled for inflation, contract maturity, and contract size in terms of spending. Hence all percentiles in the performance comparisons have been weighted by spend rather than being shown on a contract basis.

As a basis for comparison, the quartile levels for each commodity can be compared against those for the entire dataset of 120 contracts in the first column. These percentile-by-percentile distributional comparisons to the baseline show that space, C3I, and UAV contracts have systematically *higher* price growth than the set of combined contracts shown on the left; the differences are statistically significant. On the other hand, aircraft, ship, missile, and munitions contracts have systematically *lower* price growth than the set of combined contracts shown on the left, and these difference are also statistically significant.

Table 2-7. Program Contract Price Growth by Commodity: Development

Price Growth Since Original Award on Development Contracts for Currently Active MDAPs by Commodity (adjusted for inflation; weighted by spending; 2000-2015)

	Combined development contracts	Aircraft	Ships	Space	C3I	Missiles	UAV	Munitions
75th percentile	90%	39%	235%	163%	169%	27%	775%	-1%
Median	**40%**	**36%**	**44%**	**86%**	**67%**	**15%**	**109%**	**-5%**
25th percentile	20%	8%	-1%	48%	23%	-15%	92%	-8%
IQR (%-points)	*70%*	*31%*	*237%*	*116%*	*146%*	*41%*	*683%*	*7%*
N	*120*	*16*	*14*	*10*	*30*	*20*	*14*	*7*
Relative Growth		***Lower***	***Lower***	———	***Higher***	***Lower***	***Higher***	***Lower***

NOTE: Changes in price reflect the net changes in work content, cost-over-target, and final margins. Noted differences are statistically significant.

Similarly, commodity comparisons for schedule growth in development contracts are shown in Table 2-8. Here aircraft, and UAV contracts had systematically *higher* schedule growth in development than the set of combined contracts shown on the left; the difference was statistically significant. On the other hand, ship, space, missile, and munitions contracts had systematically *lower* schedule growth rates per contract than the combined set shown on the left; this difference was also statistically significant.

Table 2-8. Program Contract Schedule Growth by Commodity: Development

Schedule Growth Since Original Award on Development Contracts for Currently Active MDAPs by Commodity (weighted by spending; 2000–2015)

	Combined development contracts	Aircraft	Ships	Space	C3I	Missiles	UAV	Munitions
75th percentile	72%	71%	78%	79%	105%	39%	82%	0%
Median	**43%**	**69%**	**36%**	**35%**	**50%**	**0%**	**56%**	**0%**
25th percentile	12%	24%	0%	8%	2%	0%	51%	0%
IQR (%-points)	*60%*	*48%*	*78%*	*71%*	*102%*	*39%*	*31%*	*0%*
N	*120*	*16*	*14*	*10*	*30*	*20*	*14*	*7*
Relative Growth		***Higher***	***Lower***	***Lower***	———	***Lower***	***Higher***	***Lower***

NOTE: Noted differences are statistically significant.

CYCLE TIMES

Warfighting capabilities must not only have the needed technical performance but must be delivered in a timely fashion to address operational threats. Thus, the acquisition system must be responsive in time in addition to addressing cost and technical performance. Cycle time—the length of time between the identification and fielding of a need—therefore continues to be an area of concern in our BBP initiatives and elsewhere.

Cycle Time and Process Streamlining

GAO (2015a) recently reported on their case study of the review processes and associated effort for 15 MDAPs. To provide some context, Figure 2-26 illustrates the basic document flow for a review, with the four dark boxes being the decision makers in the chain-of-command (i.e., the PM, the PEO, the SAE, and the DAE).

Figure 2-27 details the various staff functions associated with reviewing Air Force acquisition-strategy documents—one of the critical program elements. The number of offices illustrates the diverse aspects that bear on a program's acquisition strategy and many of the other aspect of the program under review. Nine program offices surveyed by GAO (2015a) reported an average of 55 briefings with these various staff offices and decision makers in preparation for MS B review and approval. More specifically, GAO (2015a, p. 13) reported that the Air Force cases averaged over 12 months to review the programs' acquisition-strategy documents.

In response, the DoD has initiated a number of initiatives to minimize the effects on cycle time from these review processes. The DAE has directed that all acquisition-strategy documents will remain in staffing in the Office of the Secretary of Defense (OSD) for no more than 14 days. An Electronic Coordination Tool (ECT) has been implemented to enable parallel reviews for the elements shown in Figure 2-27 while enforcing limitations on staff review times. To date, 10 acquisition-strategy documents have completed their reviews and approval, with a median of 90 days and a mean of 105 days from initiation to approval. Thus, ECT reviews have thus far reduced these review times by three-quarters compared to the data from GAO's survey (2015a, p. 13). However, given the small sample sizes in both the ECT data and GAO's survey, caution is warranted in reaching final conclusions regarding the success of ECT and the other steps that have been taken.

Also, the new DoDI 5000.02 (AT&L, 2015a) encourages tailoring each program's review approach and required milestone reviews to streamline programs. It includes example models and encourages further adaptation to meet each program's specific situation.

In addition, a BBP initiative to pilot streamlined hands-on reviews has been initiated. It includes a hands-on, onsite data review in the 2–3 week period preceding a milestone decision instead of the current document-based approach to milestone reviews and decisions. Pilots are underway as we experiment with this approach.

As these initiatives are executed, analysis will be conducted to measure their effectiveness, learn from these efforts, and adjust policies to ensure best practices can be adopted when appropriate.

Figure 2-26. DoD Program Review Structure and Levels

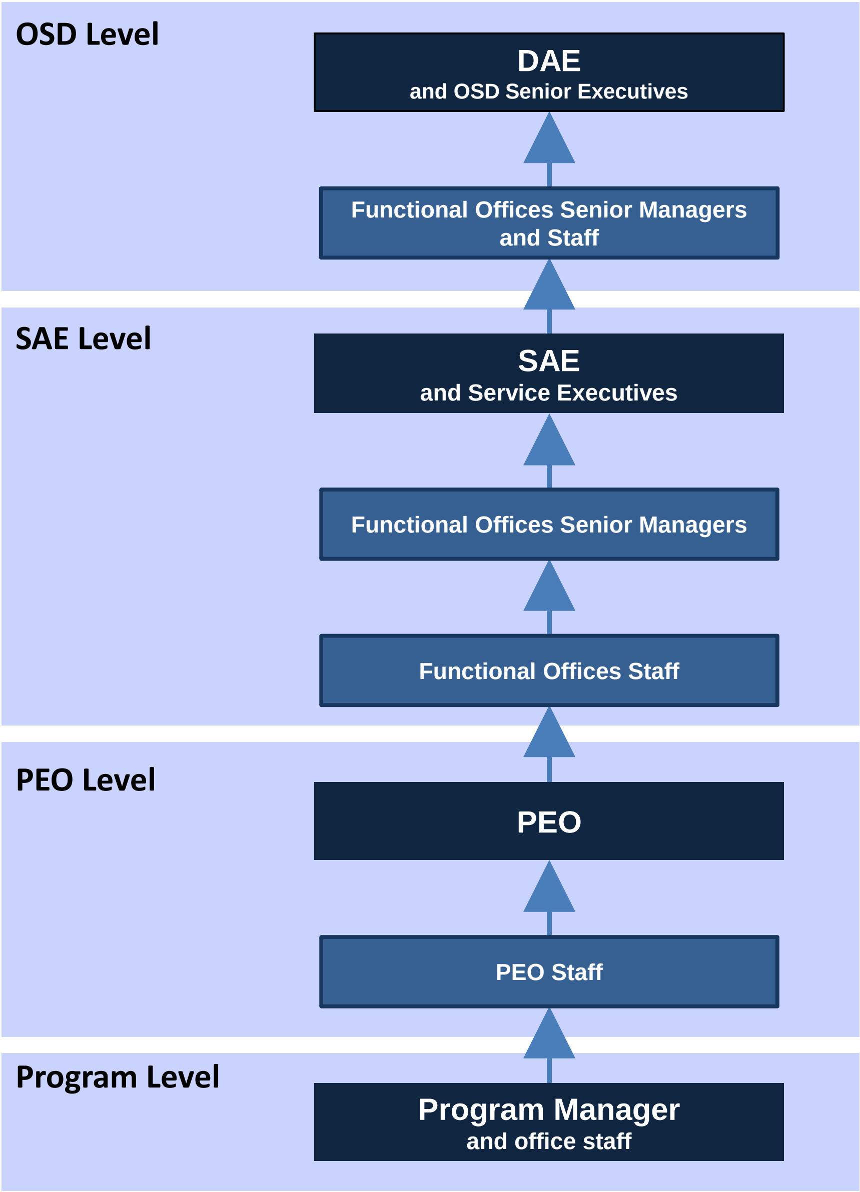


Figure 2-27. Organizations Typically Involved in the Review of an Air Force Acquisition-Strategy Document

Office of the Secretary of Defense (OSD) Level

- Defense Acquisition Executive
- Vice Chairman of the Joint Chiefs of Staff
- Under Secretary of Defense (Policy)
- Under Secretary of Defense (Comptroller)
- Under Secretary of Defense (Personnel & Readiness)
- Under Secretary of Defense (Intelligence)
- Chief Information Officer
- Director, Operational Test & Evaluation
- Director, Cost Assessment and Program Evaluation
- Director, Acquisition Resources & Analysis
- Principal Deputy Under Secretary of Defense (Acquisition, Technology, & Logistics)
- Assistant Secretary of Defense (Acquisition)
- Assistant Secretary of Defense (Logistics & Material Readiness)
- Deputy Under Secretary of Defense (Installations and Environment)
- Deputy General Counsel (Acquisition & Logistics)
- Assistant Secretary of Defense (Research & Engineering)
- Deputy Assistant Secretary of Defense, Strategic & Tactical Systems
- Deputy Assistant Secretary of Defense, Space & Intelligence
- Deputy Assistant Secretary of Defense, Communication, Command, and Control Cyber
- Director, National Geospatial-Intelligence Agency
- Deputy Director, Cost Assessment
- Director, Defense Pricing
- Director, Systems Engineering
- Director, Developmental Test & Evaluation
- Deputy Assistant Secretary of Defense, Manufacturing & Industrial Base Policy
- Director, International Cooperation
- Director, Performance Assessment and Root Cause Analysis
- Assistant Secretary of Defense (Legislative Affairs)
- Director, Defense Procurement and Acquisition Policy
- Assistant Secretary of Defense (Operational Energy Plans and Programs)

Service Acquisition Executive (SAE) Office Level

- Assistant Secretary of the Air Force for Acquisition (Service Acquisition Executive)
- Assistant Secretary of the Air Force Installations & Environment
- Air Force Logistics, Installations, & Mission Support
- Air Force Operations, Plans, & Requirements
- Air Force Intelligence, Surveillance, & Reconnaissance
- Air Force Financial Management & Comptroller
- Air Force Test & Evaluation
- Assistant Secretary of the Air Force Small Business Programs
- Assistant Secretary of the Air Force Chief Information Officer
- Assistant Secretary of the Air Force Test & Evaluation (Policy and Programs)
- Air Force Operations, Plans, & Requirements (Operational Capability Requirements)
- Air Force Logistics, Installations & Mission Support (Logistics)
- Assistant Secretary of the Air Force Installations & Environment (Logistics)
- Air Force Intelligence, Surveillance, & Reconnaissance (Strategy, Plans, Doctrine & Force Development)
- Assistant Secretary of the Air Force Chief Information Officer (Policy & Resources)
- Assistant Secretary of the Air Force Deputy General Counsel for Acquisition
- Air Force Financial Management and Comptroller Deputy Assistant Secretary (Cost and Economics)
- Air Force Financial Management and Comptroller Deputy Assistant Secretary (Budget)
- Assistant Secretary of the Air Force Directorate of Science, Technology & Engineering
- Assistant Secretary of the Air Force Directorate Management Policy & Program Integration
- Assistant Secretary of the Air Force Directorate of Contracting
- Air Force Acquisition Capability Directorate

Program Executive Office

- Program Executive Officer
- Deputy Program Executive Officer
- Functional Staff: contracting, logistics, finance, and engineering
- Program Executive Officer Execution Group

Source: GAO (2015a) presentation of DoD information.

MAIS Cycle Times

Figure 2-28 shows MAIS cycle time since 2011 as the number of years for each MAIS from MS B (or the date of FFO) to the FDD in the original estimate (i.e., the first MAR for the MAIS). Similar to last year's results, MAIS programs with MS B or FFO before 2009 had a median cycle time of five years; since then, the estimated median cycle times dropped to just above three years.[28] In other words, before 2009, half of the MAIS were planned with cycle times longer than five years. Since 2009, that estimate has dropped significantly, and no program is planned on taking longer than five years since MS B or FFO. This appears to be a direct result of the legal requirement for Critical Change Reports if the five year period is breached. Whether the DoD achieves these estimates and whether this improves acquisition performance has yet to be determined. Recall that the median schedule growth on all currently reporting MAIS since their original estimate is about three months (see Figure 2-10).

The optimal cycle time cannot be predetermined absent information on the system in question. In some cases, long cycle times may be a concern given the pace of information technology advancement. On the other hand, setting arbitrary schedule deadlines may incentivize undesirable management decisions and risky short cuts, causing failures to meet needs from end users and possibly increasing costs for subsequent upgrades.

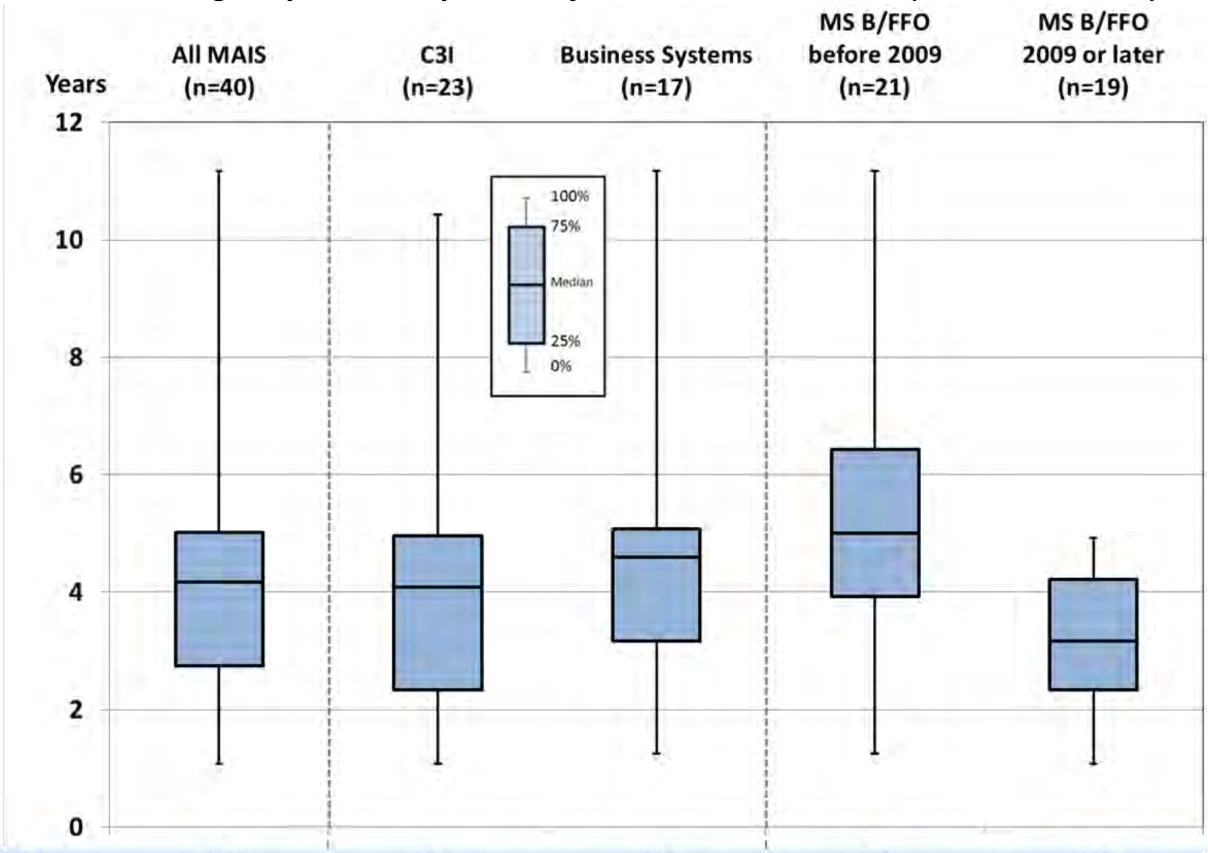
Contract Cycle Time

Contract cycle times on development and early production MDAP contracts are discussed later in the contract trends section starting on p. 70, where we discuss results indicating recent downward trends in contract cycle time.

[28]Many MAIS increments have a MS B but not MS A, so we have more consistent data using MS B. For comparison, 5 years since MS A or FFO (not MS B as shown here) to FDD is the statutory threshold beyond which a certification of variance is required. The end points of the 5-year period have changed over the years, but it is currently from MS A or Preferred Alternative Decision (PAD) to FDD.

Figure 2-28. Program Length: Information Systems

Active MAIS Originally Planned Cycle Time from MS B or FFO to FDD (2011–2014 MARs)



	All MAIS	C3I	Business Systems	MS B/FFO before 2009	MS B/FFO 2009 or later
Median	**4.2**	**4.1**	**4.6**	**5.0**	**3.2**
IQR	2.3	2.6	1.9	2.5	1.9

NOTE: Original estimates are those in the MAIS' first MAR. Included are the latest data on programs that appeared in at least one MAR from 2011 through 2014. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The IQR is the difference between the 75th and 25th percentiles. Immature programs that have not completed at least 30 percent of their original EMD schedule time were excluded to help control for maturity.

Cost and Schedule Performance: Production

Program Production Funding Growth (Quantity Adjusted)

Now examining production at the program level, the following figures summarize the unit procurement funding growth across the MDAP portfolio from the original MS B baseline and biennial changes. These use recurring unit flyaway funding data reported in the SARs and are adjusted for quantity changes since the MS B baseline.

These program-level data are for measures that (unlike PAUC and APUC) are fully adjusted for any changes in procurement quantity. These results help compare procurement unit costs at the initially estimated quantities, extrapolating data if quantities have been reduced. This approach provides a superior way of comparing what the units would have cost if we had not changed quantities by, essentially, measuring the shift in the cost-versus-quantity procurement cost curve from planned to actual.[29] In other words, we measure changes in procurement cost at the currently planned quantity to be purchased (often lower than the initial) *and* assume that the original planned quantity still was being purchased. This approach allows us to examine on a unit basis the cost of the capability to acquire those units regardless of whether we increased or decreased quantity. Of course, quantity decreases may be due to unit-cost increases, and this approach will show such cost increases clearly.

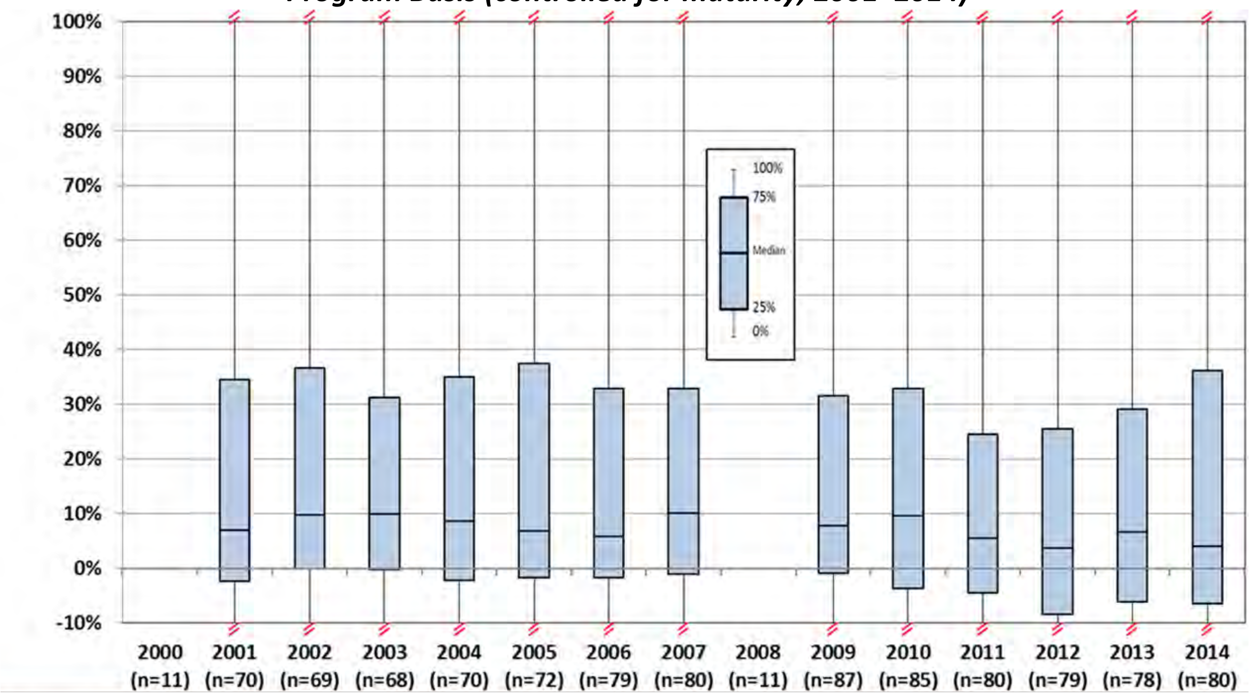
Similar to the prior RDT&E results, growth distributions in production are highly skewed, with arithmetic means higher than the medians. As noted elsewhere for the contract-level data, the overall magnitudes of production funding growth are not nearly as large as those for RDT&E. There also is considerable variability in the production funding growth across the MDAP portfolio.

Figure 2-29 shows quantity-adjusted procurement cumulative unit-funding growth over original MS B baseline for each year's MDAP portfolio on a program basis (controlled for program maturity). As with last year's results, growth has been statistically flat since at least 2001. Figure 2-30 summarizes new analysis showing that on a dollar basis (i.e., weighted by spend), the quantity-adjusted unit-funding growth is statistically lower since 2011 than in prior years (2001–2010), and 2009 and 2010 had been dropping as well. Given these are based on budget data, one would be concerned that this might just reflect budgetary pressures on programs, but these procurement measures are adjusted for any quantity changes, so regardless of the budget changes they should be reflecting actual reductions in unit costs.

[29]This basic approach for quantity adjustment is one of the standard techniques employed by the cost analysis community—see, for example, the discussions in Hough (1992), Arena et al. (2006, pp. 5–6), and Younossi et al. (2007, pp. 13-14).

Figure 2-29. Program Cost-Related Performance: Procurement

Cumulative Growth Over Original MS B Baseline of Active MDAP Planned Total (From Start to Completion) Quantity-Adjusted Unit-Procurement Recurring-Flyaway Funding: Program Basis (controlled for maturity; 2001–2014)



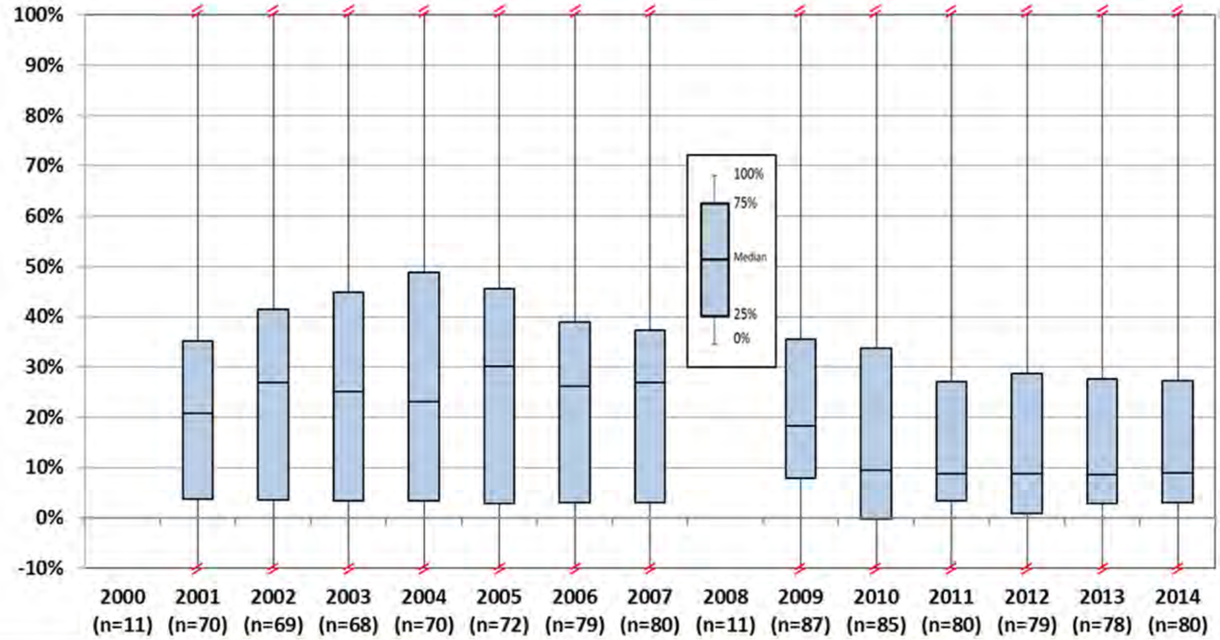
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Median		**7%**	**10%**	**10%**	**9%**	**7%**	**6%**	**10%**		**8%**	**10%**	**6%**	**4%**	**7%**	**4%**
IQR		37%	37%	31%	37%	39%	35%	34%		32%	37%	29%	34%	35%	35%

No trend across the years.

NOTE: This shows growth in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not included. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The IQR is the difference between the 75th and 25th percentiles.

Figure 2-30. Program Cost-Related Performance: Procurement (weighted by program dollars)

Cumulative Growth Over Original MS B Baseline of Active MDAP Planned Total (From Start to Completion) Quantity-Adjusted Unit-Procurement Recurring-Flyaway Funding: Dollar Basis (controlled for maturity; 2001–2014)



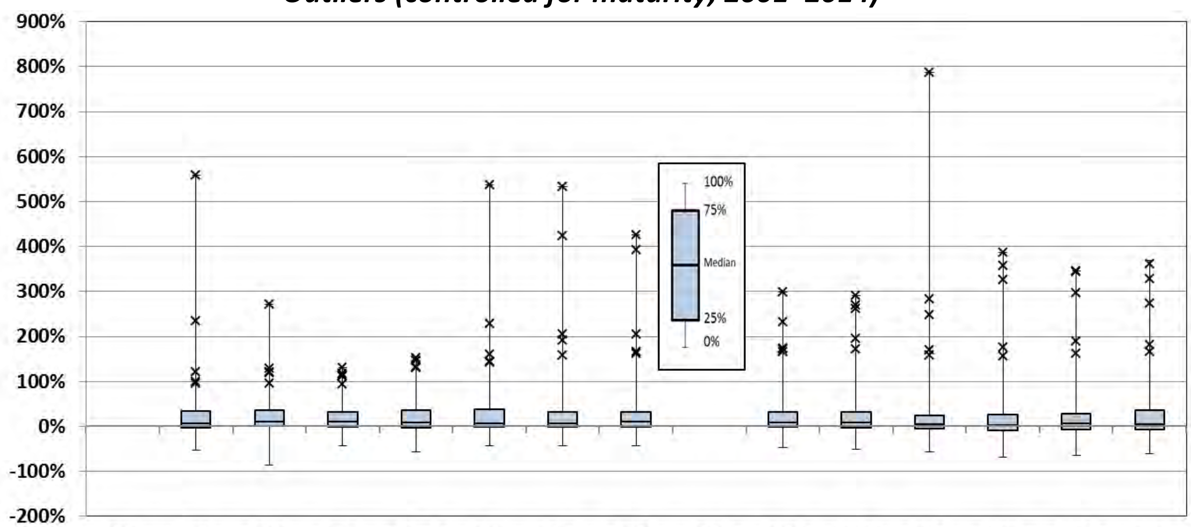
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Median		**21%**	**27%**	**25%**	**23%**	**30%**	**26%**	**27%**		**18%**	**9%**	**9%**	**9%**	**8%**	**9%**
IQR		31%	38%	42%	45%	43%	36%	34%		28%	34%	24%	28%	25%	24%
Trend										***Generally lower than before 2009.***					

NOTE: This shows growth in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not included. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The IQR is the difference between the 75th and 25th percentiles.

Figure 2-31 extends the y-axis scale to show all outliers in Figure 2-29 and Figure 2-30, and the table at the bottom identifies the five largest funding-growth programs for each year. This chart is also controlled for program maturity.

Figure 2-31. Program Cost-Related Performance Outliers: Procurement

Cumulative Growth Over Original MS B Baseline of Active MDAP Planned Total (From Start to Completion) Quantity-Adjusted Unit-Procurement Recurring-Flyaway Funding: Outliers (controlled for maturity; 2001–2014)



<table>
<tr><th></th><th>2000 (n=11)</th><th>2001 (n=70)</th><th>2002 (n=69)</th><th>2003 (n=68)</th><th>2004 (n=70)</th><th>2005 (n=72)</th><th>2006 (n=79)</th><th>2007 (n=80)</th><th>2008 (n=11)</th><th>2009 (n=87)</th><th>2010 (n=85)</th><th>2011 (n=80)</th><th>2012 (n=79)</th><th>2013 (n=78)</th><th>2014 (n=80)</th></tr>
<tr><td>Largest</td><td></td><td>SADARM Projectile</td><td>ATACMS BLK II/IIA</td><td>PAC-3 Fire Unit</td><td colspan="3">SBIRS High</td><td>NPOESS</td><td></td><td>SBIRS High</td><td>ATIRCM QRC</td><td>C-130 AMP</td><td colspan="2">EELV</td><td>AEHF</td></tr>
<tr><td>2nd Largest</td><td></td><td>ATIRCM CMWS</td><td>PAC-3 Fire Unit</td><td>GMLRS AW</td><td>EELV</td><td colspan="2">NPOESS</td><td>SBIRS High</td><td></td><td>AEHF SV 1-4</td><td colspan="2">SBIRS High</td><td colspan="2">AEHF SV 1-4</td><td>EELV</td></tr>
<tr><td>3rd Largest</td><td></td><td>PAC-3 Fire Unit</td><td>ATACMS BAT P3I</td><td colspan="2">CH-47F</td><td>GMLRS AW</td><td colspan="2">EELV</td><td></td><td>H-1 Upgrades</td><td>AEHF SV 1-4</td><td></td><td colspan="2">SBIRS High</td><td>SBIRS High</td></tr>
<tr><td>4th Largest</td><td></td><td colspan="2">GMLRS AW</td><td>EELV</td><td>GMLRS AW</td><td>EELV</td><td colspan="2">GMLRS AW</td><td></td><td colspan="2">C-130 AMP</td><td colspan="3">GMLRS AW</td><td>GMLRS AW</td></tr>
<tr><td>5th Largest</td><td></td><td>CH-47F</td><td>H-1 Upgrades</td><td>SBIRS High</td><td>PAC-3 Fire Unit</td><td colspan="2">CH-47F</td><td>C-130 AMP</td><td></td><td>GMLRS AW</td><td>EFV</td><td colspan="2">CH-47F</td><td>H-1 Upgrades</td><td>JPALS Inc 1A</td></tr>
</table>

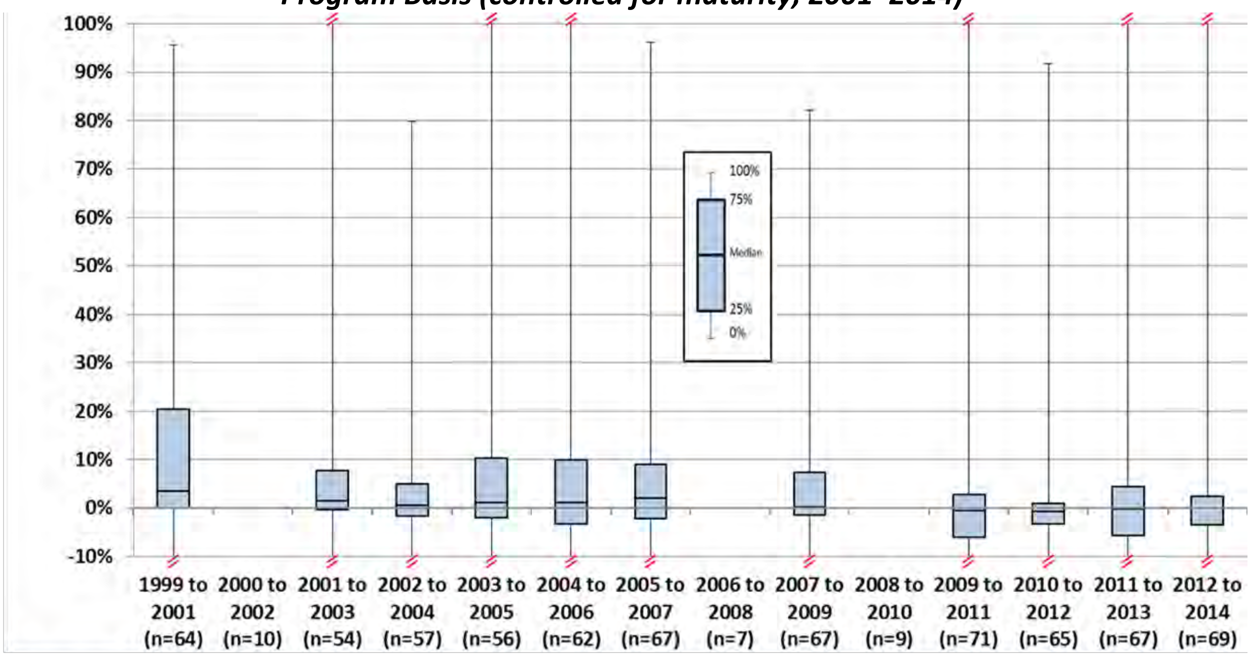
NOTE: This shows growth in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not included. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

Biennial Funding Changes: Production

Figure 2-32 shows biennial changes in total quantity-adjusted unit procurement funding (actual and planned), controlling for program maturity. Figure 2-33 shows the same but on a dollar basis. In both bases, the median unit funding growth has been zero or less since 2009. Interestingly, the median performance since 2002 is slightly lower on a dollar basis, except in the 2010–2012 period, when the median was slightly lower on a program basis. The earliest period (1999–2001) was statistically higher than nearly all subsequent periods

Figure 2-32. Program Cost-Related Biennial Performance: Procurement

Biennial Change in
Active MDAP Planned Total (From Start to Completion)
Quantity-Adjusted Unit-Procurement Recurring-Flyaway Funding:
Program Basis (controlled for maturity; 2001–2014)

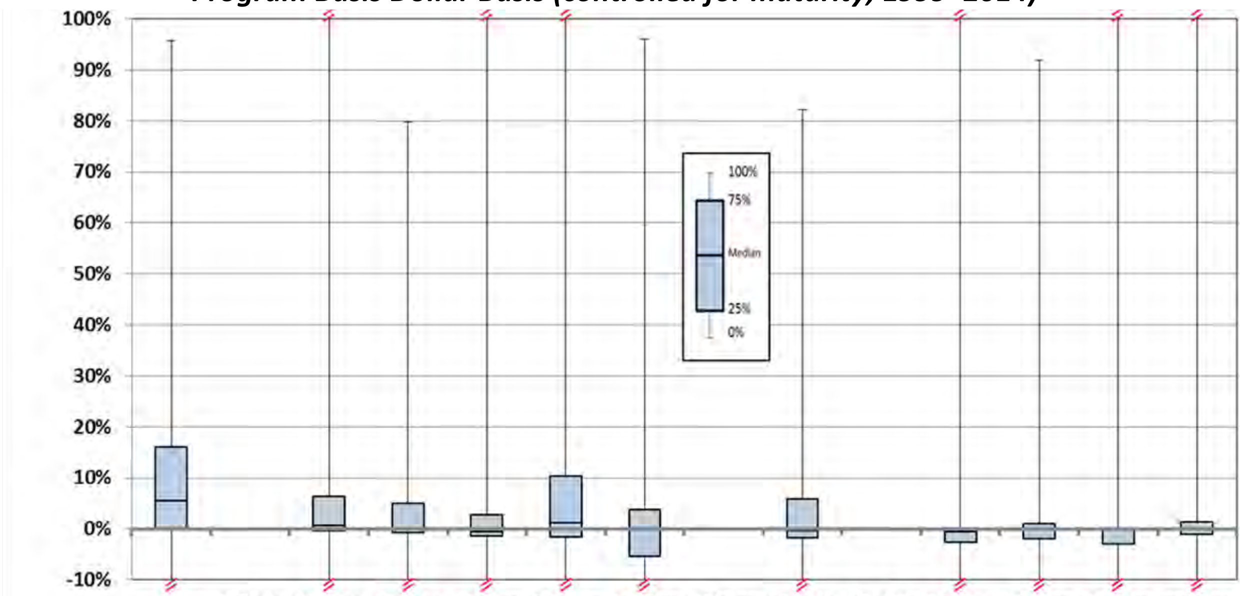


	1999 to 2001	2000 to 2002	2001 to 2003	2002 to 2004	2003 to 2005	2004 to 2006	2005 to 2007	2006 to 2008	2007 to 2009	2008 to 2010	2009 to 2011	2010 to 2012	2011 to 2013	2012 to 2014
Median	**3%**		**1%**	**1%**	**1%**	**1%**	**2%**		**0%**		**-1%**	**-1%**	**0%**	**0%**
IQR	20%		8%	7%	12%	13%	11%		9%	8%	9%	4%	10%	6%
Trends	***Higher than later periods.***										***Lower than 2001–2009.***			

NOTE: This shows biennial changes in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not included. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The IQR is the difference between the 75th and 25th percentiles.

Figure 2-33. Program Cost-Related Biennial Performance: Procurement (weighted by spend)

Biennial Change in
Active MDAP Planned Total (From Start to Completion)
Quantity-Adjusted Unit-Procurement Recurring-Flyaway Funding:
Program Basis Dollar Basis (controlled for maturity; 1999–2014)

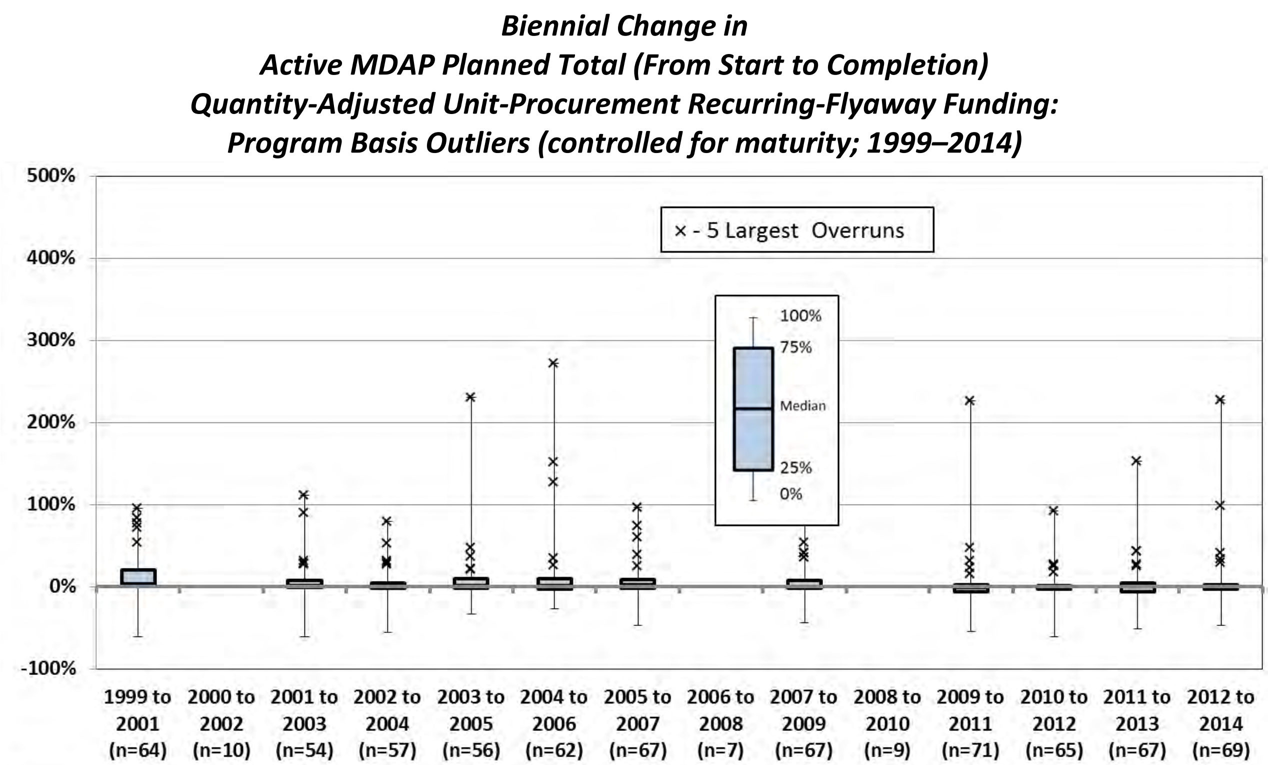


	1999 to 2001 (n=64)	2000 to 2002 (n=10)	2001 to 2003 (n=54)	2002 to 2004 (n=57)	2003 to 2005 (n=56)	2004 to 2006 (n=62)	2005 to 2007 (n=67)	2006 to 2008 (n=7)	2007 to 2009 (n=67)	2008 to 2010 (n=9)	2009 to 2011 (n=71)	2010 to 2012 (n=65)	2011 to 2013 (n=67)	2012 to 2014 (n=69)
Median	**6%**		**1%**	**0%**	**0%**	**1%**	**0%**		**0%**		**-1%**	**0%**	**0%**	**0%**
IQR	16%		7%	6%	4%	12%	9%		8%		3%	3%	3%	2%
Trend											***Lower than most years prior to 2009.***			

NOTE: This shows biennial changes in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not included. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum. The IQR is the difference between the 75th and 25th percentiles.

Figure 2-34 extends the y-axis scale to show all outliers in Figure 2-32 and Figure 2-33, and the table at the bottom identifies the five largest funding-growth programs for each year. This chart is also controlled for program maturity.

Figure 2-34. Program Cost-Related Biennial Performance Outliers: Procurement



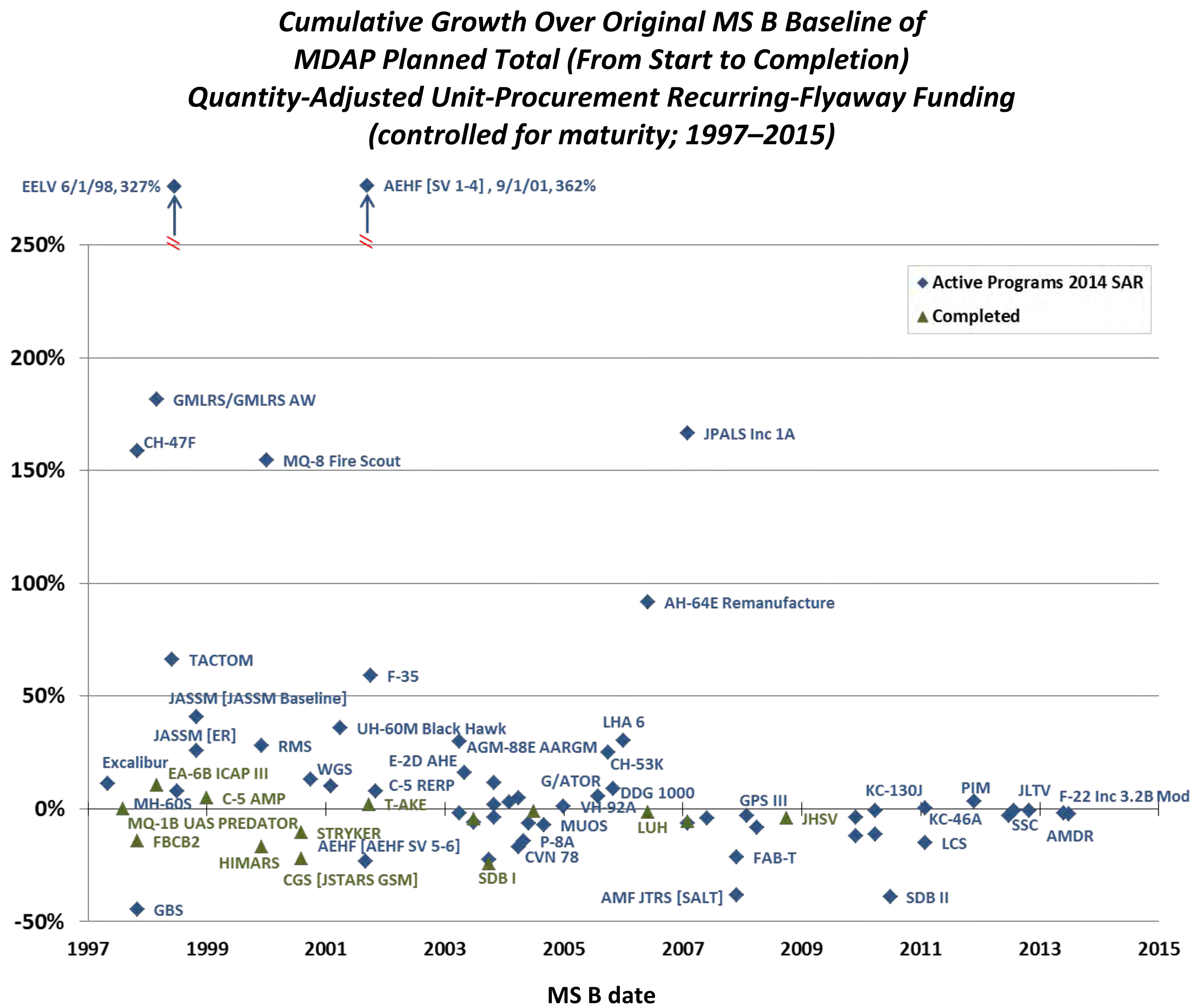
Largest	CH-47F		SBIRS High	EELV	NPOESS		C-130 AMP		B-2 RMP		C-130 AMP	JTRS HMS	JPALS Inc 1A
2nd Largest	ATIRCM CMWS		EELV	CH-47F	SBIRS High	LAND WARRIOR	EFV		AEHF SV 1-4		JTRS HMS	B-2 EHF Inc 1	VTUAV
3rd Largest	GMLRS AW		AEHF SV 1-4	SBIRS High	UH-60M	SBIRS High	NPOESS		VTUAV		Excalibur	AEHF SV 1-4	AH-64E Reman
4th Largest	MH-60R		TACTOM	AEHF SV 1-4	FCS	C-130 AMP	AEHF SV 1-4		AH-64E Reman		FAB-T	MQ-9 Reaper	AEHF SV 1-4
5th Largest	H-1 Upgrades		B-1B CMUP Computer	UH-60M	LAND WARRIOR	EFV	LHA 6		H-1 Upgrades		WGS	CH-53K	JTRS HMS

NOTE: This shows biennial changes in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not included. Boxes show second quartile, median, and third quartile; bars show first and fourth quartiles, minimum, and maximum.

Program Funding Growth by Start Date: Production

Finally as for development, we examine program procurement performance based on initiation date. Figure 2-35 plots quantity-adjusted procurement unit funding growth of all active and completed MDAPs by their original MS B date. As before, we controlled for maturity by removing newer programs.

Figure 2-35. Program Cost-Related Performance: Procurement



NOTE: This shows growth in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown.

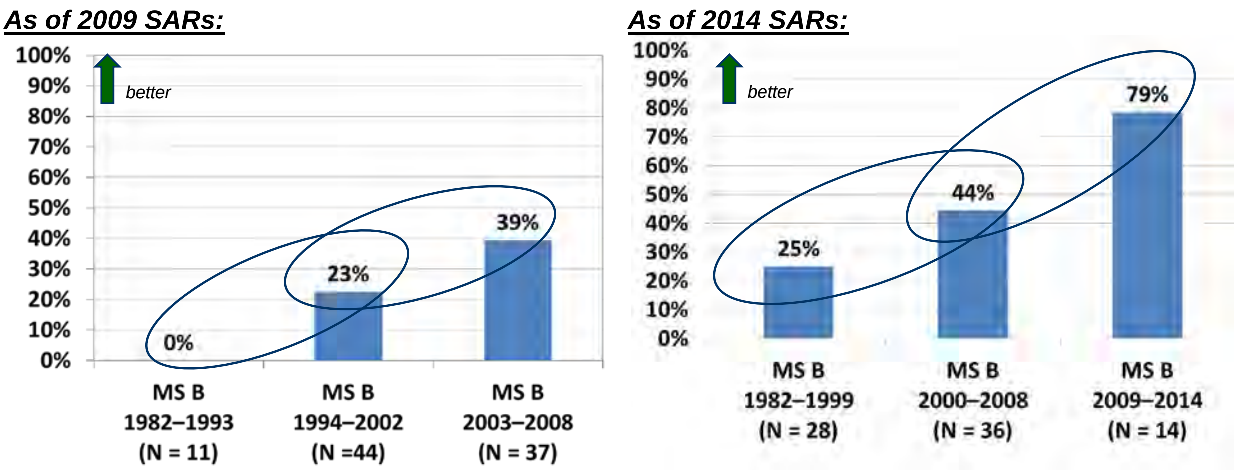
Again, there is an apparent recent improvement, but it is not statistically significant when testing back to 1997; the variation explained by any potential trend (i.e., the R^2) was too low.

However, we did find that recent programs are, in fact, improving when using a proportion test comparing programs started since 2009 to earlier programs. Figure 2-36 shows the result of separating the portfolio of active MDAPs based on program start date (i.e., their original MS B date). After deleting immature programs (i.e., those that are too new to reveal the magnitude of potential problems), we find that the proportion of active MDAPs showing quantity-adjusted unit procurement reductions (negative funding growth from original MS B baseline) is significantly higher at 79 percent for the period since 2009 than the 44 percent for the prior period 2000–2008.

To help test whether this is just a result of any remaining maturity bias, we conducted the same proportion test using earlier 2009 SARs. While the proportional differences were also significant, the percentages were lower than what we see now in the 2014 SARs. Thus, as in development, there is evidence that this is not just a maturity bias; MDAPs started since 2009 are performing better in production at controlling cost. In 2010, the DAE began requiring affordability caps on production for all MDAPs undergoing MS B reviews. In addition the Should-Cost BBP initiative was introduced in 2010.

Figure 2-36. Planned Reductions in Program Funding: Procurement

Proportions of Active MDAPs With Cost-Related Reductions Since Original MS B in Cumulative Planned Total (From Start to Completion) Quantity-Adjusted Unit-Procurement Recurring-Flyaway Funding (program basis; quantity adjusted; controlled for maturity)



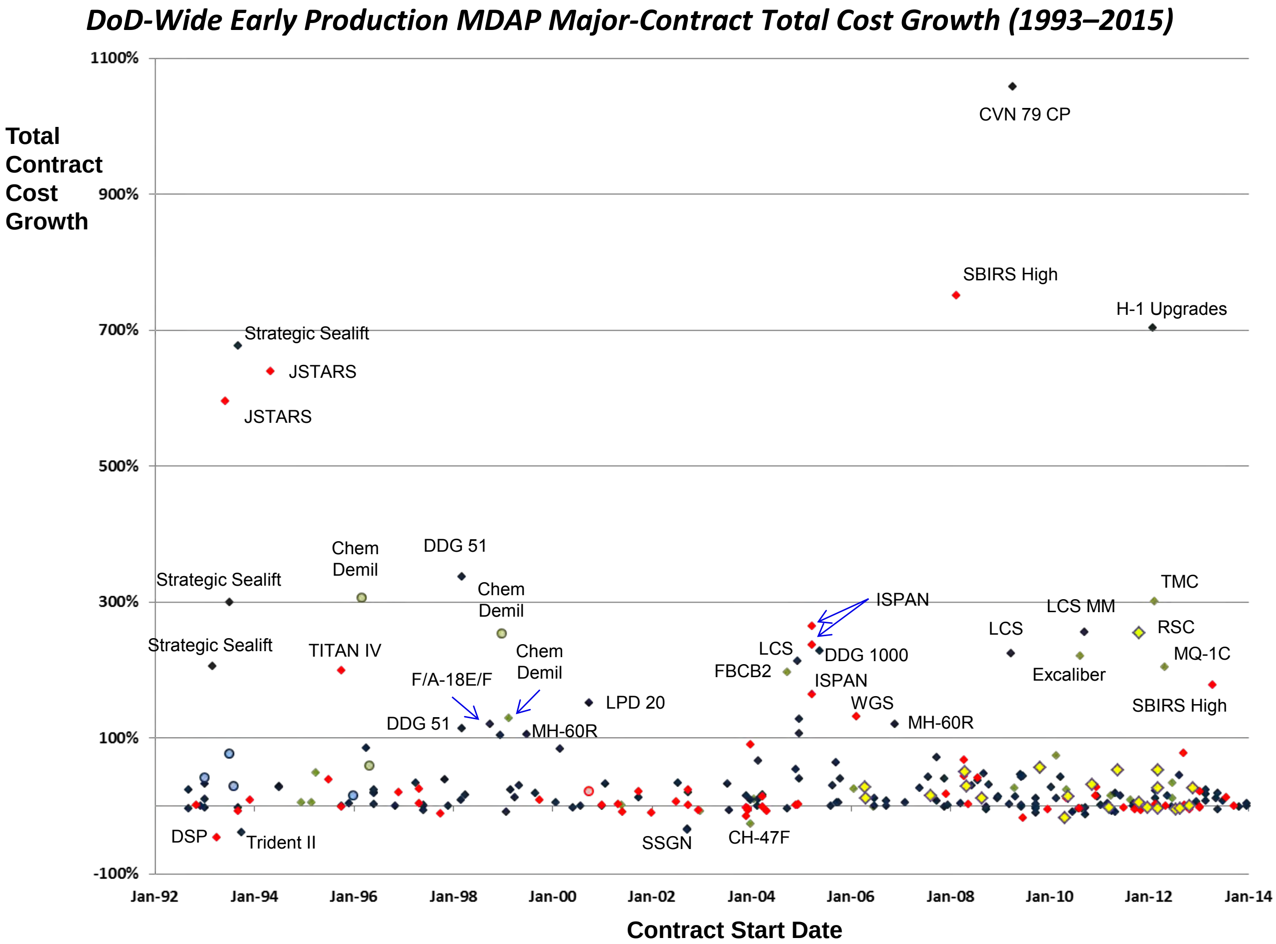
NOTE: This shows growth in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. Statistically significant differences between adjacent periods are marked with an oval. A program shows a reduction if current total unit funding is under original MS B baseline after adjusted for quantity changes since MS B baseline. Relatively new programs that have not been through at least 30 percent of their original EMD schedule are not included to help control for low maturity.

Conceptually, two behaviors different than those in development may be at work here if these are related to budget cuts instead of better program initiations. First, if quantity was reduced on an MDAP to accommodate budget reductions, we would see an increased unit funding. However, we adjusted unit funding for any quantity changes and still saw reductions. Therefore, a reduction in this case would be real. Second, if the production rate was slowed due to budget cuts, the learning curve would be slowed too. Our adjustment for quantity directly compares the initial estimated learning curve against the latest curve, so we would see this as an increase, not a decrease. The reductions we see in the data are real, and they do correlate well to the management actions taken under BBP. The data do not provide a basis for conclusively assessing cause and effect, but the trend is very encouraging.

Contract Cost Growth: Early Production

Figure 2-37 shows more than 20 years of contract data on total cost growth relative to initial contract cost targets for major MDAP early production contracts. These plots have been extensively updated since last year's report, including the addition of recent, relatively mature contracts (spending at least 30 percent of their original budget; see the discussion on avoiding maturity bias on p. 9). Outliers and selected contracts for well-known programs are identified by program name (including some duplicates for multiple large contracts for the same MDAP).

Figure 2-37. Program Contract Cost-Related Performance: Early Production



NOTE: This shows total cost growth, including any work-content changes and cost-over-target. Army programs are shown in green; Navy in blue, Air Force in red, DoD-wide in purple and yellow. Diamonds on the charts indicate growths mostly attributable to the start of the contract; circles represent significant work added later. Source cost data were reported in "then-year" dollars (unadjusted for inflation). Immature contracts that have not spent at least 30 percent of their original target cost were excluded to help control for maturity.

Contract Early Production Price and Schedule Growth by Commodity

We also compared the relative performance of early production contracts by commodity. Table 2-9 summarizes the results, controlled for inflation, contract maturity, and contract size in terms of spending. As a basis for comparison, the quartile levels for each commodity can be compared against those for the entire dataset of 159 contracts records.

Table 2-9. Program Contract Price Growth by Commodity: Early Production

Price Growth Since Original Award on Early Production Contracts for Currently Active MDAPs by Commodity (adjusted for inflation; weighted by spending; 2000–2015)

	Combined early production contracts	Aircraft	Ships	Space	C3I	Missiles	UAV
75th percentile	35%	36%	23%	188%	68%	-2%	20%
Median	**7%**	**18%**	**2%**	**7%**	**24%**	**-7%**	**13%**
25th percentile	-2%	-2%	-2%	-1%	10%	-9%	-1%
IQR (%-points)	*37%*	*38%*	*25%*	*189%*	*58%*	*7%*	*21%*
N	*159*	*33*	*48*	*13*	*16*	*11*	*18*
Relative Growth		———	***Lower***	***Higher***	***Higher***	***Lower***	———

NOTE: Changes in price reflect the net changes in work content, cost-over-target, and final margins. Ships have higher N (numbers of contracts) since they often involve separate contracts for shipyards, major ship systems, and individual ships. Noted differences are statistically significant.

Most apparent are the much lower price growth rates per contract in early production than in development. This should not be surprising. In early production, percentile-by-percentile comparisons to the entire population in the first column show that space and C3I contracts have systematically *higher* price growth and the difference is statistically significant, whereas ship and missile contracts have systematically lower price growth compared to the baseline and *the difference is statistically significant.*

Similarly, commodity comparisons for schedule growth in early production contracts are shown in Table 2-10. Here aircraft, space, missile, and UAV contracts had systematically *higher* schedule growth rates and the differences were statistically significant, whereas only ship contracts had systematically *lower* schedule growth rates when compared to the baseline and a statistically significant difference.

Table 2-10. Program Contract Schedule Growth by Commodity: Development

Schedule Growth Since Original Award Early Production Contracts by Commodity for Currently Active MDAPs (weighted by spending; 2000–2015)

	Combined early production contracts	Aircraft	Ships	Space	C3I	Missiles	UAV
75th percentile	43%	41%	30%	55%	50%	42%	59%
Median	**19%**	**28%**	**1%**	**33%**	**8%**	**34%**	**27%**
25th percentile	-37%	-30%	-37%	0%	0%	0%	-36%
IQR (%-points)	*80%*	*72%*	*67%*	*55%*	*50%*	*42%*	*95%*
N	*157*	*33*	*48*	*11*	*16*	*11*	*18*
Relative Growth		***Higher***	***Lower***	***Higher***	______	***Higher***	***Higher***

NOTE: Ships have higher N (numbers of contracts) since they often involve separate contracts for shipyards, major ship systems, and individual ships. Noted differences are statistically significant.

Contract-Level Trends and Correlates: Development and Early Production

To assess MDAP contract performance trends across DoD, we examine major contracts for MDAP programs that were active (providing SARs) at the beginning of FY 2015, controlling for inflation and removing statistical outliers.

Analytic Preliminaries

Some preliminary comments are in order before we look at performance trends in development and early production contracts. We used all major contracts for currently active MDAPs with usable EV data that are centrally available. We controlled dollarized performance attributes for inflation. To control for the distorting effects of genuine outliers on underlying price, cost, and schedule trends, we also use four common statistical tests to identify outliers to remove from the datasets (see Appendix A for further discussion). The outliers identified for the development phase are analyzed separately in detail in this report. The outliers identified for the early production phase will be analyzed in detail in a future report.

In addition to straightforward trend analysis, we compare results before and after January 2009. To help ensure these comparisons isolate the performance characteristics of interest, we controlled for contract maturity to ensure the set of contracts starting before 2009 had statistically indistinguishable completion rates compared to rates for the contracts since 2009. Also, we controlled for phase difference by analyzing development contracts separately from early production contracts. In addition, we wanted to assess any correlation of contract size to contract performance. Since it turned out that final schedule and final spending on these contracts are highly correlated, we used latest schedule length (i.e., cycle time) as a measure of contract size to prevent tautological relationships between size and growth measures that use spending (or some version of spending) in their formulae for calculation.

As for cost performance, we note that by using EV data we can decompose total contract cost growth into two independent measures: (1) the growth of the target cost through modifications that add work to a contracted effort (i.e., *work-content growth*), and (2) cost growth over that cumulative target cost on the contract (i.e., *cost-over-target*).

Work-content growth can indicate a lack of stability in engineering and system requirements and specifications at the contract level. Some analysts have long asserted that work-content growth is dominated by these low-level engineering and specification changes, not high-level KPP or KSA changes; GAO's analysis confirmed that recent programs with high cost growth seldom involved KPP changes (see GAO, 2015c). Conversely, work-content growth may reflect the use of spiral, incremental, or segmented acquisition strategies. Further analysis on outliers, such as that discussed staring on p. 44, is required to distinguish the source of work-content changes.

A net negative cost-over-target should reflect the relative success of efforts to control cost, especially (a) our Should-Cost initiative to look for innovative methods and approaches that can generate cost savings or avoidance relative to the originally expected contract cost, and (b) our efforts to increase the use of stronger formula-type incentive contracts (e.g., cost-plus-incentive-fee and fixed-price-incentive contracts) that explicitly tie cost-over-target to contractor financial results. Conversely, then, a net positive cost-over-target should indicate that post-contract Should-Cost efforts and contract type selection are not showing progress in controlling contract cost growth (independent of work-content growth).

Affordability caps should affect work-content growth and cost-over-target. They should constrain the addition of requirements that drive up costs in both development and production. Affordability caps also reinforce Should-Cost by motivating contractors and PMs to look for ways to keep costs under control. Further analysis is needed to test these linkages and look for their effects in contract performance metrics.

DoD-Wide Trends on MDAP Development and Early Production Contracts

Table 2-11 summarizes the performance trend analysis of major development and early production contracts for currently active MDAPs for price and cost growth, cycle time (schedule), schedule growth, and gross margins (the difference between price and cost) earned by the contractor. We tested for trends in two ways: whether there is a significant trend over the entire period (using regression analysis), and comparing the proportion of MDAP contracts (as a group) with work start dates either before or since 2009. We indicate whether there is a statistically significant trend or proportional difference for each performance measure. Selected illustrations of these results are included below and earlier in Figure H-13 and Figure H-11.

Table 2-11. Program Contract Performance Trends

DoD-Wide Cost- and Schedule-Related Trends on Contracts for Currently Active MDAPs (controlling for inflation; 2000–2015)

Development

	Trend	Overall Median	Significant Change in Proportions		
				Started before 2009	Started since 2009
Price growth	*none*	21%	**Yes** (% with net reductions)	29%	44% (***better***)
Total cost growth	*none*	25%	*none*		
Work content growth	*none*	1%	*none*		
Cost-over-target	**Down**	0% *	**Yes** (% below target)	37%	80% (***better***)
Margins	*none*	8%	**Yes** (% with low margins)	54%	38%
Cycle Time (years)	**Down**	6 years	**Yes** (% below median)	31%	86%
Schedule growth	*none*	38%	*none*		

Early Production

	Trend	Overall Median	Significant Change in Proportions		
				Started before 2009	Started since 2009
Price growth	*none*	−1%	**Yes** (% with net reductions)	43%	58% (***better***)
Total cost growth	*none*	4%	**Yes** (% with net reductions)	33%	52% (***better***)
Work content growth	*none*	−3%	*none*		
Cost-over-target	**Down**	0%**	**Yes** (% below target)	31%	62% (***better***)
Margins	*none*	9%	*none*		
Cycle Time (years)	**Down**	4.6 years	**Yes** (% below median)	20%	69%
Schedule growth	*none*	19%	*none*		

* Median dropped from 0% for contracts started before 2009 to −3% started since 2009.

** Median dropped from 11% for contracts started before 2009 to 0% started since 2009.

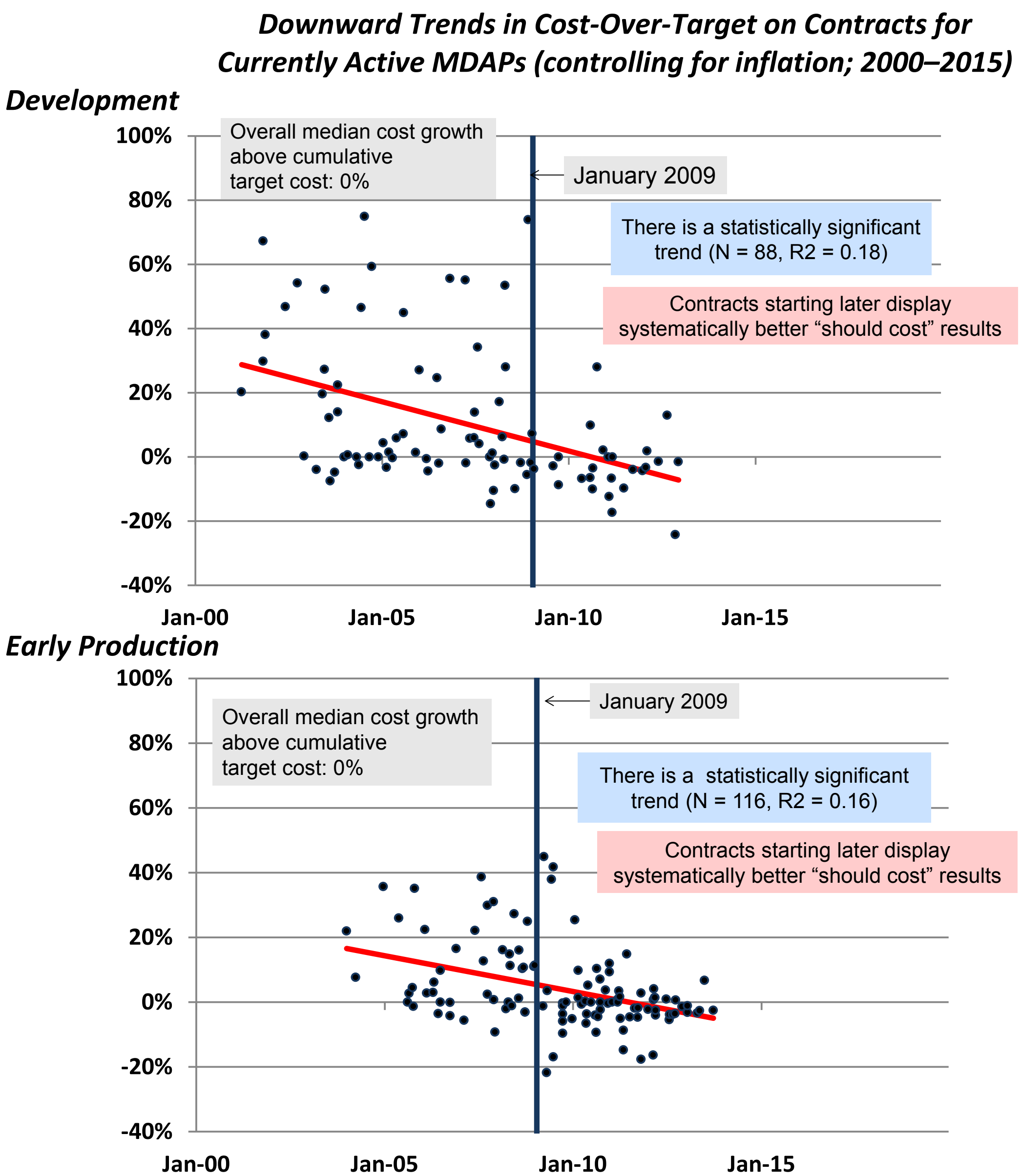
NOTE: Trends or changes were indicated only if statistically significant. Outliers were removed separately for each variable to reduce the distorting effects of outliers and obtain more reliable measures of central tendency. Changes in price reflect the net changes in work content, cost-over-target, and final margins.

Price Growth. One useful measure of cost-related performance is the *price* the DoD pays on a contract. Price differs from costs paid by the contractor in that price reflects the margin (profit or fee) on top of contractor costs. Thus, price growth reflects the net changes in work content, cost-over-target, and final margins on these contracts.

A statistically significant greater proportion of these major MDAP contracts started since 2009 generated price reductions from the original contract award. The proportions showing reductions are up from 29 percent to 44 percent in development, and up from 43 percent to 58 percent in early production. In part, this may be due to the downward trend in cost-over-target.

Cost-Over-Target. Of note, cost-over-target has dropped significantly in both the trend and proportion tests for development and early production contracts. This is evidentially the main reason why a higher proportion of MDAP contracts started since 2009 show net price reductions from original MS B baselines in both phases compared to earlier contracts. Also in terms of total cost growth (the sum of work-content growth and cost-over-target), a higher proportion of contracts started since 2009 showed a reduction than did earlier contracts—but only in early production. Figure 2-38 illustrates these improvements by plotting cost-over-target for individual contracts by their work start dates.

Figure 2-38. Program Contract Costs Related to Targets



NOTE: Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency. The differences in work-completion percentages before and since 2009 were statistically insignificant, so these results are not due to more recent contracts being immature relative to older contracts.

Given these positive trends and proportional improvements in cost-over-target and their likely implication for our BBP initiatives on affordability caps, Should-Cost, competitive environments, and contract incentives, we further examined the distribution and potential confounding factors to see whether the results still hold. Cost-over-target is highly correlated with contract cycle time. To control for the possibility that cycle time (in years) might be confounding the result, we ensured that the two subsets of contracts—those starting before 2009 and those since 2009—were statistically indistinguishable as to cycle time (schedule). We also compared the two distributions of cost-over-target when weighted by contract spending. The median cost-over-target was 0 percent for development contracts starting before 2009 but improved to −3 percent (a reduction) since 2009, and the difference is statistically significant. In early production, the median cost-over-target also dropped significantly from 11 percent before 2009 to 0 percent since 2009. Thus, the results are robust and improvement is real.

Cycle Time. Cycle time (a measure of size that could reflect factors such as complexity and the planned work pace) is the other measure that showed significant downward trends and proportional changes on both development and early production contracts. Figure 2-39 illustrates these trends by plotting cycle time for individual contracts by their work start dates.

Note that in the next chapter we developed a new abstract measure using factor analysis of not only cycle time (initial schedule) but initial contract price as well as schedule growth and cost-over-target (see discussion starting on p. 109 and analytic details starting on p. 155). That new variable could also measure (in part) complexity and also shows a statistically significant recent downward trend (see Figure 3-4 and Figure 3-7).

These results are positive from the perspective of efforts to reduce cycle time and deliver operational capabilities faster. However, this may also indicate (at least in part) that we are pursuing less-complex systems and risk losing our technical superiority.

Margins. The results on margins also require discussion. The DoD does not have an objective to lower or raise margins (i.e., we are not attacking margins or trying to achieve savings by simply lowering margins). Instead, we are working to ensure that margins change in response to performance by the contractor. Higher margins are good if we get commensurate value for them. Thus, the lack of overall margin trends in Table 2-11 is neither good nor bad because this analysis is not explicitly testing for correlations between performance measures (price, cost, and schedule growth) and margins as we did in last year’s report.

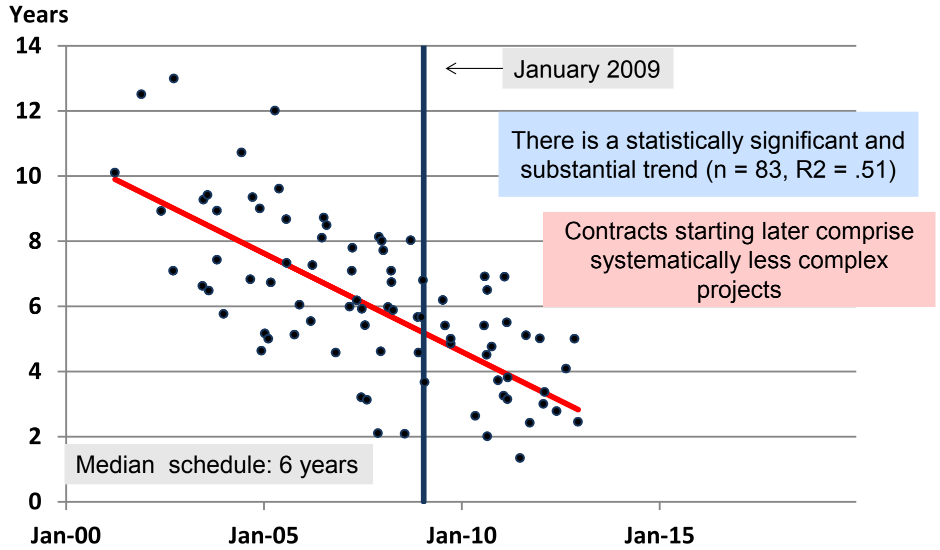
Table 2-11 shows there was a smaller share of low-margin contracts since 2009 (38 percent since 2009 versus 54 percent before 2009), and this difference is statistically significant. In other words, more contractors earned higher margins in development since 2009, illustrating that the DoD is not waging a war on profits. Quite the contrary. We are, however, waging a war on costs to the Department, and we are gaining ground. Since contractors are doing a better job of meeting cost targets (the cost-over-target results, discussed above), these higher margins may reflect the rewards for better performance.

Work-Content Growth. There was no statistically significant trend or proportional change in the net change in target cost due to scope-changing contract modifications. The median of one percent was probably lower compared to the total cost growth median of 25 percent because we removed statistical outliers in calculating these medians.

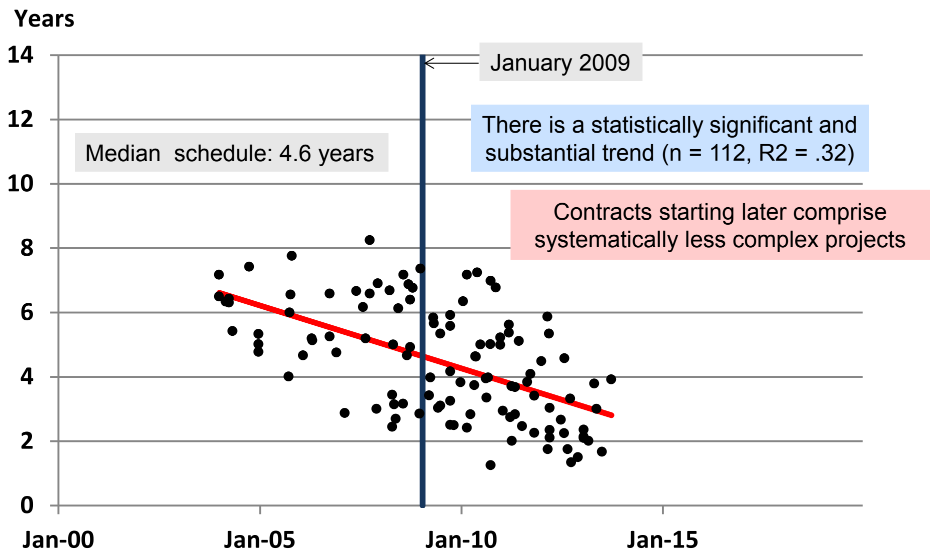
Figure 2-39. Program Contract Lengths

Downward Trends in Contract Cycle Time of Contracts for Currently Active MDAPs (2000–2015)

Development



Early Production



NOTE: Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency. The differences in work-completion percentages before and since 2009 were statistically insignificant, so these results are not due to more recent contracts being immature relative to older contracts.

Despite these improvements, other measures did not always improve. While cost-over-target performance suggests that affordability caps, Should-Cost, competitive environments, and contract-type initiatives in BBP are being implemented and producing measurable results, there is neither a statistically significant trend in price growth nor a statistically significant trend in total cost growth. This is probably because there is no statistically significant improvement in work-content growth, which tends to dominate total cost growth. On the other hand, proportional comparisons do show some signs of improvement in price growth, yet this could be because the data show we are currently undertaking less complex and less risky projects in development than in the past (see the cycle time data above and the factor analysis in Chapter 3 starting on p. 109).

Price-Growth Correlates on MDAP Development Contracts

Table 2-12 shows the factors that correlate with price growth on major development contracts (less outliers) for MDAPs that were active in early FY 2015. In addition to the correlates shown, we also tested the following variables and found that they did not correlate with price growth on this set of development contracts (i.e., they were either spurious or statistically insignificant):

- UCAs[30]
- Contract spending share of program spending
- Cost-over-target
- Share of cost growth due to work-content growth
- Share of cost growth due to cost-over-target
- Margin
- Change in margin over the contract's period of performance
- Contract spending (total dollars)
- Schedule growth
- Service (i.e., Army, Navy, Air Force, or DoD)
- Other commodities except space systems
- Quantity changes

[30] Historically since 1970 on a broader set of MDAP contract data, UCAs generally added about 7 percentage points to MDAP contract total cost growth (see AT&L 2013b, p. 44), with a much stronger correlation of 41 percentage points for ship contracts (ibid, p. 45). Price growth tends to track total cost growth (the difference being margins), so it is unclear why UCAs did not correlate with price growth on these more recent contracts. It could be that the effect was somewhat smaller than the already small 7 percentage points or the smaller sample size affected the result. Nevertheless, caution is warranted when using UCAs on development contracts because starting efforts without well-defined work requirements and cost targets weakens management control and risks wasted effort.

Table 2-12. Potential Program Contract Price Drivers: Development

Average (Median) Predicted Contributions to Price-Growth from Correlates on Development Contracts for Currently Active MDAPs (controlling for inflation; 2000–2015)

Median amount of:	Contributes these percentage-points to contract price growth
Work-content growth	3%
Cycle time (schedule length, years)	17%
If a space contract	3%
Total "average" predicted price growth:	23%

NOTE: Price growth was not adjusted for changes in work. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency. Bootstrap simulations were conducted on regression analyses to correct for bias in estimated coefficients and to calculate accurate standard errors. N = 95 contracts; R^2 = 0.89 for this regression. Regression coefficients were: 0.873 (standard error 0.0524) for work-content growth; 0.0287 for cycle time (standard error 0.00506); and 0.45 (standard error 0.201) for the space commodity effect. All other independent variables were either spurious or statistically insignificant.

As can be seen from this regression, the three statistically significant predictors of price growth were: (1) work-content growth (i.e., total growth in target costs due to contract modifications), (2) cycle time (schedule, in years), and (3) an independent commodity effect for the space systems. These three predictors account for nearly all (89 percent) of the variation in the contract price growth data in development.

To help interpret this result, we show what the regression predicts for a median contract. A median amount of work-content growth would add three percentage points to the price growth while cycle time would add 17 percentage points. In other words, median cycle time is the major correlate, and the median work-content growth has a separate effect of moderately increasing development price growth. If the contract was for a space system, this commodity factor would add another three percentage points at the median (on average).

Price-Growth Correlates on MDAP Early Production Contracts

Table 2-13 shows the factors that correlate with price growth on major early production contracts (less outliers) for MDAPs that were active in early FY 2015. In addition to the correlates shown, we also tested the following variables and found that they did not correlate with price growth on this set of early production contracts (i.e., they were either spurious or statistically insignificant):

- UCAs
- Contract spending share of program spending
- Cost-over-target
- Share of cost growth due to work-content growth
- Share of cost growth due to cost-over-target

- Margin
- Change in margin over the contract's period of performance
- Contract spending (total dollars)
- Schedule growth
- Service (i.e., Army, Navy, Air Force, or DoD)
- Other commodities except aircraft
- Quantity changes

Table 2-13. Potential Program Contract Price Drivers: Early Production

Average (Median) Predicted Contributions to Price-Growth from Correlates on Early Production Contracts for Currently Active MDAPs (controlling for inflation; 2000–2015)

Median amount of:	Contributes these percentage-points to contract price growth
Work-content growth	−3%*
Cycle time (schedule length, years)	9%
If an aircraft contract	−1%
Total "average" predicted price growth:	5%

* The median amount of work-content growth from contract modifications after adjusting for inflation was negative, hence the median predicted contribution to price growth was negative.
NOTES: Price growth was not adjusted for changes in work. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency. Bootstrap simulations were conducted on regression analyses to correct for bias in estimated coefficients and to calculate correct standard errors. N = 125 contracts; R^2 = 0.64 for this regression. Regression coefficients were: 0.849 (standard error 0.0587) for work-content growth; 0.02 for cycle time (standard error 0.00463); and −0.052 (standard error 0.0281) for the aircraft commodity effect. All other independent variables were either spurious or statistically insignificant.

As can be seen from this regression, the three statistically significant predictors of price growth were: (1) work-content growth (i.e., total growth of target costs due to contract modifications), (2) cycle time, and (3) an independent commodity effect for aircraft. These three predictors account for nearly two-thirds (64 percent) of the variation in the contract price growth data in early production. Interestingly, Factors 1 and 2 were the same as in development. However, the commodity effect in production now is for aircraft rather than space systems, and aircraft show *reduced* price growth.

To help interpret this result, we show what the regression predicts for a median contract. The median inflation-adjusted work-content growth is negative. This is why the interpreted "average" effect (measured by the median) of change in target cost is negative while the regression coefficient is positive. Thus, a median planned schedule length would be expected to result in a 9 percent-higher price growth, while a median work-content growth would reduce price growth by 3 percent. In other words, cycle time is the major correlate at the median, and work-content growth has a separate effect of reducing price growth at the median in early production. Of course, a positive inflation-adjusted work-content cost growth would be expected to increase price growth. If the contract were for an aircraft, this commodity factor would decline by another percentage point, on average.

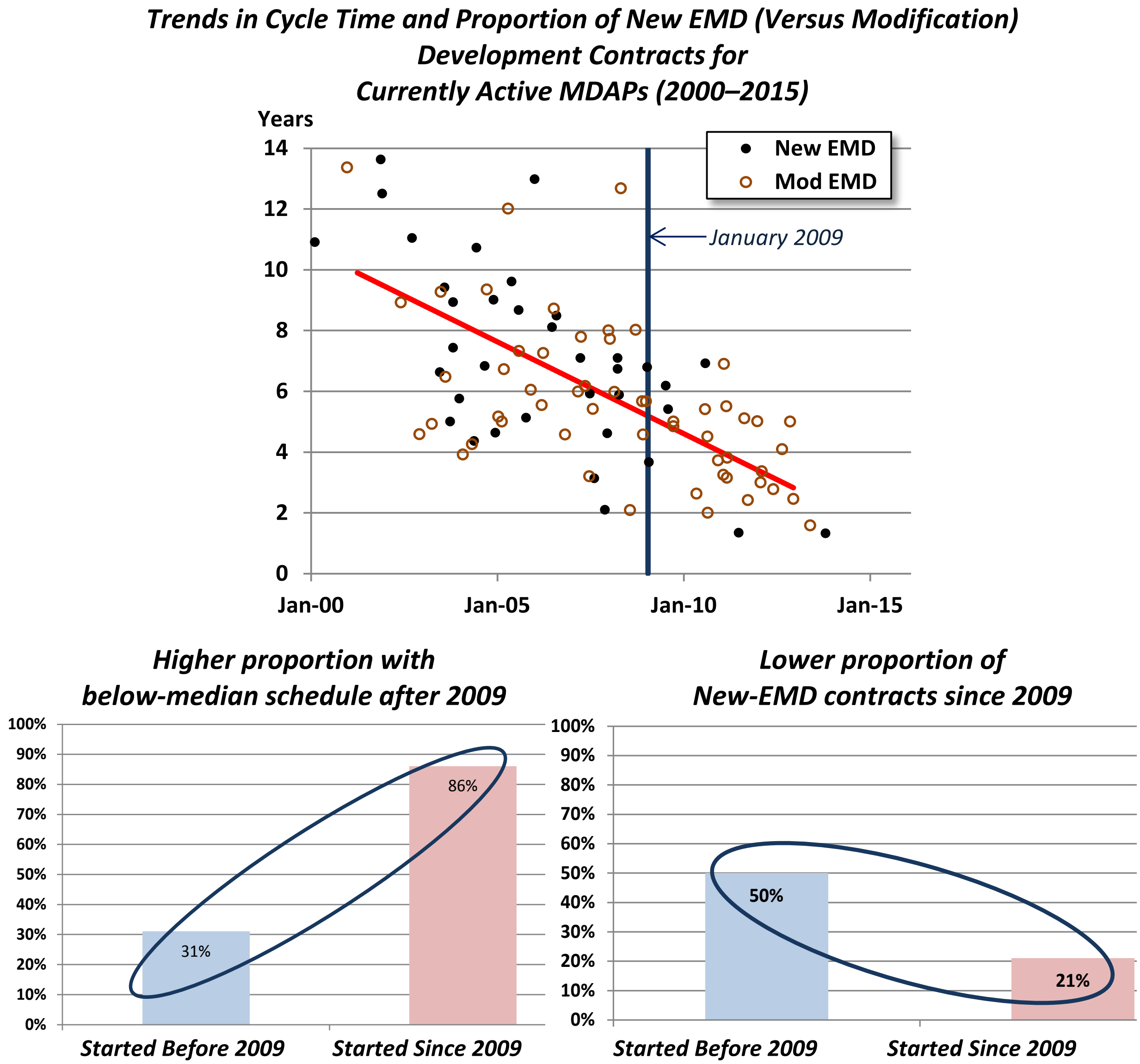
Other Potential Drivers of MDAP Contract Performance: Complexity and New Systems

One possible driver of the trends shown in Table 2-11 above may be shifts in the complexity[31] of the MDAP contract work—possibly from shifts in the relative amounts of new systems to modifications for previous systems. For example, a shorter cycle time (final schedule) could indicate (among other things) a relatively less complex effort. The data, trend analysis, and proportion tests shown in Figure 2-40 support this inference. As discussed above, contract cycle time (schedule length) had a significantly downward trend over time, and proportion tests confirm that shorter contracts represented a far greater share of development contracts since 2009 (86 percent after and 31 percent before 2009). We also examined which contracts where for new systems and which were upgrades or modifications of existing systems. Significantly fewer development contracts since 2009 were for new systems rather than system upgrades (21 percent since 2009 as contrasted with 50 percent before 2009); this can be seen visually in the upper plot of Figure 2-40 and is confirmed in the proportion test in the lower right. This trend and these shifts in proportions were all statistically significant. Thus, there is strong evidence that the reduction seen in cycle time is at least partially due to a shift in the portfolio from the acquisition of new systems to modifications. Of course, cycle time reductions could be due to other factors (e.g., higher funding levels for faster work completion; breaking up work into more contracts).

The difference in cycle time between new systems and modifications was larger for higher-spend contracts (i.e., when total contract price is larger). Figure 2-41 shows that when weighted by relative total contract spend (dollars), median cycle time for new systems was twice that for modifications (14 years compared to 6.9 years), and the values were shifted higher than when unweighted (at 7 and 5.1 years, respectively, for new and modifications).

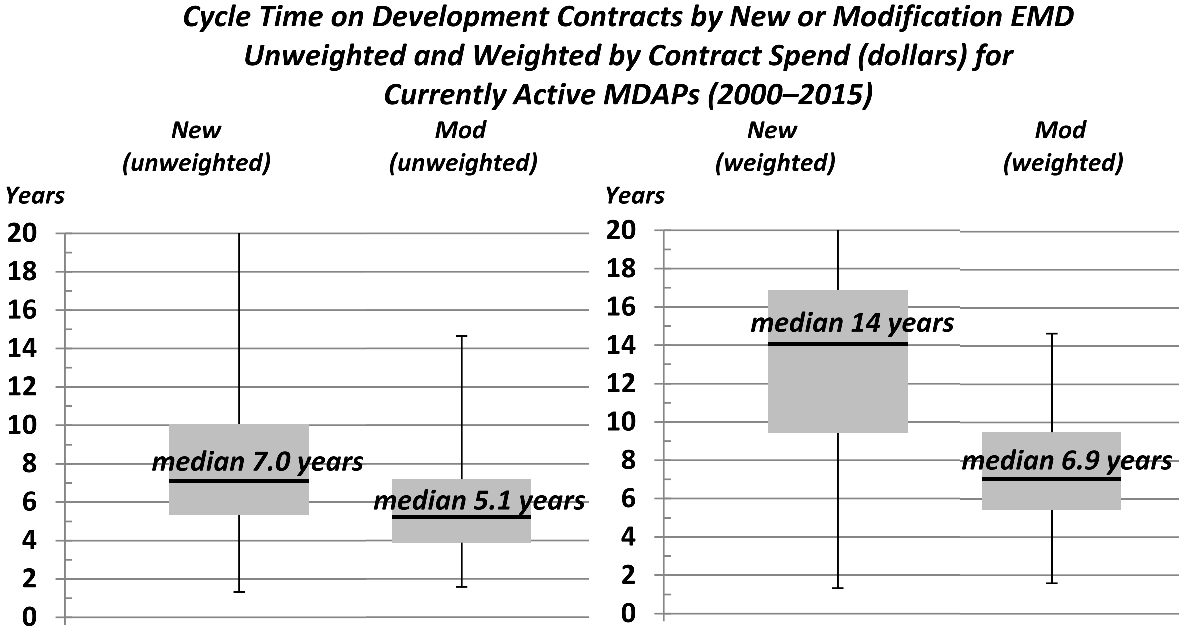
[31] Note that in this report we do not use "complexity" in the sense of unpredictability or sensitivity to initial conditions but rather the sophistication and complexity of the system and its design.

Figure 2-40. Program Contract Length Trends and New Systems: Development



NOTE: Differences are statistically significant. Statistical outliers (both high and low) were removed from the cycle time (schedule) plots to obtain more reliable measures of central tendency. Mod = modification.

Figure 2-41. Program Contract Lengths and New Systems: Spend Effects: Development



Preliminary analysis indicates that the difference between new system contracts and modification contracts also has a statistically significant effect on total spend (price) and cost-over-target. Spend is about 6 times higher at the median for new-EMD contracts, and cost-over-target is about 11 percent for new-EMD versus 0 percent for modification development contracts in this dataset. However, new versus modified systems did not have a significant effect on price growth, total cost growth, work-content growth, schedule growth, or final margins in this set of MDAP contracts. Thus, it appears that new-system contract work compared to modification work has an effect in and only in the two measures where we see overall downward trends since 2000. Further analysis is needed to control for outliers and to distinguish whether new-EMD affects the proportional improvements seen in price (on development MDAP contracts) and total cost growth (on both development and early production MDAP contracts) in Table 2-11.

Of note, later in this report we develop and employ a more sophisticated measure using multivariate factor analysis which indicates that complexity could be a major driver behind these trends (see the analysis starting on p. 109). Factor analysis allows us to reflect cost-related variables (namely, initial contract spend [price] and cost-over-target) together with schedule-related variables (namely, initial planned cycle time and schedule growth) while testing other variables. This allows us to better control for the spend effects illustrated in Figure 2-41 while trying to measure the more abstract properties of complexity and risk.

Determining the exact causes behind these trends requires further analysis, but these data indicate that complexity (for example, as evidenced by fewer new-system contracts recently) is probably one major driver of recent trends. These trend variables also align conceptually with BBP initiatives, especially for cost-over-target. However, separating BBP contributions from other effects is difficult with these data.

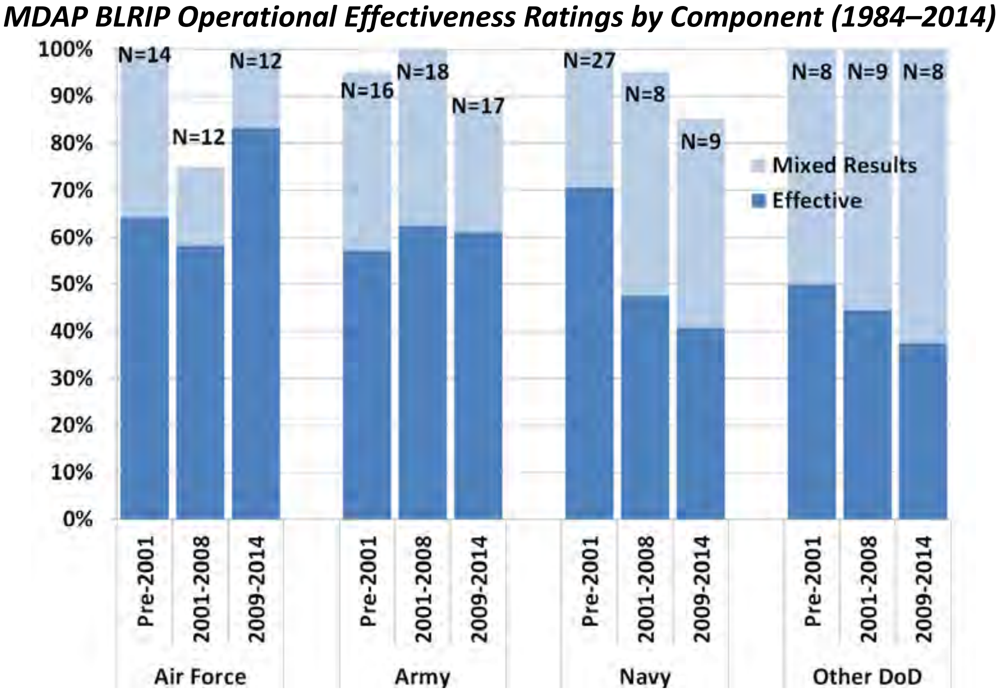
INSTITUTIONAL ANALYSES: MILITARY DEPARTMENTS

This section summarizes various acquisition performance metrics analyzed by Component. Here we examined significant program-level cost growth as exhibited by Nunn-McCurdy breaches as well as contract-level cost, price, and schedule growth.

BLRIP Operational Testing of Effectiveness and Suitability by Component

Figure 2-42 shows effectiveness ratings for the Components. The Air Force appears to have made some progress in improving BLRIP effectiveness, but the Navy appears to be slipping. Caution is warranted, however, given the small sample sizes (especially for the "Other DoD" programs). Over all time periods, more systems are rated as effective than suitable in their BLRIP operational tests. Also, Other DoD systems appear to have the worst BLRIP effectiveness and suitability ratings.

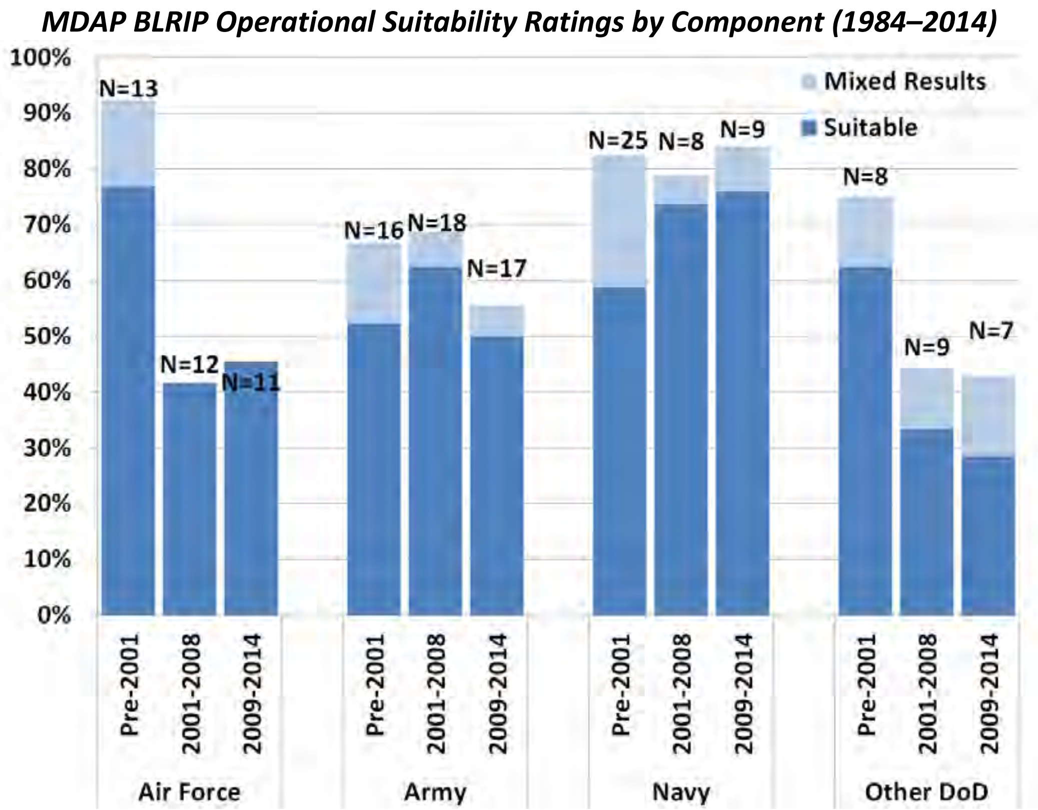
Figure 2-42. Program Technical Performance: Effectiveness by Component



Source: DOT&E reports.
NOTE: Differences are only apparent and may not be significant due to the low sample sizes (infrequent evaluations). DoD programs were Joint or other programs that are not exclusive to a single Component. Sample sizes differ between effectiveness and suitability for some Components because there was not always a definitive binary judgment for effectiveness and suitability in all reports.

Figure 2-43 shows suitability test results. Overall, the performance of systems across Components is much lower for suitability than for effectiveness. Here the Navy appears to have made some progress since 2001, but the Air Force and Other-DoD programs appear to remain at lower levels compared to the period prior to 2001. Again, caution is warranted about the precision of the percentages given the small sample. For example, just one different rating would push the Air Force and Other DoD results over the 50 percent mark.

Figure 2-43. Program Technical Performance: Suitability by Component



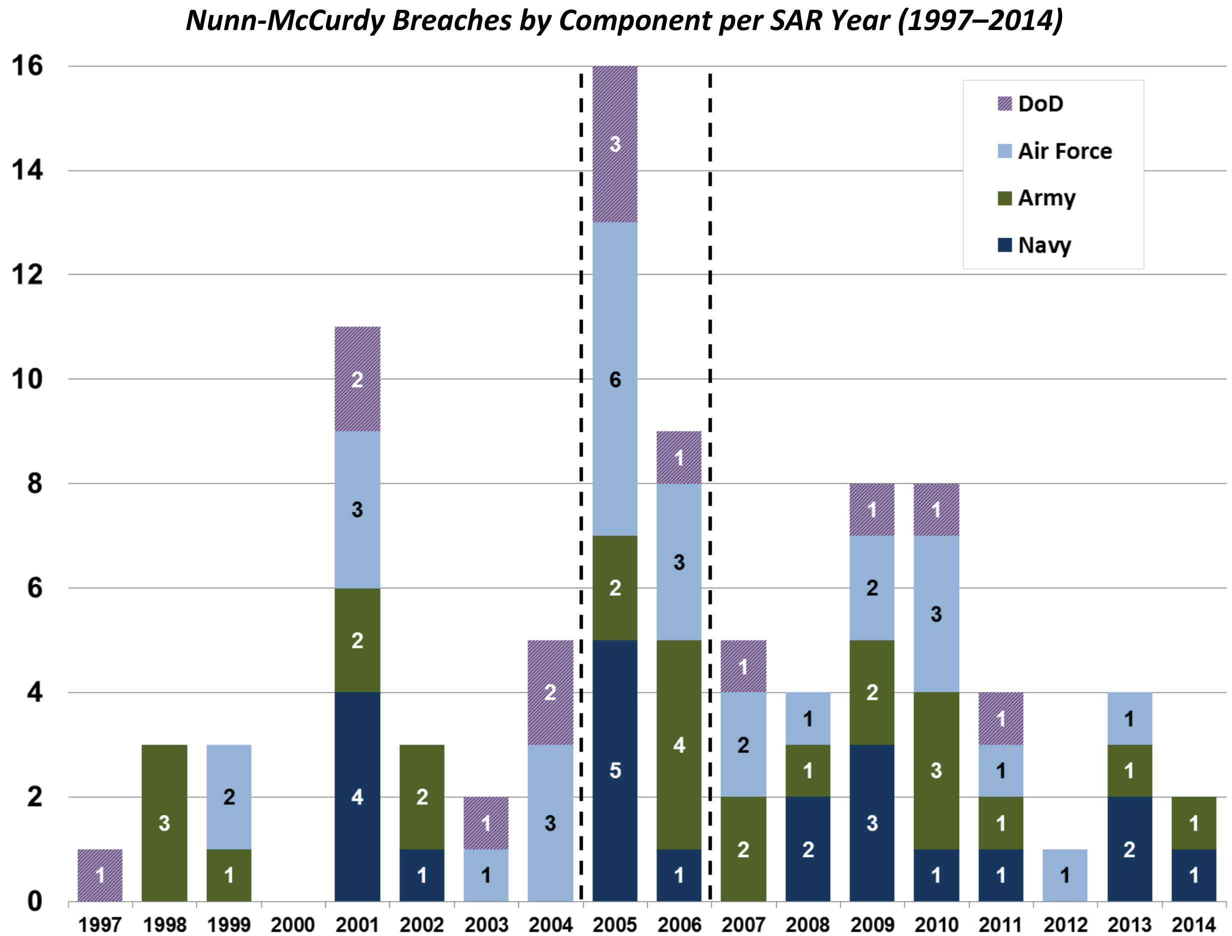
Source: DOT&E reports.

NOTE: Differences are only apparent and may not be significant due to the low sample sizes (infrequent evaluations). DoD programs were Joint or other programs that are not exclusive to a single Component. Sample sizes differ between effectiveness and suitability for some Components because there was not always a definitive binary judgment for effectiveness and suitability in all reports.

Nunn-McCurdy Cost Breaches

One measure of acquisition program cost performance is the Nunn-McCurdy breach rate by Component. Figure 2-44 shows Nunn-McCurdy breach numbers by year from 1997 to 2014. As introduced previously, this chart now aligns with the new DoD official breach list (see Table 2-2 on p. 21 above).

Figure 2-44. Program Cost-Related Performance: Nunn-McCurdy Breaches by Component



NOTE: The criteria for breaches were changed in NDAA 2006, so the counts before 2005 are different than those since 2006, and 2005 was a transition year and not comparable to either half. Breaches are determined using “base-year” dollars (i.e., adjusting for inflation). This plot includes the number of breaches in each annual SAR reporting cycle, which nominally equates to calendar year but may include updates early in the following calendar year from the President’s Budget Request. Breaches in different years for different thresholds or baselines for the same program are included in each respective year. If a program reported both a significant and critical breach in the same year, only one breach is shown here. Nunn-McCurdy breach reporting was established in the NDAA for FY 1982, but the new official AT&L list only tracks breaches back through 1997.

Table 2-14 summarizes a different analysis of Nunn-McCurdy breaches by Component. Here we do not "double count" programs that have breached multiple times. This allows us to get a sense of the tendency of programs within each Component to breach.

Historically, about a third of MDAPs had at least one breach (i.e., about two-thirds have cost growth below 15 percent). At least two-thirds of programs that breach go critical (i.e., fewer remain at the significant level), except for Army programs, which are split.

All breaches are listed regardless of cause. If a program had both a significant and a critical breach, it was only included in the "programs with critical breach" column.

As discussed in the earlier Nunn-McCurdy breach section, there are various causes of these breaches. Some programs may breach because of cancellation (e.g., Land Warrior) and some programs may have been canceled before their breaching cost growth was reported in the SARs (e.g., VH-71).

Table 2-14. Nunn-McCurdy Breaching by Component

Nunn-McCurdy MDAP Breach Rates by Component (1997–2014)

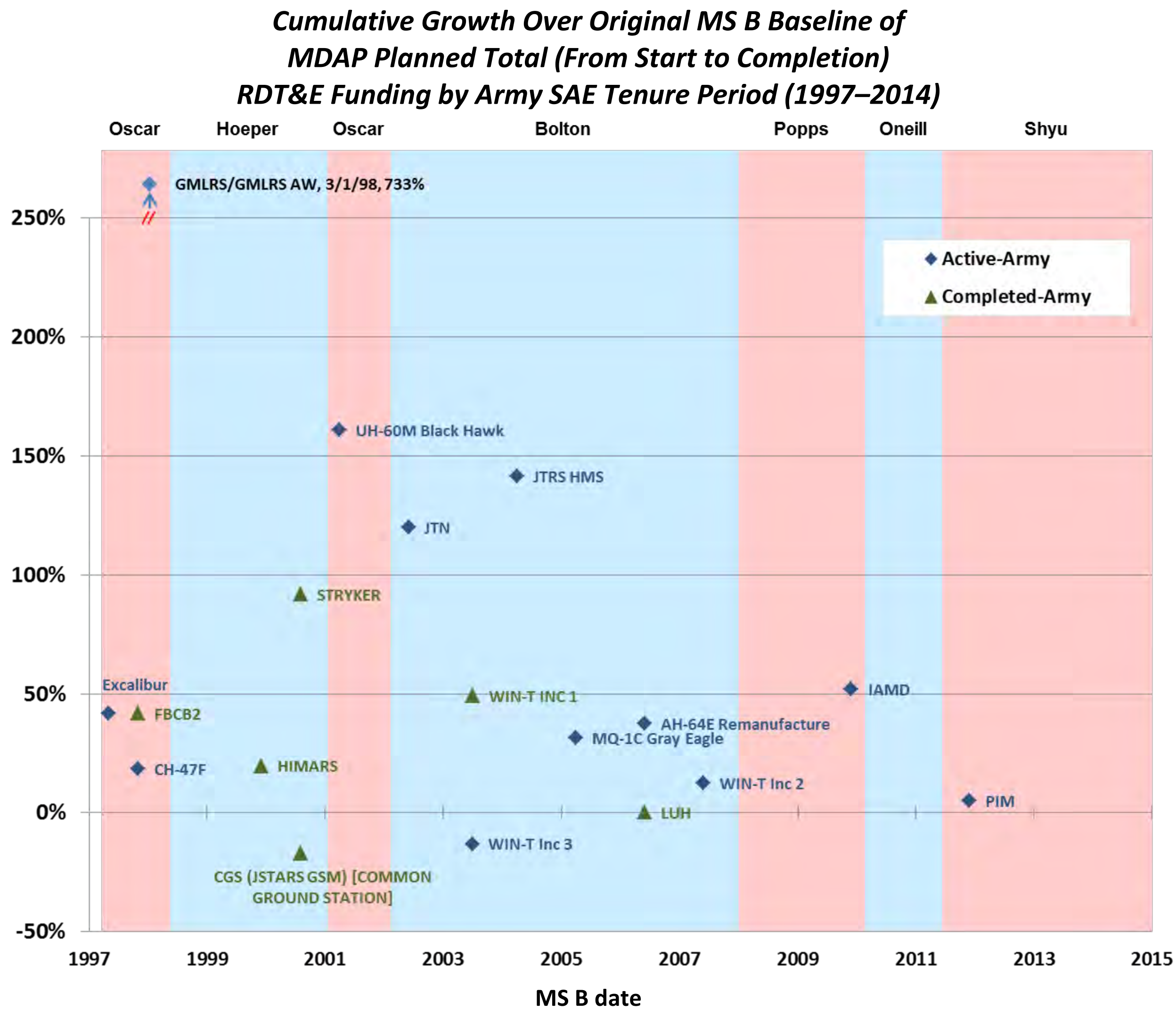
Component	Total # of Programs	# of Programs that Ever Breached	Breach Rate	# of Programs with at Most a Significant Breach	# of Programs with a Critical Breach
DoD	12	7	58%	1	6
Army	54	18	33%	8	10
Navy	64	17	27%	6	11
Air Force	57	15	26%	3	12
Total	**187**	**57**	**30%**	**18**	**39**

NOTE: The list of MDAPs by Component has been revised slightly since last year's USD(AT&L) report to align with the new official list. If a program had both a significant and critical breach, it was only included in the "programs with critical breach" column. The data are not adjusted for quantity or other variances. "DoD" programs are programs categorized as such in SAR reporting, which include Joint programs and programs (such as Chem-Demil) overseen by an organization other than the Air Force, Army, or Navy. Breaches are determined using "base-year" dollars (i.e., adjusted for inflation).

Program Cost Growth and Service Acquisition Executives

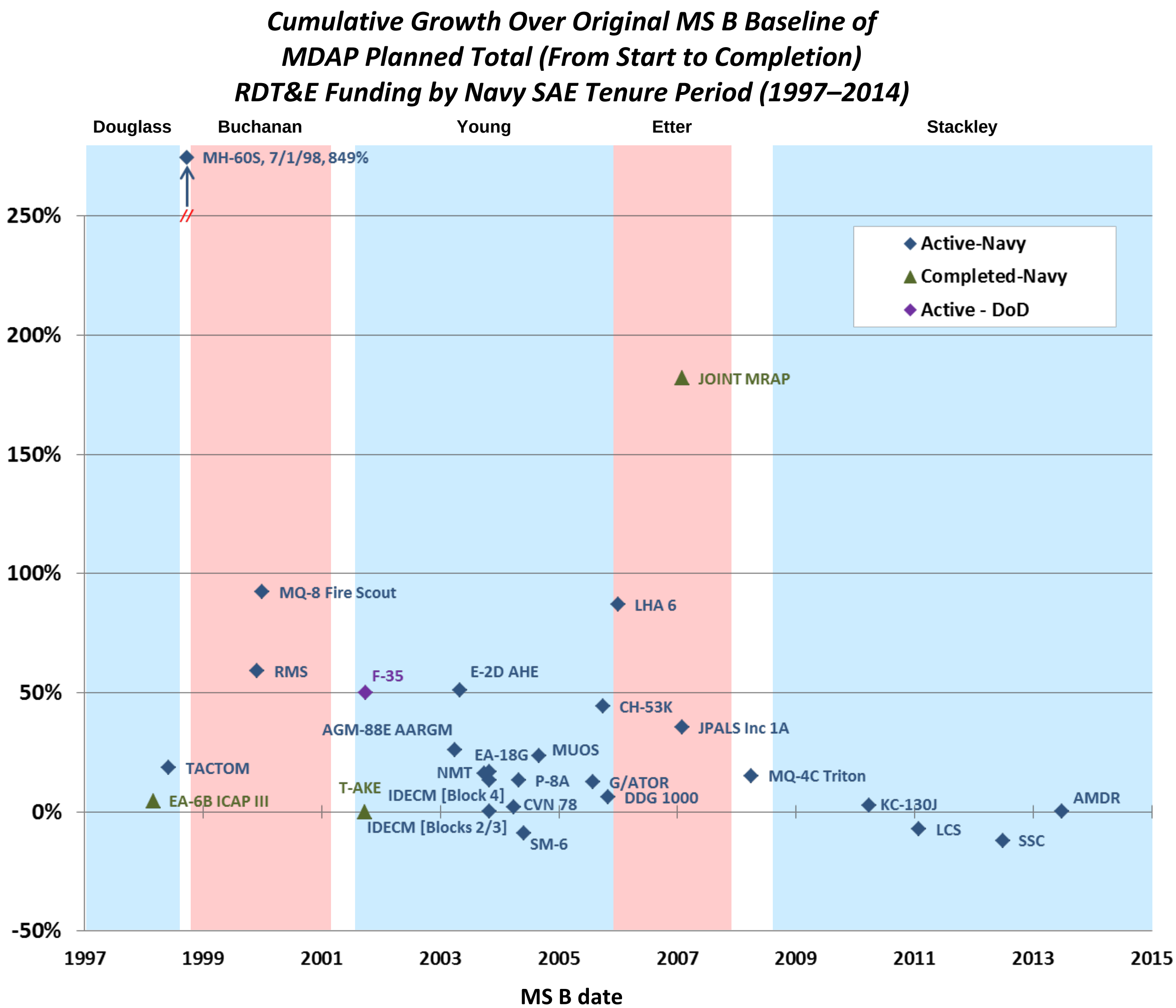
As with the earlier plots for the USD(AT&L), we plot growth in MDAP Planned Total Funding in development and procurement (respectively) for active and completed MDAPs against original baselines as reported to Congress in the SARs, identifying the person who was the SAE at the time of the MDAP's MS B approval. Figure 2-45 shows the result for Army SAEs, Figure 2-46 for Navy SAEs, and Figure 2-47 for Air Force SAEs. Recall that Figure 2-19 show total cost grown on major MDAP development contracts by Service and can serve as an early indicator of potential cost growth at the project level.

Figure 2-45. Army Program Cost-Related Development Performance Baselined in SAE Periods



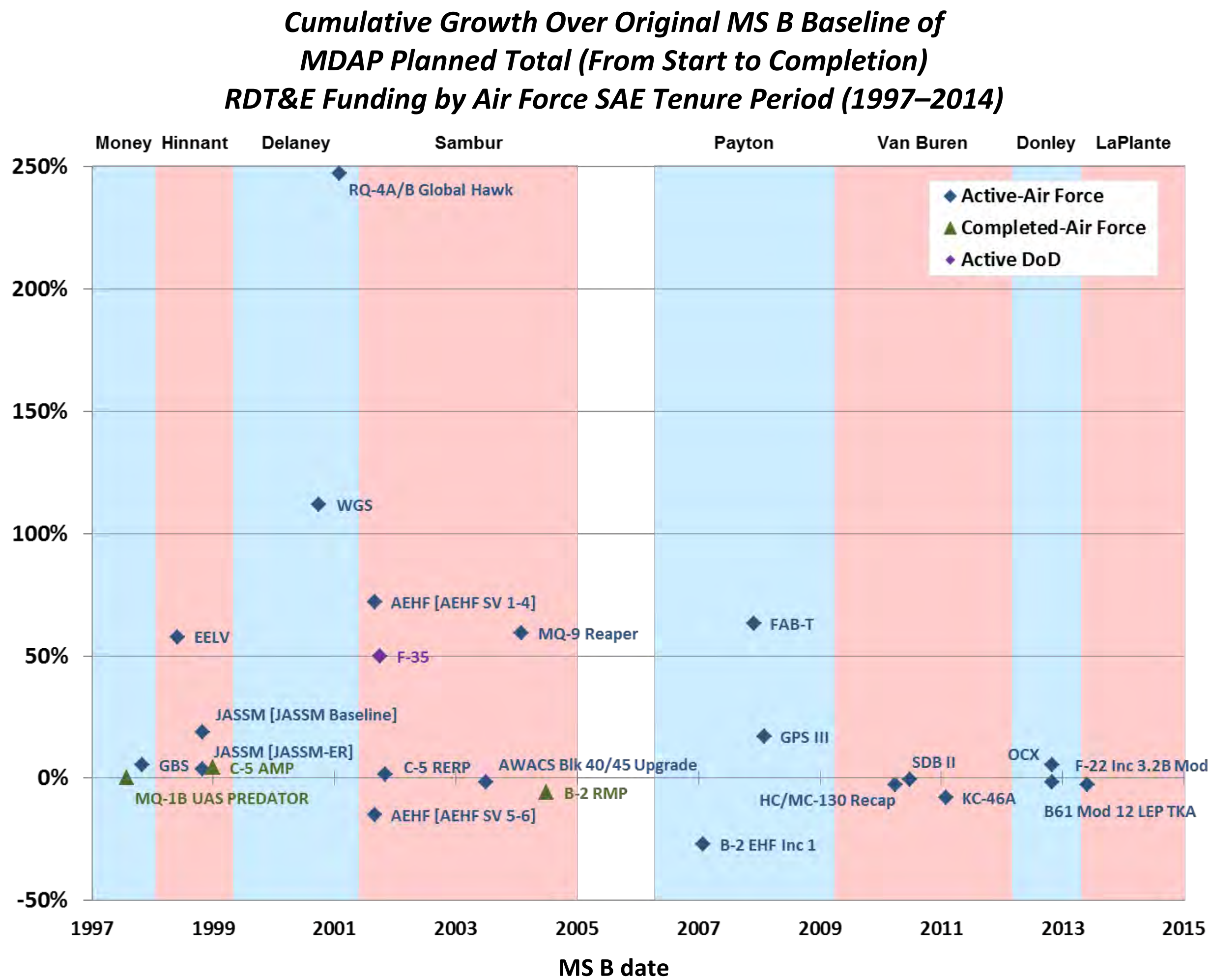
NOTE: Canceled programs include those that became inactive in the SARs before Full-Rate Production (FRP); they may show negative or artificially low RDT&E cost growth because they were canceled before they expended all of the originally planned RDT&E. Blue diamonds are programs active as of the December 2014 SAR. Any white bars between DAE shaded regions represent periods when there was no confirmed executive. Source cost data are in "base-year" dollars (adjusted for inflation). Immature programs that have not spent at least 30 percent of their original EMD schedule are excluded to help control for the naturally lower cost growth of immature programs.

Figure 2-46. Navy Program Cost-Related Development Performance Baselined in SAE Periods



NOTE: This shows total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. These are percentage changes after adjusting for inflation from original MS B baseline of actual past and estimated future funding as reported in each program's latest SAR. Total RDT&E is an insightful measure because it is necessary regardless of quantity. White bars between SAE shaded regions represent periods in which there was no confirmed executive. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown.

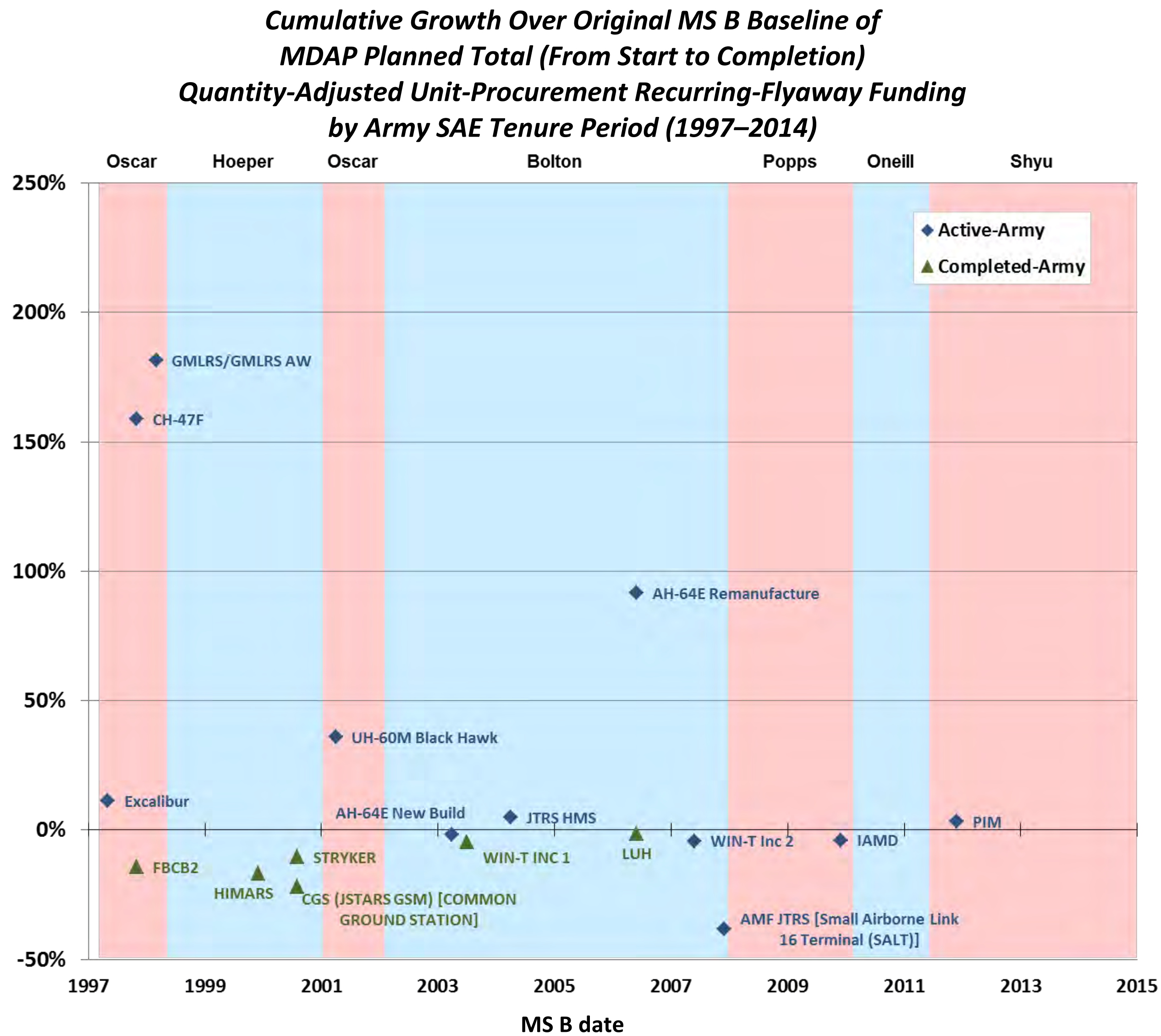
Figure 2-47. Air Force Program Cost-Related Development Performance Baselined in SAE Periods



NOTE: This shows total RDT&E funding growth independent of procurement funding and quantity changes; it reflects any work-content changes. These are percentage changes after adjusting for inflation from original MS B baseline of actual past and estimated future funding as reported in each program's latest SAR. Total RDT&E is an insightful measure because it is necessary regardless of quantity. White bars between SAE shaded regions represent periods in which there was no confirmed executive. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown.

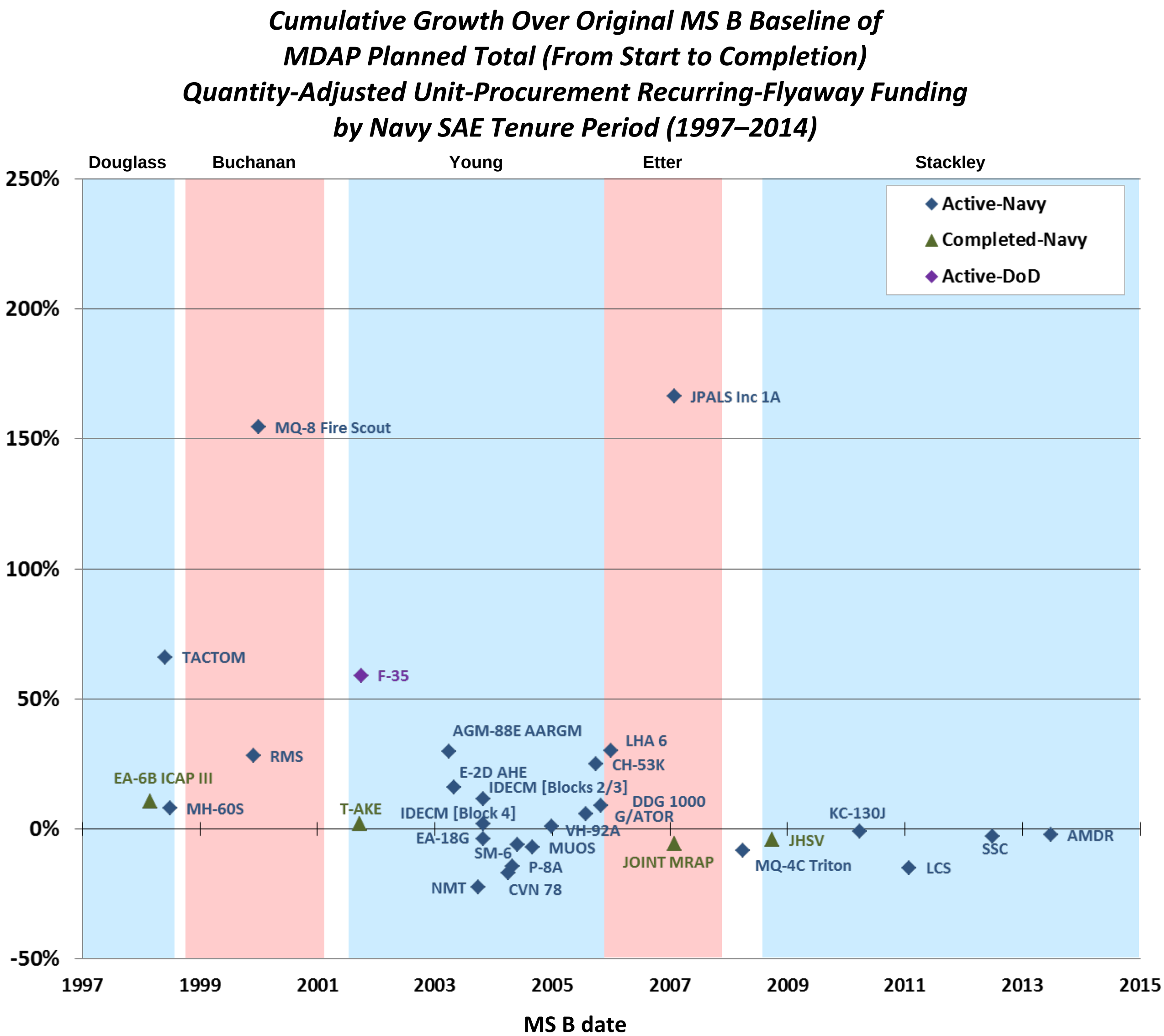
Next we plot the project-level quantity-adjusted procurement unit cost growth on MDAPs in each Military Department shown by who was the SAE in office at the time of MS B approval. Figure 2-48 shows the result for Army SAEs, Figure 2-49 for Navy SAEs, and Figure 2-50 for Air Force SAEs. Recall that Figure 2-37 show total cost grown on major MDAP early production contracts by Service and can serve as an early indicator of potential cost growth at the project level.

Figure 2-48. Army Program Cost-Related Procurement Performance Baselined in SAE Periods



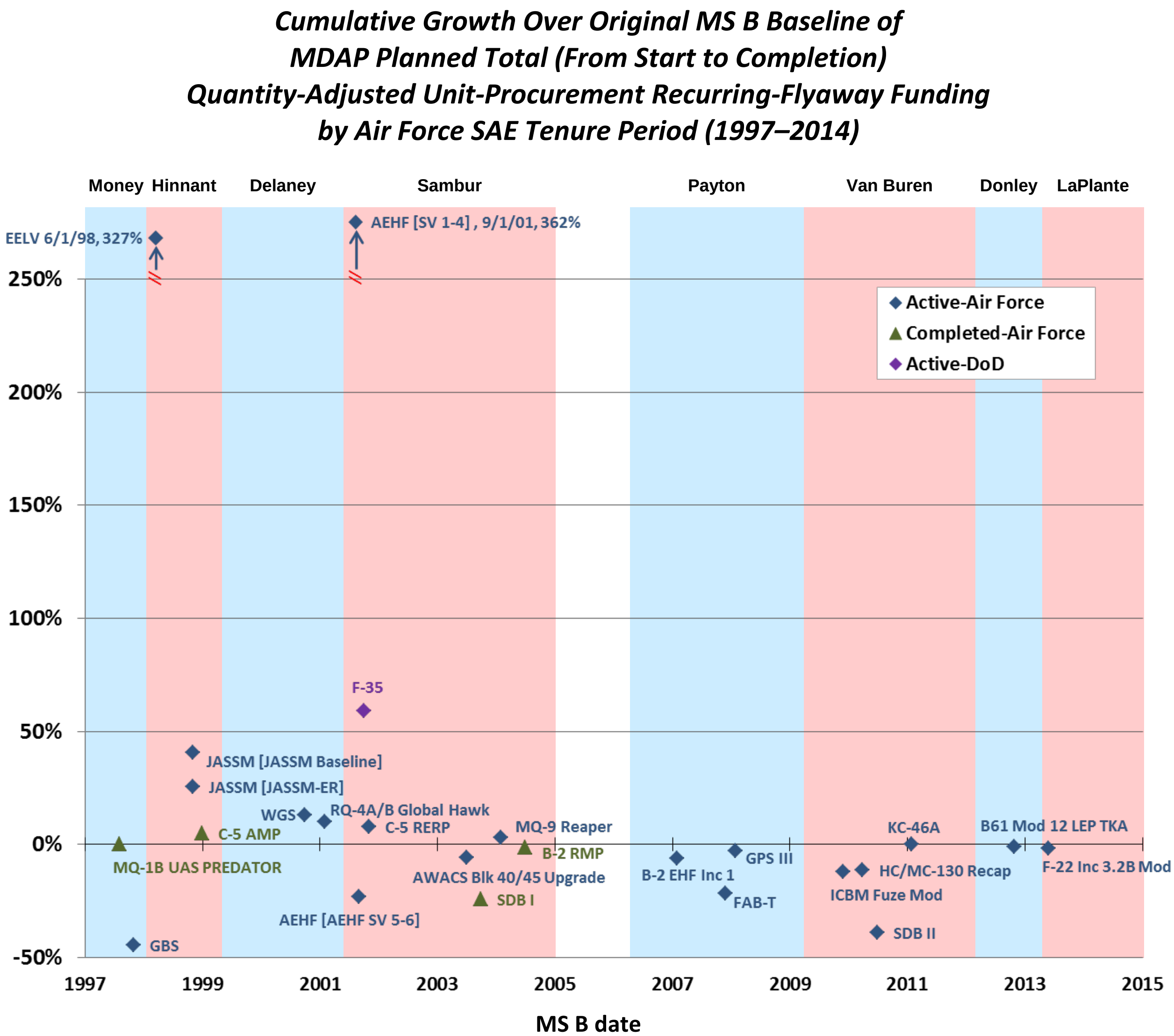
NOTE: This shows growth in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. White bars between SAE shaded regions represent periods in which there was no confirmed executive. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown.

Figure 2-49. Navy Program Cost-Related Procurement Performance Baselined in SAE Periods



NOTE: This shows growth in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. White bars between SAE shaded regions represent periods in which there was no confirmed executive. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown.

Figure 2-50. Air Force Program Cost-Related Procurement Performance Baselined in SAE Periods



NOTE: This shows growth in unit recurring flyaway funding after adjusting for quantity changes; it is independent of RDT&E funding but reflects any work-content changes. These are percentage changes after adjusting for inflation and any quantity changes from original MS B baseline of actual past and estimated needed future funding as reported in their latest SAR. White bars between SAE shaded regions represent periods in which there was no confirmed executive. Relatively new programs that have not spent at least 30 percent of their original EMD schedule are not shown.

Contract Cost Growth by Military Department

We now examine performance by Component on major contracts for MDAP programs that were active (providing SARs) at the beginning of FY 2015, controlling for inflation and removing statistical outliers. Table 2-15 summarizes trend and proportional analysis of MDAP contract performance by Component to identify areas where statistically significant improvements can be seen.

Table 2-15. Program Contract Performance Trends by Component

Component Trends on Contracts for
Currently Active MDAPs (controlling for inflation; 2000–2015)

Development

	Trend	Significant Change in Proportions		
			Started before 2009	Started since 2009
Price growth	*none*	**Army** (N=27) % with net reductions:	24%	60% (***better***)
Work content growth	*none*	*none*		
Cost-over-target	**Down**: **Army**	% below target: **Army** (N=28) **Navy** (N=27)	44% 30%	80% (***better***) 100% (***better***)
Margins	***Up***: **Army**	**Army** (N=28) % with low margins:	67%	30%
Cycle Time (years)	***Down:*** **Army, Navy, Air Force, DoD, MDA**	% below median: **Army** (N=25) **Navy** (N=28) **Air Force*** (N=16)	33% 29% 22%	100% 100% 57%
Schedule growth	*none*	*none*		

Early Production

	Trend	Significant Change in Proportions		
			Started before 2009	Started since 2009
Price growth	*none*	% with net reductions: **Army** (N=9) **Navy** (N=69) **Air Force** (N=23)	67% 48% 33%	0% (***worse***) 67% (***better***) 64% (***better***)
Work content growth	*none*	% with net reductions: **Army** (N=9) **Air Force** (N=23)	100% 22%	0% (***worse***) 71% (***better***)
Cost-over-target	**Down**: **Navy, DoD**	% below target: **Navy** (N=69) **Air Force** (N=23) **DoD** (N=6)	32% 50% 0%	56% (***better***) 86% (***better***) 43% (***better***)
Margins	*none*	*none*		
Cycle Time (years)	***Down:*** **Army, Navy, Air Force**	% below median: **Army** (N=9) **Navy** (N=61) **Air Force** (N=22) **DoD** (N=13)	0% 14% 50% 17%	71% 55% 86% 71%
Schedule growth	***Down:*** **Air Force**	% with net reductions: **Army** (N=9) **Air Force** (N=23)	33% 11%	0% (***worse***) 50% (***better***)

*DoD had a downward size trend in development, but its proportion test failed due to low sample size despite increasing from 50% to 100%. NOTES: Trends or changes were only indicated if statistically significant. Outliers were removed separately for each variable to reduce the distorting effects of outliers and obtain more reliable measures of central tendency. Price, work-content cost growth, and cost-over-target were all after adjusting for inflation. Changes in price reflect the net changes in work content, cost-over-target, and final margins. MDA is the Missile Defense Agency.

Price Growth in Development. While there were no overall trend in price growth over this period (2000–2015), the Army had more major MDAP contracts showing a net reduction in price since 2009 compared with those started before 2009. Changes in price reflect the net changes in work content, cost-over-target, and final margins.

Cost-Over-Target in Development. The Army showed a significant downward trend in cost-over-target over this period (2000–2015). Proportion tests showed that the Army and Navy had more MDAP contracts meeting or beating their cost targets since 2009 than those started before 2009.

Higher Army Margins in Development, Possibly for Better Performance. The Army had a statistically significant increase in the margins paid on development contracts. However, this appears to be justified given the Army's improvement over the period with respect to improved price growth performance and cost-over-target performance. This aligns with the DoD's effort to tie performance to margins. Also, the Army's median margin of 7.9 percent is similar to those of most of the other Components. Likewise, in proportional comparisons, the Army had a statistically significant decrease in the share of low-margin contracts since 2009 (30 percent since and 67 percent before 2009). This is entirely consistent with the trend shown above. Again, this appears justified, given the Army's improved price growth and contractor cost-over-target performance.

Cycle Time in Development. Each Component had a statistically significant downward trend in contract cycle time over this period (2000–2015). The Army, Navy, and Air Force all had significantly more contracts showing below median cycle time since 2009 than those started before 2009. DoD had more as well, but the sample size was too small to show statistical significance.

Work Content and Schedule Growth in Development. There were no statistically significant trends in work-content and schedule growth. Also, there were no significant changes in proportions with negative work-content or schedule growth since 2009.

Price Growth in Early Production. There were no significant overall trends, but proportional comparisons in early production tell a different and somewhat complicated story. The Army had a statistically significant smaller share of its early production MDAP contracts started since 2009 with price reductions (0 percent since 2009 versus 67 percent before 2009). Caution is warranted, however, in the Army's case because the sample was very small. The Navy, however, had a much higher statistically significant share of its early production contracts started since 2009 with price reductions (67 percent since 2009 and 48 percent before 2009). Likewise, the Air Force had a much higher and statistically significant share of its early production contracts started since 2009 showing price reductions (64 percent since 2009 and 33 percent before 2009). Changes in price reflect the net changes in work content, cost-over-target, and final margins.

Work-Content Growth in Early Production. While there were no overall trends, proportional comparisons for shares of contracts showing negative work-content cost growth (i.e., reductions) also tell a somewhat complicated story. The Army had a statistically significant smaller share of early production contracts started since 2009 with negative work-content growth (0.0 percent since 2009 versus 100 percent before 2009). But again, the Army's sample is small. The Air Force, however, had a larger statistically significant share of its early production contracts started since 2009 with a negative work-content growth (71 percent since 2009 versus 22 percent before 2009).

Cost-Over-Target in Early Production. Overall on early production contracts in this period (2000–2015), both Navy and DoD MDAP contracts had statistically significant downward trends for cost-over-target (possibly due to progress in Should-Cost and contract incentives). Proportional tests showed progress for the Navy, Air Force, and DoD on early production contracts with negative cost-over-target (i.e., net reductions). The Navy's share of contracts started since 2009 and which had reductions showed a statistically significant increase (56 percent since 2009 versus 32 percent before 2009). Likewise, the Air Force's share of contracts started since 2009 that had negative cost-over-target showed a statistically significant increase (86 percent since 2009 versus 50 percent before 2009). And DoD showed a statistically significant increase in its share of early production contracts started since 2009 that had cost-over-target reductions (43 percent since 2009 versus 0.0 percent before 2009).

Cycle Time in Early Production. Trends for early production contracts in terms of cycle time (measured in years) showed that the Army, Navy and Air Force all displayed statistically significant downward trends, while the DoD and the Missile Defense Agency (MDA) had no significant trends. Therefore as expected, proportions of contracts showing below-median cycle time on early production MDAP contracts increased across the board.

Schedule Growth in Early Production. While trend analysis for schedule growth during early production showed the Air Force alone had a statistically significant downward trend, proportional comparisons tell a somewhat different story. The Army had a statistically significant smaller share of early production contracts that started since 2009 and that saw net negative schedule growth (0.0 percent since 2009 and 33 percent before 2009). Again, in the Army's case, the sample was small. The Air Force, however, had a statistically significant higher share of its early production contracts started since 2009 with net negative schedule growth (50 percent since 2009 versus 11 percent before 2009).

Final Margins in Early Production. There were neither statistically significant trends in early production contract margins nor significant changes in proportions since 2009.

INSTITUTIONAL ANALYSES: PRIME CONTRACTORS

We now examine performance by major prime contractor on major contracts for MDAP programs that were active (providing SARs) at the beginning of FY 2015, controlling for inflation and removing statistical outliers.

Performance by Prime: Development

Table 2-16 summarizes preliminary analysis showing cost-over-target on development contracts by major primes since 2000. This analysis includes primes that had at least five development contracts in our larger dataset. All comparisons controlled for contract maturity and contract spending. Recall that this metric represents the degree of cost growth independent from work-content growth. It should indicate both how well contractors identify and implement cost-control measures and the effectiveness of our contract cost-growth incentives.

However, while there are statistical differences in contract cost-over-target performance across our prime contractors, we do not understand all the underlying mechanisms for these differences. There are at least two complications. First, we know that project risk and uncertainty—by almost any metric—vary substantially across commodity classes, and some of these differences in prime-contractor performance surely reflect their differences in our sample (i.e., this analysis did not control for distorting commodity effects due to the small sample sizes). Second and more generally, the contract targets themselves are not pure, unbiased estimates of project cost and schedule. In addition to expected project scale (size), they also reflect other contract terms (e.g., share lines and incentives as well as market contestability and the general negotiating environment). Also, they are, in part, the result of bidding strategies. Nevertheless, these data begin to provide performance insights relative to the cost targets agreed upon by both the government and the contractor, and they control for (i.e., are independent of) any work-content cost growth on these contracts. We plan to provide further analysis in a future report.

Table 2-16. Contractor Cost Control Related to Targets: Development

Total Cost-Over-Target on Development Contracts for Currently Active MDAPs (controlling for inflation and maturity; dollar basis; 2000–2015)

	Combined development contracts	Huntington Ingalls	General Dynamics	Raytheon	Northrop Grumman	General Atomics	Boeing	Lockheed
75th percentile	62%	-5%	13%	32%	24%	39%	40%	66%
Median	**30%**	**-7%**	**5%**	**5%**	**11%**	**24%**	**26%**	**62%***
25th percentile	6%	-7%	-4%	2%	-3%	-5%	1%	38%
IQR (% points)	*56%*	*2%*	*17%*	*30%*	*27%*	*44%*	*39%*	*28%*
N	*102*	*5*	*14*	*18*	*16*	*6*	*11*	*32*

*Driven largely by its F-35 SDD contract.

NOTE: Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency. These results do not control for commodity effects or potential biases in how the cost targets are set in the first place.

Lockheed had the highest cost-over-target at the median. A percentile-by-percentile comparison also showed that Lockheed had systematically higher cost-over-target when contracts were weighted by their respective spending over all their distributions. This difference was statistically significant, due to the large share of Lockheed's F-35 SDD contract by relative dollar size. Other differences are not statistically significant (e.g., due to low sample size).

We also compare contractors' performance in controlling schedule growth. The results are summarized in Table 2-17. All comparisons are controlled for contract maturity and weighted by contract spending. However, due to the small sample sizes, it was not possible to control for the distorting effects of any work-content growth on these contracts.

Table 2-17. Program Contract Schedule Growth by Contractor: Development

Schedule Growth Since Original Award by Prime Contractor on Development Contracts for Currently Active MDAPs (controlled for maturity; dollar basis; 2000–2015)

	Combined development contracts	Boeing	Huntington Ingalls	General Dynamics	General Atomics	Northrop Grumman	Lockheed	Raytheon
75th percentile	72%	23%	12%	60%	60%	67%	73%	108%
Median	**43%**	**0%**	**1%**	**19%**	**30%**	**52%**	**70%***	**107%****
25th percentile	12%	0%	0%	6%	22%	-2%	40%	19%
***IQR** (% points)*	*60%*	*23%*	*12%*	*54%*	*38%*	*69%*	*33%*	*89%*
N	*103*	*12*	*5*	*14*	*6*	*16*	*32*	*18*

* Driven largely by its F-35 SDD contract.
** Driven largely by its DDG 1000 SDI contract.
NOTE: These results do not control for work-content growth. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency.

Lockheed and Raytheon had very large schedule growth rates per contract at the median. Indeed, percentile-by-percentile comparisons showed that Lockheed and Raytheon contracts in development had systematically higher schedule growth over all their distributions of development contract spending. These differences were statistically significant. Lockheed's large schedule growth was due to the large dollar share of its F-35 SDD contract. Raytheon's large schedule growth was due to the large dollar share of its DDG 1000 SDI contract. Other differences are not statistically significant (e.g., due to low sample size).

Performance by Prime: Early Production

Looking at the early production contracts for currently executing MDAP programs, we also conducted preliminary analysis of the relative performance of the major prime contractors in our dataset. All comparisons were controlled for contract maturity and weighted by contract spending. The results on control of cost-over-target costs are summarized in Table 2-18 for each prime contractor that had at least five contracts in our larger dataset. As discussed above, we could not control for commodity effects due to the small sample sizes. Also, contract cost targets are not pure, unbiased estimates of project cost and schedule. In addition to expected project scale (size), they also reflect other contract terms (e.g., share lines and incentives as well as market contestability and the general negotiating environment). Also, they are, in part, the result of bidding strategies. Nevertheless, these data begin to provide performance insights relative to the cost targets agreed upon by both the government and the contractor, and they are independent of any work-content cost growth on these contracts.

Table 2-18. Contractor Cost Control Related to Targets: Early Production

Total Cost-Over-Target by Prime Contractors on Early Production Contracts for Currently Active MDAPs (controlling for inflation and maturity; dollar basis; 2000–2015)

	Combined early production contracts	General Atomics	Raytheon	Boeing	Pratt & Whitney	Austal USA	General Dynamics	Lockheed	Huntington Ingalls	BAE Systems	Northrop Grumman
75th percentile	15%	2%	5%	1%	5%	9%	4%	11%	21%	25%	53%
Median	**4%**	**-6%**	**-2%**	**0%**	**1%**	**1%**	**1%**	**4%**	**15%***	**24%****	**34%***
25th percentile	0%	-23%	-6%	-3%	-1%	1%	-4%	1%	11%	24%	0%
***IQR** (% points)*	*15%*	*24%*	*11%*	*3%*	*6%*	*8%*	*8%*	*10%*	*11%*	*2%*	*53%*
N	*140*	*12*	*12*	*19*	*6*	*5*	*19*	*36*	*13*	*5*	*13*

* Driven largely by their LPD-17 ship contracts.
** Driven largely by its DDG 1000 Advanced Gun System (AGS) contract.
NOTE: Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency. These results do not control for commodity effects or potential biases in how the cost targets are set in the first place.

Northrop Grumman, BAE Systems, and Huntington Ingalls had much higher cost-over-target at the median. Percentile-by-percentile comparisons across their entire distributions showed these prime contractors had systematically higher cost-over-target when weighting contracts by spending. These results were statistically significant. The performance of both Huntington Ingalls and Northrop Grumman was driven by the large share of their contract dollars during this period on their LPD-17 ship contracts. BAE Systems' performance was driven by the large share of its contract dollars on the DDG 1000 Advanced Gun System (AGS) contract. Other differences are not statistically significant (e.g., due to low sample size).

We also compare prime contractor abilities to control schedule growth in early performance. The results are shown in Table 2-19. As in development, due to the small sample sizes, it was not possible to control for the distorting effects of any work-content growth on these contracts.

Table 2-19. Program Contract Schedule Growth by Contractor: Early Production

Schedule Growth Since Original Award by Major Prime Contractors on Early Production for Currently Active MDAPs (controlled for maturity; dollar basis; 2000–2015)

	Combined early production contracts	General Dynamics	Huntington Ingalls	Pratt & Whitney	Boeing	Northrop Grumman	BAE Systems	Lockheed	Raytheon	Austal USA	General Atomics
75th percentile	43%	9%	29%	33%	32%	38%	55%	44%	89%	48%	122%
Median	**19%**	**0%**	**0%**	**0%**	**2%**	**26%**	**31%***	**34%*** **	**36%*** ***	**43%‡**	**55%‡‡**
25th percentile	0%	-11%	-9%	0%	0%	0%	7%	14%	20%	23%	26%
***IQR** (% points)*	*43%*	*21%*	*38%*	*33%*	*32%*	*38%*	*48%*	*29%*	*69%*	*24%*	*97%*
N	*142*	*19*	*13*	*6*	*19*	*13*	*5*	*38*	*12*	*5*	*12*

* Driven largely by a small sample size and its DDG 1000 AGS contract.
** Driven largely by its F-35 airframe LRIP contract.
*** Driven largely by its Aegis SM-6 contract.
‡ Driven largely by a small sample size and its LCS contract.
‡‡ Driven largely by its MQ-1C Gray Eagle contract.
NOTE: These results do not control for work-content growth. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency.

Lockheed, Raytheon, BAE Systems, General Atomics, and Austal USA had much higher schedule growth at the median. Percentile-by-percentile comparisons across their entire distributions showed these primes had systematically higher schedule growth when weighting contracts by dollar size. These results were statistically significant. BAE Systems' performance was due to a small sample (N) and a large share of its contract dollars being devoted to its DDG 1000 AGS contract. Lockheed's performance was driven by the high proportion of its early production contract dollars being devoted to its F-35 LRIP aircraft contract. Raytheon's result was driven by a large share of its early production contract dollars being from its Aegis SM-6 contract. Austal USA's performance was due to its small sample (N) and the large share of its contract dollars being from its LCS contract. General Atomics' performance was driven by the large share of its contract dollars being from its MQ-1C Gray Eagle contract. Other differences are not statistically significant (e.g., due to low sample size).

Superior Suppliers

As part of BBP, the three Military Departments and the Defense Logistics Agency (DLA) each established a Superior Supplier Incentive Program (SSIP) to incentivize contractor performance by recognizing the contractors that provide the greatest value to the DoD through superior performance and by informing those who perform below average. The basis for SSIP designations are contract performance assessments reported by the PM (or equivalent) to CPARS. Each assessment rates the quality, schedule, cost control, management, utilization of small businesses, regulatory compliance, and other optional aspects of the contractor's performance on a specific contract for a specified period. The contractor is allowed to review and comment on each assessment before it is finalized. This program is still evolving. The Services will meet with Tier 1 suppliers to discuss ways to reduce nonproduction-related requirements in existing contracts. These discussions may lead to the removal of non-value added procedures or requirements, which may ultimately save money for both contractors and taxpayers.

There are some differences between the methods adopted by each Component. The Navy evaluated the top 30 corporations in the first year based on the firms' contract obligations at the corporate level. Corporations then were evaluated at the company, operating division, or business-unit level. This resulted of the evaluation of 80 business units, from which 30 were selected for ranking in 2014. The Air Force and Army used data from USAspending.gov to find the 25 corporations with the highest contract obligations in their respective departments over the last 3 years. Like the Navy, the Air Force and the Army then evaluated performance at the business-unit level. DLA selected the top-40 performers from the 153 largest (in terms of contracts) suppliers of parts and commodities over the last 2 years.[32] The three Services use the last 3 years of performance data from CPARS to rate the largest firms doing business with each Service. The 3-year periods results are weighted 3,2,1 in the scoring with the most recent rated highest. The DLA's top 40 consists of the top performers in CPARS over the last 2 years and several companies that partnered with DLA on considerable cost-reduction initiatives. The Services are working to harmonize their methodologies.

Based on this analysis, the top suppliers are sorted into three tiers, with only alphabetical ordering within each tier. The very top-performing business units are in Tier 1 or Gold. The Tier 2 or Silver business units are the next highest performers. Finally, there are the Tier 3 or Bronze business unit performers. Results from the first two SSIP releases are shown in Table 2-20 (for Army suppliers), Table 2-21 (for Navy suppliers), Table 2-22 (for Air Force suppliers), and Table 2-23 (for DLA suppliers). Note that these rankings are only for the performance of the business units for the specified department or agency, and only for the time in question.

[32] Companies with primarily service-oriented contracts were not considered by DLA at this time.

Table 2-20. Army Contractors by SSIP Tiers (2014–2015)

2014	2015
Tier 1	**Tier 1**
BAE Electronic BAE Global Combat Systems Boeing Corp.— Global Services & Support Finmeccanica DRS Technologies General Electric—GE Aviation Harris Corporation Lockheed Martin—Info Systems and Global Solutions Lockheed Martin—Missiles & Fire Control Lockheed Martin—Mission Systems & Training Science Application International Corporation—Research & Development SRCTec, Inc.	Boeing Corp.—Global Services & Support Booz Allen Hamilton, Inc.* DynCorp* General Electric—GE Aviation Harris Corporation—Gov. Communication Sys. Harris Corporation—Exelis Information & Tech. Sys. Leidos Corporation—National Security Solutions* Lockheed Martin—Missiles & Fire Control ManTech International Corp.—ManTech Advanced Systems International, Inc.* Northrop Grumman— Technical Services* Raytheon—Integrated Defense Systems* Raytheon/Lockheed Martin Javelin Joint Venture* Rockwell Collins Government Systems* Textron—Bell Helicopter* Textron—Aviation* United Technologies Corp.—(UTC) Aerospace Systems*
Tier 2	**Tier 2**
Aerovironment, Inc. Boeing Military Aircraft Booz Allen Hamilton, Inc. Chemring Group PLC Cubic, Inc. General Dynamics Info. Systems and Technology Northrop Grumman Aerospace Systems Northrop Grumman Electronic Systems Northrop Grumman Information Systems Raytheon Integrated Defense Systems Raytheon Space and Airborne Systems Science Application International Corporation – Government Services Thales-Raytheon Systems	BAE Systems—Electronic Systems* BAE Systems—Intelligence & Security* BAE Systems—Platforms & Services* Boeing Corporation—Military Aircraft Cubic Finmeccanica—DRS Technologies* General Dynamics—Information Systems & Technology Lockheed Martin—Mission Systems & Training* Northrop Grumman—Electronic Systems Northrop Grumman—Information Systems Raytheon—Space & Airborne Systems Raytheon—Intelligence, Information & Services* Rolls Royce Corporation* Sierra Nevada Corporation* Thales-Raytheon Joint Venture United Technologies—Sikorsky*
Tier 3	**Tier 3**
Alliant Techsystems Defense Group BAE Systems Land & Armaments Boeing Network & Space Systems CSC North American Public Sector General Atomics Technology – Aeronautical Systems General Dynamics Combat Systems Honeywell International Aerospace Oshkosh Corporation Raytheon Intelligence, Information and Services Raytheon Missile Systems Textron Bell Helicopter Textron Systems United Technologies – Sikorsky United Technologies Aerospace Systems	AeroVironment, Inc.* Boeing Corporation—Network & Space Systems Chemring Group, PLC—Sensors & Electronics* Engility* Finmeccanica—AugustWestland* General Atomics Technology Corporation—Aeronautical Systems General Dynamics—Combat Systems Harris Corporation—Exelis C4ISR Electronics & Systems* Honeywell International—Aerospace L-3—Communications Systems* L-3—Electronic Systems* Leidos Corporations—Health & Engineering* Navistar International* Oshkosh Corporation—Oshkosh Defense Raytheon—Missile Systems Textron—Textron Systems

* New to the Tier in 2015. SOURCES: Vergun (2015); Assistant Secretary of the Army for Acquisition, Logistics, and Technology (2015). NOTE: Rankings are unordered within each tier and reflect performance of each business unit for the specified department or agency for the prior 3 years. 2015 rankings include data from FY 2012 through FY 2014.

Table 2-21. Navy Contractors by SSIP Tiers (2014 and 2015)

2014	2015
Tier 1	**Tier 1**
General Dynamics Combat Systems General Dynamics Marine Systems General Electric Aviation Lockheed Martin Mission Systems and Training Maritime Helicopter Support Company (MHSCo; Sikorsky-Lockheed joint venture) Northrop Grumman Aerospace Systems Raytheon Integrated Defense Systems Raytheon Intelligence, Information and Services Rolls Royce Defence Aerospace	BAE Systems Intelligence & Security* BAE Systems Platforms & Services* Erapsco (Spartan Corp. – Ultra Electronics JV)* General Electric (GE) Aviation Lockheed Martin Missiles & Fire Control* Lockheed Martin Mission Systems and Training Northrop Grumman Aerospace Systems Oshkosh Defense* Raytheon Intelligence, Information and Services Rolls Royce Aerospace
Tier 2	**Tier 2**
BAE Electronic Systems BAE Systems Land and Armaments Bell Helicopter Boeing Military Aircraft General Dynamics Information Systems and Technology Ingalls Shipbuilding ITT Exelis L-3 Communication Electronic Systems Northrop Grumman Electronic Systems Raytheon Missile Systems Raytheon Space and Airborne Systems Rockwell Collins Simulation	BAE Systems Electronic Systems Bell Boeing Joint Project Office* Boeing Military Aircraft Boeing Global Services & Support* General Dynamics Combat Systems* General Dynamics Information Systems and Technology General Dynamics Marine Systems* Harris Corporation, formerly Exelis C4ISR Electronics and Systems – ES ** Harris Corporation, formerly Exelis Information & Technical Services – IST** Huntington Ingalls Industries Ingalls Shipbuilding L-3 Aerospace Systems* L-3 Communications Systems L-3 Electronic Systems Lockheed Martin Information Systems & Global Solutions* MHSCo (Sikorsky-Lockheed joint venture)* Northrop Grumman Electronic Systems Northrop Grumman Information Systems* Raytheon Integrated Defense Systems* Raytheon Missile Systems Raytheon Space and Airborne Systems Rockwell Collins Government Systems Textron Systems* ViaSat* United Technologies Corporation UTC Aerospace Systems*
Tier 3	**Tier 3**
ATK Defense Austal USA Bell Boeing Joint Project Office Lockheed Martin Aeronautics Lockheed Martin Space Systems Navistar Defense Newport News Shipbuilding Sikorsky UTC Propulsion and Aerospace Systems	Austal USA, LLC General Atomics Technology Corp* Huntington Ingalls Industries Newport News Shipbuilding* Lockheed Martin Aeronautics Lockheed Martin Space Systems Orbital ATK, Inc. Defense Systems* Textron Bell Helicopter* United Technologies Corporation Pratt Whitney *** United Technologies Corporation Sikorsky ***

* New to the Tier in 2015. ** Former ITT Exelis was in the tier. *** Reorganization from 2014.

SOURCES: Assistant Secretary of the Navy for Research, Development and Acquisition (2014, 2015).

NOTE: Rankings are unordered within each tier and reflect performance of each business unit for the specified department or agency for the prior 3 years.

Table 2-22. Air Force Contractors by SSIP Tiers (2014–2015)

2014	2015
Tier 1	**Tier 1**
BAE Systems Electronic Systems Boeing Commercial Aircraft General Dynamics Aerospace L-3 Communications Systems Lockheed Martin Aeronautics Lockheed Martin Information Systems and Global Solutions Lockheed Martin Mission Systems & Training Lockheed Martin Space Systems Northrop Grumman Information Systems Rockwell Collins Commercial Systems Rolls Royce Sierra Nevada Corp United Technologies Pratt & Whitney	Boeing Commercial Aircraft Boeing Network & Space Systems* General Atomics Energy* General Dynamics Aerospace Harris Corporation, Formerly Exelis Information & Technical Services* Honeywell International Aerospace* L-3 Aerospace Systems* L-3 Communications Systems L-3 National Security Solutions* Lockheed Martin Information Systems & Global Solutions Lockheed Martin Mission Systems & Training Lockheed Martin Space Systems Northrop Grumman Information Systems Orbital ATK Flight Systems** Rolls Royce Aerospace
Tier 2	**Tier 2**
Boeing Military Aircraft Boeing Global Services & Support Boeing Network & Space Systems GE Aviation Honeywell International Aerospace L-3 Aerospace Systems L-3 National Security Solutions Lockheed Martin Missiles & Fire Control Northrop Grumman Aerospace Systems Northrop Grumman Electronic Systems Raytheon Integrated Defense Systems Raytheon Space and Airborne Systems Textron Aviation United Technologies UTC Aerospace Systems	BAE Systems Electronic Systems* BAE Systems Intelligence & Security* Boeing Global Services & Support Leidos Corporation National Security Solutions* Lockheed Martin Aeronautics* Lockheed Martin Missiles & Fire Control Orbital ATK Defense Systems** Orbital ATK Space Systems Group** Raytheon Integrated Defense Systems Raytheon Space and Airborne Systems Rockwell Collins Government Systems* Sierra Nevada Corp* Teledyne Technologies Digital Imaging* United Technologies Pratt Whitney*** UTC Aerospace Systems
Tier 3	**Tier 3**
BAE Systems Intelligence & Security Exelis C4ISR Electronics and Systems Exelis Information & Technical Services General Atomics Technology Corp Aeronautical Systems General Dynamics Information Systems and Technology Jacobs Engineering Tybrin L-3 Electronic Systems Leidos Corp (formerly SAIC) Northrop Grumman Technical Services Raytheon Intelligence, Information and Services Raytheon Missile Systems Rockwell Collins Government Systems Textron Systems United Launch Alliance (ULA) / United Launch Services (ULS) United Technologies—Sikorsky	Boeing Military Aircraft* General Atomics Aeronautical Systems General Dynamics Information Systems and Technology General Electric—GE Aviation* Harris Corporation, Formerly Exelis C4ISR Electronics and Systems* L-3 Electronic Systems Northrop Grumman Aerospace Systems* Northrop Grumman Electronic Systems* Northrop Grumman Technical Services Raytheon Intelligence, Information and Services Raytheon Missile Systems Textron Aviation* Textron Systems ULA/ULS United Technologies—Sikorsky

* New to the Tier in 2015. ** Orbital ATK was formed in 2015. *** Reorganization from 2014.

SOURCES: Vergun(2015); Haux (2015). NOTE: Rankings are unordered within each tier and reflect performance of each business unit for the specified department or agency for the prior 3 years. 2015 rankings include data from FY 2012 through FY 2014.

Table 2-23. DLA Contractors by SSIP Tiers (2014–2015)

2014–2015
Gold
American Apparel Military Uniform Company
AmerisourceBergen Drug Corporation
Atlantic Diving Supply
The Boeing Company
Burlington Industries
Foster Fuels
United Technologies Aircraft Systems
I-Solutions Direct, Inc.
Kampi Components Co.
Kovatch Mobile Equipment Corp.
Lockheed Martin Corporation
Rolls Royce Corporation, U.S.
Tennier Industries, Inc.
US Foodservice, Inc.
Theodor Willie Intertrade
Silver
3M Company
AM General LLC
American Purchasing Services, Inc. (American Medical Depot)
Canadian Commercial Corp
DMS Pharmaceutical Group, Inc.
General Dynamics Land Systems
General Electric Aviation
Husky Marketing & Supply Co.
NACCO Materials Handling Group
National Industries for the Blind
OSHKOSH Corporation
Propper International, Inc.
SourceOne Distributors
W.S. Darley & Co.
Bronze
Bell Boeing Joint Project Office
Carter Enterprises LLC
Equilon Enterprises LLC
Graybar Electric Company, Inc.
McKesson Corporation
PAPCO, Inc.
Raytheon
Science Applications International (SAIC)
Seven Seas Shipschandlers LLC
Supplycore, Inc.
Washington Gas Energy Services

SOURCES: DLA (2014, 2015).

NOTE: Rankings are unordered within each tier and reflect performance of each business unit for the specified department or agency for the prior 3 years. Rankings include data from FY 2013 and FY 2014.

(This page was left blank intentionally)

3. Broader Factors Influencing Acquisition Outcomes

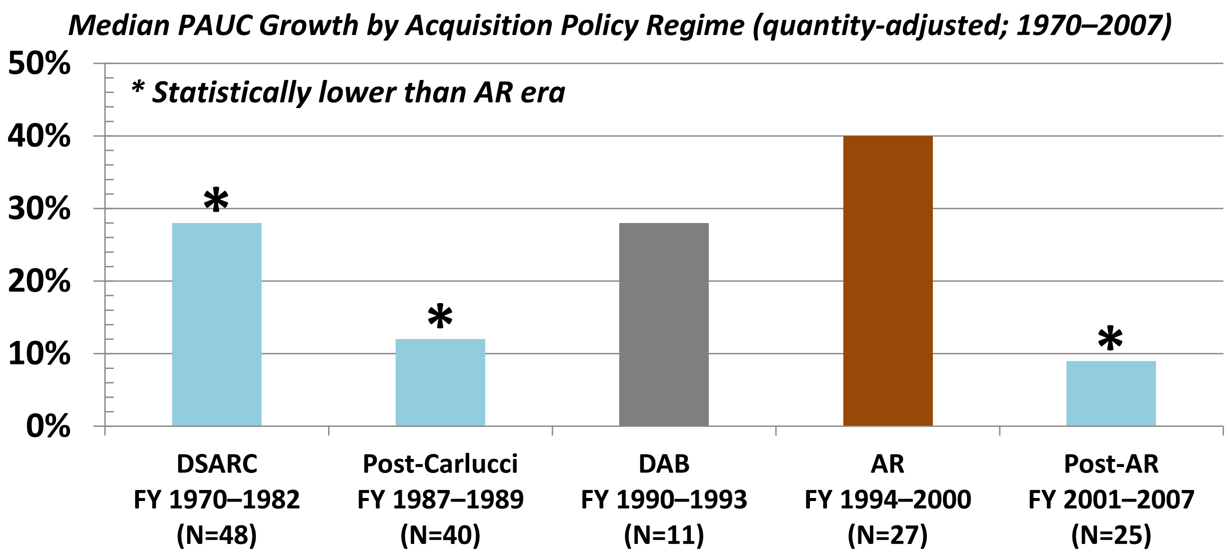
Acquisition Reform and Funding Climate Effects on Program Cost Growth

A series of recent Federally Funded Research and Development Center (FFRDC) studies at the Institute for Defense Analyses examined the potential effects of acquisition policy regimes and budget climates on MDAP cost growth.

Effects of Acquisition Policy Regimes on Program Cost Growth

First, McNicol and Wu (2014) compared PAUC growth between five broad acquisition policy regimes between 1970 and 2007 (see Figure 3-1). Programs that passed MS B or II during the reform regime of the mid-1990s had a statistically higher PAUC growth than the prior Defense Systems Acquisition Review Council (DSARC) and Post-Carlucci regimes as well as the post-AR regime of the early 2000s.

Figure 3-1. Program Cost-Related Performance in Policy Regimes

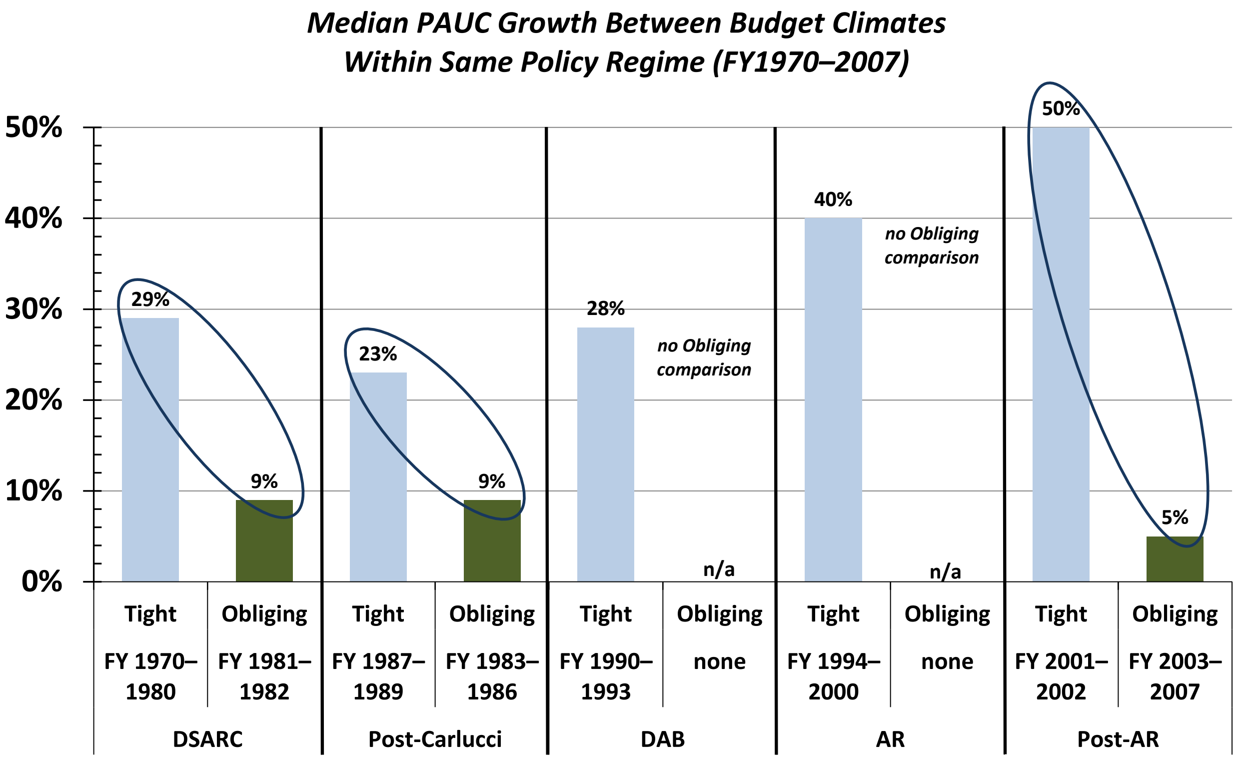


* Underlying distribution is statistically lower than that of the AR regime of 1994–2000. However, the Defense Acquisition Board (DAB) era distribution was not different enough from the AR era to be statistically significant.
Source: McNicol and Wu (2014). NOTES: The differences between the three starred era were not statistically significant. There may be further cost growth in these eras, particularly the post-AR regime. McNicol and Wu adjusted PAUC growth for any quantity changes.

Effects of Budgetary Climates on Program Cost Growth

McNicol and Wu (2014) also compared cost growth of MDAPs by MS B or II date between both acquisition regimes and budgetary climates (either relatively constrained or accommodating). They then examined the potential effect of these different policy regimes and budget climates by comparing the performance of MDAPs started in each (see Figure 3-2). Statistical tests found that PAUC growth did not correlate with policy regimes but it did with the budget climates. For each of the three policy regimes that started programs in both constrained and accommodating budget climates (i.e., DSARC, Post-Carlucci, and Post-AR in rows 1, 2, and 5), PAUC growth was significantly higher statistically for programs that passed MS B or II during a constrained year.

Figure 3-2. Effect of Budget Climates on Program Cost-Related Performance



Source: McNicol and Wu (2014).

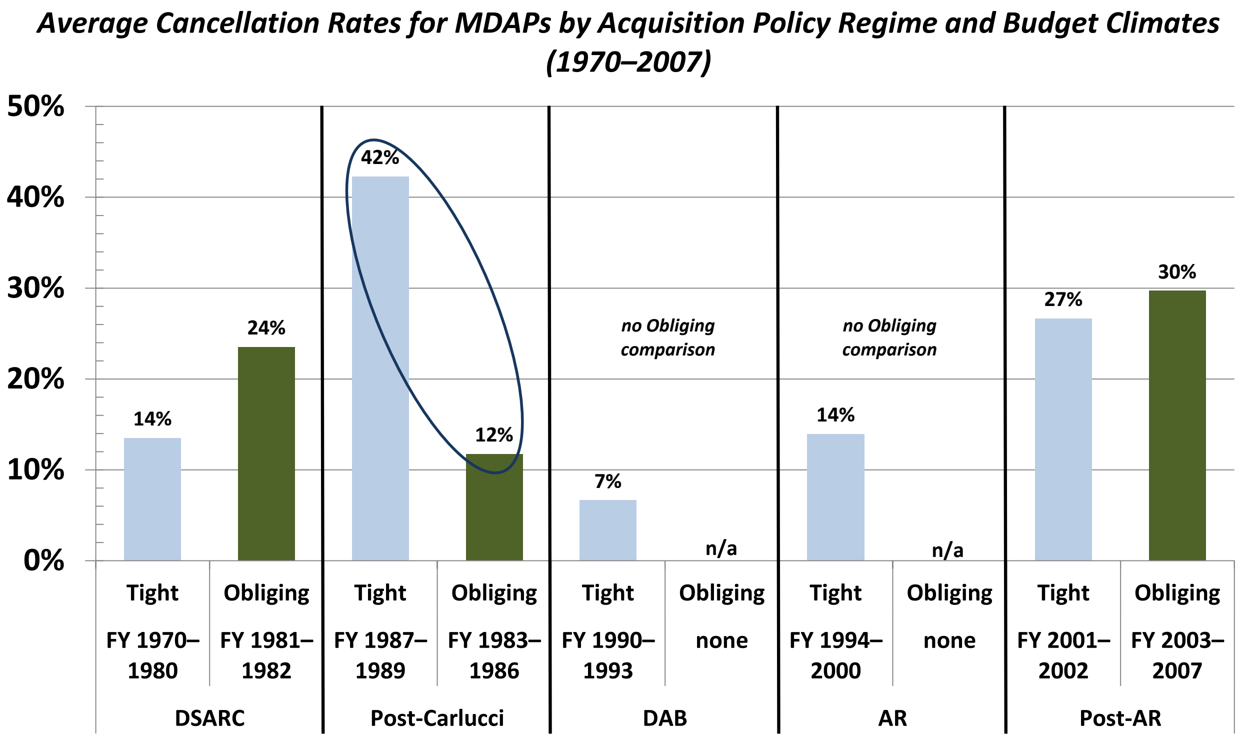
NOTES: Underlying distributions marked with ovals are statistically lower for the tight budget climate of the same acquisition regime. There were no obliging budget climates for the DAB and AR eras, so comparisons are not available in those two cases. McNicol and Wu adjusted PAUC for any changes in quantity. For consistency with the rest of our report and to better illustrate the central tendency, we report the median values instead of means from the distributions.

Effects on Program Cancellations

Further preliminary research by McNicol, Burns and Wu (2015) tested for effects of acquisition regimes, budgetary climates, and cost growth on program cancellations. As one would expect, annual cancellation rates increased statistically during relatively constrained budget periods. However, neither the budgetary climate nor the acquisition regime at program initiation (MS B or II) correlated generally with the likelihood of cancellation (see Figure 3-3, in which only one of three policy regimes with both types of budgetary climates showed a cancellation difference). Thus, tight budgets force decisions to cancel less important programs, but neither acquisition regimes nor budgetary climates at initiation seem to affect the likelihood of future cancellation. We may baseline programs too low during tight budgetary periods, but, if so, those programs generally remain about as important as other programs despite their higher cost growth.

Also, while it would seem likely that cost growth would drive cancellations, 80 percent of MDAPS with PAUC growth of at least 50 percent were not canceled. Thus cost growth alone does not appear to be the main drive of cancellations.

Figure 3-3. Effect of Budget Climates on Program Cancellation Rates



Source: McNicol, Burns and Wu (2015).

NOTE: The only statistically significant difference between cancellation rates within a policy regime was for the post-Carlucci regime, where 42 percent of those MDAPs started during tight budget periods were eventually cancelled compared to 12 percent of those MDAPs started during the subsequent obliging budget period (a factor of 3.5 times higher).

The results of these FFRDC studies suggest that budgetary climates have a larger effect on both cost growth and cancellation rates than acquisition policies. Further analysis is ongoing to better understand these correlations and to ascertain whether they hold under other program cost measures.

AFFORDABILITY

BBP established an explicit policy and process codified in DoDI 5000.02 (USD[AT&L], 2015a) for determining and ensuring the long-term affordability of the entire life-cycle costs of each system to be acquired. Affordability analysis and constraints form procurement and sustainment budget controls on the system throughout the FYDP and beyond. Constraints are determined in a top-down manner by the resources a DoD Component can allocate for a system, given inventory objectives and all other fiscal demands on the Component against a long-term future total budget projection. Constraints (especially the caps established at the Development Request for Proposals Release Decision Point before MS B) constitute a threshold for procurement and sustainment costs that cannot be exceeded by the PM. When affordability constraints cannot be met—even with aggressive cost control and reduction approaches—then technical requirements, schedule, and required quantities must be revisited (e.g., with support from the DoD Component's Configuration Steering Board). The program will be canceled if constraints still cannot be met, and the Component cannot offset cost increases by lowering the caps on other programs.[33]

It is too early to test the effect of the affordability process on outcomes from the acquisition system (e.g., whether program cancellations are made earlier and thus reduce sunk costs), but we have begun to measure the degree to which the affordability analysis and constraints has been implemented in the DoD. All three Military Services have established an affordability analysis capability, centered in their staff directorates for financial management and resource planning (i.e., G-8, N-8, and A-8 for the departments of the Army, Navy, and Air Force, respectively) in support of Service leadership decision making.

Affordability constraints have been imposed on 29 MDAPs (including pre-MDAPs that have not yet passed MS B) that have undergone major reviews between November 2010 and January 2015. Thus, about 70 percent of MDAPs programs have been assigned an affordability constraint commensurate with the policy in place at the time. The ratio is much lower for MAIS, however, at about 20 percent. The DoD is working to improve the consistency of applying the affordability policy to both MDAPs and MAIS.

In addition, USD(AT&L) tracks affordability constraints and compares program performance against those constraints on a regular basis. We are reviewing these programs and refining how best to handle each individual case given needs and budgetary constraints.

[33] Independent of affordability constraints or cost estimates, the DoDI 5000.02 instructs PMs to always look for ways to control or reduce cost. The BBP Should-Cost initiative provides organizational incentives and rewards for Components and PMs to continue looking for cost reductions below the affordability constraints while the Nunn-McCurdy breach process provides a strong organizational disincentive on excessive growth.

INFERRED MEASURES OF COMPLEXITY AND RISK IN CONTRACTS

Data related to complexity and risk are not readily available across all MDAP programs and their contracts. As we discussed earlier, we can use an existing variable such as planned schedule (cycle time) to try and infer complexity. While insightful, this simple metric can be problematic in that other factors could lead to shorter schedules besides a reduction in complexity.

Thus, to further test these results and seek a separate measure for risk, we employed the more sophisticated technique of factor analysis to develop new inferred measures of unobserved complexity and risk based on our readily available cost and schedule data (see analytic details in Appendix B starting on p. 155). Since it includes measures such as price and overruns (cost-over-target), it should conceptually be a better indicator of complexity. Note that these inferred measures are indices standardized to the distribution across all contracts in our dataset. Thus, unlike other proxies such as source-lines-of-software-code or vehicle weight, they have no interpretive scale other than to the distribution in the contracts (i.e., 0.0 is at the mean [average] and 1.0 is one standard deviation above the mean) and can have positive or negative values since they relate to the mean. Also, like these other proxies, these are not perfect but do provide useful insights.

Below are the resulting unobserved synthetic variables (factors) of complexity and risk on MDAP development and early production contracts from 2000–2015. Since we cannot prove they are, in fact, complexity and risk, we will identify them as the variables *C* and *R* but will use them to gain inferred insights into contract complexity and risk trends. We also analyzed whether these synthetic variables plus our other available contract variables correlate with contract price growth. Note that the MDAP contract dataset (used in the analysis in the prior chapter) was controlled for inflation, contract maturity, and statistical outliers so that these results would better reflect the central defense acquisition system tendency on contracts. We also controlled for phase differences by analyzing development contracts separately from early production contracts.

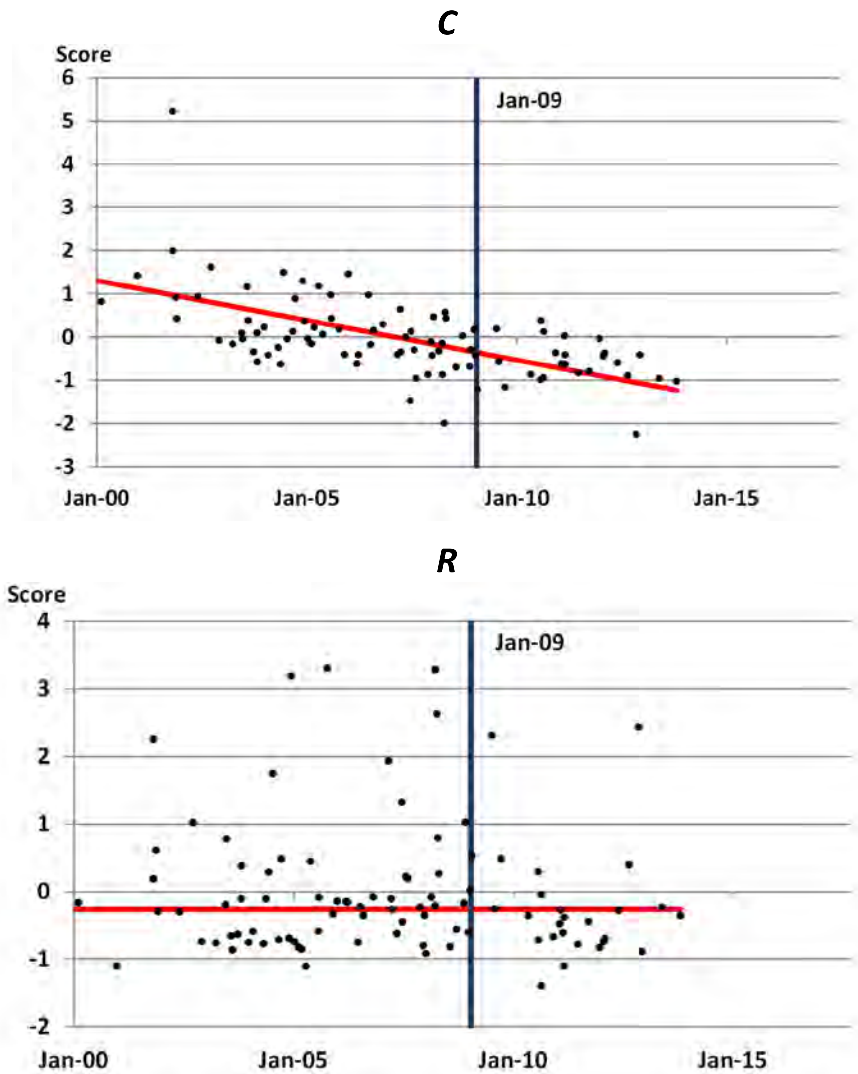
Inferred Trends in Contract Complexity and Risk: Development

Using the synthetic variables *C* and *R* developed using factor analysis, to try and see if complexity and risk have been changing over time in development. Figure 3-4 examines trends for *C* and *R*, respectively, on MDAP development contracts. Each variable's standardized value is plotted based on the contract's work start date. Over the last 15 years, there has been a consistent and statistically significant downward trend in *C* and thus potential in the complexity of our development contracts. This supports the earlier discussion that there may be a significant downward trend in development contract complexity when using planned schedule (cycle time) as the measure (see Figure 2-39 on p. 74).

For the variable *R*, however, there was no statistically observable trend in MDAP development contracts. This indicates that the DoD may not be taking more risks over time on our contracts.

However, as we saw in price and cost growth, there is wide variation in performance (i.e., the central tendency is toward low risk, but the levels vary widely).

Figure 3-4. Trends in *C* and *R* on MDAP Development Contracts for Currently Active MDAPs (2000–2015)



NOTES: The downward trend in *C* is statistically significant, but there was no statistically significant trend in *R*. N = 94 development contracts. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency.

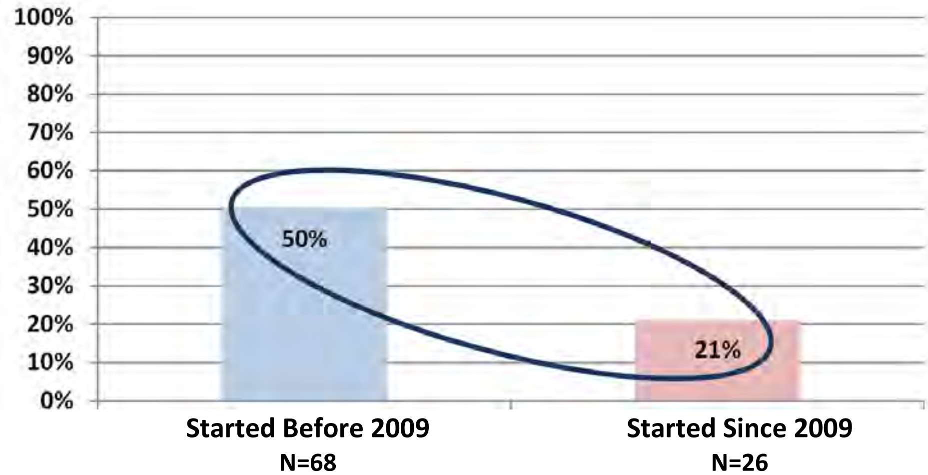
Proportion analysis also supports the result above that, as measured by the synthetic variable *C*, the complexity of development contracts may have been on a significant downward trend while risk (as inferred by the measure *R*) has been relatively flat. Figure 3-5 shows that the proportion of MDAP contracts started since 2009 with low *C* scores (below zero, which is the average for the standardized score) was substantially higher compared to earlier contracts, and the difference was statistically significant. This supports the earlier inference of a significant downward trend in development contract complexity when using planned schedule (cycle time) as a measure (see Figure 2-39 on p. 74). The share of contracts with low risk (as measured by *R*) was somewhat higher for contracts started since 2009, but the difference was not statistically significant.

Figure 3-5. Proportion of Development Contracts with Below-Average *C* and *R* Started Before and Since 2009 for Currently Active MDAPs (2000–2015)

NOTES: The proportional difference in *C* is statistically significant but not in *R*. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency.

To further investigate whether and why complexity may be decreasing over time, we examined whether there has been a shift away from starting new weapon systems towards more modifications and upgrades. Figure 3-6 compares the proportion of development contracts that are for new systems (as opposed to modifications of systems developed and produced earlier) before and since the start of 2009. Clearly, the DoD has been starting far fewer new systems (and thus more upgrades) since 2009, and the difference is statistically significant. This contractual trend is supported by the budget trend shown earlier in Figure 1-3, wherein RDT&E budget account 6.7 (Operational Systems Development for Upgrades) has been relatively stronger than budget account 6.5 (System Development and Demonstration for New Systems).

Figure 3-6. Proportion of Development Contracts for New Systems (Versus Upgrades) Before and Since 2009 on Currently Active MDAPs (2000–2015)

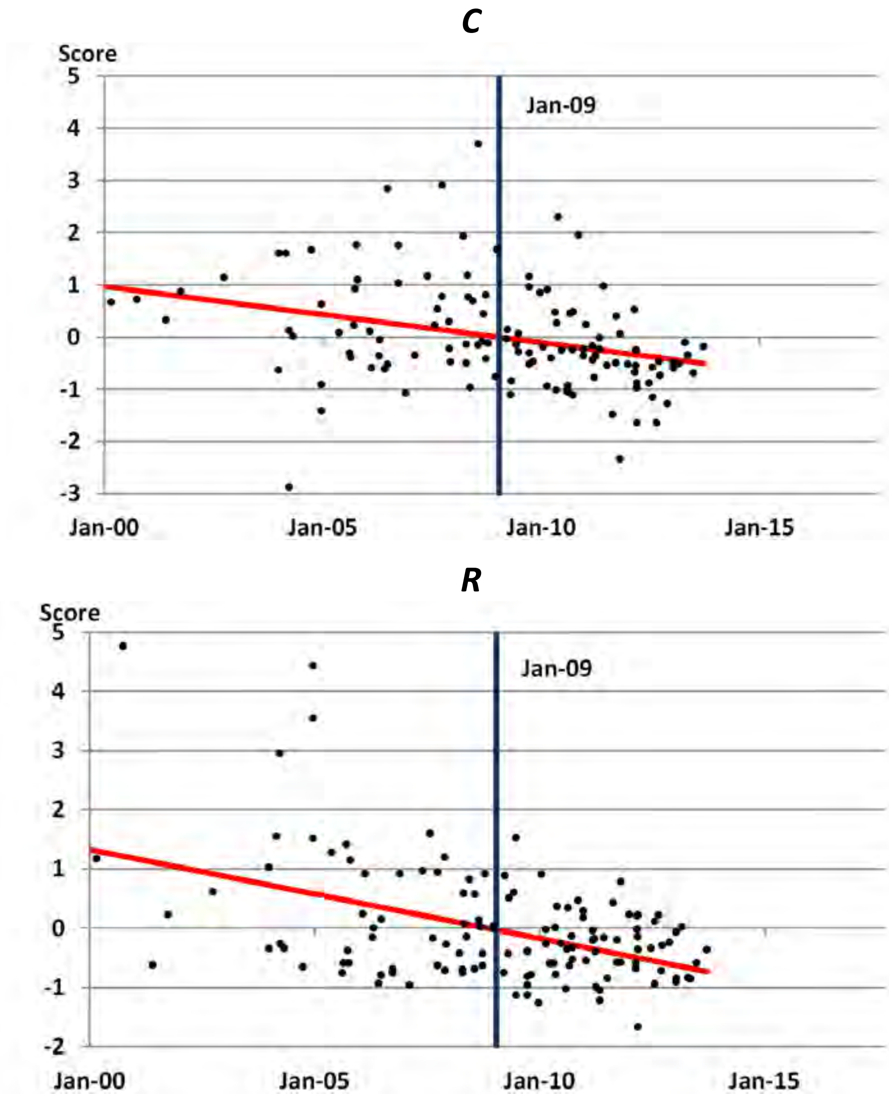


NOTES: The proportional difference is statistically significant. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency.

Inferred Trends in Contract Complexity and Risk: Early Production

Next, we used the synthetic variables *C* and *R* to try and see if complexity and risk have been changing over time in early production. Figure 3-7 examines trends for *C* and *R*, respectively, on early production MDAP contracts. Each variable's standardized value is plotted based on the contract's work start date. Over the last 15 years, there has been a consistent and statistically significant downward trend in both *C* and *R*; thus, complexity and risk in MDAP contracts started since 2009 may be declining in early production. Again, this supports the earlier discussion that there may be a significant downward trend in development contract complexity when using planned schedule (cycle time) as the measure. Also, the data on *R* may indicate we are doing a better job of reducing risk (e.g., maturing technology and resolving manufacturing risks) before going into production.

Figure 3-7. Trends in *C* and *R* on Early Production Contracts for Currently Active MDAPs (2000–2015)



NOTES: The downward trends in *C* and *R* are statistically significant. N = 130 early production contracts. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency.

Finally we made a proportional comparison of *C* and *R* in early production contracts started before and since 2009. The results of this analysis are provided in Figure 3-8. The shares of contracts with low *C* and *R* were substantially higher for contracts started since 2009, and the difference was statistically significant.

Figure 3-8. Proportion of Early Development Contracts with Below-Average *C* and *R* Started Before and Since 2009 for Currently Active MDAPs (2000–2015)

NOTES: The proportional differences in *C* and *R* are statistically significant. Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency.

Correlates of Price Growth

We next used the new synthetic variables of *C* and *R* (our inferred measures of complexity and risk derived using factor analysis), together with our other available variables in the data, to test which variables statistically predict inflation-adjusted price growth on MDAP contracts since 2000.

We first determined correlates of price growth on the 94 MDAP development contracts in our dataset. In development, there are three variables—work-content growth, *C*, and *R*—correlate with real (inflation-adjusted) price growth. All other available variables for price growth were found to be statistically insignificant or spurious.

To illustrate the results of this regression, we show in Table 3-1 the percentage-point increase in real contract price growth would be expected from the median amount of work-content growth, *C*, and *R* across the set of development contracts when controlling for spending (i.e., on a dollar basis). The average (median) amount of work-content growth, *C*, and *R* contributed 0.5, 24, and 10.5 percentage points (respectively) to price growth in our dataset of 94 development contracts. This was on top of a regression base constant of 20 percentage points; in other words, a contract with no work-content growth and average *C* and *R* indices (i.e., values of 0.0) statistically would have had a price growth of about 20 percent.

Table 3-1. Potential Program Contract Price Drivers: Development

Correlates of Contract Price Growth in Development for Currently Active MDAPs (adjusted for inflation; 2000–2015)

Median amount of:	Contributes these percentage-points to contract price growth
work-content growth	0.5
C	24
R	10.5
base constant	20
Total "average" predicted price growth:	55%

NOTES: The constant can be interpreted as predicting the amount of price growth with no work-content growth (i.e., no contract modifications), average *C* (inferred average complexity), and average *R* (inferred average risk). Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency. N = 94 development contracts; R^2 = 0.86 for this regression. Price growth was not controlled for changes in work.

We also determined correlates of price growth on early production contracts. In early production, we also found that real (inflation-adjusted) price growth depended only on work-content growth, *C*, and *R*. All other possible explanatory variables for price growth were found to be statistically insignificant or spurious.

To illustrate the results of this regression, we show in Table 3-2, the percentage-point increase in real contract price growth that would be expected from the median amount of work-content growth, complexity, and risk across the set of early production contracts when controlling for spending (i.e., on a dollar basis). The median amount of *C* and *R* contributed 4 and 3 percentage points (respectively) to price growth in our dataset of 130 early production contracts. The median amount of inflation-adjusted work-content growth was negative, so that contributed an additional −3 percentage points (i.e., a reduction at the median). This was on top of a regression base constant of 7 percentage points; in other words, a contract with no work-content growth and average complexity and risk indices (i.e., values of 0.0) statistically would have had a price growth of about 7 percent.

As expected, price growth is dramatically lower in early production than in development. Also as we would expect, the inferred contributions of complexity and risk (based on the synthetic variables *C* and *R*) to price growth were much lower in early production than in development.

Table 3-2. Potential Program Contract Price Drivers: Early Production

Correlates of Contract Price Growth in Early Production for Currently Active MDAPs (adjusted for inflation; 2000–2015)

Median amount of:	Contributes these percentage-points to contract price growth
work-content growth	–3*
C	4
R	3
base constant	7
Total "average" predicted price growth:	11%

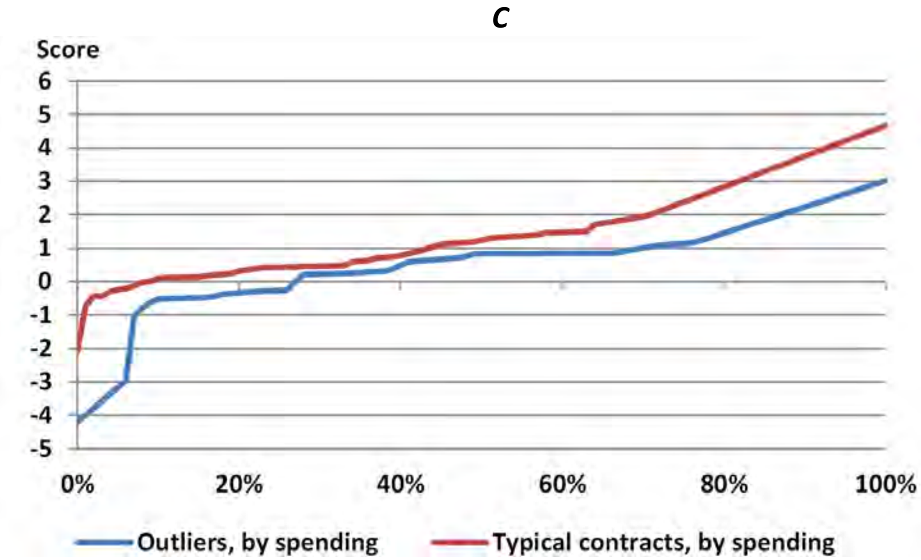
*The median amount of work-content growth from contract modifications after adjusting for inflation was negative, hence the median predicted contribution to price growth was negative.

NOTES: The constant can be interpreted as predicting the amount of price growth with no work-content growth (i.e., no contract modifications), average *C* (inferred average complexity), and average *R* (inferred average risk). Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency. N = 130 development contracts; R^2 = 0.93 for this regression. Price growth was not controlled for changes in work.

Comparing Inferred Complexity and Risk between Cost Outliers and Typical Contracts

The following analysis shows how cost-growth outlier contracts in development may differ from more typically performing contracts with respect to complexity and risk as defined by the synthetic variables *C* and *R* discussed above. Figure 3-9 graphs the percentile-by-percentile dollar distributions of outlier and typical development contracts controlling for contract size in terms of spending (dollars) and contract maturity. This shows that, dollar-for-dollar, typically performing contracts have higher *C* scores and thus may be more complex than the cost-growth outliers in this sample of development contracts. This difference is statistically significant and substantial. In other words, our cost outliers apparently may be outliers not because they are more complex but for some other reason. We intend to conduct more analysis in this area, as the outlier results are the ones we most want to avoid.

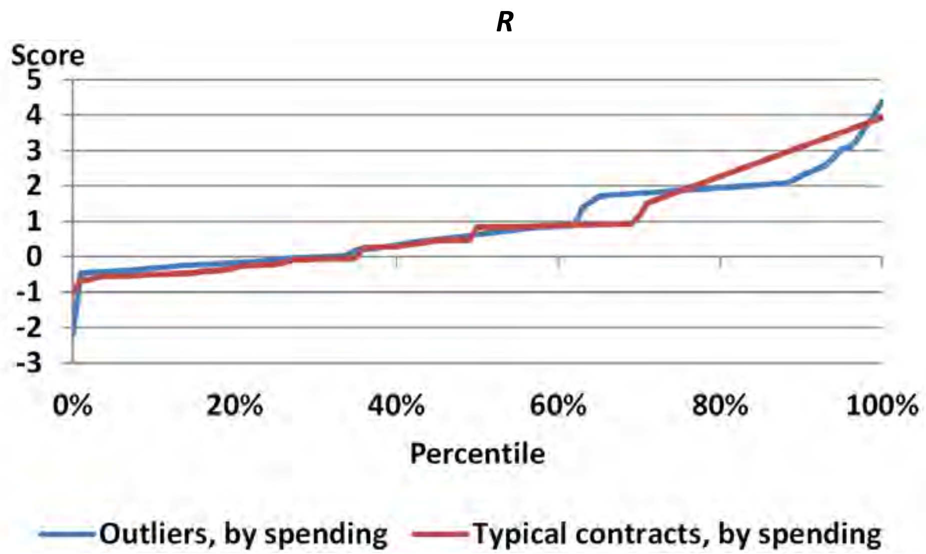
Figure 3-9. Distributions in *C*: Development-Contract Cost-Growth Outliers and Non-Outliers for Currently Active MDAPs (dollar basis; 2000–2015)



NOTE: Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency.

Similarly, Figure 3-10 compares percentile-by-percentile dollar distributions of risk for cost-growth outlier and typical development contracts. In this case, outlier contracts had systematically higher risk. The difference was statistically significant but not very large.

Figure 3-10. Distributions in *R*: Development-Contract Cost-Growth Outliers and Non-Outliers for Currently Active MDAPs (dollar basis; 2000–2015)



NOTE: Statistical outliers (both high and low) were removed to obtain more reliable measures of central tendency.

Based on the values of the synthetic variables *C* and *R*, it seems that, dollar-for-dollar, it is possible for complex contracted efforts to be better planned, estimated, and executed as evidenced by better cost control. Cost-growth outliers may have slightly higher risk, but well-controlled contracts can have high risk and manage it well.

Analysis of Subcontractor Margins

In last year's report we analyzed prime contractor final margins. In this year's report we extend our analysis to subcontractor margins and how they compare to margins on prime contracts.[34] These are the large-scale subcontractors for which we have detailed cost data—not those providing commercial items. To control for possible phase differences, we analyzed margins on development subcontracts separately from production subcontracts. In development, our data consist of 113 subcontracts with 37 associated prime contracts. In production, our data consist of 44 subcontracts with 27 associated prime contracts. All subcontracts were tier-one (i.e., directly for the prime contractor). While the results discussed below are statistically significant for the sample of MDAP contracts for which data are available, further analysis is needed to better understand the reasons for these differences.

Comparing Margins on Subcontracts to Associated Prime Contracts

Since 2001, first-tier subcontractors earned higher margins than their associated prime contractors on the same program (at the medians, about 2 percentage points higher in development and about 7 percentage points higher in production—see Figure H-20 on p. xxxiii).[35] While not uniform, the differences are statistically significant and are particularly large in production. Further analysis is needed to understand what the causes may be for these differences, but these illustrate why the DoD has been working over the past few years to motivate prime contractors to control subcontractor prices and ensure that profitability is aligned with performance—especially in production where the difference in margins is large. We also are concerned that higher subcontract margins may be motivating companies to bid on fewer prime contracts and thus reduce competition at that level.

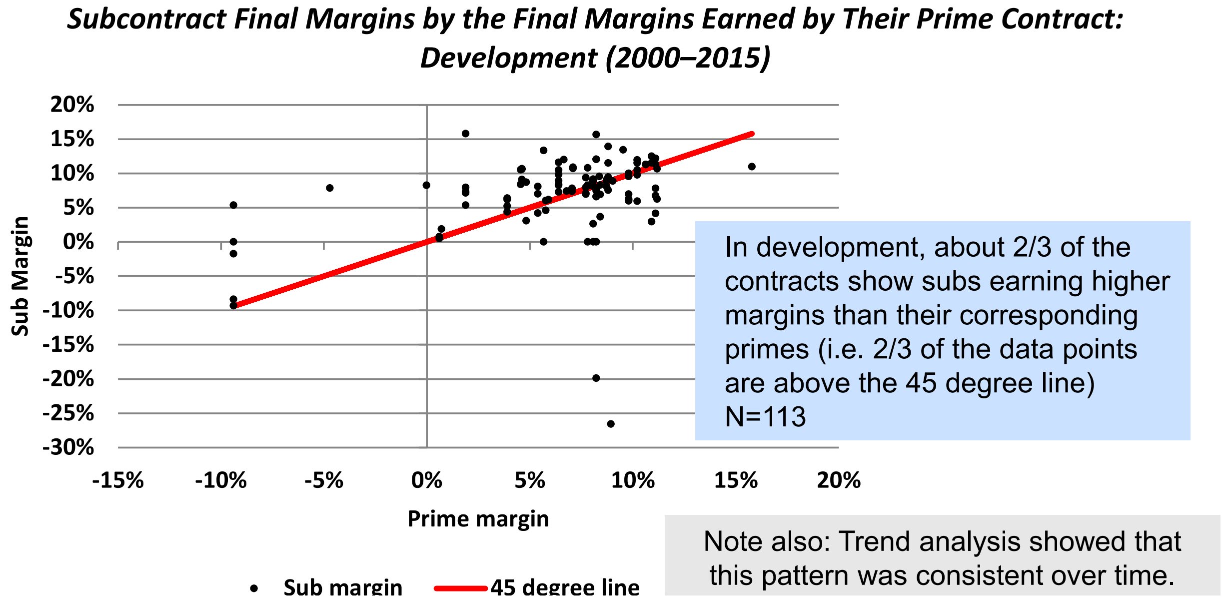
Figure 3-11 shows a scatter plot of prime margins and their associated tier-one subcontracts in development. The 45-degree line delineates the boundary in the data where prime contract and subcontract final margins would be exactly equal. Dots above the line indicate program pairs where the subcontractor's final margin was greater than the prime contractor's final margin. Symmetrically, plots below the 45-degree line indicate program pairs where the subcontractor's final margin was less than the prime contractor's final margin. Clearly a statistically significant majority of the plots (67 percent) show the subcontractors earning higher margins in development. Moreover, separate trend analysis found that this pattern has been consistent over time.

---

[34] Note that we calculated these as margins (fraction of price going to fee or profit), not markups (fraction of cost going to fee or profit). Markups at these percentages are slightly higher.

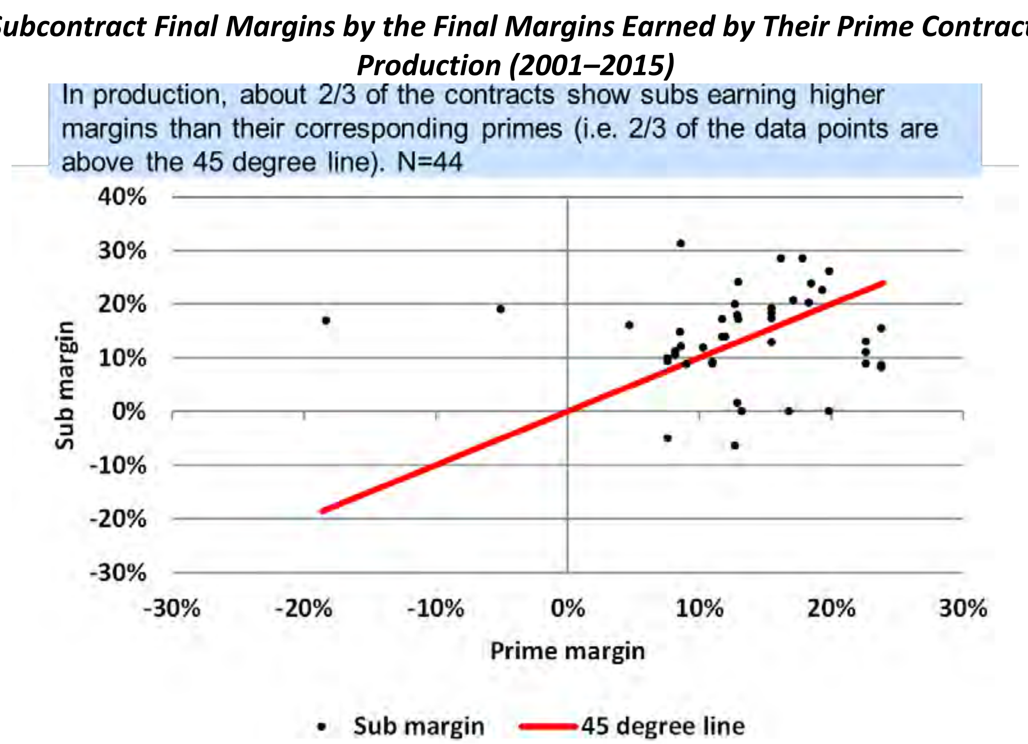
[35] Note that we calculated these as margins (fraction of price going to fee or profit), not markups (fraction of cost going to fee or profit). Markups at these percentages are slightly higher.

Figure 3-11. Paired Subcontractor and Prime Contractor Final Margins: Development



For production contracts we see a similar result. Figure 3-12 shows a scatterplot of prime margins and their associated tier-one subcontracts. The 45-degree line has exactly the same interpretation as above. Again, a statistically significant majority of the entries (64 percent) show the subcontractors earning higher margins in production than their prime contractors. As in development, separate trend analysis found that this pattern has been consistent over time.

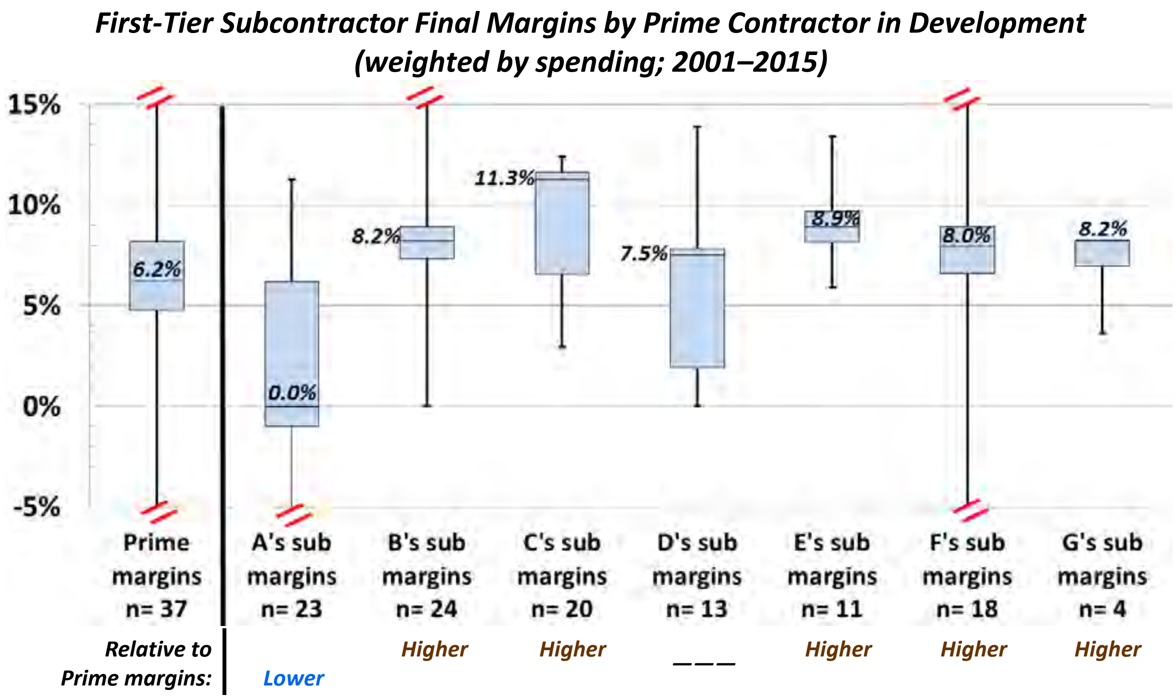
Figure 3-12. Paired Subcontractor and Prime Contractor Final Margins: Production



Subcontractor Margins by Prime Contractor

We next assessed how well specific prime contractors may have negotiated good prices with their first-tier subcontractors as evidenced by subcontractor final margins. Figure 3-13 shows the relative ability of prime contractors to control the margins earned by their subcontractors in development. As a basis to compare subcontract margins, we used the distribution of all prime contractor final margins as a group (the leftmost column of the table). Prime contractors who generally negotiate subcontracts with margins at or below what the primes generally earn may have done a better job of negotiating and controlling subcontracting costs than those that do not. Further analysis is needed.

Figure 3-13. Subcontractor Final Margins by Prime Contractor: Development

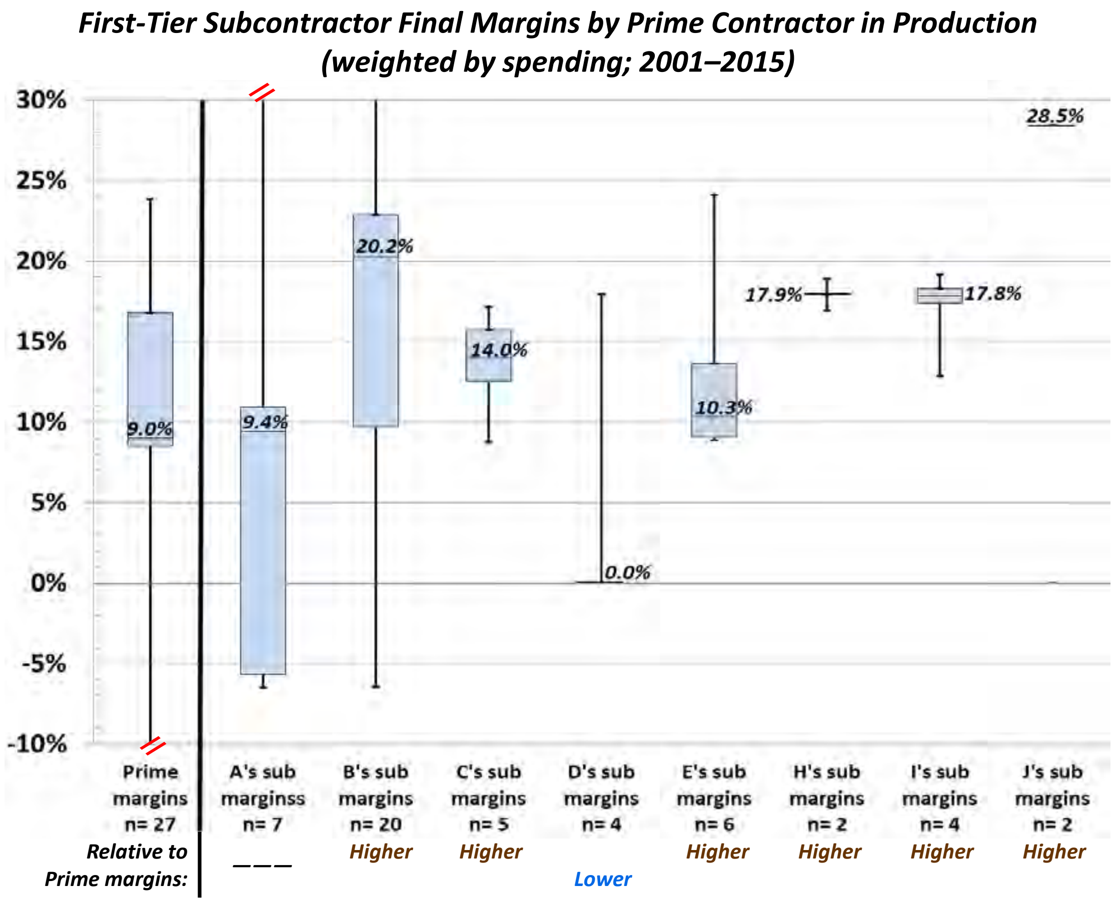


NOTES: While the sample sizes are fairly small, the relative comparisons (higher or lower) against the set of all prime contractor final margins are statistically significant. First-tier subcontractors under prime contractor A earned systematically lower final margins than all prime contractors as a group (i.e., the distribution under A is statistically lower than that under the "All prime margins" column). First-tier subcontractors under prime contractor D were statistically similar to those earned by all prime contractors. The rest of the prime contractors had first-tier subcontractors that statistically earned higher final margins than all primes as a group.

In the development dataset, only prime contractor *A*'s subcontract final margins were systematically below what the prime contractors generally earned. Subcontracts under prime contractor *D* were comparable statistically to the prime-contractor margins. The subcontract final margins for the remaining prime contractors (*B*, *C*, *E*, *F*, and *G*) were all systematically higher than what all prime contractors earned, and the differences were statistically significant. This may suggest that most prime contractors do not sufficiently control subcontractor costs, or that prime contractors lack enough information to negotiate better prices. It also suggests that there may be an incentive for contractors to prefer being a subcontractor over being a prime contractor in development phases. Alternatively, it could be that there is less required investment or work to do as a prime contractor in certain instances and thus they are willing to accept lower final margins.

Figure 3-14 shows the relative ability of prime contractors to control the margins earned by their subcontractors in production. As before, we used the distribution of the final margins that the prime contractors themselves earned (in the leftmost column of the table) as the basis to compare relative prime contractor control of subcontract margins.

Figure 3-14. Subcontractor Final Margins by Prime Contractor: Production



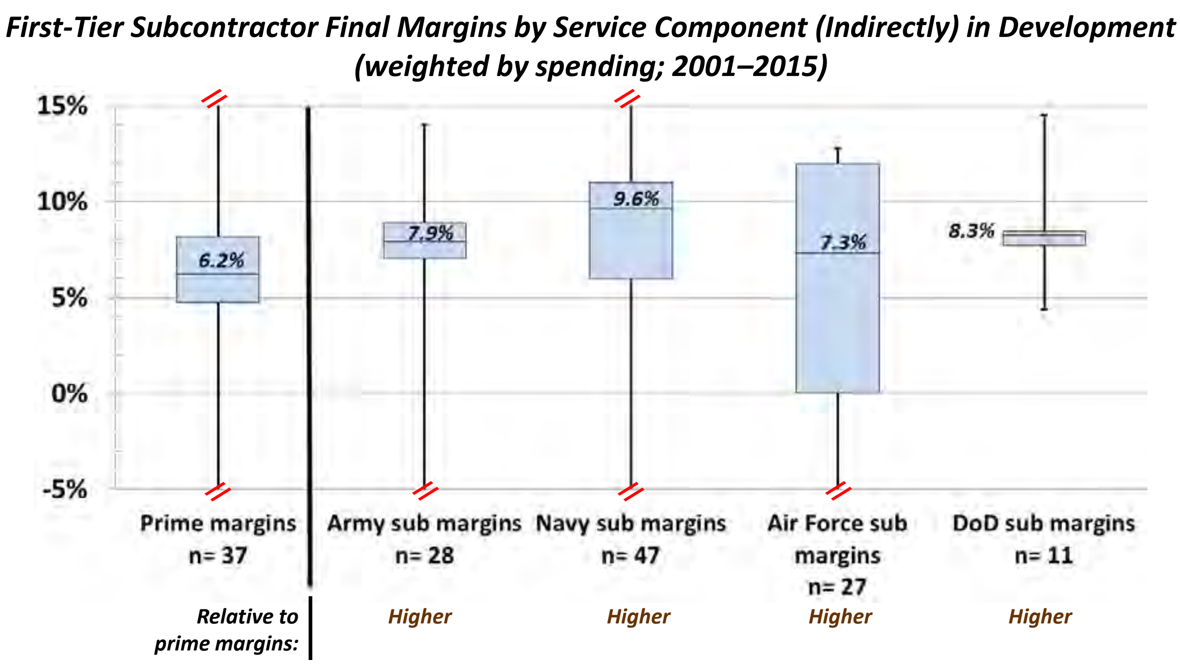
NOTES: While the sample sizes are fairly small, the relative comparisons (higher or lower) against the set of all prime contractor final margins are statistically significant. First-tier subcontractors under prime contractor *D* earned systematically lower final margins than all prime contractors as a group (i.e., the distribution under *D* is statistically lower than that under the "All prime margins" column). First-tier subcontractors under prime contractor *A* were statistically similar to those earned by all prime contractors. The rest of the prime contractors had first-tier subcontractors that statistically earned higher final margins than all primes as a group. There was insufficient data on subcontracts from prime contractors *F* and *G* from Figure 3-13 to show results here for production. Both subcontractor margins under contractor *J* had the same value shown.

In the production dataset, only for prime contractor D were tier-one subcontract final margins systematically below what prime contractors generally earned. Subcontracts under prime contractor *A* were comparable statistically to the prime-contract margins. For the remaining prime contractors in our dataset (*B*, *C*, *E*, *H*, *I*, and *J*) earned systematically higher final margins that all these prime contractors tended to earn in production, and the differences were statistically significant. As in development, this suggests that either primes do not sufficiently control subcontractor costs in production or they lack sufficient information to negotiate better prices. It also suggests that there may be an incentive for contractors to prefer being a subcontractor over being a prime contractor—particularly in production, where the margins are significantly higher and have been above 20 percent. Alternatively, it could be that there is less required investment or work to do as a prime contractor in certain instances and thus they are willing to accept lower final margins.

Subcontractor Margins by Component

We now test to see if subcontractor margins differ by DoD Component. Figure 3-15 shows the relative premiums that subcontractors earn in margins greater than that which the primes generally earn in development (again, the leftmost column), broken out by Service Component. In the development dataset, across all Service Components the subcontractors earn systematically higher margins than primes generally earn. This premium statistically was significant and substantial.

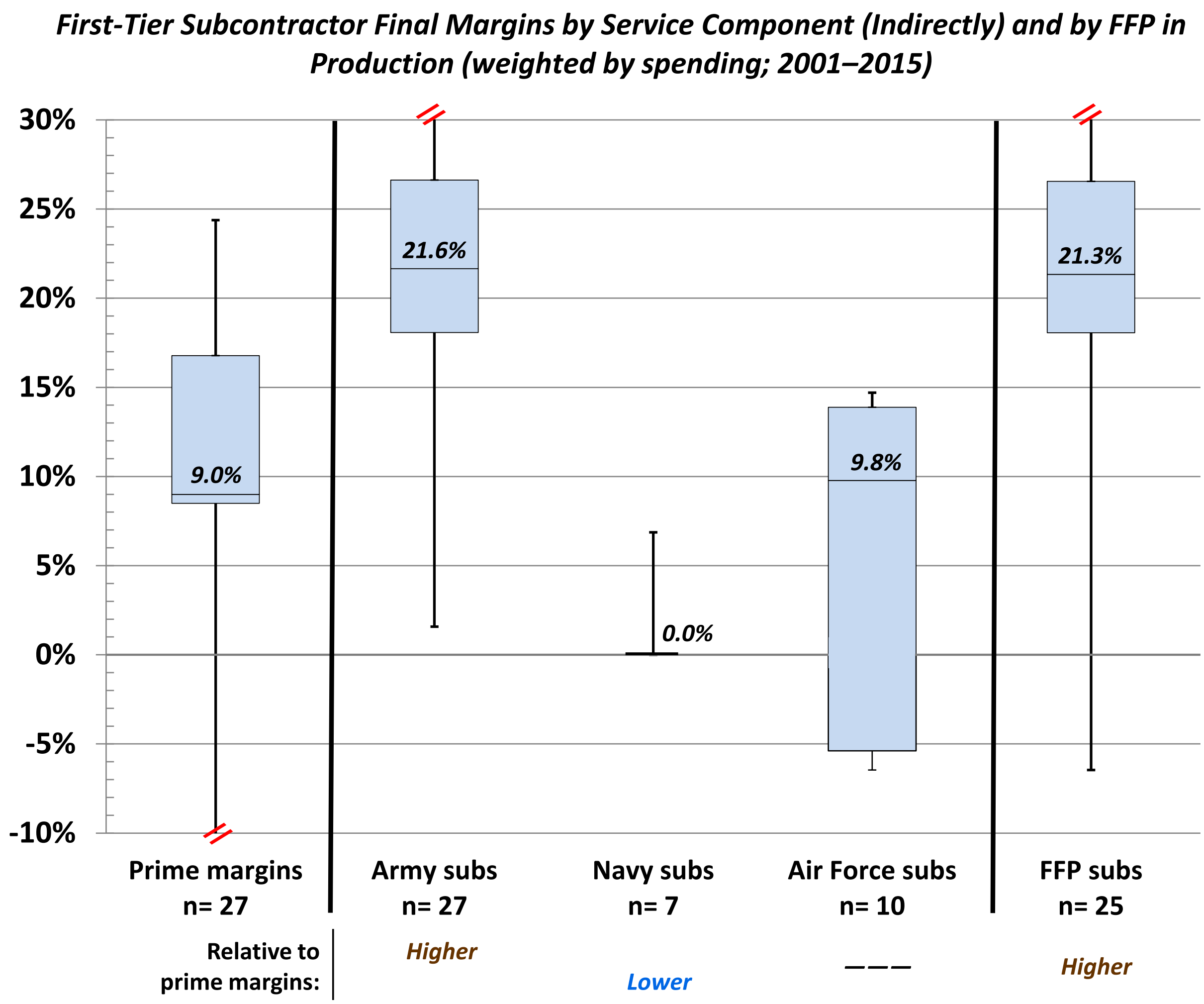
Figure 3-15. Components Whose Subcontractors Earn Higher Final Margins: Development



NOTES: While the sample sizes are fairly small, the relative comparisons (higher or lower) against the set of all prime contractor final margins are statistically significant. First-tier subcontractors for all Components earned systematically higher final margins than all prime contractors as a group (i.e., the distribution under each Component is statistically higher than that under the "Prime margins" column on the left).

Similarly, Figure 3-16 shows the relative premiums that subcontractors earn in final margins greater than that which the prime contractors generally earn in production (leftmost column), broken out by Service Component. In the production dataset, Navy subcontractors earned systematically lower final margins than the prime contractors. Air Force subcontractors earned final margins comparable to those of the prime contractors. However, Army subcontractors, and subcontractors for prime contractors who had a FFP contract, earned systematically higher final margins than the prime contractors did. These results statistically were significant and substantial.

Figure 3-16. Components Whose Subcontractors Earn Higher Final Margins: Production



NOTES: While the sample sizes are fairly small, the relative comparisons (higher or lower) against the set of all prime contractor final margins are statistically significant. The first-tier subcontracts for the Navy in this sample earned systematically lower final margins than all prime contractors as a group (i.e., the distribution is statistically lower than that under the "Prime margins" column on the left). The first-tier subcontracts in our dataset earned final margins comparable to those of all prime contractors. First-tier subcontracts for the Army as well as those using FFP contracts earned statistically higher margins. Nearly all FFP subcontracts (23 of the 25) in this dataset were for the Army, which probably explains why the Army subcontracts had systematically higher final margins.

(This page was left blank intentionally)

4. Input and Process Measures

Finally, we examine input and process measures of the defense acquisition system. While these do not directly reflect the outputs and outcomes, some are contentious topics for which it is useful to examine the available data, while others such as competition rates and small-business utilization have conceptually important ties to outcomes and thus have assigned goals.

First, we discuss the data on efforts to improve the defense acquisition workforce.

Second, we discuss two views of acquisition source selection: bid-protest rates and some data on the preponderance of LPTA source selections on services contracts.

Third, we present the results on two measures of inputs into the defense acquisition system for which the DoD has established goals: competition rates and small-business participation. While these are not direct outputs or outcomes, both should improve price and cost control as well as innovation (e.g., see last year's report for analysis related to competition effects). Also, both have significant congressional interest.

Finally, we examine government productivity and execution funding using three preliminary macro measures described below.

Defense Acquisition Workforce

Basic principles assert that workforce capabilities are key to the performance of the acquisition system. In 2012, "Increasing Professionalism" was a major addition to version 2.0 of the BBP initiatives. As a result, we begin including data on the workforce, starting with this annual report. The DoD is improving our data tracking to enable future analysis of the effects of our workforce investments in acquisition outcomes such as cost, schedule, and technical performance of our acquired systems.

Improving the Workforce

The Defense Acquisition Workforce Improvement Act of 1990 (with significant changes in 2003) established a requirement in the U.S.C. for the DoD to establish policies and procedures for the effective management (including accession, education, training, and career development) of the acquisition workforce. More recently, the Defense Acquisition Workforce Development Fund (DAWDF) was established by the FY 2008 NDAA to provide funding and support to help DoD develop the workforce capacity and capability to carry out its acquisition mission, perform appropriate oversight and achieve best value for the expenditure of public resources. This includes extending the recruiting and outreach capabilities to find highly qualified candidates and the use of DAWDF to fund hiring despite short-term execution-year O&M funding shortfalls that would otherwise impede entry-level hiring of recent graduates. To this end, the DAWDF is a partial contributor to the following three major development categories. The BBP 2.0 initiatives included higher professional standards for key leaders and more active management of career paths to maximize professional growth opportunities. BBP 3.0, implemented last spring, continues these initiatives while adding emphasis on the professionalism of the engineering and scientific workforce.

Recruiting and hiring. Workforce capability begins with sufficient size. While it is difficult to analytically calculate the minimum size of a workforce at a macro level, direct observation has shown that the significant reductions in the mid-1990s of the acquisition workforce eliminated significant capabilities that have led to performance issues. Thus, one key use of the DAWDF, together with an "insourcing" campaign, was in rebuilding the acquisition workforce. DAWDF funding has led to the hiring of more than 10,000 new civilian hires—over one-third of the 23 percent growth in the civilian acquisition workforce from FY 2008–2011. Since then, the civilian acquisition workforce has been protected and remains essentially flat despite the budget pressures. In this environment the DoD has, for the last few years, focused on improving the capability of the existing workforce rather than on size increases. This has been the purpose of the BBP 2.0 and 3.0 "building professionalism" initiatives.

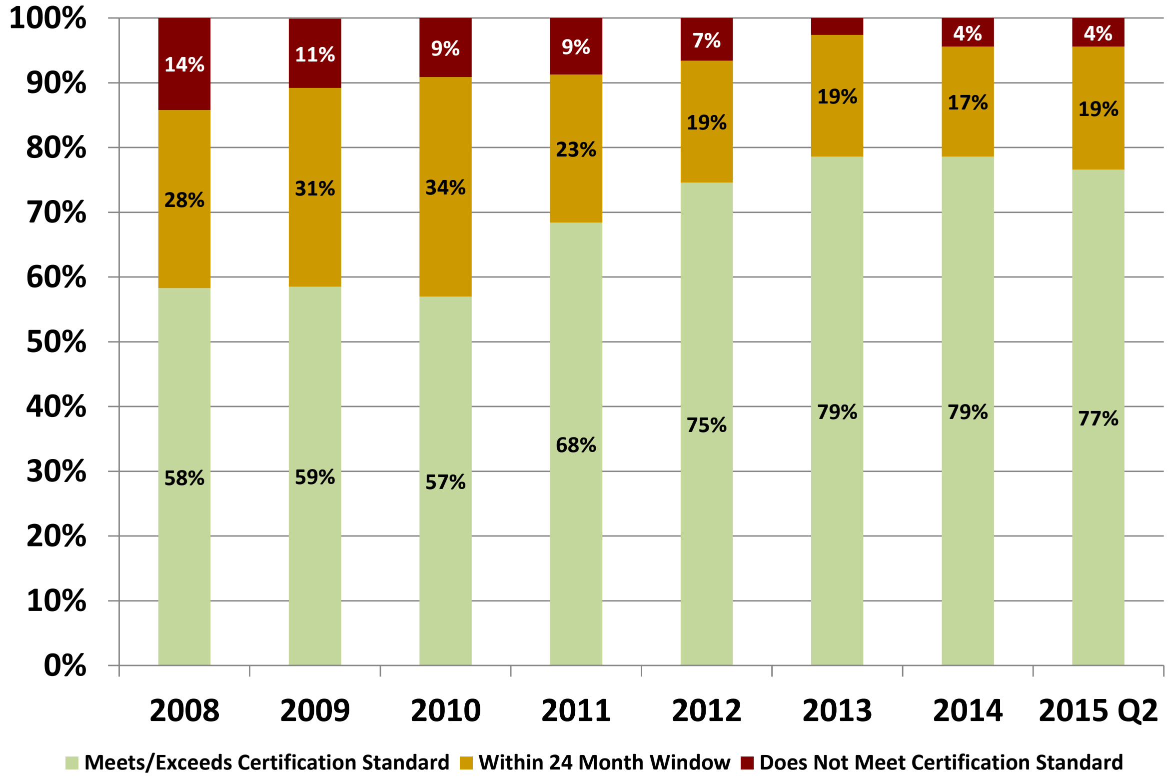
Training and development. Figure 4-1 illustrates some of the progress made in training by showing the percentage of acquisition workforce members already meeting certification standards for their current position or within the 24-month grace period since starting in the current position.[36] The fraction out of compliance has dropped from 14 percent to 4 percent while the fraction meeting standards has risen from 58 percent to 77 percent.

Retention and recognition incentives. Finally, incentives are used to enhance performance and facilitate career advancement. From FY 2010 through FY 2014, DAWDF funding supported 34,739 tuition assistance incentives, 7,277 student-loan repayments, 520 relocation incentives, 3,807 courses towards advanced academic degrees; and 6,002 incentives for recognition.[37] The education-related incentives have resulted in an increase from FY 2008 to FY 2014 in the fraction of the workforce with advanced degrees from 29 percent to 38 percent. The fraction of those with at least a Bachelor's degree has also risen from 77 percent to 83 percent.

[36] Civilian and military acquisition personnel must achieve certifications specific to their career field and position requirements. (Certifications are not available to contractor personnel.) Individuals who do not have the necessary certifications when starting a position must become certified within 24 months. While the courses required for certification depend on the specific acquisition career field, the acquisition courses cover managing risk in system acquisition, the JCIDS process, understanding statutory, regulatory and decisional contracting and acquisition law, as well as EV management and budgeting, among other topics.

[37] Details are contained in the Defense Acquisition Workforce Development Fund Annual Reports to Congress from FY 2010 through FY 2014.

Figure 4-1. Acquisition Workforce Meeting Certification Standards (2008–2015Q2)



Performance Incentives

The DoD has also expanded its awards programs and added a Should Cost and Innovation Award as part of BBP to directly recognize exceptional performance and help instill a cost-conscious culture. The David Packard Excellence in Acquisition Award (first awarded in 1997) is the DoD's highest acquisition team award and now emphasizes superior program management or successful execution of one or more BBP initiatives to reduce life-cycle costs.

The Should Cost and Innovation Award began in 2014 and recognizes outstanding commitment, innovation, and results in should-cost management.

Defense Acquisition Workforce Achievement Awards are given for excellence in 18 categories: acquisition in an expeditionary environment; auditing; financial management; cost estimating; contracting and procurement; engineering; facilities engineering; industrial and contract property management; information technology; life-cycle logistics; production, quality, and manufacturing; program management; science and technology management; test and evaluation; EV management; requirements management; services acquisition; and small business.

Key Leader Positions and Qualifications

Another major BBP workforce initiative establishes qualification standards and board certification for Key Leadership Positions (KLPs) on major acquisition programs. Not only does this further establish educational, experience, tenure, and cross-functional competency standards, but obtaining a KLP qualification serves as an important performance and career incentive.

BID PROTESTS ON DOD SOURCE SELECTIONS

A bid protest is a legal challenge to an award or solicitation for the procurement of goods or services. The GAO is a primary forum for bid protests concerning the DoD.[38] Once a protest is filed, the GAO has 100 days to issue a decision. A bid protest can have several different outcomes: it can be withdrawn by the protester, dismissed by the GAO, denied by the GAO, or sustained by the GAO.[39]

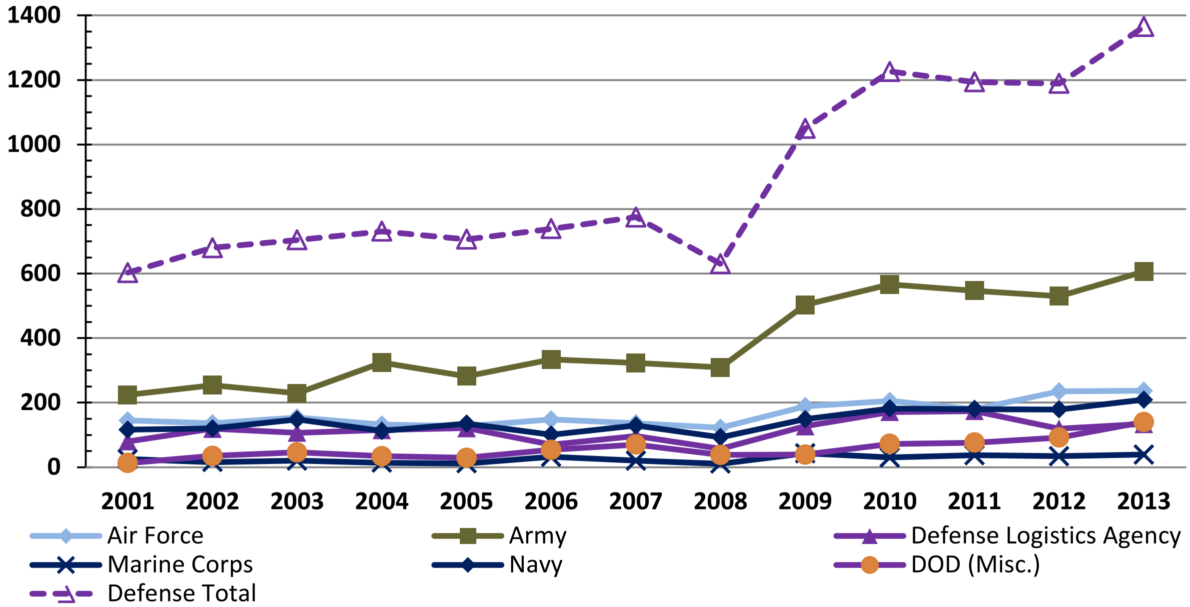
Figure 4-2 shows the total number of protests filed against the DoD from 2001–2013 (protest data were not yet available for FY 2014 for this report). The count ranged from 603 in 2001 to 1,365 in 2013. However, while there has been a sharp increase in the absolute number of protests since 2009 (especially for the Army), the sustainment rate has dropped recently and is below the federal total, which includes the DoD (see Figure 4-3). Even the individual sustainment rate for the Army (not shown) has been low and comparable to the other Components despite the rise in raw protest numbers. Thus, the increased number of protests appears to reflect external Industry strategies or competitive pressures from the declining DoD budgets rather than poor DoD source-selection performance. These results are commensurate with the Congressional Research Service's recent analysis of bid-protest rates (see Schwartz and Manuel, 2015). This result correlates with strong guidance to the program management and contracting communities over the last few years to ensure that source-selection rules are clearly defined, reflect the DoD's priorities, and are followed in execution. They provide additional guidance to ensure that source selections are well documented. Also, the BBP initiative to better define value in "best value" competitions may contribute to this result. Here we identify how much we value higher performance levels so that bidders have a clear

[38] Bid protests can also be handled by the procuring agency or the Court of Federal Claims. Thus, these data provide only a partial view into bid protest trends.

[39] Note that bid protests are an integral part of a government acquisition system that has a core value of treating all bidders fairly. (There are no protests in commercial business transactions.) They impart an incentive to ensure that solicitations and awards are conducted properly. The number of protests *sustained* provides an independent indicator of the quality of the DoD's source selections, independent of how many protests there are. See http://www.gao.gov/legal/bids/bidprotest.html as well as Schwartz and Manuel (2015) for more information on the GAO process and timeline.

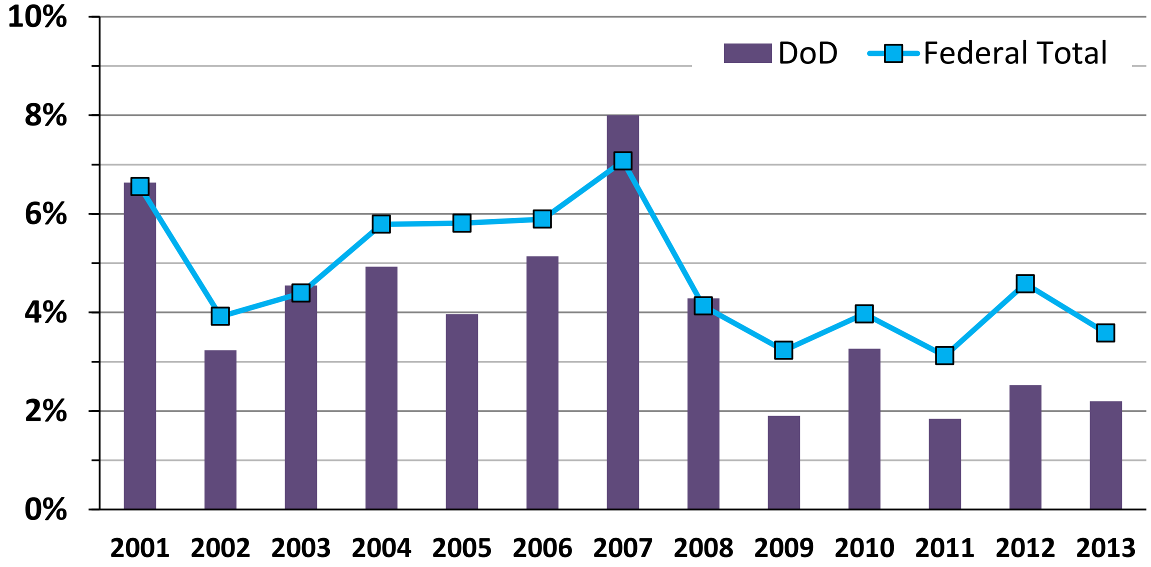
incentive to offer higher performance and the government has an objective way to evaluate differences.

Figure 4-2. Number of Protest Received by GAO (2001–2013)



SOURCE: GAO

Figure 4-3. GAO Protest Sustainment Rates (2001–2013)



Source: GAO

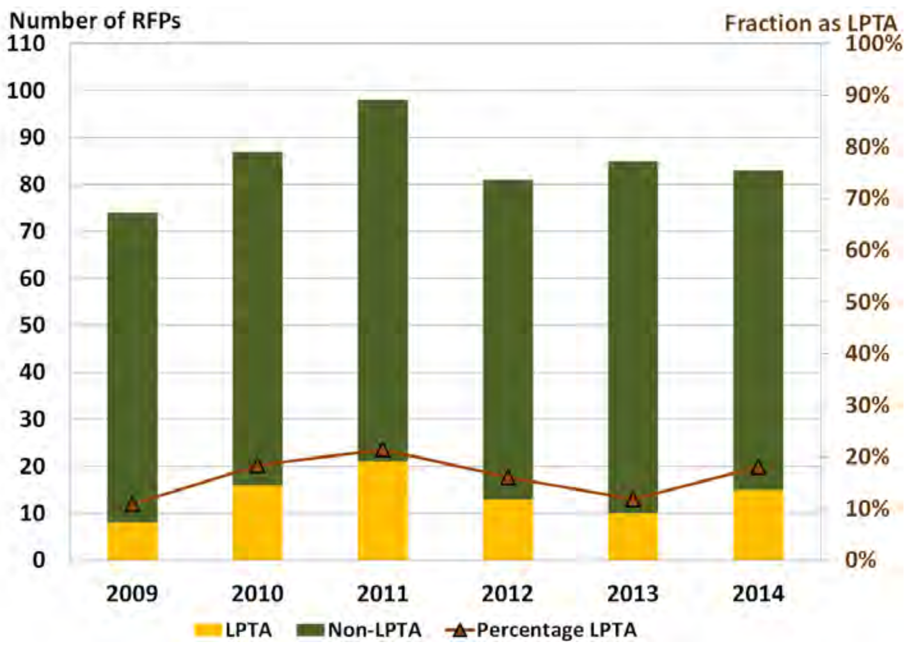
LOWEST-PRICED, TECHNICALLY ACCEPTABLE (LPTA) SOURCE SELECTIONS

The LPTA source-selection process is appropriate when best value is expected to result from selection of the technically acceptable proposal with the lowest evaluated price (FAR 15.101-2). As part of BBP, I have issued clear direction that LPTA is the appropriate source-selection process to apply only when there are well-defined requirements, the risk of unsuccessful contract performance is minimal, price is a significant factor in the source selection, and there is neither value, need, nor willingness to pay for higher performance (AT&L, 2015b).

To begin assessing the extent to which the DoD may be using LPTA inappropriately, we examined the frequency of LPTA in synopses of solicitation for four selected North American Industry Classification System (NAICS) codes for "Professional, Scientific, and Technical Services" (NAICS code 541);[40] these services generally involve capabilities for which value may need to be determined based on the proposals and capabilities offered and thus for which the use of LPTA as a source-selection method would be expected to be moderate.

Figure 4-4 shows the frequency of LPTA being cited in solicitations for these four NAICS codes. Here the annual percent of solicitations that mentioned LTPA in the solicitation ranged from 11 percent to 21 percent.

Figure 4-4. LPTA Frequency in Selected Services Solicitation Synopses (FY 2009–FY 2014)



NOTE: This analysis only examined the synopsis of these solicitations. There may be LPTA solicitations that only mention LPTA in file attachments of the solicitation.

[40] The four NAICS code examined were: 5413 (Architectural, engineering, and related services), 5414 (Specialized Design Services); 5416 (Management, Scientific, and Technical Consulting Services), and 5419 (Other professional, scientific, and technical services).

The GAO also studied the use of LPTA and found that uses of LTPA did not appear to be ill-advised (GAO, 2014b). While only a case study, this provides another data point on the use of LPTA.

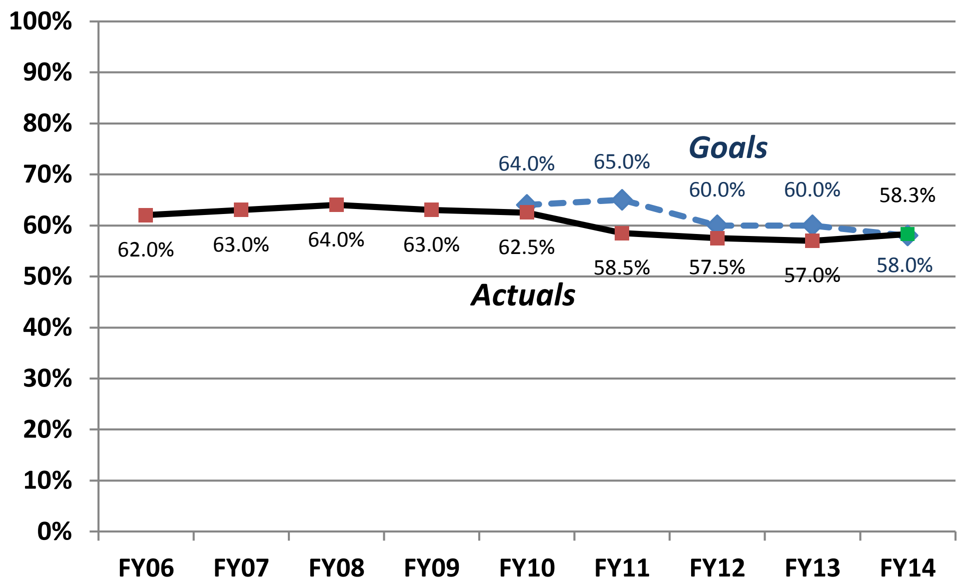
Nevertheless, in response to valid industry concerns about misuse of LPTA, my memorandum reaffirms the basic principles and guidance associated with LPTA solicitations (AT&L, 2015b).

RATES OF COMPETITIVE CONTRACTING ACROSS THE DOD FOR GOODS AND SERVICES

Competition—both head-to-head on contract competitions and environments that introduce competitive pressures in other ways—is a central tenet of our BBP initiatives. When viable, competition is, perhaps, the single best way to motivate contractors to provide the best value (i.e., the best performance at the lowest price). We have set a strategic objective to increase the percentage of spending on competed contracts from current levels.

Figure 4-5 plots the percentage of all DoD contract dollars competitively awarded from FY 2006 through FY 2014. Goals were established starting in FY 2010. Since goals were established in FY 2010, we had declining actuals until we made progress in FY 2014 at reversing the trend and meeting our goal. The budget climate has limited competitive "new starts." Starting in 2013, sequestration appears to correlate with the low point for competitive awards.

Figure 4-5. Competition Trends: Goals and Actuals (FY 2006–FY 2014)



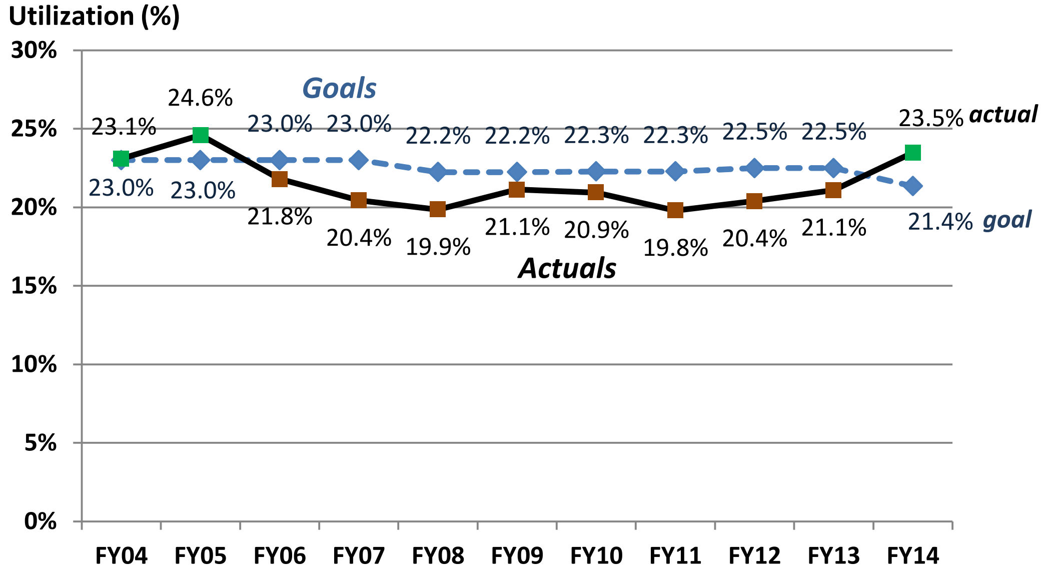
NOTE: We did not establish goals until FY 2010. Fraction of contracts competitively awarded is measured on a dollar basis.

SMALL-BUSINESS PARTICIPATION

We pursue small-business utilization goals as required by law, but more importantly because of the potential benefits from small-business contributions in both innovation and efficiency. More small-business engagement can increase the competitive nature of our solicitations, resulting in better cost and schedule performance on contracts. Small businesses also can infuse new, innovative technical solutions as capabilities are pursued for U.S. warfighters.

Figure 4-6 shows actual DoD-wide small-business utilization (measured by obligations) relative to yearly goals. Recent trends since FY 2011 have been steadily improving; we have exceeded our FY 2014 goal by 2.12 percentage points and surpassed prior years except FY 2005.

Figure 4-6. Small-Business Prime-Contracting Utilization Trends: Goals and Actuals (FY 2001–FY 2014)



PRODUCTIVITY AND EXECUTION FUNDING

There is much discussion and debate about the productivity of the defense acquisition system and whether government functions and associated staff and execution budgets can and should be trimmed. However, measuring output relative to execution costs is difficult because—unlike the commercial sector—the DoD does not have activity-based cost accounting. Also, there are no simple, accepted metrics (such as task-adjusted workforce productivity), and government management and oversight are subject to additional legal and regulatory constraints and processes (e.g., from the Buy American Act to bid protests to FAR contracting regulations established to deal with public concerns and priorities). Moreover, outsourcing of acquisition management in the 1990s (e.g., through concepts such as Total System Performance Responsibility [TSPR]) is recognized as having failed and cut too deeply into government acquisition workforce levels (e.g., recall Figure 3-1). Congress and the DoD have worked together to rebuild the acquisition workforce in recent years, so simply measuring recent trends in the relative levels of acquisition workforce or budgets to various producing activities or budgets may reflect these rebuilding efforts rather than problematic increases in execution costs or bureaucracy. We want to improve efficiency and minimize bureaucracy, but simply cutting without understanding the implications on performance may do more harm than good. The government's ability to manage technical risk in a program—and its ability to negotiate a fair and reasonable contract—depend directly on the size and composition of the government workforce. Removing this ability for short-term savings from reduced staff levels can be a very counterproductive decision.

To begin building quantitative insights into this difficult topic, we examined spending, productivity, and execution funding in the different ways shown below. None are perfect or entirely satisfactory, but together they begin to indicate that execution funding (including government "overhead") is not a major problem—at least not in the big picture. This should not deter our search for efficiencies, but it should make us cautious about doing more harm than good through false economies.

Share of Program RDT&E That Goes to Major MDAP Development Contracts

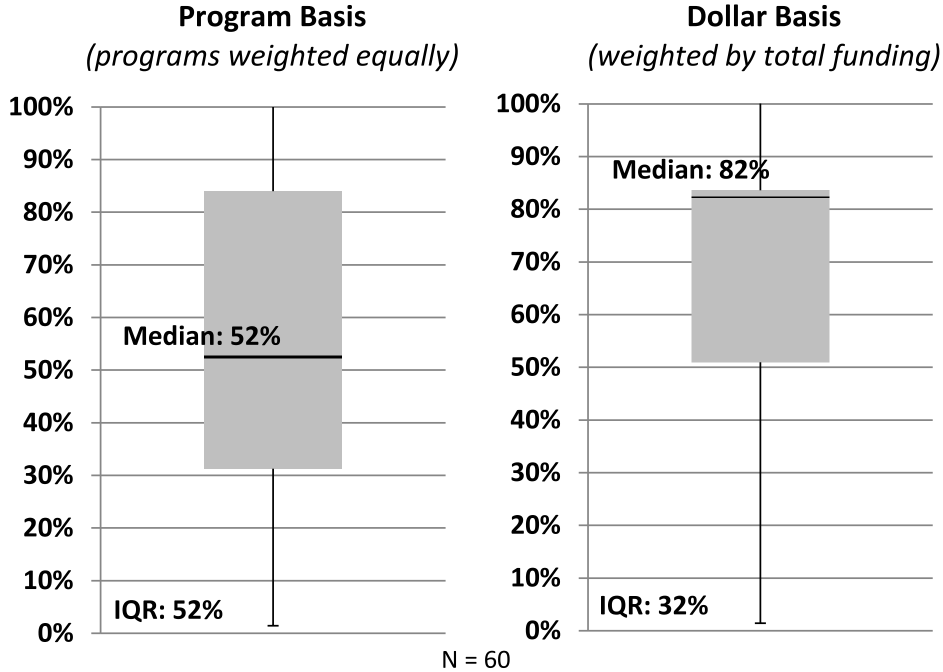
One measure of government process efficiency and execution funding is how much of program RDT&E is spent on development contracts rather than on government execution costs, which includes both true overhead and activities directly attributed to development (including program management and program-support contractors, RDT&E at government laboratories and warfare centers, and government testing) . We examine this in two different ways.

First, we examine the relative proportions of MDAP funding going out to prime contractors for RDT&E. In development, we have data on major contracts[41] and RDT&E funding for the program. These are the large contracts that are tracked with readily available EV data. Figure 4-7 summarizes the results of examining 60 active programs, looking at the major contractors' shares of program RDT&E funding on both a program and funding (dollar) basis. For half of the dollars funding these MDAPs, major contractors received more than 82 percent (see the right box-and-whisker chart in the figure). In other words, 82 percent of each RDT&E dollar "on average" (at the median) went to the major contractors for development in the program. This suggests that, on a dollar basis, government execution (both true overhead and internal RDT&E as measured by how much of the program's budget does not go to contractors on major development contracts) accounts for an average of 20 percent or less of program RDT&E funding. These figures do not count additional smaller RDT&E contracts for the program, which would increase the portion of each RDT&E dollar going to contractors for RDT&E.

[41] "Major MDAP contracts" are defined on p. 11.

This was lower on a program-by-program basis, where 52 percent of RDT&E dollars went to major contracts at the median (see the left box-and-whisker chart in the figure). This percentage probably is lower because smaller programs still have a fixed portion of management responsibilities regardless of program size. Smaller programs also may produce more RDT&E using government laboratories and warfare centers as a fraction of their total RDT&E funding. These data will be investigated further.

Figure 4-7. Minumum Program RDT&E Total Funding on Major Contracts (2000–2015)



Second, preliminary analysis corroborates that most of program RDT&E dollars go to contractors. When we examined RDT&E budget justification books for FY 2012–2015, the ratio of program management costs to total non-S&T RDT&E budget (including both large MDAPs and smaller programs) ranged from 25 percent to 30 percent.

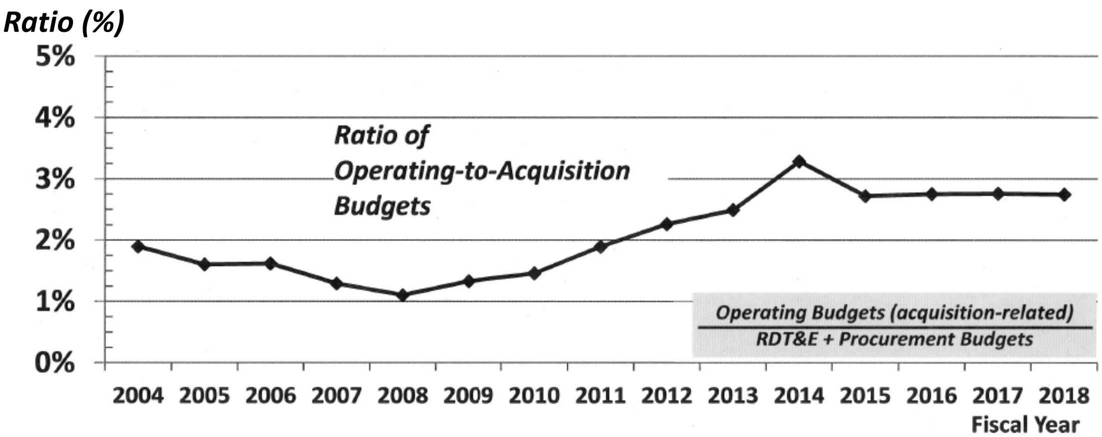
Comparing Acquisition-Related Operating Budgets to Acquisition Budgets

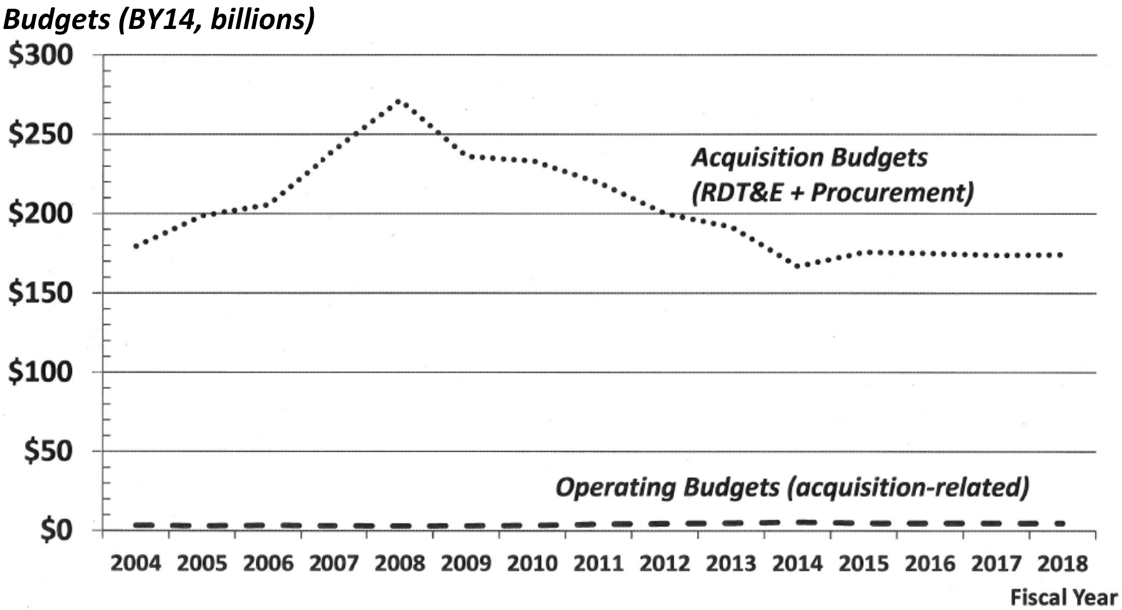
Lastly, to get a rough measure of operating budgets related to executing acquisition, we reviewed the approximately 5,000 operating-budget program elements (PEs) in the DoD budget, manually tagging those that are predominantly acquisition-oriented and that are focused on management or oversight. The budget for the Office of the USD(AT&L), for example, is one (albeit with some nonacquisition functions, such as installations and environmental oversight), as are those for DOT&E, OSD's Cost Assessment and Program Evaluation (CAPE), etc. We then compared the total of these operating budget PEs to the acquisition budgets themselves (RDT&E plus Procurement).

The results are shown in Figure 4-8. Acquisition budgets (the upper line on the lower chart) dwarf the acquisition-related operating budgets (the lower line on the lower chart). The ratio of operating budgets to acquisition budgets from FY 2004 to FY2018 in PB14 ranged from about 1.5 percent to 3.5 percent (see the upper chart). The ratio changed over the years, predominantly influenced by the large swings in acquisition budgets since 2004, although there has been a small increase in the acquisition-related operating budgets (dominated by increases in the acquisition workforce development fund but with smaller investments elsewhere to rebuild our acquisition capabilities).

While this is only a gross macro measure at the fidelity of organizational budget accounts and used an earlier PB14, the past and estimated future operating budgets are much lower than the acquisition budgets for those periods. More analysis could be done along these lines with recent data to see how these trends may have changed more recently.

Figure 4-8. Ratio of Organization Budgets to Acquisition Budgets (FY 2004–2018)





NOTES: This is earlier preliminary analysis based on PB 2014.

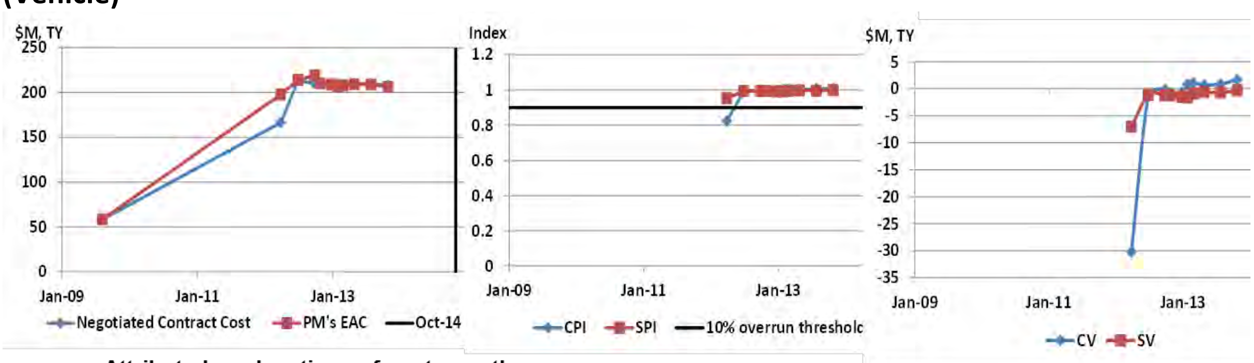
(This page was left blank intentionally)

A. Analysis of Individual Cost-Growth Outliers on MDAP Development Contracts

Figure A-1 through Figure A-27 summarize the analysis of each of the 27 MDAP contract cost-growth outlier identified earlier on p. 44. Analysis is largely based on SARs and the evidence in the EV reports.

Army Outliers

Figure A-1. Cost Growth: Paladin Integrated Management (PIM) Development Contract, Army (Vehicle)

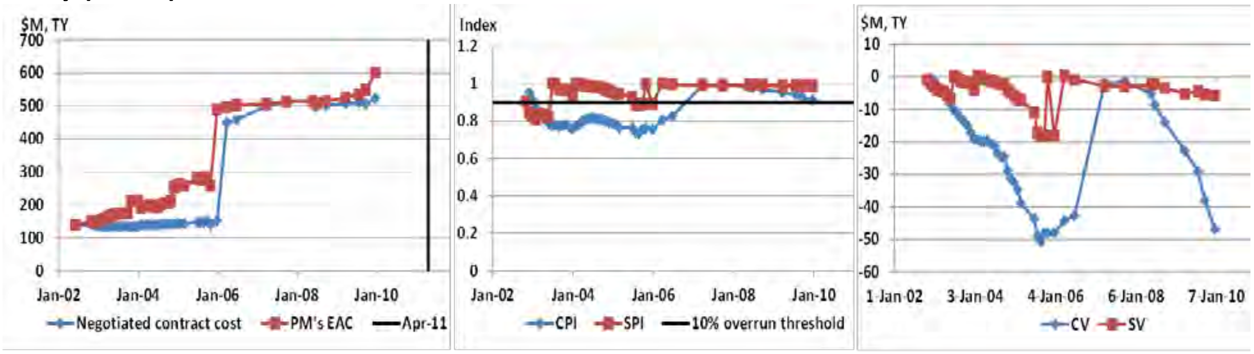


Attributed explanations of cost growth

1. Started as ACAT II program.
2. Systems engineering problems, unstable engineering or system requirements for force protection/survivability.
3. Greater than planned effort to achieve prototype design suggesting overrun rolled into contract baseline (contract management problems and underestimating cost).
4. Program restructure after being part of NLOS-C program
5. Contract restructured into a following comprehensive contract modification (CCM) CPIF contract.

- Total cost growth: 253%
- Work-content growth: 256%
- Cost-over-target: -3%
- 98% complete

Figure A-2. Cost Growth: Joint Tactical Networks (JTN) (JTRS Cluster 1) Development Contract, Army (Radio)

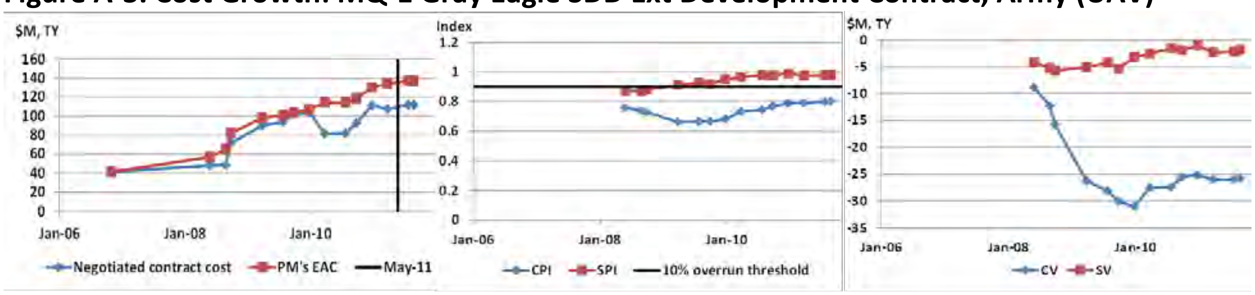


Attributed explanations of cost growth

1. Extensive rework in design of Wideband Networking Waveform.
2. OTB January 2004.
3. Restructured plan. Becomes GMR October 2005.
4. Systems engineering problems in integration, test delays, untimely software availability.
5. Unstable engineering or system requirements.

- Total cost growth: 329%
- Work-content growth: 274%
- Cost-over-target: 55%
- 95% complete

Figure A-3. Cost Growth: MQ-1 Gray Eagle SDD Ext Development Contract, Army (UAV)

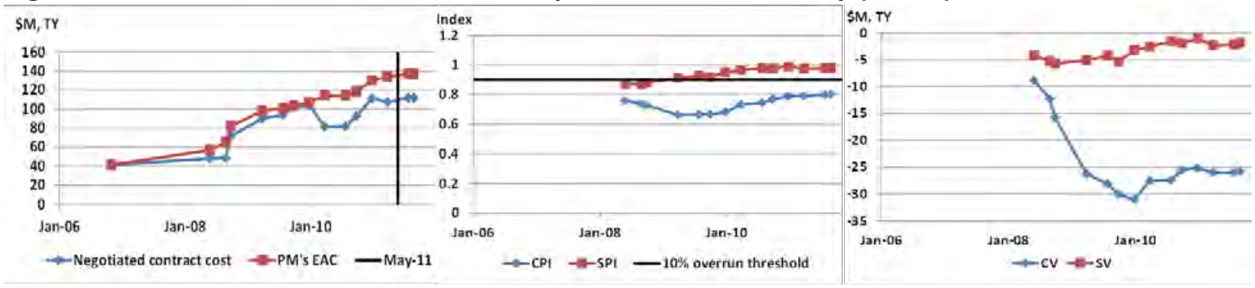


Attributed explanations of cost growth

1. Underestimated complexity of work and resolution of software discrepancy reports.
2. Complexity of development concurrently with other related developments and SE difficulties.
3. Modifications add work scope and engineering or system requirements increased.
4. Required additional network management capability, additional porting, CEA mods, and numerous ECPs.
5. Added scope for updating and testing waveform.
6. Equitable adjustment for GFE in March 2010.

- Total cost growth: 231%
- Work-content growth: 170%
- Cost-over-target: 61%
- 94% complete

Figure A-4. Cost Growth: JTN SRW Development Contract, Army (Radio)

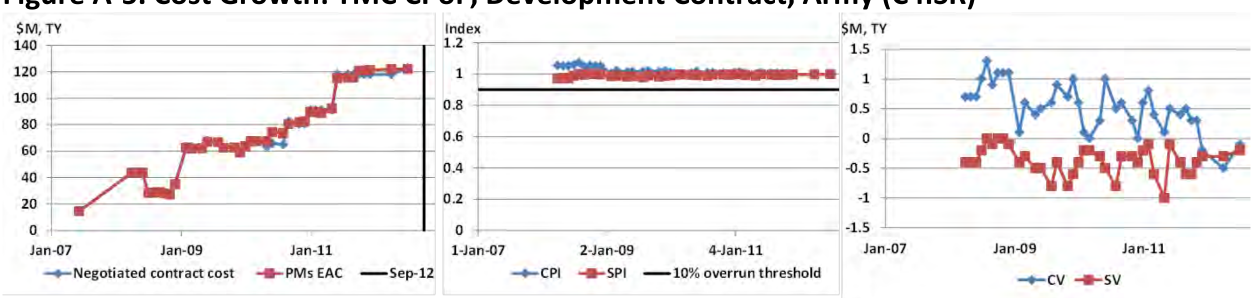


Attributed explanations of cost growth

1. Underestimated complexity of work and resolution of software discrepancy reports.
2. Complexity of development concurrently with other related developments and SE difficulties.
3. Modifications add work scope and engineering or system requirements increased.
4. Required additional network management capability, additional porting, CEA mods, and numerous ECPs.
5. Added scope for updating and testing waveform.
6. Equitable adjustment for GFE in March 2010.

- Total cost growth: 231%
- Work-content growth: 170%
- Cost-over-target: 61%
- 94% complete

Figure A-5. Cost Growth: TMC CPoF, Development Contract, Army (C4ISR)

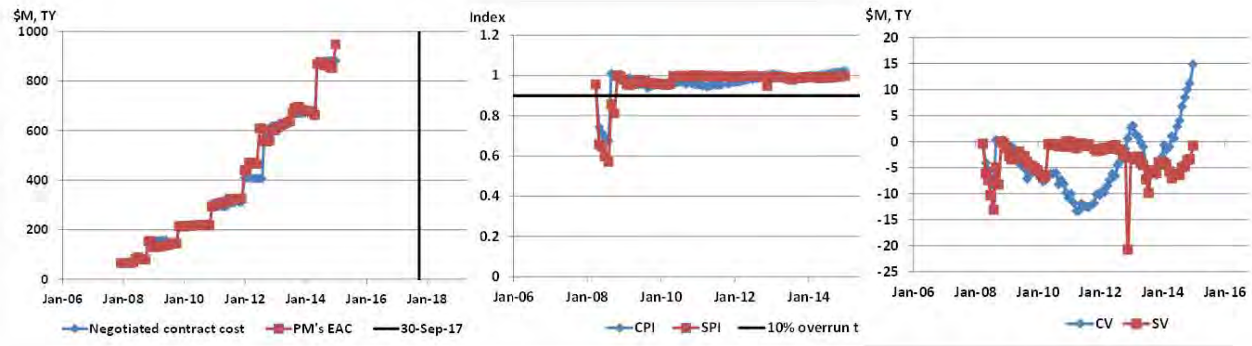


Attributed explanations of cost growth

1. This contract started after TMC was downgraded to an ACAT II program.
2. Concept began as a DARPA technology demonstration.
3. Acquisition strategy was spiral-like employing a test-fix-test process.
4. Unstable engineering or system requirements.
5. Project was developed and acquired concurrently with being used operationally.
6. Example of a type of “skunk works” project.

- Total cost growth: 736%
- Work-content growth: 742%
- Cost-over-target: -6%
- 95% complete

Figure A-6. Cost Growth: Global Combat Support System-Army, EMD Development Contract, Army (C4ISR)



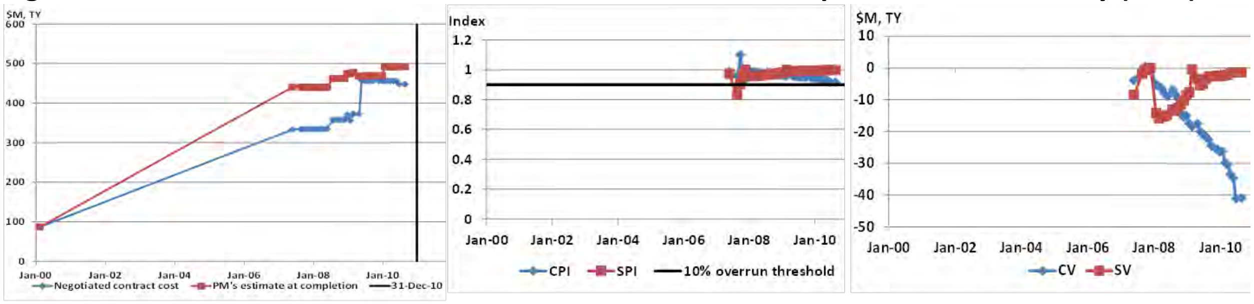
Attributed explanations of cost growth

1. Had a spiral-like acquisition strategy developing in increments or segments or versions.
2. Unstable engineering or system requirements.
3. Bad start in that APB was approved almost a year after work started on the contract.
4. There were serial schedule slips.
5. Expected large overrun indicates underestimated cost.
6. Some evidence of rolling overruns into the contract baseline suggest some contract management problems.

- Total cost growth: 1346%
- Work-content growth: 1245%
- Cost-over-target: 101%
- 95% complete

NAVY OUTLIERS

Figure A-7. Cost Growth: VTUAC MQ-8 Fire Scout, SDD Development Contract, Navy (UAV)

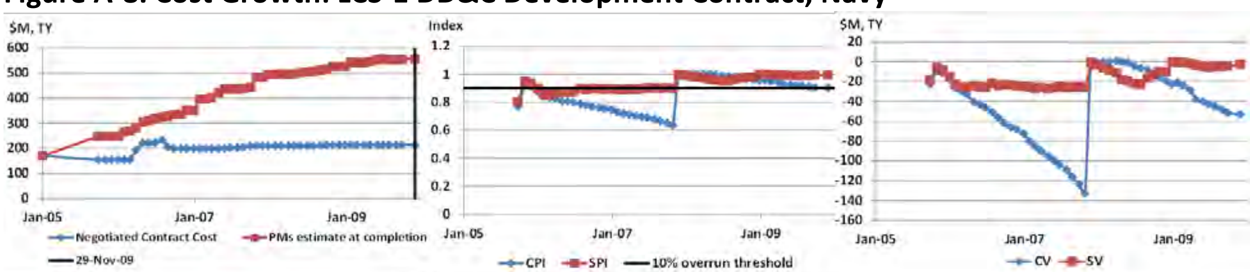


Attributed explanations of cost growth

1. Awarded as UCA in February 2000. Definitized 23 June 2006.
2. Program started as ACAT II "off-the shelf" minimal development effort.
3. There were a series of contract modifications due to performance scope growth to meet Navy engineering or system requirements.
4. There was an OTB in December 2002.
5. There was a program restructure to support change of OPEVAL platform from LCS to Frigate (FFG-8) in February 2008.

- Total cost growth: 466%
- Work-content growth: 416%
- Cost-over-target: 50%
- 98% complete

Figure A-8. Cost Growth: LCS-1 DD&C Development Contract, Navy

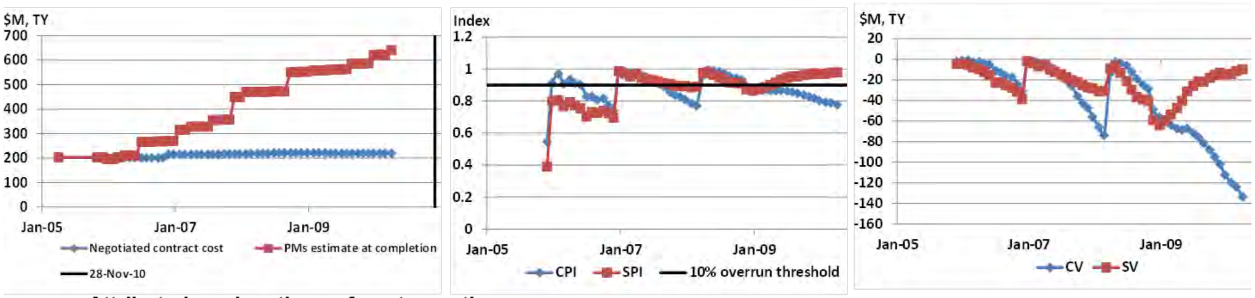


Attributed explanations of cost growth

1. There was unplanned construction to receive mission modules
2. There was additional steel and weld training due to Navy Vessel Rules
3. OTB in October 2005
4. In 2006 design changes resequenced work due to late delivery of main propulsion reduction gears and there were overruns associated with rework
5. Another OTB requested November 2007
6. In 2009, there were increased costs due to accelerated trials and corrections of deficiencies

- Total cost growth: 228%
- Work-content growth: 26%
- Cost-over-target: 202%
- 99% complete

Figure A-9. Cost Growth: LCS-2 DD&C Development Contract, Navy

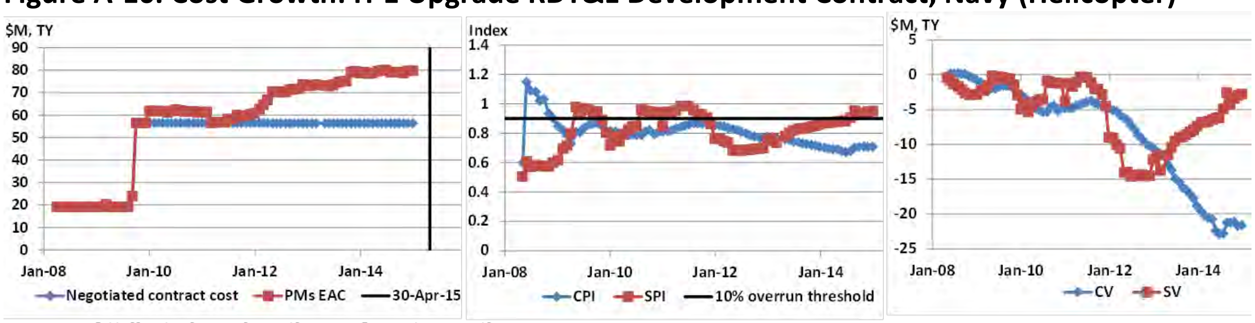


Attributed explanations of cost growth

1. Vendors late with deliveries due to Hurricane Katrina during early phase of development.
2. Design changes due to adjustments to Naval Vessel Rules.
3. December 2006: costs due to additional required systems engineering resources.
4. OTB February 2007.
5. 2007: unplanned increases in aluminum fabrication labor, increased outsourcing, rework, and relocation of work scope from Austal to BIW.
6. OTB February 2008.
7. 2009: unplanned rework, underestimated work associated with propulsion, and costs associated with correction of deficiencies at builder's trials.

- Total cost growth: 216%
- Work-content growth: 8%
- Cost-over-target: 208%
- 98% complete

Figure A-10. Cost Growth: H-1 Upgrade RDT&E Development Contract, Navy (Helicopter)

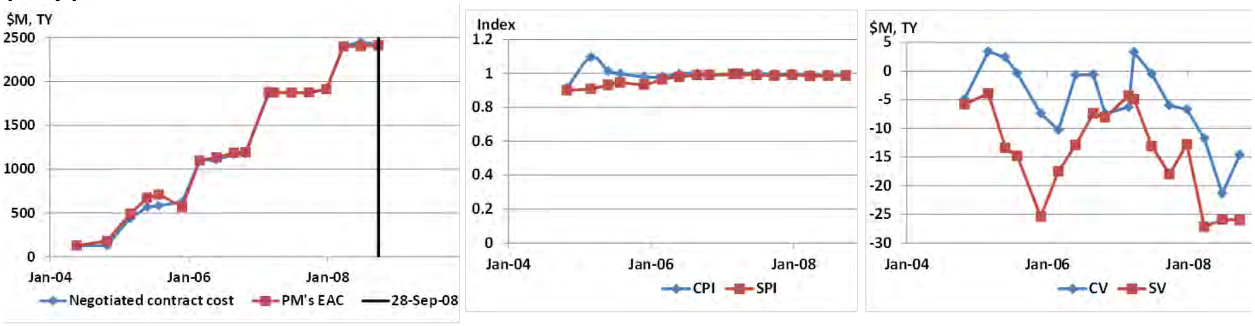


Attributed explanations of cost growth

1. September 2009: add modification to contract for phase II.
2. A spiral-like development contract.
3. 2010: model drawing conversion leads to overruns.
4. 2011: subcontractor overruns cause cost increases.
5. 2012: Manufacturing startup issues at sub contractor and increased prime contractor oversight.
6. 2013: Added resources and personnel for fabrication at subcontractor, more engineering and quality assurance support. More tooling costs.
7. 2014: increased costs for inspection completion.

- Total cost growth: 317%
- Work-content growth: 194%
- Cost-over-target: 123%
- 95% complete

Figure A-11. Cost Growth: CVN 21, Construction Preparation Development Contract, Navy (Ship)

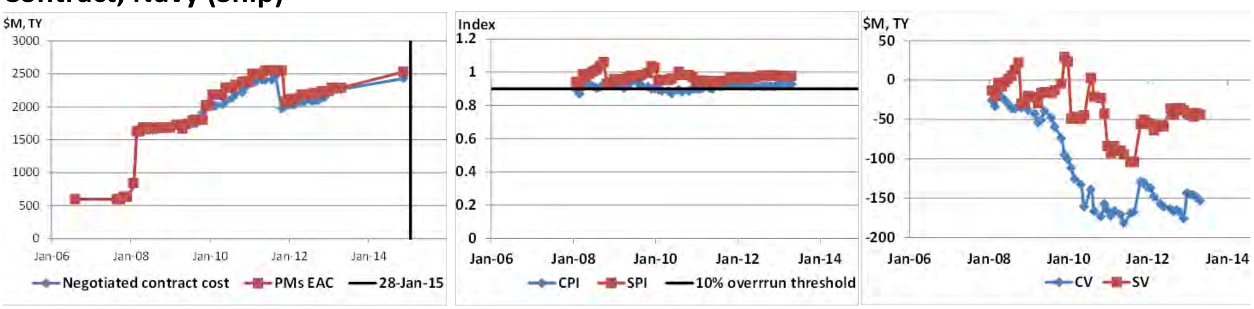


Attributed explanations of cost growth

1. 2004: Delays in engineering and design products. Inefficiencies associated with incremental approval of ship's arrangements and the maturity of ship's design
2. EVMS problems addressed in 2005
3. 2006: Rebaseline IMS and new period-of-performance to reflect a new lead-ship delivery date. Contract mods increased work scope
4. 2008: contract modifications increase work scope and extend period-of-performance due to new lead-ship delivery date and AT&L-direction

- Total cost growth: 1805%
- Work-content growth: 1818%
- Cost-over-target: -13%
- 83% complete

Figure A-12. Cost Growth: DDG 1000 (Bath Iron Works), Phase IV DD&C Development Contract, Navy (Ship)

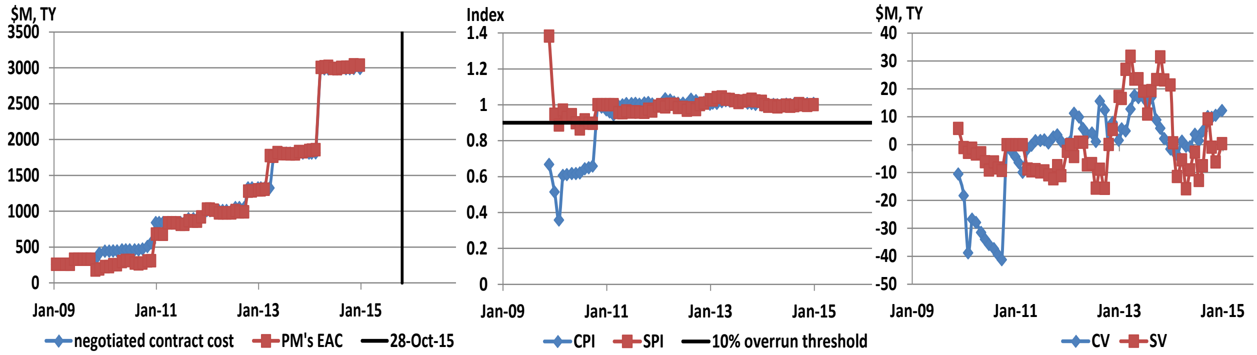


Attributed explanations of cost growth

1. Work started August 2006. IBR not until a year later, August 2007.
2. 2008: increased costs and process issues with CATIA (computer-aided 3-dimensional interactive application), late functional design, numerous engineering change orders.
3. 2009: Began transfer of some work from Ingalls Shipbuilding to BIW.
4. 2010: added electrical work, and AGS gun mount and magazine fabrication
5. 2011: PM added controls over design changes with on-site NERB (Naval Engineering Review Board) to reduce ECOs

- Total cost growth: 283%
- Work-content growth: 280%
- Cost-over-target: 3%
- 95% complete

Figure A-13. Cost Growth: CVN 79 Construction Preparation Development Contract, Navy (Ship)

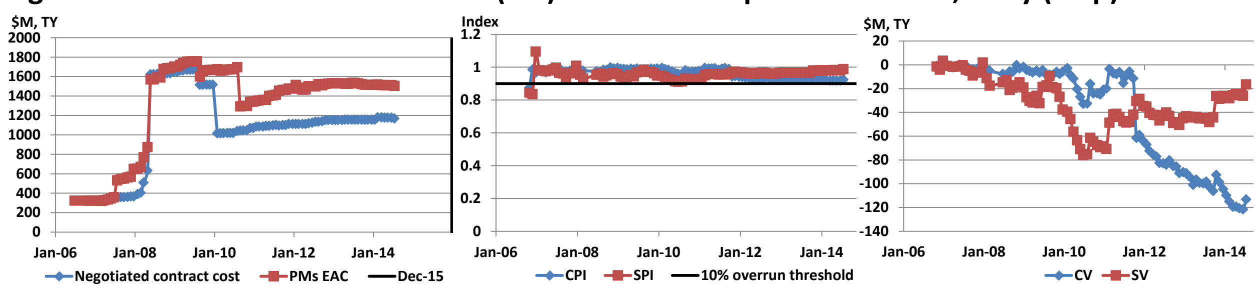


Attributed explanations of cost growth

1. December 2010 award contact extension that adds work scope.
2. December 2011 executes options that add scope.
3. January 2013 contract modifications add work scope.
4. March 2013, contract modifications add scope.
5. June 2014, contract modifications add scope.

- Total cost growth: 1078%
- Work-content growth: 1059%
- Cost-over-target: 19%
- 95% complete

Figure A-14. Cost Growth: DDG 1000 (NG) DD&C Development Contract, Navy (Ship)

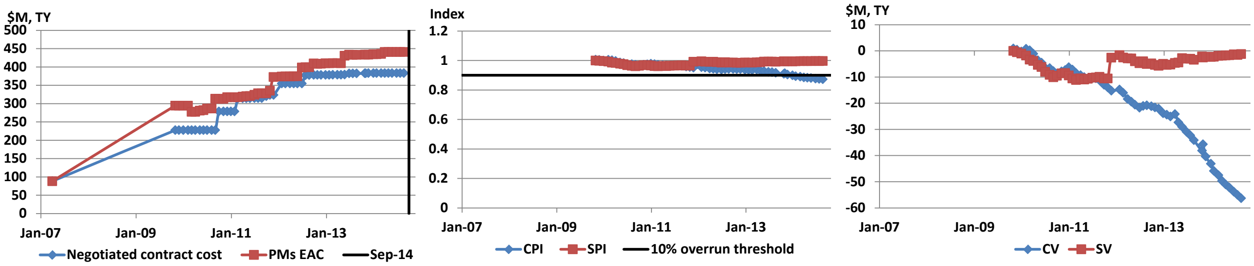


Attributed explanations of cost growth

1. 2008: Systems Engineering and design issues. Numerous ECPs.
2. 2009: Navy executing termination modification to deobligate Class Common Control equipment and add to the BIW contract. Overruns due to systems integration costs and detail design deficiencies.
3. 2012: overruns higher than expected. Rework rates for the composite deckhouse and hangar improvement through successive composite panel assemblies.

- Total cost growth 364%
- Work-content growth: 261%
- Cost-over-target: 103%
- 99% complete

Figure A-15. Cost Growth: G/ATOR EMD Development Contract, Navy

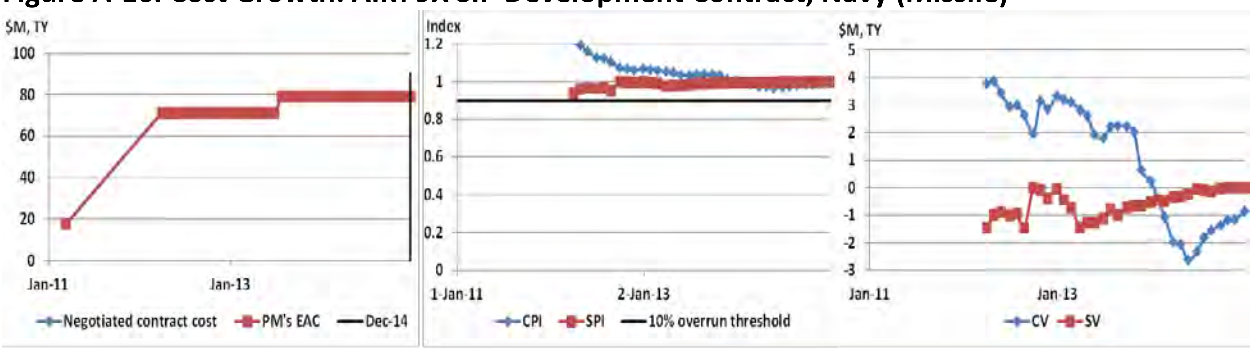


Attributed explanations of cost growth

1. Work started March 2007 as an ACAT II program. Designated ACAT I program in October 2011. APB approved 22 May 2012. This is a restructured program.
2. March 2012, contract modifications added to mitigate risk and support the replanning process.
3. June 2012: Cost grows due to unplanned additional integration tests and added manpower for software releases. Also added more work scope to mitigate risk.

- Total cost growth 400%
- Work-content growth: 334%
- Cost-over-target: 66%
- 99% complete

Figure A-16. Cost Growth: AIM 9X SIP Development Contract, Navy (Missile)



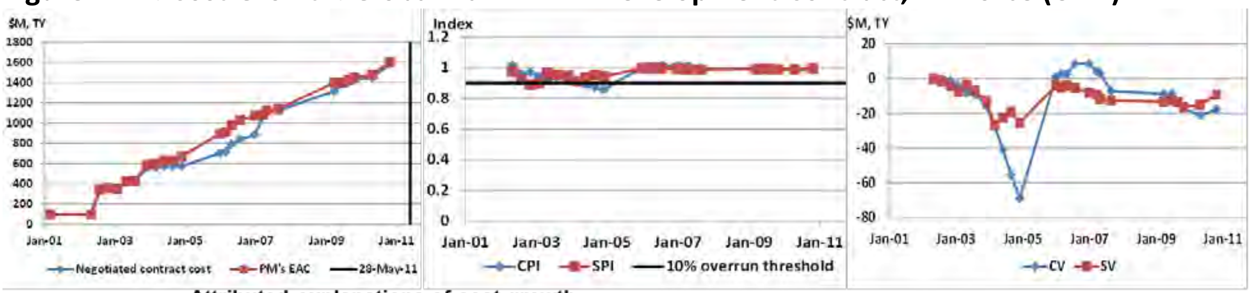
Attributed explanations of cost growth

1. Added work on Active Optical Target Detector (AOTD).
2. Contract management issue delayed submission of WBS for approval.

- Total cost growth 365%
- Work-content growth: 365%
- Cost-over-target: 0%
- 100% complete

Air Force Outliers

Figure A-17. Cost Growth: Global Hawk EMD Development Contract, Air Force (UAV)

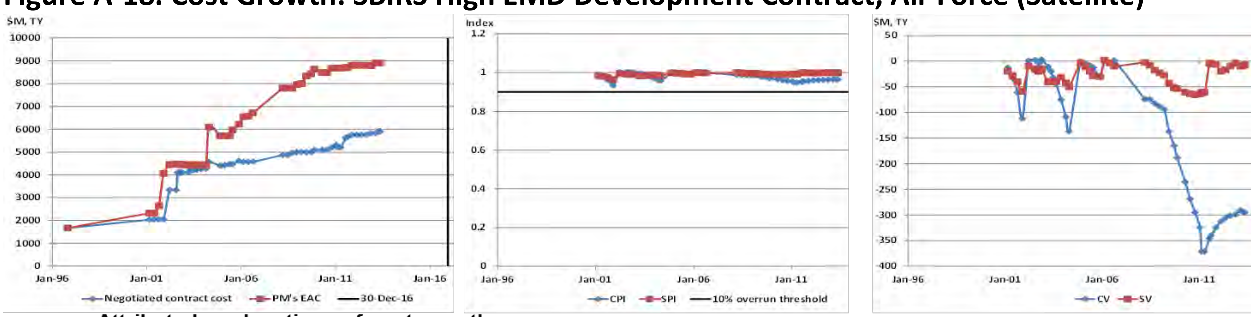


Attributed explanations of cost growth

1. This was a spiral development.
2. Transitioned from ACTD to EMD in 2001.
3. Concurrent development, production, and operations during period of performance.
4. There was a serial use of UCAs.
5. 2004: there has been an increased payload variant and problems associated with late software and sensor deliveries.
6. There was a critical Nunn-McCurdy breach November 2005.
7. Program is replanned with new APB in March 2007 with new milestones and requirements.
8. Second critical Nunn-McCurdy breach in 2011. There has been fluctuation in quantity and mix of variants. There have also been ground station and communications re-architecture issues.
9. FY15 NDAA directs a new high altitude ISR study.
10. There is some evidence of rolling overrun into contract baseline.

- Total cost growth 1599%
- Work-content growth: 1575%
- Cost-over-target: 24%
- 91% complete

Figure A-18. Cost Growth: SBIRS High EMD Development Contract, Air Force (Satellite)

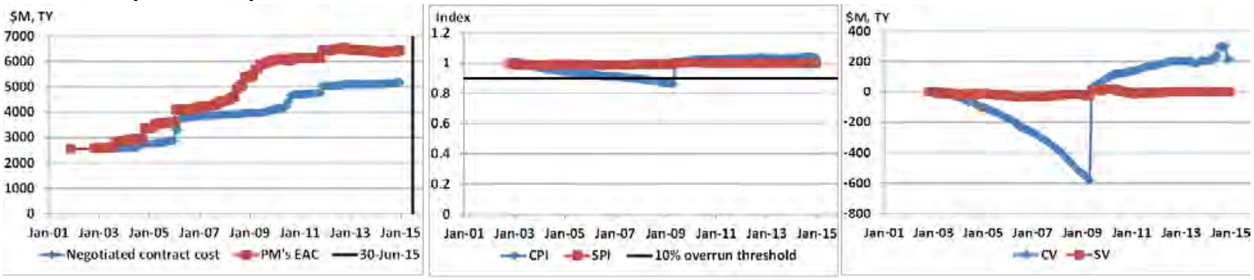


Attributed explanations of cost growth

1. 1999: APB breach triggers replan and redesign.
2. OTB in July 2000.
3. 2001: issues with software for Pointing and Control Assembly and faulty cost estimates.
4. November 2001: replan, redesign, and reorganize.
5. 2002: IBR for replan.
6. 2004: Cost growth due to problems with integration, assembly, test and checkout (AIT&C) and geosynchronous orbit payload problems.
7. April 2004: a replan OTB.

- Total cost growth 434%
- Work-content growth: 254%
- Cost-over-target: 180%
- 97% complete

Figure A-19. Cost Growth: Advanced Extremely High Frequency EMD Development Contract, Air Force (Satellite)

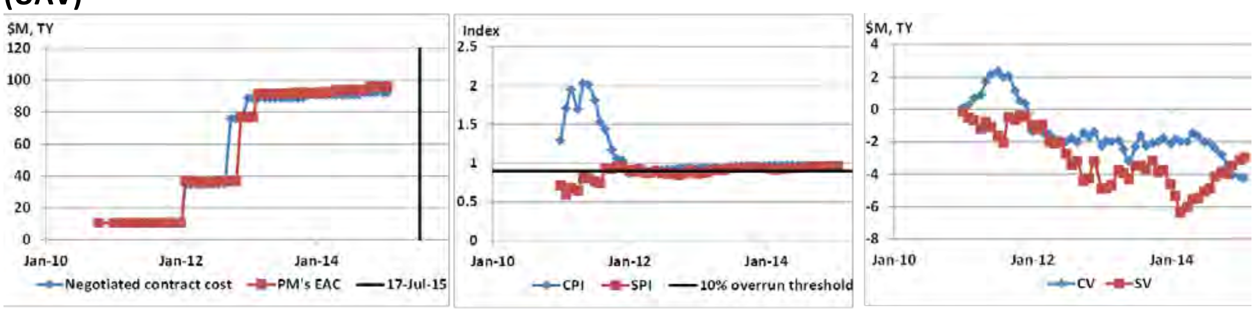


Attributed explanations of cost growth

1. Work started November 2001.
2. There were numerous ECPs, replans, and overruns. 2004: ASIC requirements are verified. There were digital processor subsystem (DPS) design issues, other design issues, and material cost growth.
3. 2006: SV has assembly, integration and test (AI&T) issues. There were unplanned ground support equipment tasks, increased complexity, and additional redesign and testing.
4. 2009: OTB to rebaseline that includes replanning remaining work.

- Total cost growth 153%
- Work-content growth: 104%
- Cost-over-target: 49%
- 99% complete

Figure A-20. Cost Growth: MQ-9 Targeting Improvements Development Contract, Air Force (UAV)

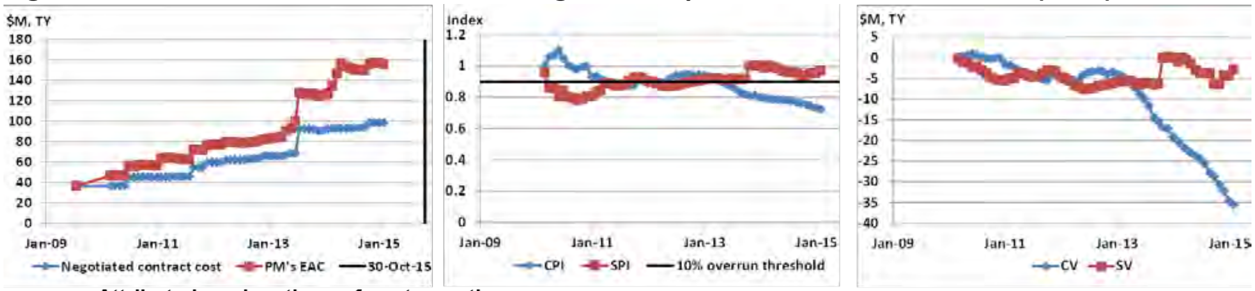


Attributed explanations of cost growth

1. Work started October 2010.
2. February 2012: APB signed.
3. Cost growth within Tri-Beam Emission and Receiver (TBEAR) laser development.
4. Systems engineering problems with integration that affects design and production.
5. Delays in development of TBEAR.
6. Unstable engineering or system requirements.
7. Some evidence of EV and contract management issues (rolling overrun into baseline).

- Total cost growth 797%
- Work-content growth: 760%
- Cost-over-target: 37%
- 93% complete

Figure A-21. Cost Growth: MQ-9 SDD Bridge Development Contract, Air Force (UAV)

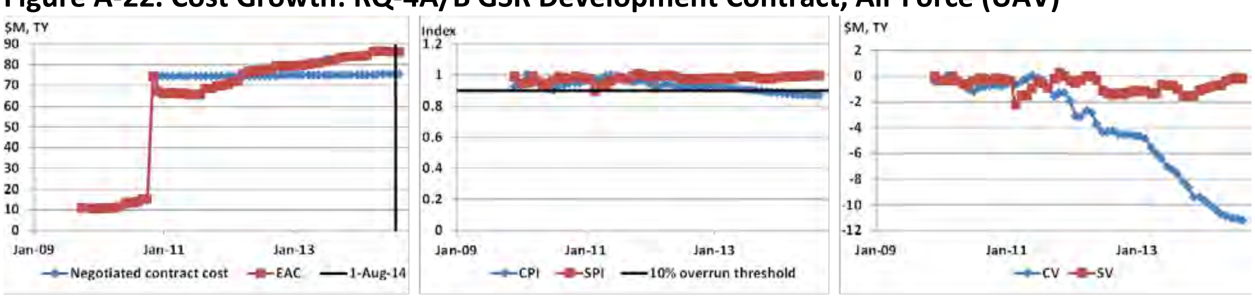


Attributed explanations of cost growth

1. Work started July 2009.
2. Cost growth due to overruns, rebaselining, and contract modifications.
3. 2013: unplanned and unscheduled on-wing certification events, test eval support after software changes identified in previous tests.
4. 2011: overrun modification.
5. February 2014: overrun modification.
6. October 2014: overrun modification.
7. October 2011: multiple design changes to forward bay and heavyweight landing gear re-tests.
8. 2013: accelerated retrofit, realign environmental testing with prime, underforecasts of activities needed to complete.

- Total cost growth 327%
- Work-content growth: 169%
- Cost-over-target: 158%
- 87% complete

Figure A-22. Cost Growth: RQ-4A/B GSR Development Contract, Air Force (UAV)

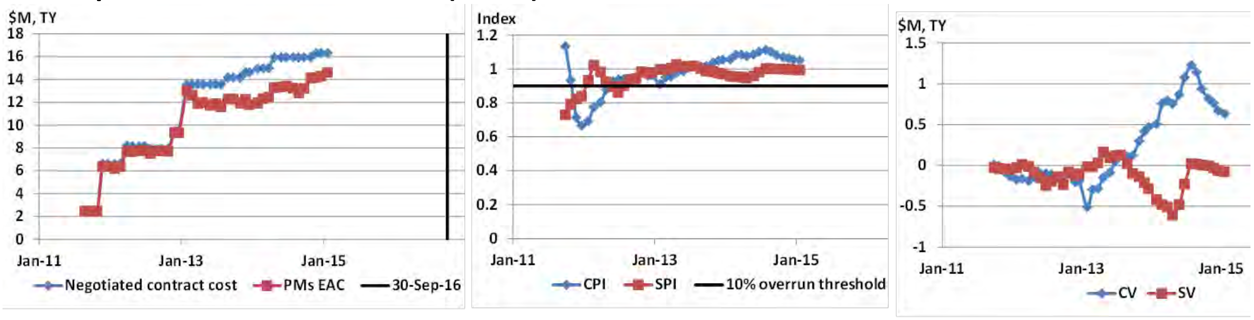


Attributed explanations of cost growth

1. Work started September 2009.
2. A spiral-like development. Phase 0 for study phase through SRR (Systems Requirements Review) then Phase IA for development, program management, engineering, and testing.
3. 2011: underestimated management costs due to additional needs for reviews.
4. 2012: add broadband hardware and software licenses.
5. 2013: Addition of GSRA Phase IA supplier software licenses. Increased rates as result of Northrop Grumman corporate structuring. Overruns due to underestimation of systems engineering costs.

- Total cost growth 682%
- Work-content growth: 586%
- Cost-over-target: 96%
- 99% complete

Figure A-23. Cost Growth: Mission Planning System Increment 4, MEPEC II for F-22 Development Contract, Air Force (C4ISR)

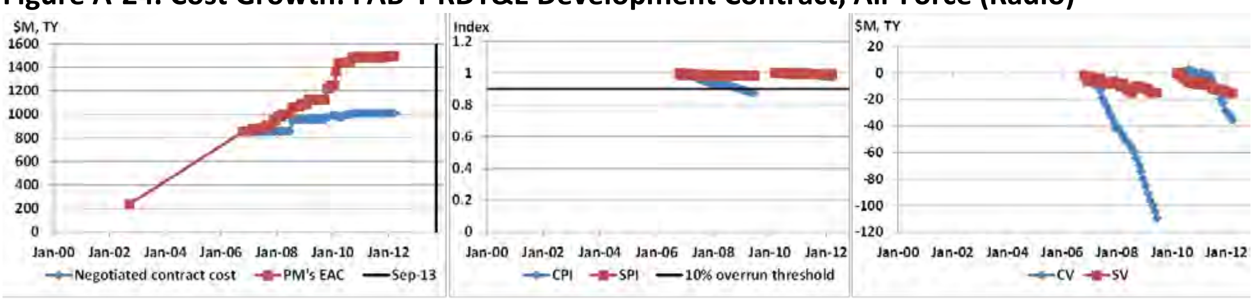


Attributed explanations of cost growth

1. Work started August 2011.
2. A spiral-like development with version releases.
3. Unstable engineering or system requirements.
4. 2011: New APB and critical schedule change.
5. November 2011: contract modification for more work.
6. 2013: contract modification for more work.

- Total cost growth 493%
- Work-content growth: 563%
- Cost-over-target: -70%
- 90% complete

Figure A-24. Cost Growth: FAB-T RDT&E Development Contract, Air Force (Radio)



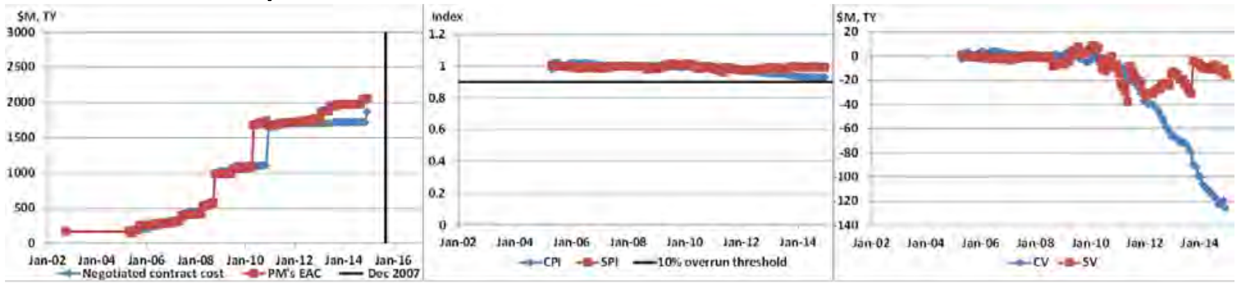
Attributed explanations of cost growth

1. 2002-2006: added scope and there was a replan. No EV reporting for 4 years.
2. 2007: Design qualification issues that had to be reworked.
3. 2009: Higher costs than expected to meet Engineering Development Model (EDM) schedule. There were hardware qualifications problems that required rework and there were integration problems.
4. 2010: There was an OTB.
5. 2011: More integration issues.
6. 2012: Converts to FFP contract.

- Total cost growth 533%
- Work-content growth: 327%
- Cost-over-target: 206%
- 92% complete

DoD Outliers

Figure A-25. Cost Growth: Chem-Demil (Assembled Chemical Weapons Alternatives (ACWA)), Pueblo CAPP Development Contract, DoD

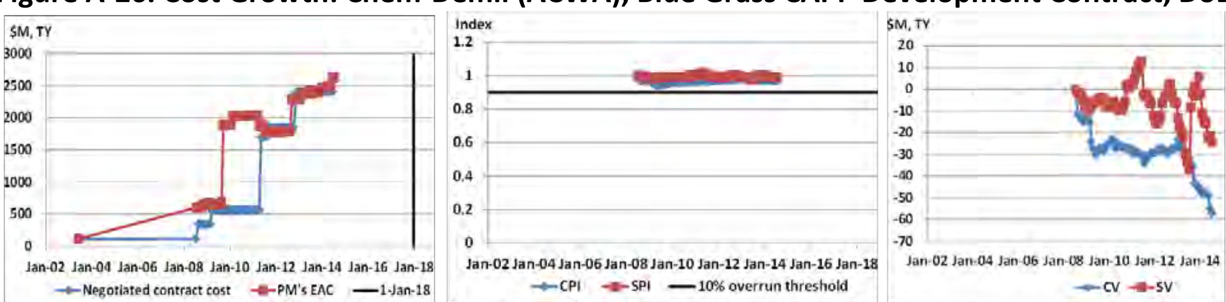


Attributed explanations of cost growth

1. Work started September 2002
2. A spiral-like development with a multi-phase task order contract
3. Unstable engineering or system requirements
4. 2004: SWO due to concerns about contract structure
5. 2005: Redesign definitized
6. 2007: Additional engineering for risk reduction and redesign. Expect award of Construction Stage 3 in FY2008
7. 2009: cost for systemization proposal work is higher than planned. Systemization Part 1 awarded
8. 2010: Systemization Part 2 awarded
9. There is some evidence of rolling overruns into contract baselines
10. 2013: added engineering costs to support construction and systemization
11. 2014: Discovery of plant deficiencies that must be fixed to complete systemization
12. Critical Nunn-McCurdy breaches in August 2006 and December 2010

- Total cost growth 1116%
- Work-content growth: 1004%
- Cost-over-target: 112%
- 89% complete

Figure A-26. Cost Growth: Chem-Demil (ACWA), Blue Grass CAPP Development Contract, DoD



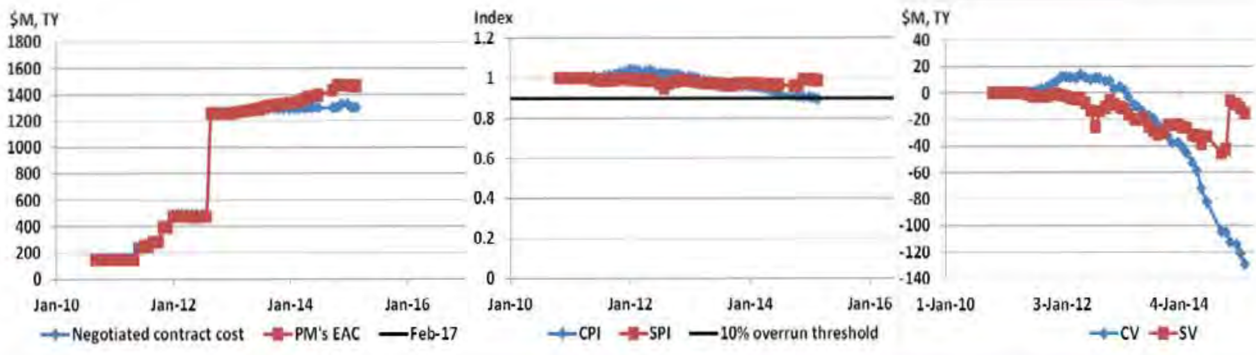
Attributed explanations of cost growth

1. Work started June 2003.
2. A spiral-like development with a multi-phase task order contract.
3. Unstable engineering or system requirements.
4. 2004: evolution and complexity of design required rework. Technical tests were not meeting interim milestones. SWO in September 2004.
5. 2006: Replanning.
6. New APB April 2007.
7. 2009: Design rework overruns and part 1 for "First of a kind" work (FOAK).
8. Construction Phase IV definitized 31 March 2011.
9. 2013: added engineering for construction optimization, added engineering overtime to speed up turnarounds for construction changes. There is an increase in business, safety and systemization support of construction.
10. 2014: deficiencies in construction must be fixed to complete contract. Evidence of contract management issues (EV).
11. Critical Nunn-McCurdy breaches in August 2006 and December 2010.

- Total cost growth 2197%
- Work-content growth: 2165%
- Cost-over-target: 32%
- 70% complete

MDA OUTLIERS

Figure A-27. Cost Growth: Aegis SM-3 Block IIA Development Contract, MDA (C4ISR)



Attributed explanations of cost growth

1. Significant work added to contract. An especially large amount of work added in August 2012.
2. Also underestimated costs of given work leading to cost overruns above target costs.

- Total cost growth 918%
- Work-content growth: 803%
- Cost-over-target: 115%
- 83% complete

(This page was left blank intentionally)

B. Statistical Analysis Details

Generally, statistical analyses conducted for this report involved both parametric and nonparametric tests, as discussed below. We also employed factor analysis to develop new inferred measures of complexity and risk on MDAP contracts.

Supporting Sample Analysis for Regressions

In our linear multivariate regression analyses, we conducted supporting sample analysis tests for normality of residuals (Smirnov-Kolmogorov and Shapiro-Wilk tests), heteroskedasticity (Cook-Weisberg test), multicollinearity (variance inflation factor test), omitted variables (Ramsey Regression Equation Specification Error Test [RESET]), and correct model specification (Linktest). We also used bootstrap simulations to obtain unbiased coefficient estimates, correct standard errors, and correct confidence intervals.

Single Variable Analysis Tests

Single variable analyses allowed us to focus on differences by a single factor (e.g., phase, contract type, cost or price growth, schedule growth, or final margin). Nonparametric tests (Wilcoxon rank-sum and Kolmogorov-Smirnov) were used to test for statistical significance between populations, and the median was used as the measure of central tendency because the distributions were skewed. The Chi-Squared test was used to determine statistical significance for categorical variables.

Eliminating Outliers in Trend Analysis

Outliers can overly bias trend analysis, leading to trends that are too dependent on one or two significant outliers. As a result, the following four tests were employed on our analysis of cost-growth, price-growth, schedule-growth, and final-margin trends:

- Studentized residual test
- Leverage test
- Cook's Distance Test
- DFITS (Difference in Fits, Standardized) Test [also called DFFITS]

For example, our MDAP dataset that includes price and margin had a small sample of major programs from before the year 2000. These tests determined that those temporal outliers (along with a small number of others) should be rejected in our trend analysis.

Interpreting Box-and-Whisker Diagrams and Percentile Plots

Throughout this report, the so-called "box and whisker" charts (described in Figure B-1) help visualize the distribution of a particular variable. The gray boxes show the second and third quartiles (i.e., the 25th to the 50th percentile, and the 50th to the 75th percentile). The minimum and maximum are shown with a small bar at the end of the vertical line (or may run off the chart in some instances). The median (50th percentile, where half of the occurrences are above it and half below) is the *best measure of central tendency* in the data because the distributions are skewed. Note that the quartiles do not convey the actual distributions within the quartiles. As seen by the illustrations on the left of the figure, these distributions can be "lumpy" or nonuniform, but the charts do provide a quick visual for comparing two distributions. The charts also convey a sense of how much of the distribution is, say, negative or larger than a value of interest.

Next to some box-and-whisker charts we show the actual distributions so that we can see the distributions within each quartile. Figure B-2 shows how these percentile charts compare to the box-and-whisker charts, providing not only the quartiles but the whole distribution for each percentage. In this case, we show the box-and-whisker for the black line.

Figure B-1. Key to Reading the "Box and Whisker" Charts

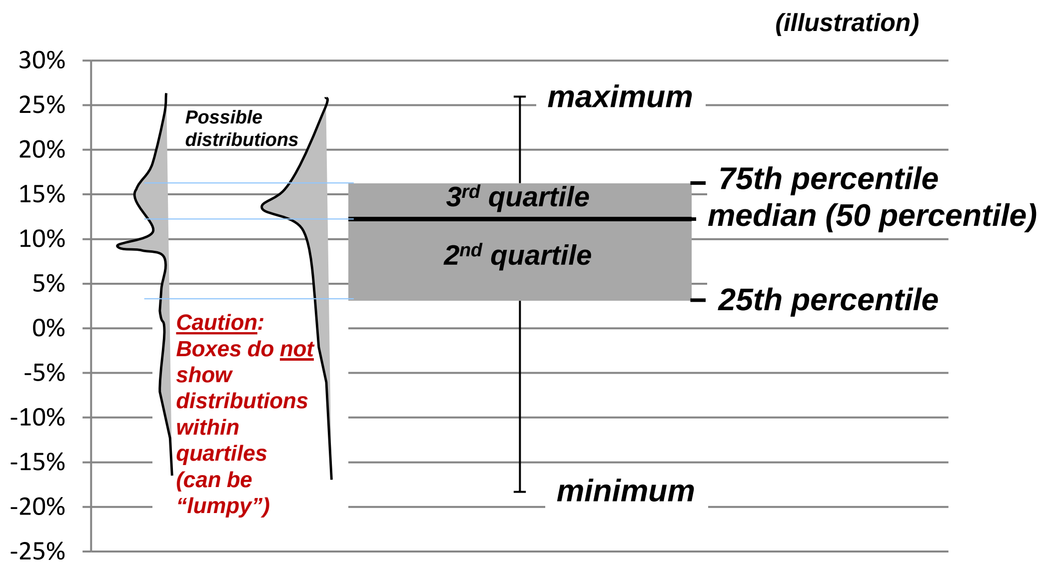
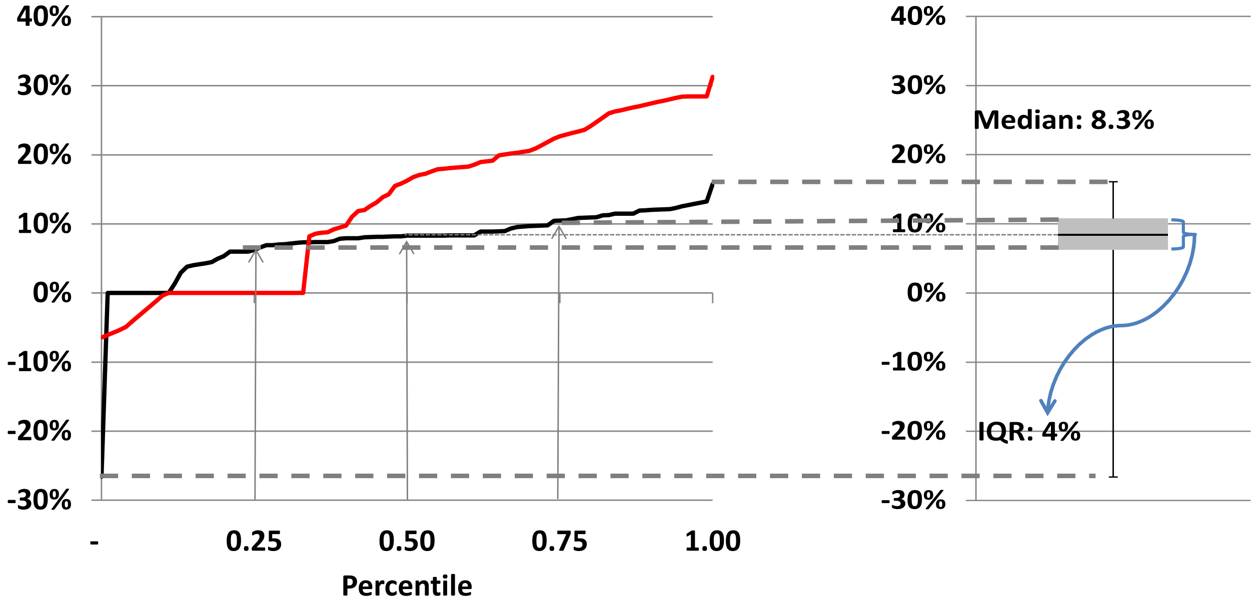


Figure B-2. Comparing Percentile and "Box and Whisker" Charts



Factor Analysis of Complexity and Risk

Complexity and risk are real attributes of major MDAP contracts in development and early production. However, they are not directly observable in readily available contract or program cost and schedule data. In earlier analysis, we often used either contract schedule or planned period of performance as a proxy for complexity. The rationale was that complex projects required a great many well-thought-out planned activities that had to be sequenced and synchronized in time. Thus, the complexity of a project should be reflected in its planned schedule (observed and measured in years). This proved useful, but to further test the results and seek a measure for risk we employed a more sophisticated factor analysis to develop new measures based on our readily available cost and schedule data.[42]

In theory, complexity and risk should be related to (and thus highly correlated with) the *stock* (static) variables of initial contract planned spending (dollars) and planned schedule, as well as

[42] Factor analysis can be used to identify underlying (hidden) factors based on a linear polynomial combination of the observable variables. Interpreting the resulting factors requires heuristic inference based on theoretical behaviors and expectations. Just as our prior use of one variable as a surrogate measure for another (e.g., using cycle time to measure complexity), there is no guarantee that the resulting factors constitute the interpretations we apply to them (in our case, that one is complexity and the other is risk), but observing the constituents of those factors and how the behave helps to provide credence to these inferences. It also helps to confirm prior results of other surrogate variables (i.e., whether complexity is going down as seen by reducing cycle time).

with the *flow* (dynamic) variables of cost-over-target (cost growth over the final contract cost target) and schedule growth.[43]

Factor analysis of the two observed stock variables and the two observed dynamic variables revealed two primary linear combinations of the four variables. When we examined the linear combinations for these factors and compared them to theory, we asserted that they were reasonable measure of complexity and risk (i.e., that they were weighted by their correlation to the two unobserved factors of complexity and risk):

- One linear combination of the four variables was heavily weighted toward initial contract spending and even more heavily weighted toward planned contract schedule. This linear combination most appropriately described the *complexity* of the project and is labeled *C*.
- The other linear combination was very heavily weighted toward the flow variables (cost-over-target and schedule growth). Since these two variables reflect the realization of "unknown unknowns" during contract execution, heavily weighting them most appropriately measures, ex post, the project's attribute of *risk* and is labeled *R*.

These two vectors were constructed to be completely independent of each other (i.e., orthogonal) and to be standard scores (i.e., both *C* and *R* sample scores have an arithmetic mean [average] of 0.0 and a standard deviation of 1.0). These calculated measures of *C* and *R* then could be used to conduct analysis of our dataset of major MDAP contracts for currently active programs.

Having identified the linear combinations and weightings of observed variables, factor analysis was continued to calculate two vectors of *C* and *R* corresponding to all the observed records in our MDAP contract dataset from 2000–2015.

Regression then was conducted on the data to measure the central tendency of *C* and *R* on these MDAP contracts. We should note that the dataset was controlled for inflation, for contract maturity, and for statistical outliers so that our regression results and trend analysis would not be distorted unduly and would better reflect the central tendency, true regression relationships, and true trends behind the observed data and the constructed factor data. We also controlled for phase by analyzing development contracts separately from early production contracts.

[43] "Stock" variables are measured at one point in time (e.g., at contract initiation in our case of planned contract cost and schedule). "Flow" variables are measured over time (i.e., throughout contract execution in our cases of cost growth over evolving contract target cost [cost-over-target] and schedule growth over time).

Figure B-3. Factor Analysis of *C* and *R* for Development

Observed variables	*C*	*R*
Cumulative growth of contract cost over total target cost	0.28	0.60
Cumulative contract schedule growth	-0.23	0.63
Initial value of contract spend (price)	0.46	0.19
Initial contract schedule	0.54	-0.13

***C* and *R* are the unobserved factors revealed through the observed variables by factor analysis.**

Weights or scoring coefficients based on the regression method and varimax rotated factors.

These weights are used to score each contract for *C* and *R*. This score is a standard score constructed to have an average of 0 and a standard deviation of 1.

NOTE: Conceptually, complexity should be most correlated with the stock values of initial price and schedule; whereas, risk should be most correlated with the dynamic growth rates of cumulative cost growth over total target cost and cumulative schedule growth rate. Thus, we use the variables C and R to denote the unobserved factors that align best with the inferred properties. Varimax searches for a rotation (i.e., a linear combination) of the original factors such that the variance of the loadings is maximized.

Figure B-4. Factor Analysis of *C* and *R* for Early Production

Observed variables	*C*	*R*
Cumulative growth of contract cost over total target cost	0.11	0.78
Cumulative contract schedule growth	-0.34	0.52
Initial value of contract spend (price)	0.45	0.09
Initial contract schedule	0.52	0.09

***C* and *R* are the unobserved factors revealed through the observed variables by factor analysis.**

Weights or scoring coefficients based on the regression method and varimax rotated factors.

These weights are used to score each contract for *C* and *R*. Each is a standard score constructed to have an average of 0 and a standard deviation of 1.

NOTE: Conceptually, complexity should be most correlated with the stock values of initial price and schedule; whereas, risk should be most correlated with the dynamic growth rates of cumulative cost growth over total target cost and cumulative schedule growth rate. Thus, we use the variables C and R to denote the unobserved factors that align best with the inferred properties.

(This page was left blank intentionally)

C. Abbreviations

ACAT—Acquisition Category

ACWA—Assembled Chemical Weapons Alternatives

AEHF—Advanced Extremely High Frequency

AIM—Air Intercept Missile

AoA—Analysis of Alternatives

APUC—Average Procurement Unit Cost

AR—Acquisition Reform

AT&L—Acquisition, Technology, and Logistics

ATIRCM/CMWS—Advanced Threat Infrared Countermeasure/Common Missile Warning System

BA—budget activity

BBP—Better Buying Power

BIW—Bath Iron Works

BLRIP—beyond low-rate initial production

BY—base year

C3I— Command, Control, Communications, and Intelligence

C4ISR—Command, Control, Communications, Computers, Intelligence, Surveillance, and Reconnaissance

CAPE—Cost Assessment and Program Evaluation

CAPP—Chemical Agent-Destruction Pilot Plant

CBB—Contract Budget Base

Chem-Demil—chemical (weapons) demilitarization

CMWS—Common Missile Warning System

CP—construction preparation

CPARS—Contractor Performance Assessment Reporting System

CPoF—Command Post of the Future

CY—calendar year

DAB—Defense Acquisition Board

DAE—Defense Acquisition Executive

DAWDF—Defense Acquisition Workforce Development Fund

DD&C—Detail Design and Construction

DDG—destroyer, guided-missile

DFITS—Difference in Fits, Standardized (also called DFFITS)

DLA—Defense Logistics Agency

DoD—Department of Defense

DoDI—Department of Defense Instruction

DOT&E—Director, Operational Test and Evaluation

DSARC—Defense Systems Acquisition Review Council

EAC—estimate at completion

ECSS—Expeditionary Combat Support System

ECT—Electronic Coordination Tool

EELV—Evolved Expendable Launch Vehicle

EMD—Engineering, Manufacturing and Development

EV—earned value

FAB-T—Family of Advanced Beyond-line-of-sight Terminals

FAR—Federal Acquisition Regulation

FCS—Future Combat Systems

FDD—Full-Deployment Decision

FFO—(date) funds first obligated

FFP—firm-fixed-price

FFRDC—Federally Funded Research and Development Center

FOC—Full Operational Capability

FPDS-NG—Federal Procurement Data System-Next Generation

FRP—full-rate production

FY—Fiscal Year

FYDP—Future-Years Defense Program

GAO—Government Accountability Office

GCSS-A—Global Combat Support System—Army

GMR—Ground Mobile Radios

GPS—Global Positioning System

GSR— Ground-Station Re-architecture

HMS—Handheld, Manpack, and Small Form-Fit

IQR—interquartile range

JCIDS—Joint Capabilities Integration and Development System

JLENS—Joint Land Attack Cruise Missile Defense Elevated Netted Sensor System

JPALS—Joint Precision Approach and Landing System

JSF—Joint Strike Fighter

JSOW—Joint Stand-off Weapon

JTN—Joint Tactical Networks

JTRS—Joint Tactical Radio System

JTRS GMR—Joint Tactical Radio System Ground Mobile Radios

JTRS HMS—Joint Tactical Radio System Handheld, Manpack, and Small Form-Fit (radio)

KLP—Key Leadership Position

KSA—Key System Attributes

KPP—Key Performance Parameter

LCS—Littoral Combat Ship

LPTA—Lowest-Price, Technically Acceptable

LRIP—Low-Rate Initial Production

MAIS—Major Automated Information Systems

MAR—MAIS Annual Report

MDA—Missile Defense Agency

MDAP—Major Defense Acquisition Program

Mod—modification

MPS—Mission Planning System

MS—Milestone

NAICS—North American Industry Classification System

NDAA—National Defense Authorization Act

O&M—Operations and Maintenance

OCO—Overseas Contingency Operations

OODA—observe, orient, decide, and act

OSD—Office of the Secretary of Defense

PAD—Preferred Alternative Decision

PARCA—Performance Assessments and Root Cause Analyses

PAUC—Program Acquisition Unit Cost

PB—President's budget

PE—program element

PEO—Program Executive Officer

PIM—Paladin Integrated Management

PM—program manager

PM EAC—Program manager's estimate at completion

PSC—product service code

R&D—research and development

RDT&E—Research Development Test and Evaluation

RESET—Ramsey Regression Equation Specification Error Test

RMS—Remote Minehunting System

SAE—Service Acquisition Executive

SAIC—Science Applications International Corp.

SAR—Selected Acquisition Report

S&T—science and technology

SBIRS—Space-Based Infrared System

SDD—System Development and Demonstration

SM—Standard Missile

SRW—Soldier Radio Waveform

SSIP—Superior Supplier Incentive Program

TMC—Tactical Mission Command

TSPR—Total System Performance Responsibility

TY—then year (not adjusted for inflation)

UAV—unmanned aerial vehicle

UCA—undefinitized contract action

UTC—United Technologies Corporation

ULA—United Launch Alliance

ULS—United Launch Services

U.S.—United States

U.S.C.—United States Code

USD—Under Secretary of Defense

USD(AT&L)—Under Secretary of Defense, Acquisition, Technology, and Logistics

VTUAV—Vertical-Takeoff-and-Landing Tactical Unmanned Aerial Vehicle (Fire Scout)

WGS—Wideband Global Satellite

WIN-T—Warfighter Information Network–Tactical

WSARA—Weapon System Acquisition Reform Act (of 2009)

ZBN—Z-variant, build new

D. Indices

Index of Figures

INDEX OF TABLES

E. References

Assistant Secretary of the Army for Acquisition, Logistics, and Technology, "Top Defense Firms Recognized for Superior Performance," Army News Service, U.S. Department of the Army, Washington, DC, July 14, 2015. Last accessed on July 22, 2015: http://www.army.mil/article/152251/Top_defense_firms_recognized_for_superior_performance/

Assistant Secretary of the Navy for Research, Development and Acquisition, "Department of the Navy Names Superior Suppliers for 2014," Navy News Service, U.S. Department of the Navy, Story Number: NNS140613-20, Washington, DC, June 13, 2014. Last accessed May 28, 2015: http://www.navy.mil/submit/display.asp?story_id=81669

_____,"Department of the Navy Names Superior Suppliers for 2015," Navy News Service, U.S. Department of the Navy, Story Number: NNS150714-15, Washington, DC, July 14, 2015. Last accessed July 23, 2015: http://www.navy.mil/submit/display.asp?story_id=90129

Defense Logistics Agency (DLA), "Defense Logistics Agency announces superior suppliers for 2013, 2014," DLA Public Affairs, Fort Belvoir, VA, December 11, 2014. Last accessed August 10, 2015: http://www.dla.mil/dla_media_center/pages/newsarticle201412110930.aspx

_____, "Defense Logistics Agency Announces Superior Suppliers for 2015," DLA News Center, DLA Public Affairs, Fort Belvoir, VA, July 15, 2015. Last accessed on July 22, 2015: http://www.dla.mil/DLA_Media_Center/Pages/newsarticle201507151030.aspx

Director of Program Assessments and Root Cause Analyses, "PARCA Root Cause Analysis for the Wideband Global Satellite (WGS) Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 17, 2010a.

_____, "PARCA Root Cause Analysis for the Advanced Threat Infrared Countermeasure/Common Missile Warning System (ATIRCM/CMWS) Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 19, 2010b.

_____, "PARCA Root Cause Analysis for the Remote Minehunting System (RMS)," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 20, 2010c.

_____, "PARCA Root Cause Analysis for the Apache Longbow Block III Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 21, 2010d.

_____, "PARCA Root Cause Analysis for the DDG-1000 Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 21, 2010e.

_____, "PARCA Root Cause Analysis for the Joint Strike Fighter (JSF) [F-35] Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 25, 2010f.

_____, "PARCA Root Cause Analysis for the Excalibur Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, December 2, 2010g.

_____, "PARCA Root Cause Analysis for the Enterprise Resource Planning Systems," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, February 23, 2011a. *[Includes GCSS-MC, Navy ERP, and early analysis of the Air Force ECSS.]*

_____, "PARCA Root Cause Analysis of the Assembled Chemical Weapons Alternatives (ACWA) Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 18, 2011b.

_____, "PARCA Root Cause Analysis of the [RQ-4A/B] Global Hawk Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 23, 2011c.

_____, "PARCA Root Cause Analysis of the Joint Tactical Radio System Ground Mobile Radios (JTRS GMR) Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, September 12, 2011d.

_____, "PARCA Root Cause Analysis of the Family of Advanced Beyond Line-of-Sight Terminals (FAB-T) Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, October 13, 2011e.

_____, "Results for PARCA's Root Cause Analysis of the Joint Land Attack Cruise Missile Defense Elevated Netted Sensor System [JLENS] Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 11, 2012a.

_____, "PARCA's Discretionary Root Cause Analysis of the [P-8A] Poseidon Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 11, 2012b.

_____, "PARCA's Root Cause Analysis of the Evolved Expendable Launch Vehicle [EELV] Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, June 21, 2012c.

_____, "Root Cause Analysis of the [Air Force] Expeditionary Combat Support System [ECSS] Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, August 28, 2013.

_____, "PARCA Root Cause Analysis of the Joint Precision Approach and Landing System Increment 1A (JPALS Inc. 1A) Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 23, 2014a.

_____, "PARCA Root Cause Analysis of the MQ-8 Vertical Takeoff and Landing Tactical Unmanned Aerial Vehicle Fire Scout (VTUAV) Program," Memorandum for the Under Secretary of Defense for AT&L, U.S. Department of Defense, Washington, DC, May 28, 2014b.

Government Accountability Office (GAO), *Defense Acquisitions: Defense Acquisitions Assessments of Selected Weapon Programs*, Washington DC, GAO-14-340SP, March 31, 2014a.

_____, *Defense Contracting: Factors DOD Considers When Choosing Best Value Processes Are Consistent with Guidance for Selected Acquisitions*, Washington, DC, GAO-14-584, July 30, 2014b.

_____, DoD Should Streamline Its Decision-Making Process for Weapon Systems to Reduce Inefficiencies, Washington, DC, GAO-15-192, February 2015a.

_____, *Defense Acquisitions: Defense Acquisitions Assessments of Selected Weapon Programs*, Washington DC, GAO-15-342SP, March 12, 2015b.

_____, *Defense Acquisition Process: Military Service Chiefs' Concerns Reflect Need to Better Define Requirements before Programs Start*, Washington, DC, GAO-15-469, June 2015c.

Haux, Hailey. Secretary of the Air Force Public Affairs Command Information, "AF Releases SSIP Results to DoD Contractors," U.S. Department of the Air Force, Washington, DC, July 14, 2015. Last accessed July 22, 2015: http://www.af.mil/News/ArticleDisplay/tabid/223/Article/608569/af-releases-ssip-results-to-dod-contractors.aspx

McNicol, David L., and Linda Wu, *Evidence on the Effect of DoD Acquisition Policy and Process on Cost Growth of Major Defense Acquisition Programs*, Alexandria, VA: Institute for Defense Analyses, P-5126, September 2014. Last accessed July 6, 2015: http://www.dtic.mil/dtic/tr/fulltext/u2/a609472.pdf

McNicol, David L., Sara K. Burns and Linda Wu, *Evidence on the Effect of DoD Acquisition Policy and Process and Funding Climate on Cancellations of Major Defense Acquisitions Programs*, Alexandria, VA: Institute for Defense Analyses, Paper P-5218, May 2015.

Remington, Kaye, and Julien Pollack, *Tools for Complex Projects*, Surrey, England: Gower Publishing, Ltd., 2007.

Schwartz, Moshe, *The Nunn-McCurdy Act: Background, Analysis, and Issues for Congress*, Congressional Research Service, June 21, 2010.

Schwartz, Moshe, and Kate M. Manuel, *GAO Bid Protests: Trends and Analysis*, Congressional Research Service, July 21, 2015.

Under Secretary of Defense for Acquisition, Technology, and Logistics (AT&L), "Implementation Directive for Better Buying Power—Obtaining Greater Efficiency and Productivity in Defense Spending," Memorandum for Secretaries of the Military Departments and Directors of the Defense Agencies, U.S. Department of Defense, Washington, DC, November 3, 2010.

_____, "Implementation Directive for Better Buying Power 2.0—Achieving Greater Efficiency and Productivity in Defense Spending," Memorandum for Secretaries of the Military Departments, Deputy Chief Management Officer, Department of Defense Chief Information Officer, Directors of the Defense Agencies, and AT&L Direct Reports, U.S. Department of Defense, Washington, DC, April 24, 2013a.

_____, *Performance of the Defense Acquisition System, 2013 Annual Report,* U.S. Department of Defense, Washington, DC, June 28, 2013b.

_____, *Performance of the Defense Acquisition System, 2014 Annual Report,* U.S. Department of Defense, Washington, DC, June 13, 2014.

_____, *Operation of the Defense Acquisition System*, Instruction 5000.02, U.S. Department of Defense, Washington, DC, January 7, 2015a.

_____, "Appropriate Use of Lowest Priced Technically Acceptable Source Selection Process and Associated Contract Type," U.S. Department of Defense, Washington, DC, March 4, 2015b. http://bbp.dau.mil/docs/Appropriate_Use_of_Lowest_Priced_Technically_Acceptable_Source_Selec_Process_Assoc_Con_Type.pdf

U.S. Department of Defense (DoD), Contractor Performance Assessment Reporting System (CPARS) Policy Guide, Washington, DC, June 2011. https://www.cpars.gov/cparsfiles/pdfs/DoD-CPARS-Guide.pdf

U.S. General Services Administration, *Product and Service Codes Manual—Federal Procurement Data System*, Office of Government-wide Policy, Aug. 2011.

Vergun, David, "Key defense suppliers recognized for stellar delivery to warfighters," U.S. Department of the Army, Washington, DC, February 9, 2015. Last accessed February 23, 2015: http://www.army.mil/article/142449/Key_defense_suppliers_recognized_for_stellar_delivery_to_warfighters/