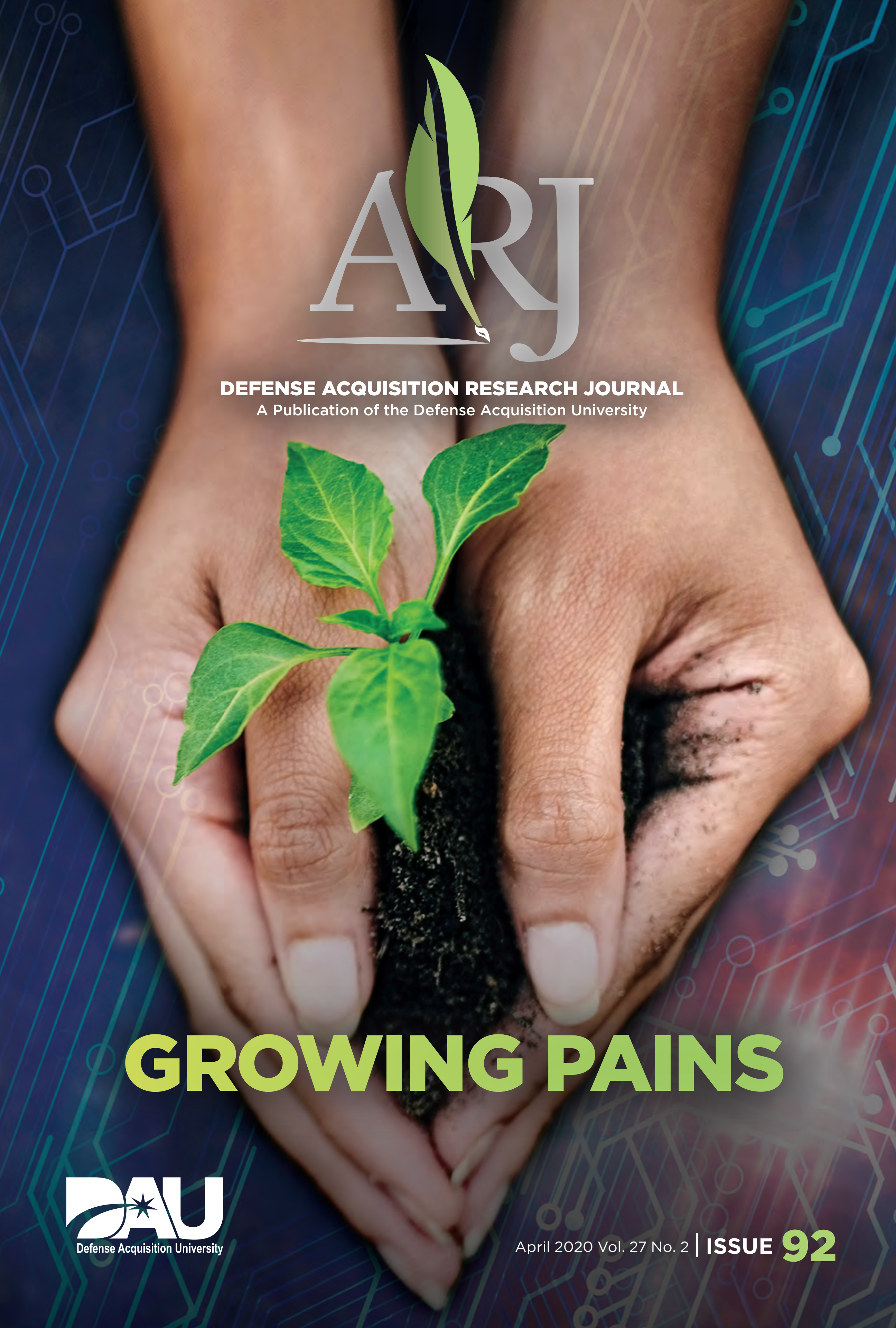


ARJ

DEFENSE ACQUISITION RESEARCH JOURNAL

A Publication of the Defense Acquisition University

GROWING PAINS

Defense Acquisition University

April 2020 Vol. 27 No. 2 | ISSUE 92

Software Productivity Trends and Issues

David M. Tate

Analysis of Military Construction Cost Growth in USAF Major Defense Acquisition Programs

Capt Emily E. Angell, USAF, Edward D. White, Jonathan D. Ritschel, and Alfred E. Thal, Jr.

Inflation and Price Escalation Adjustments in Estimating Program Costs: F-35 Case Study

Stanley A. Horowitz and Bruce R. Harmon

ARTICLE LIST

ARJ EXTRA

The Defense Acquisition Professional Reading List

To Provide and Maintain a Navy: 1775–1945

Written by CAPT Richard L. Wright, USN (Ret.)

Reviewed by Brad Martin, Senior Policy Researcher, RAND Corporation

DEFENSE ACQUISITION RESEARCH JOURNAL

A Publication of the Defense Acquisition University

ISSN 2156-8391 (print) ISSN 2156-8405 (online)
DOI: https://doi.org/10.22594/dau.042020-92.27.02

The *Defense Acquisition Research Journal*, formerly the *Defense Acquisition Review Journal*, is published quarterly by the Defense Acquisition University (DAU) Press and is an official publication of the Department of Defense. Postage is paid at the U.S. Postal facility, Fort Belvoir, VA, and at additional U.S. Postal facilities. Postmaster, send address changes to: Editor, *Defense Acquisition Research Journal*, DAU Press, 9820 Belvoir Road, Suite 3, Fort Belvoir, VA 22060-5565. The journal-level DOI is: https://doi.org/10.22594/dauARJ.issn.2156-8391. Some photos appearing in this publication may be digitally enhanced.

DEFENSE ACQUISITION RESEARCH JOURNAL

A Publication of the Defense Acquisition University

Managing Editor,
Chief of Visual Arts & Press
Norene L. Johnson

Assistant Editor
Emily Beliles

Art Director
Michael Bubar-Krukowski

Production Manager
Frances Battle

Graphic Designer,
Digital Publications
Nina Austin

Technical Editor
Collie J. Johnson

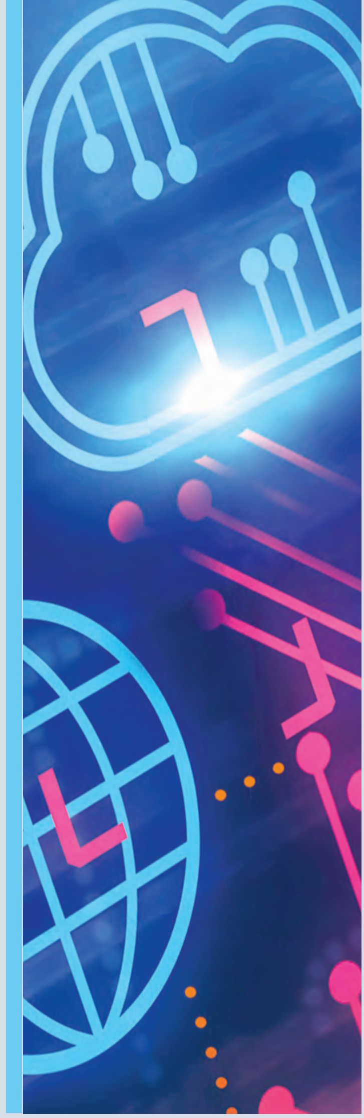
Associate Editor
Michael Shoemaker

Copy Editor,
Circulation Manager
Debbie Gonzalez

Multimedia Assistant
Noelia Gamboa

Editing, Design, and Layout
Chickasaw Nation Industries
The C3 Group

CONTENTS | Featured Research



194

Inflation and Price Escalation Adjustments in Estimating Program Costs: F-35 Case Study

Stanley A. Horowitz and Bruce R. Harmon

This article illustrates the importance of basing estimates of future program prices on historical price increases of similar systems. In the case of tactical aircraft, using general inflation rates to predict future aircraft prices is likely to lead to serious underestimates of future cost.

CONTENTS | Featured Research

FROM THE CHAIRMAN AND EXECUTIVE EDITOR

Dr. Larrie D. Ferreiro

The theme of this edition of the *Defense Acquisition Research Journal* is "Growing Pains," an apt description of the problems that many defense acquisition programs face with cost and schedule.

The first research paper in this issue, "Software Productivity Trends and Issues" by David M. Tate, examines recent trends in defense software supply, demand, and productivity to estimate the severity of capacity bottleneck, then briefly discusses potential long-term actions available to the Department of Defense to mitigate that bottleneck. In the second article, "Analysis of Military Construction Cost Growth in USAF Major Defense Acquisition Programs," the authors Capt Emily E. Angell, USAF, Edward D. White, Jonathan D. Ritschel, and Alfred E. Thal, Jr., use descriptive and inferential statistics to identify cost growth of military construction at the programmatic level, and describe how they change over time. The findings of this study may help determine allocation of resources in developing cost estimates. The third article is "Inflation and Price Escalation Adjustments in Estimating Program Costs: F-35 Case Study" by Stanley A. Horowitz and Bruce R. Harmon. It illustrates the importance of basing estimates of future program prices on historical price increases of similar systems. The authors caution that in the case of tactical aircraft, using general inflation rates to predict future aircraft prices is likely to lead to serious underestimates of future cost.

This issue's Current Research Resources in Defense Acquisition focuses on Acquisition Reform. It contains descriptions of several key resources, along with links to the DAU Knowledge Repository sites.

The featured reading in this issue's Defense Acquisition Professional Reading List is *To Provide and Maintain a Navy: 1775–1945* by CAPT Richard L. Wright, USN (Ret.), and reviewed by Brad Martin.

Dr. Larrie D. Ferreiro
Chairman and Executive Editor
Defense ARJ

From the Assistant Editor

Emily Beliles

The Managing Editor asked me to share a few words of advice to prospective authors who hope to get published in the *Defense Acquisition Research Journal (ARJ)*. While I am not a subject matter expert by any means, I do have some experience with what types of articles fare well in the review process.

Firstly, take a look at the Research Agenda—this issue includes quite a few updates. See whether your article fits within the purview of the *Defense ARJ*. You might want to reach out to one of the POCs listed on the Defense Acquisition University Research website and ask for their feedback on your article before you submit it.

Secondly, carefully read our Guidelines for Contributors, which is included at the end of every issue and available on our website as well. Ask yourself whether your article meets the requirements. Does it qualify as research? Does it involve the creation of new knowledge? Is it relevant to the Defense acquisition workforce? Does it meet all the critical characteristics of empirical research articles?

Finally, take care to submit all of the required documents listed on the submission checklist posted on our website. This includes ensuring that your article is formatted correctly and that your citations and references comply with the *Publication Manual of the American Psychological Association* style guide (6th edition). Even small errors can slow down your submission process and unnecessarily delay your article from being processed.

Please feel free to reach out if you have any further questions about prospective articles. I wish you the best of luck in getting your work published and hope to see your article come across my desk soon.

DAU CENTER FOR DEFENSE ACQUISITION

RESEARCH AGENDA 2020

This Research Agenda is intended to make researchers aware of the topics that are, or should be, of particular concern to the broad defense acquisition community in the government, academic, and industrial sectors. It is compiled using inputs from Subject Matter Experts (SMEs) across those sectors. These topics are periodically vetted and updated as needed to ensure they address current areas of strategic interest.

The purpose of conducting research in these areas is to provide solid, empirically based findings to create a broad body of knowledge that can inform the development of policies, procedures, and processes in defense acquisition, and to help shape the thought leadership for the acquisition community. These research topics should be considered guidelines to help investigators form their own research questions. Some questions may cross topics and thus appear in multiple research areas.

Potential researchers are encouraged to contact the DAU Director of Research (research@dau.edu) to suggest additional research questions and topics. They are also encouraged to contact the Point(s) of Contact (POC), who may be able to provide general guidance as to current areas of interest, potential sources of information, etc. Contact information for the POCs is available on the DAU Research website at ***https://www.dau.edu/library/research/p/Research-Areas***.

Affordability and Cost Growth

- Define or bound “affordability” in the defense portfolio. What is it? How will we know if something is affordable or unaffordable?
- What means are there (or can be developed) to measure, manage, and control “affordability” at the Program Office level? At the industry level? How do we determine its effectiveness?
- What means are there (or can be developed) to measure, manage, and control “Should Cost” estimates at the Service, Component, Program Executive, Program Office, and industry levels? How do we determine their effectiveness?
- What means are there (or can be developed) to evaluate and compare incentives for achieving “Should Cost” at the Service, Component, Program Executive, Program Office, and industry levels?
- Recent acquisition studies have noted the vast number of programs and projects that don’t successfully make it through the acquisition system and are subsequently cancelled. What would systematic root cause analyses reveal about the underlying reasons, whether and how these cancellations are detrimental, and what acquisition leaders might do to rectify problems?
- Do joint programs—at the inter-Service and international levels—result in cost growth or cost savings compared with single-Service (or single-nation) acquisition? What are the specific mechanisms for cost savings or growth at each stage of acquisition? Do the data support “jointness” across the board, or only at specific stages of a program, e.g., only at Research and Development (R&D), or only with specific aspects, e.g., critical systems or logistics?
- Can we compare systems with significantly increased capability developed in the commercial market to Department of Defense (DoD)-developed systems of similar characteristics?
- Is there a misalignment between industry and government priorities that causes the cost of such systems to grow significantly faster than inflation?
- If so, can we identify why this misalignment arises? What relationship (if any) does it have to industry's required focus on shareholder value and/or profit, versus the government's charter to deliver specific capabilities for the least total ownership costs?

Industrial Productivity and Innovation

Industry insight and oversight

- What means are there (or can be developed) to measure the level of insight and/or control that government has over subcontractors?
- What means are there (or can be developed) to measure costs of enforcement (e.g., auditors) versus actual savings from enforcement?
- What means are there (or can be developed) to evaluate and compare incentives for subcontractor/supply chain competition and efficiencies?
- What means are there (or can be developed) to evaluate and compare market-based incentives with regulatory incentives?
- How can we perform institutional analyses of the behaviors of acquisition organizations that incentivize productivity?
- What means are there (or can be developed) to evaluate and compare the barriers of entry for SMEs in defense acquisition versus other industrial sectors?

- Is there a way to measure how and where market incentives are more effective than regulation, and vice versa?
- Do we have (or can we develop) methods to measure the effect of government requirements on increased overhead costs, at both government and industrial levels?
- Examine the possibilities to rationalize and balance the portfolio of capabilities through buying larger quantities of common systems/subsystems/components across Defense Agencies and Services. Are there examples from commercial procurement and international defense acquisition that have produced positive outcomes?
- Can principal-agent theory be used to analyze defense procurement realities? How?
- What means are there (or can be developed) to measure the effect on defense acquisition costs of maintaining the industrial base in various sectors?
- What means are there (or can be developed) of measuring the effect of utilizing defense industrial infrastructure for commercial manufacture, in particular in growth industries? In other words, can we measure the effect of using defense manufacturing to expand the buyer base?
- What means are there (or can be developed) to measure the breadth and depth of the industrial base in various sectors that go beyond a simple head count of providers?
- Has change in the industrial base resulted in actual change in output? How is that measured?

Independent Research and Development

- What means do we require to measure the cost-effectiveness or Return on Investment (ROI) for DoD-reimbursed Independent Research and Development (IR&D)?
- Can we properly account for sales and revenues that are products of IR&D?
- Can we properly account for the barriers to entry for SMEs in terms of IR&D?
- Examine industry trends in IR&D, for example, percentage of revenue devoted to IR&D, collaboration with academia. How do they vary by industry sector, in particular, those associated with defense acquisition?
- What means are there (or can be developed) to measure the ROI for DoD-reimbursed IR&D versus directly funded defense R&D?
- What incentive structures are required to get industry to focus on and fund disruptive technologies?
- What has been the impact of IR&D on developing disruptive technologies?

Competition

Measuring the effects of competition

- What means are there (or can be developed) to measure the effect on defense acquisition costs of maintaining an industrial base in various sectors?
- What means are there (or can be developed) of measuring the effect of utilizing defense industrial infrastructure for commercial manufacture, in particular in growth industries? In other words, can we measure the effect of using defense manufacturing to expand the buyer base?

- What means are there (or can be developed) to determine the degree of openness that exists in competitive awards?
- What are the different effects of the two best value source selection processes (tradeoff versus lowest price technically acceptable) on program cost, schedule, and performance?

Strategic competition

- Is there evidence that competition between system portfolios is an effective means of controlling price and costs?
- Does lack of competition automatically mean higher prices? For example, is there evidence that sole source can result in lower overall administrative costs at both the government and industry levels, to the effect of lowering total costs?
- What are long-term historical trends for competition guidance and practice in defense acquisition policies and practices?
- To what extent are contracts being awarded noncompetitively by congressional mandate, for policy interest reasons? What is the effect on contract price and performance?
- What means are there (or can be developed) to determine the degree to which competitive program costs are negatively affected by laws and regulations such as the Berry Amendment, Buy American Act, etc.?
- The DoD should have enormous buying power and the ability to influence supplier prices. Is this the case? Examine the potential change in cost performance due to greater centralization of buying organizations or strategies.

Effects of industrial base

- What are the effects on program cost, schedule, and performance of having more or fewer competitors? What measures are there to determine these effects?
- What means are there (or can be developed) to measure the breadth and depth of the industrial base in various sectors, that go beyond a simple head count of providers?
- Has the change in industrial base resulted in actual change in output? How is that measured?

Competitive contracting

- Commercial industry often cultivates long-term, exclusive (noncompetitive) supply chain relationships. Does this model have any application to defense acquisition? Under what conditions/circumstances?
- What is the effect on program cost performance of awards based on varying levels of competition: (a) “Effective Competition” (two or more offers; (b) “Ineffective Competition” (only one offer received in response to competitive solicitation; (c) “Split Awards” versus winner take all; and (d) “Sole Source.”

Improve DoD outreach for technology and products from global markets

- How have militaries in the past benefitted from global technology development?
- How/why have militaries missed the largest technological advances?

- What are the key areas that require DoD focus and attention in the coming years to maintain or enhance the technological advantage of its weapons systems and equipment?
- What types of efforts should DoD consider pursuing to increase the breadth and depth of technology push efforts in DoD acquisition programs?
- How effectively are DoD's global Science and Technology (S&T) investments transitioned into DoD acquisition programs?
- Are managers of DoD's applied R&D (i.e., acquisition program) investments effectively pursuing and using sources of global technology to affordably meet current and future DoD acquisition program requirements? If not, what steps could DoD take to improve its performance in these two areas?
- What are the strengths and weaknesses of DoD's global defense technology investment approach as compared to the approaches used by other nations?
- What are the strengths and weaknesses of DoD's global defense technology investment approach as compared to the approaches used by the private sector—both domestic and foreign entities (companies, universities, private-public partnerships, think tanks, etc.)?
- How does DoD currently assess the relative benefits and risks associated with global versus U.S. sourcing of key technologies used in DoD acquisition programs? How could DoD improve its policies and procedures in this area to enhance the benefits of global technology sourcing while minimizing potential risks?
- How could current DoD/U.S. Government Technology Security and Foreign Disclosure (TSFD) decision-making policies and processes be improved to help DoD better balance the benefits and risks associated with potential global sourcing of key technologies used in current and future DoD acquisition programs?
- How do DoD primes and key subcontractors currently assess the relative benefits and risks associated with global versus U.S. sourcing of key technologies used in DoD acquisition programs? How could they improve their contractor policies and procedures in this area to enhance the benefits of global technology sourcing while minimizing potential risks?
- How could current U.S. Government Export Control system decision-making policies and processes be improved to help DoD better balance the benefits and risks associated with potential global sourcing of key technologies used in current and future DoD acquisition programs?

Comparative studies

- Compare the industrial policies of military acquisition in different nations and the policy impacts on acquisition outcomes.
- Compare the cost and contract performance of highly regulated public utilities with nonregulated "natural monopolies," e.g., military satellites, warship building.
- Compare contracting/competition practices between DoD and complex, custom-built commercial products (e.g., offshore oil platforms).
- Compare program cost performance in various market sectors: highly competitive (multiple offerors), limited (two of three offerors), or monopoly?
- Compare the cost and contract performance of military acquisition programs in nations having single "purple" acquisition organizations with those having Service-level acquisition agencies.

Acquisition of Services

Metrics

- What metrics are currently collected and available on services acquisition:
  - Within the Department of Defense?
  - Within the U.S. Government?
  - Outside of the U.S. Government?
- What and how much do these metrics tell us about services acquisition in general and about the specific programs for which the metrics are collected?
- What are the possible metrics that could be used in evaluating services acquisition programs?
  - How many metrics should be used?
  - What is the efficacy of each metric?
  - What is the predictive power of each metric?
  - What is the interdependence (overlap) between metrics?
- How do we collect data for services acquisition metrics?
  - What is being done with the data that are currently being collected?
  - Are the data that are being collected on services acquisition reliable?
  - Is the collection process affecting the data that are collected for services acquisition?
- How do we measure the impact of different government requirements on overhead costs and rates on services contracts?

Industrial base

- What is the right amount of contracted services for government organizations?
  - What are the parameters that effect Make, Buy decisions in government services?
  - How do the different parameters interact and affect government force management and industry research availability?
- What are the advantages, disadvantages, and impacts of capping pass-through costs, and how do they change with the value of the pass-through costs?
- For Base Operations and Support (BOS) contracts, is there a best size? Should large BOS contracts be broken up? What are the parameters that should be considered?
- In the management of large services contracts, what is the best organization? Is the System Program Office a good model? What parameters should be used in evaluating the advantages and disadvantages of an organization to manage large services contracts?
- What effect does strategic sourcing and category management have on small business if the small business is a strategic source or whether the small business is not a strategic source?
- Do the on-ramping and off-ramping requirements of some service contracts have an effect on the industrial base? If so, what are the impacts?

Industry practices

- What private sector business practices, other than maximizing profit, can the government effectively use to incentivize performance and otherwise improve business relationships with vendors?
- What are the best methods for evaluating different incentives to encourage small businesses to participate in government services contracts?
- What potential benefits can the government achieve from long-term supply chain relationships? What are the disadvantages?
- What benefits does industry get from the use of category managers and functional domain experts, and can the government achieve the same benefits?
- How can the government best capture, validate, and use demand management strategies?
- Are current services acquisition taxonomies comprehensive, or can they be improved?

Make/buy

- What methods can best be used to define the cost value relationship in different classes of service contracts?
- Can we develop a method for determining the “should cost” of different services?
- Can we define and bound affordability of specific services?
- What are the characteristics of “inherently governmental” activities, and how can we evaluate the value of these services based on comparable characteristics in a competitive labor market?
- In services contracts, what are the inherent life-cycle costs, and how do we capture the life-cycle costs in make/buy decision making?
- In the case of government services contracting, what are the factors that contribute to less-than-optimum make/buy decision making?

Category management/strategic sourcing

- What effect does strategic sourcing/category management have on competition?
    - Effects on short term versus long term.
    - Effects on competition outside of the strategic sourcing/category management area of consideration.
- What metrics do different industries use for measuring the effectiveness of their supply chain management?
- Would the centralization of services acquisition contracts have measurable impacts on cost performance? Why or why not?
- What are the fundamental differences between the services taxonomy and the category management taxonomy, and are there means and good reasons to align the two taxonomies?

Contract management/efficacy

- What are the best ways to address the services parts of contracts that include both services and products (goods)?
- In the management of services contracts, what are the non-value-added tasks, and are there realistic ways to reduce the impact of these tasks on our process?

- When funds for services are provided via pass-throughs (i.e., from another organization), how are the requirements tracked, validated, and reviewed?
- Do Undefinitized Contract Actions have an effect on contractor pricing and willingness, or lack of willingness to provide support during proposal analysis?
- For multi-award, Indefinite-Delivery, Indefinite-Quantity (IDIQ)-type contracts, is there a method for optimizing the different characteristics (number of vendors, timelines, on-ramping, off-ramping, etc.) of these contracts?

Policy

- What current government policies inhibit alignment of contractors' approaches from aligning with the government's services acquisition programs?

Administrative Processes

- What means are there (or can be developed) to measure the efficiency and effectiveness of DoD oversight, at the Component, Service, and Office of the Secretary of Defense levels?
- What measures are there (or can be developed) to evaluate and compare the costs of oversight versus the cost savings from improved processes?
- What means are there (or can be developed) to empirically establish oversight process metrics as a basis for comparison? Can these be used to establish the relationship of oversight to cost/schedule/performance outcomes?
- What means are there (or can be developed) to study the organizational and governance frameworks, resulting in successful change management?
- To what extent (investment and performance) can scenario/simulations testing improve the delivery of complex projects?
- Is there a comparative statistical divergence between organizational honesty (reality) and contractual relationships (intent) in tendering?
- How does one formulate relational contracting frameworks to better account and manage risk and liability in a collaborative environment?

Human Capital of Acquisition Workforce

- What means are there (or can be developed) to measure ROI for acquisition workforce training?
- What elements of the Professional Military Education framework can be applied to the professionalization of the civilian defense acquisition workforce?
- What factors contribute to the management and successful delivery of modern complex project management, including performance over the project life cycle?
- What behavioral leadership characteristics can be commonly observed in successful complex projects, contrasted against unsuccessful complex projects?
- What is the functional role of talent management in building organizational sustainability, performance, and leadership?
- How do we create incentives in the acquisition workforce (management, career, social, organizational) that provide real cost reductions?

Defense Business Systems

Organizational structure and culture in support of Agile software development methodologies

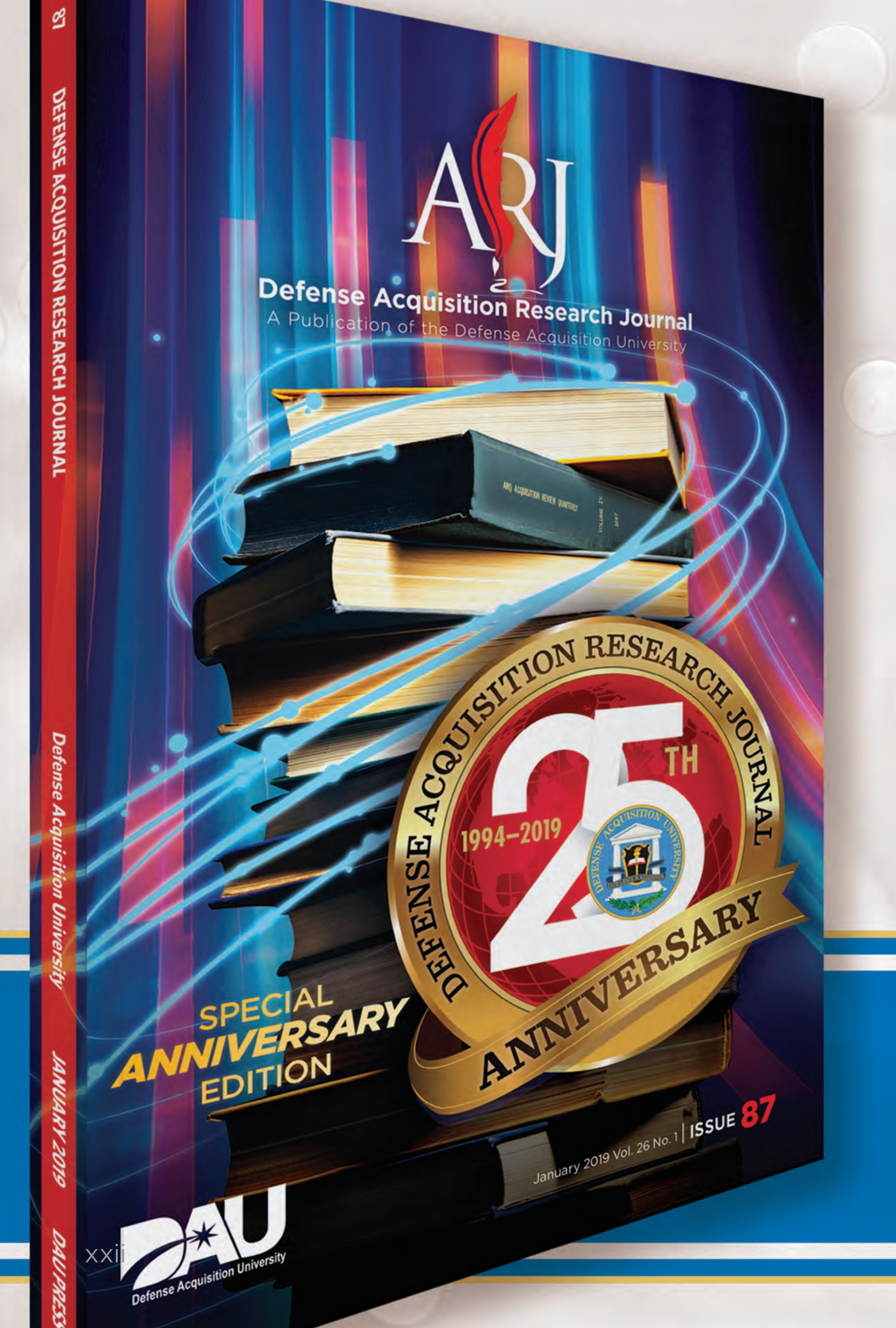
- At the beginning of the Business Capability Acquisition Cycle (BCAC) process, various steps are used to ensure accurate requirements are thoroughly documented and supported throughout the software development life cycle. How can these documentation requirements and processes be streamlined to support more direct-line communication between the end-user and software engineers? What are the hurdles to implement these changes and how are they overcome? What are the effects of these changes on the organization or agency?
- Regarding new starts, how can the BCAC be modified specifically to support Agile development? How are these changes advantageous or disadvantageous to the customer and organization? Would these changes be helpful or detrimental with R&D versus a concurrent design and engineering software project?
- Generally, readiness review briefings within the BCAC are used to determine if a project is at an acceptable state to go to the next step in the process. If software is developed and released to production within a single Sprint (potentially every 2 weeks), how are Test Readiness Reviews, Systems Requirements Reviews, and Production Readiness Reviews handled? How have the changes made to these events made them more or less relevant?
- How are organizations and agencies structured to support concurrent software design and development? What organizational structure would support R&D and non-R&D Information Technology (IT) capabilities?
- What steps are used to choose Agile as the default software development process versus any other software development methodology (e.g., Waterfall, Spiral, or Incremental) for your organization? What are the effects on project cost, schedule, and performance?
- Within DoD agencies and military branches, has the adaption of Agile resulted in faster deployment of new IT capabilities to the customer? How is this determined and measured?
- Industry often produces software using Agile. The DoD's BCAC process can be a process that produces an abundance of bureaucracy counter to Agile principles. How does hiring a contractor to implement or maintain IT capabilities and introducing Agile software development methods within a BCAC non-Agile process create conflict? How are these conflicts resolved or reconciled?
- How is IT engineering investment and innovation supported throughout DoD? What organizational or cultural aspects of an agency are specific to that support?

Defense Acquisition and Society

- To what extent should the DoD use the defense acquisition process to effectuate various social policies? The existing procurement regime favors a dizzying array of private interests ranging from organized labor; domestic manufacturers and firms located in areas of high unemployment; small businesses, including disadvantaged and women-owned firms; blind, severely handicapped, and prison industries; and, most recently, environmentally friendly vendors. Affirmatively steering the government's business from the open marketplace to preferred providers adds complexity, thus increasing transaction costs throughout the procurement process, which absorbs scarce resources. (Source: IBM Center for the Business of Government, http://www.businessofgovernment.org)

- How significant are the transaction costs resulting from the administration's commitment to transparency (generally, and specifically in the context of stimulus or recovery spending)? In a representative democracy, transparency is critical. But transparency is expensive and time-consuming, and the additional resources required to comply with the recently enhanced disclosure standards remain an unfunded mandate. Thus, the existing acquisition workforce must devote scarce resources to an (admittedly legitimate) end other than the pursuit of value for money or customer satisfaction. Is there an optimal balance or a point of diminishing returns? In other words, at what point does the cost of developing transparent systems and measures exceed the benefits of that transparency? (Source: IBM Center for the Business of Government, http://www.businessofgovernment.org)

Potential authors are encouraged to peruse the DAU Research website (***https://www.dau.edu/library/research/p/Research-Areas***) for information on contacting the POC for each content area.


ARJ
Defense Acquisition Research Journal
A Publication of the Defense Acquisition University
DEFENSE ACQUISITION RESEARCH JOURNAL
25TH
1994–2019
ANNIVERSARY
SPECIAL
ANNIVERSARY
EDITION
January 2019 Vol. 26 No. 1 | ISSUE 87
DAU
Defense Acquisition University
81
DEFENSE ACQUISITION RESEARCH JOURNAL
Defense Acquisition University
JANUARY 2019

2019 PLATINUM WINNER

MARCOM
AWARD FOR
PEER REVIEWED
ACADEMIC JOURNAL

HERMES
AWARD FOR
PRINT MEDIA | PUBLICATION

APEX®
AWARD FOR
PUBLICATION EXCELLENCE

Norene Johnson, Emily Beliles,
and Michael Bubar-Krukowski

Defense Acquisition University Press

Fort Belvoir, VA

2019
APEX®
AWARDS FOR
PUBLICATION EXCELLENCE

We’re on the Web at:
http://www.dau.edu/library/arj

ISSUE 92
APRIL 2020
VOL. 27 NO. 2

Image designed by Michael Bubar-Krukowski

SOFTWARE PRODUCTIVITY TRENDS and ISSUES

David M. Tate

The Department of Defense is experiencing an explosive increase in its demand for software implemented features in weapon systems. The combination of exponential increases in computing power and similar advances in memory density and speed has made software mediated implementation of system features increasingly attractive. In the meantime, defense software productivity and industrial base capacity have not been growing as quickly as demand. This article uses the limited data that exist regarding defense software supply, demand, and productivity trends to estimate the severity of the capacity bottleneck, then briefly discusses the potential actions available to the Department to mitigate that bottleneck in the long run.

DOI: *https://doi.org/10.22594/dau.19-838.27.02*
Keywords: *Defense Industrial Base, Capacity, Demand Growth, Supply Shortfall, Cleared Workforce*

Malthus on Software

The Scottish cleric and economist Thomas Robert Malthus famously noted that, when there is enough food to go around, population growth is exponential. Since Malthus could not envision any means whereby food production could also grow exponentially, given the constraints of arable land and property ownership, he predicted that the inevitable result would be a population limited by recurring poverty and starvation.

Malthus was wrong about food, at least in his time, but could he be right about defense software? Any exponential growth in demand without a commensurate exponential growth in supply will soon be frustrated. Rapidly growing demand for new software, combined with the need to sustain the new code going forward, places considerable stress on the productive capacity of the defense software industrial base. The ability to keep up will depend on just how fast demand is growing, how quickly the Department of Defense (DoD) can grow the industrial base, and how quickly the productivity of individual software developers improves over time. To determine whether DoD should worry, we looked at each of those factors in turn.

Unfortunately, data on defense software demand and the defense software industrial base are extremely sparse. The last comprehensive attempt to estimate the total demand for national security software[1] and the capacity of the industrial base to meet that demand is more than a decade out of date (Chao, 2006). Data collection for the DoD software effort consists primarily of Software Resources Data Reports (SRDR) (Office of the Secretary of Defense, 2019), which are generally limited to major programs, not currently curated or normalized to a degree that can support DoD-level analysis, and subject to severe data quality issues (Arnold, Braxton, & Wingrove, 2015; Morin, 2017). At the same time, measuring software productivity is notoriously difficult, and such standards as exist are neither widely used nor consistently applied (Card, 2006; Krishnan, Kriebel, Kekre, & Mukhopadhyay, 2000).

In this article, we extrapolate from that last comprehensive baseline using a variety of more recent data sources to make plausible assumptions regarding changes in demand, workforce, and productivity. We find that even conservative estimates of these trends lead to a conclusion that, absent significant changes, future defense capabilities will be severely limited by the available productive capacity of the cleared software workforce.

How Fast Is Defense Demand for Software Growing?

Among the data available to support analysis of demand growth in defense software, some strong indicators are available:

- The National Research Council (2010) wrote "The extent of the DoD code in service has been increasing by more than an order of magnitude every decade, and a similar growth pattern has been exhibited within individual, long-lived military systems." One order of magnitude per decade is approximately 25% annual growth.

- The Aerospace Vehicle Systems Institute (n.d.) states that source lines of code (SLOC) in aircraft (both military and commercial) have been doubling approximately every 4 years. This corresponds to an annual growth rate of about 18%.

- The Department of the Army (2011) estimated that the volume of code under Army depot maintenance (either post-deployment or post-production support) had increased from 5 million to 240 million SLOC between 1980 and 2009. This corresponds to about 15% annual growth.

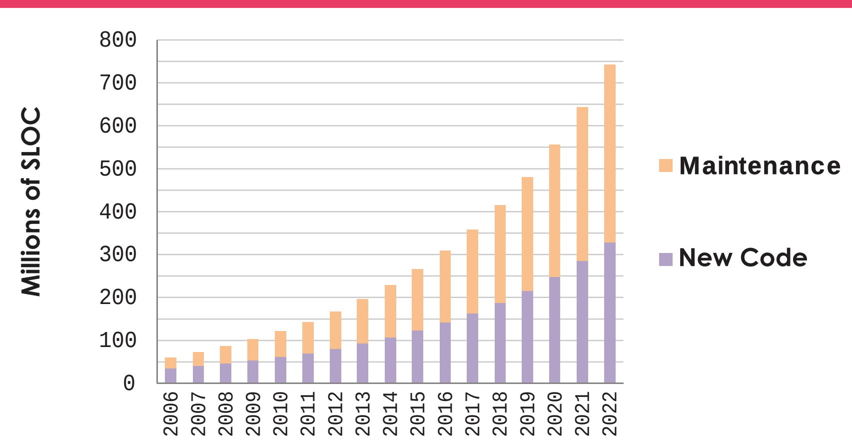
- Dvorak (2009) stated that National Aeronautics and Space Administration unmanned space systems SLOC have also increased by an order of magnitude every 10 years, with manned systems SLOC growing even faster.

Taken together, these suggest an annual growth rate of at least 15% for the amount of software being developed and maintained for defense purposes, with 25% or more annual growth possible. Annual growth of 15 to 25% means doubling every 3 to 5 years, on top of which is the added workload of maintaining the growing base of deployed code.

To forecast future demands for new code and software maintenance, we also need to know the current size of the code base and the current annual demand. Surprisingly, this information is apparently not being tracked. The most recent nationwide demand estimate we were able to find (Chao, 2006) concluded that the 2006 requirement for national security software was about 35 million lines of new code and 25 million lines of maintenance code. We can apply the "20% per year" rule of thumb for maintenance effort to infer a deployed 2006 base of about 125 million lines of code. We will base our analysis on those assumptions: 125 million source lines of code (MSLOC) under maintenance in 2006, 35 MSLOC of new code required in 2006, and annual demand for new code growing at 15% annually from that time forward. For maintenance effort, we assume that annual maintenance effort on the installed base is equivalent to 20% of the development effort of the base, and that half of the maintenance effort results in more new code to be maintained.[2] In addition, some fraction of the installed code base is retired every year. We will assume that 10% of the installed base is retired each year, exactly offsetting the new code generated by maintenance. As we will see, the conclusions of this investigation that follow are not sensitive to the exact parameter values chosen here, or the estimate of the current installed code base.

Figure 1 shows the projected growth in annual demand for defense software under these assumptions, separated into new code and maintenance of existing code. Bear in mind that this is a projection of unconstrained demand—how much DoD is expected to want to buy, if it is available at prices comparable to historical prices.

FIGURE 1. FORECAST OF DOD SOFTWARE DEMAND



It is worth noting that, under these assumptions, the total size of the deployed code base under maintenance is projected to be more than 1 billion SLOC by 2018, and more than 3 billion SLOC by 2025. Figure 1 shows only the new effort each year, not the deployed base.

The Supply of Defense Software

Chao (2006) estimated both the size of the defense software workforce and the productivity of that workforce. The productive capacity of the industrial base is the product of those two factors. We will attempt to update these estimates using such data as are available. For purposes of this analysis, we will accept Chao's estimates that 68,000 software developers possessed security clearances in 2006, with a productive capacity of 75 MSLOC per year. That implies a productivity at that time of roughly 1,100 SLOC per developer in 2006, or (equivalently) 900 developers required per MSLOC, as our baseline.

The Size of the Workforce

The Bureau of Labor Statistics (BLS, 2017) estimates that from 2010 to 2015, total employment of software developers[3] grew almost 30%, or about 5.3% annually. However, it forecasts that rate to decline sharply going forward, averaging only about 1.6% per year over the decade of 2014–2024 (BLS, 2016–2017). The defense software industrial base will need to grow more quickly than that to keep pace with established demand growth.

"Any scarcity of cleared software talent should translate into rising salaries and benefits for workers with those skills, providing incentive for more and more workers to enter the industry.

Any scarcity of cleared software talent should translate into rising salaries and benefits for workers with those skills, providing incentive for more and more workers to enter the industry. In a free and liquid market, we would expect this to happen fairly quickly. Unfortunately, some aspects of this particular market might be problematic. The first is the requirement that workers be U.S. citizens with security clearances. This not only dramatically restricts the pool of potential entrants; it also creates a licensure bottleneck for individuals seeking to join the labor force. In April 2019, the Director of the National Background Investigations Bureau reported that the backlog for security clearance investigations had been reduced by nearly one-third but was still nearly half a million cases. Processing times remained long, with initial investigations averaging 234 days for Secret clearances and 468 days for Top Secret (Kyzer, 2019). Hiccups in overall growth of the cleared workforce in the last decade have also emerged, driven by government response to high-profile leaks and worker response to the breach of Office of Personnel Management personal information files (Kyzer, 2015). Defense software employers are also facing tough competition from the commercial sector, which is experiencing an explosion of demand for software to power the expanding role of the Internet in daily life. While other industries can supplement U.S. graduates with offshore or immigrant labor, that solution is unavailable to the defense sector under current regulations.

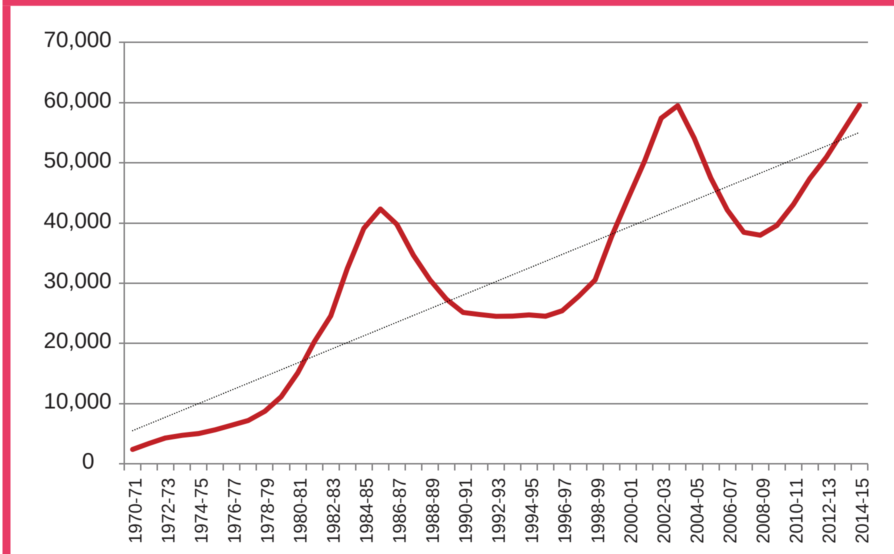
Another barrier to market corrections is that the most urgent scarcities seem to be at the high end of the experience scale. Chao (2006) found that (at least in 2006) no general shortage of programmers existed, but a significant shortage (with corresponding salary premium) of relatively senior software project managers, architects, and developers was already apparent. At the tip of the pyramid, they cited a cadre of 500–600 "elite" individuals who play a disproportionate role in project success.

Finally, it is not clear that DoD *wants* the market to correct itself through increases in compensation. Contractor labor rates are closely monitored by DoD, and the government pushes back when they rise too quickly. Senior software talent in the general economy can be as highly compensated as senior management executives. Arrington (2010a) reported that "[a Google employee] was recently offered a counter offer he couldn't refuse (except he did). He was offered a 15% raise on his $150,000 mid-level developer salary, quadruple the stock benefits and...wait for it...a $500,000 cash bonus to stay for a year. He took the Facebook offer anyway." (Note that $150,000 for a mid-level developer is already well above industry norms.) Arrington

(2010b) also reported that Google had paid a top software engineer $3.5 million to turn down an offer from Facebook. Reimbursable federal contractor labor costs are capped by the provisions of Section 702 of the Bipartisan Budget Act of 2013 (Pub. L. 113-67); companies choosing to pay salaries higher than that cap ($525,000 in 2018) must take the difference out of profit (Office of Federal Procurement Policy, 2018). This provides a disincentive to employing top software talent on federal contracts. For federal civilian employees, the permitted salaries are even more tightly constrained.

On the supply side, what does the educational pipeline for software look like? The number of bachelor's degrees conferred each year in computer and information sciences has shown a striking cyclical pattern over the past 4 decades (Figure 2). The general trend has been a baseline increase of about 1,000 degrees per year, with superimposed boom-and-bust cycles. We are currently on the upswing of a boom cycle, with more than 60,000 degrees conferred per year.

FIGURE 2. ANNUAL COMPUTER SCIENCE AND INFORMATION SCIENCES BACHELOR S DEGREES CONFERRED



Note. Only every other academic year is labeled. Source: National Center for Education Statistics (2012, 2016).

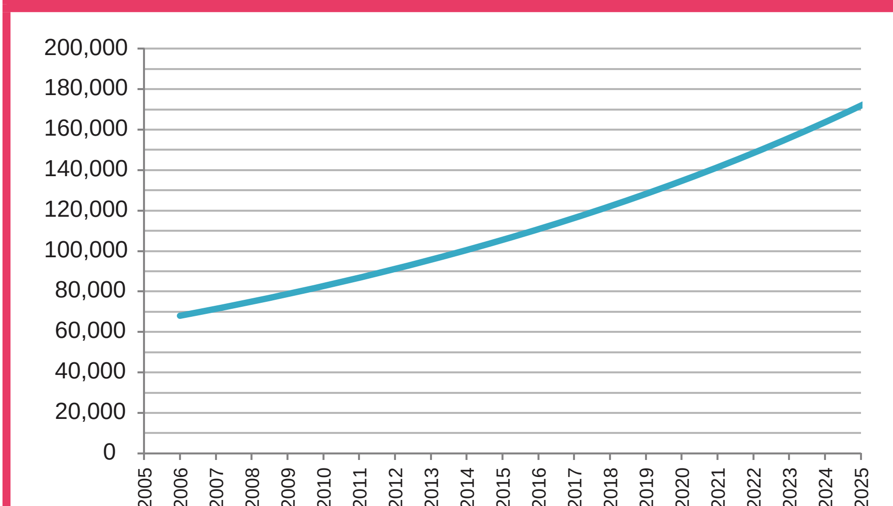
In addition to this pool of potential defense software developers, the educational pipeline for software developers also includes nontraditional educational options. More than 16,000 students graduated from "coding boot camp" programs in 2015, and that number has been growing rapidly over the few years that such programs have existed (Lauerman, 2015).

This suggests that as many as 80,000 potential developers are graduating per year. In 2006, the cleared software workforce made up 7% of the national software workforce and 16% of the overall cleared workforce (Chao, 2006). Again, optimistically, if 10% of new graduates (college and boot camp combined) end up in the cleared software workforce, that would currently be about 8,000 per year, which could grow to 10,000 per year in a couple of years. This corresponds to between 5% and 10% annual growth. For purposes of our baseline analysis, we will assume annual workforce growth of 5%, comparable to recent growth in software developers and well above the forecast national average for the software industry.

"More than 16,000 students graduated from "coding boot camp" programs in 2015, and that number has been growing rapidly over the few years that such programs have existed (Lauerman, 2015).

As noted in the preceding discussion, in 2006, roughly 68,000 cleared software developers were employed in the defense industrial base. If we assume 5% annual growth in the national security software developer workforce starting in 2006, that would translate to about 120,000 people today, reaching 150,000 by 2023. Figure 3 shows this projected growth over time.

FIGURE 3. FORECAST OF CLEARED SOFTWARE WORKFORCE SIZE



The Productivity of Defense Software Developers

Malthus was wrong about hunger in England, in large part because the technology for food production improved enormously over the next few centuries, making individual farmers much more productive and bringing marginal land into productive use. A comparable technology revolution in software productivity could offset the growth in software demand, even if the workforce grows only slowly. Are there signs of such growth in individual productivity?

In 2000, Jones estimated defense software productivity at 4.2 function points (FP) per staff month (SM); in 2013, his estimate was 6.75 FP/SM. That corresponds to just under 4% annual productivity improvement. This is in line with other historical estimates of software productivity growth. For example, Longstreet (2001) estimated about 4% annual productivity growth (FP per hour) from 1970 to 2000 industry-wide. These estimates are based on FP, rather than on MSLOC. Since the number of FP per line of code has been growing historically (Jones, 2013), productivity growth in terms of MSLOC would be somewhat lower, but we will optimistically estimate MSLOC productivity growth at 4% as well.

Of course, DoD may not yet have realized all of the productivity enhancement that can be had using current technology. The potential for leap-ahead productivity improvements—analogous to the farming breakthroughs that Malthus failed to foresee—is discussed in the Recommendations section.

Supply vs. Demand

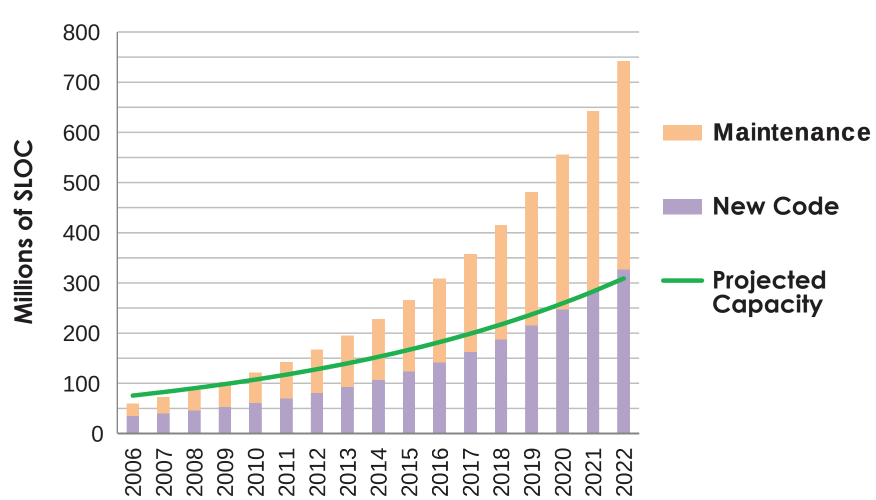
We now have all of the pieces we need for an end-to-end estimate of future productive capacity versus projected unconstrained demand. Figure 4 shows that, even using generally optimistic assumptions and estimates, demand already exceeds the capacity of the industrial base. According to this forecast, DoD will soon also reach the point of neither being able to produce all of the new code desired (without maintenance), nor to maintain all existing code (with no new development). The projected 2020 workforce of 135,000 developers would be less than half of the 290,000 developers required to write and maintain all of the code desired up to that point.

Revisiting the assumptions behind this forecast, we have assumed:

- 15% annual growth in demand for new code
- 5% annual defense software workforce expansion
- 4% annual productivity growth

- A workforce of 68,000 in 2006
- Demand for 35 MSLOC in 2006
- An installed base of 125 MSLOC in 2006
- Productive capacity of 75 MSLOC in 2006
- 20% annual maintenance effort
- 50% of maintenance resulting in new code
- 10% annual retirement of software in the base

FIGURE 4. FORECAST OF SUPPLY VS. UNCONSTRAINED DEMAND



Most of these assumptions could be fairly described as optimistic, given the available historical data. Varying the parameters changes the details of the forecast, but the bottom line remains the same. For example, if we assume that productivity growth post-2006 will be 8% instead of 4%, we get the results in Figure 5. Software development is still capacity-constrained in this case, but not as severely. Conversely, if we keep productivity growth at 4% but allow the workforce to grow by 10% per year, we get the results in Figure 6.

Necessarily, the reverse is also true—if annual demand growth is closer to 20%, or the 2006 installed base was significantly larger than 125 MSLOC,[4] or cleared workforce growth stagnates, then all of these pictures would look much worse. Assuming less optimistic values for the annual maintenance fraction (40%), or the proportion of maintenance that generates new code (>50%) (Galorath, n.d.), would also lower the estimated future capacity significantly.

FIGURE 5. FORECAST OF SUPPLY VS. DEMAND AT 8% PRODUCTIVITY GROWTH

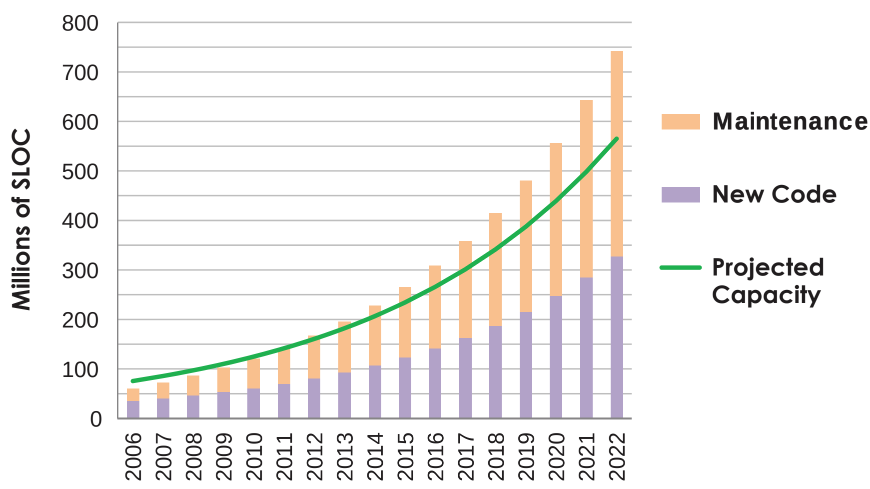
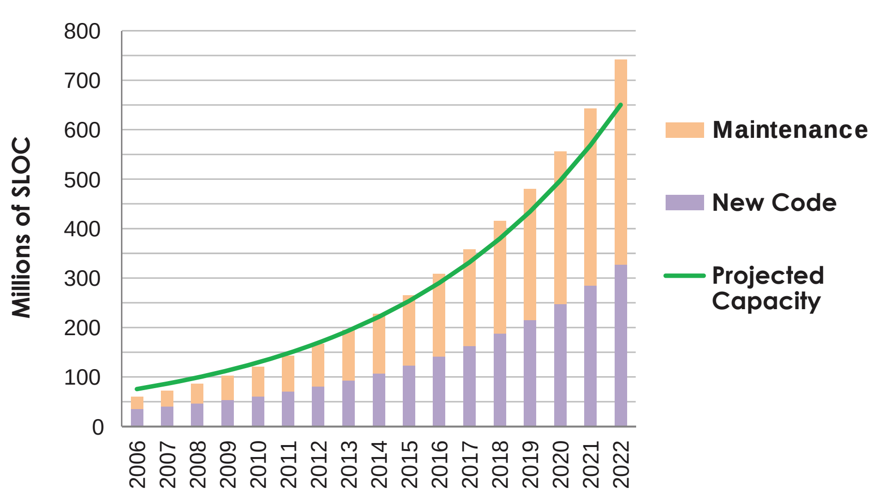


FIGURE 6. FORECAST OF SUPPLY VS. DEMAND ASSUMING 10% WORKFORCE GROWTH



If This Were Correct, Wouldn't Someone Have Noticed?

Is it really possible that the nation could be suffering a (possibly severe) shortage of software developers in the defense sector without anyone noticing? What symptoms should analysts look for?

Barnow, Trutko, and Piatak (2013) list 16 separate employer actions that might indicate a labor shortage. These include increased recruiting expenditures, increased use of overtime, new on-the-job training programs, relaxing minimum qualifications, etc. These are in addition to the operational symptoms of resource shortage, such as increased development times, lower-than-predicted staffing levels, and higher ratios of systems engineering/program management costs to touch labor costs.

Evidence of these indicators is already present in the defense sector.

- Chao (2006) found that senior software architects and project managers in the cleared software sector earned at least 50% more than their counterparts in the general economy. They took this to indicate that those particular skills were already in short supply throughout the defense industrial base.
- Lucero (2009) found that many defense software positions were being filled by personnel with no formal software engineering training (on-the-job training).
- The Government Accountability Office (GAO) quoted Air Force officials at Ogden Air Logistics Complex as expressing "concerns about [...] personnel gaps, largely in software maintenance" (GAO, 2017, Vacancies).
- As of late 2019, ClearanceJobs.com had nearly 30,000 job postings for software developers, software engineers, or software managers. This was nearly half of all listings at that site.
- Salaries for cleared information technology program/project managers rose 10% in one year between 2013 and 2014, faster than any other category and passing engineers as the highest compensated cleared occupation group (ClearanceJobs.com, 2014, Salary Rise).

- BLS estimates the national unemployment rate for technology professionals at only 2.9% (ClearanceJobs.com, 2014, Vacancies).
- Nearly half of recent ClearanceJobs.com survey respondents have been in their current job less than 3 years (Kyzer, 2017, Churn).

Barnow et al. (2013) also note that measuring occupational shortages is difficult, in part because occupational vacancy data are not generally available in the United States. Also, available reporting uses job classification systems that are based on outdated industrial models and are too broad to be useful for many purposes. It would be very interesting to look at (for instance) how the aggregate cost per staff month of defense software development has changed over the past decade, as reflected in SRDR reporting of major programs.

What Are the Policy Options?

We identify several available short-term and long-term policy options associated with both the supply and demand for defense software. In the long run, success may depend on breakthroughs in the last and most speculative of the options—investment in the transition from a craft labor model to an industrial automation model of software development.

Option 1: Moderate Demand

The obvious short-term solution to a scarcity of software productive capacity is to ask for less software. At the present time, it seems unlikely that the defense establishment would be willing or able to accomplish this. Software is viewed as vital to any hope of maintaining the United States' traditional technological advantage in military capability. A significant overall reduction in software demand would also require the federal agencies that procure national security software to cooperate effectively to optimize the allocation of software development capacity to the most important, software-intensive programs. Given that these agencies struggle to allocate resources efficiently within and among their own acquisition portfolios, this seems like a stretch. The results, then, would be a less-efficient allocation of software resources to capabilities, an associated effective loss of software productivity, and failure to reap the potential benefits of software-mediated capabilities.

In the longer term, natural factors limit the growth in demand for software. Defense budgets do not grow without limit, so the exponential growth in software demand reflects, to some extent, substitution of software for other categories of expenditure—primarily analog hardware and human labor. There are natural limits to that process. Regardless of the underlying desire for software-mediated capabilities, DoD cannot procure more software than it can afford, or than the industrial base is able to provide.

"If rapid response to a rapidly changing world is one of the motivations for implementing capabilities in software, it makes no sense to pursue designs whose complex software will require 20 years or more to design, build, and test.

Perhaps just as important, the size and complexity of the software in a system affects how long it takes to develop and field that system. If rapid response to a rapidly changing world is one of the motivations for implementing capabilities in software, it makes no sense to pursue designs whose complex software will require 20 years or more to design, build, and test. Prior analysis of the dependence of development cycle times on software content assumed development times unconstrained by industrial base issues (Tate, 2016). If Major Defense Acquisition Program/Major Automated Information System (MDAP/MAIS) software projects are now subject to chronic resource shortfalls, those past lead-time estimates were optimistic. Increased demand for software-mediated functions thus has a twofold negative effect on schedules: first, by adding work to the critical development path of each program; and second, by starving the programs of the resources necessary to do the work on the critical path. From a policy perspective, it does not seem practical for DoD or Congress to mandate reduced use of software overall, or to set limits on the amount of software in any one program. Not only would those policies be counterproductive, they would also be unenforceable, and prone to wasteful gaming by the Services and defense contractor base. Demand-side policy options appear to be unhelpful here.

Option 2: Grow the Workforce

From a policy perspective, several plausible mechanisms are available to increase the effective growth rate of the defense software base:

- Encourage students to pursue software education, both through traditional college degrees and nontraditional (e.g., boot camp) training programs. Incentives could include low-interest loans, direct subsidies/scholarships, loan forgiveness, etc. These could be made contingent on a minimum tenure of employment in the defense sector. Incentives like this have been successful in increasing primary care physician recruitment in underserved areas (Verma et al., 2016).

- Continue to invest in improving the throughput of the security clearance process, especially for software workers. While this has been a priority for DoD in recent years, progress has been slow (Kyzer, 2019).

- Relax barriers to employing foreign nationals. The software industry has thoroughly globalized, but the defense sector is not permitted to take advantage of that at present. As we shall see in the discussion that follows, increased use of open source software accomplishes this implicitly without relaxing security standards.

- Adjust restrictions on allowable contract costs for software talent.

The first three of these options would tend to reduce the price of defense software by increasing supply, thus somewhat offsetting the investment required. Allowing higher reimbursable salaries for key software professionals looks like it would tend to increase the cost of any given system—but it might not. It might improve efficiency and increase supply by enough to offset the higher cost per hour of that expert labor. It might also improve the availability, timeliness, and quality of delivered systems.

Option 3: Improve Productivity Dramatically

Multiple drivers of significant productivity improvement have permeated the commercial software world over the past few decades. These include computer-aided software engineering (CASE) tools (Krishnan et al., 2000), automated test environments (National Institute of Standards and Technology, 2002), improved programming languages (Jones, 2013, Table 14.2),[5] agile (and similar) development processes, open source ecosystems (Lerner, 2010), and modular open system architectures. The defense software base has shared in the benefits of CASE tools and automated test environments, and to some extent from open source software (Wheeler, 2010). Improved productivity through programming language modernization was temporarily delayed during the 1990s by the mandate to write defense software in Ada, and continues to be hampered by the large installed base of defense software in obsolescent languages such as COBOL, FORTRAN, and HAL/C. DoD has not yet leveraged agile development practices or modular open architectures to a significant degree (Defense Innovation Board [DIB], 2019).

Definitions of "agile development" invariably lead to arguments among both advocates and skeptics, but in general the phrase refers to a strategy of rapid, small-scale, incremental development and release of software functionality, driven not by prespecified final requirements or specifications, but rather by close, iterative interaction with future users of the software. The key features here are:

- Small—features are added in many small increments, rather than a few large blocks/versions/updates
- Rapid—new releases happen on a scale of weeks, not months or years
- No fixed requirements—users, developers, and other stakeholders together explore the space of potential features and discover which are the most useful
- Interactive—stakeholders and developers work as a collaborative partnership, rather than as customer and vendor, with developers in self-organizing teams

All of these key features pose problems for traditional DoD acquisition (Broadus, 2013). Having many small incremental releases of functionality breaks the logistics system whereby new software releases are coordinated and deployed to far-flung operational units. The absence of fixed formal requirements is antithetic to the Joint Requirements Oversight Council (JROC) mission of specifying formal, validated requirements with threshold levels. It may also cause legal and practical headaches for the writers of requests for proposals and the awarders of contracts, not to mention cost and schedule estimators. The interaction between developers and users requires active, ongoing participation of uniformed and civilian personnel who would traditionally never get near the system under development until (perhaps) Operational Test and Evaluation. That ongoing collaboration might last for years.

Open source software refers to software that is collaboratively developed and maintained by a community of volunteer contributors. Examples of thriving open source ecosystems include the Linux operating system, the Apache web-hosting platform, the FreeRTOS real-time operating system for embedded systems, the R and Python programming environments, the emacs document editor, and the MySQL relational database. The collaborative nature of the communities of developers working with these tools can lead to enormous total effort. For example, the Linux Foundation estimated in 2008 that the total cost to develop from scratch the Fedora 9 distribution of Linux (including the Linux kernel itself) would have been more than $12 billion (McPherson, Proffitt, & Hale-Evans, 2008). That was nearly a decade of additional development ago.

The other dominant recent development in the commercial world that has generated significant productivity gains is the use of modular open system architectures. Stephen Welby (2014), during his time as Deputy Assistant Secretary of Defense for Systems Engineering, described these as "technical architectures that leverage technical standards to support a modular, loosely coupled and highly cohesive system structure" (p. 3). Note that modularity and openness are distinct concepts, each of which contributes separately to potential productivity improvements. Modularity is about the way the software's functions are organized into composable units. Openness is about who can see, modify, publish, or use the code. Not all modular architectures are open; not all open systems are modular. However, a synergy exists between the two ideas—modularity increases the efficiency of individual contributions to the open code base, while openness allows more competition and participation.

For DoD, the following key features could drive enhanced productivity:

- Composable software modules that can be combined in many ways *without modification* to execute more complex functions
- Well-defined, standardized, documented interfaces for these modules
- Universal transparent access to (nearly) all of the source code
- Extensive rights to modify or enhance existing source code
- A large base of independent agents actively engaged in developing, improving, and maintaining the software without being directly paid by the eventual users

Modularity would allow large parts of the code base to be reused in new applications with little or no modification, greatly reducing development times. It would also make future upgrades faster and cheaper. Judicious use of open source software, on the other hand, would enable DoD to take advantage of a huge body of well-maintained and inexpensive software that already exists. Studies have shown that the community approach to development results in a higher level of scrutiny—and thus, generally lower defect rates—for frequently used modules in such ecosystems (Brockmeier, 2003). Similarly, software assurance and cybersecurity can sometimes be easier for open source software than for proprietary software (Wheeler, 2010).

Open source ecosystems also provide a potential indirect mechanism for opening defense software development to the noncleared workforce. Any defense software that is based on Linux, or written in Python, or implemented using FreeRTOS, is leveraging the efforts of thousands of developers outside the usual defense workforce. In the end, this might be the best argument in favor of open source—that it promises not only significant productivity gains from leveraging existing commercial software, but also the largest available effective expansion of the defense software workforce.

Recommendations

Thus far, we have seen estimates of supply and demand, some optimistic yet sobering forecasts, and an enumeration of possible policy options. These lead naturally to three principal recommendations:

Recommendation 1: Collect data.

Study the industrial base; measure the effective demand; measure the installed code base and ongoing maintenance efforts; measure industrial base capacity and productivity. Indeed, formulating sensible strategy without this basic information is impossible.

The forecasts in this article are built on sparse data from inconsistent sources, because those are the only data that exist. An improved update to the Chao (2006) investigation of the state of the defense software industrial base is long overdue and would enable DoD to replace the very uncertain estimates in this article with actionable facts.

Recommendation 2: Adopt commercially proven, productivity-enhancing acquisition models.

In recent decades, DoD has bet that the boom in commercial software is a rising tide that would necessarily lift defense software productivity as well. As documented throughout this article, this turned out to be only partially true; the needs and culture of DoD acquisition are sufficiently different from those of the commercial world that some productivity advances arising in the commercial sector did not automatically translate to the defense sector. For example, agile development methods, though greatly desired by both DoD and external advisors (DIB, 2019), will require radical changes in requirements management, stakeholder involvement in development, and acquisition planning and budgeting processes (Broadus, 2013). Similar institutional and cultural barriers prevail that oppose or seek to limit expanded use of open source software and modular open systems architecture. DoD leadership have been pushing in this direction (DoD, 2017), but there is considerable institutional inertia and active resistance to be overcome, both within government and within the industrial base.

Recommendation 2A: In particular, embrace open source software ecosystems. Of the existing productivity enhancers, this is the only one that might potentially provide both immediate rapid productivity improvement and an effective expansion of the workforce. Evidence from the commercial world suggests that embracing open source software would not only be cost-effective, it might be necessary in order to keep up with the pace of technology change and threat evolution. Furthermore, the early stages of developing such ecosystems might well not look much like progress.

Recommendation 2B: Enforce modular architectures and pursue data rights to enable fast and efficient insertion of future upgrades into legacy platforms and systems. One of the best ways to improve future productivity is to build systems that are easy to modify. When major platforms are in service for decades, the ability to host the essential capabilities of the future is more important than the requirement that they deliver maximum capability the day they are fielded (Patel & Fischerkeller, 2013).

Recommendation 3: Fund basic productivity research the way DoD used to do.

The federal government, and in particular DoD, played an enormous role in funding and guiding the development of core software technologies that enabled U.S. dominance in that industry for the first few decades of the computer age (Mowery, 1999). Without additional fundamental improvements in software productivity, DoD's ability to field needed capabilities quickly enough to keep pace with changing threats will be limited by the time it takes to develop the software that implements those capabilities. DoD is already struggling to keep pace; major leaps forward are needed.

In particular, focus research on freeing software development from its current 19th-century industrial model. At present, each software application written is the custom hand-tooled product of skilled craftsmen, analogous to the way automobiles were made before Henry Ford revolutionized that industry. In the long run, the key productivity breakthrough must be the automation of software development as a process, enabling mass production and industrialization. Software product lines (Hinchey, 2018) are a start, but the transformational end-state, with autonomous systems with only high-level human guidance writing software from scratch with the dependability required of defense systems, is currently still in the realm of science fiction.

References

Aerospace Vehicle Systems Institute. (n.d.). SAVI–System Architecture Virtual Integration. Retrieved from http://savi.avsi.aero/about-savi/savi-motivation/exponential-system-complexity/

Arnold, R., Braxton, P., & Wingrove, K. (2015, June). The continual pursuit of the one true Software Resources Data Reporting (SRDR) database. In *Proceedings of the 2015 International Cost Estimating and Analysis Association (ICEEA) Professional Development and Training Workshop*, San Diego, CA.

Arrington, M. (2010a). *Google making extraordinary counteroffers to stop flow of employees to Facebook.* Retrieved from https://techcrunch.com/2010/09/01/google-making-extraordinary-counteroffers-to-stop-flow-of-employees-to-facebook/

Arrington, M. (2010b). *Google offers staff engineer $3.5 million to turn down Facebook offer.* Retrieved from https://techcrunch.com/2010/11/11/google-offers-staff-engineer-3-5-million-to-turn-down-facebook-offer/

Barnow, B. S., Trutko, J., & Piatak, J. S. (2013). How do we know occupational labor shortages exist? *Employment Research, 20*(2), 4–6. doi:10.17848/1075-8445.20(2)-2

Broadus, W. (2013). The challenges of being agile in DoD. *Defense AT&L, 42*(1), 5–9.

Brockmeier, J. (2003). *Comparing free and proprietary software defect rates.* Retrieved from https://lwn.net/Articles/22623/

Bureau of Labor Statistics. (2016–2017). Software developers, job outlook. *Occupational Outlook Handbook, 2016-17 Edition.* Retrieved February 17, 2017, from https://www.bls.gov/ooh/computer-and-information-technology/software-developers.htm

Bureau of Labor Statistics. (2017). *Occupational employment statistics.* Retrieved from https://www.bls.gov/oes/tables.htm

Card, D. N. (2006). The challenge of productivity measurement. In *Proceedings of the Pacific Northwest Software Quality Conference 2006,* Portland, OR.

Chao, P. (2006, October). *An assessment of the national security software industrial base* [Briefing]. Retrieved from https://www.csis.org/analysis/assessment-national-security-software-industrial-base

ClearanceJobs.com. (n.d.). *Largest cleared career network.* Retrieved from https://www.clearancejobs.com/jobs?keywords=software

ClearanceJobs.com. (2014). *In the clear: Compensation decline for security-cleared professionals levels off despite strong headwinds.* Retrieved from https://about.clearancejobs.com/hubfs/pdfs/ClearanceJobs_Compensation_Survey_2014.pdf

Defense Innovation Board. (2019). *Software is never done: Refactoring the acquisition code for competitive advantage.* Report of the Defense Innovation Board. Retrieved from https://media.defense.gov/2019/Mar/26/2002105909/-1/-1/0/SWAP.REPORT_MAIN.BODY.3.21.19.PDF

Department of the Army. (2011). *Army organic industrial base strategic plan 2012–2022.* Retrieved from https://www.army.mil/e2/c/downloads/276549.pdf

Department of Defense. (2017, February 23). *DoD announces the launch of "Code.mil," an experiment in open source* [DoD News Release No: NR-077-17]. Retrieved from https://www.defense.gov/News/News-Releases/News-Release-View/Article/1092364/dod-announces-the-launch-of-codemil-an-experiment-in-open-source?source=GovDelivery

Dvorak, D. L. (Ed.) (2009). *NASA study on flight software complexity.* Pasadena, CA: NASA Jet Propulsion Laboratory, California Institute of Technology.

Galorath. (n.d.). *SEER® by Galorath*. Retrieved from http://galorath.com/software_maintenance_cost

Government Accountability Office. (2017). *Depot maintenance: Executed workload and maintenance operations at DOD depots* (Report No. GAO-17-82R). Washington, DC: Author.

Hinchey, M. (2018). Dynamic software product lines. In *Software Technology: 10 Years of Innovation in IEEE Computer.* Hoboken, NJ: Wiley-IEEE Press.

Jones, C. (2000). *Software assessments, benchmarks, and best practices.* Boston, MA: Addison-Wesley.

Jones, C. (2013). Function points as a universal software metric. *ACM SIGSOFT Software Engineering Notes, 38*(4), 1–27. doi:10.1145/2492248.2492268

Krishnan, M. S., Kriebel, C. H., Kekre, S., & Mukhopadhyay, T. (2000). An empirical analysis of productivity and quality in software products. *Management Science, 46*(6), 745–759.

Kyzer, L. (2015). *How the OPM breach affected hiring for clearance jobs.* Retrieved from https://www.govexec.com/management/2015/10/how-opm-breach-has-affected-hiring-clearance-jobs/123123/

Kyzer, L. (2017). *Clearance salary trends in 2017.* Retrieved from https://news.clearancejobs.com/2017/01/19/clearance-salary-trends-2017/

Kyzer, L. (2019). *Progress is finally being made on security clearance backlog.* Retrieved from https://www.govexec.com/management/2019/04/progress-finally-being-made-security-clearance-backlog/156481/

Lauerman, J. (2015). *Coding boot camp enrollment soars as students seek tech jobs.* Retrieved from https://www.bloomberg.com/news/articles/2015-06-08/coding-boot-camp-enrollment-soars-as-students-seek-tech-jobs

Lerner, J. (2010). The commingled code: Open source and economic development. *MIT Press Books, 1*(1).

Longstreet, D. (2001). Software productivity since 1970. *Function Point Training and Analysis Manual.* Retrieved from http://www.softwaremetrics.com/Articles/history.htm

Lucero, D. (2009, April). *Software sustainment challenges in defense acquisition.* Paper presented at the AIAA Infotech@Aerospace Conference, Seattle, WA. doi:10.2514/6.2009-1816

McPherson, A., Proffitt, B., & Hale-Evans, R. (2008). *Estimating the total cost of a Linux distribution.* Retrieved from https://www.linux.com/publications/estimating-total-cost-linux-distribution

Morin, J. (2017). *DoD cost analysis data improvement* [Memorandum]. Washington, DC: Office of the Director, Cost Assessment and Program Evaluation.

Mowery, D. C. (1999). The computer software industry. In D. C. Mowery & R. R. Nelson (Eds.), *Sources of Industrial Leadership: Studies of seven industries* (pp. 133–168). Cambridge, UK: Cambridge University Press.

National Center for Education Statistics. (2012). *Digest of Education Statistics: 2012* [Data Set]. Retrieved from https://nces.ed.gov/programs/digest/d12/tables/dt12_349.asp

National Center for Education Statistics. (2016). *Digest of Education Statistics: 2016* [Data Set]. Retrieved from https://nces.ed.gov/programs/digest/d16/tables/dt16_322.10.asp?current=yes

National Institute of Standards & Technology. (2002). *The economic impacts of inadequate infrastructure for software testing: Final report.* Gaithersburg, MD: Acquisition and Assistance Division.

National Research Council, Committee for Advancing Software-Intensive Systems Producibility. (2010). *Critical code: Software producibility for defense*. Washington, DC: The National Academies Press. doi:10.17226/12979

Office of Federal Procurement Policy. (2018). *Contractor compensation cap for contracts awarded on or after June 24, 2014.* Retrieved from https://www.whitehouse.gov/wp-content/uploads/2017/11/ContractorCompensationCapContractsAwardedafterJune24.pdf

Office of the Secretary of Defense. (2019). *The Software Resources Data Report (SRDR) implementation guidance.* Washington, DC: Office of Cost Assessment and Program Evaluation.

Patel, P. R., & Fischerkeller, M. P. (2013). *Prepare to be wrong: Assessing and designing for adaptability, flexibility, and responsiveness* (IDA Paper P-5005). Alexandria, VA: Institute for Defense Analyses.

Tate, D. M. (2016). Software development may drive future acquisition cycle times. *IDA Research Insights* (NS D-8053). Retrieved from https://www.ida.org/research-and-publications/publications/all/s/so/software-development-may-drive-future-acquisition-cycle-times

Verma, P., Ford, J. A., Stuart, A., Howe, A., Everington, S., & Steel, N. (2016). A systematic review of strategies to recruit and retain primary care doctors. *BMC Health Services Research, 16*(126).

Welby, S. P. (2014, October). Modular open systems architecture in DoD acquisition. In *Proceedings of the 17th Annual National Defense Industrial Association (NDIA) Systems Engineering Conference,* Springfield, VA.

Wheeler, D. A. (2010, April). *Open source software (OSS/FLOSS) and security.* Powerpoint presentation to the International Workshop on Free/Open Source Software Technologies, Riyadh, Saudi Arabia. Retrieved from https://www.dwheeler.com/essays/oss_security_saudi_arabia.pdf

Endnotes

[1] "National security software" is taken to include software used for national security applications and missions by the Department of Defense, Department of Energy, and Department of Homeland Security. At present, defense software is by far the largest portion of national security software.

[2] Jones (2013) estimates the maintenance costs of a nominal 1000-function point application at closer to 40% per year over the first 5 years. Using that estimate would result in a smaller 2006 deployed code base estimate, but much faster growth in that base in subsequent years.

[3] BLS occupation codes 15-1132 (Software developers, applications) and 15-1133 (Software developers, system software), total employment as of May 2010 and May 2015, retrieved from Occupational Employment Statistics on the BLS website, at https://www.bls.gov/oes/tables.htm.

[4] Given that the Army alone claimed to have 240 MSLOC under sustainment in 2009, 125 MSLOC defense-wide in 2006 seems improbably low.

[5] For our purposes, "improved" simply means more Function Points (or lines of code) of product per staff month of effort, on average.

Acknowledgments

This work was performed under the IDA Central Research Program, supported by Department of Defense contract HQ0034-14-D-0001.

Author Biography

Dr. David Tate

joined the research staff of Institute for Defense Analyses' Cost Analysis and Research Division in 2000. Prior to that, he was a Senior Operations Research Analyst (Telecom) at Decision-Science Applications and an Assistant Professor of Industrial Engineering at the University of Pittsburgh. Dr. Tate holds bachelor's degrees in Philosophy and Mathematical Sciences from the Johns Hopkins University, and MS and PhD degrees in Operations Research from Cornell University.

(E-mail address: dtate@ida.org)

Image designed by Michael Bubar-Krukowski

Analysis of Military Construction COST GROWTH in USAF MAJOR DEFENSE ACQUISITION PROGRAMS

Capt Emily E. Angell, USAF, Edward D. White, Jonathan D. Ritschel, and Alfred E. Thal, Jr.

This study uses descriptive and inferential statistics to identify cost growth of military construction (MILCON) at the programmatic level, while bridging the gap between Selected Acquisition Report (SAR) estimates and actual project costs. Findings of this study aid the cost community with appropriate allocation of resources in developing these estimates. Overall, Major Defense Acquisition Programs (MDAP) appear to experience more negative growth (cost savings) in MILCON estimates on reviewed SARs—typically less than 0.2% of the total program cost. SAR estimates became more accurate from the first to last SAR in comparison to the total MILCON programmed for all projects within a program. However, the last SAR's median MILCON cost estimate was approximately $31 million underestimated on projects currently authorized and appropriated for MDAPs. Preliminary research was restricted to 32 programs of which only 10 had authorized and accessible projects for comparison. Initial results suggest building on this exploratory analysis.

DOI: *https://doi.org/10.22594/dau.19-840.27.02*
Keywords: *ACAT I, Air Force Military Construction (MILCON), Selected Acquisition Report (SAR), Cost Growth, Major Defense Acquisition Program (MDAP)*

A myriad of factors within and outside of the Department of Defense (DoD)'s control can affect the differences often reported between initial cost estimates and final costs of military construction (MILCON) projects. However, recent MILCON projects with cost overruns have raised congressional concerns regarding the quality of DoD MILCON cost estimating practices, emphasizing the importance of an accurate cost estimate (Government Accountability Office [GAO], 2018). MILCON cost overruns are the increase of actual funds required to complete a project that has already been authorized and appropriated for execution at a lower budgetary level. In contrast, MILCON cost growth refers to the increase in cost estimates for a project or program over time (a program can consist of several projects); it can also represent a positive difference between an estimate at a given time and actual costs. Although previous studies have focused on MILCON cost overruns for projects, there appears to be no published studies documenting MILCON cost growth at a programmatic level. This article addresses this shortfall and investigates whether similar MILCON cost overruns occur at the program level.

Background

When the U.S. Air Force (USAF) acquires new programs, MILCON project requirements often accompany an MDAP, or Major Defense Acquisition Program (2017). By statute enacted in 2017 (10 U.S.C. § 2430), MDAPs are categorized as Acquisition Category I (ACAT, 2018) programs if they meet any of the following threshold criteria:

1. Total eventual expenditure of research, development, test and evaluation costs greater than $480 million (fiscal year 2014 constant dollars)

2. Total eventual expenditure of procurement costs greater than $2.79 billion (fiscal year 2014 constant dollars)

3. Specifically designated by milestone decision authority as special interest

All MDAPs are required to submit a periodic status report to Congress containing cost, schedule, and technical information; this report is called the Selected Acquisition Report (2018) and is prepared by each respective program office. The annual reporting for a particular program may be terminated by the Under Secretary of Defense (Acquisition and

Sustainment) when 90% of expected production deliveries have been made or planned acquisition expense has been disbursed (SAR, 2018). Until such time, reporting must continue periodically.

Title 10 U.S.C. § 2432 (2010) and Air Force Instruction (AFI) 32-1021 (USAF, 2016) mandate that all anticipated system-specific MILCON costs be estimated in every SAR for all MDAPs, if applicable (SAR, 2010). Project cost estimates are typically prepared by civil engineer units at bases or headquarters where new facilities are expected throughout the life of the program acquisition. The program office is responsible for submitting an accumulated programmatic MILCON cost estimate in each SAR submitted to Congress. According to AFI 32-1021 (USAF, 2016), MILCON project development and cost estimation begins with base civil engineer units using a DD Form 1391, Military Construction Project Data, to explain and to justify the project through all levels of the Air Force, Office of the Secretary of Defense, Office of Management and Budget, and Congress. Each of these forms includes the cost estimate for a single project, which assists in the use of parametric estimating tools with historical cost data where applicable.

Congress has historically scrutinized the DoD for MILCON cost overruns of projects from the time of funding appropriation through project completion. The GAO has reviewed MILCON project processes along with specific in-depth case studies for projects of interest for some time. The first GAO study to focus on the cost estimating of MILCON projects was reported in 1981 and concerned the variability to actual costs. The latest GAO study on MILCON cost estimating was reported in 2018 and focused on the reliability of the estimates.

The first GAO (1981) fieldwork study analyzed a broad sample of 83 MILCON projects from Fiscal Years (FY) 1978–1980; these projects represented a variety of facility types in various stages of cost overruns, cost underruns, and close to budget amounts. They found that most projects were estimated at least 18

months prior to project bidding for contract and that it was not unusual for the contract amount to differ from the estimated amount that was submitted to Congress for budget. This is an important recognition considering the MILCON costs reported in SARs are inevitably estimated more than 18 months prior to contract bidding. Additionally, GAO found that even with the most accurate information at 100% complete design, the actual cost is still influenced by bidder interest in a particular project, fluctuations in labor and material costs, changes in requirements or design after budget submission, and changes in site location for geographical and/or environmental reasons.

> **"Congress has historically scrutinized the DoD for MILCON cost overruns of projects from the time of funding appropriation through project completion.**

Concerned with constrained fiscal resources and the military's ability to effectively plan, estimate, and execute MILCON projects, Congress recently directed the Comptroller General of the United States to review and to report on DoD's MILCON cost estimating procedures. This mandate resulted in the 2018 GAO study, which analyzed MILCON appropriations from FYs 2005–2016 totaling $66 billion for all DoD MILCON projects during those 12 years. By the end of FY 2016, DoD had obligated $60.9 billion (92%) and expended $55 billion (83%). Research specific to FYs 2010–2016 discovered that DoD achieved $4.2 billion in MILCON project savings of which $1.6 billion had been reprogrammed to fund emergency projects—projects that did not receive the full requested appropriation, or projects needing additional funding. Regarding cost overruns, GAO (2018) stated that some differences between initial estimates and final costs for MILCON projects can be attributed to factors outside of DoD's control, such as unforeseen environmental and site conditions.

In a broader perspective, the construction project literature review identified many possible factors or causes for project cost overruns (Federle & Pigneri, 1993; Flyvberg, Holm, & Buhl, 2002; GAO, 1981; Giegerich, 2002; Harbuck, 2004; Jahren & Ashe, 1990; Thal, Cook, & White, 2010; Trost & Oberlender, 2003; Zentner, 1996). These articles span from 1981 to 2010 and cover a plethora of industry projects such as MILCON, transportation infrastructure and highways, nuclear construction, and naval facilities. Table 1 outlines a list of factors that were commonly identified in these articles as variables that can affect construction cost overruns.

TABLE 1. FACTORS AFFECTING CONSTRUCTION PROJECT COST OVERRUNS

	GAO, 1981	Jahren & Ashe, 1990	Federle & Pigneri, 1993	Zentner, 1996	Flyvberg, Holm, & Buhl, 2002	Giegerich, 2002	Trost & Oberlender, 2003	Harbuck, 2004	Thal, Cook, & White, 2010
Unforeseen Changes									
Changes in Scope/Requirements or Change Orders	√	√		√		√		√	
Changes in Schedule or Delays						√		√	
Changes in Anticipated Bid Opening Date	√								
Changes in Site Location	√								
Bidding Environment and Contractor Behavior									
Contract Bidder Interest in Project or Number of Bids	√	√	√				√	√	√
Ratio/Difference: Low Bid to Government/Engineer Estimate		√	√						√
Contractor History or Unsatisfactory Performance			√			√			
Disputes or Claims		√				√			
Bid Range: Highest to Lowest Bid			√						
Design Process									
Changes, Errors, or Ambiguity in Design	√							√	
Design Effort or Funds Available for Design			√	√		√			√
Design Complexity						√			
Design Length									√
External Factors									
Fluctuations in Labor/Material Costs or Economics	√						√		
Local Government/Permitting Agencies or Politics					√			√	
Project Features									
Construction Type		√	√		√				√
Location or Site Requirements			√		√		√		
Size		√							
Construction Duration/Length			√						
Estimation Process									
Cost Information Available							√		
Estimator Team Experience							√		
Estimate Effort or Time Allowed to Prepare Estimate							√		
Leadership									
Improper Scope Definition				√					
Lack of Estimate Accountability				√					
Strategic Misrepresentation					√				
Supervision Effort/Management Involvement			√	√		√	√		

Note. Those highlighted in green reflect three or more references indicating a similar factor in affecting cost overruns.

Even though MILCON cost estimates might represent a fraction of the total estimate associated with the actual acquisition of an MDAP, MILCON estimation still represents a vital process to investigate. Given the authors could find no published analysis concerning MILCON cost growth at a programmatic level, this study takes an empirical approach to determine whether cost growth or underruns statistically exist with respect to estimates. Specifically, this study addresses three questions:

1. What is the typical growth in program-level MILCON cost estimates for MDAPs led by the USAF?
2. Which variables or factors are statistically associated with program-level MILCON cost growth?
3. What, if any, is the association between SAR reported program-level estimates and actual project-level costs as of the current date of data?

Databases and Methods

To address these questions, two databases were created. The first captures program MILCON cost estimates from the SARs (with the implicit assumption that correct information is being recorded therein). The majority of this MILCON data initially originated from an internal Air Force Life

Cycle Management Center (AFLCMC) database of all SARs from 1966–2015. This is because the Defense Acquisition Management Information Retrieval (DAMIR) system (the current authoritative source for SARs) only contains automated SAR records from December 1997 to the present. The AFLCMC database derives from the original SAR sources dating back to 1966 (pre-DAMIR).

The AFLCMC database has 120 Air Force-led acquisition programs with 1,330 total SAR records. This study narrowed the 120 programs to 41 that contained at least one MILCON cost estimate. Seven of these acquisition programs were cancelled according to the AFLCMC database. This left the study with 34 programs to analyze. Because the internal AFLCMC database was last updated in December 2015, this study updated the SAR information for these 34 programs, resulting in 13 additional SARs. Lastly, two programs were excluded from analysis since the reporting duration from the first to last SAR was less than 12 months.

Therefore, the first database has 32 Air Force-led programs (as indicated in the SARs) with 444 associated SARs. Table 2 summarizes the data inclusions and exclusions taken to arrive at these 32 programs, while Table 3 displays these final 32 programs, commodity types, and total years of SAR reporting for each program. Of these programs, the mean and median SAR reporting times are 13 and 10.5 years, respectively. The dominant commodity type is aircraft, with 18 of the 32 programs (approximately 56%). All MILCON cost estimates and total program cost estimates were normalized from program base years to constant year 2018 using the appropriate inflation factors from the Secretary of the Air Force/Financial Management Cost and Economics (SAF/FMCE, 2018), Directorate of Economics and Business Management.

TABLE 2. PROGRAM MILCON DATABASE INCLUSIONS/EXCLUSIONS

Criteria	Δ Programs	Δ Reports	Total Programs	Total Reports	Years Included
Initial SAR data provided by AFLCMC	+ 120	+ 1,330	120	1,330	1966-2015
MILCON not reported in any SAR for the program	- 79	- 836	41	494	1966-2015
Acquisition program cancelled	- 7	- 59	34	435	1966-2015
Latest SARs added from DAMIR		+ 13	34	448	1966-2017
First to last SAR spans less than 12 months	- 2	- 4	32	444	1966-2017

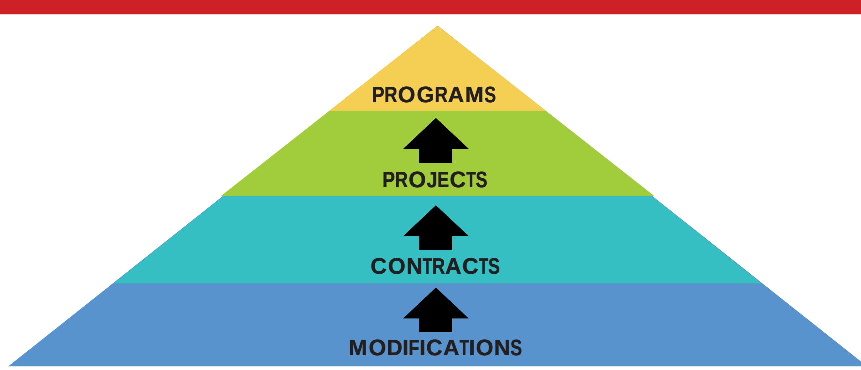
TABLE 3. 32 PROGRAMS INCLUDED IN PROGRAM MILCON DATABASE

	Weapon System Type	Total Years Reported
Short-Range Attack Missile (AGM-69A)	Missile	9
Minuteman III (LGM30G) Intercontinental Ballistic Missile	Launch Vehicle	11
A-7D Corsair II Carrier-Capable Subsonic Light Attack Aircraft	Aircraft	7
F-111 A/D/E/F Tactical Fighter Bomber	Aircraft	7
E-4 (Advanced Airborne Command Post; National Emergency Airborne Command Post)	Aircraft	10
AGM-86B (Air-Launched Cruise Missile)	Missile	9
Ground Launched Cruise Missile (BGM-109G)	Missile	12
KC-10A Aerial Refueling Tanker Aircraft	Aircraft	9
Global Positioning System Satellite Block I/II/IIA	Satellite	14
C-5B Military Transport Aircraft	Aircraft	7
Defense Meteorological Satellite Program (DMSP)	Satellite	16
Defense Support Program (DSP)	Satellite	14
Inertial Upper Stage (IUS)	Launch Vehicle	11
Advanced Cruise Missile (AGM-129A)	Missile	9
Peacekeeper (LGM-118A) Four-Stage Intercontinental Ballistic Missile	Launch Vehicle	9
C-17 Military Transport Aircraft	Aircraft	26
E-8A Joint Surveillance and Target Attack Radar System (JSTARS)	Aircraft	19
Titan IV (Complementary Expendable Launch Vehicle)	Launch Vehicle	17
F-22 All-Weather Stealth Tactical Fighter Aircraft	Aircraft	25
B-2A Spirit Heavy Strategic Bomber	Aircraft	10
Military Strategic and Tactical Relay (MILSTAR) Terminals	Electronic	8
National Airspace System (NAS)	Electronic	23
T-6A/B Joint Primary Aircraft Training System (JPATS)	Aircraft	21
C-130J Military Transport Aircraft	Aircraft	22
Space-Based Infrared System (SBIRS)	Satellite	22
C-5 Reliability Enhancement and Re-engining Program (RERP)	Aircraft	16
Global Hawk (RQ4)	Aircraft	14
C-27J Joint Cargo Aircraft (JCA)	Aircraft	10
Reaper (MQ9) Unmanned Aerial Vehicle	Aircraft	9
HC/MC-130J Personnel Recovery Aircraft	Aircraft	8
KC-46A Military Aerial Refueling and Strategic Military Transport Aircraft	Aircraft	7
Combat Rescue Helicopter (HH60W)	Aircraft	4

The second database consists of MILCON data from projects using the Automated Civil Engineer System – Project Management (ACES-PM). Examples of MILCON projects may include mission training complexes, aircraft hangars, or information system complexes. ACES-PM provides data for every individual project associated with a selected acquisition program. Key data include project cost information including appropriation, obligation, and expenditure. Since ACES-PM was fielded in 2000, this leaves a limited scope for project comparison with programs, which is acknowledged as a study limitation.

From the original 32 programs within the first database, only 11 programs included SAR estimates after 2000, when ACES-PM was fielded. Of these 11 programs, one program had more than 85% of its projects still in the design or ready-to-advertise status, and was therefore excluded from actual cost analysis. The remaining had less than 40% of the projects still in design or ready-to-advertise status. The final count of programs analyzed at the project level was 10, with nine categorized as aircraft and one categorized as a satellite. The total number of projects for these programs consisted of 216. The aircraft commodities included cargo, fighter, helicopter, tanker, trainer, and unmanned aerial vehicle programs. MILCON project data were pulled from ACES-PM and were current as of October 2018. (Note: due to confidentiality requirements, the authors cannot name the specific 10 programs of the listed 32 in Table 3 for this second database.) The information obtained from ACES-PM included contract data, contract modification data, and project data. The contract modification data were amalgamated into contract data and then subsequently amalgamated into program data. Figure 1 graphically depicts this process. As with the program SARs, all monetary values were normalized to constant year 2018 using the SAF/FMCE inflation factors

FIGURE 1. PROJECT MILCON DATABASE ROLL-UP PROCESS



The analysis incorporated a mixture of descriptive and inferential statistics to address the three primary research questions. Various reporting intervals of SAR reports were compared to the final SAR's MILCON cost estimate to analyze growth in the form of amounts and percentages. These intervals were from the start of reporting, after a quarter of reports had been submitted (25th percentile), at the median point of submitted reports (50th percentile), and after three-quarters of reports had been submitted (75th percentile). Descriptive measures consisted of the mean, median, standard deviation, maximum, and minimum values for the various reporting intervals of SAR reports.

The program-level roll-up from the Project MILCON Database was integrated with the Program MILCON Database for the 10 available programs to tie the SAR report variables to the project variables and actual MILCON costs. Cost growth was analyzed at the various stages of SAR reporting similar to the process just described. The primary difference in this cost growth analysis is that all SAR reporting stages were compared to programmed amounts, obligation amounts for projects with construction complete, and obligation amounts for projects with construction at least underway instead of the last SAR cost estimate reported.

Because both databases have relatively small sample sizes, continuous variables of percentile cost growth were converted into categorical binary variables, or dummy variables, to test for statistical dependency (via contingency table analysis). Three dummy variables were created for each measurement of cost growth to indicate (a) positive cost growth, or estimates increasing over time, (b) at least +/- 1% cost growth, or an increase or decrease of estimates over time by at least 1%, and (c) at least +/- 2% cost growth, or an increase or decrease of estimates over time by at least 2%. Table 4 lists these variables as well as other variables considered for analysis.

TABLE 4. PROGRAM MILCON DATABASE VARIABLES

Cost Growth Variables for Descriptive Statistics	Dependent Cost Growth Variables for Contingency Tables	Independent Cost Growth Variables for Contingency Tables
Growth First to Last SAR ($ and %)	Growth First to Last SAR (positive %, >|1%| and >|2%|)	Commodity Type
Growth 1st/2nd/3rd Quartile Report to Last SAR ($ and %)	Growth 1st/2nd/3rd Quartile Report to Last SAR (positive %, >|1%| and >|2%|)	Prototype
Growth Mean to Last SAR ($ and %)	Growth Mean to Last SAR (positive %, >|1%| and >|2%|)	Modification
Growth Median to Last SAR ($ and %)	Growth Median to Last SAR (positive %, >|1%| and >|2%|)	Base Year
Growth Minimum to Last SAR ($ and %)		Mean MILCON Cost to Program Cost Ratio
Growth Maximum to Last SAR ($ and %)		MILCON Cost Estimate on Last SAR
		Total Program Estimate on Last SAR

Although categorical variables can be tested for dependency through contingency tables, a relatively large sample is required for a Pearson's Chi-Squared test and the associated odds ratio. For small samples, Fisher's Exact Test is more appropriate and presents a conditional exact inference. An exact inference does not rely on assumptions that parameters hold true through infinity, but it is an exact calculation of a p-value given the data presented (Agresti, 1992). Because both of our databases had relatively small sample sizes, Fisher's Exact Test was used to test for variable dependency significance. The reader is directed to McDonald (2014, pp. 77–85) for more details regarding the use of Fisher's Exact Test. The next section highlights the results of the descriptive and inferential analysis of cost growth for MILCON MDAP programs.

Results

The first set of results ascertain the typical growth in program-level MILCON cost estimates for USAF-led MDAPs. Figure 2 displays cost growth as a percentage of total acquisition program costs from the first SAR cost estimate to the last SAR cost estimate for the 32 programs. The majority of programs (78%) show cost growth or cost savings within a 2% difference from the original estimate or a 0% cost growth, indicating minor program MILCON estimate changes over time. Table 5 displays mean and median cost growth in dollar value and percentage from the first SAR estimate, median SAR estimate, mean SAR estimate value, and median SAR estimate value to the last SAR estimate. Dissimilarities between means and medians reflect outliers (both positively and negatively) present throughout the phases of SAR reporting and when observing the dollar value or percentage. Notably, the percentage of cost growth shows less skewing and is used to analyze typical cost growth from cost estimates.

FIGURE 2. MILCON COST GROWTH AS PERCENTAGE OF TOTAL PROGRAM COST FROM FIRST TO LAST SAR COST ESTIMATES

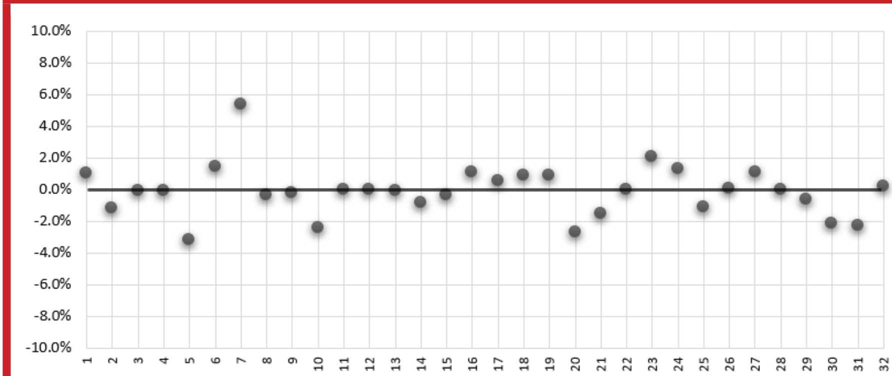


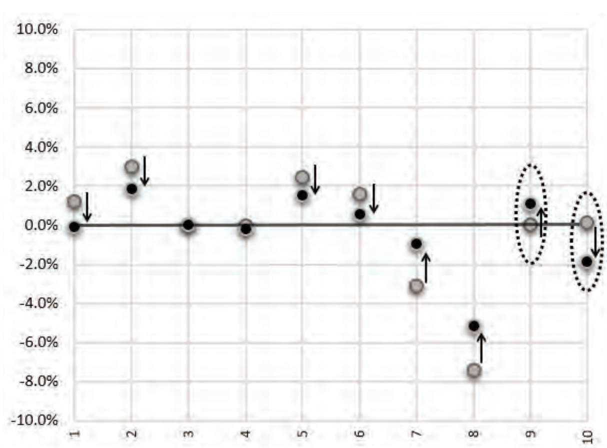
TABLE 5. DESCRIPTIVE STATISTICS OF MILCON COST GROWTH TO LAST SAR ESTIMATE (32 PROGRAMS)

	Mean ($M)	Median ($M)	Mean (%)	Median (%)
First Report to Last SAR Estimate	-$28.499	-$0.129	-0.11%	-0.03%
Median Report to Last SAR Estimate	$8.242	$0.000	-0.16%	0.00%
Mean Value to Last SAR Estimate	-$6.182	-$0.431	-0.14%	-0.04%
Median Value to Last SAR Estimate	$7.625	$0.000	-0.06%	0.00%

Utilizing a sample of 32 programs and comparing estimates to the final SAR's MILCON cost estimate yields a typical cost growth of MILCON estimates reported for USAF MDAPs on SARs, which is relatively small in comparison to the total program cost. Table 5's mean and median percentages indicate that cost growth percentages range from -0.16% to 0.00% of the total acquisition program cost reported on the last SAR. Due to the mean and median percentages leaning toward negative values, the central tendency for MILCON cost growth among MDAPs appears to be cost savings.

With respect to actual MILCON costs as reported by ACES-PM, Figure 3 highlights cost growth as a percentage to total acquisition program cost for the 10 programs in the second database. Each of the programs has two data points, which represent (a) the cost growth from the first reported SAR (gray), and (b) the cost growth from last reported SAR (black). It was anticipated that the cost growth percentages would move inward to the 0% cost growth target line from the first SAR to the last SAR as true (not estimated) MILCON costs were recorded and SAR cost estimates were updated to reflect these. This was not the case for two of these programs.

FIGURE 3. COMPARISON OF MILCON COST GROWTH FROM FIRST (GRAY CIRCLES) AND LAST (BLACK CIRCLES) SAR ESTIMATES TO ACTUAL MILCON COSTS



Note. Circled areas highlight where last estimates exceeded first estimates.

Table 6 outlines the same descriptive statistics as Table 5 with the exception of measuring cost growth against programmed and obligated amounts derived from accumulated actual projects instead of measuring cost growth against the last reported SAR estimate. The eighth listed program on the x-axis of Figure 3 has a significantly lower programmed and obligated amount than on the reported SAR estimates, which is skewing Table 5's means towards cost savings. This could be caused by unprogrammed projects that are still needed for the future or an improperly high estimate when reporting MILCON estimates in the SARs. Due to the small sample size of 10, this program was not removed for analysis. For the purpose of measuring central tendency values, the median may depict a better measurement for this dataset.

TABLE 6. DESCRIPTIVE STATISTICS OF MILCON COST GROWTH TO ACTUAL COSTS

	Mean ($M)	Median ($M)	Mean (%)	Median (%)
First Report to Programmed Amount	-$122.420	$30.394	0.43%	1.05%
Last Report to Programmed Amount	-$71.179	$31.662	0.33%	0.48%
Mean Value to Programmed Amount	-$74.819	$22.915	0.28%	0.51%
First Report to Obligated Amount (Construction Complete)	-$231.938	$3.756	- 0.37%	0.10%
Last Report to Obligated Amount (Construction Complete)	-$164.903	$20.346	- 0.47%	- 0.15%
Mean Report to Obligated Amount (Construction Complete)	-$184.337	-$5.084	- 0.51%	- 0.07%
First Report to Obligated Amount (Construction Underway)	-$198.090	$3.756	- 0.21%	0.10%
Last Report to Obligated Amount (Construction Underway)	-$146.850	-$2.774	- 0.32%	- 0.03%
Mean Report to Obligated Amount (Construction Underway)	-$150.489	-$2.017	- 0.36%	- 0.02%

Utilizing programmed amounts as a measurement of actual costs as of October 22, 2018, the median cost growth percentage from SAR reports ranges from 0.48% to 1.05% of the last reported total acquisition cost on a SAR. In dollar values, the median cost growth from SAR reports to programmed actual costs ranges from $22.92 million to $31.66 million. While the percentage of total acquisition program cost is relatively small, the dollar values appear significant when considering multiple acquisition programs that may encounter these cost growths from the reported MILCON estimate on SARs.

Both obligation amount measurements of actual costs display median central tendencies of less cost growth and even depicting cost savings. The median cost growth percentage from SAR reports ranges from -0.15% to 0.10% of the last reported total acquisition cost on the MDAP SAR. The median dollar amount of cost growth ranges from -$5.08 million to $20.35 million. These values may be smaller than the programmed amount measurement because the obligation amount does not include projects that have not begun construction yet, nor incorporate total costs for projects with construction still underway or not completely financially closed out.

Shifting to the inferential part of the analysis, those results are now presented that address the second question: which variables or factors are statistically associated with program-level MILCON cost growth?

Drawing from both the Program MILCON Database and Project MILCON Database, dichotomous (dummy) variables were utilized in contingency table analysis to identify potential dependent variables, which showed significance in Fisher's Exact Test with a *p*-value less than 0.10. This analysis was performed using JMP Pro 13, predictive analytics software that elevates statistical discovery.

A significant right tail shows that the tested cost growth is more probable if the tested independent dummy variable is indicated with a "1" than if it is a "0." For example, a right tail for the *"≥ 15 Years of SAR Reports"* dummy variable tested against positive cost growth indicates that positive cost growth is more probable if the program has 15 or more years of SAR reports. A significant left tail shows that the opposite is more probable. Continuing with the first example, a significant left tail for the *"≥ 15 Years of SAR Reports"* dummy variable tested against positive cost growth indicates that positive cost growth is less probable if the program has 15 or more years of SAR reports. For the purpose of this study, all of the contingency table tests use one-tailed hypotheses to determine directionality of the variables' dependency.

"It can be expected that greater deviations of cost growths or savings in comparison to total acquisition costs would occur on larger MILCON estimates with smaller total acquisition costs.

Table 7 illustrates the results for the 32 programs, while Table 8 highlights the results using ACE-PM data for the subset of 10 programs. (Note: For brevity, these tables reflect the results of multiple 2 x 2 Fisher Exact tests.) Significance measurements of *p*-values are marked with asterisks (*). One asterisk indicates a significant Fisher's Exact *p*-value of 0.10 or less, two asterisks indicate a *p*-value of 0.05 or less, and three asterisks indicate the highest significance with a *p*-value of 0.01 or less. Additionally, the right- and left-tailed significance is marked to show whether the independent factor tested more probable (right tail) or the opposite tested more probable (left tail). Due to the exploratory nature of this study, spurious findings are possible. Therefore, those findings with a *p*-value level less than 0.01 or by the number of significant (*p*-value 0.10 or less) Fisher's Exact tests are the ones the analysis primarily addresses in significance.

TABLE 7. TOP SIGNIFICANT FACTORS FOR COST GROWTH TO LAST SAR (32 PROGRAMS)

Table Legend:
- * *p*-value < 0.10
- ** *p*-value < 0.05
- *** *p*-value < 0.01
- L left-tail significance
- R right-tail significance

	≥15 Years of Reports	Aircraft	Missile	<0.5% Avg MILCON % to Total	>5% Avg MILCON % to Total	<$10M MILCON on Last SAR	<$50M MILCON on Last SAR	>$10B Total Program on Last SAR
First to Last (Positive Growth)	R ***					L **		
First to Last (>|1%| Growth)				L **	R *	L *	L *	R *
First to Last (>|2%| Growth)		R *		L *				R *
Q1 to Last (Positive Growth)						L **		
Q1 to Last (>|1%| Growth)				L **	R **			
Q1 to Last (>|2%| Growth)								
Q2 to Last (Positive Growth)		R ***					L ***	R **
Q2 to Last (>|1%| Growth)			R *		R **			
Q2 to Last (>|2%| Growth)								
Q3 to Last (Positive Growth)	R **	R **						R *
Q3 to Last (>|1%| Growth)		L *	R **		R **			
Q3 to Last (>|2%| Growth)		L *	R **		R **			
MED to Last (Positive Growth)	R *	R **						
MED to Last (>|1%| Growth)								
MED to Last (>|2%| Growth)								
MEAN to Last (Positive Growth)	R *	R **					L *	
MEAN to Last (>|1%| Growth)								
MEAN to Last (>|2%| Growth)								
Total Significant Contingency Tables	4	7	3	3	5	3	3	4

TABLE 8. TOP SIGNIFICANT FACTORS FOR COST GROWTH TO PROGRAMMED AMOUNTS (10 PROGRAMS)

Table Legend:
- * *p*-value < 0.10
- ** *p*-value < 0.05
- *** *p*-value < 0.01
- L left-tail significance
- R right-tail significance

	<4 Bases with Projects	<$50M Programmed for Projects	>$400M Programmed for Projects	<10 Different Companies with Project Contracts	<10,000 Contracted Performance Period Days	>$10B Total Program on Last SAR	<$10M in Contract Modifications	<$3M in Contract Modifications	<200 Contract Modifications	<50 Contract Modifications	≥75% of Projects w/Contract Modifications
First to Programmed (Positive Growth)											
First to Programmed (>|1%| Growth)	L *	L ***								L ***	
First to Programmed (>|2%| Growth)		L *	R *	L **	L *					L *	
Q1 to Programmed (Positive Growth)											
Q1 to Programmed (>|1%| Growth)		L *	R *	L **	L *					L *	
Q1 to Programmed (>|2%| Growth)	L **		R **	L ***	L **						
Q2 to Programmed (Positive Growth)											
Q2 to Programmed (>|1%| Growth)		L *	R *	L **	L *					L *	
Q2 to Programmed (>|2%| Growth)	L *		R ***	L **	L ***	R **	L *	L **	L **		
Q3 to Programmed (Positive Growth)											R *
Q3 to Programmed (>|1%| Growth)			R *		L *						
Q3 to Programmed (>|2%| Growth)											
Last to Programmed (Positive Growth)											R *
Last to Programmed (>|1%| Growth)			R *		L *						
Last to Programmed (>|2%| Growth)											
MED to Programmed (Positive Growth)											R *
MED to Programmed (>|1%| Growth)		L *	R *	L **	L *					L *	
MED to Programmed (>|2%| Growth)	L *		R ***	L **	L ***	R **	L *	L **	L **		
MEAN to Programmed (Positive Growth)											
MEAN to Programmed (>|1%| Growth)				L *							
MEAN to Programmed (>|2%| Growth)	L *		R ***	L **	L ***	R **	L *	L **	L **		
Total Significant Contingency Tables	5	5	10	9	10	3	3	3	3	5	3

One predictor variable with a high frequency of significance among the various reporting intervals of SARs was cost growth for programs that had MILCON estimates averaging more than 5% of the total program costs. All five of the significant average MILCON % dummy variables with respect to +/- 1% and +/- 2% cost growth contained significant right tails. This means that cost deviation of more than 1% or 2% of the total acquisition cost is more probable for programs averaging MILCON estimates more than 5% of the total program cost. It can be expected that greater deviations of cost growths or savings in comparison to total acquisition costs would occur on larger MILCON estimates with smaller total acquisition costs.

The other predictor variable with the most counts of significant tests among varying reporting intervals of SARs was cost growth for the aircraft commodity. The four significant aircraft commodity tests (as shown in Table 7, under the aircraft column) with respect to positive cost growth contained significant right tails, which means that positive cost growth is more probable for aircraft programs than nonaircraft programs. This could be due to higher total acquisition costs of aircraft programs compared to nonaircraft programs. The average total acquisition cost for aircraft programs was $7.8 billion, whereas nonaircraft programs averaged $1.6 billion. In summary, positive cost growth in MILCON estimates is more likely for aircraft programs, but the growth is probably less than 1% of the total program cost.

Focusing on the project level (Table 8), the predictor variable, which was one of the most frequently significant among various reporting intervals of SARs tested against programmed amounts, was cost growth for programs with more than $400 million of MILCON funds programmed for projects. All 10 of these significant tests with respect to +/- 1% and +/- 2% cost growth contained significant right tails, which means cost deviation of more than 1% or 2% of the total acquisition program cost is more probable for programs that currently have more than $400 million cumulatively programmed for MILCON projects. Perhaps a larger dollar amount programmed for MILCON projects shows increases in planned projects' costs or shows that new projects were added to the mission requirement for the acquisition program, thereby deviating SAR estimates by more than 1% or 2% of the total program cost.

A similar significant predictor variable was cost growth for programs with fewer than 10,000 cumulative performance-period days contracted for projects. This variable is a summation value from all contracts for all projects within a program, consisting of a cumulative number of days on contract for performance periods. All 10 of these significant tests with respect to +/- 1% and +/- 2% cost growth contained significant left tails, which means cost deviation of more than 1% or 2% of the total acquisition cost is more probable for programs with 10,000 or more cumulative performance-period days on contracts for all projects within the program. This finding suggests that programs requiring more performance-period days cumulatively across all projects for the program are more likely to experience changes in costs up or down from the original SAR estimates.

Another most significant predictor variable was cost growth for programs having project contracts with fewer than 10 different companies. All nine of these significant tests against +/- 1% and +/- 2% cost growth contained significant left tails, which means cost deviation of more than 1% or 2% of the total acquisition cost is less probable for programs having project contracts with fewer than 10 different companies. This finding also suggests that programs working with 10 or more companies are more likely to experience increased costs from the original SAR estimates.

Discussion and Conclusions

The study turns now to answering the three questions posed earlier. The first question concerned the typical growth in program-level MILCON cost estimates for MDAPs led by the USAF. Analysis showed that growth deviations decreased over reporting time with the mean SAR estimate being $6.2 million greater than the MILCON cost estimate on the last report. Using the median, the typical SAR estimate was only $431 thousand greater than the MILCON estimate from the last report. This equates to a cost savings of 0.04% of the total program cost on the last SAR report.

Considering cost growth from the first to the last MILCON SAR estimate, the typical amount was -$28.5 million, with the median cost growth being -$129 thousand, thereby suggesting cost savings as the typical trend for MILCON in MDAPs led by the Air Force. Utilizing a percentage to total program costs, the mean cost growth from first to last SAR is -0.11% of the total program cost, and the median cost growth across a program's span of SARs is -0.03% of the total program cost.

The second question concerned which variables or factors are statistically associated with program-level MILCON cost growth. First, aircraft commodities tend to drive positive cost growth for MILCON projects but not by more than 1% of the total program cost. Second, a higher average percentage of MILCON cost estimates reported on SARs for a program compared to the total program cost estimate can drive cost growth or savings by more than 1% or 2% of the total program cost. Third, more funds cumulatively programmed for projects within a program may drive cost growth or savings by more than 1% or 2% of the total program cost. Fourth, higher cumulative performance-period days on contracts across all projects within a program may indicate cost growth or savings by more than 1% or 2% of the total program cost. Lastly, having more companies contracted for projects within a program (greater than 9 as seen in Table 8) may drive cost growth or savings by more than 1% or 2% of the total program cost.

Several other factors appeared significant and future studies should investigate them as possible drivers to MILCON cost growth in MDAPs. The number of bases authorized for projects within a program, the number of contract modifications, and the monetary value of contract modifications may affect the size of cost growth in comparison to total program costs. Additionally, the number of years between the first and last MILCON SAR estimate and the percentage of projects with contract modifications may drive positive cost growth.

The last question concerned the association, if any, between SAR-reported program-level estimates and actual project-level costs. The Project MILCON Database with 10 programs had considerably fewer sample programs than the first database of 32 programs, but it allowed analysis of actual cost growth from projects that have been completed or at the minimum have been authorized for programming as of October 22, 2018. With various MILCON requirements for different programs and commodities, dollar values varied greatly across programs. For the purpose of analyzing the association between cost estimates on the SARs and actual costs from projects, percentages of cost growth were used. Zero percent cost growth suggests perfect estimation with no disconnect between SAR reportings and actual costs.

Analyzing the median cost growth percentage from all reporting intervals of SARs to the current programmed amount, results range from 0.48% to 1.05% of the total program cost. This suggests that the SAR estimates were slightly underestimated to what has been programmed for projects within the acquisition program. The median cost growth percentages compared to the current obligation amounts range from -0.15% to 0.10% of the total program cost. This proposes that the SAR estimates are generally closer to what has been already obligated on projects and could remain more accurate if no other obligations were made toward the programmed amounts. This course of action is highly unlikely in the authors' opinion.

Previously mentioned references share commonalities with the findings of this study. Federle and Pigneri (1993) found that the duration of the construction project can affect cost overruns for the project. The study found that the cumulative days of contracted performance were significant regarding cost growth at the programmatic level of MILCON. Four studies from Table 1 showed that the type of project or construction affected the cost overrun of the project. This study found that MILCON projects for aircraft acquisition programs were more likely to experience cost growth than the nonaircraft MDAPs when testing at the programmatic level. Table 1 also showed five studies that found changes in requirements or the presence of change orders to be an indication of cost overruns in construction projects. This study found both the monetary value and the number of contract modifications tested relatively significant for MILCON in acquisition programs. For example, in Table 8, the analysis suggests that programs with fewer contract modifications or total amount of modifications have a greater chance of staying within 2% of the final programmed budgeted amount from about the halfway completion point and onward. Lastly, three studies reported that the number of project performance locations drives cost overruns. From Table 8, we see that the number of different locations required for the program tested significant as well.

With numerous published studies regarding MILCON project overruns and general construction overrun factors, MILCON cost growth for Air Force MDAPs had yet to be analyzed in a published forum. Although using only a small representative sample of acquisition programs, this study found typical MILCON cost growth to be negative, which indicates more cost savings than cost growth across SAR MILCON estimates. The savings are typically less than 0.2% of the total program cost, which implies minimal impact to MDAP decisions regarding the weapon system as a whole. However, this finding contradicts MILCON cost overruns as reported previously by the GAO.

The early MILCON estimates from SARs compared to current programmed or obligated values for projects suggests a slight disconnect in estimating in the SAR reports. Though estimates got more accurate from the first to the last SAR for most programs, Table 6 shows the last SAR's median MILCON cost estimate was approximately $31 million underestimated compared to projects currently authorized and appropriated for the programs. Though the median cost growth percentage from last SAR to programmed amount is only 0.48% of the total acquisition program's cost, the dollar value can add up and impact budgetary decisions about scarce resources.

The analyses presented in this article help the cost community identify the characteristics of MILCON projects that have historically deviated the most from the estimate. Consequently, the cost community can make better resource decisions in allocating time and effort in developing these estimates. For example, based on analysis findings, an aircraft-associated MILCON project with more than 10,000 anticipated cumulative performance-period days should have more cost-estimating resources allocated to it. Additionally, decision makers for these types of programs should require robust justification and evidence supporting these estimates.

In conclusion, the results, in addition to the differences between the mean and median values, suggest two macro statistical findings. One, the positive median values suggest that the typical project is experiencing cost overruns, which agrees with the GAO findings from 1981 to 2018. Two, in contrast to the first macro finding, the negative mean values suggest a few projects costing much less than expected. Therefore, when pooling all the projects together, the overall program is showing a cost savings when assessing SAR cost estimates over time. Going forward, future studies should build upon this study with further data from ACES-PM to ascertain whether the trend detected here continues.

References

Acquisition Category. (2018). In *AcqNotes* [Online encyclopedia]. Retrieved from http://acqnotes.com/acqnote/acquisitions/acquisition-category

Agresti, A. (1992). A survey of exact inference for contingency tables. *Statistical Science, 7*(1), 131–177.

Federle, M.O., & Pigneri, S.C. (1993). Predictive model of cost overruns. In *Transactions of AACE International: 1993* (pp. L 7.1–7.9). Morgantown, WV: AACE International.

Flyvberg, B., Holm, M. S., & Buhl, S. (2002). Underestimating costs in public works projects: Error or lie? *Journal of the American Planning Association, 68*(3), 279–295. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/01944360208976273

Giegerich, D. B. (2002). Early warning signs of troubled projects. In *AACE International Transactions: 2002* (pp. CDR 2.1–2.8). Morgantown, WV: AACE International.

Government Accountability Office. (1981). *Why actual costs of military construction projects vary from their estimates* (LCD-81-17). Retrieved from https://www.gao.gov/products/LCD-81-17

Government Accountability Office. (2018). *Action needed to increase the reliability of construction cost estimates* (GAO-18-101). Retrieved from https://www.gao.gov/products/GAO-18-101

Harbuck, R. H. (2004). Competitive bidding for highway construction projects. In *AACE International Transactions: 2004* (pp. EST 9.1–9.4). Morgantown, WV: AACE International.

Jahren, C. T., & Ashe, A. M. (1990). Predictors of cost-overrun rates. *Journal of Construction Engineering and Management, 116*(3), 548–552.

Major Defense Acquisition Program, 10 U.S.C. § 2430. (2017). Retrieved from https://www.law.cornell.edu/uscode/text/10/2430

McDonald, J. H. (2014). *Handbook of biological statistics* (3rd ed.). Baltimore, MD: Sparky House Publishing.

Secretary of the Air Force/Financial Management Cost and Economics. (2018). Fiscal year 2018 inflation tables. In *AF Portal* [Online portal].Retrieved from https://www.my.af.mil/gcss-af/USAF/ep/contentView.do?contentType=EDITORIAL&contentId=c6925EC1756020FB5E044080020E329A9&channelPageId=s6925EC1350790FB5E044080020E329A9&programId=t6925EC2D91620FB5E044080020E329A9

Selected Acquisition Reports, 10 U.S.C. § 2432. (2010). Retrieved from https://www.gpo.gov/fdsys/granule/USCODE-2010-title10/USCODE-2010-title10-subtitleA-partIV-chap144-sec2432

Selected Acquisition Report. (2018). In *AcqNotes* [Online encyclopedia]. Retrieved from http://acqnotes.com/acqnote/acquisitions/selected-acquisition-report-sar

Thal, A. E, Cook, J. J., & White, E. D. (2010). Estimation of cost contingency for Air Force construction projects. *Journal of Construction Engineering and Management, 136*(11), 1181–1188.

Trost, S. M., & Oberlender, G. D. (2003). Predicting accuracy of early cost estimates using factor analysis and multivariate regression. *Journal of Construction Engineering and Management, 129*(2), 198–204.

U.S. Air Force. (2016). *Planning and programming military construction (MILCON) projects* (AFI 32-1021). Washington, DC: Author.

Zentner, R. S. (1996). Accurate estimates start with clear pictures. In *Transactions of AACE International: 1996* (C&S/M&C 10.1–10.6). Morgantown, WV: AACE International.

Author Biographies

Capt Emily E. Angell, USAF,

is a cost analyst at the Air Force Cost Analysis Agency, Joint Base Andrews, Maryland. She received her Bachelor of Applied Science (BAS) in Supervision and Management from Daytona State College and a Master of Science (MS) in Cost Analysis from the Air Force Institute of Technology (AFIT).

(E-mail address: Emily.Angell.1@us.af.mil)

Dr. Edward D. White

is a Professor of Statistics in the Department of Mathematics and Statistics at AFIT. He received his BS in Mathematics from the University of Tampa, MAS from The Ohio State University, and PhD in Statistics from Texas A&M University. His primary research interests include statistical modeling, simulation, and data analytics.

(E-mail address: Edward.White@afit.edu)

Dr. Jonathan D. Ritschel

is an Assistant Professor of Cost Analysis in the Department of Systems Engineering and Management at AFIT. He received his BBA in Accountancy from the University of Notre Dame, his MS in Cost Analysis from AFIT, and his PhD in Economics from George Mason University. Dr. Ritschel's research interests include public choice, the effects of acquisition reforms on cost growth in DoD weapon systems, and economic institutional analysis.

(E-mail address: Jonathan.Ritschel@afit.edu)

Dr. Alfred E. Thal, Jr.,

is an Associate Professor of Engineering Management at AFIT. He earned a BS in Civil Engineering from Texas Tech University, an MS in Engineering Management from AFIT, and a PhD in Environmental Engineering from the University of Oklahoma. Dr. Thal's research interests include innovation, project management, process improvement, facility and infrastructure management, and general environmental/engineering management.

(E-mail address: Al.Thal@afit.edu)

Inflation and Price ESCALATION ADJUSTMENTS in ESTIMATING PROGRAM COSTS: F-35 CASE STUDY

Stanley A. Horowitz and Bruce R. Harmon

Applying price indexes presents a challenge in estimating the costs of new defense systems. An inappropriate price index—one not closely linked to the inputs to the systems being costed—can introduce errors in both development of cost estimating relationships (CER) and in development of out-year budgets. To help cost analysts understand the impacts of different price indexes, this article applies two sets of price indexes to the F-35 program. Using hedonic price indexes derived from CERs, the authors isolate changes in price due to factors other than changes in quality by developing a "Baseline" CER model using data on historical tactical aircraft programs available early in the F-35 program. The focus of the work is to improve estimates of acquisition costs. All the data used in the econometric analysis are acquisition cost data. Better cost estimates should improve projections of budget requirements.

DOI: *https://doi.org/10.22594/dau.19-836.27.02*
Keywords: *Cost Estimating Relationships, CER, Cost Growth, Hedonic Price Index, Fighter Aircraft*

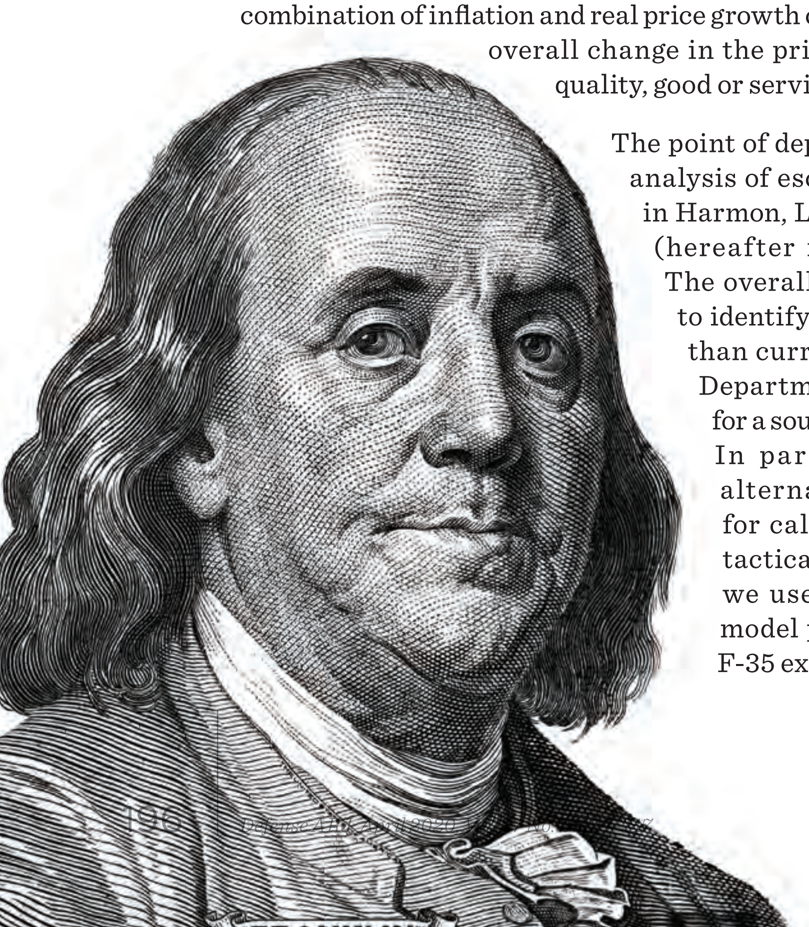
Image designed by Michael Bubar-Krukowski

THE UNITED STATES OF AMERICA
HC 16190782 A
ONE HUNDRED DOLLARS
100

Background

The application of price indexes presents a substantial challenge in estimating the costs of new defense systems. The problem is twofold. First, the analyst must use a price index when normalizing historical cost data to a common point in time (where the normalized costs are referred to as "base year" (BY) dollars in defense acquisitions. These data can then be used to help estimate the costs of future systems. Second, as budget requirements for future acquisitions are in "then-year" (TY) dollars (or more generally, "nominal" dollars), BY dollar estimates must be escalated to TY dollars using a price index. Using an inappropriate price index can introduce errors in both of these steps. In this article, we apply two sets of price indexes to a cost estimating problem—the F-35 Joint Strike Fighter (JSF) procurement program. The purpose is to help cost analysts and others involved in the acquisition process understand the impacts of different price indexes and to provide guidance in their choice.

In general, price indexes isolate changes in price due to factors other than quality changes. These changes can be categorized into changes due to general inflation, changes in the overall price level in the economy (subsequently often just called "inflation"), and real price growth—price changes for a particular class of products relative to inflation. The combination of inflation and real price growth constitute price escalation—overall change in the price of a specified, constant quality, good or service.



The point of departure for this work is the analysis of escalation indexes presented in Harmon, Levine, and Horowitz (2014) (hereafter referred to as "D-5112"). The overall goal of that research was to identify a price index that is better than current indexes at meeting the Department of Defense (DoD) need for a sound basis for cost estimation. In particular, we explored an alternative "hedonic" approach for calculating price indexes for tactical aircraft. In this analysis, we used updates to the hedonic model presented in D-5112 in the F-35 example.

Hedonic analysis involves estimating the relationship between the characteristics of a product being purchased and the price of the product. It allows analysts to understand the extent to which price variations over time are due to changes in the characteristics of the product and the extent to which price variations reflect changes in the price of a consistent product. Price escalation is meant to measure changes in prices for consistent products. Hedonic analysis allows one to determine how much of a gross change in price is due to changes in product characteristics and how much is due to escalation in the price of a consistent product.

"The combination of inflation and real price growth constitute price escalation — overall change in the price of a specified, constant quality, good or service.

The main conclusion is not that hedonic development of product-specific escalation indexes should be used universally. Rather, it is that cost analysts should be attentive to possible differences between inflation and escalation, and the implications of using inflation as a proxy for escalation when it is not a good one.

The F-35 Cost Estimating Problem

The F-35 program has experienced significant program cost growth since its October 2001 Milestone (MS) B decision that initiated Engineering and Manufacturing Development (EMD). A substantial portion of this cost growth has been in its unit recurring flyaway (URF) cost, with much of this attributed to the incorrect application of price indexes (Arnold et al., 2010). Given the tactical aircraft focus of the Institute for Defense Analyses (IDA)'s previous hedonic models, the F-35 makes for a suitable case study.

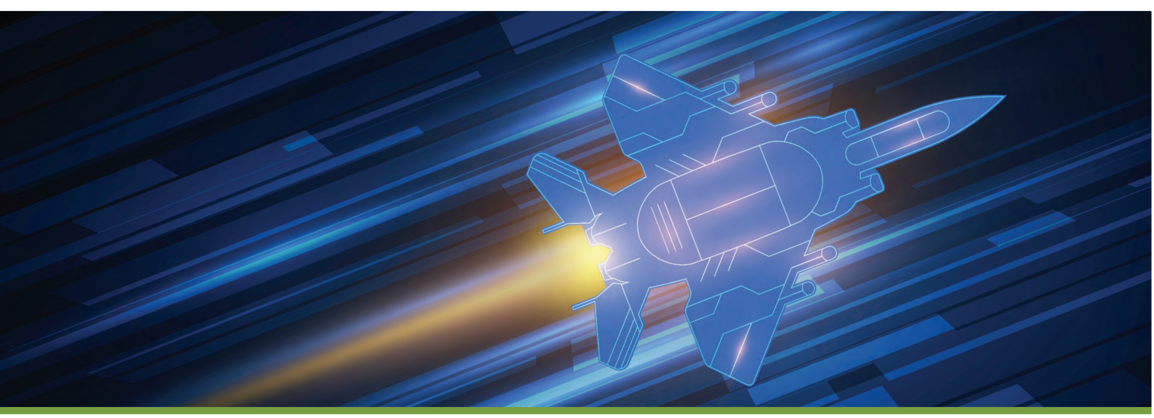
We used information available at MS B to develop models for exploring the effects of escalation adjustments on estimated F-35 URF costs. The resulting estimated costs can then be compared to several benchmarks, including cost estimates produced by the JSF Program Office (JPO) and observed URF costs for F-35s procured from 2007 through 2013. From this exercise, we draw lessons for future cost estimating practice. Although the authors did not directly consult the JPO in developing this analysis, we had substantive interactions with the JPO over the years on topics related to F-35 costs and cost growth (for example, Arnold et al., 2010).

Hedonic Price Index Models for Tactical Aircraft

In this section, we review past work on hedonic price index models and present updates developed specifically for the F-35 cost estimation problem. The estimation of the hedonic indexes for tactical aircraft builds upon tools that cost estimators have used for years. The basic setup is:

nominal system unit price = f(year, quality variables, other control variables)

The hedonic index application has commonalities with cost estimating relationships (CER), which also model system costs as a function of quality variables and quantity relationships (to capture learning and production rate effects). The hedonic index estimation differs from past cost estimating practice in that the price index is estimated simultaneously with other model parameters and the dependent variable is expressed in TY (nominal) dollars. In CER development, adjustments needed to normalize historical cost data to BY dollars used as the dependent variable are often performed using a general deflator based on an index of overall inflation. An example of such an index is the National Defense Budget Estimates "Green Book," published by the Office of the Under Secretary of Defense (Comptroller), or OUSD(C).[1] For commodities such as tactical aircraft, a given observed price may reflect both inflation and relative price changes. A key reason for relative price changes is that the underlying mix of goods and services that goes into producing military aircraft is different from the broader mix that drives overall inflation. Another reason is variation in the quantity purchased. Typically, normalization to a common quantity (e.g., first unit or 100th unit)[2] is performed using BY dollars prior to CER estimation. Thus, another unique aspect of our modeling is the simultaneous estimation of CER and learning curve parameters, as well as production rate effects.



The hedonic analysis described in D-5112 used the direct time-dummy variable approach formulated by Triplett (2006), an early developer of hedonic analysis. The update to the earlier analyses also used this approach, along with the same set of explanatory variables (Table 1). Five quality variables describe the aircraft, two quantity variables capture the cost effects of learning and production rate, and the time-dummy variables identify each fiscal year in which the aircraft were procured. The hedonic index is defined by the expression $b_t^{D_t}$,where D_t is a 1/0 dummy variable with a value of 1 for fiscal year t, and b_t is the estimated index for that year. BY dollars are calculated as BY $\text{dollars}_t = \frac{\text{TY dollars}_t}{b_t^{D_t}}$. In the application of the Green Book index, the index (where the BY value equals 1) replaces the $b_t^{D_t}$ expression in calculating BY dollars.[3]

TABLE 1. EXPLANATORY VARIABLES
Quality variables
Empty weight in pounds
Maximum speed in knots
Advanced materials as percentage of structure weight
Dummy variable for 5th generation aircraft[a]
Dummy variable for Short Take-Off and Vertical Landing (STOVL) aircraft[b]
Quantity variables
Cumulative production
Lot size (number of aircraft produced in a year)
Time-dummy variables

[a]5th-generation aircraft are characterized by stealth, internal weapons carriage, avionics with information fusion and support of net-centric operations. In the D-5112 sample, the F-22 and F-35 A/B/C were classified as 5th-generation aircraft; in the update, we added the F-117. [b]The A/V-8B and F-35B aircraft with STOVL capability needed for operations from small aircraft carriers and short unimproved airfields.

The database used in regression estimation contains pooled cross-section and time-series data, often called "panel data" in the econometrics literature, where each panel is an aircraft program. The cost metric of interest is the URF cost. In D-5112, the time series included 40 fiscal years (FYs 1973–2013), with 2012 as the base year; the cross-sections (panels) consisted of the 11 aircraft programs' original designs plus derivatives of these designs from series or block changes. In model estimation, the quality changes associated with the series/block changes are captured in the changes in empty weight over time. Production rate effects were calculated by estimating the annual fixed cost for each program.[4] Learning spillovers due to commonality between the EA 18G and F/A 18E/F and between the F-35 variants were included in the model.[5] We also accounted for loss of learning due to series/block changes.[6]

Updating Hedonic Price Index Models for Tactical Aircraft

For the current analyses, we made multiple changes to the previous work, including several versions of the model meant to capture different aspects of the F-35 cost estimating problem. Our primary focus is on the "Baseline" F-35 model; the intent was to use information available for the MS B (October 2001) cost estimate. As the FY 2002 budget materials were released earlier in 2001, we used data through FY 2002. Eliminating the newer data means that we dropped the EA-18G from the data sample along with the three F-35 variants (F-35A, F-35B, and F-35C); also, the F-22A program is truncated. This left the F-22A as the sole 5th-generation aircraft with only two data points (2001 and 2002). In order to include another 5th-generation aircraft, we added the F-117A[7] to the updated sample.

In addition to the original series aircraft, derivative follow-on aircraft were relevant for the F-14A (F-14A+ and F-14B), F-15A (F-15C, F-15C Multistage Improvement Program, and F-15E), F-16A (F 16C Blocks 25/30/50), F/A-18A (F/A-18C and F/A-18C Night Attack), and A/V-8B (A/V-8B Night Attack and A/V-8B Radar).[8] As these derivative aircraft were produced serially, they were included in the same panel as the original design. We use 2002 as the BY price index; this was also the BY for the F-35 MS B estimates and the associated URF goal.

In addition to the Baseline model, we estimated other model variations to address different aspects of the F-35 cost estimating problem. The Green Book model replaces the statistically estimated hedonic index with the procurement budget index published in the FY 2002 National Defense Budget Estimates. This would be more typical of the approach used in CER estimation. All hedonic model variations follow the "Full CER Hedonic Model" approach from D-5112. We also estimated a "Full Information" model, using complete actual data through 2013. The purpose of that model is to provide a close comparison with the model included in D-5112.[9] A slight modification of this model excludes the F-35—the "Full Information less F-35" variation provides hedonic index values through 2013 without using any information from F-35 program cost experience. Unlike in the D-5112 and Full Information models, the Baseline model does not generate price index values from 2003 through 2013; instead, a methodology is presented in which model results are extrapolated to produce estimated index values through 2013. The "Full Information" models are presented to help us understand how costs actually evolved after 2002. They do not contribute to the quantitative analysis of how better projections could have been made in 2002.

Model Estimation and Results

This section presents regression results for the different model variations. Comparisons are shown between these models and the Full CER Hedonic Model described in D-5112. As the functional form of the models is the same, we do not repeat the detailed exposition presented in D-5112—instead, the differences in the regression results are highlighted.

We estimate the model parameters using maximum likelihood estimation. The models are fit using the nonlinear optimization package within Microsoft Excel. The distribution of errors is assumed to be multiplicative/lognormal—this is analogous to estimating a log-log regression using linear regression.

Table 2 presents key regression metrics and parameter estimates for the five models.

TABLE 2. COMPARISON OF REGRESSION RESULTS

Metric	FY 1973–FY 2002		FY 1973–FY 2013		
	Baseline	Green Book	D-5112	Full Information	Full Information Less F-35
Price index used	Hedonic	Green Book	Hedonic	Hedonic	Hedonic
Number of data points	117	117	150	159	143
Parameters estimated	41	11	55	54	53
Adjusted R^2	0.97	0.84	0.97	0.97	0.97
Standard error	0.09	0.20	0.09	0.09	0.09
Quality coefficients					
Empty Weight[a]	0.78	0.75	0.83	0.84	0.81
Maximum Speed[a]	0.29	0.08	0.30	0.28	0.26
Advanced Materials[b]	1.95	1.86	1.67	1.63	1.77
5th-Generation[b]	1.24	1.44	1.11	1.16	1.14
STOVL Capability[b]	1.00	1.00	1.10	1.05	1.00
1st unit cost (T1), FY02$M					
F-14A	240	119	271	261	261
F-15A	196	94	218	207	209
F-16A	97	50	109	104	104
F/A-18A	140	73	158	153	153
F-117A	187	128	189c	192	192
A/V-8B	81	49	94	88	87
F/A-18E	197	101	219	210	213
F-22A	370	212	368	367	365
F-35A	235[c]	144[c]	233	234	233[c]
F-35B	246[c]	154[c]	267	259	246[c]
F-35C	278[c]	169[c]	276	277	277[c]
Learning curve slope	84.5%	88.1%	83.9%	84.1%	84.1%
Escalation growth rate: 73–02	7.4%	4.5%	7.6%	7.5%	7.5%
Escalation growth rate: 02–13	N/A	2.1%[d]	3.6%	3.5%	3.2%

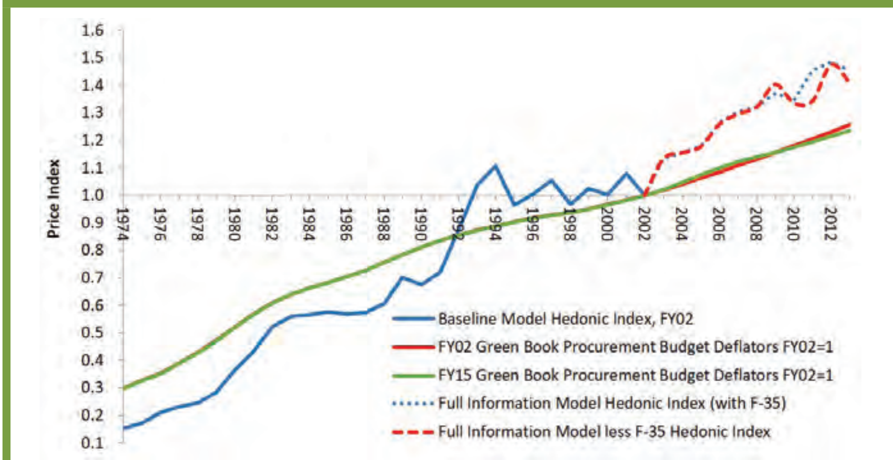
[a] The coefficients on these variables enter the model in the form xb. [b] The coefficients on these variables enter the model in the form bx. [c]Out-of-sample estimates. [d]Extrapolated from projections in the FY 2002 Green Book.

The regression fits for the models in which a hedonic index is estimated are comparable. Restricting the index to that prescribed in the 2002 Green Book results in a significantly worse model fit. The learning curve slopes are similar for the hedonic models, but the slope is substantially shallower for the Green Book model (88% vs. 84%)—again, this is consistent with the embedded underestimation of escalation when normalizing the data to constant year dollars. Systematically lower constant dollar costs in the earlier years mean that the estimated learning effect is blunted. The steeper learning slope is also consistent with values of fighter/attack aircraft learning curve coefficients estimated using labor hour costs in previous studies (Harmon, 2010; Resetar, Rogers, & Hess, 1991; Younossi, Kennedy, & Graser, 2001).

Coefficients on weight, speed, and materials composition are relatively stable across the models and are consistent with those reported in past CER studies (Harmon, 2010; Harmon, Nelson, & Arnold, 1991; Resetar et al., 1991; Younossi et al., 2001). Unit prices increase with weight, maximum speed, and more advanced materials. The one exception is the speed variable in the Green Book model—as the aircraft with the highest maximum speeds (the F-15 and F-14) appear early in the sample, the underestimates of aircraft inflation associated with the model tend to bias its parameter estimate downward. Estimates for the 5th-generation and short take-off and vertical landing (STOVL) aircraft effects change some when the F-117 is introduced into the sample. The 5th-generation factor increases from 1.11 to 1.16, while the STOVL factor decreases from 1.10 to 1.05. When the F-35 is excluded from the regression, the STOVL factor goes to 1.00—this reflects the influence of the F-35B (which is a 5th-generation STOVL aircraft), with the A/V-8B the only other STOVL aircraft in the sample.[10] The range of 5th-generation premiums for the hedonic models is generally consistent with values from an earlier IDA paper on the cost of stealth (Nelson, Harmon, Bontz, & Devers, 2001), although the 1.24 factor for the Baseline model is somewhat higher than expected. The 1.44 factor estimated with the Green Book model is clearly too high—the bias is a mirror image of the maximum speed coefficient, where underestimated escalation and newer 5th-generation aircraft interact. Thus, if there is a relationship between time and the values of the quality variables, a systematic bias in the price escalation used will result in a related bias in the coefficients on the quality variables. Also note that the analogous cost drivers in the historical studies are usually estimated using labor hour data, eliminating the possibility of bias from price escalation.

Estimated first unit variable costs (T1s) for each initial Mission-Design-Series (MDS) (usually the "A" series) are calculated using the quality coefficients, the regression intercept, and the values of the quality variables for each MDS. Table 2 shows the T1s for all relevant MDS, including "out-of-sample" cases in which the MDS was not used in model estimation. These cases are the F-35 variants, with the exception of the F-117A, which was not used in estimating the D-5112 model. For the models using the hedonic indexes, the out-of-sample estimates were close to the values calculated using the models that included those MDS. The exception is the F-35B, where the more complex STOVL capabilities were not well captured in the models not using the F-35 data. Even in this case, the out-of-sample F-35B T1s are only around 5% lower than the estimates from the other hedonic models. The T1s from the Green Book models are all substantially lower than those from the hedonic models. This is consistent with the shallower learning curve for the Green Book model, where the real prices of the initial lots are systematically underestimated because of biased escalation. Figure 1 shows the escalation indexes for a selection of the regression models.[11] Also included for comparison is the FY 2015 Green Book index.

FIGURE 1. COMPARISON OF PRICE INDEXES



These indexes are portrayed in the price growth rates shown in Table 2. Of most interest for the F-35 estimating exercise are the Baseline and Green Book models. The other models are included for comparison purposes as well as to provide escalation estimates through 2013. No 2002–2013 escalation is associated with the Baseline model; one of the goals of our analyses is to suggest a methodology for extrapolating forward growth rates from the Baseline model hedonic index. Also note how little the Green Book inflation changed from the FY 2002 forecasts (including extrapolations from FY 2007 to FY 2013) through the actuals reflected in the latest FY 2015 values.

Normalizing the data using the Green Book index results in a constant-dollar cost data set and associated model that systematically underestimates costs in the earlier years and overestimates costs in the later years. In addition to introducing bias in the quality parameters, using the Green Book index also results in a shallower learning curve. This behavior is not evident in the Baseline model. Clearly, in both the distortion of the parameter estimates and the systematic errors in estimating the actual data, a naïve application of price indexes can be problematic.

F-35 Cost Estimating Applications

F-35 URF estimates generated by the Baseline and Green Book models are compared against three sets of benchmarks:

- MS B program cost estimates and subsequent cost estimates associated with the 2009 "Nunn-McCurdy" unit cost breach,[12] in BY 2002 dollars
- Actual TY dollar budget values for the 2008–2013 FY lots
- The latest program cost estimate as reported in the December 2013 Selected Acquisition Report (SAR), reported in TY dollars

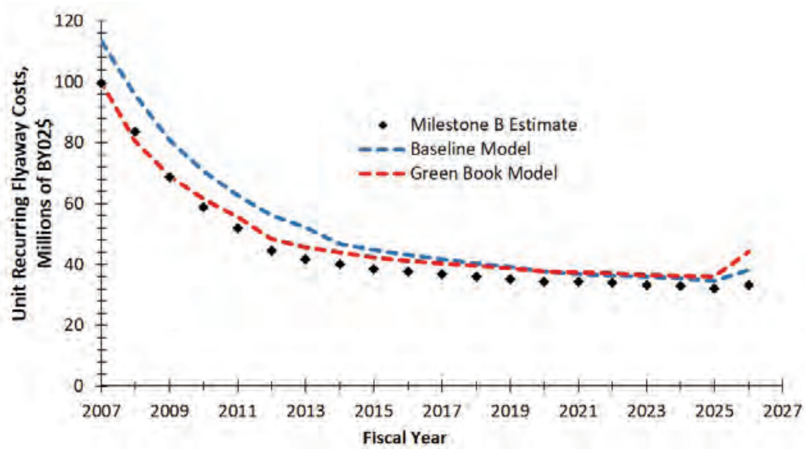
To do this, the Baseline and Green Book models are used to produce BY 2002 cost estimates for each scenario. For comparisons with the TY actuals and estimates, we use either an index calculated from the historical hedonic index ("projected hedonic index") or the Green Book index. The BY 2002 estimate comparisons demonstrate the effect of different price indexes on the structure of the CER model, while the TY dollar estimates also show the effect of the different indexes in projecting BY estimates forward. Budget projections reflect expected costs. When TY costs are underestimated, budget projections will be too low.

> **"Weight growth in all F-35 variants was a driver of cost growth between MS B paper designs and the current designs reflecting the aircraft as produced.**

F-35 MS B and Nunn-McCurdy Breach Estimates

MS B estimates are the initial benchmarks used for budgeting and for calculating program cost growth. As both models take into account production rate and learning, they can produce an analog of the MS B estimate using the quantities and production schedule associated with the October 2001 program. The IDA model estimates in this application do not carry explicit assumptions regarding future (post-2002) escalation—they are in BY 2002 dollars as directly calculated by the model. Figure 2 shows comparisons between the MS B URF estimates (all F-35 variants combined) and those generated by the Baseline and Green Book models using MS B input values.

FIGURE 2. COMPARISONS OF MILESTONE B AND MODEL ESTIMATES FOR ALL F 35 VARIANTS



The estimates from the two models converge as a result of the shallower learning slope of the Green Book model. Both models produce estimates above the program MS B URF estimate. However, they are substantially below the 2009 SAR estimates that triggered the Nunn-McCurdy breach. Many elements of F-35 cost growth are not captured in the earlier model estimates. Data from Arnold et al. (2010) allow us to isolate and deconstruct the URF portion of the cost growth.[13]

Weight growth in all F-35 variants was a driver of cost growth between MS B paper designs and the current designs reflecting the aircraft as produced. Almost all weight growth attributable to redesign was evident by the 2009 Nunn-McCurdy breach and reflected in the production lots.[14] As empty weight is an input to the models, the weight growth must be taken into account when comparing model outputs to the MS B estimates and subsequent cost growth. Another change affecting cost model application is the decrease in commonality between variants (F-35A/F-35B/F-35C) since MS B. Current commonality is reflected in the "spillover" parameter affecting learning across variants estimated as part of the Full Information model. The cost effects of commonality have been estimated by the JSF program using a detailed assignment of the learning quantities depending on common component applications. As we cannot reproduce such a detailed analysis, we make use of the spillover parameter instead—for the MS B estimate, we increase its value to reflect higher commonality assumed at that point.

Table 3 shows the MS B URF estimate, a buildup of cost growth drivers to the 2009 estimate as derived from Arnold et al., and comparisons with the model estimates. Model estimates presented include calculations with MS B inputs, and with inputs reflecting contemporary values for empty weight and commonality (learning spillovers).

TABLE 3. F 35 PROGRAM GROWTH TRACK FROM MILESTONE B TO 2009 SAR AND MODEL ESTIMATE COMPARISONS

Metric	F-35 Program URF Cost, in Millions of BY 2002$			
	Cost Growth Increment	Cumulative Cost Growth	Baseline Model	Green Book Model
MS B Estimate		40.7		
Major Subcontractor Fee	1.5	42.2		
Change in Materials Manufacturing Efficiency	3.0	45.2	47.3[a]	44.6[a]
Design-Negated Affordability and Production Efficiency Plans	3.0	48.2		
Aircraft Weight Growth	3.0	51.2	52.1[b]	48.4[b]
Change in Buy Profiles (2009 SAR)	2.5	53.7		
Escalation Rates (2009 SAR Estimate)	7.0	60.7		

[a]MS B weight and commonality. [b]Contemporary weight and commonality.

We orient the model outputs in the table to reflect how they relate to the cost growth elements from the MS B estimates. Elements that represent underestimates based on a departure from business as usual (i.e., the historical database) are included above the model estimates calculated with the MS B weight and commonality assumptions. The estimates reflecting updated weight and commonality are in line with cost growth through the Aircraft Weight Growth row. Not accounted for in this application of the IDA model estimates are cost increases due to buy profile changes (a reduction in quantities and a stretch-out of the procurement schedule) and a misapplication of escalation rates for future costs.[15] The last cost growth element is informative of our research question. Instead of using contractor-specific labor rate escalation, the JPO used OUSD(C) Green Book inflation when converting constant dollar estimates to TY dollar estimates.

From Arnold et al. (2010, p. 12):

> However, at the time of Milestone B, the Defense Contract Management Agency (DCMA) and Lockheed Martin had already agreed to a Forward Pricing Rate Agreement (FPRA) that increased rates more than the OUSD(C) escalation indices...therefore, the fully burdened labor rates turned out to be significantly higher than those used in the JPO Milestone B [estimate].

The preferred methodology reflected in the 2009 JPO cost estimate is to escalate estimated constant year costs to TY dollars using escalation rates appropriate to the different cost elements. The OUSD(C) index is then used to de-escalate the TY dollars to BY dollars, which are, in turn, reported in the SARs and used as a basis for cost growth calculations. This correction of the original methodology is responsible for the $7 million unit cost growth due to escalation rates shown in Table 3. Analogous steps are not reflected in the BY 2002 model estimates in Table 3; thus, the constant year model estimates presented for comparison are conceptually similar to the JPO's MS B estimates, reflecting the same error.[16] The next sections focus on model-generated TY estimates in the context of more up-to-date F-35 estimates.

F-35 Actual Budget Values

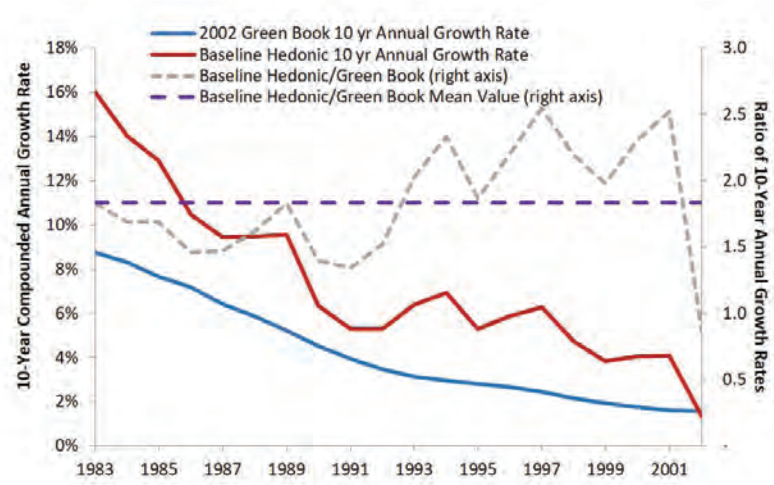
This section compares model-generated estimates with actual historical costs. The emphasis is on the results from the Baseline model. The budget experience is taken from Navy and Air Force President's Budget (PB) Justification Books, "Exhibit P-5, Cost Analysis" sheets. In collecting these data, we used the values in the latest PB in which they appeared; e.g., for the FY 2013 lot, data presented in the FY 2015 PB submission were used. For this exercise, the unadjusted TY URF values were used.

For the Baseline model, we developed the projected hedonic index to generate TY estimates through FY 2013. We also included results for the Green Book model, where the FY 2002 Green Book index (including extrapolations through FY 2013) is applied. The hedonic indexes generated by the Full Information and Full Information Less F-35 models were used for comparison purposes only. For model inputs, we used contemporary values for the quality variables and the procurement profiles reflected in the budget data.

The projected hedonic index is based on the relationship between the FY 2002 Green Book and Baseline hedonic indexes; it has the advantage of using only information through 2002 while taking into account the systematically higher escalation rates associated with the hedonic indexes vs. the Green Book rates.

To calculate the projected hedonic index, we first define the relationship between the Green Book index and the hedonic index using data through 2002, as estimated by the Baseline model. Given the year-to-year volatility of the hedonic index, we do this by comparing 10-year compounded annual growth rates. These data are shown in Figure 3.

FIGURE 3. COMPARISON OF BASELINE HEDONIC AND GREEN BOOK INDEX GROWTH RATES, 1983 2002



Examination of the data shows that the hedonic and Green Book indexes relate to one another most consistently through a multiplicative factor vice an additive adjustment. We use the calculated average ratio (mean value) of 1.83 shown in the figure as a conversion factor on the 2003–2013 Green Book values to arrive at the projected hedonic index. This is shown along with the other indexes in Figure 4.

FIGURE 4. COMPARISON OF HEDONIC AND GREEN BOOK INDEXES, 2002 2013

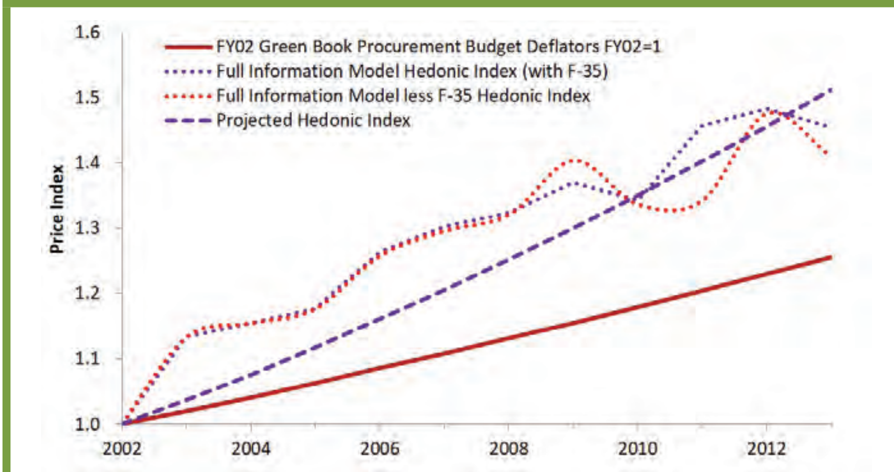


Figure 5 compares the URF estimates associated with the two models and three escalation index assumptions with the budget actuals.

FIGURE 5. COMPARISON OF MODEL ESTIMATES WITH BUDGET ACTUALS, ALL F 35 VARIANTS

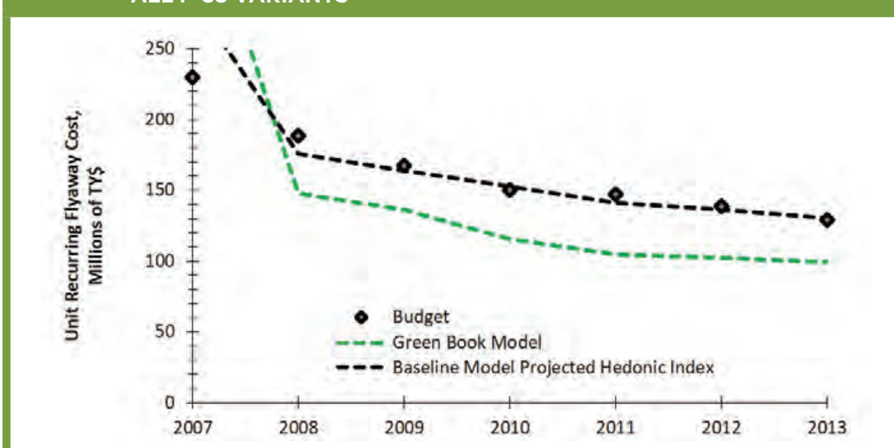


Table 4 compares the estimated URF costs with the budget actuals calculated for the 2007–2013 budget years, broken out by variants.

The results show that the Baseline model estimates, when projected forward using the hedonic index, come close to the actual budget values for 2007–2013; estimates depending on the Green Book index consistently underestimate the budget URF costs. However, the Baseline model tends to miss the costs for the individual variants, with the F-35B underestimated and the F-35C overestimated. This result is consistent with the differences in parameter estimates between the Baseline and Full Information models, which are, in turn, a result of the more complex STOVL implementation of the F-35B relative to the A/V-8B that is not completely captured in weight differences.

TABLE 4. COMPARISON OF ESTIMATES OF 2007 2013 URF COSTS, MILLIONS OF TY$

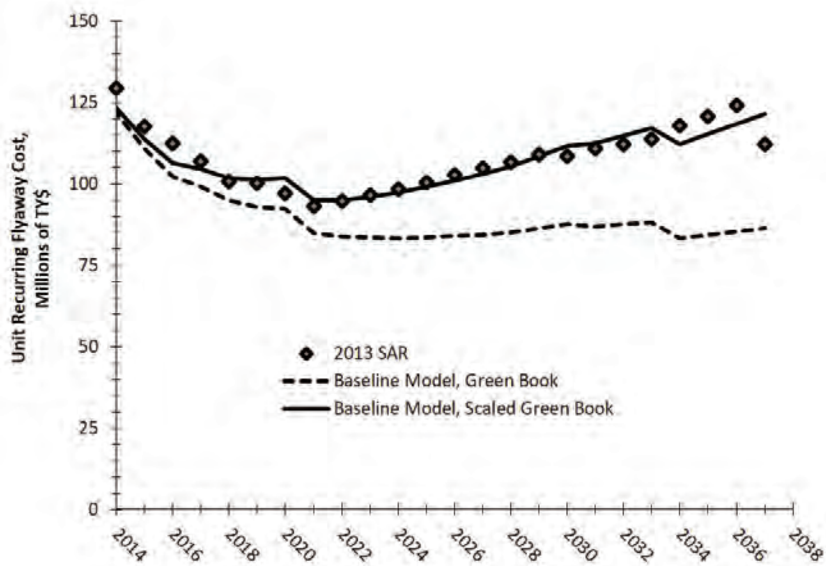
Variants	Actual Budget	Baseline Model, Projected Hedonic Index	Green Book Model and Index
All Variants	149	147	115
F-35A	139	137	110
F-35B	160	152	121
F-35C	167	175	124

F-35 2013 SAR/PB 2015 Estimates

This section takes a somewhat different approach to the F-35 estimating problem. The question we want to answer is this: which scaling of the FY 2015 Green Book index results in the closest fit to the latest JPO estimates? While the previous F-35 estimating exercises took the data available in 2002 as given, in this case we assume contemporary data for escalation projections. To address this question, we use only the Baseline model with the projected hedonic index as presented earlier. For 2014 onwards, we scale the FY 2015 Green Book index by a multiplier analogous to the factor used to calculate the projected hedonic index. The multiplier is determined by scaling the Green Book index such that the model-estimated totals for 2014–2037 are the same as those reported in the SAR. The resulting factor is 1.75—comparing directly with the 1.83 factor used to calculate the projected hedonic index. This analysis is shown graphically in Figure 6.

If the estimates are projected using the unadjusted Green Book index, the 2014–2037 URF estimate is $88 million versus $106 million reported in the SAR. This shows the impact of the different indexes on projected costs, isolated from their influence on defining the CER model.

FIGURE 6. COMPARISON OF MODEL ESTIMATES WITH THE 2013 SAR ESTIMATES, ALL F 35 VARIANTS



Summary and Conclusions

This article describes different approaches to estimating expected price growth in defense system costs. The comparison of cost estimates based on escalation predictions derived from hedonic modeling with F-35 budget actuals through FY 2013 is particularly interesting. Although the model inputs reflect the latest F-35 aircraft characteristics and program parameters, in terms of the structure of the model and escalation projections, the models are defined by the information that was available at MS B. As the hedonic index is directly estimated only for the historic period, a methodology to project forward escalation rates associated with the hedonic index is applied. This example shows the close correlation between the Baseline hedonic model estimates and the budget actuals. The lower estimates from the Green Book model are due to two factors: the underestimates of escalation from FY 2002–FY 2013 and biases introduced into the model parameters because of underestimates of escalation in the historical period.

Looking out to FY 2037, we find that projecting escalation using our approach closely mimics the more detailed buildup of input-specific escalation rates used by the JPO. This is in contrast to projections using Green Book escalation, which result in an $18 million underestimate in unit costs.

We demonstrate the effect of different escalation methodologies using top-level CER models. Cost analysts usually build up their estimates from a more detailed level. However, issues regarding the proper application of price indexes, for both normalizing historical data and making projections, are equally valid in more typical cost-estimating environments. For example, rates of price growth for raw material inputs, propulsion systems, electronic components, and labor inputs are likely to be different from those of general inflation. In our last example, we calculated overall escalation rates implied in the JPO estimates for the rest of the F-35 program; we found these escalation rates to be consistent with those projected using values from the historical hedonic price index.

The main point is not the superiority of hedonic development of escalation indexes. Rather, it is that cost analysts should be attentive to possible differences between inflation and escalation, and the implications of using inflation as a proxy for escalation when it is not a good one.

References

Arnold, S. A., Byun, J. S., Cloud, H. A., Gallo, A. O., Gonwa, M. W., Harmon, B. R., … Bronson, P. F. (2010). *WSARA 2009: Joint Strike Fighter root cause analysis* (IDA Paper P-4612). Alexandria, VA: Institute for Defense Analyses.

Harmon, B. R. (2010). *Cost estimating techniques for tactical aircraft manufacturing labor* (IDA Paper P-4490). Alexandria, VA: Institute for Defense Analyses.

Harmon, B. R., Levine, D. B., & Horowitz, S. A. (2014). *Inflation adjustments for defense acquisition* (IDA Document D-5112). Alexandria, VA: Institute for Defense Analyses.

Harmon, B. R., Nelson, J. R., & Arnold, S. A. (1991). *Unit cost implications of new materials: Preliminary analyses of airframe experience (revised)* (IDA Document D-908-REV). Alexandria, VA: Institute for Defense Analyses.

Nelson, R., Harmon, B. R., Bontz, R. J., & Devers, W. C. (2001). *Cost estimating for modern combat aircraft: Adjusting existing databases and methods to include low-observable cost considerations* (IDA Paper P-3528). Alexandria, VA: Institute for Defense Analyses.

Office of the Under Secretary of Defense (Comptroller). (2001). *National defense budget estimates for FY 2002.* Retrieved from https://comptroller.defense.gov/Portals/45/Documents/defbudget/Docs/fy2002_greenbook.pdf

Office of the Under Secretary of Defense (Comptroller). (2014). *National defense budget estimates for FY 2015.* Retrieved from https://comptroller.defense.gov/Portals/45/Documents/defbudget/fy2015/FY15_Green_Book.pdf

Resetar, S. A., Rogers, J. C., & Hess, R. W. (1991). *Advanced airframe structural materials: A primer and cost estimating methodology* (R-4016-AF). Retrieved from https://www.rand.org/pubs/reports/R4016.html

Triplett, J. E. (2006). *Handbook on hedonic indexes and quality adjustments in price indexes: Special application to information technology products.* Paris, France: Organisation for Economic Co-operation and Development.

Younossi, O., Kennedy, M., & Graser, J. C. (2001). *Military airframe costs: The effects of advanced materials and manufacturing processes* (MR-1370-AF). Retrieved from https://www.rand.org/pubs/monograph_reports/MR1370.html

Endnotes

[1] The National Defense Budget Estimates is commonly referred to as "The Green Book"—a reference source for data associated within current DoD budget estimates.

[2] Although unit prices are also sensitive to production rate, this typically has not been taken into account.

[3] If the values for the Green Book escalation index were the same as the hedonic price index, all other model parameters would also be the same.

[4] Fixed costs for each program were estimated as a function of the estimated maximum variable costs.

[5] Learning spillovers are captured by estimating parameters that assign some portion of the cumulative quantity across related aircraft (e.g., F-35A, F-35B, and F-35C). Theoretically the parameter has a lower bound of 0 (no spillovers found, learning only occurs within a single variant) and 1 (all variants proceed down a single learning curve). Although these parameters can be estimated analytically based on aircraft characteristics, we estimated them as part of overall model estimation (for the F-35 we used data through 2013).

[6] This is accounted for by a parameter that decrements cumulative quantity at each block change. A common parameter is estimated across the entire data sample, where a block change results in a decrement of learning quantity of 37%.

[7] Stealth technology is the prime feature of 5th-generation aircraft and the F-117. The F-117 differs from newer examples of 5th-generation aircraft in having less sophisticated electronic systems.

[8] Military aircraft are described by Mission-Design-Series (MDS). For the F-14A, for example, the mission is fighter (F), the design is 14, and the original series is A. The aircraft in column headings of Table 1 are new designs, with the exception of the F/A-18E, which was a major change from the previous F/A-18s; the three F-35 variants are being built for different missions and produced in parallel.

[9] The model in D-5112 used data through 2012 and did not include the F-117A.

[10] This does not mean that STOVL capabilities are free in the model; holding all else equal, STOVL aircraft will tend to be heavier and have more advanced materials than a conventional aircraft. Also note that in model estimation, the coefficient on the STOVL dummy was restricted to ≥ 1.00.

[11] The published FY 2002 Green Book deflators include only projections through FY 2007—beyond FY 2007, we use the 2.1% inflation rate evident in the FY 2004 to FY 2007 projections.

[12] A Nunn-McCurdy unit cost breach (10 U.S.C. § 2433a, "Critical cost growth in major defense acquisition programs") occurs when cost growth in program or acquisition unit cost surpasses 15%.

[13] The 2009 F-35 Nunn-McCurdy breach was driven by cost growth in EMD and nonrecurring procurement as well as by URF.

[14] We used the latest available weight status to characterize the F-35 variants as procured. These values were fixed across the procurement lots and do not include any weight growth margin.

[15] Both of these effects are addressed in the later benchmark comparisons.

[16] Although it would be possible to capture the 2009 procurement profile and escalation application effects in the modelling exercise, we address only these issues in the context of more up-to-date cost data.

Acknowledgments

Thank you to Sarah K. Burns and David M. Tate for performing technical review of this document, and to Steve Miller, Carol Moore, and Garret Summers of the Office of the Secretary of Defense, Cost Assessment and Program Evaluation for inspiring and commenting on the analysis.

Author Biographies

Mr. Stanley Horowitz

is an Assistant Director at the Institute for Defense Analyses (IDA). Much of his work involves personnel compensation and management policies, and analysis of readiness. Recently he studied inflation indexes in DoD. He directed studies of Reserve Component readiness, costing, training, and volunteerism. In 2015, he received the Goodpaster Award for Excellence in Research from IDA. Mr. Horowitz studied economics at M.I.T. and the University of Chicago.

(E-mail address: shorowit@ida.org)

Mr. Bruce Harmon

works for the Institute for Defense Analyses, where he has been a professional research staff member for over 25 years. Mr. Harmon has extensive experience in modeling the costs and schedules of various aerospace systems, as well as in analyses of other acquisition issues. He is a PhD candidate in Economics at American University, Washington, DC.

(E-mail address: bharmon@ida.org)

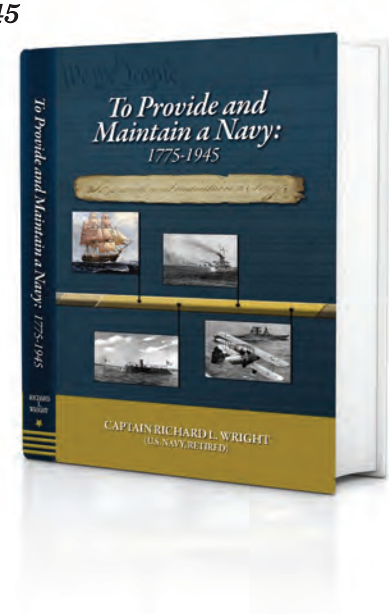
PROFESSIONAL READING LIST

The Defense Acquisition Professional Reading List is intended to enrich the knowledge and understanding of the civilian, military, contractor, and industrial workforce who participate in the entire defense acquisition enterprise. These book recommendations are designed to complement the education and training vital to developing essential competencies and skills of the acquisition workforce. Each issue of the *Defense Acquisition Research Journal* will include one or more reviews of suggested books, with more available on our website at http://dau.edu/library.

We encourage our readers to submit book reviews they believe should be required reading for the defense acquisition professional. The books themselves should be in print or generally available to a wide audience; address subjects and themes that have broad applicability to defense acquisition professionals; and provide context for the reader, not prescriptive practices. Book reviews should be 450 words or fewer, describe the book and its major ideas, and explain its relevancy to defense acquisition. Please send your reviews to the managing editor, *Defense Acquisition Research Journal* at *DefenseARJ@dau.edu*.

Featured Books

To Provide and Maintain a Navy: 1775–1945



Author: CAPT Richard L. Wright, USN (Ret.)

Publisher: Strategic Insight, Ltd.

Copyright Date: 2019

Hard/Softcover: 276 pages

ISBN-13: 9780578420967

Available from: https://www.stratsight.com/online-store

Reviewed by: Brad Martin, Senior Policy Researcher, RAND Corporation

This book is worth reading by defense acquisition professionals. The events discussed go past simple historical interest, both in describing parallel situations to the present day and in identifying characteristics that have informed current structures.

The book credits several contributors in addition to the named author. While it does not read like a committee report, it is difficult to identify a single intellectual theme extending throughout. This may relate to a critical problem with the book's attempt to cover a large subject in some—but not too much—detail while at the same time making general points. This is difficult to achieve, resulting in some places where detail appears to be lacking and in other places a search for generality where it is not clear that generality exists. For example, the history of the Union Navy in the Civil War is fascinating, but with some attempts to describe tactics, some attempts to describe specific battles, then some general conclusions about industrial base. There's a reasonable argument that the Civil War offered a relatively unique set of circumstances, interesting in its own right, but not necessarily relevant to subsequent events. The chapter does not support a general point, nor does it cover the history in detail.

The focus of the work is on surface ships, which is certainly defensible for most of the U.S. Navy's history. However, the discussion of naval aviation is largely a discussion of aircraft carrier design rather than a description of aviation's role or aircraft capabilities. Submarine programs do receive attention, but here again the discussion is largely limited to submarines that were being designed and procured. The interplay between plans, concepts, and capabilities does not receive much discussion. Since the actual readiness to fight in war depends to a large degree on how well capabilities meet operational concepts, I believe the reader would not receive a complete picture of naval capability development—particularly in the critical interwar period—just from reading this book.

However, these shortcomings are largely the result of an ambitious attempt to cover a large subject in limited space. If those limitations are borne in mind, the book does provide a good sense of factors within the U.S. political system, economy, and culture that have influenced the Navy throughout its history. It is also well-referenced, well-edited, and generous with striking illustrations. For the acquisition professional looking for an overview of the U.S. Navy's history from the perspective of industrial base and capability development, this book is an excellent resource.

Current Research Resources in DEFENSE ACQUISITION

ACQUISITION REFORM

Each issue of the *Defense Acquisition Research Journal* will bring to the attention of the defense acquisition community a topic of current research, which has been undertaken by the DAU Knowledge Repository (KR) librarian team in collaboration with DAU's Director of Research. Both government civilian and military Defense Acquisition Workforce (DAW) readers will be able to access papers publicly and from licensed resources on the DAU KR website: https://identity.dau.edu/EmpowerIDWebIdPForms/Login/KRsite.

Nongovernment DAW readers should be able to use their local knowledge management centers/libraries to download, borrow, or obtain copies. We regret that DAU cannot furnish downloads or copies.

Defense Acquisition Research Journal readers are encouraged to submit proposed topics for future research by the DAU KR librarian team. Please send your suggestion with a short write-up (less than 100 words) explaining the topic's relevance to current defense acquisition to: Managing Editor, *Defense Acquisition Research Journal*, **DefenseARJ@dau.edu**.

An Analysis of Major Acquisition Reforms Through Text Mining and Grounded Theory Design

Amanda L. McGowin

Summary:

Dozens of experts and senior leaders within the acquisition community have published their notions on the reasons for cost growth; nevertheless, legislation has yet to eradicate this presumed conundrum. For this reason, this research is aimed at identifying existing trends within past major Defense Acquisition Reform legislation, as well as in a compendium of views from leaders within the Defense Acquisition community on the efficacy of acquisition reform, to determine the possible disconnect.

APA Citation:

McGowin, A. L. (2018, March). *An analysis of major acquisition reforms through text mining and grounded theory design* (AFIT-ENV-MS-18-M-224). Air Force Institute of Technology. Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/1056519.pdf

A Roadmap to the Section 809 Panel Reports

Advisory Panel on Streamlining and Codifying Acquisition Regulations

Summary:

The Section 809 Panel has proposed 98 recommendations, which collectively provide a strategy for moving defense acquisition into the 21st century. The recommendations work in concert to support the first of five concepts, namely "adapt at the speed of a changing world." This Roadmap organizes the recommendations thematically around the remaining four concepts (namely leverage the dynamic marketplace, allocate resources effectively, enable the workforce, and simplify acquisition), to demonstrate how suggested changes to acquisition practices can produce desired outcomes.

APA Citation:

Advisory Panel on Streamlining and Codifying Acquisition Regulations (2019, February). *A roadmap to the section 809 panel reports.* Retrieved from https://discover.dtic.mil/wp-content/uploads/809-Panel-2019/Roadmap/Sec809Panel_Roadmap_Feb2019.pdf

Acquisition Reform in the FY2016–FY2018 National Defense Authorization Acts (NDAAs)

Moshe Schwartz

Summary:

This report provides a brief overview of selected acquisition-related provisions found in the National Defense Authorization Acts for FY2016 through FY2018. This report also discusses one of the more controversial and extensive legislative changes made in recent years affecting acquisition: the breakup of the office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, as well as the shift of authority from that office to the military departments.

APA Citation:

Schwartz, M. (2018, January). *Acquisition reform in the FY2016–FY2018 National Defense Authorization Acts (NDAAs)*. Congressional Research Service. Retrieved from https://congressional.proquest.com/congressional/docview/t21.d22.crs-2018-fdt-0019?accountid=40390

A Systems Theory Based Examination of Failure in Acquisition System Reform

Charles B. Keating, Joseph M. Bradley, Polinpapilinho F. Katina, and Craig Arndt

Summary:

Following a brief introduction to Systems Theory, this paper reports on efforts to (a) briefly examine the current state of the defense acquisition system and programs, focused on successes, failures, major reform themes, and critical challenges for moving forward; (b) mapping of systems pathologies to provide a different "Systems Theory"-based perspective of acquisition system reform as well as acquisition system development; and (c) suggest implications for acquisition system development based on contributions from Systems Theory.

APA Citation:

Keating, C. B., Bradley, J. M., Katina, P. F., & Arndt, C. (2018). *A Systems theory based examination of failure in acquisition system reform.* Naval Postgraduate School. Retrieved from https://calhoun.nps.edu/bitstream/handle/10945/58785/SYM-AM-18-100-AP_Keating.pdf?sequence=1

DoD Acquisition Reform: Leadership Attention Needed to Effectively Implement Changes to Acquisition Oversight

Shelby S. Oakley

Summary:

This report addresses (a) the progress DoD has made implementing selected oversight reforms related to major defense acquisition programs; (b) how DoD has used middle-tier acquisition pathways; and (c) challenges DoD faces related to reform implementation. GAO reviewed five reforms: milestone decision authority designation; cost, fielding, and performance goals; independent technical risk assessments; restructuring of acquisition oversight offices; and middle-tier acquisition.

APA Citation:

Oakley, S. S. (2019, June). *DoD acquisition reform: Leadership attention needed to effectively implement changes to acquisition oversight* (GAO-19-439). U.S. Government Accountability Office. Retrieved from https://www.gao.gov/assets/700/699527.pdf

Defense ARJ Guidelines
FOR CONTRIBUTORS

The *Defense Acquisition Research Journal (ARJ)* is a scholarly peer-reviewed journal published by the Defense Acquisition University (DAU). All submissions receive a double-blind review to ensure impartial evaluation.

IN GENERAL

We welcome submissions describing original research or case histories from anyone involved in the defense acquisition process. Defense acquisition is broadly defined as any actions, processes, or techniques relevant to the conceptualization, initiation, design, development, testing, contracting, production, deployment, logistics support, modification, and disposal of weapons and other systems, supplies, or services needed for a nation's defense and security, or intended for use to support military missions.

Research involves the creation of new knowledge. This generally requires either original analysis of material from primary sources, including program documents, policy papers, memoranda, surveys, interviews, etc.; or analysis of new data collected by the researcher. Articles are characterized by a systematic inquiry into a subject to establish facts or test theories that have implications for the development of acquisition policy and/or process.

The *Defense ARJ* also welcomes case history submissions from anyone involved in the defense acquisition process. Case histories differ from case studies, which are primarily intended for classroom and pedagogical use. Case histories must be based on defense acquisition programs or efforts. Cases from all acquisition career fields and/or phases of the acquisition life cycle will be considered. They may be decision-based, descriptive, or explanatory in nature. Cases must be sufficiently focused and complete (i.e., not open-ended like classroom case studies) with relevant analysis and conclusions. All cases must be factual and authentic. Fictional cases will not be considered.

We encourage prospective writers to coauthor, adding depth to manuscripts. We recommend that junior researchers select a mentor who has been previously published or has expertise in the manuscript's subject. Authors should be familiar with the style and format of previous *Defense ARJs* and adhere to the use of endnotes versus footnotes, formatting of reference lists, and the use of designated style guides. It is also the responsibility of the corresponding author to furnish any required government agency/employer clearances with each submission.

MANUSCRIPTS

Manuscripts should reflect research of empirically supported experience in one or more of the areas of acquisition discussed above. The *Defense ARJ* is a scholarly research journal and as such does not publish position papers, essays, or other writings not supported by research firmly based in empirical data. Authors should clearly state in their submission whether they are submitting a research article or a case history. The requirements for each are outlined below.

Research Articles

Empirical research findings are based on acquired knowledge and experience versus results founded on theory and belief. Critical characteristics of empirical research articles:

- clearly state the question,
- define the research methodology,

- describe the research instruments (e.g., program documentation, surveys, interviews),
- describe the limitations of the research (e.g., access to data, sample size),
- summarize protocols to protect human subjects (e.g., in surveys and interviews), if applicable,
- ensure results are clearly described, both quantitatively and qualitatively,
- determine if results are generalizable to the defense acquisition community
- determine if the study can be replicated, and
- discuss suggestions for future research (if applicable).

Research articles may be published either in print and online, or as a Web-only version. Articles that are 5,000 words or fewer (excluding abstracts, references, and endnotes) will be considered for print as well as Web publication. Articles between 5,000 and 10,000 words will be considered for Web only publication, with a two-sentence summary included in the print version of the *Defense ARJ*. In no case should article submissions exceed 10,000 words.

Case Histories

Care should be taken not to disclose any personally identifiable information regarding research participants or organizations involved unless written consent has been obtained. If names of the involved organization and participants are changed for confidentiality, this should be highlighted in an endnote. Authors are required to state in writing that they have complied with APA ethical standards. A copy of the APA Ethical Principles may be obtained at http://www.apa.org/ethics/.

All case histories, if accepted, will receive a double-blind review as do all manuscripts submitted to the *Defense ARJ*.

Each case history should contain the following components:

- Introduction
- Background

- Characters
- Situation/problem
- Analysis
- Conclusions
- References

Book Reviews

Defense ARJ readers are encouraged to submit book reviews they believe should be required reading for the defense acquisition professional. The reviews should be 500 words or fewer describing the book and its major ideas, and explaining why it is relevant to defense acquisition. In general, book reviews should reflect specific in-depth knowledge and understanding that is uniquely applicable to the acquisition and life cycle of large complex defense systems and services. Please include the title, ISBN number, and all necessary identifying information for the book that you are reviewing as well as your current title or position for the byline.

Audience and Writing Style

The readers of the *Defense ARJ* are primarily practitioners within the defense acquisition community. Authors should therefore strive to demonstrate, clearly and concisely, how their work affects this community. At the same time, do not take an overly scholarly approach in either content or language.

Format

Please submit your manuscript according to the submissions guidelines below, with references in APA format (author date-page number form of citation) as outlined in the latest edition of the *Publication Manual of the American Psychological Association*. References should include Digital Object Identifier (DOI) numbers when available. The author(s) should not use automatic reference/bibliography fields in text or references as they can be error-prone. Any fields should be converted to static text before submission, and the document should be stripped of any outline formatting. All headings should conform to APA style. For all other style questions, please refer to the latest edition of the *Chicago Manual of Style*.

Contributors are encouraged to seek the advice of a reference librarian in completing citation of government documents because standard formulas of citations may provide incomplete information in reference to government works. Helpful guidance is also available in *The Complete Guide to Citing Government Information Resources: A Manual for Writers and Librarians* (Garner & Smith, 1993), Bethesda, MD: Congressional Information Service.

The author (or corresponding author in cases of multiple authors) should attach a cover letter to the manuscript that provides all of the authors' names, mailing and e-mail addresses, as well as telephone numbers. The letter should verify that (1) the submission is an original product of the author(s); (2) all the named authors materially contributed to the research and writing of the paper; (3) the submission has not been previously published in another journal (monographs and conference proceedings serve as exceptions to this policy and are eligible for consideration for publication in the *Defense ARJ*); (4) it is not under consideration by another journal for publication. If the manuscript is a case history, the author must state that they have complied with APA ethical standards in conducting their work. A copy of the APA Ethical Principles may be obtained at http://www.apa.org/ethics/. Finally, the corresponding author as well as each coauthor is required to sign the copyright release form available at our website: www.dau.edu/library/arj.

COPYRIGHT

The *Defense ARJ* is a publication of the United States Government and as such is not copyrighted. We will not accept copyrighted manuscripts that require special posting requirements or restrictions. If we do publish your copyrighted article, we will print only the usual caveats. The work of federal employees undertaken as part of their official duties is not subject to copyright except in rare cases.

Web-only publications will be held to the same high standards and scrutiny as articles that appear in the printed version of the journal and will be posted to the DAU website at www.dau.edu.

In citing the work of others, please be precise when following the author date-page number format. It is the contributor's responsibility to obtain permission from a copyright holder if the proposed use exceeds the fair use provisions of the law (see the latest edition of *Circular 92: Copyright Law of the United States of America and Related Laws Contained in Title 17 of the United States Code,* Washington, DC: U.S. Government Printing Office). Contributors will be required to submit a copy of the writer's permission to the managing editor before publication.

We reserve the right to decline any article that fails to meet the following copyright requirements:

- The author cannot obtain permission to use previously copyrighted material (e.g., graphs or illustrations) in the article.
- The author will not allow DAU to post the article in our *Defense ARJ* issue on our Internet homepage.
- The author requires that the usual copyright notices be posted with the article.
- To publish the article requires copyright payment by the DAU Press.

SUBMISSION

All manuscript submissions should include the following:

- Completed submission checklist
- Completed copyright release form
- Cover letter containing the complete mailing address, e-mail address, and telephone number for each author
- Biographical sketch for each author (70 words or fewer)
- Headshot for each author saved as a 300 dpi (dots per inch) high resolution JPEG or Tiff file no smaller than 5x7 inches with a plain background in business dress for men (shirt, tie, and jacket) and business appropriate attire for women. All active duty military should submit headshots in Class A uniforms. Please note: low-resolution images from Web, PowerPoint, or Word will not be accepted due to low image quality.
- One copy of the typed manuscript, including:
    - Title (12 words or fewer)
    - Abstract (150 to 250 words)
    - Two-sentence summary
    - Keywords (5 words or fewer—please include descriptive words that do not appear in the manuscript title, to make the article easier to find)

- Figures and tables saved as separate individual files and appropriately labeled

The manuscript should be submitted in Microsoft Word (please do not send PDFs), double-spaced Times New Roman, 12-point font size (5,000 words or fewer for the printed edition and 10,000 words or fewer for online-only content excluding abstracts, figures, tables, and references).

Figures or tables should not be inserted or embedded into the text, but submitted as separate files in the original software format in which they were created. For additional information on the preparation of figures or tables, refer to the Scientific Illustration Committee, 1988, Illustrating Science: Standards for Publication, Bethesda, MD: Council of Biology Editors, Inc. Restructure briefing charts and slides to look similar to those in previous issues of the *Defense ARJ*.

All forms are available at our website: www.dau.edu/library/arj. Submissions should be sent electronically, as appropriately labeled files, to the *Defense ARJ* managing editor at: DefenseARJ@dau.edu.

Defense ARJ
PRINT SCHEDULE

The *Defense ARJ* is published in quarterly theme editions. All submissions are due by the first day of the month. See print schedule below.

Author Deadline	Issue
July	January
October	April
January	July
April	October

In most cases, the author will be notified that the submission has been received within 48 hours of its arrival. Following an initial review, submissions will be referred to peer reviewers and for subsequent consideration by the Executive Editor, ***Defense ARJ***.

Contributors may direct their questions to the Managing Editor, *Defense ARJ*, at the address shown below, or by calling 703-805-3801 (fax: 703-805-2917), or via the Internet at norene.johnson@dau.edu.

The DAU Homepage can be accessed at:
https://www.dau.edu

DEPARTMENT OF DEFENSE

DEFENSE ACQUISITION UNIVERSITY

ATTN: DAU PRESS (*Defense ARJ*)

9820 BELVOIR RD STE 3

FORT BELVOIR, VA 22060-5565


CALL FOR
AUTHORS

We are currently soliciting articles and subject matter experts for the 2020-2021 *Defense Acquisition Research Journal (ARJ)* print year.

Please see our guidelines for contributors for submission deadlines.

Even if your agency does not require you to publish, consider these career-enhancing possibilities:

- Share your acquisition research results with the Acquisition and Sustainment (A&S) community.
- Change the way Department of Defense (DoD) does business.
- Help others avoid pitfalls with lessons learned or best practices from your project or program.
- Teach others with a step-by-step tutorial on a process or approach.
- Share new information that your program has uncovered or discovered through the implementation of new initiatives.
- Condense your graduate project into something beneficial to acquisition professionals.

ENJOY THESE BENEFITS:

- Earn 30 continuous learning points for publishing in a refereed (peer reviewed) journal.
- Earn a promotion or an award.
- Become part of a focus group sharing similar interests.
- Become a nationally recognized expert in your field or specialty.
- Be asked to speak at a conference or symposium.

We welcome submissions from anyone involved with or interested in the defense acquisition process—the conceptualization, initiation, design, testing, contracting, production, deployment, logistics support, modification, and disposal of weapons and other systems, supplies, or services (including construction) needed by the DoD, or intended for use to support military missions.

For submissions, please contact the *Defense ARJ* managing editor at ***DefenseARJ@dau.edu*** and visit the *Defense ARJ* Submissions page at ***https://www.dau.edu/library/arj/p/Defense-ARJ-Submissions***.

Defense Acquisition Research Journal SURVEY

Please rate this publication based on the following scores:

5 — Exceptional   4 — Great   3 — Good   2 — Fair   1 — Poor

Please circle the appropriate response.

1. How would you rate the overall publication?	5	4	3	2	1
2. How would you rate the design of the publication?	5	4	3	2	1

	True	False
a) This publication is easy to read		
b) This publication is useful to my career		
c) This publication contributes to my job effectiveness		
d) I read most of this publication		
e) I recommend this publication to others in the acquisition field		

If hand written, please write legibly.

3. What topics would you like to see get more coverage in future ***Defense ARJs***?

4. What topics would you like to see get less coverage in future ***Defense ARJs***?

5. Provide any constructive criticism to help us to improve this publication:

6. Please provide e-mail address for follow up (optional):

YOU CAN ALSO FIND THE SURVEY ONLINE AT

https://www.dau.edu/library/arj

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST-CLASS MAIL PERMIT NO. 12 FORT BELVOIR,VA

POSTAGE WILL BE PAID BY ADDRESSEE

DEFENSE ACQUISITION UNIVERSITY
ATTN: DAU PRESS
9820 BELVOIR RD STE 3
FORT BELVOIR VA 22060-9910

DEFENSE ACQUISITION RESEARCH JOURNAL

A Publication of the Defense Acquisition University

We're on the Web at:
http://www.dau.edu/library/arj

ARJ

DEFENSE ACQUISITION RESEARCH JOURNAL

A Publication of the Defense Acquisition University

Current. Connected. Innovative.

HERMES CREATIVE AWARDS

2019 PLATINUM WINNER

WINNER

2019 APEX®

AWARDS FOR PUBLICATION EXCELLENCE

2019 MARCOM

PLATINUM

WINNER