INSTITUTE FOR DEFENSE ANALYSES

Further Evidence on the Effect of Acquisition Policy and Process on Cost Growth of Major Defense Acquisition Programs

(Revised)

David L. McNicol
David M. Tate
Sarah K. Burns
Linda Wu

August 2016

Approved for public release;
distribution is unlimited.

IDA Paper P-5330, Revised
Log: H 16-000168

INSTITUTE FOR DEFENSE ANALYSES
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

The Institute for Defense Analyses is a non-profit corporation that operates three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.

About this Publication

This work was conducted by the Institute for Defense Analyses (IDA) under contract HQ0034-14-D-0001, Project AY-7-405202, "Root Cause Analysis and Performance Assessment Methods and Analyses (Cost Growth Studies)," for the Director, Performance Assessments and Root Cause Analyses (D,PARCA). The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgments

Thank you to Mark Husband of Defense Acquisition University, Philip M. Lurie and Prashant R. Patel of IDA for performing technical review of this document.

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper P-5330, (Revised)

Further Evidence on the Effect of Acquisition Policy and Process on Cost Growth of Major Defense Acquisition Programs

(Revised)

David L. McNicol
David M. Tate
Sarah K. Burns
Linda Wu

Preface

This paper is a revised version of the IDA paper with the same publication number (P-5330) and same title released in June 2016. In the original version of this paper, one of the sources of historical cost growth data in acquisition programs was a DoD-sponsored study by another research organization published in 1996. Within several years of its publication, some of the data associated with that study were found—by the originating organization—not to be useful for analysis because cost growth was not measured from the relevant baseline or the program had been cancelled before it went into production. As a result of that discovery, the originating organization withdrew the data. However, because of a bureaucratic mishap, in the mid-2000s the uncorrected and previously withdrawn study was posted on the originating organization's website. When the IDA author of this paper requested a hard copy of that study, the originating organization provided a copy containing the incorrect data. The originating organization discovered this mishap in July 2016, removed the publication from its website, and notified the IDA author of this paper. Consequently, the IDA research team removed the flawed data from the database used for P-5330 and redid all of the statistical computations and the tables and graphs. This revised version of P-5330 incorporates those changes. The paper's principal findings, on the association of changes in DoD acquisition policies and processes and cost growth, remain unchanged.

Executive Summary

Institute for Defense Analyses (IDA) Paper P-5126[1] reported two significant findings. First, Major Defense Acquisition Programs (MDAPs) that entered Engineering and Manufacturing Development (EMD) during "bust" phases of the Department of Defense (DoD) bust-boom funding cycle on average had much higher cost growth than those that entered EMD during "boom" climates. Second, the paper found that additional reforms after those introduced in mid-1969 by then Deputy Secretary of Defense David Packard did not significantly reduce cost growth.

The latter conclusion leaves open the possibility that the Packard reforms reduced cost growth compared to the record of the 1960s. This is not simply an historical question because the main features of today's Office of the Secretary of Defense (OSD)-level acquisition oversight process remain those of the process installed by Packard in mid-1969. Moreover, the issue is salient now because of its implications for ongoing discussions of reform of the DoD weapon system acquisition process.

The database available for P-5126 did not contain cost growth estimates for any MDAPs that entered EMD during the 1960s, so that paper could not compare cost growth pre- and post-Packard reforms. This paper does so, using cost growth data for programs that entered EMD in the 1960s from a paper published by IDA in 1992.[2] It also uses a different cost growth metric and employs additional statistical tests.

The table on the following page reports Average Procurement Unit Cost (APUC) growth experienced by MDAPs that entered their EMD phase during six periods. Each of these periods is from a bust portion of the DoD bust-boom funding cycle; the different periods are marked by major changes in acquisition policies or processes. It is important to note that APUC growth is computed by comparing the EMD baseline value for APUC—which can be thought of as a goal or a prediction—to the actual APUC, normalized to the initial baseline quantity (or, for ongoing programs, to the December 2012 Selected Acquisition Reports (SARs), which were the most recent available when this project began). The APUC growth figures shown in bust periods in the table are the

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

2 Karen W. Tyson et al., "The Effects of Management Initiatives on the Costs and Schedules of Defense Acquisition Programs, Vol. I: Main Report," IDA Paper P-2722 (Alexandria, VA: Institute for Defense Analyses, 1992).

quantity normalized average for the MDAPs in that acquisition regime, binned by the year the MDAP entered EMD. This is done on the hypothesis that the acquisition policies and processes in place when an MDAP enters EMD have an effect on the amount of cost growth it experiences in the future.

Average APUC Growth by Acquisition Regime for MDAPs that Entered EMD during a Bust Funding Climate

Acquisition Regime	Period (FY)	Average APUC Growth*
McNamara-Clifford	1964–1969	90% (16)
Defense Systems Acquisition Review Council (DSARC)	1970–1980	36% (49)
Post-Carlucci DSARC	1987–1989	42% (10)
Defense Acquisition Board (DAB)	1990–1993	35% (11)
Acquisition Reform (AR)	1994–2000	78% (27)
DAB post-AR	2001–2002	113% (6)

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

* Normalized for changes in quantity.

A plausible reading of the averages in this table is as follows: Packard's radically new acquisition phases and his more highly structured process were successful in reducing APUC growth, which fell to less than half its average level during the 1960s. Perhaps encouraged by Packard's success and public distaste for cost growth, acquisition reform efforts persisted, but had no appreciable further effect on average cost growth prior to the AR years. The reduction of OSD oversight of MDAPs during the AR era coincided with the return of average APUC growth to nearly its 1960s level. In sum, the Packard reforms of late fiscal year (FY) 1969 appear to have reduced APUC growth; they were not significantly improved upon in this respect through the bust years that followed; and the AR years were associated with higher APUC growth, which may be related to a reduction of OSD-level oversight.

The question for the statistical analysis in an exploratory context like this is: Can cause reasonably be ascribed to the period-to-period changes in APUC growth, or are those changes more likely simply random fluctuations in the data? The statistical analysis presented in this paper supports two of the three points offered above—the Packard reforms did significantly reduce APUC growth and the further reforms introduced post-Packard and pre-AR did not yield significant further reductions in APUC growth. The results on the third point are less clear-cut, but on balance suggest that the higher APUC of the AR years is statistically significant.

Statistical analysis also sheds some light on what the 1969 Packard reforms achieved and why high APUC growth returned during the AR years. In addition to average APUC growth, the table below reports the number of MDAPs in the cohort that experienced APUC growth at least one standard deviation (S) above the sample mean ($\bar{X}$). The sample mean is 57.4 percent and the standard deviation is 85.4 percent, so one standard deviation above the mean is 143 percent. ($\bar{X}$ and S are computed for the bust periods only.) This is an arbitrary breakpoint adopted because it proves to be useful. MDAPs with this level of APUC growth will be called "extremely high cost growth" programs.

Average APUC Growth by Acquisition Regime and the Number of High Cost Growth MDAPs in Each Cohort, Bust Funding Climates

Acquisition Regime	Period (FY)	Average APUC Growth*	$\geq \bar{X} + S$
McNamara-Clifford	1964–1969	90% (16)	3
DSARC	1970–1980	36% (49)	0
Post-Carlucci DSARC	1987–1989	42% (10)	1
DAB	1990–1993	35% (11)	0
AR	1994–2000	78% (27)	7
DAB post-AR	2001–2002	113% (6)	1

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

* Normalized for changes in quantity.

The striking feature of these data is the paucity of extremely high cost growth programs after the introduction of the Packard reforms in 1969 and before AR. A total of 70 programs for which we have APUC growth estimates entered EMD during the 18 years of the DSARC, Post-Carlucci DSARC, and DAB periods in bust funding climates. Only one of these has an estimated quantity normalized APUC growth from the EMD baseline of at least 143 percent. (Note that MDAPs that were cancelled or truncated are not included in the database, although the inclusion of cancellations for which we have cost growth estimates would not alter this point.) The other side of this coin is the greater frequency of extremely high cost growth systems in the McNamara-Clifford years and during the AR period. Three of 16 programs of the McNamara-Clifford years showed extremely high cost growth, as did seven of 27 MDAPs that passed Milestone (MS) II/B during the AR years.

Statistical analysis supports the impression left by inspection of the data:

- The frequency of extremely high cost growth programs was significantly higher in the McNamara-Clifford years than in the DSARC period.
- The frequency of extremely high cost growth programs also was significantly higher during the AR years than during the DSARC period.

This pattern did not appear in parallel statistical tests using two other breakpoints (50 percent and 100 percent). That is, statistically, it is the extremely high cost growth systems that stand out.

Finally, we considered the pattern in average APUC growth across the six acquisition periods if the 12 extremely high cost growth programs are removed. The means of the truncated distributions are presented in the table below. Pair-wise tests found the average APUC growth for the AR years (without the extremely high cost growth systems) to be significantly lower than the averages for the McNamara-Clifford and DSARC periods. None of the other differences was statistically significant and a test of the table as a whole did not reveal significant differences. It appears, then, that the significant differences in average APUC growth mainly stem from the significantly higher proportion of extremely high cost growth system during the McNamara-Clifford and AR periods.

Average APUC Growth by Acquisition Regime during Bust Funding Climates, with Extremely High Cost Growth Systems Removed

Acquisition Regime	Period (FY)	Average APUC Growth*
McNamara-Clifford	1964–1969	47% (13)
DSARC	1970–1980	36% (49)
Post Carlucci DSARC	1987–1989	20% (9)
DAB	1990–1993	35% (11)
AR	1994–2000	18% (20)
DAB post AR	2001–2002	40% (5)

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

* Normalized for changes in quantity

This conclusion adds a level of detail to the interpretation of the APUC growth data offered above. Packard's radically new acquisition phases and his more highly structured process were successful in preventing instances of extremely high cost growth and, for this reason, significantly reduced average APUC growth. The relaxation of OSD oversight of MDAPs during the AR era saw a return of a significant number of extremely high cost growth systems and, for that reason, average APUC growth returned to nearly its 1960s level. In sum, the Packard reforms of late FY 1969 worked well in essentially eliminating instances of extremely high cost growth and in that way reduced average APUC growth, they were not significantly improved upon in this respect through the early 2000s, and the relaxation of OSD-level oversight of the AR years was associated with a significant number of extremely high cost growth programs and therefore of higher average APUC growth.

The Under Secretary of Defense (Acquisition, Technology and Logistics) (USD(AT&L)) can use the DAB process to bring MDAPs into conformance with acquisition policy at MS II/B. Among other things, programs should have employed the appropriate contracting mechanism, should have a sound test plan, should not proceed until the technologies to be employed are reasonably mature, should rest on realistic programmatic assumptions, and should be fully funded to a realistic cost estimate. It is not surprising, then, to find that (except in the AR years when OSD-level oversight was relaxed) the DSARC process and its successor the DAB process largely eliminated instances of extreme cost growth. This might be due to direct OSD-level modification of particular MDAPs. Alternatively, the certainty of reviews by the DSARC/DAB might have prompted the Services to avoid in the programs they proposed the characteristics that cause high cost growth. The best way to gain a deeper insight into the matter probably is to compare closely the AR period with the DSARC and to examine the extremely high cost growth programs.

It is surprising that the statistically significant differences are found only for the extremely high cost growth systems. The description of the process certainly suggests that it also should have an effect on programs with smaller but still very substantial cost growth. This finding, however, does not necessarily imply that the OSD-level process has no effect. Instead, the statistical finding as such is that the fairly rudimentary OSD-level process of the McNamara-Clifford years apparently did about as well as its more elaborate successors except on extremely high cost growth systems.

Contents

Supporting data files provided on CD (inside back cover):

- Data and Computations for P-5126 Cost Growth Estimates.xlsx
- IDA Paper P-2722-VOL-1.pdf
- IDA Paper P-2722 Vol 1_Main Report Appendix A Tables 1-10.xlsx
- Main Database V 4.1 2Aug2016.xlsx
- McCrillis Briefing.pdf
- Program Notes.docx

A. Introduction

Institute for Defense Analyses (IDA) Paper P-5126[1] reported two significant findings. First, Major Defense Acquisition Programs (MDAPs) that entered Engineering and Manufacturing Development (EMD) during "bust" funding climates on average had much higher cost growth than those that entered EMD during "boom" climates. Second, the paper found that additional reforms after those introduced in mid-1969 by then Deputy Secretary of Defense David Packard had not significantly reduced cost growth.

As P-5126 noted, the latter conclusion leaves open the possibility that the Packard reforms reduced cost growth compared to the record of the 1960s. If in fact they did, the conclusion of P-5126 would have to be amended to read: The introduction in 1969 of effective Office of the Secretary of Defense (OSD)-level oversight of major acquisition programs reduced cost growth, but the additional reforms of the 1970s, 1980s, and early 1990s did not result in further reductions. Along the same line, it is of interest to revisit the mixed evidence P-5126 found on the effect on cost growth of less active OSD-level oversight of 1994–2000. The crucial question is whether there is statistical evidence that cost growth decreased when OSD-level controls were imposed and also increased when those controls were relaxed.

This is not simply an historical question, because the main features of today's OSD-level acquisition oversight process remain those of the process installed by Packard in mid-1969. Moreover, the issue is salient now because of its implications for ongoing discussions of reform of the DoD weapon system acquisition process.

The database available for P-5126 did not contain cost growth estimates for any MDAPs that entered EMD during the 1960s, so that paper could not compare cost growth pre- and post-Packard. This paper uses cost growth data for programs that entered EMD in the 1960s from a 1992 IDA paper.[2] This paper also uses a different cost growth metric and employs additional statistical tests.

The next section briefly describes the OSD-level acquisition oversight introduced by Robert McNamara in the mid-1960s and the changes made to it in 1969 by Packard. It is necessary to do this because the McNamara reforms are no longer part of the collective memory of the DoD acquisition community. Subsequent sections then turn to the statistical analysis and the conclusions it suggests.

---

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

[2] Karen Tyson et al., "The Effects of Management Initiatives on the Costs and Schedules of Defense Acquisition Programs, Vol. I: Main Report," IDA Paper P-2722 (Alexandria, VA: Institute for Defense Analyses, 1992).

B. Origins of the OSD-Level Acquisition Oversight Process

From the creation of the National Security Establishment in 1947 through 1960, OSD had no institutionalized process for the oversight of major weapon system acquisitions.[3] The origins of the OSD-level process for overseeing major weapon system acquisitions lie in initiatives taken by McNamara, of which the following are especially relevant for current purposes:[4]

- Promulgation of policy on contract types
- Establishment of milestone decision points and the Development Concept Paper
- Active oversight of ongoing MDAPs[5]

These initiatives were an embryonic OSD-level acquisition oversight process.

McNamara directed the use of Total Package Procurement (TPP) when it was judged to be practicable and, when not, Fixed Price Incentive Fee (FPIF) or Cost Plus Incentive Fee (CPIF) contracts.[6] By 1966, McNamara had concluded that TPP contracts were in fact not a practicable way to acquire major weapon systems, although acquisition policy apparently still had a tilt towards fixed price contracts, even for development. Packard picked up on this topic where McNamara left off. He ruled out the use of TPP and discouraged the use of FPIF for development contracts in favor of CPIF. (Cost Plus Award Fee may not have been included in the contracting play book yet.) As a general matter, Packard's policy was to match contract terms to the riskiness of the acquisition.

---

3 The Secretary of Defense could, and on occasion did, act to cancel or initiate major acquisitions. Major acquisition programs were also subject to review during the budget cycle by the Office of the Assistant Secretary of Defense (Comptroller) and the Office of Management and Budget. Additionally, a major building block of McNamara's process began operating in 1959. See William D. O'Neil and Gene H. Porter, "What to Buy? The Role of Director of Defense Research and Engineering (DDR&E)—Lessons from the 1970s," IDA Paper P-4675 (Alexandria, VA: Institute for Defense Analyses, January 2011), 25.

4 These categories are abstracted from pages 35–45 of J. Ronald Fox, *Defense Acquisition Reform, 1969 to 2009*: *An Elusive Goal* (Washington, DC: U.S. Army Center of Military History, 2011). Fox also notes that McNamara moved to consolidate acquisition functions in defense agencies—e.g., the agency that became the Defense Logistics Agency—and promoted the use by program managers of particular management tools such as PERT and earned value.

5 There are several cases—most notably the F-111—in which McNamara played a very active role in the oversight of the program. These cases almost certainly are exceptions, but the literature survey done for this paper uncovered little about how the process worked in the more typical cases. Adding to the confusion, the sources consulted suggest that during the McNamara years a major acquisition program might arise in the Planning, Programming, and Budgeting System (PPBS) or in the acquisition process.

6 Gordon Adams, Paul Murphy, and William Grey Rosenau, *Controlling Weapons Costs: Can Pentagon Reforms Work?* (New York: Council on Economic Priorities, 1983), 19–20, as cited in Fox, *Defense Acquisition Reform, 1969 to 2009*, 38. A TPP contract is one that covers EMD, at least a significant portion of procurement, and at least part of the support of the system (for example, depot maintenance.)

Packard's establishment of the Defense Systems Acquisition Review Council (DSARC) often is seen as the hallmark of his 1969 reforms. The notion of milestone reviews, however, entered the OSD-level acquisition process in 1964 with issuance of DoD Directive (DoDD) 3200.9, Initiation of Engineering and Operational Systems Development.[7] This original version of the directive set one point at which OSD—in principle, the Secretary of Defense—approval was required for an acquisition program to proceed. In 1965, a second decision point was added and the Director, Defense Research and Engineering (DDR&E) instituted the precursor of the Development Concept Paper (DCP), which, starting in 1968, was required to initiate any major development project. DDR&E coordinated initial DCPs with concerned OSD offices (and probably the Joint Staff and other Services) and was what now would be called the Milestone Decision Authority (MDA) for the initial DCP.[8] Once approved by DDR&E, the proposed new start went to the Secretary of Defense, although the sources consulted do not indicate whether it went as a separate action or as part of the Service's budget submission. It is also not clear which OSD official was the MDA for the second milestone.

Viewed against this background, the establishment of the DSARC was an evolutionary step. The Development Concept Paper was renamed the Decision Coordinating Paper (retaining the acronym) to reflect the broader scope of the new milestone definitions. The MDA at Milestone (MS) I and MS II was DDR&E; the Assistant Secretary of Defense for Installations and Logistics was the MDA for MS III. Decisions at the DSARC level were advisory to the Secretary and Deputy Secretary of Defense but, apart from exceptional cases, they probably got to that level by way of the Service's proposed budgets (and the Comptroller was the backstop enforcer of the requirement for milestone approval before a program could advance to the next stage.)

OSD had a much larger role in oversight of major acquisition programs under the DSARC process than it did pre-1961. The picture in contrast to the McNamara years is less clear-cut. On one hand, under the new acquisition directives, the Secretary of Defense, while retaining full legal authority over acquisition programs, would act through the established acquisition process except in extraordinary circumstances, which in comparison to cases such as the F-111 implied less OSD-level control over acquisitions. On the other hand, the DSARC had a greater substantive scope for the more typical

---

7 The first version of DoDD 3200.9 was issued in 1964. A revision that made provision for the Contract Definition Phase was issued July 1, 1965. See Thomas K. Glennan, Jr., "Policies for Military Research and Development," RAND Paper P-3253 (Santa Monica CA: The RAND Corporation, November 1965), 12. O'Neil and Porter, "What to Buy?," 25-47, sketches how the process evolved and worked during the 1960s.

8 On the first point, see O'Neil and Porter, "What to Buy?," 30. On the second, see C. W. Borklund, *The Department of Defense* (New York: Frederick A. Praeger, 1969), 83.

program and was more tightly organized. For the large majority of major acquisition programs, then, the new DSARC process probably was more effective.[9]

The last of the McNamara acquisition initiatives to be considered here is the substance of the milestones.[10] The 1965 version of the DoDD 3200.9 process had three phases. The first of these "was called *concept formulation*. During concept formulation OSD and the Service(s) involved assured themselves that they were buying the right system to meet real needs and that the technology was fully ready."[11] Concept formulation typically was initiated by a Service, but involved DDR&E and the Office of the Assistant Secretary of Defense for Systems Analysis (OASD(SA)), and included what would now be called an Analysis of Alternatives led by OASD(SA). It also apparently included what would later be called a Mission Element Need Statement and also the main parts of an Acquisition Strategy and plans for oversight of the program as it proceeded.

Approval to proceed from the Concept Formulation phase authorized the Service sponsoring the program to fund at least one company to prepare a definitized contract proposal. The OSD (milestone) review of these proposals was the basis for award of a contract, usually to a single source, for development and procurement of the system. That is to say, the second of DoDD 3200.9's milestones combined what now would be called MS B and MS C authority.

Packard's reforms separated the decision to allow the program to enter EMD from the decision to enter the Production phase (now MS C) and required OSD-level approval of each decision. Packard also established a new Validation Phase, which has at various times since been called Demonstration and Validation, Program Development and Risk Reduction and, currently, Technology Maturation and Risk Reduction. MS I (now MS A) authorized entry into this phase. DoDD 3200.9's Contract Definition phase was collapsed into the new and broader Validation phase. These changes were more revolutionary than evolutionary.[12]

---

[9] Clark A. Murdock, *Defense Policy Formation* (Albany: State University of New York Press, 1974), 155–179, disagrees with this judgment. Murdock is primarily concerned with Systems Analysis and resource allocation, but also comments specifically on the acquisition process. In particular, he notes that the new Decision Coordinating Paper did not provide "any mechanism for ongoing managerial control." This is accurate in that the Packard reforms placed management of the programs in the hands of the Services. It is incomplete in that the Services were responsible for staying within what would later be called the Acquisition Program Baseline, and the MDA was enjoined to act in cases in which they did not.

[10] Fox, *Defense Acquisition Reform, 1969 to 2009*, 57, provides a useful schematic comparison of the DoDD 3200.9 milestones and those of Packard's DoDD 5000.1/DoDI 5000.2.

[11] O'Neil and Porter, "What to Buy?," 30.

[12] The revision of DoDI 5000.2, issued October 23, 2000, formally established MSs A, B, and C (in place of MSs I, II, and III) as the main decision points for an MDAP. The definitions are such that MS B is placed several months earlier in the process than MS II.

The provisional judgment offered here is that Packard's acquisition reforms provide a plausible reason for expecting program outcomes, measured by cost growth, schedule slips, and performance shortfalls, to be better than what was achieved during the McNamara-Clifford years. This judgment does not imply that DoD was doing a better job of deciding what to buy, but only that, as a result of the Packard reforms, OSD became more effective in oversight of acquisition programs from MS II through the completion of procurement.

C. Statistical Analysis of Average Cost Growth

The statistical analysis presented here rests on definitions of periods delimited by major changes in acquisition policy and process. Two of these already encountered are labeled "McNamara-Clifford" and "DSARC." Another part of the scaffolding of the analysis is budget climate. Two climates are distinguished—"bust" and "boom." Finally, the analysis rests on a set of conventions concerning which MDAPs are included in the database and the way in which cost growth is measured.

These topics are developed in P-5126 and a companion to it on cancellation of MDAPs.[13] For the convenience of those who have not read these papers, background material on acquisition periods and funding climates is included in Appendix A. Appendix B states the conventions used in assembling the database and identifies the sources of the data used. Readers who are generally familiar with the OSD-level acquisition process and various acquisition reform efforts can use Appendix A and Appendix B as references. Those with less familiarity with the acquisition process probably should read at least Appendix A before proceeding with the main text.

This section considers whether there are statistically significant differences in average unit cost growth across the successive acquisition regimes, using Average Procurement Unit Cost (APUC) as the measure of cost growth. "APUC growth" means growth in APUC in program base year dollars normalized to the baseline quantity approved at MS II/B. Attention in this section and most of the one that follows is limited to MDAPs that entered EMD during bust periods. Results for boom periods are briefly mentioned at the end of Section D and results on the comparison of average APUC growth in bust and boom periods are noted in Section E.

Table 1 reports average APUC growth experienced by MDAPs that passed MS II/B during each of the six acquisition regimes in a bust period. It is important to bear in mind that APUC growth is computed by comparing the MS II/B baseline value for APUC—which can be thought of as a goal or a prediction—to the actual APUC, normalized to the

---

[13] David L. McNicol, Sarah K. Burns, and Linda Wu, "Evidence on the Effect of DoD Acquisition Policy and Process and Funding Climate on Cancellation of Major Defense Acquisition Programs," IDA Paper P-5218 (Alexandria, VA: Institute for Defense Analyses, May 2015).

MS II/B quantity[14] (or, for ongoing programs, to the projected APUC in the December 2012 Selected Acquisition Reports (SARs), which were the most recent available when this project began).[15] The APUC growth figures shown are the quantity normalized average for the MDAPs in that acquisition regime, binned by the year the MDAP passed MS II/B. This is done on the hypothesis that the acquisition policies and processes in place when an MDAP passes MS II/B, particularly the rigor of the MS II/B review, have an effect on the amount of cost growth it experiences in the future.

Table 1. Average APUC Growth by Acquisition Regime for MDAPs that Entered EMD during a Bust Funding Climate

Acquisition Regime	Period (FY)	Average APUC Growth*
McNamara-Clifford	1964–1969	90% (16)
DSARC	1970–1980	36% (49)
Post Carlucci DSARC	1987–1989	42% (10)
Defense Acquisition Board (DAB)	1990–1993	35% (11)
Acquisition Reform (AR)	1994–2000	78% (27)
DAB post AR	2001–2002	113% (6)

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

* Normalized for changes in quantity.

A plausible reading of the averages in Table 1 is as follows: Packard's radically new acquisition phases and his more highly structured process were successful in reducing APUC growth, which fell to less than half the average level it had during the 1960s. Perhaps encouraged by Packard's success and public distaste for cost growth, acquisition reform efforts persisted, but had no appreciable further effect on average cost growth prior to the AR years. Reduction of OSD oversight during the AR era coincided with the return of average APUC growth to nearly its 1960s level. In sum, the Packard reforms of late FY 1969 appear to have reduced APUC growth; they were not significantly improved upon in this respect through the bust years that followed; and the AR years were associated with higher APUC growth, which may be related to a reduction of OSD-level oversight.

---

[14] About three-quarters of the MDAPs that passed MS II/B in the period FY 1988–FY 2007 acquired at least 90 percent of their MS II/B baseline quantity. The median program acquired 100 percent and the average program acquired 111 percent. See McNicol, Burns, and Wu, "Evidence on the Effect of DoD Acquisition Policy and Process," 7–8.

[15] We follow the convention of not including in the database any MDAP that was not at least five years beyond EMD (so that cost growth would have time to appear). The most recent SARs available when P-5126 was written were those for December 2012. Consequently, MDAPs that passed MS B during FY 2007 were the most recent included in the database.

The question for the statistical analysis in an exploratory context is: Can cause reasonably be ascribed to the period-to-period changes in APUC growth, or are those changes simply random fluctuations in the data?

It is useful to break this question into three parts. First, is the difference between the average APUC growth post Packard reforms (36 percent) and the average for FY 1964–FY 1969 (90 percent) statistically significant? The tests used found this difference to be statistically significant at the 6 percent level.[16] It is worth noting these reductions can be attributed in part to the policies on contract type that Packard instituted. Four of the 16 programs in the data set for FY 1964–FY 1969 used TPP, and one used a Firm Fixed Price (FFP) development contract. The average APUC growth for these five contracts was 131 percent; the average cost growth for the remaining FY 1961–FY 1969 programs was 72 percent.[17] TPP and FFP contracts were less commonly used during FY 1970–FY 1980, but three of the 70 MDAPs for which we have APUC growth data and that passed MS II/B during this period used a TPP contract and two (both subsequently cancelled) used an FFP development contract.

Second, are the differences in average APUC growth for the three periods between McNamara-Clifford and AR statistically significant? The tests used did not reveal any statistically significant differences between the averages of APUC growth in these three periods.[18] This implies that the lower average APUC growth (35 percent) of MDAPs that passed MS II during the DAB years (FY 1990–FY 1993), for example, cannot be attributed confidently to the full implementation of the DAB in 1990, because a change of this size has a considerable probability of occurring by chance.

Third, and finally, were the AR years associated with significantly higher average APUC growth? The results in this case were mixed. One test indicated that average APUC growth over the AR years was higher than it was for the three preceding periods

---

16 The Mann-Whitney U test rejected the null hypothesis (P=0.060) that the sample for the DSARC period and that for McNamara-Clifford period were drawn from the same population. (n_1=49, n_2 = 16, U = 516). A two-tail t-test assuming unequal sample variances found the difference in the means to be significant (P=0.084). The Kalmogorov-Smirnov (K-S) test showed that APUC growth estimates for the McNamara-Clifford period probably are not normally distributed. The result of the t-test, even with the correction for unequal variances, is therefore somewhat suspect.

17 For further discussion of TPP and FFP development contracts, see Tyson et al., "The Effects of Management Initiatives," Chapter X; David L. McNicol, *Cost Growth in Major Weapon Procurement Programs*, 2nd ed. (Alexandria, VA: Institute for Defense Analyses, 2004), 53 and 57–59; and O'Neil and Porter, "What to Buy?," 29–31.

18 We compared the three periods using Analysis of Variance (ANOVA). ANOVA failed to reject the null hypothesis that the observations in the three periods were drawn from identical normal populations (P = 0.905).. The K-S test found that the samples were consistent with the assumption that each was drawn from the same normal distribution.

combined. That result, however, was not confirmed by another test.[19] This is similar to the result found in P-5126 and it occurs for the same reason—the variability of APUC growth in the AR period was too large for the differences in the means to be statistically significant.

The Bayesian analysis presented in Appendix C provides a stronger result for the AR years. It finds clear evidence that both the McNamara-Clifford period (FY 1964–FY 1969) and the AR years (FY 1994–FY 2000) had a much higher probability of high cost growth than did the bust climate portions of the three intervening periods (DSARC, Post-Carlucci DSARC, and DAB).

Returning to the interpretation of Table 1 offered above, the statistical analysis of average APUC growth supports the first two of the points offered—the Packard reforms did reduce APUC growth and the further reforms introduced post-Packard and pre-AR did not yield significant further reductions in APUC growth. The results on the third point are not clear-cut. The statistical tests reported above do not support attributing the high mean APUC growth during FY 1994–FY 2000 to acquisition reform, but the results presented in Appendix C do support such an interpretation.

D. Statistical Analysis of the Proportion of Extremely High APUC Growth Programs

The preceding section looked for effects of acquisition policy and process in differences between successive periods in the average APUC growth of MDAPs that passed MS II/B during them. Although reasonable, framing the analysis in this way glosses over the possibility—explored in this section—that acquisition policy and process mainly work by influencing the proportion of MDAPs that experience extremely high cost growth.

Some relevant data are provided in Table 2. The average APUC growth figures are the same as those presented in Table 1. In addition, Table 2 reports the number of MDAPs in the cohort that experienced three different levels of APUC growth—50 percent, 100 percent, and one standard deviation (S) above the sample mean ($\bar{X}$). The sample mean is 57.4 percent and the standard deviation is 85.4 percent, so one standard deviation beyond the mean rounds to 143 percent. ($\bar{X}$ and S are computed for the bust periods only.) In what follows, MDAPs in the last of the categories will be called "extremely high cost growth" programs. These are arbitrary breaks adopted because they

[19] A two-tail t-test assuming unequal variances found the difference to be significant. (P=0.082.) The K-S test rejected the null hypothesis that the observations for either the AR period or the three preceding periods combined were normally distributed. A Mann-Whitney U test did not find a significant difference between the average APUC growth of the AR years and that for the three preceding periods combined.

proved to be useful. Note that the figures for the number of systems in the right tail are not additive. For example, of the 16 MDAPS that entered EMD during the period FY 1964–FY 1969, nine had APUC growth of at least 50 percent. Of these nine, five had APUC growth of more than 100 percent, and of the five, three had APUC growth of more than 143 percent.

The striking feature of the data in Table 2 is the paucity of extremely high cost growth programs after the introduction of the Packard reforms in 1969 and before AR. A total of 70 programs in our sample passed MS II during the 18 years of the DSARC, Post-Carlucci DSARC, and DAB periods in bust funding climates. Only one of these has an estimated quantity normalized APUC growth from the MS II baseline of at least 143 percent.[20] The other side of this coin is the greater frequency of extremely high cost growth systems in the McNamara-Clifford years and during the AR period. Three of 16 programs of the McNamara-Clifford years showed extremely high cost growth, as did seven of 27 MDAPs that passed MS II during the AR years.

Table 2. Average APUC Growth by Acquisition Regime and the Number of High Cost Growth MDAPs in Each Cohort, Bust Funding Climates

Acquisition Regime	Period (FY)	Average APUC Growth*	≥ 50%	≥ 100%	$\geq \bar{X} + S$
McNamara-Clifford	1964–1969	90% (16)	9	5	3
DSARC	1970–1980	36% (49)	17	5	0
Post-Carlucci DSARC	1987–1989	42% (10)	3	3	1
DAB	1990–1993	35% (11)	5	1	0
Acquisition Reform (AR)	1994–2000	78% (27)	11	7	7
DAB post-AR	2001–2002	113% (6)	2	1	1

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

* Normalized for changes in quantity.

Statistical analysis gives substantially the conclusions suggested by inspection of the data in Table 2. The three periods between McNamara-Clifford and AR will be referred to as the DSARC/DAB period. We found that:

- The frequency of extremely high cost growth programs was significantly higher in the McNamara-Clifford years than in the DSARC/DAB period.

[20] This is the FGM-148A Javelin. Roland also had a very high APUC growth (308 percent) but was placed on the cancelled list. Roland was developed during the mid-1960s by a French-German consortium. In 1975, the US Army decided to develop and procure a US version. The planned procurement was severely reduced, but enough was acquired to equip one Army National Guard battalion. This does not fully meet the definition of a cancellation but was judged to be closer to a cancellation than to a truncation of the program.

- The frequency of extremely high cost growth programs also was significantly higher during the AR years than during the DSARC/DAB period.[21]

In contrast to the results of the preceding section, both the McNamara-Clifford period and the AR period, then, stand out as having a significantly larger proportion of extremely high cost programs.

Table 3 lists the extremely high cost growth systems. Twelve of the 13 passed MS II/B during bust climates. Helicopters (2), satellite programs (3), and launch vehicles (2) are over-represented but do not dominate the list, particularly for the 1960s.

Table 3. Extremely High Cost Growth Systems

System Name	MS II/B FY	APUC
Bust Climates		
AGM-69 Short Range Attack Missile (SRAM)	1967	4.56
MIM-23 Hawk (Improved Hawk)	1965	2.07
M47 Dragon Guided Missile	1966	1.72
FGM-148A Javelin Advanced Anti-Tank Weapon System	1989	1.59
Space Based IR Sensor (SBIRS) High	1997	3.90
Evolved Expendable Launch Vehicle (EELV)	1998	3.42
Global Broadcast Service (GBS)	1998	2.60
Guided Multiple Launch Rocket System (GMLRS)	1998	2.15
H-1 Upgrades	1996	1.97
CH-47F (Improved Cargo Helicopter)	1998	1.81
Patriot Advanced Capability-3 (PAC-3)	1994	1.49
Advanced Extremely High Frequency (AEHF) Satellite	2001	4.78
Boom Climates		
Titan IV Expendable Launch Vehicle (ELV)	1985	1.49

We also explored whether the proportions of systems with cost growth of at least 50 percent or 100 percent might show the same pattern across acquisition periods as the extremely high cost growth systems. Analyses parallel with those just described with

[21] These statements are based on results for Fisher's Exact Test (FET): (1) P=0.019 in the comparison of McNamara-Clifford to the DSARC/DAB; and (2) P < 0.001 for the comparison of FY 1994–FY 2000 with the DSARC/DAB.

observations of at least 50 percent APUC growth and 100 percent APUC growth showed no significant differences across the acquisition periods.[22]

Appendix D presents results obtained from a technique (quantile regression) that compares the APUC growth distributions across acquisition regimes at several points. The comparison reported used deciles. The results were consistent with those stated above in two respects: (1) there were no significant differences across the six acquisition periods in the central portions of the distribution (4th through the 7th deciles), and (2) the McNamara-Clifford and AR periods had significantly fatter right tails. It also is interesting to note that there is some evidence that the left tails of these two periods were somewhat fatter than those of other periods; that is, McNamara-Clifford and AR had higher highs and perhaps higher lows.

Finally, we considered the pattern in average APUC growth across the bust phases of the six acquisition periods if the 12 extremely high cost growth programs are removed. The means of the truncated distributions are presented in Table 4. Pair-wise tests found the average APUC growth for the AR years (without the extremely high cost growth systems) to be significantly lower than the averages for the McNamara-Clifford and DSARC periods. None of the other differences was statistically significant and a test of the table as a whole did not reveal significant differences.[23] It appears, then, that the significant differences in average APUC growth reported in Section C stem from the significantly higher proportion of extremely high cost growth system during the McNamara-Clifford and AR periods.

---

[22] Fisher's Exact Test was used. The test with the breakpoint at 100 percent excluded the observations of at least 143 percent, and the test for observations of at least 50 percent excluded both the observations of at least 143 percent and those at least 100 percent but less than 143 percent.

[23] The comparisons were made using the t-test of the differences between the means of two independent samples. ANOVA for the table as a whole yielded P = 0.41. K-S found the observations (excluding extremely high cost growth programs) for four of the periods to be consistent with having been drawn from normal distributions. Anderson-Darling (A-D) found this same result for the other two periods (FY 1987–FY 1989 and FY 2001–FY 2002). A-D was used for those periods because of their smaller sample size.

Table 4. Average APUC Growth by Acquisition Regime for Bust Funding Climates, Excluding Extremely High Cost Growth Programs

Acquisition Regime	Period (FY)	Average APUC Growth*
McNamara-Clifford	1964–1969	47% (13)
DSARC	1970–1980	36% (49)
Post Carlucci DSARC	1987–1989	29% (9)
DAB	1990–1993	35% (11)
Acquisition Reform	1994–2000	18% (20)
DAB post-AR	2001–2002	40% (5)

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

* Normalized for changes in quantity.

Appendix E presents an analysis of the boom case that parallels that of this and the preceding section for the bust case. There was no indication of significant association between acquisition period and average APUC growth. The proportion of MDAPs in the far right tail of the distributions was significantly higher for the DSARC bust period than the DSARC boom period, but not significantly higher in the other cases.

E. Interpretation of the Statistical Results

The conclusions of the preceding section add a level of detail to the interpretation of the APUC growth data offered in Section C (p. 6). Packard's radically new acquisition phases and his more highly structured process were almost completely successful in preventing instances of extremely high cost growth and, for this reason, significantly reduced average APUC growth. The relaxation of OSD oversight of MDAPs during the AR era saw a return of a significant number of extremely high cost growth systems and, for that reason, average APUC growth returned to nearly its 1960s level. In sum, the Packard reforms of late FY 1969 worked well in essentially eliminating instances of extremely high cost growth and in that way reduced average APUC growth, they were not significantly improved upon in this respect through the early 2000s, and the relaxation of OSD-level oversight of the AR years was associated with a significant number of extremely high cost growth programs and therefore of higher average APUC growth.

The DAB process is a mechanism the Under Secretary of Defense for Acquisition, Technology and Logistics can use to bring MDAPs into conformance with acquisition policy at MS II/B. Among other things, programs should use the appropriate contracting mechanism, should have a sound test plan, should not proceed until the technologies to be employed are reasonably mature, should rest on realistic programmatic assumptions, and should be fully funded to a realistic cost estimate. It is not surprising, then, to find that (except in the AR years when OSD-level oversight was relaxed) the DSARC process and its successor, the DAB process, largely eliminated instances of extreme cost growth. This might be due to direct OSD-level modification of particular MDAPs. Alternatively, the certainty of reviews by the DSARC/DAB might have prompted the Services to avoid in the programs they proposed the characteristics that cause high cost growth. The best way to gain a deeper insight into the matter probably is to compare closely the AR period with the DSARC and to examine the extremely high cost growth programs.

It is surprising that the statistically significant differences are found only for the extremely high cost growth systems. The description of the process certainly suggests that it also should have an effect on programs with smaller but still very substantial cost growth. This finding, however, does not necessarily imply that the OSD-level process has no effect. Instead, the statistical finding as such is that the fairly rudimentary OSD-level process of the McNamara-Clifford years did as well as its more elaborate successors except on extremely high cost growth systems.

It is, finally, important to note that this paper has been concerned almost entirely with cost growth of MDAPs that passed MS II/B in bust periods. A complete summary also would need to take into account parallel analyses for the boom periods and the comparisons of cost growth in bust and boom periods for a given acquisition regime (Appendix E). That task, however, is postponed to a subsequent study.

Appendix A.
Background

This appendix provides brief descriptions of (1) each of the acquisition policy and process periods used, and (2) the funding climates. The material is drawn mainly from IDA P-5126 and IDA P-5218 (footnote 1, p. 1 and footnote 14, p. 6). Information on the data used is provided in Appendix B.

A. Acquisition Policy and Process Periods

The first steps towards installing an Office of the Secretary of Defense (OSD)-level acquisition oversight process were taken by then Secretary of Defense Robert McNamara in 1964. There have since been many dozens of changes in acquisition policy or process made by senior OSD officials and on the order of a dozen major changes directed by the Congress. We organize the large number of policy and process changes by identifying time periods during which the main features of acquisition policy and process remained approximately the same. These are as follows:

1. The McNamara-Clifford years, FY 1964–FY 1969
2. The Defense Systems Acquisition Review Council (DSARC), FY 1970–FY 1982
3. The Post-Carlucci Initiatives DSARC, FY 1983–FY 1989
4. The Defense Acquisition Board (DAB), FY 1990–FY 1993
5. Acquisition Reform (AR), FY 1994–FY 2000
6. The DAB Post-AR, FY 2001 to date

The best known of McNamara's changes in OSD decision-making processes was the creation of the Office of Systems Analysis (OSA) and the Planning, Programming, and Budgeting System (PPBS), which moved decisions on what major systems to procure to the Secretary of Defense level. OSA and PPBS so overshadowed development of an OSD-level acquisition oversight process that the latter is no longer part of the collective memory of the DoD acquisition community. Nonetheless, DoD Directive (DoDD) 3200.9, first issued in 1964, specified a milestone review for major acquisition programs, and McNamara was personally involved in the oversight of some acquisition programs.

The second acquisition period is marked by the set of reforms initiated by David Packard, then Deputy Secretary of Defense, in mid-1969. The shorthand used here to identify these changes is the "establishment of the DSARC." The DSARC process had all of the basic features that are still central to OSD-level oversight of MDAPs—three milestones,[1] with definitions similar to those used now; formal milestone reviews; a body to advise the Milestone Decision Authority (MDA); and a document (the Decision Coordinating Paper) that described the basic features of the program, provided a vehicle for staff inputs, and set down the cost, schedule, and performance goals that the program was to meet. In January 1972, the DSARC process was expanded to include an independent cost estimate at MS II and MS III provided by the newly established Cost Analysis Improvement Group (CAIG).[2]

The vehicle for the transition from the first phase of the DSARC (FY 1970–FY 1982) to the second (FY 1983–FY 1989) was the Carlucci Initiatives, named after then Deputy Secretary of Defense Frank Carlucci. These were developed during calendar 1981 and implemented during FY 1982–FY 1983. The Carlucci Initiatives did not involve any major changes in the DSARC process or in the policies Packard had established; in these terms the Carlucci Initiatives were more reaffirmation than change. The Carlucci Initiatives, however, included measures designed to coordinate decisions on MDAPs made in the PPBS and those made in the DSARC process. Important statutory changes were also made during 1982 and 1983.

The DSARC was followed by the DAB.[3] While the DAB itself bears a strong family resemblance to the DSARC, the statute that created it also created the position that is now called Under Secretary of Defense for Acquisition, Technology and Logistics (USD(AT&L)) and Service Acquisition Executives who reported to the USD(AT&L); designated USD(AT&L) as the MDA for most MDAPs; and removed the Service chiefs from the acquisition chain of command. In addition, the statute created the position of

---

1 The milestones were designated by Roman numerals until about 2000. Milestone (MS) I allowed a program to begin what was called Program Demonstration and Validation. MS II approval was required for a program to begin EMD, and MS III approval permitted the program to go into full rate production. The revision of DoDI 5000.2, issued October 23, 2000, formally established MSs A, B, and C (in place of MSs I, II, and III) as the main decision points for an MDAP. The definitions are such that MS B is placed several months earlier in the process than MS II.

2 The CAIG was established by a memorandum signed by Melvin Laird on January 25, 1972. A December 7, 1971 memorandum signed by David Packard directed the Military Departments to begin using "independent parametric cost analysis." See Donald Srull, ed., *The Cost Analysis Improvement Group: A History* (McLean, VA: Logistics Management Institute (now LMI), 1998), 47–48. Since the implementation of the Weapon Systems Acquisition Reform Act of 2009 (WSARA), the independent cost estimates have been provided by the Cost Assessment Deputate of the Office of Cost Assessment and Program Evaluation (CAPE).

3 There is some uncertainty about when the Post-Carlucci Reforms DSARC should end and the DAB period should begin. The relevant statutes were passed in December 1985 (first quarter FY 1986), and the DAB began functioning under that name in late FY 1987 or early FY 1988; however, DoD did not implement the full set of reforms required by statute until 1990. We have for that reason set the line at 1990.

Vice Chairman of the Joint Chiefs of Staff (VCJCS) and prompted a new requirements process centered on the VCJCS.

Acquisition Reform (FY 1994–FY 2000) was to a large extent intended to put acquisition of major weapon systems on a more commercial basis and make it easier for firms outside the defense sector to sell to DoD. Contracts were structured so that defense contractors assumed more responsibility for system performance; correspondingly, DoD's role in contract management decreased. OSD oversight of MDAPs also was relaxed somewhat. Substantial cuts in acquisition staffs at both the OSD level and Service Headquarters level were made, and senior decision makers took a more permissive attitude towards cost growth.

The post-AR DAB period was marked by the arrival of a new administration in January 2001. This period saw no overt rejection of Acquisition Reform but most of its initiatives were no longer pursued. The new administration seemed implicitly to favor a return to the *status quo ante* in OSD-level oversight of acquisition programs. In 2005, however, the Congress strengthened the Nunn-McCurdy Act and, in 2009, passed WSARA, which made several changes that may prove to be consequential.

B. Funding Climates

The amount appropriated to DoD each year for procurement over the period FY 1960 through FY 2010 is shown in Figure A-1. The data for the mid- to late-1960s are misleading. A large part of the increase in procurement funding during FY 1964–FY 1969 was for munitions and procurement to replace systems lost in combat, particularly aircraft. Insofar as modernization of weapon systems was concerned, there was little or no boom associated with the Vietnam War.[4] Recognizing this, the period FY 1964 until about 2010 saw two complete bust-boom cycles in DoD procurement funding.

---

4 These comments are based on an unpublished IDA working database drawn from various US government sources. We are indebted to Dr. Daniel Cuda for providing these data.

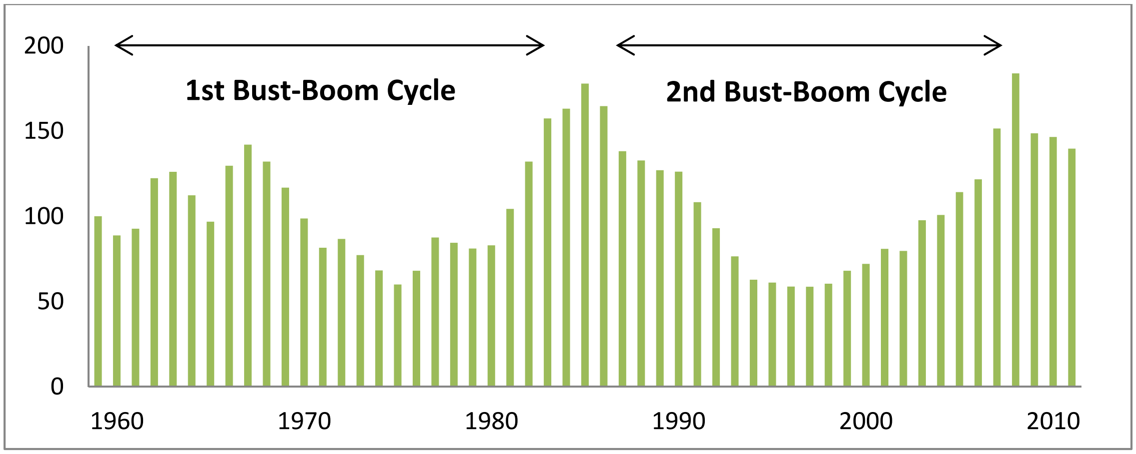

Figure A-1. DoD Budget Authority for Procurement in Billions of Constant FY 2014 Dollars

The boom phase of the first cycle was prompted by the invasion of Afghanistan by the Soviet Union in late December 1979 (that is, towards the end of the first quarter of FY 1980.) The budget that President Carter submitted to the Congress a few weeks later requested a substantial increase in DoD funding, including funding for procurement. Statements associated with the FY 1981 budget request promised further increases in future years. President Reagan, who took office in January 1981, requested increases in DoD funding that went well beyond what the Carter Administration had planned. By FY 1985, DoD funding for procurement was more than 70 percent above what it had been in FY 1981. The end of the Carter-Reagan defense buildup was marked by the passage in December 1985 of the Gramm-Rudman-Hollings Act, which provided for sequestration to satisfy budget targets.

The second bust phase continued—for a total of 16 years—through FY 2002. Procurement funding for FY 1987 was down about 16 percent from its FY 1986 level and dropped modestly further over the next three fiscal years. Additional large drops came after the fall of the Berlin Wall on November 9, 1989 (first quarter FY 1990) and the collapse of the Soviet Union on December 25, 1991 (first quarter FY 1992). By the mid-1990s DoD procurement funding had returned to about the Post-Vietnam War levels of the mid-1970s and was little more than one-third of its peak (FY 1985) level.

The defense spending boom that followed was prompted by the Al Qaeda attacks of September 11, 2001 and the subsequent wars in Afghanistan and Iraq. DoD funding for procurement increased in FY 2003 by about 20 percent over the preceding year and continued to increase through FY 2008. FY 2009 probably should be counted as the last year of this second boom period because the Great Recession began in its first quarter and the withdrawal of US troops from Iraq began in July of 2009. Expectations of a decline in defense spending developed at that time, and procurement funding in FY 2009 dropped by about 20 percent from the FY 2008 level. Further declines came after the Budget Control Act of 2011 was signed into law in August 2011.

It is important to realize that the turns in the funds appropriated for procurement are a lagging indicator of a change in budget climate. Funding and programmatic decisions embedded in the MS II/B baselines of MDAPs are made at least one or two years before the President's Budgets in question are submitted to the Congress. Consequently, those decisions necessarily reflect expectations held by decision makers about the future DoD budget climate, and the breakpoints between different budget climates should mark the points at which there were major shifts in expectations.

We used three events to identify the breakpoints between funding climates: (1) the invasion of Afghanistan by the USSR in late December 1979; (2) passage of the Gramm-Rudman-Hollings Act in December 1985; and (3) the terrorist attack on the United States on September 11, 2001. Senior decision makers could reasonably expect each of the events identified to signal major and sustained changes in the defense funding climate, which in fact they did.

Analysis of surrounding presidential statements and budgetary events led to the following breakpoints between the funding climates:

- FY 1981 as the first year of the Carter-Reagan buildup,
- FY 1986 as the final year of the Carter-Reagan buildup, and
- FY 2003 as the first year of the post-9/11 defense buildup.

The funding climates will be referred to as Relatively Constrained (or bust) and Relatively Accommodating (or boom).

Appendix B.
The Data

A compact disc (CD) in a pocket on the inside back cover of this paper contains the database used in this research. The CD also contains several other files that help to document or explain the data. This appendix provides background information that facilitates access to and use of the files on the CD.

A. Cost Growth Metric and Ground Rules

The cost growth metric used in this study is quantity normalized growth in Average Procurement Unit Cost (APUC) in program base year dollars. P-5126 used Program Acquisition Unit Cost (PAUC), which includes Research, Development, Test, and Evaluation (RDT&E) funding as well as procurement funding. APUC does not include RDT&E funding. Whether changes in acquisition policy and process had a significant association with growth in RDT&E cost is left as the subject of a future paper. In most instances the APUC growth figure used is measured from the MS II/B baseline.

Each of the programs for which the database includes an APUC growth estimate completed EMD, went into production, and fielded at least some units to operating forces. Programs that were cancelled (or programs that never proceeded beyond EMD or which were not an MDAP) were not used in the statistical analyses. One reason to not include cancelled programs is that it is generally very difficult, and sometimes impossible, to get a reasonable estimate of the unit cost growth they experienced. A better reason is that the relevant outcome metric for these programs is "cancellation." Another paper in this series (P-5218; see footnote 14, p. 6) examines the extent to which changes in acquisition policy and process are associated with differences in cancellation rates.

The database used in this research contains an estimate of APUC growth for 174 MDAPs that entered EMD during FY 1964–FY 2007. We follow the convention of not including in the database any MDAP that was not at least five years beyond EMD (so that cost growth would have time to appear.) The most recent Selected Acquisition Reports (SARs) available when P-5126 was written were those for December 2012. Consequently, MDAPs that passed MS B during FY 2007 were the most recent included in the database.

The estimates mainly are drawn from the database developed for P-5218, which in turn evolved from the database for P-5126. The cost growth observations for FY 1964–

FY 1969, however, and a few of the observations for FY 1970–FY 1986 are drawn from other studies, as is discussed below.

Figure B-1 is a histogram of the cost growth estimates in the database used for this paper. Forty-four of the MDAPs in the sample had negative quantity adjusted APUC growth; that is, their actual APUC normalized to the MS II/B quantity (or the most recent estimate of the actual) was less than that predicted at MS II/B. At the other end of the spectrum, seven programs had APUC growth in excess of 200 percent. Only 28 percent (48) of the MDAPs in the sample have an APUC growth that falls in the range of 20 to 40 percent. An important message of this figure is, then, that insofar as APUC growth is concerned, an MDAP with average cost growth is not typical.

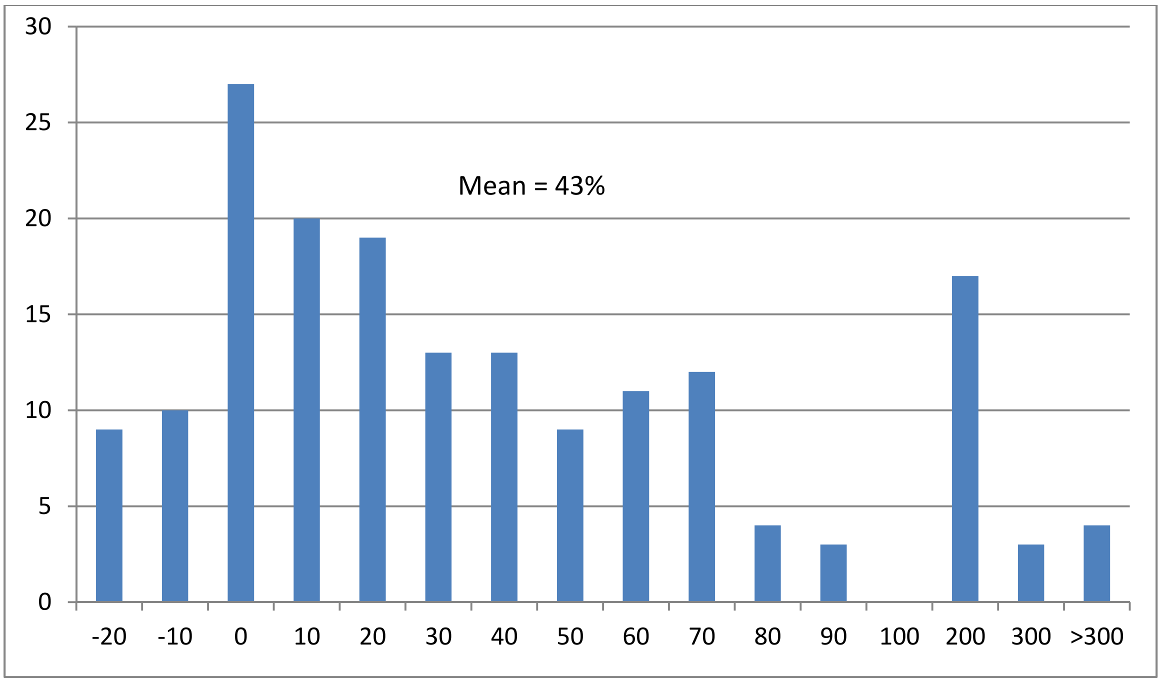

Figure B-1. Histogram of Quantity Normalized APUC Growth from MS II/B Baseline

B. Business Rules

Almost all of the data used in this research were taken directly or indirectly from SARs. SARs filed in FY 1997 and subsequent years are available through the Defense Acquisition Management Information Retrieval (DAMIR) system. Many SARS filed before FY 1997 are available on an Office of the Under Secretary of Defense (Acquisition, Technology and Logistics) (OUSD(AT&L)) SIPRNet site. These two sources provided SARs under 345 distinct labels.

Not all of these distinct labels are distinct programs. Three steps are needed to get from the list of distinct SAR labels to a list of MDAPs:

1. During the 1970s, each Component involved in a joint program sometimes filed a SAR. These SARs reported the same program data. The database used in this research includes only the data reported (for the entire program) in the SAR filed by the lead Component.
2. The program name used on the SAR often changes over the acquisition cycle for a given program. For example, the OH-58D Kiowa Warrior first reported as the Army Helicopter Improvement Program (AHIP). In most cases the database uses the name under which the last (or, for ongoing programs, most recent) SAR was filed.
3. Multiple MDAPs that have passed MS II/B are sometimes combined into a single MDAP. Conversely, a single MDAP that has passed MS II/B is sometimes split into two or more separate MDAPs. If the data permit (and they often did not), our rule was to maintain the program(s) as they had been defined at MS II/B.

For the reasons noted above, the database does not include any MDAPs that passed MS B after FY 2007. In addition, the following were excluded from the main database:

- Major Automated Information Systems (MAIS),
- Chemical Demilitarization Programs,
- Ballistic Missile Defense programs managed by the Ballistic Missile Defense Agency and its predecessors,
- Programs that filed a SAR but never were designated as an MDAP, and
- Programs cancelled before they passed MS II/B or before they were designated as an MDAP.

These exclusions were indicated by the purpose of the analysis, which is to gauge the effect of different OSD-level acquisition regimes and funding climates on MDAP outcomes. The database then should include only programs subject to OSD-level acquisition policy and process, and, to at least a significant extent, the excluded programs differed from the MDAP norm.

The exclusions resulted in a main database that includes 307 MDAPs. Programs that filed SARs before FY 2007 and were excluded from the main database are placed in a separate spreadsheet.

The great majority of the MDAPs in the database passed MS II/B at the OSD level. A few, however, entered at MS III/C, obtained both MS II/B and Low Rate Initial Production authority in a single OSD-level review, or passed MS II/B at the Service level and later became Acquisition Category (ACAT) I programs. These cases are noted in the database for MDAPs for programs that became MDAPs in FY 1989 or later, but not reliably noted for programs begun earlier.

Finally, it proved to be necessary to adopt a clear criterion for program cancellation. In the database a program is classified as cancelled if:

- The program did not result in production of any fully configured end items, or
- Any fully configured end items produced were used only for testing and development.

Application of this definition was not clear-cut for six programs that passed MS II/B at the Service level, later filed SARs, and subsequently were cancelled. We retained on the list of cancelled programs the five that had been designated as an ACAT I program and excluded the one that had not.[1]

Two other programs were counted as cancelled, although they did not exactly satisfy the criteria stated. The C-27J was included on the list of cancelled programs because the 21 C-27Js produced were placed directly in long-term storage and later transferred to Special Operations Command and the US Coast Guard. Roland was included, although the system was produced in the United States in limited quantities and issued to a single National Guard battalion, which falls into a gray area between issue of the system to active duty units and its use only for development, experiment, and training.

The file "Program Notes" (on the CD) provides some information bearing on each of the 57 cancellations we identified. We found 16 additional programs that filed one or more SARs during FY 1959–FY 2007 and were cancelled. These 16 were not included on the list of cancelled programs because they were either cancelled before passing MS II/B, were never designated an ACAT I program, or were cancelled after they fell below the ACAT I level. They appear as numbers 58–73 in "Program Notes."

C. Coverage

As was noted above, the database includes 57 MDAPs that were cancelled (as an ACAT I program) after passing MS II/B and includes 250 programs that went into production. We have a cost growth estimate (RDT&E, APUC, or PAUC, or usually all three) for 174 MDAPs. The database contains 76 MDAPs that were not cancelled but for which we do not have an APUC growth estimate. Overall, the database reports a cost

---

[1] AN/WQR-Advanced Deployable System, AQM-127A Supersonic Low Altitude Target, Advanced Seal Delivery System, ASM-135A Air-Launched Anti-Satellite System, and Land Warrior. Extended Range Munition was cancelled before it was designated an ACAT I program.

growth estimate for about 70 percent of the MDAPs that went into production. Table B-1 reports the relevant data broken down by the nine time periods used in the statistical analysis.

Table B-1. Percent of MDAPs Not Cancelled without an APUC Estimate, by Time Periods Used in the Statistical Analyses

Period (FY)	No. of MDAPs with an APUC Estimate	No. of MDAPs Not Cancelled without an APUC Estimate	Number of MDAPs Not Cancelled	Percent of MDAPs Not Cancelled without a Cost Estimate	No. of MDAPS Cancelled	No. of MDAPs in the Category
Bust						
1964–1969	16	5	21	24%	3	24
1970–1980	49	16	65	25%	11	76
1987–1989	10	3	13	23%	11	24
1990–1993	11	3	14	21%	1	15
1994–2000	27	10	37	27%	6	43
2001–2002	6	6	12	50%	4	16
Total	119	43	162	27%	36	198
Boom						
1981–1982	7	8	15	53%	4	19
1983–1986	31	18	49	37%	6	55
2003–2007	17	7	24	29%	11	35
Total	55	33	88	38%	21	109
Grand Total	**174**	**76**	**250**	**30%**	**57**	**307**

D. Sources of Cost Growth Estimates

The APUC growth estimates used in this paper were taken from four sources:

- The Program Analysis and Evaluation (PA&E) Cost Growth Database.
- Karen Tyson et al., "The Effects of Management Initiatives on the Costs and Schedules of Defense Acquisition Programs, Vol. I: Main Report," IDA Paper P-2722 (Alexandria, VA: Institute for Defense Analyses, 1992).

- Communication from the RAND Corporation providing updates to the FY 2015 SARs of estimates for six MDAPs published in a 1996 study[2] of APUC growth estimates normalized to the MS II/B baseline.
- David L McNicol and Linda Wu, "Evidence on the Effect of DoD Acquisition Policy and Process on Cost Growth Major Defense Acquisition Programs," IDA Paper P-5126 (Alexandria, VA: Institute for Defense Analyses, September 2014).

These four sources use the same definitions of the relevant cost terms and are based on SAR data. Each also, in most instances, measures cost growth from the MS II/B baseline when it is available and reports quantity normalized unit cost growth. Thus, an APUC estimate from, for example, Tyson et al. (hereafter P-2722) means the same thing as an APUC estimate from the other three sources.

The PA&E cost growth database is documented in a briefing by John McCrillis given at the 2003 Annual DoD Cost Analysis Symposium. The briefing is included on the CD provided on the inside back cover of this paper. The CD also includes the main volume of P-2722, and an Excel workbook with the data. The next section of this appendix describes how the P-5126 cost growth estimates were made.

The PA&E estimates were constructed through a detailed examination of the SAR variances. The IDA P-2722, IDA P-5126, and RAND estimates were made with data at a much more aggregated level. The methods used were essentially the same, but it is reasonable to assume that they differ in detailed ways not captured by the general characterization each offers of the method used. P-2722 did not in all cases follow the business rules used in P-5126 and this paper.

There were several MDAPs from the 1960s and 1970s for which we had two APUC growth estimates. The decisions on which of alternative estimates to use was entirely rules-based. The PA&E database did not provide estimates for MDAPs that entered EMD during FY 1964–FY 1969. The APUC growth estimates used for FY 1960–FY 1969 are from P-2722. In addition to the SAR data, P-2722's estimates in many cases reflected other sources of information, including material provided by the program office and contractors. For FY 1970 and beyond, we used the PA&E estimate in all cases in which the last SAR for the program had been filed by the time of the final update of the PA&E database (which used the December 2004 SARs). In a few cases, P-2722 had a cost growth estimate for a program not included in the PA&E database. In these instances we

[2] Jeanne M. Jarvaise, Jeffrey A. Drezner, and Dan Norton, "The Defense System Cost Performance Database: Cost Growth Using Selected Acquisition Reports," RAND Report MR-625-OSD (Santa Monica, CA: The RAND Corporation, 1996).

used the estimate from P-2722 if the program was reported complete in the most recent SARs used in making the estimate, and otherwise used the RAND estimate.

As is described in the next section, P-5126 made estimates of unit cost growth for 58 MDAPs using the program's final SAR or for programs then still underway in the December 2004 SAR. Table B-2 summarizes the number of estimates drawn from each of the sources.

Table B-2. Sources of the APUC Growth Estimates Used in Different Periods

Period (FY)	PA&E	P-2722	RAND	P-5126	Total
1964–1969	0	16	0	0	16
1970–1979	36	8	2	0	46
1980–1989	45	0	4	2	51
1990–1999	7	0	0	29	36
2000–2007	0	0	0	25	25
Total	**88**	**24**	**6**	**58**	**174**

E. Computation of the P-5126 Estimates

Fifty-eight of the MDAPs in the PA&E cost growth database were still ongoing at the time of the final PA&E update (that is, when the December 2004 SARs were filed). These had to be replaced with estimates that incorporated data from what were then the most recent SARs when this project began (those for December 2012). This section briefly describes how those estimates were made. The relevant data and computations are in the Data and Computations for P-5126 Cost Growth Estimates.xlsx file on the CD.

1. RDT&E

The SARs report fully configured units acquired with RDT&E funds and those acquired with procurement funds. Only the former are used in computing quantity-adjusted RDT&E cost growth. Our procedure was simply to compute the ratio of the Current Estimate (CE) of RDT&E cost and the baseline RDT&E cost (both in program base year dollars) and scale that by the ratio of baseline quantity to CE quantity. Suppose, for example, that the number of fully configured units purchased with RDT&E funds has increased from four to five and that CE RDT&E cost is 50 percent larger than the baseline cost. Our computation of unit RDT&E cost growth is then (4/5) x 1.5 - 1,or 20 percent.

2. APUC

The method used to normalize APUC for quantity change depended, first, on the extent to which quantity changed between MS II/B and the final SAR and, second, on whether a useable estimate of the slope of the learning curve was available.

a. No Quantity Change (NQC)

The SAR CE quantity was within ±1 percent of the MS II/B quantity for 13 of the MDAPs for which estimates were required. No quantity normalization is needed for these programs; their APUC growth is computed by dividing the CE APUC in the final SAR (or the December 2012 SAR for an ongoing program) by the MS II/B APUC and subtracting 1. The APUC growth for SBIRS-High (SBIRS-H) also falls under this heading. The total number of SBIRS-H satellites to be acquired decreased from five (at MS II) to four (the December 2012 SAR). The decrease, however, was in a satellite purchased with RDT&E funds, and we did not put these on a learning curve. There was no change in the number of SBIRS-H satellites purchased with procurement funds. Finally, although the PAC-3 quantity change fell outside the ±1 percent boundary, data limitations made it necessary to compute the PAC-3 APUC growth as the ratio of the CE and MS II APUCs.

b. Defense Acquisition Management Information Retrieval System (DAMIRS) Learning Curve (DLC)

The DoD contractor staff for DAMIRS provided us with their estimates of learning curve parameters that we were able to use to compute APUC growth for 14 MDAPs that passed MS II/B during FY 1989–FY 2001. We refer to these as the DAMIRS Learning Curve (DLC) APUC growth estimates. For each of these, we took the CE APUC growth in program base year dollars from the last SAR for the program or the December 2012 SAR (for still ongoing programs). The task was to normalize this APUC estimate to the MS II/B quantity, which was done as follows:

- We used the learning curve to compute the recurring flyaway cost at the MS II/B baseline quantity.
- The CE estimates of RDT&E and non-recurring flyaway cost were taken from the final SAR for the program or from the December 2012 SAR (for still ongoing programs).
- Support costs paid for with procurement dollars are, for many programs, primarily initial spares and support equipment, although other items may also fall into this category. Initial spares and support equipment normally scale with the number of units of the system purchased. For that reason, we used the CE

support cost reported in the last or most recent SAR scaled to the MS II/B baseline quantity.

c. Calibrated Learning Curve (CLC)

Twenty-nine MDAPs did not have a PA&E estimate or estimated learning curve parameters, and their CE quantity was significantly different from the MS II/B quantity. The approach we used in those cases rested on a cost progress curve of the conventional form:

$$C = TQ^{\beta} \quad (1)$$

In this expression, C is recurring flyaway cost, T is first unit cost, Q is cumulative production, and β is the cost progress parameter. We solved this and used the CE for recurring flyaway to get:

$$\hat{T} = CQ^{-\beta} \quad (2)$$

This will be referred to as the calibrated learning curve (CLC) method. A value of β = 0.94 was used for each of the programs. From this point, the computations were the same as those for MDAPs for which DAMIRS staff provided the learning curve parameters.

3. PAUC

Quantity PAUC is simply the sum of quantity normalized RDT&E and Procurement, divided by the baseline quantity, less 1.

F. Summary

Table B-3 below provides an overview of the number of estimates in P-5126 made with each of the methods.

Table B-3. Sources of the Quantity Normalized Unit Cost Growth Estimates Used in Different Periods

Period (FY)	NQC	DLC	CLC	Total
1989–2001	6	14	17	37
2002–2007	8	0	13	21
Total	**14**	**14**	**30**	**58**

G. Comparison of the PA&E and CLC PAUC Growth Estimates

It is pertinent to ask how the P-5126 estimates compare to those for the same programs in the PA&E database. This is primarily a question about FY 2002–FY 2007, which has a notably low average PAUC growth (10 percent) and for which about two-thirds of the PAUC growth estimates were made with the CLC method. It is secondarily a question about FY 1989–FY 2001 in relation to FY 1970–FY 1988. In the earlier period, all of the PAUC growth estimates were from PA&E, while in the later, about two-thirds were made using either the DLC or the CLC method.

The obvious approach to this issue is to compare the PA&E PAUC growth for systems that have been completed with PAUC growth for those same systems computed using the DLC and the CLC methods. Unfortunately, there are no MDAPs that have been completed and for which we have both a PA&E PAUC growth estimate and the data needed to compute a DLC or CLC estimate.

The best we can do is to examine the 23 MDAPs that passed MS II/B during FY 1989–FY 2001 and for which we have a PA&E PAUC growth estimate, a DLC estimate, and a CLC estimate. As was noted above, the PA&E estimates were most recently updated with the 2004 SARs. The DLC and CLC estimates, in contrast, incorporated more recent data—either the final SAR for the program or, for ongoing programs, the December 2012 SAR. Consequently, in most cases we would expect the DLC and CLC PAUC growth estimates to be larger than the corresponding PA&E estimate. That is the test: A method fails if it yields estimates that are "too often" and by "too much" less than the PA&E estimates. Clearly, this is a weak test.

The relevant estimates are presented in Table B-4. The comparison of the PA&E estimates and CLC estimates is on the left, and the comparison of the PA&E and DLC estimates is on the right. The CLC estimates are larger than the PA&E estimates for 17 of the 23 MDAPs—in most cases, considerably larger. They are smaller in six cases (shaded rows). In all but one of these cases (Joint Direct Attack Munition, or JDAM) the differences are absolutely or relatively small. The average of CLC APUC growth estimates is 77 percent in comparison to an average of 60 percent for the PA&E estimates. The DLC estimates exhibit the same pattern. The average of the DLC estimates is 73 percent, and four of them (shaded rows) are less than the PA&E estimate for the program—three by a substantial amount.

Table B-4. Comparison of PA&E, CLC, and DLC PAUC Growth Estimates for 23 MDAPs

Program	PA&E	CLC	Program	PA&E	DLC
LONGBOW APACHE	78%	117%	LONGBOW APACHE	78%	133%
F-22	41%	71%	F-22	41%	55%
F/A-18E/F	6%	12%	F/A-18E/F	6%	9%
BRADLEY UPGRADE	39%	54%	BRADLEY UPGRADE	39%	86%
MIDS	30%	72%	MIDS	30%	68%
CEC	48%	62%	CEC	48%	62%
H-1 UPGRADES	124%	192%	H-1 UPGRADES	124%	197%
LPD 17	43%	71%	LPD 17	43%	72%
CH-47F	147%	173%	CH-47F	147%	156%
GMLRS/GMLRS AW	125%	249%	GMLRS/GMLRS AW	125%	243%
MH-60S	62%	69%	MH-60S	62%	70%
Tactical Tomahawk	24%	28%	Tactical Tomahawk	24%	27%
GBS	10%	31%	GBS	10%	33%
Stryker	21%	25%	Stryker	21%	22%
UH-60M Black Hawk	49%	62%	UH-60M Black Hawk	49%	61%
WGS	28%	55%	WGS	28%	42%
C-130J	70%	84%	C-130J	70%	70%
JPATS	43%	40%	JPATS	43%	44%
SSN 774	35%	33%	SSN 774	35%	37%
JDAM	18%	-10%	JDAM	18%	-13%
JAVELIN	229%	197%	JAVELIN	229%	134%
MH-60R	95%	74%	MH-60R	95%	80%
NAS	25%	21%	NAS	25%	1%
Average	**60%**	**77%**		**60%**	**73%**

Note: The PA&E estimates were updated only through the 2004 SARs. The CLC and DLC estimates incorporate information from the last SAR for the program or the December 2012 SAR (for still ongoing programs).

Appendix C. Bayesian Contingency Analysis of Extremely High Cost Growth Frequency

A. Introduction

Section C introduced the concept of extremely high cost growth, defined as a percent cost growth more than one standard deviation above the sample mean. This appendix presents a Bayesian analysis of the hypothesis that programs passing Milestone II/B during the DSARC/DAB regimes[1] were less likely to experience extreme cost growth than were programs in the other observed periods. In particular, we explore the hypothesis that the DSARC/DAB probability of extreme cost growth was lower than the rates in surrounding periods by at least Q percentage points, for various values of Q.

B. Bayesian Contingency Analysis

The basic framework for a Bayesian contingency analysis is as follows. We first group the data into the periods to be compared, and note the number of observations and "successes" in each period. In the present case, success denotes cost growth more than one standard deviation above the sample mean. The resulting data are shown in Table C-1.

Table C-1. Occurrence of Extremely High Cost Growth in Each Period

Acquisition Regime	Period (FY)	Number of Programs	$\geq \bar{X} + S$
McNamara-Clifford	1964–1969	16	3
DSARC/DAB bust	1970–1993	70	1
AR and Post-AR DAB	1994–2002	33	8

For the three periods, we model the number of successes as independent Binomial random variables with known sample sizes n_1, n_2, and n_3, and unknown success probabilities p_1, p_2, and p_3. We then assign a prior distribution on the values of p_1, p_2, and p_3, compute the posterior distributions of each given the observed data, and use those posterior distributions to calculate the probabilities of various assertions of interest, such

[1] For purposes of this appendix, "DSARC/DAB" refers to the periods identified in the main text as "DSARC," "Post-Carlucci DSARC," and "DAB."

as whether $p_2 < p_1$ and $p_2 < p_3$. Prior distributions for Binomial probabilities are generally chosen to have Beta distributions, which have the convenient property that the resulting posterior distribution will also have a Beta distribution, with parameters that have a natural interpretation in terms of how much information the past data provide, relative to future observations.

Our null hypothesis is that there is no difference in success rates across the three periods. We assign identical Beta prior distributions to each of p_1, p_2, and p_3. Given the null hypothesis, it is reasonable to select the parameters of the Beta distribution so that the expected value is the overall success rate in the sample (12/119 = 0.10). In addition, we wish the prior distribution to have a high variance, corresponding to low certainty regarding the true success rate. Those considerations led to the choice of identical Beta(0.1, 0.9) priors, which correspond to having already observed 1/10 of a success and 9/10 of a failure. (Equivalently, one can think of this as having observed a single past trial and being 90% certain that it was a failure.)

C. Results

Given a Beta(a,b) prior distribution, the posterior distribution after observing a further m successes and k failures is Beta(a+m,b+k). Given the observed data, the posterior distributions of p_1, p_2, and p_3 are thus Beta(3.1,13.9), Beta(1.1,69.9), and Beta(8.1,25.9), respectively. The estimated posterior probabilities for the three periods are $\hat{p}_1 = 0.182$, $\hat{p}_2 = 0.015$, and $\hat{p}_3 = 0.238$, the means of the posterior distributions. The prior distribution and posterior distributions for each period are displayed graphically in Figure C-1.

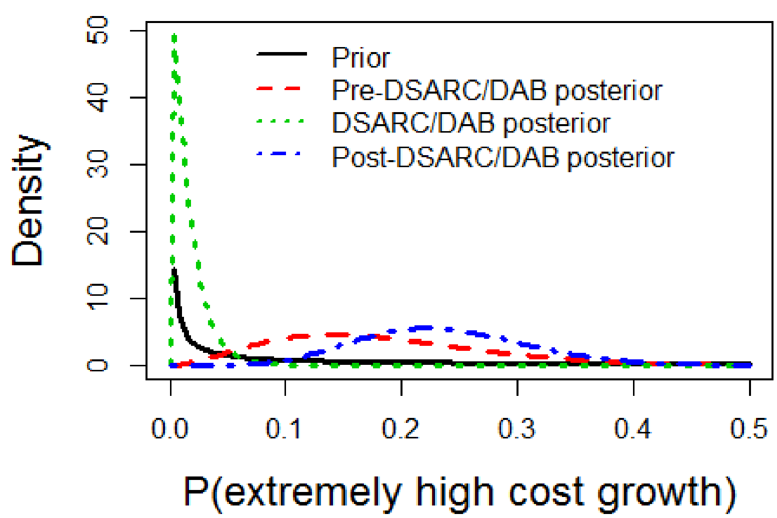

Figure C-1. Prior and Posterior Distributions for P (extremely high cost growth)

We define the event L_Q to mean that p_2 is less than both p_1 and p_3 by at least Q—that is, that the probability of extremely high cost growth was lower in the DSARC/DAB period by at least 100Q percentage points than in either the preceding or following periods. To assess the probability of L_Q for various values of Q, we used Monte Carlo simulation to draw 100,000 samples from the posterior distributions of p_1, p_2, and p_3 respectively, and tested the frequency with which event L_Q occurred. The results are shown in Table C-2.

Table C-2. Probability that DSARC/DAB Extremely High Cost Growth Rate Was Lower by at Least Q

Q	$P(L_Q)$
0	.995
0.01	.990
0.02	.981
0.05	.923
0.1	.723
0.2	.195

We can interpret this as follows. Given the assumed prior distributions:

- The probability that the DSARC/DAB years were less likely to produce extreme cost growth than either other period is 99 percent.
- The probability that the DSARC/DAB rate was at least 5 percentage points lower than in either other period is greater than 90 percent.
- There is a more than 70 percent chance that the DSARC/DAB rate was lower than both other periods by at least 10 percentage points.
- It is plausible (probability ~20%) that the DSARC/DAB rate was lower than both other periods by at least 20 percentage points, despite the fact that the overall observed rate of extremely high cost growth was only ~10 percent.

Appendix D.
Quantile Regressions

Quantile regression can be used to investigate systematically which parts of a distribution of outcomes—APUC growth, in this case—are associated with changes in treatment groups—acquisition policy and process periods for bust climates. We estimated a decile regression using the 119 APUC observations in the database for bust periods. The observations were grouped into the same six periods used in the main text. Standard errors were estimated using the Stata bootstrap procedure with 20 replications.

The results are presented in Table D-1 on the following page. The deciles will be referred to as q10, q20, and so on. Note first that the DSARC period is used as the reference case. The constant terms are the deciles for the distribution of the DSARC observations. For example, the constant term for q50 (0.31) is the median of the distribution of the observations for the DSARC period, and 20 percent of the observations are less than the constant term for the q20 (-0.01). The estimated coefficients reported are the difference between the q^{th} decile for the period and the q^{th} decile for the DSARC, that is, the constant term. Thus, the median for the McNamara-Clifford period (q50) is 0.58 (= 0.27 + 0.31).

The results have three main features. First, there are no statistically significant differences among the acquisition regimes in the middle of the distribution, in particular at q40, q50, q60, and q70. This is consistent with the results reported in the main text.

Second, there are statistically significant differences in the 8th and 9th deciles. For q80, both the McNamara-Clifford period and the Acquisition Reform (AR) period have significantly higher average APUC growth than does the DSARC period. For q90, both AR and the bust portion of the DAB Post-AR period (FY 2001–FY 2002) are significantly higher than the DSARC, but the McNamara-Clifford period is not. The results reported in the main text get the opposite result—significantly higher APUC growth in the McNamara-Clifford years and ambiguous results for AR and the DAB Post-AR periods.

Third, there is some evidence in the decile regression of significant differences in q20 and q30. In particular, in q20 both McNamara-Clifford and DAB Post-AR are significantly higher than the DSARC period and in q30 DAB Post-AR is significantly higher. These results are only hints, however, since for both q20 and q30 the 95 percent confidence interval on the constant term is wide.

Table D-1. Results of the Quantile Regression

VARIABLES	q10	q20	q30	q40	q50	q60	q70	q80	q90
R_1	0.0900	0.210**	0.180	0.150	0.270	0.160	0.600	0.540	1.070
	(0.101)	(0.0987)	(0.139)	(0.160)	(0.234)	(0.301)	(0.514)	(0.865)	(1.398)
R_3	0.0500	-0.01000	-0.01000	-0.160	-0.01000	-0.110	-0.230	0.420	0.0900
	(0.136)	(0.0915)	(0.123)	(0.168)	(0.289)	(0.345)	(0.466)	(0.445)	(0.319)
R_4	0.130	0.0900**	0.0300	-0.110	-0.140	0.0900	0.01000	-0.0200	-0.120
	(0.178)	(0.0445)	(0.126)	(0.168)	(0.197)	(0.193)	(0.162)	(0.193)	(0.254)
R_5	-0.0900	-0.0500	0.0700	0.0400	0.0400	0.0700	0.0900	1.140*	1.600*
	(0.0847)	(0.0944)	(0.137)	(0.146)	(0.159)	(0.377)	(0.607)	(0.645)	(0.823)
R_6	0.180	0.420***	0.350**	0.210	0.150	0.0300	0.0500	-0.0700	3.780*
	(0.175)	(0.159)	(0.152)	(0.177)	(0.998)	(0.970)	(1.785)	(2.173)	(2.181)
Constant	-0.0900***	-0.01000	0.0600	0.210**	0.310***	0.430***	0.550***	0.670***	1.00***
	(0.0295)	(0.0379)	(0.0591)	(0.104)	(0.0876)	(0.0850)	(0.100)	(0.0926)	(0.160)

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

R_1 McNamara-Clifford FY 1964–FY 1969

R_3 Post-Carlucci DSARC (bust phase) FY 1987–FY 1987

R_4 DAB FY 1990–FY 1993

R_5 Acquisitions Reform (AR) (FY 1994–FY 2000)

R_6 DAB Post AR (bust phase) FY 2001–FY 2003

Appendix E. Statistical Analyses for Boom Climates

This appendix first briefly reports the results of statistical tests of the association between acquisition regime and quantity normalized APUC growth during boom climates. It then turns to comparison of APUC growth for a given acquisition regime between the two funding climates.

A. Acquisition Regime and APUC Growth in Boom Climates

The relevant summary data are presented in Table E-1. This corresponds to the similar table for bust periods in the main body of this paper (Table 2, p. 9). Analysis of Variance (ANOVA) accepted the null hypothesis that the observations for the three regimes are drawn from the same normal distribution (P=0.58). K-S, however, indicated that the observations for FY 1983–FY 1986 and probably those for FY 2003–FY 2007 were not drawn from a normal distribution. Consequentially, the Mann-Whitney Test was computed for each of the pair-wise comparisons of the three boom period-acquisition regime combinations. In each case, the Mann-Whitney test failed to reject the null hypothesis that the two samples being compared were from identical distributions. The conclusion drawn from these results is that changes in acquisition policy and process from one boom period to the next are not associated with statistically significant changes in average APUC growth.

Table E-1. Average APUC Growth by Acquisition Regime and the Number of High Cost Growth MDAPs in Each Cohort, Boom Climates

Acquisition Regime	Period (FY)	Average APUC Growth*	≥11%	≥22%	≥ 45%
DSARC	1981–1982	20% (7)	4	3	2
Post Carlucci DSARC	1983–1986	15% (31)	13	9	5
DAB post-AR	2003–2007	7% (17)	7	2	1

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

* Normalized for changes in quantity.

Because of the results obtained for the bust climates it is appropriate also to ask if there are significant differences across the three boom climate acquisition regimes in the right tail of the distributions. The breakpoints used in Table E-1 were chosen to be about at the same relative point of the right tail of the distribution of APUC growth estimates for the boom climates as the break points used for the bust climates. Recall that the mean APUC growth for the bust climates is 57.3 percent and the first break point used was 50 percent, which is approximately 87 percent of 57.3 percent. The mean APUC growth for the boom periods is 13 percent, so the lowest breakpoint used for the boom periods is 11 percent (= 0.87x13 percent). The middle interval starts at twice the lowest breakpoint (11 percent and 22 percent vice 50 percent and 100 percent), and the higher is the mean plus the sample standard deviation (13 percent plus 31 percent, rounded to 45 percent.)

Fisher's Exact Test was applied for the data in Table E-1. The results failed to reject the null hypothesis of no dependence on the proportion of high cost observations on the funding climate-acquisition regime. Loosely, Fisher's Exact Test provides no indication of differences among the right-hand tails of the distributions of APUC growth across acquisition regimes in boom funding climates. This is the opposite of the result found for bust climates, but we have no observations in a boom funding climate for the two acquisition regimes with a significantly higher proportion of extremely high cost growth programs—McNamara-Clifford and AR.

B. Association of Funding Climate and APUC Growth

The second question addressed in this appendix is whether the mean APUC growth of MDAPs that passed MS II/B in bust periods is significantly higher than that of MDAPs that passed MS II/B in boom periods. The relevant mean APUC growth figures are presented in Table E-2.

Table E-2. Average APUC Growth by Acquisition Regime and Funding Climate

	Relatively Constrained Funding Climate (Bust)		Relatively Accommodating Funding Climate (Boom)	
Acquisition Regime	**Period**	**APUC Growth**	**Period**	**APUC Growth**
McNamara-Clifford	1964–1969	90% (16)	**Not observed**	
DSARC	1970–1980	36% (49)	1981–1982	20% (7)
Post Carlucci DSARC	1987–1989	42% (10)	1983–1986	15% (31)
DAB	1990–1993	35% (11)	**Not observed**	
Acquisition Reform	1994–2000	78% (27)		
DAB post-AR	2001–2002	113% (6)	2003–2007	7% (17)

The differences between the relevant row entries are significant for a simple t-test: for 1983–1986 versus 1987–1989 (P=0.095) and for 2003–2007 versus 2001–2002 (P=0.021). The first of these results is suspect because the distribution of APUC growth rates for 1983–1986 is non-normal. The second is suspect because the sample variance for the 2001–2002 observations was much larger than that for 2003–2007. The results of the t-test was confirmed by the Mann-Whitney test for FY 2003–FY 2007 versus FY 2001–FY 2002 (P=0.006), but not for FY 1983–FY 1986 versus FY 1987–FY 1989. The difference between the average APUC growth in FY 1970–FY 1980 and FY 1981–FY 1982, although substantial (36 percent versus 20 percent), was not significant by either the t-test or the Mann-Whitney test.

In view of the results obtained for the bust climates, it is also necessary to ask whether the right tails of the bust distributions are "fatter" than the right tails of the corresponding boom climate distributions. The data we used to examine this question are presented in Table E-3.

Table E-3. Number of Observations in Three Segments of the Right Tail in Bust and Boom Periods for Three Acquisition Regimes

Acquisition Regime	Period	Average APUC Growth	A	B	C
DSARC	1970–1980	36% (49)	12	5	0
	1981-1982	20% (7)	1	1	2
Post Carlucci DSARC	1987–1989	42% (10)	1	2	1
	1983–1986	15% (31)	4	4	5
DAB post AR	2001–2002	113% (6)	1	0	1
	2003–2007	7% (17)	5	1	1

Bust Climate **Boom Climate**

A: Bust Climate 50% ≤ x < 100% Boom Climate 11% ≤ x < 22%

B: Bust Climate 100% ≤ x < 143% Boom Climate 22% ≤ x < 50%

C: Bust Climate ≥ 143% Boom Climate ≥ 50%

Fisher's Exact Test found no differences for any of the segments in comparing the two climates for the DAB post AR and the Post-Carlucci DSARC periods. This is the case shown in Figure E-1: the shape of the distributions remains the same in the two climates, but the mean is shifted to the right in the bust climate. Fisher's Exact Test did find a statistically significant difference for the DSARC period (P = 0.001) for segment C, but not for the other two segments. In this case, in the bust period, not only was the mean higher, but the right tail was significantly fatter relative to the distribution for the boom period.

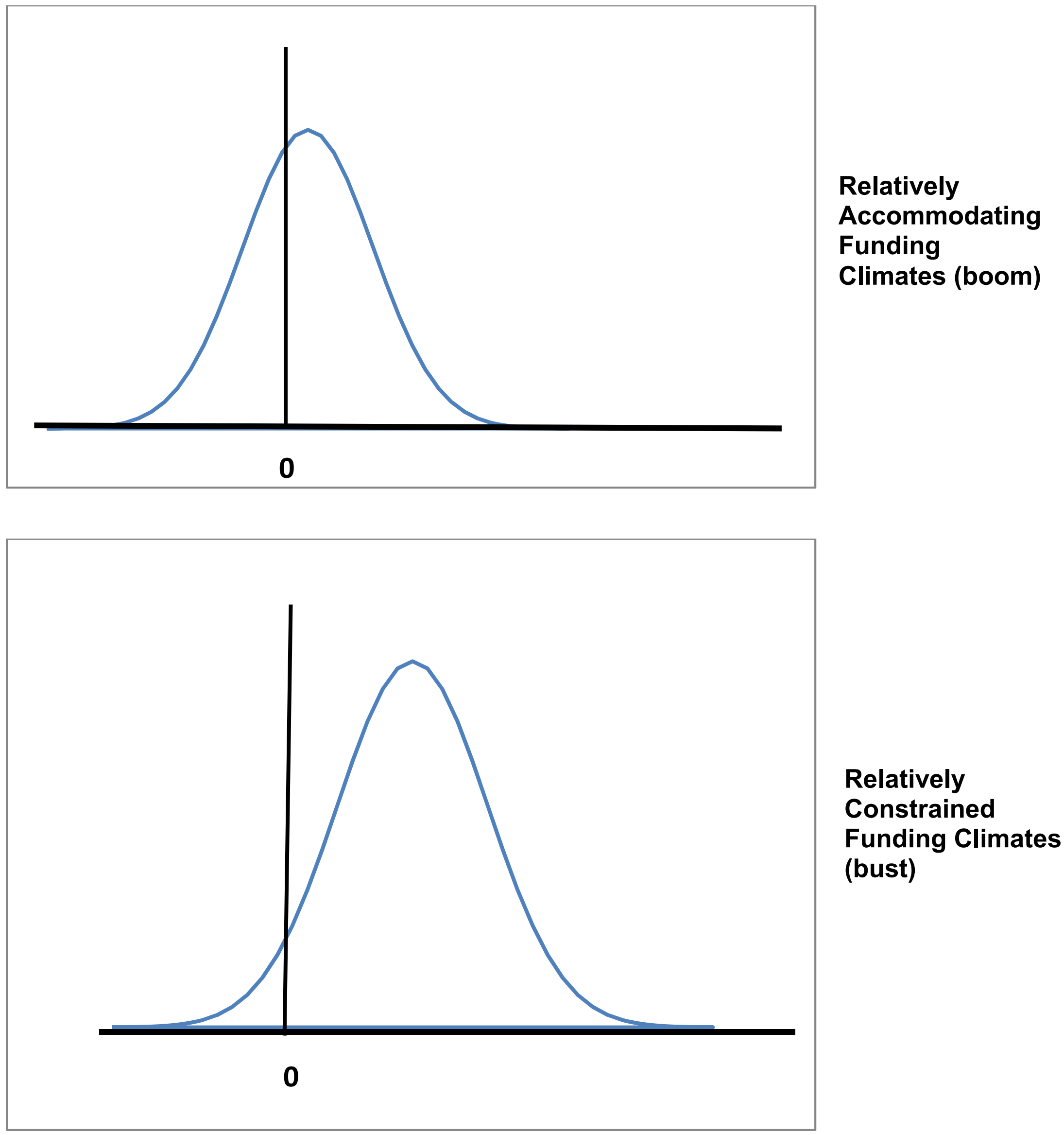

Figure E-1. Notional Shift in the Distribution of APUC Growth between Funding Climates

Illustrations

Figures

Tables

References

Adams, Gordon, Paul Murphy, and William Grey Rosenau. *Controlling Weapons Costs: Can Pentagon Reforms Work?* New York: Council on Economic Priorities, 1983.

Borklund, C. W. *The Department of Defense*. New York: Frederick A. Praeger, 1969.

Fox, J. Ronald. *Defense Acquisition Reform, 1969 to 2009: An Elusive Goal*. Washington, DC: U.S. Army Center of Military History, 2011.

Glennan, Thomas K., Jr. "Policies for Military Research and Development." RAND Paper P-3253. Santa Monica CA: The RAND Corporation, November 1965.

Jarvaise, Jeanne M., Jeffrey A. Drezner, and Dan Norton. "The Defense System Cost Performance Database: Cost Growth Using Selected Acquisition Reports." RAND Report MR-625-OSD. Santa Monica, CA: The RAND Corporation, 1996.

McNicol, David L. *Cost Growth in Major Weapon Procurement* Programs, 2nd edition. Alexandria, VA: Institute for Defense Analyses, 2004.

McNicol, David L., Sarah K. Burns, and Linda Wu. "Evidence on the Effect of DoD Acquisition Policy and Process and Funding Climate on Cancellation of Major Defense Acquisition Programs." IDA Paper P-5218. Alexandria, VA: Institute for Defense Analyses, May 2015.

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

Murdock, Clark A. *Defense Policy Formation*. Albany, NY: State University of New York Press, 1974.

O'Neil, William D., and Gene H. Porter. "What to Buy? The Role of Director of Defense Research and Engineering (DDR&E)—Lessons from the 1970s." IDA Paper P-4675. Alexandria, VA: Institute for Defense Analyses, January 2011.

Srull, Donald, ed. *The Cost Analysis Improvement Group: A History*. McLean, VA: Logistics Management Institute (now LMI), 1998.

Tyson, Karen W., Neang I. Om, D. Calvin Gogerty, and J. Richard Nelson. "The Effects of Management Initiatives on the Costs and Schedules of Defense Acquisition Programs, Vol. I: Main Report." IDA Paper P-2722. Alexandria, VA: Institute for Defense Analyses, 1992.

United States Department of Defense, Office of the Under Secretary of Defense (Comptroller). *National Defense Budget Estimates for FY 2014*. Washington, DC: Department of Defense, May 2013.

Abbreviations

ACAT	Acquisition Category
A-D	Anderson-Darling
AHIP	Army Helicopter Improvement Program
ANOVA	Analysis of Variance
APUC	Average Procurement Unit Cost
AR	Acquisition Reform
CAIG	Cost Analysis Improvement Group
CAPE	Cost Assessment and Program Evaluation
CD	Compact Disc
CE	Current Estimate
CLC	Calibrated Learning Curve
CPIF	Cost Plus Incentive Fee
DAB	Defense Acquisition Board
DAMIRS	Defense Acquisition Management Information Retrieval System
DCP	Development Concept Paper/Design Coordinating Paper
DDR&E	Director, Defense Research and Engineering
DLC	DAMIRS Learning Curve
DoD	Department of Defense
DoDD	Department of Defense Directive
DoDI	Department of Defense Instruction
DSARC	Defense Systems Acquisition Review Council
EMD	Engineering and Manufacturing Development
FET	Fisher's Exact Test
FFP	Firm Fixed Price
FPIF	Fixed Price Incentive Fee
FY	Fiscal Year
IDA	Institute for Defense Analyses
JDAM	Joint Direct Attack Munition
K-S	Kolmogorov-Smirnov
MAIS	Major Automated Information System

MDA	Milestone Decision Authority
MDAP	Major Defense Acquisition Program
MS	Milestone
NQC	No Quantity Change
OASD(SA)	Office of the Assistant Secretary of Defense for Systems Analysis
OSD	Office of the Secretary of Defense
PA&E	Program Analysis and Evaluation
PAUC	Program Acquisition Unit Cost
PPBS	Planning, Programming, and Budgeting System
RDT&E	Research, Development, Test and Evaluation
SAR	Selected Acquisition Report
TPP	Total Package Procurement
US	United States
USD(AT&L)	Under Secretary of Defense (Acquisition, Technology and Logistics)
VCJCS	Vice Chairman, Joint Chiefs of Staff
WSARA	Weapon Systems Acquisition Reform Act of 2009

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
xx-08-2016	Final	Sept 2015 - May 2016, Aug 2016

4. TITLE AND SUBTITLE
Further Evidence on the Effect of Acquisition Policy and Process on Cost Growth of Major Defense Acquisition Programs (Revised)

5a. CONTRACT NUMBER
HQ0034-14-D-0001

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
McNicol, David, L.
Tate, David, M.
Burns, Sarah, K.
Wu, Linda

5d. PROJECT NUMBER

5e. TASK NUMBER
AY-7-405202

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses
4850 Mark Center Drive
Alexandria, VA 22311-1882

8. PERFORMING ORGANIZATION REPORT NUMBER
IDA Paper P-5330, Revised
Log #: H 16-000168 (Main Paper)
Log #: H 16-000173 (Support Data CD)

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Director, Performance Assessments and Root Cause Analyses
Office of the Assistant Secretary of Defense for Acquisition
Room 3C889A
3620 Defense Pentagon
Washington, DC 20301-3620

10. SPONSOR/MONITOR'S ACRONYM(S)
D,PARCA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Supersedes IDA Paper P-5330, June 2016.

14. ABSTRACT
Institute for Defense Analyses Paper P-5126 found that additional acquisition reforms after those introduced in mid-1969 by then Deputy Secretary of Defense David Packard did not significantly reduce cost growth on Major Defense Acquisition Programs (MDAPs). While a useful insight into the effects of changes in acquisition policy and process, that conclusion leaves open the possibility that the Packard reforms reduced cost growth compared to the record of the 1960s, which is the issue examined in this paper. The paper finds that average cost growth of MDAPs that entered Engineering and Manufacturing Development during Fiscal Year (FY) 1970–FY 1980 was significantly lower than the average of those that entered during FY 1964–FY 1969. Further statistical analysis also suggests that it was significantly lower than the average during FY 1994–FY 2000 when Office of the Secretary of Defense (OSD)-level oversight of MDAPs was less stringent. These stand as instances of a significant association between changes in OSD-level oversight and cost growth. The paper also provides evidence that average cost growth in FY 1964–FY 1969 and FY 1994–FY 2000 was particularly high largely because the proportion of MDAPs that experienced extremely high cost growth was significantly larger than it was in other periods.

15. SUBJECT TERMS
Acquisition reform, acquisition policy, acquisition process, cancellation, termination, cost growth, APUC growth, funding climate

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Bliss, Gary, R.
Unclassified	Unclassified	Unclassified	Same as Report	53	**19b. TELEPHONE NUMBER** *(Include area code)* (571) 256-0646

Reset