3-2022

Analysis of Cost and Schedule Estimation Trends for Major Defense Acquisition Programs

Sammantha J. Jones

ANALYSIS OF COST AND SCHEDULE ESTIMATION TRENDS FOR MAJOR DEFENSE ACQUISITION PROGRAMS

Thesis

Sammantha J. Jones, 1st Lieutenant, USAF

AFIT-ENV-MS-22-M-215

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENV-MS-22-M-215

ANALYSIS OF COST AND SCHEDULE ESTIMATION TRENDS FOR MAJOR DEFENSE ACQUISITION PROGRAMS

THESIS

Presented to the Faculty

Department of Mathematics and Statistics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Cost Analysis

Sammantha J. Jones, BS

Lieutenant, USAF

March 2022

AFIT-ENV-MS-22-M-215

ANALYSIS OF COST AND SCHEDULE ESTIMATION TRENDS FOR MAJOR DEFENSE ACQUISITION PROGRAMS

Sammantha J. Jones, BS

1st Lieutenant, USAF

Committee Membership:

Dr. Edward D. White
Chairman

Dr. Jonathan D. Ritschel
Member

Mr. Shawn Valentine
Member

AFIT-ENV-MS-22-M-215

Abstract

This study uses both descriptive and inferential techniques to investigate trends in cost and schedule estimates created by the Department of Defense for their Major Defense Acquisition Programs (MDAPs) throughout the last five decades. For schedule growth percentages, we did not identify any statistically significant trend regarding increasing or decreasing schedule changes or variances of schedule estimates throughout the five decades analyzed. Examining the overall cost growth of MDAPs, a statistically significant difference between the Cost Growth Factors (CGFs) calculated between the 1990s and the 2010s was found, with the 2010s exhibiting lower CGFs. A downward trend in the variances of overall CGFs throughout the decades was identified in three out of four of the analyses conducted. Lastly, differences were detected in the program acquisition unit cost (PAUC) CGFs between the 1990s and the 2010s, and slight differences between the 1970s and the 2010s. Additionally, PAUC CGFs also displayed differences in variances. Overall, we found no trends identifying that the DoD is improving accuracy of cost or schedule estimates for MDAPs; however, we did identify a statistically consistent reduction in the variances of overall CGFs for MDAPs through the five decades. This finding appears to be the first documented case known to us.

Acknowledgments

I cannot express enough gratitude to my advisor and committee chairman, Dr. Tony White, for his continuous patience, support, and direction during the formation and execution of this thesis. Furthermore, the completion of this study would not have been possible without the guidance provided by Dr. Jonathan Ritschel and Mr. Shawn Valentine. Lastly, and most importantly, I would like to thank my husband and son for their endless support during the long hours and late nights spent completing this study.

Sammantha J. Jones

Table of Contents

Page

List of Figures

Page

List of Tables

ANALYSIS OF COST AND SCHEDULE ESTIMATION TRENDS FOR MAJOR DEFENSE ACQUISITION PROGRAMS

I. Introduction

General Issue

Major Defense Acquisition Programs (MDAPs) are essential for the development and production of military aircraft, satellites, missiles, and several other large investment items that U.S. military operations require. These investments are not your customary purchases though; they are acquisitions with large economic risks tied to them. Within these programs there is a historical precedent for underestimating program costs and schedules. In 2017 the RAND Corporation even recommended that the acquisition community approached early cost estimates with skepticism (Light et al., 2017). MDAPs include acquisitions that require more than $300M for research, development, test, and evaluation (RDT&E) or $1.8 billion in total expenditures in base year 1990 constant dollars (House of Representatives, Congress, 2011).

Cost growth in MDAPs is common; however, when there is dramatic growth within programs, this can lead to what is known as a Nunn-McCurdy Breach. From 1997 to 2016, 58 out of 189, or 36% of MDAPs experienced cost growth large enough to cause these breaches. Out of these 58 breaches 18 were significant and 40 were critical (OUSD, 2016, p. 65). Significant breaches occur when current estimates meet or exceed 15% of the current baseline estimate or 30% of the original baseline estimate of an acquisition program. Critical breaches occur at the 25% and 50% levels respectively ([USC05] 10 USC 2433: Unit cost reports, n.d.). Although there is an expectation that MDAPs exhibit small amounts of cost growth, these cases require congressional notification and a thorough review of the program.

MDAPs that experience Nunn-McCurdy breaches are extreme examples of cost growth. But due to their programmatic costs, even a small percentage of cost growth can add millions of additional funding needs for the programs. Schedule growth in MDAPs can also lead to readiness issues and apprehension for military and congressional leadership. Because of the funding and readiness issues, there have been efforts over the last several decades to reduce cost and schedule growth within MDAPs. These efforts include sweeping reforms, changes in business practices, updates to record keeping requirements, and adjustments in the overall structure of how MDAPs are executed and their records maintained (Fox et al., 2012, Dwyer et al., 2020).

Over the last few decades there has been extensive analyses conducted on Department of Defense's (DoD) MDAPs. Various organizations such as the Congressional Research Service, DoD, Government Accountability Office (GAO), or even contracted organizations such as RAND or the Institute for Defense Analyses (IDA) have conducted these studies, providing a plethora of information regarding MDAPs. In 2016 the DoD published their last annual acquisition system performance report. In this report they analyzed MDAPs through a variety of different lenses to include cost and schedule growth, cost performance overall, cost performance broken out by development and production, cost growth by military departments, cost growth by contractors, and a few other viewpoints (Department of Defense, 2016).

The 2016 report claims that there has been a continuing improvement in the fields of defense acquisitions, however their analyses provide several various micro-level insights into the cost and schedule growth of DoD MDAPs. While these micro-level assessments are extremely important to understanding what is happening in MDAPs, their study does not provide a macro-level analysis that truly identified whether the overall cost and schedule growth of MDAPs have changed (Department of Defense, 2016).

Problem Statement & Research Focus

The goal of this study is to identify if the DoD has been successful in increasing their ability to accurately estimate the cost and schedule of MDAPs over the last five decades. This thesis provides a comprehensive analysis on overall cost growth, program acquisition unit cost (PAUC) growth, and schedule growth for the DoD's largest acquisitions, MDAPs. For overall cost growth and PAUC cost growth, we used calculated cost growth factors (CGF) as our metric to complete the analysis, and for schedule growth, we calculated the percentage change in the estimated date for the program to reach initial operational capabilities (IOC).

Besides MDAP cost, this research also focuses on schedule estimation. Schedule variation and growth is also extremely important when it comes to MDAPs. Although not looked at with as much scrutiny as cost estimates, schedule estimates, especially delays in fielding programs, can lead to readiness issues. This has become highlighted by leadership who have recently started trying to accelerate the acquisition process. According to DoDD 5000.01, one of the established policy objectives for the Defense Acquisition System it to deliver performance at the speed of relevance (Department of Defense, 2017).

The inspiration for this study came from two papers published through RAND in the mid-2000s. One was completed by Arena et al. (2006) titled "Historical Cost Growth of Completed Weapons System Programs", while the second was a follow-on study completed by Younossi et al. (2007) titled "Is Weapons System Cost Growth Increasing." Both papers provide insights into cost growth in MDAPs and Younossi's study reviewed developmental cost growth data across three decades (1970s through 1990s). This perspective of MDAPs across the decades allows for an insightful macro-level review of how these factors are changing across time (Arena et al., 2006; Younossi et al., 2007).

Although the RAND studies were the inspiration for this research, this thesis deviates from those in two major ways. First, we included schedule growth into our analysis, and secondly, we did not delineate our sample into development and procurement costs. Although dividing cost growth into development and procurement is a common practice when analyzing MDAPs, we wanted to look at the overall cost and schedule growth holistically for this study. There are other deviations between the analyses, such as how the data is presented and the type of inferential analyses utilized, but the overall goal of this paper is to investigate cost (and schedule) growth over the last five decades and to identify whether the DoD has seen a decrease in cost (and schedule) growth over the last five decades (Younossi et al., 2007; Arena et al., 2006).

Research Question

This paper has one overarching question applied to all three analyses within this thesis and that is: how have DoD cost and schedule estimates changed over the last five decades? We answer this question by first analyzing schedule growth percentages, then move onto the overall CGFs of MDAPs, and lastly, we investigate the quantity adjusted CGFs.

Methodology

The data used for this research originates from Selected Acquisition Report (SAR) data archived in the CADE SAR Database. We then aggregated the data into groups by decade and investigated to identify whether there has been a change in schedule and cost estimations within DoD MDAPs over the last 50 years from the 1970s through the 2010s.

The data analyses included descriptive statistics such as the mean and median cost and schedule growths of the MDAPs by decade, the interquartile ranges (IQRs), standard deviations, and coefficient of variations (CVs). We also investigated the calculated cost growth factors (CGFs) and schedule growth percentages using inferential statistics such as the non-parametric Kruskal-Wallis Test to determine statistically significant differences in the data across the decades.

Preview

This thesis looks at whether the DoD has decreased their cost and schedule growth for MDAPs over the last 50 years. We do these four different ways; by looking at MDAPs holistically with completed and ongoing programs combined, by reviewing only completed programs, then looking at currently ongoing programs, and lastly by examining MDAPs at the MS B +5 years point in their lifecycle. As we move forward in this thesis, Chapter II summarizes the available literature on the DoD acquisition process, MDAPs, and completed studies regarding cost and schedule growth in these programs. After evaluating the information that is already available, we dive into the methodology of this analysis in Chapter III. This chapter outlines the various tests conducted and why they are important to the overall study. Chapter IV, the analysis and results section of the thesis, communicates the various descriptive statistics and inferential results of the testing. After our analysis, we finish this thesis with a review of our findings and recommendations for future research in Chapter V.

II. Literature Review

Chapter Overview

The purpose of this chapter is to review completed research regarding DoD cost and schedule growth. We begin by describing MDAPS, their execution process, and why they are important. Since several entities, most notably Congress, typically review MDAPs annually, we then investigate and review completed studies pertaining to cost growth and schedule slippage to ascertain salient findings relevant to this thesis' research questions.

Introduction to MDAPs and SARs

MDAPs are the DoD's largest investments and constitute a large proportion of the DoD portfolio relative to their program numbers. Currently, the GAO reports annually on DoD weapon systems based on their total cost and acquisition status. Of the 107 programs evaluated in their 2021 report, 84 were MDAPs. These 84 MDAPs have a total planned investment of $1.79 trillion Fiscal Year (FY) 2021 dollars. The GAO has been completing annual reports for the last 19 years and are able to provide historical data and information about trends within the MDAPs. Accordingly, they have reported consistent cost growth in the DoD's MDAP portfolio for the last 15 years. They attribute the most dramatic cost changes to quantity changes (Government Accountability Office, 2021).

To provide an additional perspective on the financial size of the DoD's acquisition program, we can also look at the amount of funding requested annually for procurement purchases. In February of 2020, the Office of the Undersecretary of Defense, Comptroller submitted a $131.7 billion dollar procurement requirement in the FY21 president's budget. This funds request was purely for purchase of investment items and does not include additional funding sources required to pay for MDAPs including research & development (R&D) funding and operations and maintenance (O&M) required to realize a large DoD acquisition (Office of the Under Secretary of Defense (Comptroller), 2020b).

The cost of MDAPs drives the overall interest and scrutiny reserved to these programs; however, although cost drives most of the attention, schedule variations are also critical when in terms of military readiness. MDAPs, also referred to as Acquisition Category (ACAT) I programs, are broken down into three sub-categories: ACAT IB, ACAT IC, and ACAT ID. Categories are decided based on the total expenditure required for the acquisition and their overall importance to the DoD and various services. The program's category is what drives who the Milestone Decision Authority (MDA) is. ACAT IDs require that the Defense Acquisition Executive be the MDA, ACAT IBs require MDA approval from the Service Acquisition Executive allowing each individual services to manage the programs, and ACAT ICs are programs that require MDA decisions to come from Component Acquisition Executive or the Head of the DoD component leading the program (Defense Acquisition University, n.d.).

Over the history of DoD acquisition programs there have been numerous attempts to increase the reliability of MDAP estimates. There have been several reforms that have caused the Office of the Secretary of Defense (OSD) acquisition process oversight to teeter back and forth between centralized and decentralized management. The McNamara Innovations, initiated in the 1960s when Robert S. McNamara filled the position of Secretary of Defense, began centralizing the acquisition process. During this period, the establishment of the Office of Systems Analysis occurred, which gave the Secretary of Defense a team of over one hundred professionals to prepare cost estimates and cost-benefit analysis to aid in DoD's decision-making process. This increased oversight at the OSD level also drove innovation at the service level. The various services in turn created their own systems analysis offices to assist in ensuring compliance prior to OSD level MDAP submission while simultaneously the schoolhouse for program managers implemented courses regarding planning, controlling, and streamlining the acquisition process (Dwyer et al., 2020).

Although McNamara spent almost a decade attempting to reign in cost and schedule growth within the acquisition program, it was not enough, and the GAO found that a majority of MDAPs started in the 1960s were not meeting performance standards and 38 ongoing programs were already at 50% above their original estimates (Fox et al., 2012, p. 40). As the new Secretary of Defense, Melvin Laird, took control of the DoD, he was able to bring David Packard, of Hewlett-Packard (HP), in as his deputy and appointed him as the lead of research, development, and procurement for the DoD. To assist in his endeavor, Packard created the Defense Systems Acquisition Review Council (DSARC) and an Industry Advisory Council (IAC). The IAC recommended that program management would be best executed by the individual services as they had better expertise of their individual requirements. The role of the DSARC became that of an intermediary between the services and OSD. The services would build and manage their own programs and then submit their program objectives to OSD for periodic review at certain milestones (Fox et al., 2012).

Since those initial reforms in the 1960's and 1970's there has been a steady back and forth of centralization and decentralization within the acquisition process. Following the McNamara Reforms and the DSARC came the Acquisition Improvement Program, the

Defense Acquisition Board, the Mandate for Change and Transformation, the Weapons System Acquisition Reform Acts, and the Restructuring of the AT&L all having been leveraged to assist in strengthening the DoD acquisition process. To assist in their own studies, Dwyer et al. (2020) categorized these various eras of oversight into two distinct groups: centralized and decentralized. Table 1. illustrates the overall outline of this classification effort.

Table 1. Cycles of Acquisition Reform (Dwyer et al., 2020)

Oversight Approach	Years	Reform Cycle
Centralized	1970-1980	Defense System Acquisition Reform Council
	1990-1993	Defense Acquisition Board
	2008-2016	Weapon Systems Acquistion Reform Act
Decentralized	1967-1969	McNamara Reforms
	1987-1989	Acquisition Improvement Program
	1994-2007	Mandate for Change and Transformation
	2017-present	Restructuring AT&L

Along with changes in the overall management of MDAPs, there has also been dramatic changes in the procedures used for the development and lifecycle of MDAPs. Figure 1. shows how the acquisition process looked during the McNamara era of the 1960's versus the changes administered during the Packard era of the early 70's. During the Packard Era we can see the introduction of two additional milestone decision points. The implementation of these additional milestones was to provide increased oversight in the DoD acquisition process. It gave the various services additional evaluation points throughout a program's lifecycle to ensure that they met development, cost, and schedule targets prior to moving on to the next stage (Fox et al., 2012).

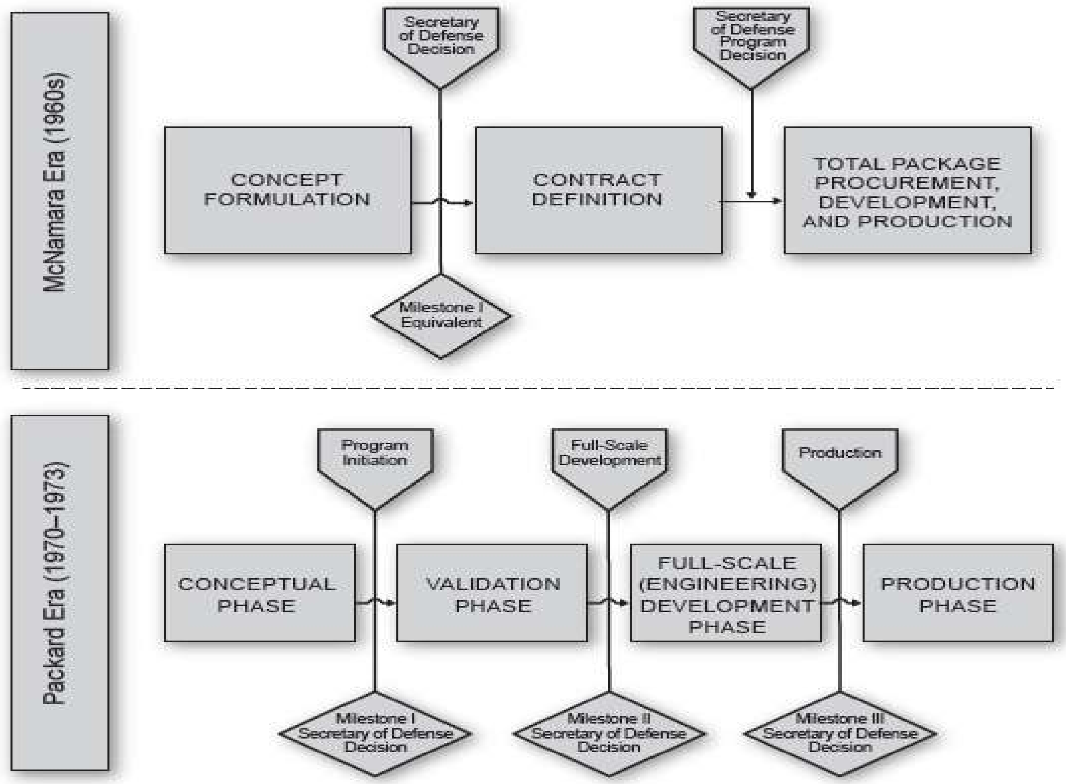


Figure 1. System Acquisition Process (Fox et al., 2012)

Throughout the decades there were slight modifications to the various phases and milestones and their definitions. For example, the validation phase shown in Figure 1 changed to "Demonstration and Validation" in 1977, then to "Concept Demonstration/Validation" in 1987, and then transformed yet again to "Program Definition and Risk Reduction" in 1996. Although there were small terminology changes like this made to Packard's original outline, each of the milestone and phase designations remained the same until 2000. In 2000 this outline saw its largest modifications. The overall outline remained the same; however, the milestone and phase designations both transformed from numeric to alphabetical ones; I became A, II became B, and III became C.

Although the designations and titles transformed, the basic premise of each milestone and phase remained the same. For example, the original Milestone I definition "Approval to Begin a New Acquisition Program" changed to "Approval to Enter the Concept and Technology Development." Both milestones were defined as the decision points where the program moved from an idea into an actuality (Fox et al., 2012). In contrast to Figure 1, which illustrates Packard's original acquisition process framework, Figure 2 illustrates the organization and structure of these new designations. This is the current standard still used by the DoD (AcqNotes LLC, 2021).

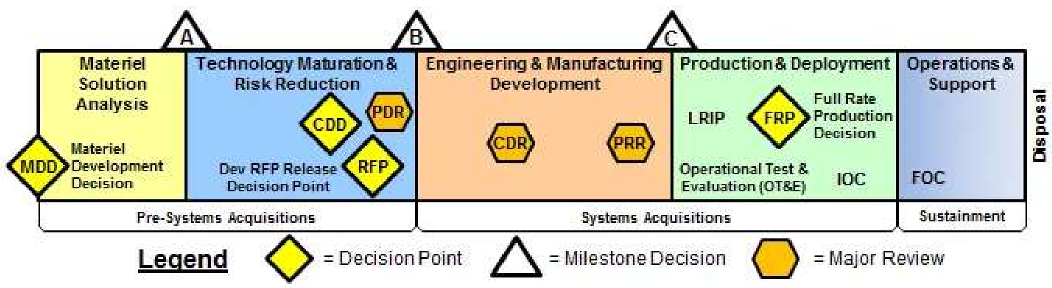


Figure 2. Major Capability Acquisition Process (AcqNotes LLC, 2021)

Milestone decisions are critical in the MDAP lifecycle. Embedded into the acquisition process to evaluate whether a program is ready to transition to the next stage, these decision points provide oversight that prevent programs from moving forward prior to adequate development, designing, and planning. As already discussed, and outlined in the DoDI 5000.85, there are currently three milestone decision points in the acquisition process: A, B, and C. Milestone A is the entry point of a program into the technology maturation and risk reduction (TMRR) phase. Approval at this milestone require the completion of various assessments such as the analysis of alternatives (AoA), an independent cost estimate (ICE), and an independent technical risk assessment (ITRA). The program manager will submit an initial “should cost” estimate, along with an acquisition strategy, an outline of assumptions and programs risk during the milestone A review. The respective approval authority reviews and approves prior to the program moving to the next phase. The TMRR phase is critical in reducing cost risk regarding technology, engineering, integration, and the acquisition program life cycle. This phase is key in developing a more detailed acquisition strategy and taking the program from concept to a preliminary design. This phase is also critical as it leads into the Milestone B (MS B) (Defense Acquisition University, n.d.).

MS B is where a program enters engineering, manufacturing, and development (EMD). This phase is where the idea transforms into a viable end item for use. There are many programs that begin at MS B. For programs that require newer technology MS B is when estimates start to become realistic as the TMRR in complete and the creation of credible cost comparisons can start. Because of its place in the MDAP cycle, MS B is the most reasonable point in the estimation cycle to begin any cost or schedule growth analyses. Many previously published studies use MS B as the starting point of their analysis on MDAP cost or schedule variations (Younossi et al., 2007; McNicol, 2018; Dwyer et al., 2020; Defense Acquisition University, n.d.).

Milestone C (MS C) is where production officially begins. After the MDAPs pass the MS C, they enter Low-Rate Initial Production, which tests the manufacturing process and leads into Initial Operational Test & Evaluation and then to Full Rate Production. During MS C programs also enter Initial Operational Capability (IOC), which is a critical step for MDAPs. IOC happens when a few of the organization slated to receive the product have received and are able to deploy and maintain the products. A successful IOC leads to fully deploying the system. The achievement of full operational capability occurs when the deployment of the product to all the organizations occurs, and they can use and maintain the systems (Defense Acquisition University, n.d.). When it comes to evaluating the schedule growth of programs in this paper, we utilize the IOC as the cutoff for the calculations. We further define these calculations in Chapter III.

Relevant Research

In 2007 RAND (Younossi et al., 2007) published a study that focused on the cost growth of DoD weapons system. Overall, there were 76 programs included in this study, and the data was evaluated in a few different ways. The first part of this analysis looked at completed MDAPs and analyzed the total CGF while separating and investigating the sample by Development CGFs and Procurement CGFs. Additionally, this study divided the MDAPs by commodity type to find the various CGFs and then aggregated them into three categories: Aircraft and Helicopters, Missiles and Electronics, and Space. He took a quick glance as the average development CGF for completed programs across the 1970s through the 1990s. and then explored the development costs of completed and ongoing programs. Lastly, Younossi et al. (2007) analyzed development cost growth weighted by commodity alongside completed, completed and ongoing, and completed programs at the MS B +5 years mark. Although his chart shows a downward trend for the development CGF of completed programs, he highlights in his article that completed programs from the 1990s that may suffer from maturity bias. In this context the maturity bias comes from completed programs from the 1970 having longer lifespans than those completed from the 1990s. These longer lifespans can drive higher cost growth, while the shorter lifespans can indicate shorter acquisition turn-around and less opportunity for realized cost growth. Figure 3 illustrates the overall findings of this study.

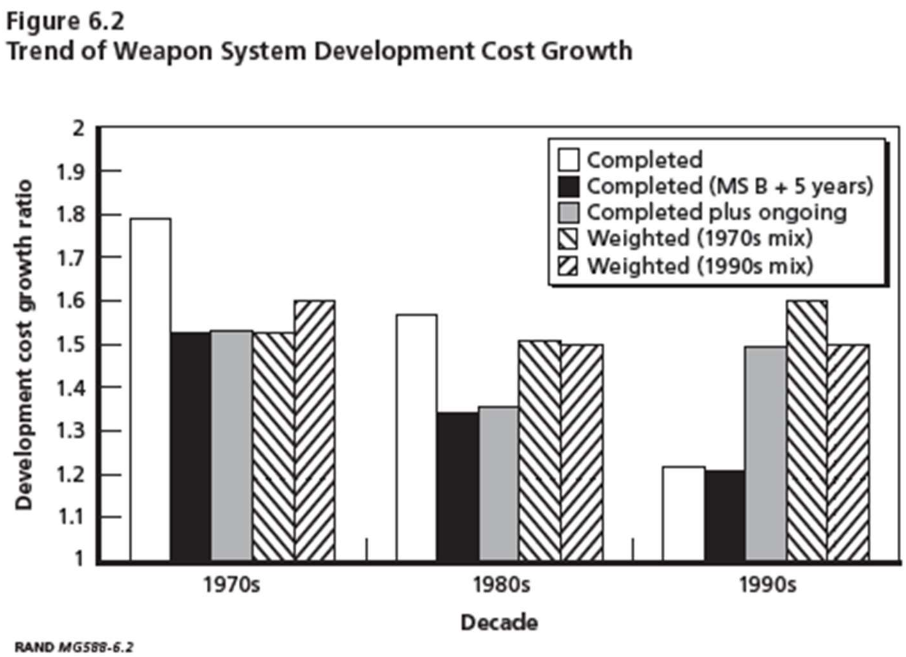


Figure 3. Trend of Weapon System Development Cost Growth (Younossi et al., 2007)

Calculating the CGF in completed programs in this study consisted of taking the final cost and dividing by the programs estimated cost. For incomplete programs, this study utilized the MS B plus five years method which entailed using only ongoing programs that had aged at least five years since reaching MS B. To calculate the CGF for ongoing programs Younossi et al. (2007) divided the cost estimate obtained from the SAR available at five years past MS B by the estimate at MS B. The study identified that over the three decades investigated, there were no significant improvement in cost growth.

Kozlak et al. (2017) reported on the investigation of cost growth of DoD aircraft MDAPs. This study evaluated the CGFs of 30 aircraft programs though their acquisition life cycle. In this analysis the data was aggregated into three overarching categories, Development, Procurement, and Total. They then broke the programs down to analyze cost growth at various stages: Critical Design Review (CDR), First Flight (FF), Development Test & Evaluation (DT&E), Initial Operational Capability (IOC), and Full Operational Capability (FOC). In this analysis, Kozlak et al. (2017) identified that procurement costs drove cost growth the most within each program. There was a trend of a large spike at FF observed and they also note that a program realizes about 91 percent of its overall growth about 6.5 years after reaching MS B (Kozlak et al., 2017).

Over the last decade, the Naval Postgraduate School has completed a series of studies on the DoD acquisition process. Dwyer et al. (2020) focused on whether the implementation of various acquisition reforms had any impact on cycle times in MDAPs. The cycles that were analyzed were the MS B, Milestone C, and the IOC. Through this analysis the reform cycles were separated by each individual reform which included the McNamara Reform, the Defense Systems Acquisition Reform Council, the Acquisition Improvement Programs, the Defense Acquisition Board, the Mandate for Change and Transformation, and the Weapons System Acquisition Reform Act (See Table 1. for breakout).

This study determined that looking at the MDAPs during each individual reform cycle did not yield any substantial evidence of impact on cycle times. However, the reforms were then aggregated into two groups; centralized and decentralized. Although they did not find any statistically significant change in cycle time based on these groupings, they did find that there was a statistically significant difference in cycle time growth between the two groups, with centralized oversight leading to more accurate schedules. They suggested that the increased oversight of these programs during the times of centralization could be attributed to the decrease in cycle time growth (Dwyer et al., 2020).

There are a few studies that focus purely on the Weapons Systems Acquisition Reform Act of 2009 (WSARA), which was implemented "to improve the organization and procedures of the Department of Defense for the acquisition of major weapon systems, and for other purposes" (Weapon Systems Acquisition Reform Act of 2009, P.L. No. 111 - 23, 2009). In a GAO report published in 2012, they assert that the WSARA had positive impacts on the overall cost and schedule estimation process by forcing program offices to scrutinize the estimates more heavily (Government Accountability Office, 2012). Other analyses completed by Banford and Weidman (2014) and Ritschel et al. (2019) both indicated that the implementation of WSARA has had positive impacts on the cost variations and cost overruns in MDAPs. In the journal publication from Ritschel et al. (2019), there was a decrease in cost overruns for research and development (R&D) contracts. These findings did not translate to contracts outside of the R&D scope, where variables falling under the political scope were found to be the primary drivers of cost overruns. These studies provide critical insight into the WSARAs impact on the cost overruns in the DoD acquisitions but are relatively focused with Banford and Weidman evaluating Army programs from 2005-2015 and Ritschel et al. investigating Air Force programs from 2002 to 2015.

Through the Institute for Defense Analysis (IDA), there have been several studies commissioned that investigated cost growth of MDAPS. McNicol et al. (2016) studied Average Procurement Unit Cost (APUC) growth pre-Packard and post-Packard Reforms. That study's database consisted of MDAPs from 1964 to 2001 grouped into six categories based on the "acquisition regime". The study determined that Packard's reforms had a successful impact in curbing dramatic cost growth in MDAPS and decreasing APUC cost growth. However, during the 1990s, alongside a reduction in oversight, the DoD saw a spike in APUC growth.

Looking at the reform cycles provides a great insight, however, another analysis from the IDA, discussed the impact on funding climates relationship on PAUC growth. McNicol (2018) investigated MDAPs during boom-and-bust funding climates. The implications identified in the analysis state that PAUC growth during boom cycle were due to program changes while PAUC growth during bust cycles were primarily attributable errors in estimation. Another trend noted is the increase of program cancellations during times of funding shortfalls within the DoD.

Overall, there have been many attempts to analyze cost and schedule growth in MDAPs, however many of these studies evaluate programs at a micro level looking at individual program categories, impacts of specific legislation on cost or schedule variations, or studies focused on either cost growth or schedule separately. This thesis aims to create a DoD wide cost and schedule analysis that provides a macro look at MDAPs over the last five decades.

Summary

This chapter provided a brief outline of the DoD MDAP acquisition history and process while defining many of the important phases and milestones that are important in the MDAP lifecycle. Additionally, we reviewed and discussed research completed regarding DoD cost and schedule estimation accuracy and growth. Moving into Chapter III we examine our data and explain the methodology used for the investigation of our cost growth factors and schedule growth percentages.

III. Methodology

Chapter Overview

This chapter begins by discussing the limitations and assumptions used for reviewing our data. We then move onto how we collected the data, where it was sourced from, and how the data was normalized for use in the analyses. The last part of this chapter dives into the tools used to complete the data analyses. We investigate MDAPS through three separate lenses regarding cost and schedule growth: program schedule growth, overall program Cost Growth Factor (CGF), and the PAUC CGF.

Assumptions & Limitations

There have been great strides in consolidating and normalizing MDAP data, most notably the creation of the Cost Assessment Data Enterprise (CADE) SAR database. Available since February of 2019, the CADE SAR database is a consolidation of DAMIR SAR data and non-DAMIR legacy SARs. The legacy SARs originate "from recognized meritorious Service databases previously keystroked from authoritative 1969-1996 paper SARs," most notably the RAND constructed and maintained database (OSD CAPE, 2019). Although the CADE SAR database provides users with a consolidated source for SAR data, the data that is in the system is only as good as the data that was available on the original SAR documentation. SAR data overall is known to have missing or invalid data and this problem has transferred into the CADE database. Just like other studies before this, these discrepancies forced us to exclude certain data points which in turn caused our sample size to vary (Hough, 1992). For the program data included in this study, it we assumed that the data from CADE is accurate.

Programs not inputted into the CADE SAR database (i.e., programs that require additional security clearances) along with programs with missing data points were omitted from this analysis. The first part of the data analysis describes more in depth the conditions for inclusion and exclusion criteria used regarding the SARs found in the CADE SAR database.

Data Analysis

Prior to any data analysis, we needed to decide what metrics or factors to use to evaluate schedule and cost estimation growth within MDAPs. To do this we chose to calculate the percentage of estimated schedule growth for MDAPs based on their MS B to IOC time, and cost growth factors based on their estimated cost at MS B versus the cost reported on their last reported SAR.

We started by finding the percentage of schedule estimate growth experienced by a program. This calculation started by first finding the difference between the last reported IOC estimate and the IOC estimate from the MS B SAR. We then divided this difference by the total amount of time between the reported MS B and IOC identified on the latest SAR. Equation 1 displays this.

$$\frac{\textit{Current IOC Estimate - Estimated IOC at MS B}}{\textit{Current IOC Estimate - Actual MS B Date}} \tag{1}$$

To analyze the total program cost growth, we took the current total cost estimate and divided it by the estimated total program cost at the MS B or equivalent mark. Equation 2 displays this calculation that generated the CGFs for our analysis. With this equation a CGF of one equates to no change in the estimates from MS B and the latest SAR, less than one shows an estimate reduction and greater than one shows an increase in the overall cost estimate of the program. This CGF calculation has been utilized previously in cost growth studies conducted by Arena et al. (2006), Younossi et al. (2007), and Kozlak et al. (2017)

$$\frac{\textit{Cost Estimate from Latest SAR}}{\textit{Cost Estimate from MS B SAR}} \quad \textbf{(2)}$$

While it is important to analyze the overall total program CGF, the last approach used to analyze the data was identifying cost growth at the unit level. As already discussed, quantity changes drive some cost growth within MDAPs. These quantity variations can cause dramatic changes and should be considered, analyzing the PAUC CGF accounts for these changes in quantity. To analyze the PAUC changes, we divided the total number of units estimated on the MS B SAR by the total cost estimate on the same SAR. The total number of units includes development and production units. Then we calculated the current PAUC by taking the quantity reported on the latest SAR and dividing that by the latest cost estimate. Equations 3 and 4 highlight these calculations. After those two values were determined, we then divided the latest PAUC by the MS B PAUC to find the CGF for this variable which is the same CGF equation as shown in Equation 2.

$$\frac{\textit{Total \# of Units estimated on MS B SAR}}{\textit{Cost Estimate from MS B SAR}} \quad \textbf{(3)}$$

$$\frac{\textit{Total \# of Units Estimated on Latest SAR}}{\textit{Cost Estimate from Latest SAR}} \quad \textbf{(4)}$$

After identifying the factors needed to evaluate schedule and cost growth, the next step was to collect the data needed for these calculations. The data used for this analysis came from the CADE system. Using the SAR Unit Cost Report along with the Current and Baseline Estimate report and the CADE SAR Data listing, we identified 409 SARs. Because these three reports were inconsistent in providing a thorough listing of all SARs, with each listing having a few SARs that were not on the other reports, we created a separate database, and then consolidated and crosschecked the various listings. This database contained all the SARs identified through these three reports and became the starting point of the data analysis.

The very first step in the analysis was to exclude programs not considered MDAPs. This included removing programs classified in CADE as Pre-MDAP, Other, Special Interest, Major Automated Information System (MAIS), Major System, or Department of Energy Programs. There were also programs annotated on the CADE SAR listing; however, they were not on the SAR Unit Cost Report or the Current and Baselines Report. Although CADE holds records of these SARs, there was no cost or schedule data retrievable for analysis on these programs within CADE. For that reason, we automatically excluded these programs from the analysis.

Some programs that had SARs listed in CADE had either been terminated, transitioned, completely restructured, or re-categorized. When programs are classified under these terms this could mean several things; the programs may have been altered and no longer meet the financial requirements to be considered an ACAT I program. Or in the case of the WIN-T program, the restricting of the program subsequently divided it into three separate WIN-T programs. Programs categorized as transitioned, restructured, or recategorized were excluded from the overall analysis. However, if they led to the creation of a new MDAP there were exceptions. For example, the WIN-T, after being broken into three separate programs, drove the creation of one MDAP that met the requirements to be included into our dataset: the WIN-T increment 2.

Another aspect of SAR data that caused program exclusion was missing milestone data. This missing milestone data means there were no MS B and/or IOC dates available in CADE. Without these dates it was impossible to calculate cost or schedule growth as there was no starting estimate available for analysis. Table 2 shows the overarching inclusion and exclusion listing and the number of programs affected.

The final exclusion criteria for our sample involved adjusting for the maturity level of the younger MDAPs within our sample. Programs that were less than five years old were left out of the analysis. This is because newer programs tend to not have had the opportunity to realize cost and schedule adjustments in their short lifespans. For example, within our schedule sample the mean time it takes for an MDAP to move from MS B to IOC is 8.6 years with 108 of the 120 taking more than five years to reach IOC. Although 10% of our schedule growth sample did reach IOC prior to reaching the 5-year mark, to limit the maturity bias presented by these younger MDAPs, we excluded them from our sample. For our sample this led to the exclusion of seven MDAPs that reached MS B in 2017 or later. This exclusion method was also employed by Younossi et al. (2007) in his study of weapons system cost growth.

Table 2. SAR Inclusion and Exclusion Table

SAR Sample Inclusion & Exclusion Table	
Total Number of SARS available in CADE	409
Programs Classified as Terminated	26
Transitioned or Restructured Programs	11
SAR not Classified as MDAPs*	17
SARs w/ no data available in CADE**	25
SARs with Missing Milestone B Data***	129
Programs < 5 years since MS B	7
Final MDAP SAR Sample	194

*This includes Pre-MDAP, Other, Special Interest, MAIS, Major System, and DoE Program classifications.
**These programs were listed in CADE but had no cost or schedule data avaialable for analysis.
***These programs did not have any MS B data available as a starting point for the cost and schedule growth analysis.

After completing the SAR evaluation and building the initial database, we further parsed the data into three separate databases to explore schedule growth, overall program cost growth, and PAUC cost growth individually. This resulted in further reductions, specific to each growth exploration. To calculate the change in schedule growth, we used two main milestones: MS B and IOC. We then calculated the schedule growth percentage by finding the difference in the IOC estimate at MS B versus the realized MS B to IOC timeframe and dividing that difference by the total actual (or estimated in the case of ongoing MDAPs) time that the programs took to make it from MS B to IOC. This provided us with the percentage of growth seen by each individual MDAP. To get this calculation not only did the SARs need to have the data already discussed, but they also needed to have IOC estimates. Out of the 194 programs that remained from the overall exclusion criteria, 74 more programs did not have IOC estimates available in CADE and were not used in the schedule analysis. This left an overall sample size of 120 programs for evaluation of schedule variations shown in Table 3.

Table 3. Schedule Estimation Sample

Schedule Estimation Sample Size	
MDAP SAR Sample	194
Programs with no IOC Estimates in CADE	74
Final Schedule Estimation Size Sample	120

After finalizing the schedule estimation sample, we moved on to the overall cost growth sample. Finding the CGF consisted of taking the latest SAR cost estimate and dividing it by the estimate at MS B. If a program had no cost data available at the MS B or on the latest SAR, they were excluded from the sample. Overall, as Table 4 outlines, there were 11 additional programs from the final MDAP sample of 194 excluded from the overall CGF portion of the analysis leaving us with a sample of 183 programs.

Table 4.Overall Cost Growth Sample

Overall Cost Estimation Sample Size	
MDAP SAR Sample	194
Programs with missing Cost Data*	11
Final Overall Cost Estimation Sample Size	183
*Missing MS B MS C, or final cost estimate data	

The last data set created identified the PAUC Cost Growth. This allows us to view cost growth while controlling for quantity changes in cost growth. To create this, we only needed the cost estimation data from MS B and from the latest SAR, but we also had to have the quantity data from those SARs. Some SARs within the CADE SAR database did not have this information available and were therefore excluded from this analysis. Table 5 shows the final sample used for the PAUC cost growth analysis.

Table 5. PAUC Cost Growth Sample

PAUC Estimation Sample Size	
MDAP SAR Sample	194
Programs with missing Cost Data*	11
Programs with missing Quantity Data	18
Final PAUC estimation Sample Size	165
*Missing MS B or final cost estimate data	

After finalizing our three databases, we normalized our data, specifically the cost data. Since these programs can take years, even decades to complete, there are many times that their costs are re-baselined to a different FY. There were several programs that had their estimates at MS B set to an earlier FY, while the current estimates were in a different FY. To ensure consistency, we used the current base years for the programs. However, this meant that the programs that faced multiple base years throughout their lifespan needed to be normalized. To do this all the estimates with multiple base years had their earlier cost estimates brought forward using Secretary of the Air Force (SAF) raw inflation indices. During this process, we created a table using the SAF provided inflation calculator tool, using the original estimates base year. Then the inflation factor was identified for the current estimates base year. For example, if a program started in 1989 and then re-baselined in 1995, we created an inflation table with a base year of 1989, and then the raw average inflation index for 1995 was used to normalize the dollar amount.

After we began our analysis by analyzing all the data together, we then split those data points into competed and ongoing programs. As highlighted in Table 6, looking at the schedule growth of programs there are 70 programs that have been completed and 50 that are still ongoing. For overall CGF and PAUC CGF, there were 118 and 102 completed programs, while ongoing programs accounted for 65 and 63 programs, respectively.

Table 6. Completed vs. Ongoing Program Breakout

Completed Vs. Ongoing Programs	
Schedule Difference (Yr) - Completed vs. Ongoing	
Completed Programs	70
Ongoing Programs	50
Total	**120**
Overall CGF - Completed vs. Ongoing	
Completed Programs	118
Ongoing Programs	65
Total	**183**
PAUC CGF - Completed vs. Ongoing	
Completed Programs	102
Ongoing Programs	63
Total	**165**

The last statistical analysis conducted investigated maturity bias. To provide us a maturity bias adjusted view of how MDAPs have changed over the decades we used the three original databases identified in Tables 3-5 and then searched the estimates reported on the SARs five years after the programs reached MS B. Some programs did not last for five years after reaching MS B (i.e., IOC or FOC was reached within 1-4 years after reaching MS B) and were thus excluded from the schedule, overall cost and PAUC cost estimates. Although our sample sizes were reduced slightly due to these circumstances, there we no dramatic sample losses. Theses exclusions are revisited as we go through the statistical analysis in Chapter IV.

Statistical Analysis

For this study, we investigated our MDAP's schedule and cost growth through three different lenses. First, we presented the summary statistics for each individual data set, then conducted a set of non-parametric analyses, and finished our assessments by evaluating the overall variance of the data sets.

The summary statistics consisted of means, medians, standard deviations, interquartile ranges (IQRs), and Coefficients of Variations (CVs) of the observations broken out by decade. When presenting the summary statistics, we investigated trends in the data points, or data that stood out dramatically compared to rest of the data. We performed this investigation two different ways: through numerical tables containing the data points and by using Box Plots to visually analyze the data. The goal was to keep all the data that we could in the overall analysis, however, in the case of extreme outliers, those MDAPs were excluded.

The next part of the analyses consisted of inferential tests, which strove to identify whether there were differences in the data between the decades. As we shift from descriptive statistics and begin analyzing our data with inferential tests there are a few more exclusions applied to the sample. For our completed data, the 2010 decade only had two observations available; these were not enough data points to use within our analysis. For ongoing programs, there were also two more additional exclusions that were performed as there was only one observation a piece in schedule growth percentage for the 1980s and overall CGF for the 1970s.

The standard F-test conducted under an Analysis of Variance (ANOVA) was originally thought to be the best methodology to compare the data. However, after assessing the data for a normal distribution pattern, it was determined that the assumption of a normal distribution could not be substantiated. Even after temporarily eliminating outliers and reassessing, this assumption still could not be met. Since the assumption of normality could not be met, non-parametric analyses were employed.

To test whether there was a statistically significant change in MDAP cost and schedule growth across the decades we analyzed the data with the non-parametric tests. The Kruskal-Wallis test and the Wilcoxon Rank Sum test (herein referred to as just the Wilcoxon test) were employed through JMP 12. The Kruskal-Wallis and Wilcoxon tests allowed us to test the data without assuming a normal distribution for any of the samples. The Kruskal-Wallis, considered the non-parametric equivalent of the one-way ANOVA, compares the mean ranks (MRs) or the medians of the data sets rather than the means. When comparing medians, the distributions of the data must have the same shape; however, our data does not have the same shape of distributions throughout the decades. Since our data does not have distributions with the same shape, we will use the MRs for our comparisons (Laerd Statistics, 2014). The Wilcoxon test is used for the same purpose; however, the Wilcoxon is employed when there are only two groups of data tested. As we continue forward in the analysis, there is one data set, the schedule growth percentage of ongoing programs, that only has observations for two decades available for comparison and thus uses the Wilcoxon test rather than the Kruskal-Wallis.

The Kruskal-Wallis and Wilcoxon tests were tested using an alpha level of 0.05. If the *p*-value is less than .05, then we will reject the null hypothesis in favor of the alternative hypothesis. The following hypotheses are tested utilizing the Kruskal-Wallis test:

$$H_0: MR_{70s} = MR_{80s} = MR_{90s} = MR_{00s} = MR_{10s}$$

$H_{a:}$ At least one decade's MR ≠ the other MRs

This null and alternative hypothesis are applicable to all 12 analyses (three for all data, three for completed MDAPs, three for ongoing MDAPs, and the final three for the maturity bias adjusted data). Because we investigated schedule, overall cost growth, and PAUC cost growth separately, we modify the interpretation of the null hypothesis. For the schedule growth analysis, we have the null hypothesis as:

H_0: There has been no change in schedule estimation growth for MDAPs over the last five decades.

This null hypothesis is applied to the analysis of the schedule growth of all programs, completed, ongoing MDAPS, and maturity bias adjusted analysis. For overall cost growth, we have the null hypothesis as:

H_0: There has been no change in overall cost estimation growth for MDAPs in the last five decades.

This null hypothesis is again applied to the four analyses of the overall CGF: all programs, completed, ongoing MDAPS, and maturity bias adjusted. Lastly, for PAUC cost growth, we have the null hypothesis as:

H_0: There has been no change in the PAUC cost estimation over the last five decades.

This null hypothesis is again applied to the four various analyses of PAUC cost growth.

Unlike the one-way ANOVA, which has the Tukey multiple comparison test available to categorize the various groups of data and aggregate them into groups with like means, the Kruskal-Wallis uses a different approach. The Kruskal-Wallis test itself does not indicate what groups of data are statistically different, only that there is a difference. To determine the decades that differ from one another, we used the Steel-Dwass multi-comparison test. The Steel-Dwass test does not aggregate the various data sets into groups with like data (or means in the case of the Tukey analysis). The Steel-Dwass test utilizes a pairwise comparison to find whether there are statistical differences between two sets of data. The null and alternative hypotheses for the Steel-Dwass are:

$$H_0: MR_1 = MR_2$$

$$H_a: MR_1 \neq MR_2$$

In the Steel-Dwass, two decades at a time are compared. For example, the 1990s will be compared to the 1980, then the 1990s will be compared the 2000s. The Steel-Dwass cycles through all the pairwise comparisons and identifies whether there is a statistical difference between two of the decades. To test this hypothesis, we again used an alpha of 0.05. This is because the Steel-Dwass test within JMP already applies the Bonferroni correction method for conserving the overall level of significance (*Re: Does JMP use Bonferroni correction for ANOVA and Kruskal-Wallis,* 2014).

The last part of our inferential analysis is to compare the variances of the decades. This is to help us assess whether the variance of cost and schedule estimation has changed through the decades. The Brown-Forsythe was employed to test whether the variances were different across the decades. The Brown-Forsythe analyzes standard deviations based on the medians rather than the means of the data in an attempt to minimize the effect of outliers or skewness in the data.

Since our data is not normally distributed, utilizing the Brown-Forsythe test provides a more robust results versus the Levene Test. Although the variance test cannot tell us how accurate we are at cost estimating, it can assist us in determining whether our estimates or estimation errors are consistent across the decades (Brown & Forsythe, 1974, Stephanie, 2015).

If we concluded that the variances were statistically different through the decades, we reviewed the standard deviations to find whether there were patterns of the variances decreasing or increasing through the decades. These tests are again tested with an alpha of 0.05 against the *p*-value produced by the test within JMP. The null and alternative hypotheses of these test are:

H_0: There are no differences in the variances across the decades.

H_a: At least one variance is different that the variances of the other decades.

Summary

The beginning of this chapter discussed the limitation and assumptions we had to make during the data collection process. We then proceeded to discuss the data collected, where it was sourced from, how our samples were created, and how the data was normalized for use in the analyses. The last part of this chapter dives into the tools used to complete the data analyses. We investigate MDAPs through three separate lenses regarding cost and schedule growth: program schedule growth, overall program Cost

Growth Factor (CGF), and the PAUC CGF. The next chapter details our findings and results.

IV. Analysis and Results

Chapter Overview

Chapter IV presents the findings of are analyses. We begin our investigation of cost and schedule growth by looking at the data of ongoing and completed MDAPs together. After our initial evaluation, we then divided the data into separate categories of completed and ongoing MDAPs for further investigation. The last portion of the analysis we look at is the schedule and cost growth estimates at the MS B +5 years points, which provided us a view of estimate growth adjusted for maturity bias.

Analysis of Ongoing and Completed MDAPs

The first analysis we assessed was that containing all our data. This included our ongoing and completed programs with no regards to program maturity. This provides a raw look at how DoD MDAP's cost and schedule estimates stand up against one another across the last five decades. As we begin looking at all this data combined, we start our analysis by exploring the estimated schedule growth percentages realized in these acquisition programs.

Analysis of Schedule Growth Percentage for Ongoing and Completed Programs

Table 7. Schedule Growth Percentage Summary Statistics – Ongoing & Completed Programs

Schedule Growth Percentage Summary Statistics - Ongoing & Completed MDAPs						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	12	24%	24%	0.15	0.27	0.65
1980	21	21%	20%	0.20	0.36	0.93
1990	27	32%	21%	0.28	0.44	0.87
2000	35	32%	21%	0.50	0.31	1.58
2010	25	13%	8%	0.31	0.28	2.34

Table 7 provides a quick outline of the summary statistics for the schedule growth percentages of MDAPs for the last five decades. As we can see the 1990s and 2000s have the highest mean schedule growth, while the 1970s has the highest median growth. The 2010s have the highest Coefficient of Variation (CV) showing that this data is the least consistent data.

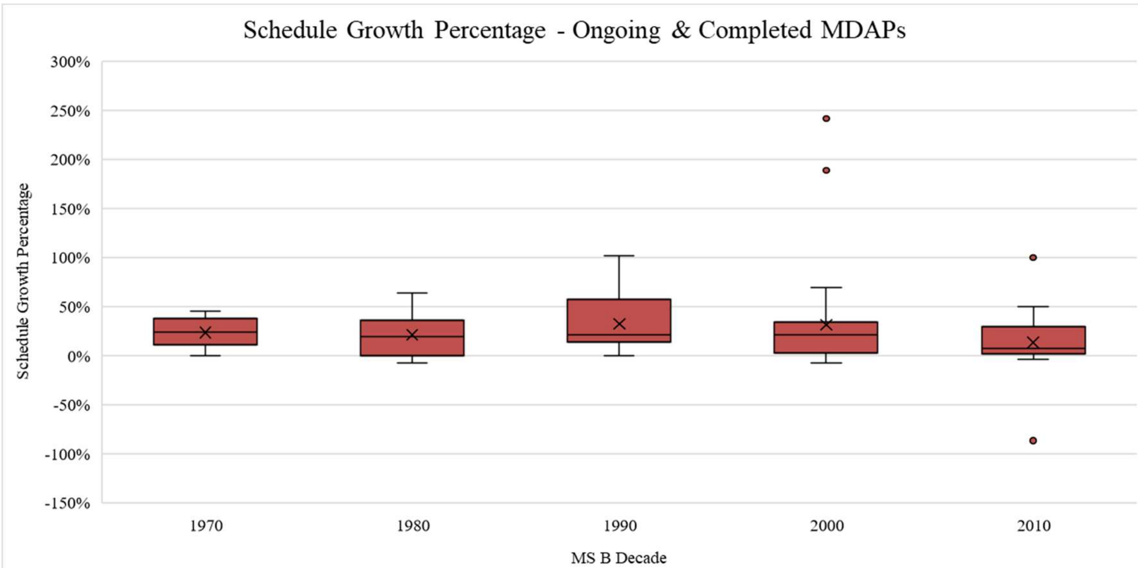


Figure 4. Box Plot – Schedule Growth Percentage

Figure 4 presents the visualization of the data outlined in Table 7. As we can see the 2000s and the 2010s are the only decades with any outliers; they also both have the two highest CVs. Although identified as outliers, we had no reason to exclude them and consequently left them in through the rest of the analysis. Table 8 shows the accompanying results of the Kruskal-Wallis test for schedule growth. As evident by the *p*-value of 0.2123, we fail to reject the null hypothesis and assume that there is no statistical difference in the schedule growth percentage through the last five decades.

Table 8. Kruskal-Wallis Test Results – Schedule Growth Percentage

Kruskal-Wallis Test		
ChiSquare	DF	Prob>ChiSq
5.8283	4	0.2123

Table 9. Variance Test Results – Schedule Growth Percentage

Schedule Growth Percentage Variance Tests - Ongoing & Completed MDAPs				
Test	F Ratio	DFNum	DFDen	Prob > F
Brown-Forsythe	0.6618	4	115	0.6198

Regarding the variance test shown in Table 9, we see that although the 2000s and 2010s had higher CVs, they did not have statistically significant differences in the variances according to the results of the Brown-Forsythe test. This shows that the data has comparable variances throughout the five decades analyzed.

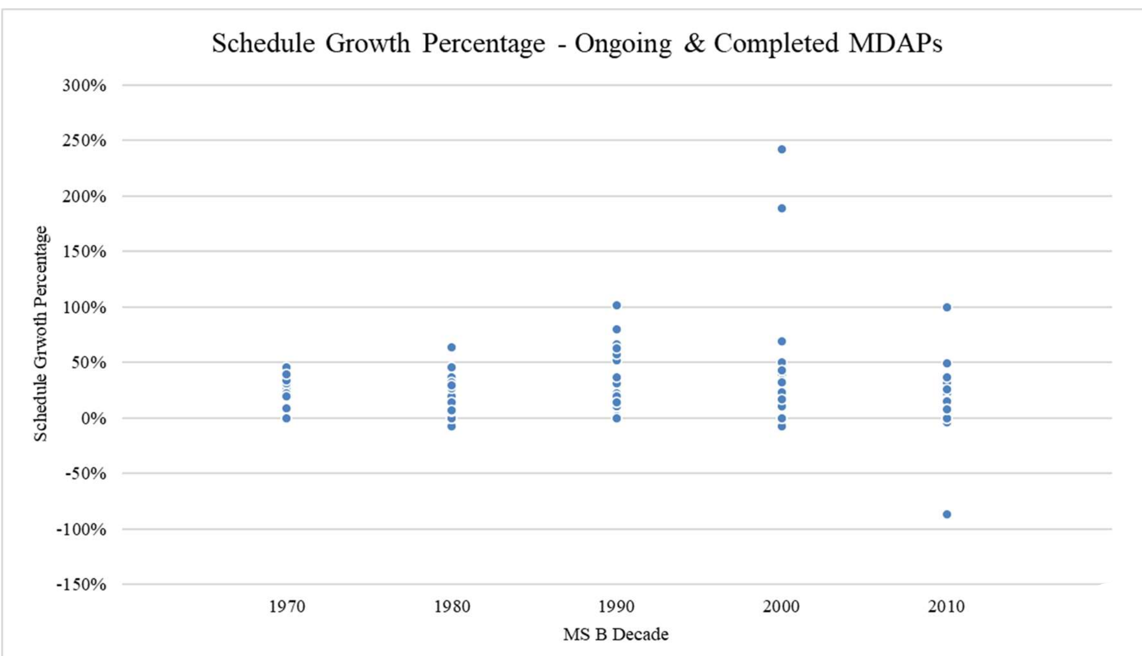


Figure 5. Schedule Growth Percentage of Ongoing & Completed Programs

Figure 5 provides us with another visualization of the schedule growth percentages seen by the ongoing and completed MDAPs. Even when plotting all the data points throughout the decades, the schedule growth percentages are decently consistent, with many of the MDAPs realizing between 0% and 50% estimate growth.

The next portion of our analysis reviewed overall CGF for ongoing and completed MDAPs. As we can observe in Table 10, the 1970s and 1980s have the highest CVs, the highest standard deviations, and the highest means. Exploring the 1980s a little further we also note that this decade had the highest mean CGF, the lowest median CGF, and a standard deviation over three times larger than the next highest standard deviation.

Analysis of Overall CGFs for Ongoing and Completed Programs

Table 10. Overall CGF Summary Statistics – Ongoing & Completed Programs

Overall CGF Summary Statistics - All Programs						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	29	2.83	1.44	3.62	2.88	1.28
1980	46	3.27	0.99	11.94	1.26	3.65
1990	37	1.66	1.26	1.17	1.51	0.71
2000	42	1.33	1.12	0.71	0.53	0.54
2010	29	1.14	1.02	0.37	0.25	0.32

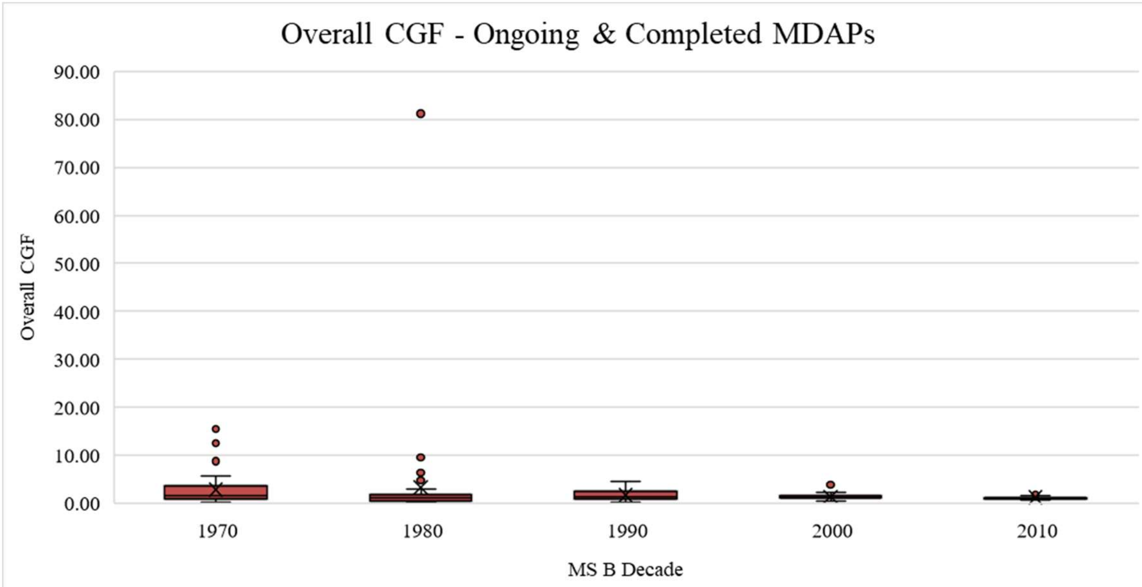


Figure 6. Box Plot - Overall CGF

Looking at the Box Plot in Figure 6, we can see why the 1980s data in Table 10 shows such a high standard deviation. The 1980s has one program that realized an unusual amount of cost growth compared to all of the other MDAPs. This program was the DDG 51 MDAP, which was a Naval Ship. Originally it was estimated that they would purchase 14 ships; however, the most recent SAR shows them purchasing 95. Since this program has experienced abnormal growth, prior to continuing with our inferential anlaysis this outlier was removed.

Table 11. Overall CGF Summary Statistics – Ongoing & Completed Programs (Excluding DDG 51)

Overall CGF Summary Statistics - Ongoing & Completed (Excluding DDG 51)						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	29	2.83	1.44	3.62	2.88	1.28
1980	45	1.54	0.98	2.11	1.24	1.37
1990	37	1.66	1.26	1.17	1.51	0.71
2000	42	1.33	1.12	0.71	0.53	0.54
2010	29	1.14	1.02	0.37	0.25	0.32

After removing the DDG-51 MDAP from the data, the 1980s mean dropped dramatically along with the standard deviation and the CV. Table 11 oultines these changes in the data. This drove the 1970s to have the highest summary statistics excluding the CV.

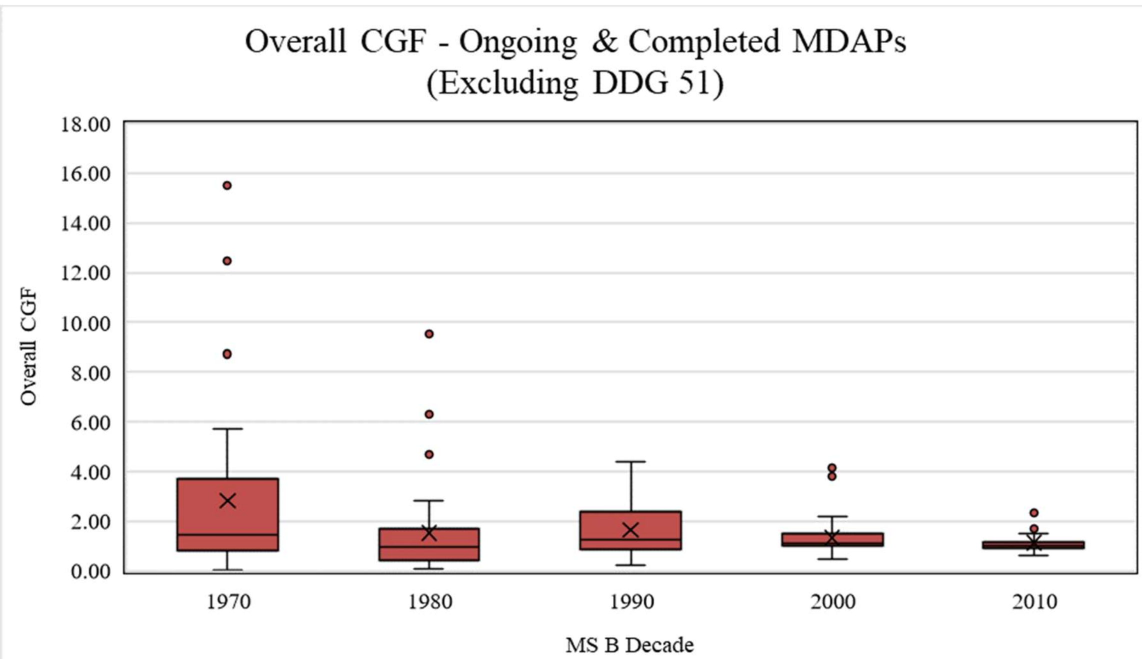


Figure 7. Box and Whisher Chart - Overall CGF (Excluding DDG 51)

The Box Plot in Figure 7, which excludes the DDG 51 program, provides a significantly enhanced picture of the distibution of the overall CGF of MDAPs over the last five decades. As we can see there are still a few outliers in the 1970s, 1980s, 2000, and 2010s however, none were as dramatic as the DDG 51 program.

Table 12. Kruskal-Wallis Test Results – Overall CGF

Kruskal-Wallis Test		
ChiSquare	DF	Prob>ChiSq
8.3	4	0.0812

When running the Kruskal-Wallis test to see if there were any differences in the overall CGF across the decades, we found that there were no statistical differences in the overall CGF through the decades. The *p*-value of 0.08, although not significant, could be viewed as moderately significant and may be impacted by having a larger sample or less variability within the data.

Table 13. Variance Test Results – Overall CGF

Overall CGF Variance Tests - Ongoing & Completed MDAPs (Excluding DDG 51)				
Test	F Ratio	DFNum	DFDen	Prob > F
Brown-Forsythe	5.13	4	177	0.0006

Although there were no statistical differences in the overall CGF thorough the decades, when testing the variance, we can see that the Brown-Forsythe test indicated that there are statistically significant differences in the variances across the decades.

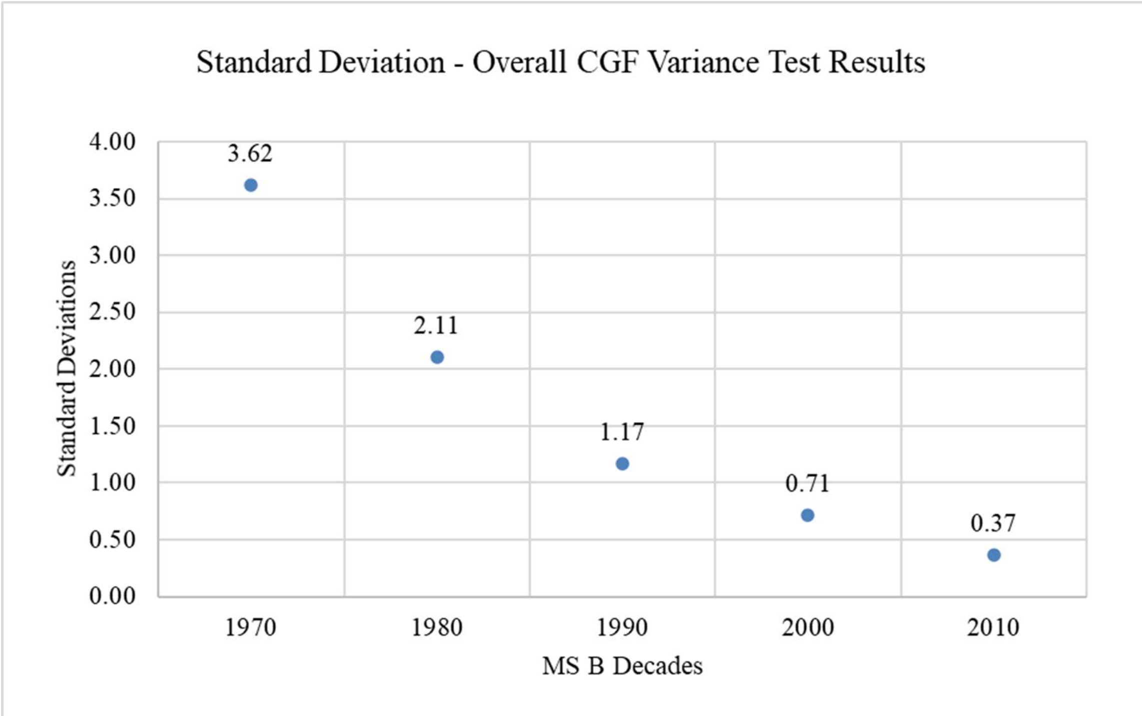


Figure 8. Standard Deviations - Overall CGF of Ongoing & Completed Programs

Figure 8 displays how the standard deviations compare to one each and assists us in visualizing where the differences lie. As we can see the standard deviations of the data have been consistently decreasing throughout the decades, with the 1970s standard deviation of 3.62 being higher than all the other standard deviations.

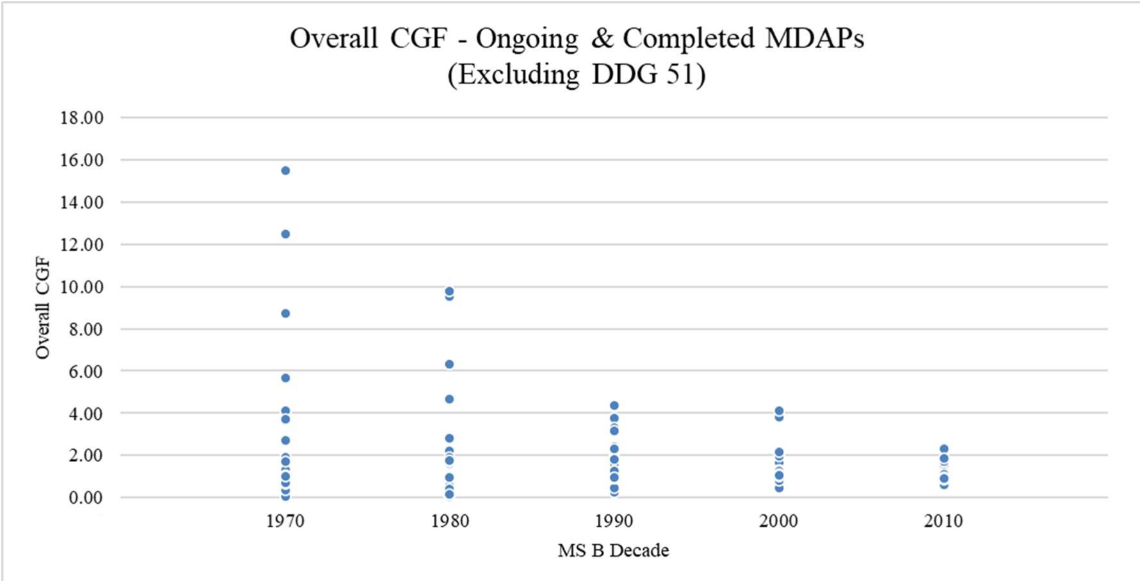


Figure 9. Overall CGF of Ongoing & Completed Programs

The overall CGFs plotted in Figure 9 tells the same story as that of the variance tests. Although the five decades did not show a statistically significant difference in the overall CGF, we can clearly see that the variability of the overall CGFs throughout has steadily decreased since the 1970s.

Analysis of PAUC CGFs for Ongoing and Completed Programs

Table 14. PAUC CGF Summary Statistics – Ongoing & Completed Programs

	PAUC CGF Summary Statistics - Ongoing & Completed MDAPs					
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	25	1.55	1.36	0.91	1.10	0.59
1980	35	2.39	1.10	3.22	0.94	1.35
1990	37	2.15	1.26	2.70	0.99	1.25
2000	39	2.04	1.11	5.46	0.39	2.67
2010	29	1.01	1.02	0.14	0.19	0.14

Our last look for ongoing and completed MDAPs together had us investigating quantity adjusted cost growth of MDAPs by analyzing the PAUC CGF. When reviewing the summary statistics shown in Table 14, we can see that the 1980s holds the highest mean PAUC CGF, but the 1990s has the highest median PAUC CGF. The 2000s though have the highest standard deviation and highest CV.

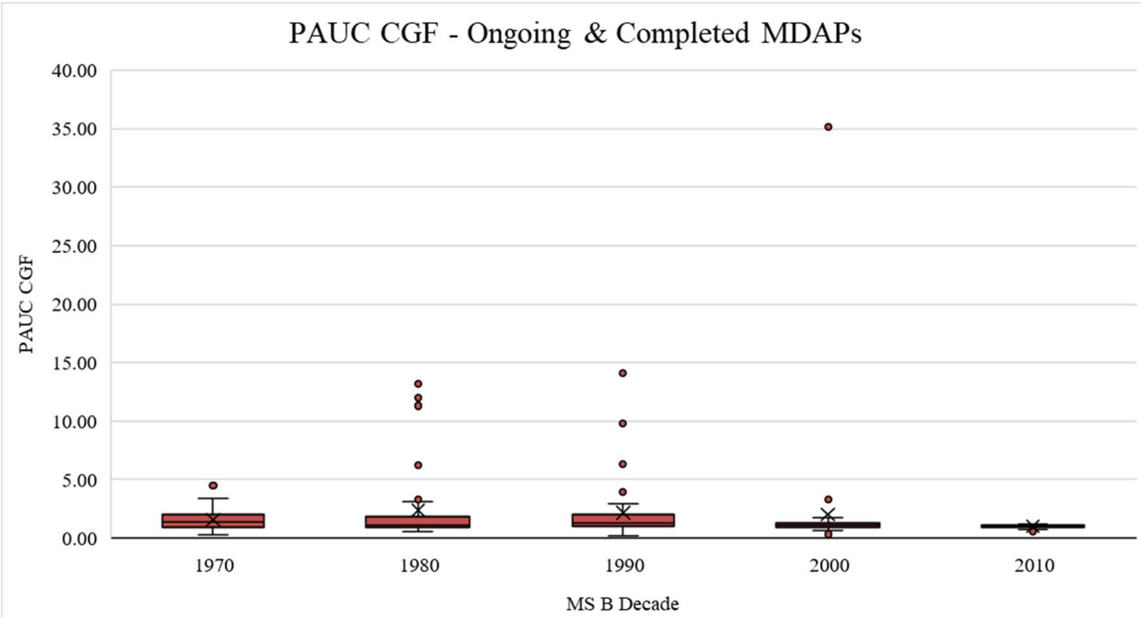


Figure 10. Box Plot – PAUC CGF

Looking at the data plotted in Figure 10 further reveals how the PAUC CGFs appear when compared to one another. As we can see the 1980s and 1990s have the highest number of overall outliers above the IQR, but the 2000s boasts the highest outlier of all the decades. This outlier is the C-130 AMP MDAP, which originally planned to acquire 519 units, but reported only purchasing nine on its most recent SAR. This drove PAUC estimated cost from $7.26 million dollars per unit to $255.18 million dollars per unit. Since this outlier is markedly different from the outliers in any other decade, we removed this program prior to conducting its inferential analysis.

Table 15. PAUC CGF Summary Statistics – Ongoing & Completed Programs (Excluding C-130 AMP)

PAUC CGF Summary Statistics - Ongoing & Completed MDAPs (Excluding C-130 AMP)						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	25	1.55	1.36	0.91	1.10	0.59
1980	35	2.39	1.10	3.22	0.94	1.35
1990	37	2.15	1.26	2.70	0.99	1.25
2000	38	1.17	1.11	0.46	0.38	0.39
2010	29	1.01	1.02	0.14	0.19	0.14

As we can see in Table 15, excluding the C-130 program from our analysis reduced the 2000's standard deviation and CV dramatically, moving them from the highest to the second lowest, with only the 2010s showing a lower standard deviation and overall variance within its data.

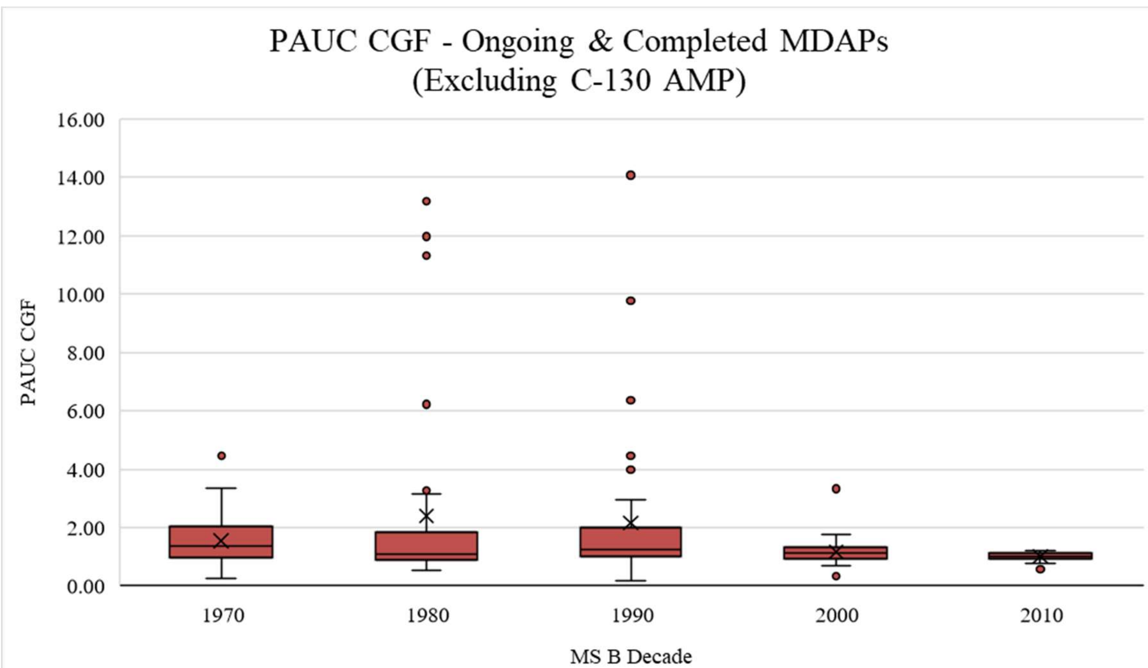


Figure 11. Box Plot – PAUC CGF (Excluding C-130 AMP)

Removing the C-130 AMP program from our dataset also leaves us a clearer look at the IQRs in each decade. As we can see in Figure 11, both the 1980s and 1990s have five outliers while the rest of decades have one or two. Although the 1980s and 1990s do have higher outliers than the other decades, neither of the two decades have any outliers that are extraordinarily different from the rest of the data as we saw with the C-130 AMP. For this reason these outliers were not excluded from our anlaysis.

Table 16. Kruskal-Wallis Test Result- PAUC CGF

Kruskal-Wallis Test		
ChiSquare	**DF**	**Prob>ChiSq**
10.6936	4	0.0302

The results for the PAUC CGF Kruskal-Wallis test revealed a *p*-value of 0.0302 which is lower than our alpha of 0.05. This leads us to reject our null hypothesis and assume that there are differences between the mean ranks (MRs) of the PAUC CGFs across the decades.

Table 17. Steel-Dwass Results – PAUC CGF

Steel-Dwass Results	
Observation Level	**P-Value**
2010-1990	**0.0411**
2010-1970	**0.0505**

Since our Kruskal-Wallis test leads us assume that there were differences between MRs of the various decades, we further investigated by conducting the Steel-Dwass test to see what decades differed. Although the Steel-Dwass test checks all the pairwise comparisons, Table 17 outlines the results of the two comparisons that identified statistically significant differences. With our alpha set at 0.05, the decades of 2010-1990 show that there is a statistical difference as the *p*-value of 0.0411 which is less than our alpha. Our second finding identified 2010 compared to the 1970s. Although not technically below the alpha of 0.05, the *p*-value of 0.0505 was also close enough to the alpha that it is worth reporting as a moderate finding.

Table 18. Variance Test Results – PAUC CGF

PAUC CGF Variance Tests - Ongoing & Completed MDAPs (Excluding C-130 AMP)				
Test	**F Ratio**	**DFNum**	**DFDen**	**Prob > F**
Brown-Forsythe	3.4306	4	159	0.0101

The next portion of our inferential analysis is to test the variances of the PAUC CGF between the decades. In Table 18 we can see that the results of the Brown-Forsythe test yielded a *p*-value of 0.0101, which is below the alpha of 0.05. This low *p*-value revealed a statistically significant differences in the variances across the decades.

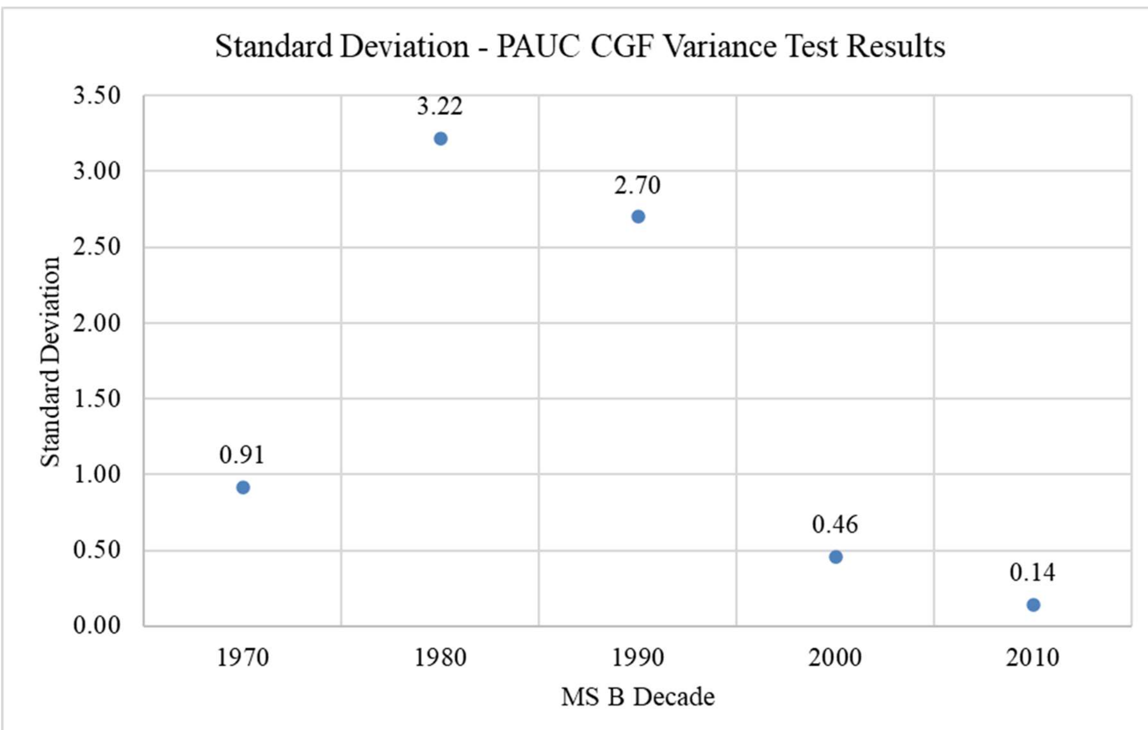


Figure 12. Standard Deviations - PAUC CGF of Ongoing & Completed Programs

Looking at the standard deviations plotted in Figure 12 we can see that the 1980s and 1990s have much higher standard deviations than the other three decades while the 2000s and 2010s are decreasing.

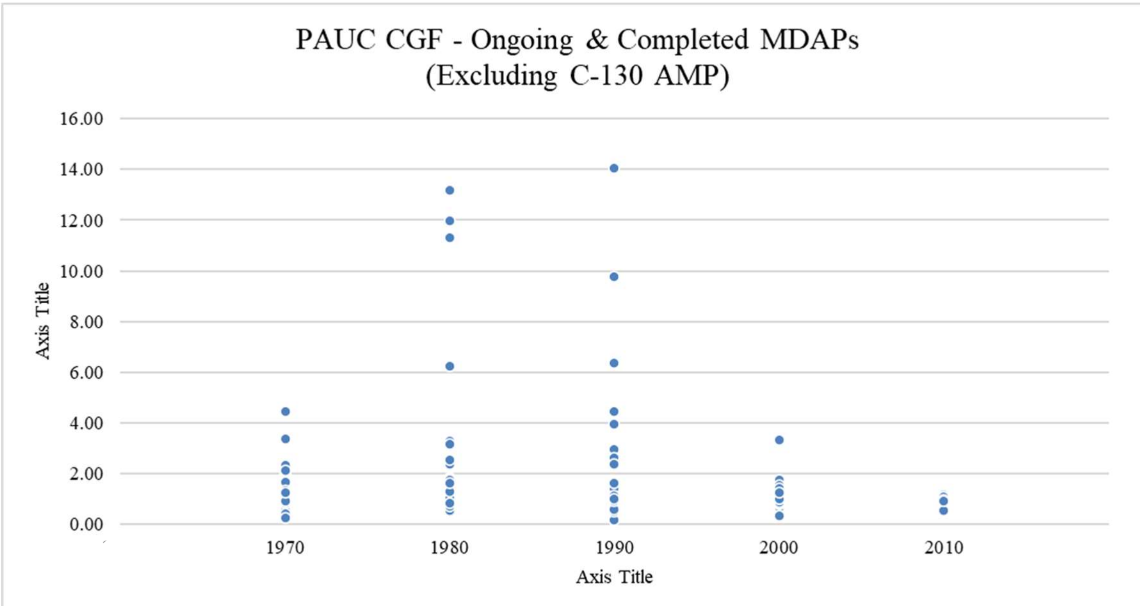


Figure 13. PAUC CGF of Ongoing & Completed Programs

The plotted PAUC CGFs in Figure 13 reiterate the story that the statistical analysis has conveyed with one small exception. The 1980s and 1990s show the most dispersion within its data while, at least since the 1990s, the PAUC CGFs have shifted downwards and have realized less variance. Now although the 1980s has a higher mean, standard deviation, and higher CV, it was not identified as statistically different than any of the other decades regarding the MRs. When reviewing the medians though, you can see that the 1970s, along with the 1990s is higher than the others. Along with the higher medians, the 1970s and 1990s also have higher MRs than the 2010 decade, which is why the Steel-Dwass pairwise comparisons identified these decades as statistically different from one another.

Overall, the 1980s provides an interesting story. Using the Kruskal-Wallis and analyzing the MRs, we can identify that the 1980s are not statistically different than any other decade, but when it comes to variances, the 1980s provided a starting point in which the variances for each decade started decreasing. According to the Brown-Forsythe test this higher standard deviation contributed to a statistically significant difference in the variances.

Regarding the schedule growth percentages and CGFs for ongoing and completed programs combined, we identified that there were statistically significant differences in the PAUC CGFs. These differences were between the 1990s and 2010s at a significance level of 0.0411 and again (at a moderate level) between the 1970s and 2010s at a significance of .0505. Within these analyses, we also found that there were differences in the variances of CGFs across the decades. These differences were highlighted in the overall CGFs and the PAUC CGFs. Overall CGFS saw downward trends in the variances since the 1970s, while the PAUC CGFs saw an initial increase in the 1980s then downward movement from the 1980s through the 2010s.

Analysis of Completed MDAPs

Although analyzing all the data together provides an interesting baseline and noteworthy insights, partitioning the data into completed and ongoing programs to analyze each separately is critical. This separation allows us to have an increased granularity within the analyses of the schedule and cost growth in MDAPs. To begin, we separated out and investigated the completed MDAPs and started exploring our schedule growth percentage.

Analysis of Schedule Growth Percentages for Completed Programs

Table 19. Schedule Growth Percentage Summary Statistics – Completed Programs

Schedule Growth Percentage Summary Statistics - Completed MDAPs						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	12	24%	24%	0.15	0.27	0.65
1980	20	20%	20%	0.19	0.34	0.97
1990	23	27%	18%	0.26	0.23	0.99
2000	13	34%	11%	0.66	0.38	1.95
2010	2	4%	4%	0.06	0.08	1.41

When inspecting the summary statistics of the schedule growth of completed programs identified in Table 19, throughout the decades, there are a few interesting trends to point out. We can see the mean schedule growth percentage bounces up and down with no discernable pattern while the medians of each decade do seem to consistently trend downward. Although the medians are decreasing, the standard deviations and CVs are growing; that is until the 2010s.

The 2010s show a decrease across the board, however this decade did not provide enough observations to provide any substantial insight. Since there are only two MDAPs observed as completed during this decade, we excluded the 2010s from this section of the analyses and from the following two analyses looking at overall CGFs and PAUC CGFs.

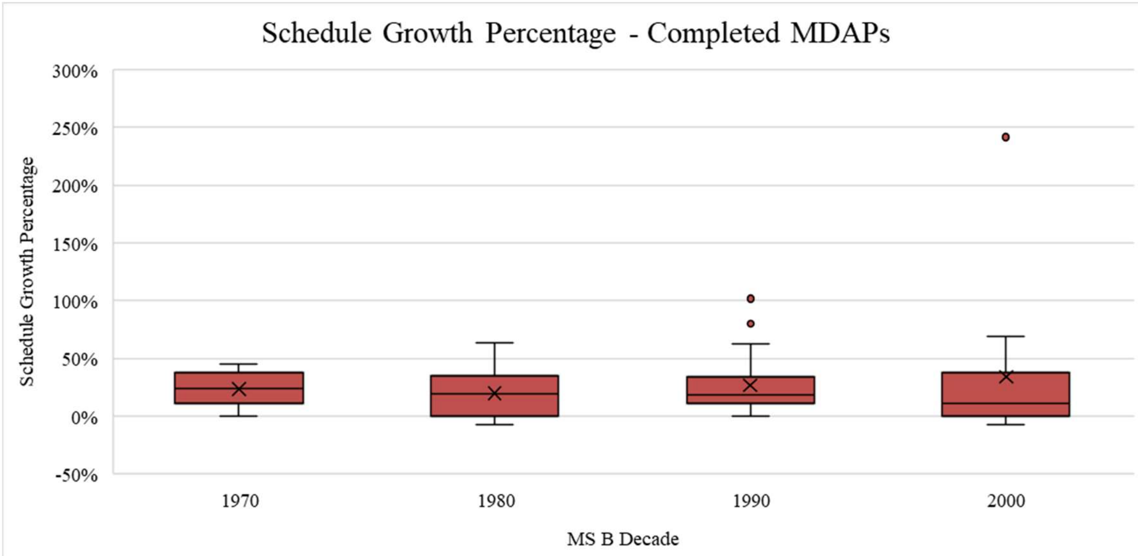


Figure 14. Box Plot – Schedule Growth Percentage

As we have already concluded that there are not enough completed programs in the 2010s to be included in the part of the completed programs section of the analysis, we went ahead and excluded that decade from our Box Plot. With the remaining data we can see in Figure 14 that there is an obvious outlier within the 2000s. This outlier belongs the Joint Mine Resistant Ambush Protection (MRAP) vehicles MDAP. Since this is a singular outlier that is markedly different that the rest of the data, it was excluded from the remainder of the inferential analyses.

Table 20. Schedule Growth Percentage Summary Statistics – Completed Programs (Excluding Joint MRAP)

Schedule Growth Percentage Summary Statistics - Completed MPADs (Excluding Joint MRAP)						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	12	24%	24%	0.15	0.27	0.65
1980	20	20%	20%	0.19	0.34	0.97
1990	23	27%	18%	0.26	0.23	0.99
2000	12	17%	11%	0.22	0.30	1.35

Although we excluded the largest outlier from the data analysis, Table 20 still shows that the 2000s have the highest CV, although this decade does possess the lowest mean and median schedule growth percentages. Overall, there does not seem to be any discernable pattern regarding the mean, standard deviations, and overall IQRs, but it does seem as through the medians are decreasing and the CVs are increasing.

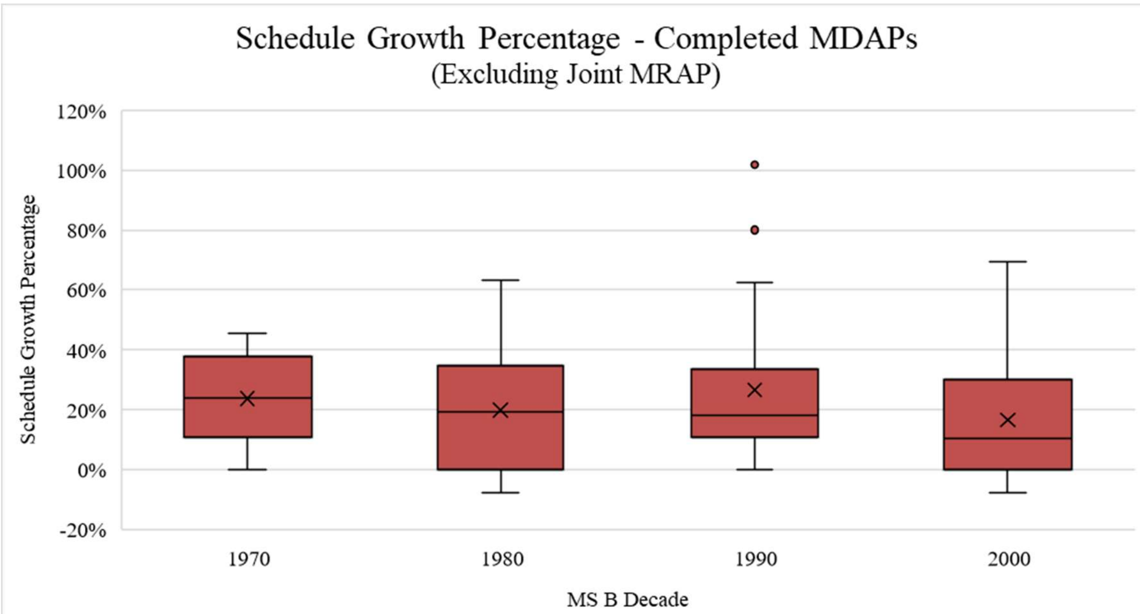


Figure 15. Box Plot – Schedule Growth Percentage (Excluding Joint MRAP)

After excluding the Joint MRAP, the remaining data generated the Box Plot in Figure 15. Examining this chart yields no discernable patterns between the decades in regards to the schedule growth percentages. Athough there are a couple of outliers in the 1990s, they do not severely skew the standard devation or the CV of the decade and thus are included in the inferential analysis.

Table 21. Kruskal-Wallis Test Results – Schedule Growth Percentage

Kruskal-Wallis Test		
ChiSquare	DF	Prob>ChiSq
2.2569	3	0.5208

The results of the Kruskal-Wallis test, as shown in Table 21, provide us a *p*-value of 0.5208, which leads us to assume that null hypothesis is correct and that there are no differences in the MRs of MDAPs through the decades.

Table 22. Variance Test Results – Schedule Growth Percentage

Schedule Growth Percentage Variance Tests - Completed MDAPs (Excluding Joint MRAP)				
Test	F Ratio	DFNum	DFDen	Prob > F
Brown-Forsythe	1.1548	3	64	0.334

Moving on, we conducted our variance test for the schedule growth percentages of completed programs. Using the Brown-Forsythe variance test, we can see in Table 22 that the resulting *p*-value is 0.334. Since this test's *p*-value was higher than our alpha of 0.05 we again assume that our null hypothesis is accurate and that there are no differences in the variances of schedule growth percentages through the decades.

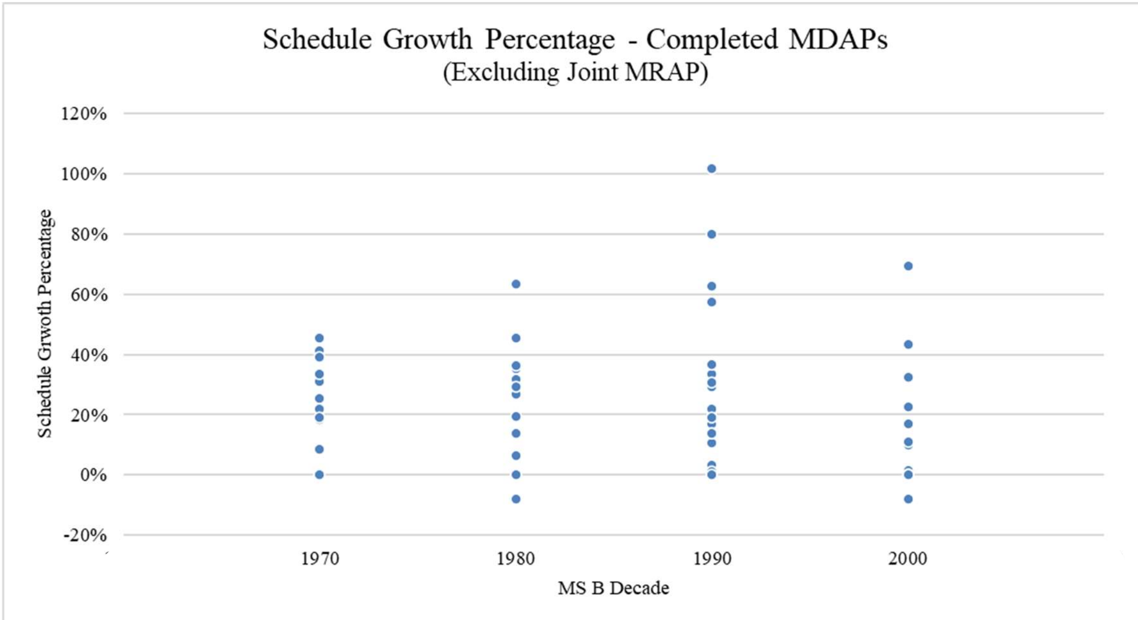


Figure 16. Schedule Growth Percentage of Completed Programs

The plotted data points shown in Figure 16. allow us to visually distinguish how the various completed programs compare. Although the 1990s have MDAPs with the highest schedule growth percentage, they also have a large cluster of programs that saw under 40% schedule growth which contribute to the lower median seen within the decade. The 2000s have lower mean and median data, but its overall dispersion of the data points from below 0% (indicating negative schedule growth) to over 60% generated higher IQRs and CVs for the decade.

Analysis of Overall CGFs for Completed Programs

Table 23. Overall CGF Summary Statistics – Completed Programs

	Overall CGF Summary Statistics - Completed					
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	28	2.83	1.37	3.69	2.91	1.30
1980	41	1.50	0.98	2.15	1.31	1.43
1990	29	1.41	1.01	1.00	1.27	0.71
2000	18	1.14	1.11	0.44	0.25	0.39

Shifting our attention to overall CGF of completed programs, we started examining the data by reviewing the summary statistics outlined in Table 23. As stated before, the 2000s did not have enough observations to warrant inclusion in this analysis, so we begin by automatically excluding them from the data set. Looking at the rest of the data we can see that the means, standard deviations and IQRs all seem to be decreasing through the decades. However, it is not obvious whether these differences are statistically significant.

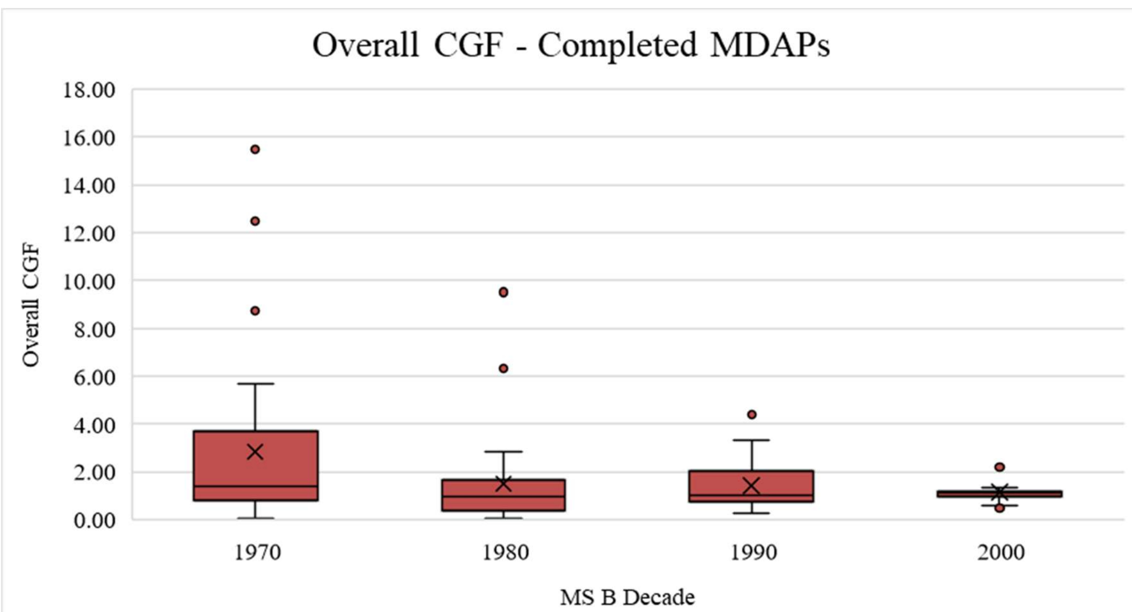


Figure 17. Box Plot – Overall CGF

When conducting the visual analysis of the IQRs using the Box Plot in Figure 17, there are a few trends to note. As identified in Table 23, the 2000s show the least amount of variance while the 1970s has the most outliers of all the decades. When viewing the IQRs of all the data across the decades there appears to be a downward trend in the extent and scope of outliers within each decade. While the outliers seems to be decreasing through the decades, the median CGFs consistantly fall under 2.0 and show no signs of increasing or decreasing consistantly across time. Lastly, it was identified in Table 23 that the means are trending downwards; however, when conducting a visual inspection there doesn't look to be a distinct difference between the means of the 80s, 90s, and 2000s. Only the 1970s appear to have a different mean than the other decades.

Table 24. Kruskal-Wallis Test Results – Overall CGF

Kruskal-Wallis Test		
ChiSquare	DF	Prob>ChiSq
5.6448	3	0.1302

The results of the Kruskal-Wallis test, identified in Table 24, provided us with a *p*-value of 0.1302, which is greater than our alpha of 0.05. For this inferential test we failed to reject the null hypothesis and move forward assuming that there are no differences in the MRs of the data between the decades.

Table 25. Variance Test Results – Overall CGF

Overall CGF Variance Tests - Completed MDAPs				
Test	F Ratio	DFNum	DFDen	Prob > F
Brown-Forsythe	3.1729	3	112	0.027

The next step in our analysis led us to test the overall CGF variances across the decades. As shown in Table 25, the variance test calculated a *p*-value of 0.027 which is lower than our alpha of 0.05. This conclusion led us to reject the null hypothesis and assume that there are statistically significant differences in the variances of the data through the decades.

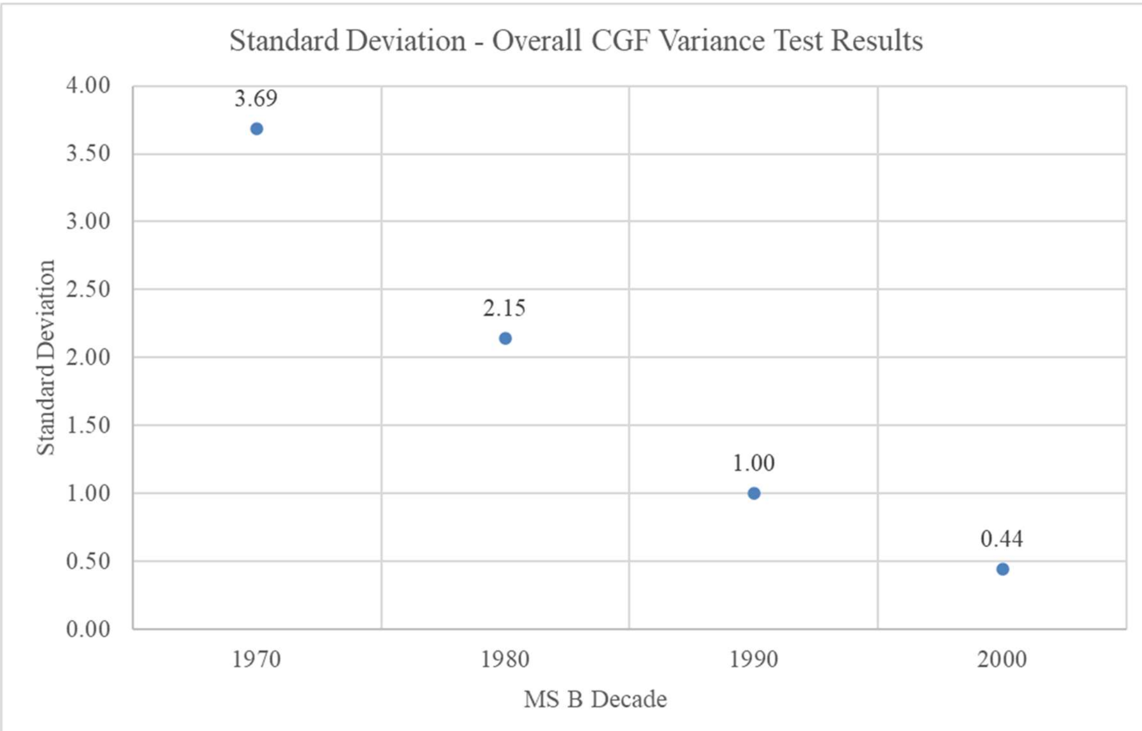


Figure 18. Standard Deviations - Overall CGF of Completed Programs

Plotting the standards deviations of the decades illustrates how the variances have changed over time. Figure 18 reveals that there has been a steady decrease in the standard deviations of overall CGFs for completed MDAPs.

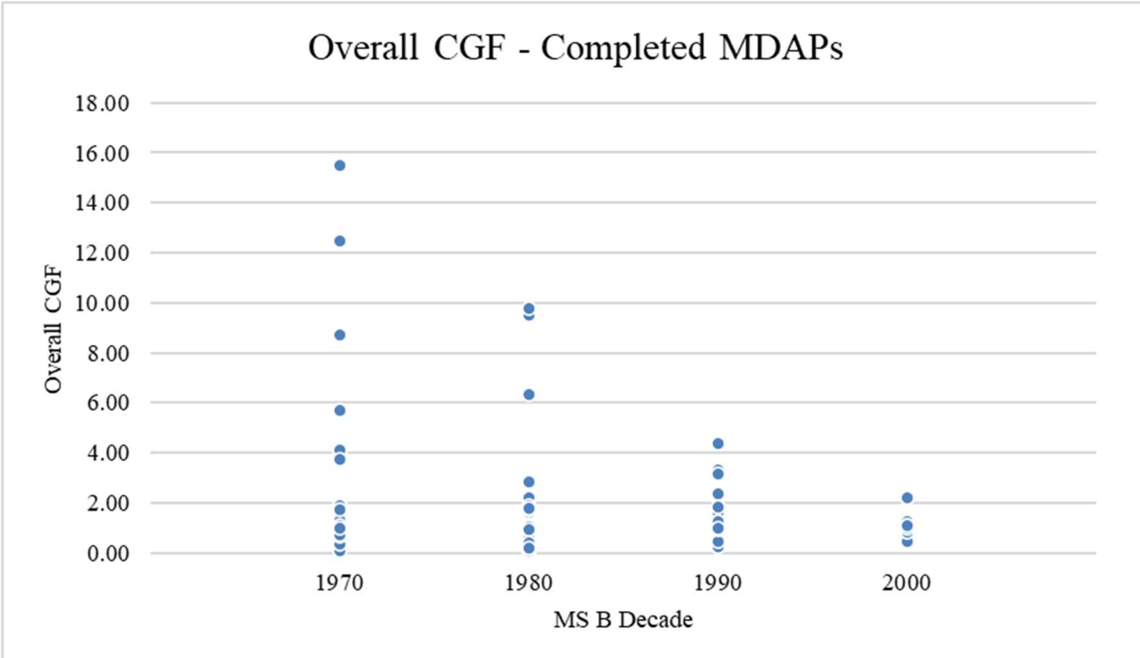


Figure 19. Overall CGF of Completed Programs

Figure 19 leads us to our last visual inspection of the overall CGF datapoints for the completed MDAPs in our sample. As we have already discussed there is a visually noticeable downward trend in the dispersion of the CGFs through the decades.

Analysis of PAUC CGFs for Completed Programs

Table 26. PAUC CGF Summary Statistics – Completed Programs

PAUC CGF Summary Statistics - Completed						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	25	1.55	1.36	0.91	1.10	0.59
1980	31	2.03	1.05	2.83	0.90	1.40
1990	29	2.29	1.26	2.88	1.37	1.26
2000	15	3.45	1.09	8.79	0.25	2.55

The next dataset we studied were the quantity adjusted PAUC CGFs. Again, we excluded the 2010s due to the relatively small sample size of completed programs contained within the decade. While the 2010s were excluded, the remaining decades (the 1970s through the 2000s) were included and their summary statistics are outlined in

Table 26. Within this data there appears to be an increase in the PAUC CGF means and the standard deviations throughout the decades. The standard deviation of the 2000s is over twice as high as the next highest standard deviation belonging to the 1990s. This peculiar discrepancy required further investigation.

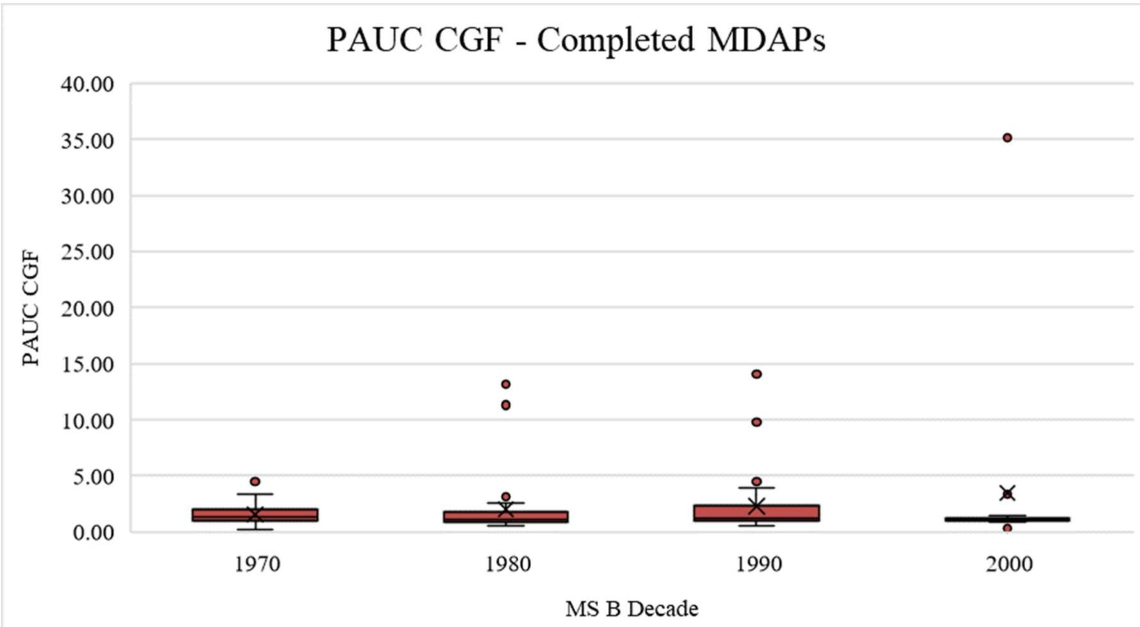


Figure 20. Box and Whisher Chart – PAUC CGF

Upon further investigation using the Box Plot in Figure 20 we identified what appears to be the source of the irregular standard deviation. This data point belongs to the C-130 AMP program, which we already determined needed to be excluded from the anlaysis due to its extreme value. Because of this exclusion we reran the summary statistics and completed the inferential anlaysis without the C-130 AMP MDAP included in our data.

Table 27. PAUC CGF Summary Statistics – Completed Programs (Excluding C130 AMP)

PAUC CGF Summary Statistics - Completed (Excluding C-130 AMP)						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	25	1.55	1.36	0.91	1.10	0.59
1980	31	2.03	1.05	2.83	0.90	1.40
1990	29	2.29	1.26	2.88	1.37	1.26
2000	14	1.19	1.07	0.66	0.19	0.56

With the exclusion of the C-130 AMP, there is a substantial drop in the mean, standard deviation, and the CV for the PAUC CGF data from the 2000s, leaving the 1990s to take the lead in these statistics. Table 27 outlines the final summary statistics for the PAUC CGFs.

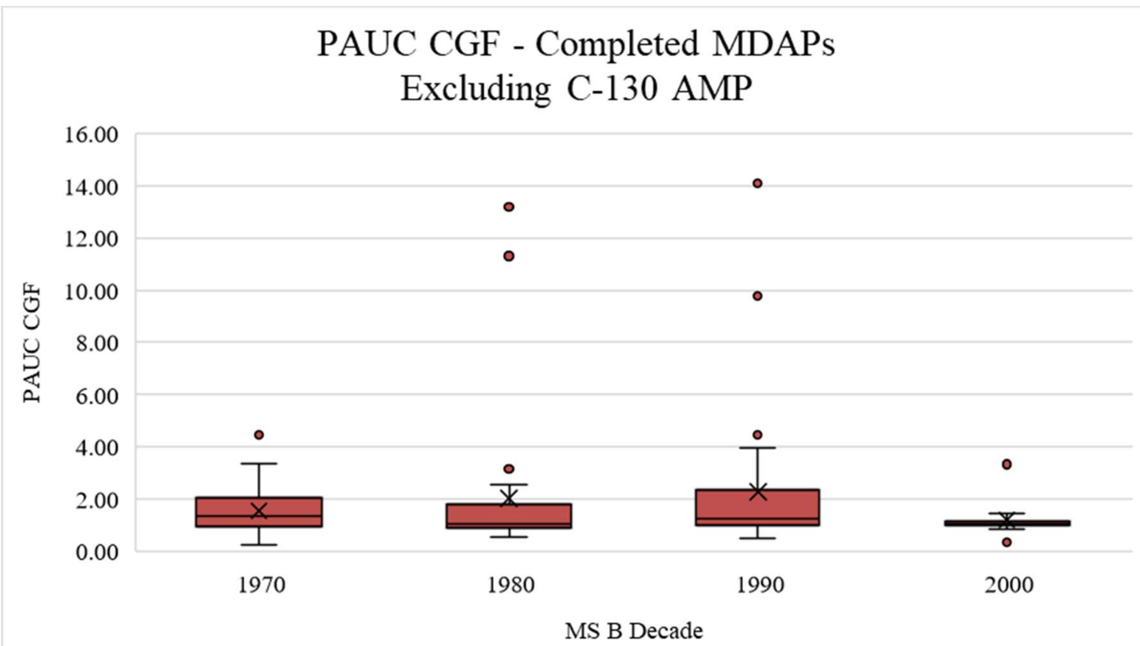


Figure 21. Box Plot – PAUC CGF (Excluding C-130 AMP)

Examining Figure 21 we identified a few additional outliers noticable in 1980s and 1990s. However, these datapoints are neither singular outliers nor are they extreme enough to be excluded from the rest of the anlaysis.

Table 28. Kruskal-Wallis Test Results –PAUC CGF

Kruskal-Wallis Test		
ChiSquare	**DF**	**Prob>ChiSq**
3.2777	3	0.3508

Table 28 displays the results of the Kruskal-Wallis test identifying a *p*-value of 0.3508. Since this value is higher than our alpha of 0.05, we assume that there were no statistical differences in PAUC CGFs between the various decades.

Table 29. Variance Test Results – PAUC CGF

PAUC CGF Variance Tests - Completed MDAPs (Excluding C-130 AMP)				
Test	**F Ratio**	**DFNum**	**DFDen**	**Prob > F**
Brown-Forsythe	0.9340	3	95	0.4275

Shifting our attention to Table 29, we identified that the results of the Brown-Forsythe variance test provided a *p*-value of 0.4275 leading us to assume that our null hypothesis is correct and that the variances throughout the decades do not differ.

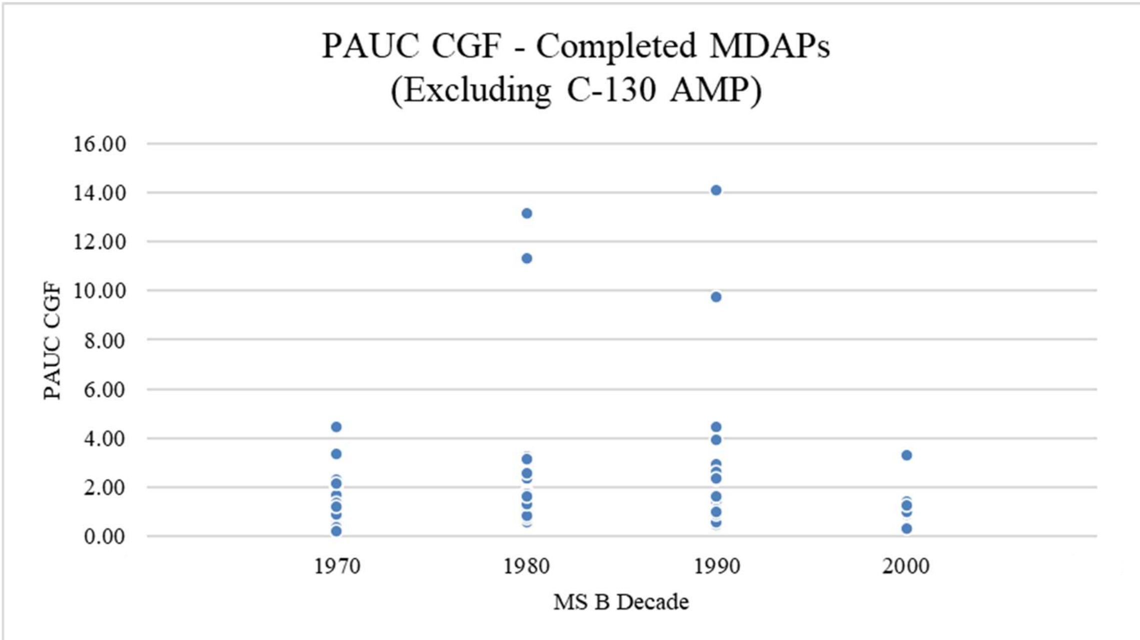


Figure 22. PAUC CGF of Completed Programs

Taking one last glance at the PAUC CGFs of completed programs in Figure 22, we can again see that the 1980s and 1990s have outliers that are higher than the 1970s and 2000s factors. Minus a few exceptions, most of the PAUC CGFs seem clustered under 4.0 through all the decades which, when combined with the results of the Kruskal-Wallis and variances tests, show that these estimates for completed MDAPs in our sample were relatively 44consistent throughout the 40 years analyzed.

Our analysis of completed programs did not result in any statistically significant differences between the decades in the schedule growth percentages, overall CGFs, or PAUC CGFs utilizing our non-parametric analysis. However, we did find that there has been a statistically significant difference between the variances of the overall CGFs identified for MDAPs. Exploring these differences led to the conclusion that these variances have realized a downward trend from the 1970 though the 2000s.

Analysis of Ongoing MDAPs

The next portion of this study evaluates the cost and schedule growth of ongoing MDAPs. As we investigate the data contained within the sample of ongoing programs, we noticed that there were a few programs that were quite a bit older than the rest of our observations. The 2000s and 2010s contained the bulk of our ongoing programs, but there are a couple of programs from the 1980s and 1990s that are still open and ongoing.

Analysis of Schedule Growth Percentages for Ongoing Programs

Table 30. Schedule Summary Statistics – Ongoing Programs

Schedule Growth Percentage Summary Statistics - Ongoing MDAPs						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	0					
1980	1	45%			0.00	
1990	4	65%	12%	0.12	0.23	0.18
2000	22	30%	39%	0.39	0.31	1.29
2010	23	14%	32%	0.32	0.30	2.29

As we begin our analysis of ongoing MDAPs, we started with exploring schedule growth percentages. From the summary statistics shown in Table 30 we quickly identified that neither the 1970s, 1980s, nor the 1990s had enough observations to have any statistical power in our analysis. Since there were not at least five observations in any of these decades we excluded them from the rest of our analysis.

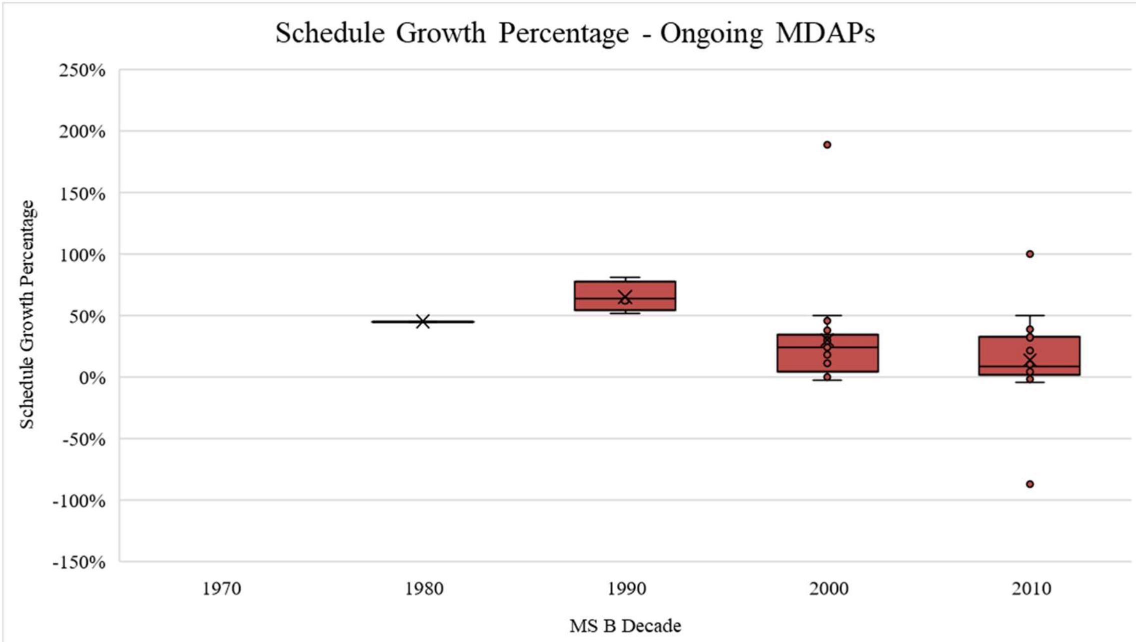


Figure 23. Box Plot – Schedule Growth Percentage

Although they are not included in the analysis, Figure 23 illustrates what schedule growth looks like for all of the onoing programs identified. As we can see the one program from the 1980s, the Sense and Destroy Armor (SADARM) munitions, has seen less than 50% schedule growth during its lifespan and the ongoing MDAPs from the 1980s have an IQR higher than the programs in the 2000s and 2010s. Interestingly though, both the 2000s and 2010s seem to have outliers that have higher realized schedule growth than the ongoing programs from the 1980s and 1990s.

Table 31. Schedule Growth Percentage Summary Statistics – Ongoing Programs

Schedule Growth Percentage Summary Statistics - Ongoing MDAPs						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
2000	22	30%	39%	0.39	0.31	1.29
2010	23	14%	32%	0.32	0.30	2.29

Although it is interesting to examine the entirety of our data together, as already identified, the only decades with enough observations to conduct a meaningful statistical analysis are the 2000s and 2010s. Between these two decades we can see in Table 31 that both the mean and median schedule growth percentages have decreased from the programs categorized in the 2000s versus the 2010s, but the mean percentages saw a more striking decrease than the median schedule growth. The standard deviations and the IQRs seem to be decently similar, but the 2010s, while having a slightly lower standard deviation, does have a bit higher CV.

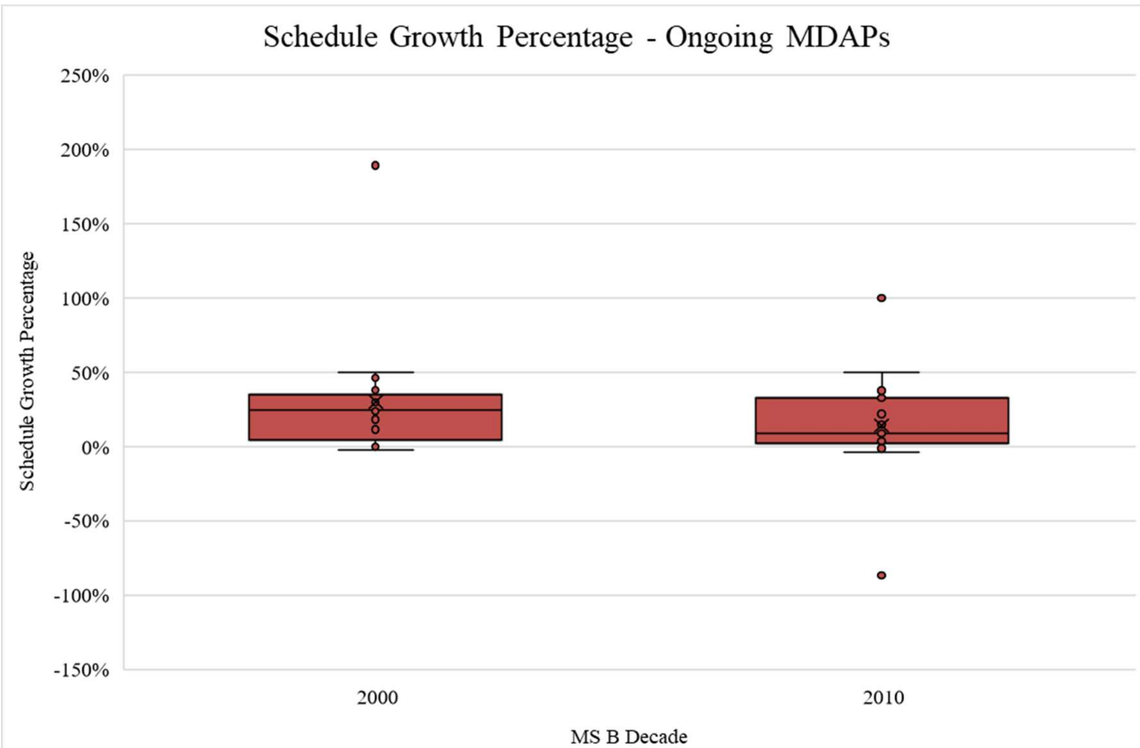


Figure 24. Box Plot – Schedule Growth Percentage

Looking at the Box Plot in Figure 24 we can see that both decades have outliers, one seen in the 2000s, and two visible in the 2010s. Because both decades have outliers, and neither decade contains one that is radically abnormal, these programs were left in the rest of the anlaysis. Aside from the outliers, the IQRs provide an interesting sight. They seem to be extremely similar in their locations in regards to their upper and lower limits, however, the means seem to be in noticeably different locations within the IQRs.

Table 32. Wilcoxon Test Results – Schedule Growth Percentage

Wilcoxon Test		
ChiSquare	**DF**	**Prob>ChiSq**
2.6025	1	0.1067

Although the summary statistics and Box Plot provide some interesting insights into the schedule growth of MDAPs, the inferential statistical tests tell us whether there is a statistically significant difference between the MRs of the decades. Since there are only two decades analyzed in this test, we shifted from the Kruskal-Wallis Test to the Wilcoxon Rank Sum test. The Wilcoxon test, displayed in Table 32, provided us with a p-value of 0.1067, leading us to assume that there is not statistically significant difference between the MRs across the two decades.

Table 33. Variance Test Results – Schedule Growth Percentage

Schedule Growth Percentage Variance Tests - Ongoing MDAPs				
Test	**F Ratio**	**DFNum**	**DFDen**	**Prob > F**
Brown-Forsythe	0.0058	1	43	0.9398

Shifting our attention to the variance tests results outlined Table 33, we can see that the Brown-Forsythe variance test produced a *p*-value above 0.9, revealing that there is no statistically significant difference between the variances of the two the decades analyzed.

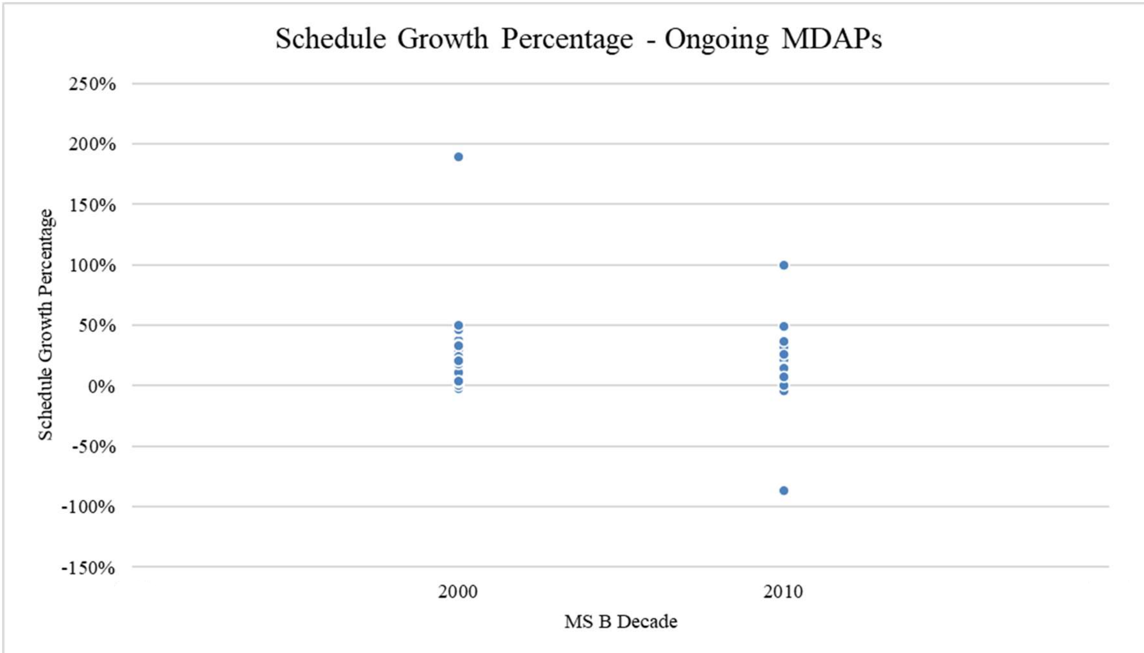


Figure 25. Schedule Growth Percentage of Ongoing Programs

Much like the Box Plot displayed in Figure 24, the schedule growth percentages plotted in Figure 25 shows that most of the programs schedule growth falls between 0 and 50% schedule growth, with both decades containing a at least program that stands out from the cluster of datapoints in each individual decade. For the schedule growth percentage of ongoing programs, there is not a visually or statistically significant difference in the data from the 2000s and 2010s.

Analysis of Overall CGFs for Ongoing Programs

Table 34. Overall CGF Summary Statistics – Ongoing Programs

Overall CGF Summary Statistics - Ongoing MDAPs						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1980	5	17.79	1.14	35.51	42.01	2.00
1990	8	2.56	2.79	1.35	2.24	0.53
2000	24	1.47	1.21	0.84	0.64	0.57
2010	27	1.16	1.04	0.37	0.24	0.32

The next portion of our study led us to examine the overall CGFs of ongoing MDAPs. As we began looking at the summary statistics shown in Table 34, the 1980s immediately stood out against the data from the other decades. The mean for this decade, at 17.79 is considerably higher than the next three decades. The standard deviation, IQR, and CV are also much higher than the summary statistics identified for the next three decades. Interestingly though, the median is less than both the 1990s and 2000s.

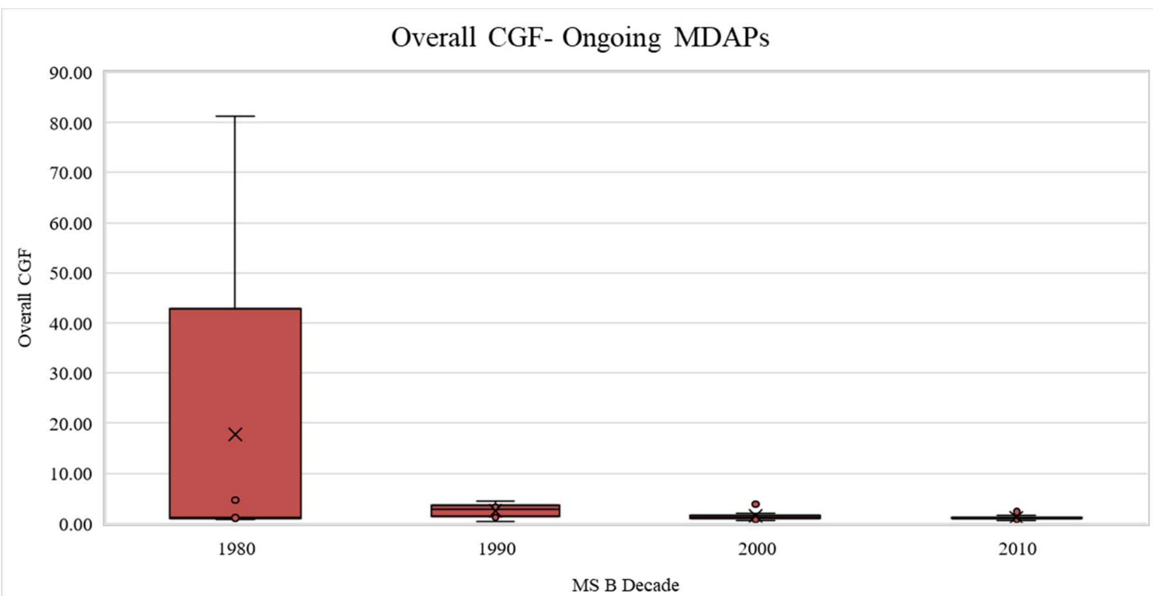


Figure 265. Box Plot – Overall CGF

In Figure 25 we can visually identify that the 1980s IQR is dramatically larger than the IQRs for the other decades. The driving force of this is the DDG-51 program that was identified in the first section of this analysis as a major oulier that needed to be excluded. With this information we again decided to exclude this MDAP, which reduced our sample from the 1980s to four observations. Excluding the DDG-51 programs from our data subsequently drove us to exclud the entirety of the 1980s data from the statistical anlaysis since there were no longer enough observations.

Table 35. Overall CGF Summary Statistics – Ongoing Programs (Excluding 1980s Programs)

Overall CGF Summary Statistics - Ongoing MDAPs						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1990	8	2.56	2.79	1.35	2.24	0.53
2000	24	1.47	1.21	0.84	0.64	0.57
2010	27	1.16	1.04	0.37	0.24	0.32

With the exclusion of the 1980's data, we were left with the overall cost growth data for three decades. Looking at the summary statistics for this data, located in Table 35, we can see that the 1990s has the highest mean and median data, along with highest standard deviations and IQRs. Although the 90s takes the lead with most of the data, the 2000s has the highest CV. The 2010s obtained the lowest position for all the summary statistics presented in Table 35. Although there seems to be a downward trend regarding the means and median, we next conduct the Kruskal-Wallis test to determine whether the differences in the MRs of the CGFs are statistically significant.

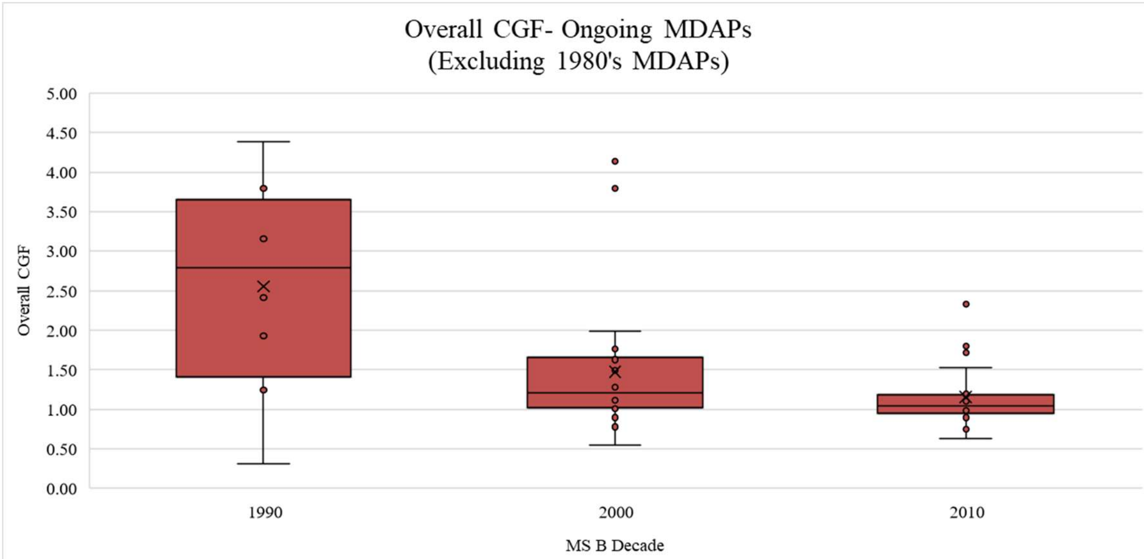


Figure 27. Box Plot – Overall CGF (Excluding 1980s MDAPs)

Continuing with our analysis of the summary statatistics, Figure 27 illustrates the data identified in Table 35. We can see that the 1990's overall CGF for ongoing programs has the highest mean and median along with the largest IQR range. The 2000s and the 2010s have much smaller IQRs, but both have outliers within their datasets outside of their respective IQRs. Although these two decades have a few outliers, none are substantial enough to be excluded from our investigation of the inferential statistics.

Table 36. Kruskal-Wallis Test Results – Overall CGF

Kruskal-Wallis Test		
ChiSquare	DF	Prob>ChiSq
11.1786	2	0.0069

Since we again have three decades of data to compare, we shifted back to using the Kruskal-Wallis test. Table 36 shows that the *p*-value produced from this test is 0.0069 which is less than our alpha 0.05 and led us to reject our null hypothesis and assume that there are differences in the MRs between the decades.

Table 37. Steel-Dwass Results – Overall CGF

Steel-Dwass Results	
Observation Level	**P-Value**
1990-2010	**0.0096**

Using the Steel-Dwass test we further investigated what decades were different from one another. Using this test, we found that there were statistically significant differences between the 1990s and 2010s. In Table 37 we can see that the *p*-value calculated using the Steel-Dwass equals 0.0096, which again is less than our alpha of 0.05.

Table 38. Variance Test Results – Overall CGF

Overall CGF Variance Tests - Ongoing MDAPs				
Test	**F Ratio**	**DFNum**	**DFDen**	**Prob > F**
Brown-Forsythe	0.3155	2		0.002

Table 38 displays the results Brown-Forsythe variance tests for the overall CGF of ongoing programs produced *p*-values of 0.002, which was lower than our alpha of 0.05. These results show that there is a statistically significant differences in the variance of the data across the decades.

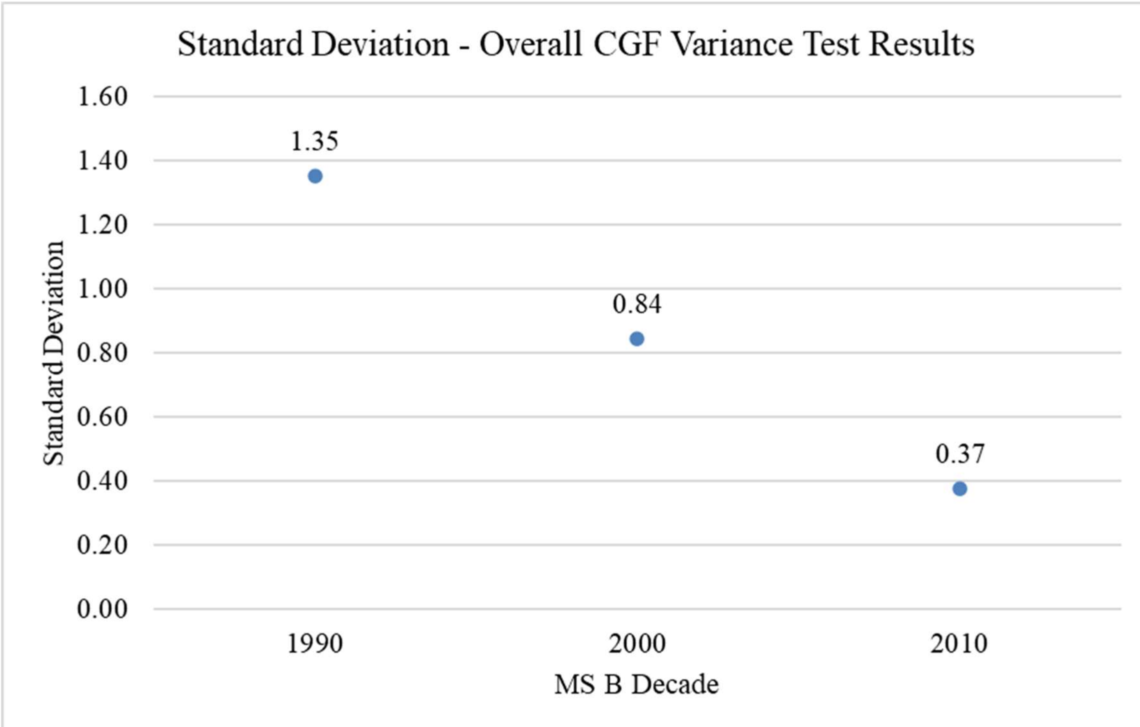


Figure 28. Overall CGF of Ongoing Programs

Since our Brown-Forsythe variance test identified that there were statistically significant differences in the standard deviations, plotting the data provides insight into where these differences occur and whether there are any trends. As we can see in Figure 28, the standard deviations have been decreasing since the 1990s.

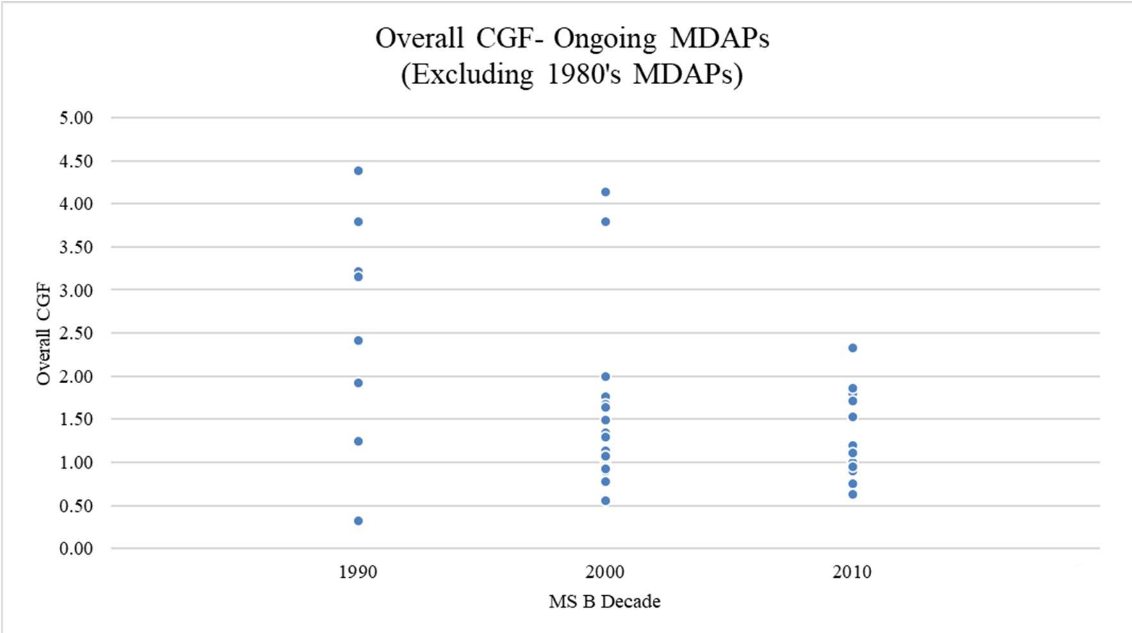


Figure 29. Overall CGF of Ongoing Programs

Prior to moving on to the next piece of our analysis, we took one last visual look at the distribution of the overall CGFs for each decade. As we reviewed the overall CGF plotted in Figure 29, we identified that 1990s and 2000s see highs and lows that are relatively close; however, the 2000s has a cluster of data points under the 2.0 while the 1990s display more evenly distributed datapoints between 0 and 4.5. This more even distribution contrasts with the clusters that we see in the 2000s and 2010s.

Analysis of PAUC CGFs for Ongoing Programs

Table 39. PAUC CGF Summary Statistics – Ongoing Programs

PAUC CGF Summary Statistics - Ongoing MDAPs						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1980	4	5.19	3.83	1.25	9.32	0.98
1990	8	1.67	1.25	1.96	0.97	1.17
2000	24	1.16	1.15	0.30	0.44	0.26
2010	27	1.01	1.02	0.14	0.19	0.14

The last portion of our investigation of ongoing programs reviewed the PAUC CGF of MDAPs. Although we had data from every decade, there are only four data points in the 1980s as shown in Table 39. Again, since there were so few observations available for the 1980s, we excluded this decade from our overall analysis.

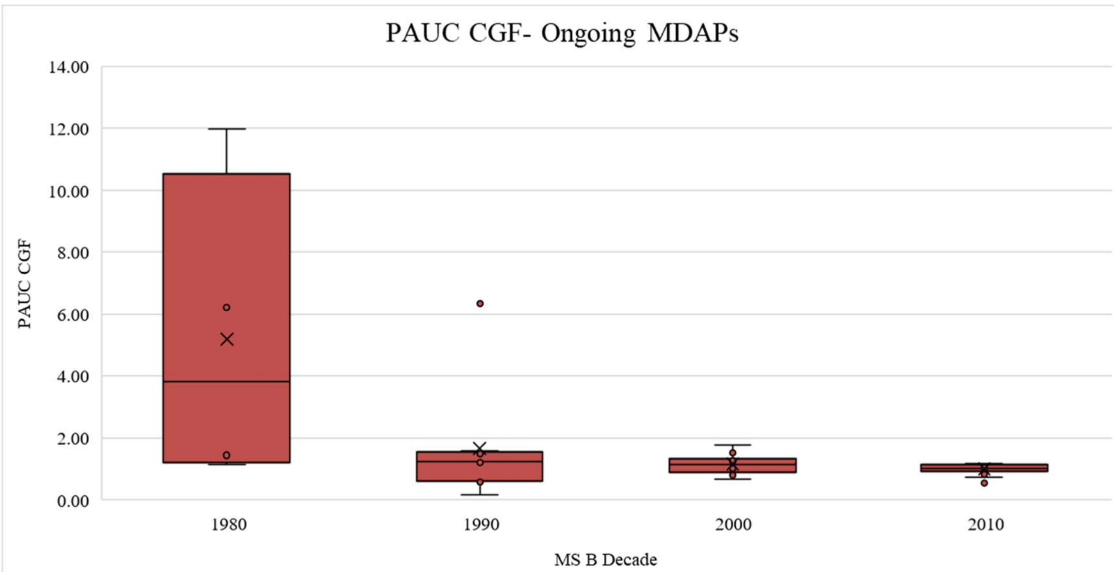


Figure 30. Box Plot – PAUC CGF

Before completely excluding the 1980s from our analysis, the Box Plot in Figure 30 was created to view that data against the rest of the decades. It looks as though the 1980s has quite a bit of range within the PAUC CGFs for the four ongoing programs. The 1980s also has a higher mean and median than those of the other next three decades.

Figure 30 also led us to identify that the 1990s had a substantial outlier within its data. This datapoint is tied to the National Security Space Launch (NSSL) program. Due to its relative location in comparison to the rest of the data from the 1990s, 2000s, and 2010s, this outlier was excluded from the analysis of summary statistics and inferential statistics for our PAUC CGF of ongoing programs.

Table 40. PAUC CGF Summary Statistics – Ongoing Programs (Excluding NSSL)

Overall CGF Summary Statistics - Ongoing MDAPs (Excluding NSSL)						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1990	7	1.00	1.21	0.53	0.91	0.54
2000	24	1.16	1.15	0.30	0.44	0.26
2010	27	1.01	1.02	0.14	0.19	0.14

After excluding the 1980s data and the NSSL program from the 1990s, we began our investigation of the PAUC CGFs of ongoing programs by looking at the summary statistics outlines in Table 40. Although there does not seem to be any patterns regarding the mean data, it does seem as through the medians, standard deviations, IQRs, and CVs are generally getting smaller through the decades. Moving our attention to Figure 31, we can see how the data outlined for each decade in Table 40 compares to one another.

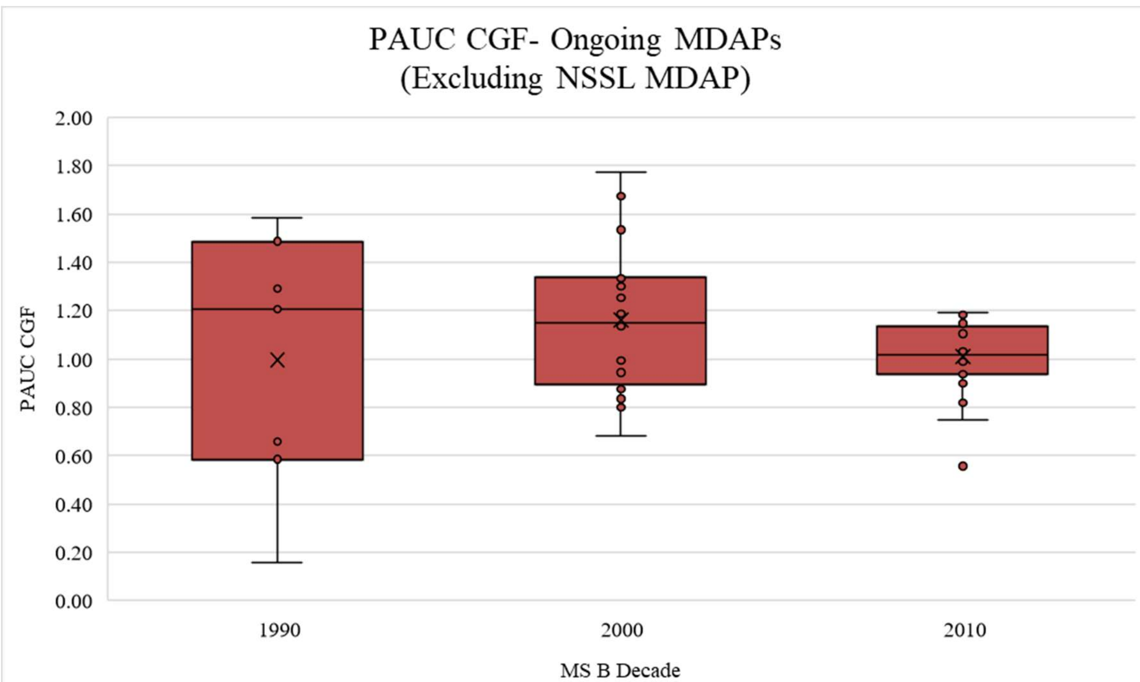


Figure 31. Box Plot – PAUC CGF (Excluding NSSL)

The 1990s, as already identified, has the largest IQR, ranging from under 0.2 to around 1.6, but no longer has any outliers. The 2010s has the smallest IQR, however there is one datapoint in the 2010s that is an outlier for the decade. Although it lies outside the IQR, it does not stand out as irregular against the rest of the data and was included in the analysis.

Table 41. Kruskal-Wallis Test Results – PAUC CGF

Kruskal-Wallis Test		
ChiSquare	DF	Prob>ChiSq
2.722	2	0.2564

Moving to our Kruskal-Wallis test, we found that the *p*-value produced was 0.2564 which is presented in Table 41. This *p*-value is higher than our alpha of 0.05 which leads us to assume that there are no differences in the MRs of the PAUC CGF through the three decades investigated.

Table 42. Variance Test Results – PAUC CGF

PAUC CGF Variance Tests - Ongoing MDAPs (Excluding NSSL)				
Test	F Ratio	DFNum	DFDen	Prob > F
Brown-Forsythe	6.9731	2	56	0.0001

The last test conducted for our ongoing MDAPs was the Brown-Forsythe variance test. The results of this test, shown in Table 42, revealed a *p*-value of 0.0001. This *p*-value, which is lower than our alpha of 0.05, led us to reject the null hypothesis and assume that there were differences between the variances across the three decades.

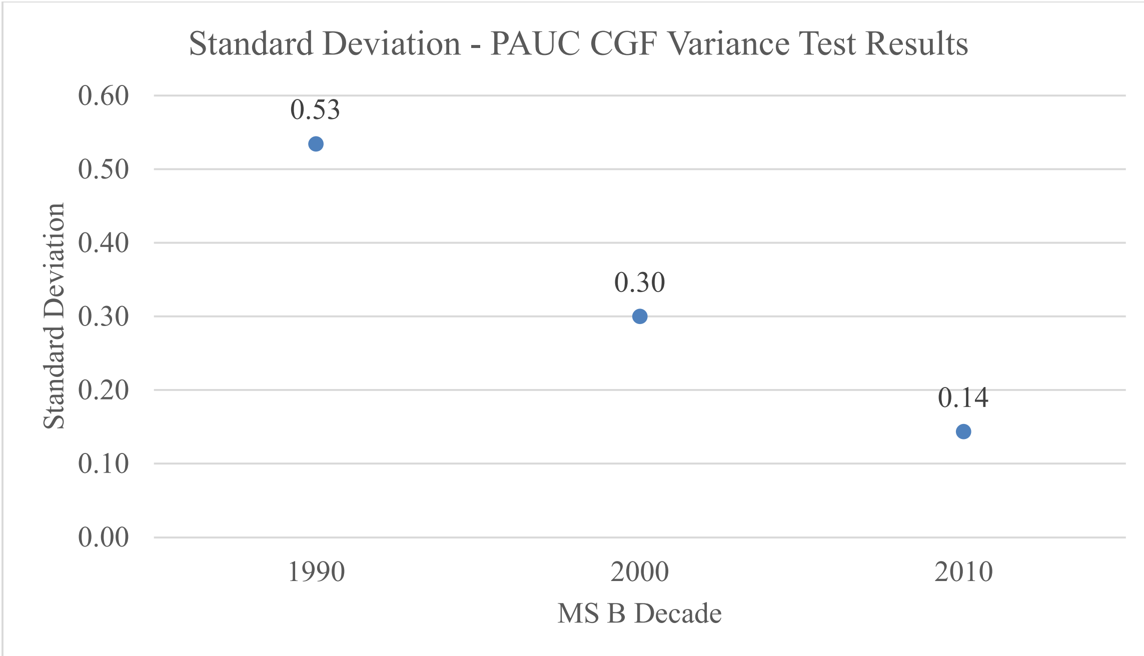


Figure 32. Standard Deviations - PAUC CGF of Ongoing Programs

The Brown-Forsythe variance test indicated that there was a statistically significant difference in the variances across the decades. Looking at Figure 32, we can see that the standard deviations for the data in each decade are decreasing across time.

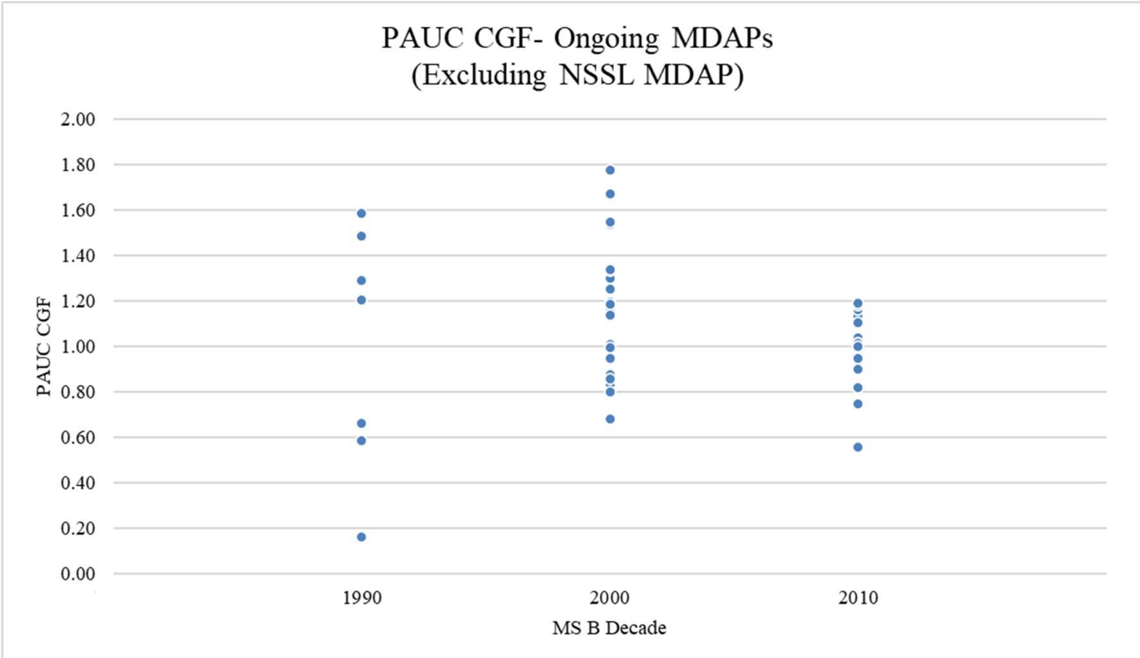


Figure 33. PAUC CGF of Ongoing Programs

Figure 33 provides our last look at how the PAUC CGFs of each decade compare with one another. Although the PAUC CGF of ongoing MDAPs (minus the NSSL program) are under 1.8 for the last three decades, we can see that the spread of the growth factors is declining each decade. This visual observation corresponds with the results of our inferential analysis that there are not any differences in the MRs through the decades. Although there are no statistically significant differences in the MRs of the data, looking at the spread of the growth factors in each decade we can see that the datapoints cluster closer together throughout the decades, which supports the results of the variance test conducted.

Analyzing our samples of ongoing programs, we identified a statistically significant difference in the overall CGFs of MDAPs between the 1990s and 2010s, with the 1990s having a higher MR than that identified for the 2010s. Schedule growth percentages and

PAUC CGFs did not have the same trend identified within their respective datasets. Differences in variances were identified in the overall CGFs and PAUC CGFs of ongoing programs, with both sets of data showing decreasing variances from the 1990s to the 2010s.

Analysis of MDAPs at MS B + 5 Years

After investigating cost and schedule estimation growth of MDAPs through the three previously identified lenses: all programs, completed programs, and ongoing programs, we now shift our attention to the last section of this study. For these last three analyses, we investigated MDAP's cost and schedule growth from MS B to the MS B +5 years point in their lifespans. This analysis allows us to examine a large majority of our programs against one another without the comparisons being distorted due to maturity bias.

Analysis of Schedule Growth Percentages for MDAPs at MS B +5 Years

Table 43. Schedule Growth Percentage Summary Statistics – MS B + 5 Years

Schedule Growth Percentage Summary Statistics - MS B + 5 Years						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	12	14%	9%	0.15	0.24	1.04
1980	21	10%	5%	0.14	0.17	1.42
1990	27	14%	7%	0.22	0.21	1.54
2000	35	25%	9%	0.52	0.21	2.07
2010	25	9%	4%	0.29	0.19	3.14

We began the maturity bias adjusted portion of our study by looking at the schedule growth of MDAPs at the MS B +5 years point. Looking at the summary statistics in Table 43 we quickly identified that the 2010s had the lowest mean and median schedule growth percentages, however this decade also has the highest CV. Although the CVs of our samples are increasing throughout the decades, the means, medians and IQRs have no discernable pattern of growth or retreat.

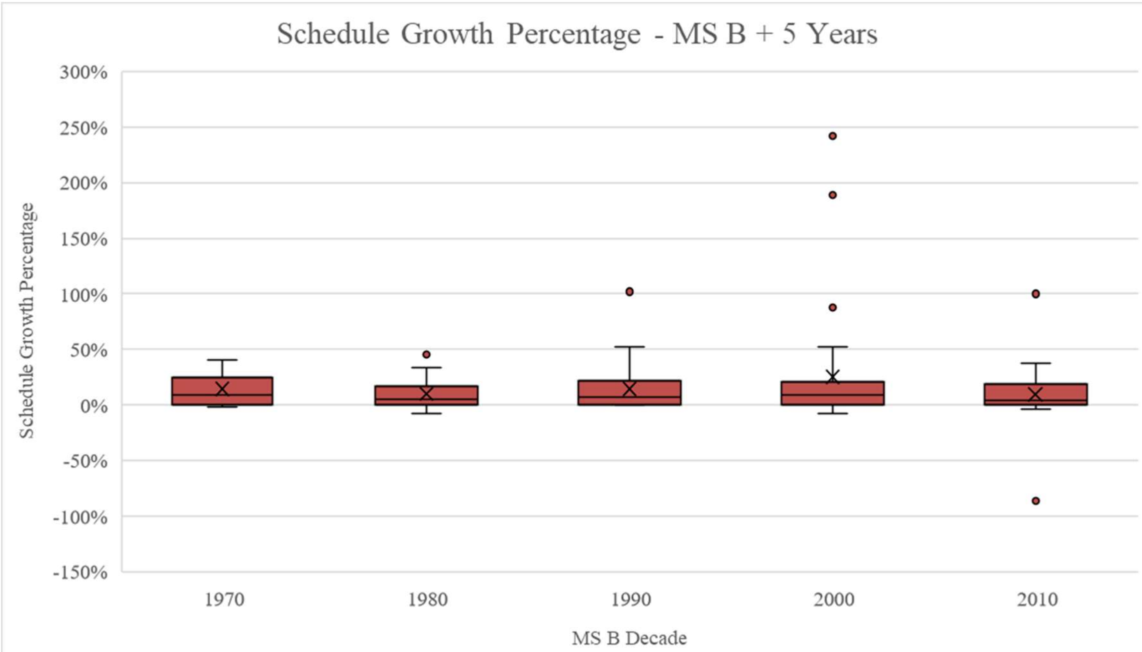


Figure 34. Box Plot – Schedule Growth Percentage

Reviewing the data using the Box Plot in Figure 34, it is noticeable that the IQRs, means, and medians seem to be relatively close which we identified utilizing the summary statistics in Table 43. One observation that does stand out is the number of outliers and their distance from the IQRs throughout the decades, especially in the 2000s. Although there are a some outliers identified, since this data has already been normalized at the MS B +5 years, we decided to not exclude any additional data points from the analysis.

Table 44. Kruskal-Wallis Test Results – Schedule Growth Percentage

Wilcoxon Test		
ChiSquare	DF	Prob>ChiSq
1.3005	4	0.8613

The results of the Kruskal-Wallis test, presented in Table 44, show a *p*-value of 0.8613, leading us to accept the null hypothesis and assume that there are no differences in the MRs of schedule growth percentages through the five decades at the MS B +5 years point in the MDAPs lifecycles.

Table 45. Variance Test Results – Schedule Growth Percentage

Schedule Growth Percentage Variance Tests - MS B + 5 Years				
Test	F Ratio	DFNum	DFDen	Prob > F
Brown-Forsythe	0.9037	4	115	0.4644

Looking at the results of Brown-Forsythe test presented in Table 45, we can identify that there are no statistically significant differences in the variances of the data through the decades.

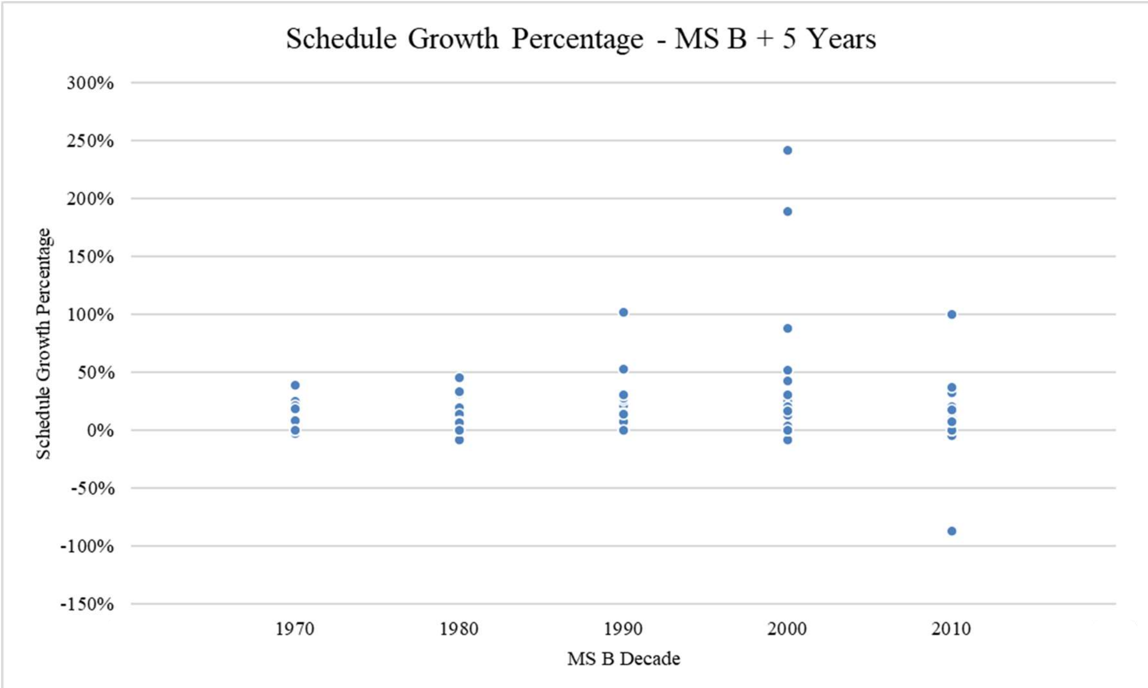


Figure 35. Schedule Growth Percentage - MS B +5 Years

Figure 35. provides us with our last visual inspection of schedule growth for our MS B +5 years data. A majority of the datapoints through all five decades fall between 0% and 50% growth. The clusters of datapoints through the decades coincide with the results of our inferential analysis that there are not any statistical differences in the MRs of the decades nor the variances of the data.

Analysis of Overall CGFs for MDAPs at MS B +5 Years

Table 46. Overall CGF Summary Statistics – MS B + 5 Years

Overall CGF Summary Statistics - MS B + 5 Years						
MS B Decade	Observations	Mean	Median	Std Deviation	IQR	CV
1970	28	2.09	1.31	2.70	1.17	1.30
1980	37	1.44	1.01	1.74	0.28	1.21
1990	33	1.37	1.10	1.24	0.63	0.90
2000	42	1.16	1.10	0.45	0.25	0.38
2010	29	1.13	1.02	0.37	0.23	0.33

Next, we turn our attention to the overall CGFs of MDAPs at the MS B +5 years point. Regarding the overall CGF summary statistics presented in Table 46, it seems as though the mean CGFs are reducing through the decades, however, the medians do not follow that same trend. Standard deviations and CVs also seem to be getting smaller through the decades.

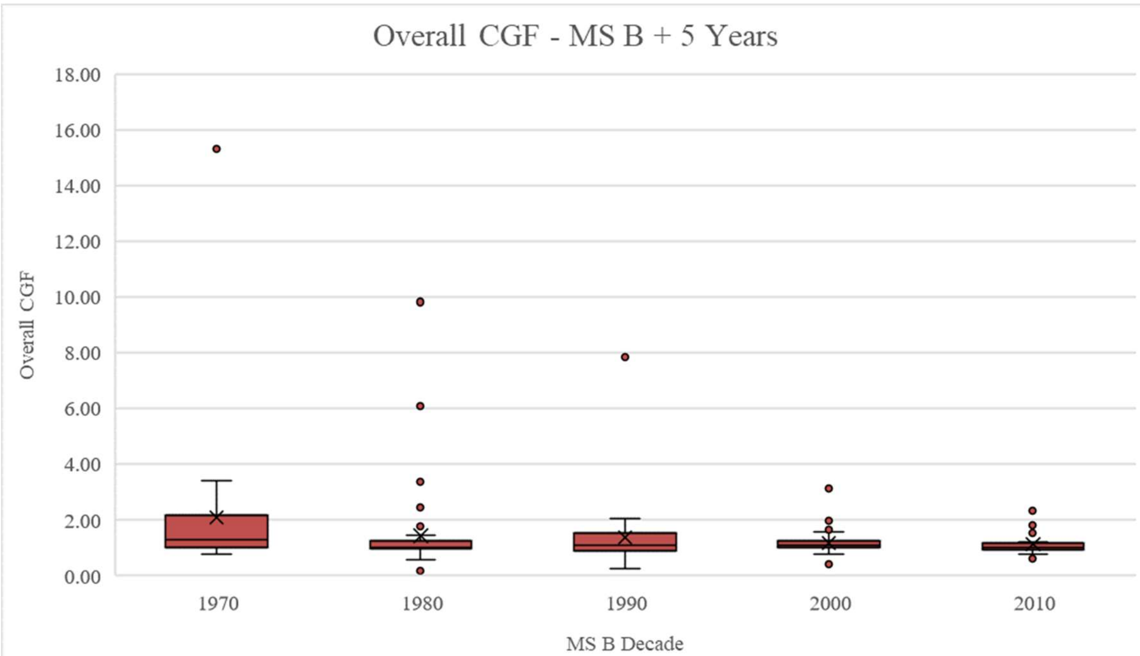


Figure 36. Box Plot – Overall CGF

Illustrating the overall CGFs using the Box Plot in Figure 36, we can quickly identify that each decade has outliers, however it does appear as through the distance of the outliers from the individual decades IQRs is decreasing over the course of the five decades analyzed.

Table 47. Kruskal-Wallis Test Results – Overall CGF

Kruskal-Wallis Test		
ChiSquare	DF	Prob>ChiSq
8.0749	4	0.0889

Table 47 displays the results of the Kruskal-Wallis test. This test resulted in a *p*-value of 0.0889 leading us to accept the null hypothesis and assume that there are no differences in the MRs of the overall CGFs.

Table 48. Variance Test Results – Overall CGF

Overall CGF Variance Tests - MS B + Years				
Test	F Ratio	DFNum	DFDen	Prob > F
Brown-Forsythe	1.9139	4	164	0.1105

Shifting our attention to the variance test results in Table 48, we find that the Brown-Forsythe test provided a *p*-value of 0.1105. This led us to accept our null hypothesis that there are no differences in the variances.

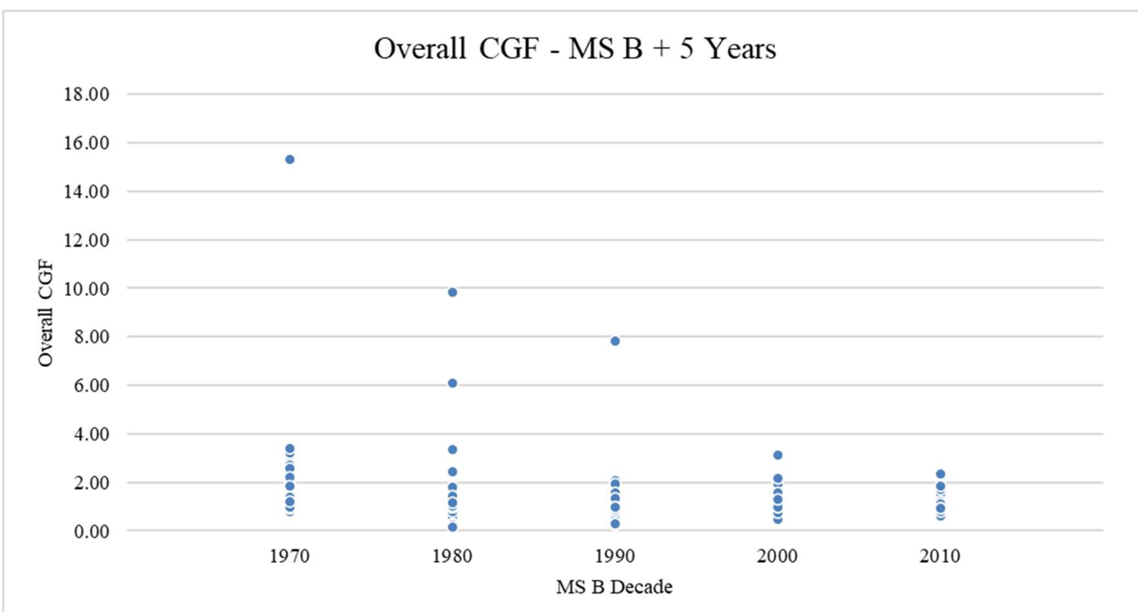


Figure 37. Overall CGF – MS B +5 Years

Figure 37 provided us one last opportunity to conduct a visual inspection of the data. Again, we can see that the outliers for each decade are lessening in their distances from the rest of the data within the decades which was highlighted by the decrease in CVs across the decades. Although this may not be statistically significant it is an interesting trend to note.

Analysis of PAUC CGFs for MDAPs at MS B +5 Years

Table 49. PAUC CGF Summary Statistics – MS B + 5 Years

PAUC CGF Summary Statistics - MS B + 5 Years						
MS B Decade	**Observations**	**Mean**	**Median**	**Std Deviation**	**IQR**	**CV**
1970	26	1.37	1.06	0.83	0.72	0.61
1980	31	1.66	1.10	1.93	0.47	1.16
1990	33	1.44	1.19	1.33	0.45	0.92
2000	38	1.08	1.07	0.30	0.14	0.27
2010	27	0.98	1.00	0.13	0.12	0.14

The last piece of this study examined the PAUC CGF of MDAPs at the MS B +5 years mark. From the summary statistics outlined in Table 49, we can identify that the mean CGFs have increased and decreased with no discernable pattern which is the same as the median data, standard deviations, and the CVs. The IQRs however, have steadily decreased in size over the last five decades.

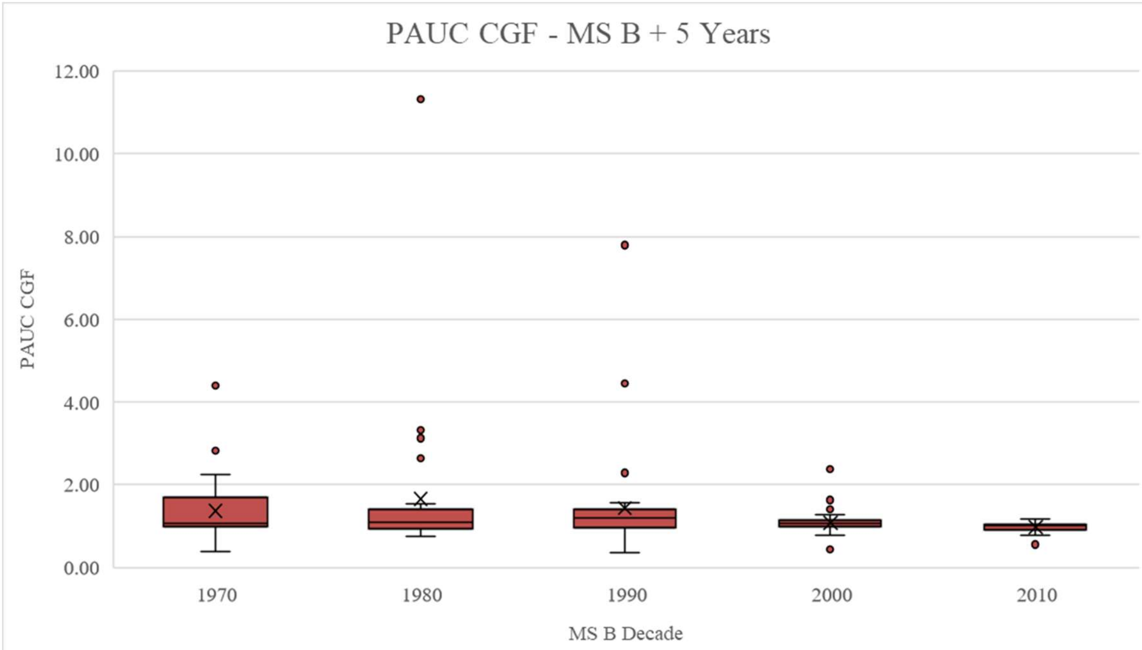


Figure 38. Box Plot – PAUC CGF

In Figure 38 we can see how the summary statistics of the decades appear against one another. There are outliers in each decade, but the 1980s features the outlier that is furthest away from the IQR. It is also visible that the IQRs are getting smaller through the decades, but the means and medians do not appear to shift too much.

Table 50. Kruskal-Wallis Test Results –PAUC CGF

Kruskal-Wallis Test		
ChiSquare	DF	Prob>ChiSq
8.0948	4	0.0882

The Kruskal-Wallis test produced a *p*-value of 0.0882 as shown in Table 50. Since this is higher than our alpha of 0.05, we assume that our null hypothesis is correct and there are no differences in the MRs of the PAUC CGFs through the last five decades.

Table 51. Variance Test Results – PAUC CGF

PAUC CGF Variance Test - MS B + 5 Years				
Test	**F Ratio**	**DFNum**	**DFDen**	**Prob > F**
Brown-Forsythe	2.0756	4	150	0.0868

The Brown-Forsythe variance test produced a *p*-value of 0.0868 which is higher than our alpha of 0.05. These results, shown in Table 51, indicated that there is no statistically significant difference between the variances across the five decades.

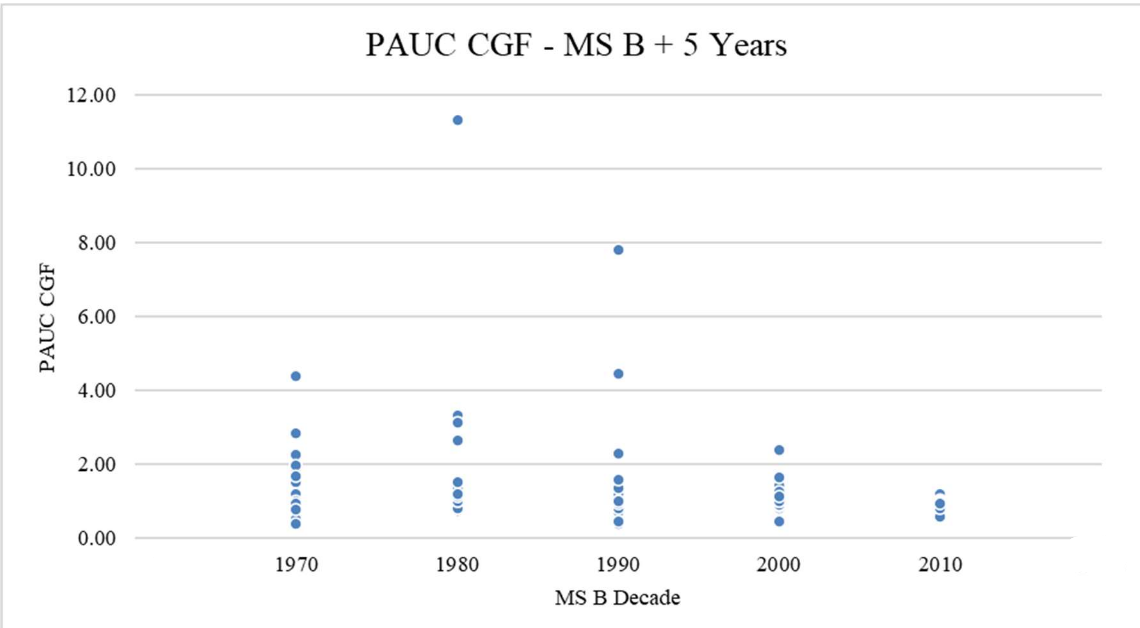


Figure 39. PAUC CGF – MS B + 5 Years

The PAUC CGFs plotted in Figure 39 show that many of the growth factors across all five decades fall below 2 at the MS B +5 years mark, which corresponds with the means and medians identified in the summary statistics

The analyses completed for MDAPs at the MS B +5 years mark did not result in any statistically significant difference in growth factors (or percentages) or variances regarding schedule, overall CGFs, or PAUC CGFs between the five decades when using a 95 percent confidence interval, or 0.05 alpha, between the five decades. This conclusion may highlight the impact that maturity bias play in analyzing cost growth. Since it was identified that there were differences in the PAUC CGFs of ongoing and completed programs and ongoing programs from the 1990s and 2010s, these results show that programs from the 1990s didn't realize their PAUC cost growth until after the MS B +5 years mark.

Review of Results

The previous sections of this chapter provided in depth analyses of our various samples, however the inferential tests completed are critical pieces in answering the research question. The two main tests results drawn from the inferential test were the conclusions of the Kruskal-Wallis (or Wilcoxon Rank Sum) and the Brown-Forsythe tests. For each of these tests, if the data was found to be statistically significant, additional information was included in the review. This additional information included the results of the Steel-Dwass multiple comparisons and the trend of how the standard deviations changed from decade to decade.

Table 52. Overall Results of Inferential Statistical Analysis

Results of Inferential Analysis of Cost and Schedule Estimation Growth for DoD MDAPs				
	Kruskal-Wallis/Wilcoxon	Steel-Dwass Results	B-F	Standard Deviation Trend**
All Programs				
Schedule	No		No	
Overall CGF	No		**Yes**	↓↓↓↓
PAUC CGF	**Yes**	**1990 ≠ 2010**	**Yes**	↑↓↓↓
		1970 ≠ 2010*		
Completed				
Schedule	No		No	
Overall CGF	No		**Yes**	↓↓↓
PAUC CGF	No		No	
Ongoing				
Schedule	No		No	
Overall CGF	**Yes**	**1990 ≠ 2010**	**Yes**	↓↓
PAUC CGF	No		**Yes**	↓↓
MS B +5 Years				
Schedule	No		No	
Overall CGF	No		No	
PAUC CGF	No		No	

*Differences between 2010 and 1970 had an alpha of .0505 which were not technically significant with our alpha of 0.05, but were identified as moderately significant.
**Arrow represent direction the standard deviation shifted from the previous decade within the analysis

Table 52 provides the overview of all the results of the inferential statistical analyses conducted for this study. This summary table allows us to quickly identify that there were two instances where cost growth estimates had differences between the decades. When analyzing all the ongoing and completed programs together the MRs of the PAUC CGFs of 2010 differs from the MRs of the PAUC CGFs from the 1990s, and moderately differ from those in the 1970s. Both the 1970s and 1990s possessed median CGFs that were higher than those seen in the 2010s. For ongoing MDAPs, there were differences in the MRs of the overall CGFs belonging to the 1990s. These differences occurred between the MRs of the overall CGFs in the 2000s and 2010s in comparison to the 1990s, with the 90s containing median CGFs that were higher than the next two decades.

Shifting our attention from the differences between the MRs of cost and schedule estimate growth across the decades, we reviewed the differences in variances across the decades. There are differences in the variances for overall CGF and PAUC CGFs for ongoing and completed programs. The overall CGF variance decreased through the five decades reviewed, but the PAUC CGF increased in the 1980s and then subsequently decreased each decade after that. When analyzing completed programs, the overall CGF of MDAPs had differences in the variance, decreasing across the four decades in the sample. The PAUC CGF of ongoing programs also displayed differences in variances across the three decades contained in the analysis. For the variance of this data, there is a decrease in the standard deviation for each decade.

Summary

Chapter IV outlined the findings of the analyses conducted. We started our investigation of cost and schedule growth by looking at the data of ongoing and completed MDAPs together and then divided the data into separate categories, completed and ongoing MDAPs for further investigation. The last portion of the analysis we investigated the schedule and cost growth estimates at the MS B +5 years points, which provided us with a view of estimate growth adjusted for maturity bias. As we move forward to Chapter V, we reexamine our results and address how these results answer our research question and compare to other completed studies.

V. Conclusions and Recommendations

Chapter Overview

This chapter reiterates the results from our analysis and addresses our original research question in a concise manner. After discussing the overarching results, we discuss how these results compare with previously completed studies, and then discuss recommendations for additional research topics regarding MDAP cost and schedule estimates.

Conclusions of Research

When analyzing all the MDAPs, ongoing and completed together, we found that there were no statistically significant differences in the MRs of schedule estimates or MDAPs overall CGFs across the five decades. We did find that there were differences between the PAUC CGFs between the 1990s and the 2010s, with the MRs of the PAUC CGFs being higher in the 1990s than in the 2010s. For both the overall CGFs and the PAUC CGFs we found that the variances decreased across the five decades analyzed. The overall CGFs of ongoing and completed programs have decreased since the 1970s onwards and the PAUC CGFs decreased from the 1980s onward.

As we moved our focus to completed MDAPs, we did not find any statistically significant differences in the MRs of growth seen in schedule, overall CGFs, or PAUC CGFs over the decades analyzed. It is important to note that for completed programs that we excluded the 2010s from all the analyses since there were not enough completed programs in the decade to be included. Although there were no differences in the cost and schedule growth of completed MDAPs from the 1970s to the 2000s, we identified that the variances for the overall CGFs of completed MDAPs decreased consistently each decade from the 1970s through the 2000s.

Investigating cost and schedule growth of ongoing programs yielded no statistically significant results involving schedule estimates of PAUC CGFs, however, there were significant results found regarding the overall CGFs. For overall CGFs for MDAPs we identified that the MRs of programs from the 1990s were higher than those from the 2010s. When reviewing the summary statistics, you can see that the means and medians for these programs are also decreasing through the decades further supporting this finding. Along with the differences in the MRs for overall CGFs, we identified that the variances also saw consistent decreases across the three decades analyzed. The PAUC CGFs also had differences within the variances across the three decades with the standard deviations from the 1990s to the 2010s consistently decreasing.

For the last part for our study, we looked at MDAPs with estimates at the MS B +5 years mark. Using this snapshot in the MDAPs lifetime we were able to account for maturity bias in cost and schedule growth for MDAPs. For this part of the analysis there were no statistically significant differences in the schedule growth percentages, overall CGFs, or PAUC CGFs across the five decades. There were also no differences in the variances of cost and schedule estimate growth at the MS B +5 years mark across the five decades investigated in this study.

Throughout the various inferential analyses contained within this study there was no consistent evidence that the DoD has improved in their ability to accurately estimate the costs and schedules of MDAPs across the last five decades. There were a couple instances in which the 1990s were statistically different than the 2010s regarding cost estimates. In both instances the 1990s had higher MRs, than the 2010s, however this hints more to the notion that the 1990s were the anomaly and saw more cost growth than the other decades, rather than the DoD has gotten better since the 1970s.

Research Question Revisited

How have DoD cost and schedule estimates changed over the last five decades?

We were able to identify two cases in which cost growth was statistically different between decades: the PAUC CGF of ongoing and completed MDAPs and the overall CGF of ongoing MDAPs. In both instances the 1990s had higher MRs than the 2010s; however, this does not necessarily highlight any upward or downward trends regarding schedule or cost growth in MDAP estimates, but it does suggest that the 1990s may have been an abnormal decade for DoD cost estimators.

Regarding the variance of cost and schedule growth through the decades, we identified that there were a few instances where this aspect of cost estimating did get better over the decades. Specifically, when looking at the variances of ongoing and completed programs combined, we can see that the variances of the overall CGFs and PAUC CGFs have both decreased consistently since the 1970s and 1980s, respectively. Although these results are not continuously highlighted through the rest of the analyses, there are three more cases in which the variances are identified as statistically different through the decades tested.

As illustrated in Figure 40, the standard deviations of all four analyses of overall CGFs show downward trends. As already identified the overall CGFs of ongoing and completed programs show a statistically significant difference in the variances.

According to our inferential analyses, the overall CGFs of completed programs and the overall CGFs of ongoing programs also show statistically significant difference in their variances. These visibly perceptible downward trends in the variances of overall CGFs may suggest improvements in cost estimating.

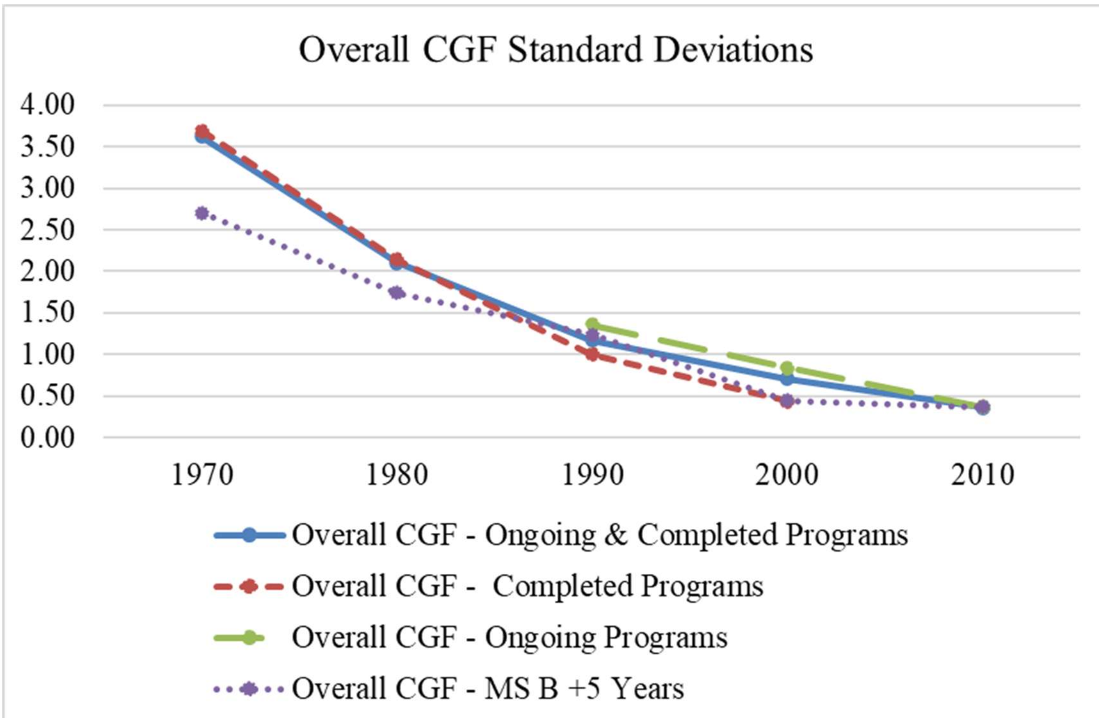


Figure 40. Standard Deviations of Overall CGFs

Figure 41 illustrates the trends seen in the PAUC CGF's standard deviations identified for the four analyses. For our PAUC CGFs we identified that there were statistically significant differences between the decades when evaluating ongoing and completed programs together, as noted above. We can see that there was an initial increase in the standard deviation from the 1970s to the 1980s, but a consistent downwards trend since the 1980s. There was also a statistically significant difference in the standard deviations of the PAUC CGFs of ongoing programs. The standard deviations realized for ongoing program are much lower overall than those seen by those seen in the three other categories and thus do not visually stand out in Figure 41. Although the PAUC CGFs do not see trends as stable as those seen by the overall CGFs, there does seem to be downward movement in the standards deviations since the 1990s for all four analyses, and since the 1980s for two out of the four.

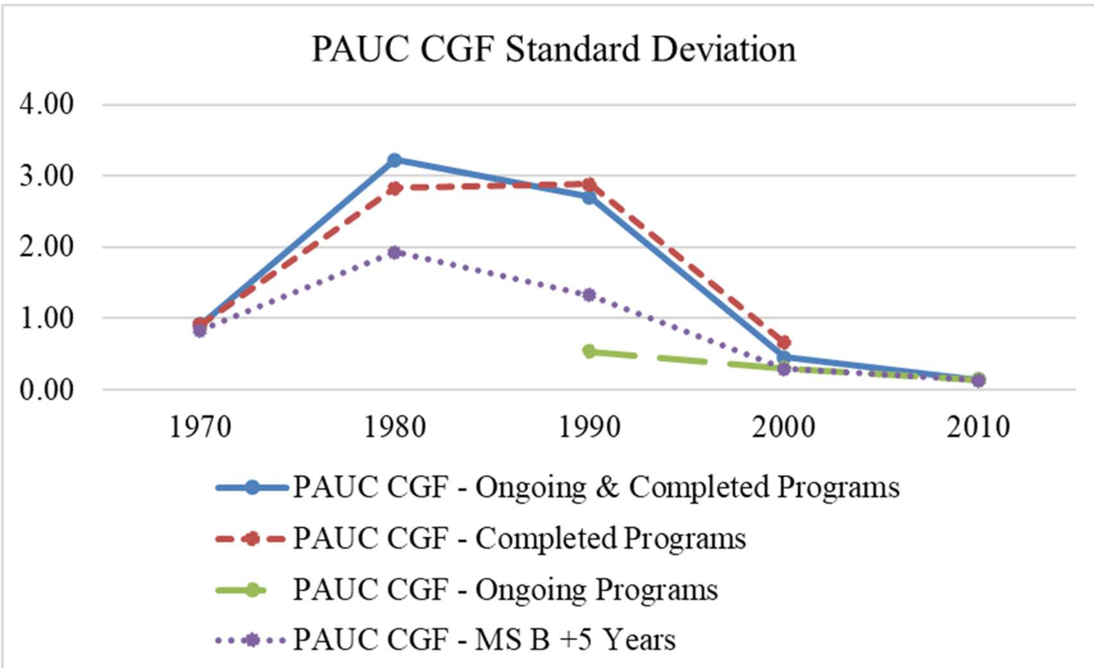


Figure 41. Standard Deviations of PAUC CGFs

Final Thoughts

Our analysis of schedule estimation growth did not culminate in any statistically significant differences across the five decades reviewed in this analysis regardless of how we categorized the data (i.e., completed and ongoing, completed, ongoing, or MS B +5 years). Through descriptive statistics we identified that MDAPs experienced a mean of 16% schedule growth at the MS B +5 years mark, and currently completed and ongoing MDAPs realized a mean schedule growth of 25%. Although there were no identifiable trends pointing to the DoD improving their schedule estimation accuracy, they have managed to stay relatively consistent in their schedule estimates across the last five decades. If the DoD is interested in increasing the accuracy of their schedule estimates, utilizing this data as a starting point in their analyses may prove beneficial.

The schedule growth results identified from the descriptive statistics differ slightly from those identified by Dwyer et al. (2020). In that analysis, they identified that the mean cycle time growth for MDAPs was 31.2%; however, their sample contained 221 MDAPs, versus this study that only had a sample size of 120. The way that cycle time growth was calculated in Dwyer's study also differs from the way that our schedule growth percentage was calculated as the default for that study was to use the developmental estimates and then to use production estimates when the developmental estimates were unavailable. In contrast we used the IOC estimates taken from the original MS B estimate and that from the latest SAR.

Although there were slight differences in the cycle-time growth and schedule estimation growth calculations, neither study found statistically significant results across reform eras as defined by Dwyer et al. (2020) or the decades as looked at in this study. Comparatively when looking at reform eras aggregated into centralized and decentralized categories, Dwyer et al. (2020) did find statistically significant difference in cycle time growth.

Investigating the overall CGF of DoD MDAPs did yield a noteworthy trend regarding the variance of CGFs across the decades. While analyzing the variance of completed and ongoing, completed, and ongoing we consistently identified that the overall CGF decreased across the decades at a significant level (*p*-values of 0.0006, 0.027, and 0.002, respectively).

These results provide a stark contrast to variances found by Younossi et al. (2007) while analyzing developmental CGFs from the 1970s through the 1990s. Within the sample studied, they identified that the variances were increasing through the decades. However, there are a few caveats to these differences. Firstly, Younossi only looked at the developmental CGF versus the CGF for the entire MDAP. Furthermore, the sample used for the 2007 study was also slightly different in size than the sample used in this thesis. Lastly, this growth in variances identified occurred prior to outliers being excluded from the data. Excluding the outliers led to different result regarding the variance trend. Rather than seeing the upward trend as seen with the outliers, excluding them led to a slight growth in the standard deviation from the 1970s to the 1980s then a drop from the 1980s to the 1990s.

In our investigation of PAUC CGFs we identified that there were statistically significant differences between MRs of the PAUC CGFs from the 1990s versus the 2010s and moderately significant differences between those from 1970s to the 2010s. There were also two of the four analysis that identified a trend in the variances of CGFs across time. There has been a downward trend in the variances of PAUC CGFs since the 1980s when reviewing ongoing and completed programs simultaneously. This downward trend was again identified when looking at ongoing programs, which found that the variances in PAUC CGFs have been decreasing since the 1990s. McNichol et al. (2016) investigated PAUC CGFs and found a few statistically significant differences in PAUC CGFs across time.

At the beginning of the 2018 study, we identified that there were statistically significant differences in PAUC Cost growth beginning during the Packard Reform, but found that since then, there have not been any other statistically significant differences in PAUC Cost growth based on reform cycles. These reductions from the McNamara era to the Packard Reforms did however continue through the end of 2009, which was also the end of the period investigated. These results support our findings. Although the reform cycles do not align perfectly with our decade groupings, they are similar. Like McNichol et al. (2016), we did not identify any statistically significant differences from the 1970s through the 2000s. It was not until the 2010s, which were not included in the McNicol Study, where differences were identified.

In the investigation of PAUC cost growth during funding climates, categorized as busts and booms, McNichol et al. (2018) identified that there were no statistically significant difference in the variances of the various funding climates from the 1970s through 1989, however they did identify differences in the variances of cost growth percentages during the Defense Acquisition Board (DAB) era, which took place from 1990-1993 and again from 2001-2009 based on whether the funding climate was a boom or bust. In the analyses the boom cycle consisted of MDAPs that reached MS B from 2003-2009 and all the other years were classified as bust funding climates. Boom Cycles saw PAUC growth at 2% and the bust cycles saw PAUC cost growth percentage of 40%. The *p*-value produced by the two-tailed t-test in was <0.001 highlighting the statistically significant differences. Although they did not track any lasting trends in the variances, using the results from our analysis we can support the finding of an increase in variances for program from the 1990s, while also highlighting the decreasing of the PAUC CGFs used in our study over the decades.

Recommendations for Future Research

There are many ways to inspect cost and schedule growth in MDAPs. This paper examines cost and schedule growth on a macro-level across the last five decades, but there are still many micro-level analyses that can be completed. Looking at schedule and cost growth by commodity could provide additional insight into trends in cost and schedule growth for MDAPs.

Although there have been various time series studies completed regarding funding climates and the introduction of certain best business practices and legislation, there are still some niches that researchers can explore regarding how these factors were impacted by such events.

Analyzing cost and schedule growth is an enduring process and even macro-level studies should be completed on a regular basis to ensure that our perspectives on this topic do not become outdated. Re-evaluating and recompleting this analysis in the next 10 to 20 years is highly recommended as it will provide a stronger assessment of the cost and schedule growth of MDAPs that are currently ongoing.

Summary

This chapter reiterated the results from our analysis and addressed our original research question "How have DoD cost and schedule estimates changed over the last five decades?". The results of our statistical analysis identified that there are trends showing that the variances of the overall CGFs for MDAPs are decreasing through the five decades investigated. There were statistically significant differences between the overall CGFs for ongoing MDAPs and PAUC CGFs for ongoing and completed MDAPs, however the difference was identified to be sporadic rather than consistent across time. Both showed differences between the 1990s and 2010s. For schedule growth percentages, there were no statistically significant trends showing that schedule estimations were changing across time. After discussing the overarching results, we reviewed how these results compared with previous studies and then assessed recommendations for additional research topics regarding MDAP cost and schedule estimation.

Appendix A – Kruskal Wallis JMP Output

Kruskal-Wallis - Ongoing & Completed MDAPs - Schedule Growth Percentage					
Level	**Count**	**Score Sum**	**Expected Score**	**Score Mean**	**(Mean-Mean0)/Std0**
1970	12	786.5	726	65.5417	0.525
1980	21	1223.5	1270.5	58.2619	-0.322
1990	27	1899.5	1633.5	70.3519	1.671
2000	35	2153	2117.5	61.5143	0.202
2010	25	1197.5	1512.5	47.9	-2.035
ChiSquare	**DF**	**Prob>ChiSq**			
5.8283	4	**0.2123**			

Kruskal-Wallis - Ongoing & Completed MDAPs - Overall CGF					
Level	**Count**	**Score Sum**	**Expected Score**	**Score Mean**	**(Mean-Mean0)/Std0**
1970	29	3074	2653.5	106	1.615
1980	45	3392	4117.5	75.378	-2.365
1990	37	3671.5	3385.5	99.23	0.998
2000	42	4082	3843	97.19	0.797
2010	29	2433.5	2653.5	83.914	-0.844
ChiSquare	**DF**	**Prob>ChiSq**			
8.3	4	**0.0812**			

Kruskal-Wallis - Ongoing & Completed MDAPs - PAUC CGF					
Level	**Count**	**Score Sum**	**Expected Score**	**Score Mean**	**(Mean-Mean0)/Std0**
1970	25	2389.5	2062.5	95.58	1.494
1980	35	3025	2887.5	86.4286	0.55
1990	37	3461	3052.5	93.5405	1.605
2000	38	2876.5	3135	75.6974	-1.006
2010	29	1778	2392.5	61.3103	-2.647
ChiSquare	**DF**	**Prob>ChiSq**			
10.6936	4	**0.0302**			

Kruskal-Wallis - Completed MDAPs - Schedule Growth Percentage					
Level	**Count**	**Score Sum**	**Expected Score**	**Score Mean**	**(Mean-Mean0)/Std0**
1970	12	458.5	408	38.2083	0.82
1980	20	654	680	32.7	-0.35
1990	23	834	782	36.2609	0.682
2000	12	331.5	408	27.625	-1.247
ChiSquare	DF	Prob>ChiSq			
2.2569	**3**	**0.5208**			

Kruskal-Wallis - Completed MDAPs - Overall CGF					
Level	**Count**	**Score Sum**	**Expected Score**	**Score Mean**	**(Mean-Mean0)/Std0**
1970	28	1931.5	1638	68.9821	1.89
1980	41	2036	2398.5	49.6585	-2.091
1990	29	1736.5	1696.5	59.8793	0.252
2000	18	1082	1053	60.1111	0.217
ChiSquare	DF	Prob>ChiSq			
5.6448	**3**	**0.1302**			

Kruskal-Wallis - Completed MDAPs - PAUC CGF					
Level	**Count**	**Score Sum**	**Expected Score**	**Score Mean**	**(Mean-Mean0)/Std0**
1970	25	1336	1250	53.44	0.689
1980	31	1434	1550	46.2581	-0.872
1990	29	1606	1450	55.3793	1.196
2000	14	574	700	41	-1.261
ChiSquare	DF	Prob>ChiSq			
3.2777	**3**	**0.3508**			

Kruskal-Wallis - Ongoing MDAPs - Schedule Growth Percentage					
Level	**Count**	**Score Sum**	**Expected Score**	**Score Mean**	**(Mean-Mean0)/Std0**
2000	22	577	506	26.2273	1.602
2010	23	458	529	19.913	-1.602
ChiSquare	**DF**	**Prob>ChiSq**			
2.6025	1	0.1067			

Kruskal-Wallis - Ongoing MDAPs - Overall CGF					
Level	**Count**	**Score Sum**	**Expected Score**	**Score Mean**	**(Mean-Mean0)/Std0**
1990	8	364	240	45.5	2.735
2000	24	756.5	720	31.5208	0.556
2010	27	649.5	810	24.0556	-2.435
ChiSquare	**DF**	**Prob>ChiSq**			
9.9423	2	0.0069			

Kruskal-Wallis - Ongoing MDAPs - PAUC CGF					
Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
1990	7	210	206.5	30	0.072
2000	24	805.5	708	33.5625	1.532
2010	27	695.5	796.5	25.7593	-1.567
ChiSquare	DF	Prob>ChiSq			
2.722	2	0.2564			

Kruskal-Wallis - MS B +5 Years - Schedule Growth Percentage					
Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
1970	12	800.5	726	66.7083	0.655
1980	21	1158.5	1270.5	55.1667	-0.779
1990	27	1639	1633.5	60.7037	0.032
2000	35	2218	2117.5	63.3714	0.584
2010	25	1444	1512.5	57.76	-0.445
ChiSquare	DF	Prob>ChiSq			
1.3005	4	**0.8613**			

Kruskal-Wallis - MS B +5 Years - Overall CGF					
Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
1970	28	2983	2380	106.536	2.551
1980	37	2826	3145	76.378	-1.212
1990	33	2804.5	2805	84.985	0
2000	42	3606	3570	85.857	0.129
2010	29	2145.5	2465	73.983	-1.332
ChiSquare	DF	Prob>ChiSq			
8.0749	4	**0.0889**			

Kruskal-Wallis - MS B +5 Years - PAUC CGF					
Level	Count	Score Sum	Expected Score	Score Mean	(Mean-Mean0)/Std0
1970	26	2174	2028	83.6154	0.698
1980	31	2605.5	2418	84.0484	0.837
1990	33	2884.5	2574	87.4091	1.357
2000	38	2869	2964	75.5	-0.393
2010	27	1557	2106	57.6667	-2.591
ChiSquare	DF	Prob>ChiSq			
8.0948	4	**0.0882**			

Appendix B – Steel-Dwass JMP Output

Steel - Dwass - Ongoing & Completed MDAPs - Schedule Growth Percentage								
Level	**- Level**	**Score Mean Difference**	**Std Err Dif**	**Z**	**p-Value**	**Hodges-Lehmann**	**Lower CL**	**Upper CL**
1990	1980	4.8677	4.0661	1.1972	0.7531	0.08	-0.13	0.3
2000	1980	1.3333	4.4916	0.2969	0.9983	0.01	-0.16	0.19
1990	1970	1.2037	3.9524	0.3046	0.9981	0.03	-0.18	0.34
1980	1970	-1.3095	3.4810	-0.3762	0.9958	-0.02	-0.25	0.19
2000	1970	-1.3429	4.5806	-0.2932	0.9984	-0.015	-0.21	0.19
2010	1980	-3.4171	3.9653	-0.8618	0.9108	-0.06	-0.27	0.08
2000	1990	-4.3302	4.6177	-0.9377	0.8822	-0.06	-0.25	0.12
2010	1970	-5.8583	3.7971	-1.5428	0.5344	-0.11	-0.31	0.08
2010	2000	-6.5829	4.5692	-1.4407	0.6012	-0.08	-0.25	0.05
2010	1990	-10.0533	4.2039	-2.3915	0.1176	-0.14	-0.35	0.02

Steel - Dwass - Ongoing & Completed MDAPs - Overall CGF								
Level	**- Level**	**Score Mean Difference**	**Std Err Dif**	**Z**	**p-Value**	**Hodges-Lehmann**	**Lower CL**	**Upper CL**
2000	1980	11.1857	5.4185	2.0644	0.2357	0.2	-0.09	0.62
1990	1980	9.8991	5.2848	1.8731	0.3319	0.32	-0.13	0.96
2010	1980	5.2169	5.1204	1.0188	0.8468	0.09	-0.16	0.56
2000	1990	-2.1351	5.1736	-0.4127	0.9939	-0.1	-0.76	0.32
1990	1970	-3.3830	4.7607	-0.7106	0.9542	-0.23	-1.19	0.54
2000	1970	-5.0714	4.9824	-1.0179	0.8473	-0.22	-1.11	0.27
2010	1990	-5.9664	4.7601	-1.2534	0.7198	-0.16	-1.04	0.19
2010	1970	-6.1379	4.4340	-1.3843	0.6378	-0.29	-1.7	0.18
2010	2000	-7.2865	4.9816	-1.4627	0.5869	-0.09	-0.3	0.1
1980	1970	-10.6322	5.1209	-2.0762	0.2304	-0.55	-1.46	0.12

Steel - Dwass - Ongoing & Completed MDAPs - PAUC CGF								
Level	**- Level**	**Score Mean Difference**	**Std Err Dif**	**Z**	**p-Value**	**Hodges-Lehmann**	**Lower CL**	**Upper CL**
1990	1980	1.8069	4.9330	0.3663	0.9962	0.06	-0.45	0.48
1990	1970	-0.2681	4.6700	-0.0574	1.0000	-0.02	-0.6	0.58
1980	1970	-1.6800	4.5728	-0.3674	0.9961	-0.08	-0.59	0.58
2000	1980	-4.3635	4.9696	-0.8781	0.9051	-0.09	-0.59	0.16
2010	1980	-7.2512	4.6734	-1.5516	0.5287	-0.14	-0.7	0.07
2010	2000	-7.5390	4.8020	-1.5700	0.5168	-0.09	-0.25	0.06
2000	1990	-8.6682	5.0327	-1.7224	0.4200	-0.17	-0.55	0.1
2000	1970	-9.2511	4.7195	-1.9602	0.2857	-0.27	-0.73	0.1
2010	1970	-11.6938	4.2924	-2.7243	0.0505	-0.42	-0.88	0.01
2010	1990	-13.3169	4.7587	-2.7984	0.0411	-0.26	-0.59	0

Steel - Dwass - Completed MDAPs - Schedule Percentage Growth								
Level	**- Level**	**Score Mean Difference**	**Std Err Dif**	**Z**	**p-Value**	**Hodges-Lehmann**	**Lower CL**	**Upper CL**
1990	1980	2.3370	3.8296	0.6102	0.9289	0.03	-0.14	0.21
1990	1970	-1.6486	3.6447	-0.4523	0.9692	-0.025	-0.2	0.19
2000	1980	-2.0000	3.3980	-0.5886	0.9356	-0.04	-0.27	0.17
1980	1970	-2.0667	3.4068	-0.6066	0.9300	-0.04	-0.25	0.15
2000	1970	-3.4167	2.8742	-1.1888	0.6340	-0.085	-0.31	0.15
2000	1990	-5.0091	3.6421	-1.3753	0.5149	-0.085	-0.29	0.11

Steel - Dwass - Completed MDAPs - Overall CGF								
Level	**- Level**	**Score Mean Difference**	**Std Err Dif**	**Z**	**p-Value**	**Hodges-Lehmann**	**Lower CL**	**Upper CL**
2000	1980	6.7954	4.8556	1.3995	0.4996	0.15	-0.42	0.6
1990	1980	6.2994	4.9375	1.2758	0.5784	0.22	-0.25	0.76
2000	1990	-0.3602	4.1131	-0.0876	0.9998	-0.01	-0.69	0.34
2000	1970	-4.7004	4.0540	-1.1595	0.6525	-0.305	-1.91	0.27
1990	1970	-5.0191	4.3974	-1.1414	0.6638	-0.31	-1.23	0.32
1980	1970	-10.1577	4.9183	-2.0653	0.1646	-0.54	-1.34	0.1

Steel - Dwass - Completed MDAPs - PAUC CGF								
Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
1990	1980	5.1724	4.5089	1.1472	0.6602	0.15	-0.23	0.59
1990	1970	1.1917	4.2922	0.2777	0.9925	0.04	-0.54	0.69
2000	1980	-0.9850	4.2279	-0.2330	0.9955	-0.03	-0.72	0.24
1980	1970	-3.4323	4.3837	-0.7830	0.8622	-0.14	-0.63	0.44
2000	1970	-6.0171	3.8054	-1.5812	0.3894	-0.31	-0.95	0.17
2000	1990	-6.4606	4.0847	-1.5817	0.3891	-0.23	-1.2	0.1

Steel - Dwass - Ongoing MDAPs - Schedule Growth Percentage								
Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
2010	2000	-6.2698	3.9140	-1.6019	0.1092	-0.09	-0.23	0.02

Steel - Dwass - Ongoing MDAPs - Overall CGF								
Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
2010	2000	-6.6898	4.1690	-1.6046	0.2436	-0.13	-0.48	0.07
2000	1990	-8.0833	3.8294	-2.1109	0.0876	-1.335	-2.38	0.25
2010	1990	-12.0718	4.1233	-2.9277	0.0096	-1.455	-2.59	-0.22

Steel - Dwass - Ongoing MDAPs - PAUC CGF								
Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
2000	1990	2.1220	3.9045	0.5435	0.8498	0.125	-0.4	0.69
2010	1990	-2.6984	4.2208	-0.6393	0.7984	-0.13	-0.47	0.51
2010	2000	-6.6898	4.1683	-1.6049	0.2435	-0.13	-0.31	0.05

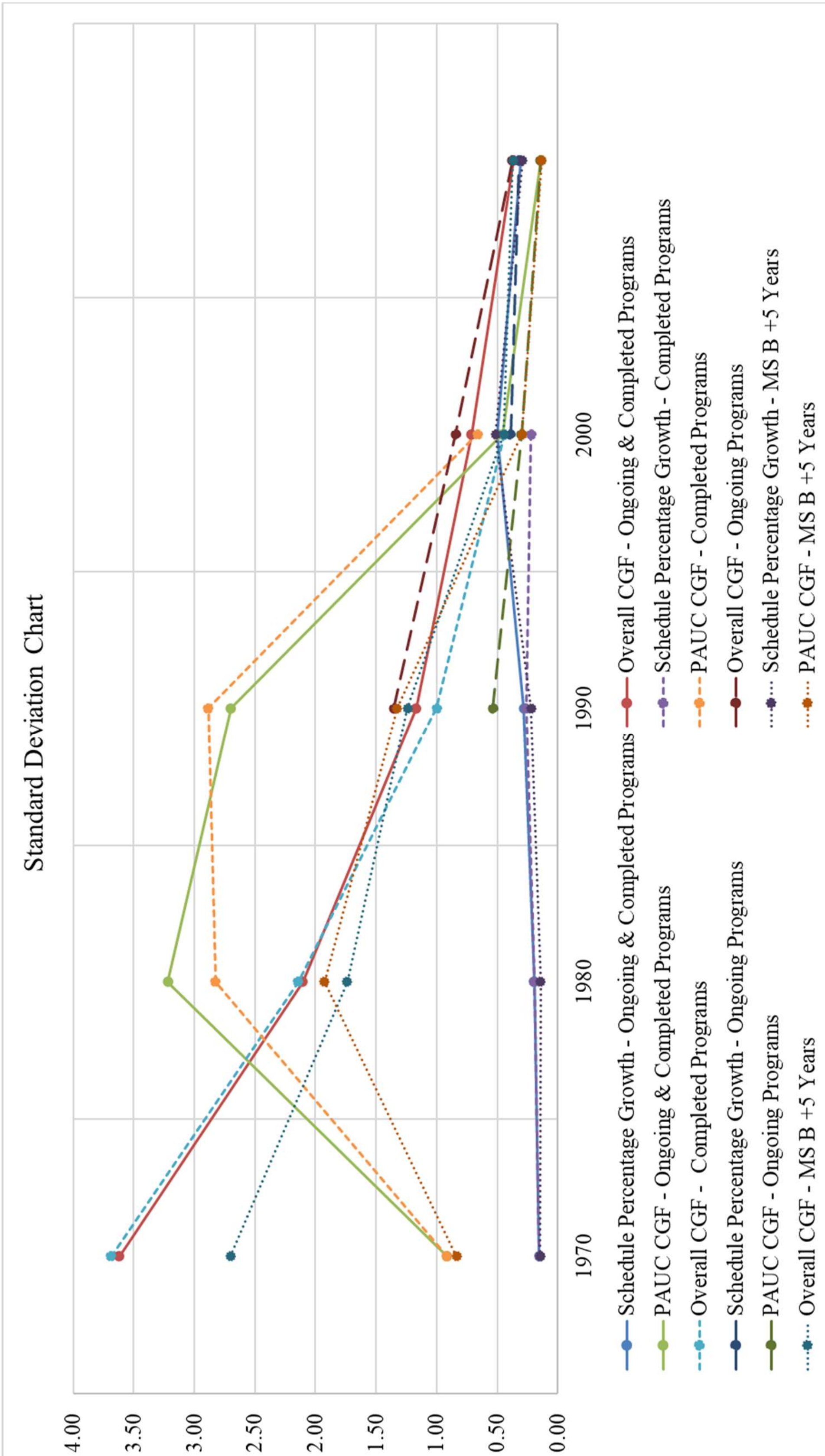
Steel - Dwass - MS B +5 Years- Scheduel Growth Percentage								
Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
2000	1980	3.6952	4.4122	0.8375	0.9190	0.01	-0.06	0.15
1990	1980	2.2857	3.9807	0.5742	0.9788	0	-0.07	0.14
2000	1990	1.2466	4.5416	0.2745	0.9988	0	-0.08	0.13
2010	1980	0.8762	3.9366	0.2226	0.9995	0	-0.11	0.11
2000	1970	-0.8393	4.5373	-0.1850	0.9997	0	-0.18	0.16
2010	1990	-0.9244	4.1722	-0.2216	0.9995	0	-0.13	0.07
1990	1970	-1.7454	3.9107	-0.4463	0.9918	-0.01	-0.19	0.13
2010	2000	-2.7086	4.5367	-0.5970	0.9756	-0.02	-0.15	0.07
2010	1970	-3.2067	3.7899	-0.8461	0.9161	-0.03	-0.22	0.1
1980	1970	-3.2083	3.4491	-0.9302	0.8853	-0.03	-0.2	0.1

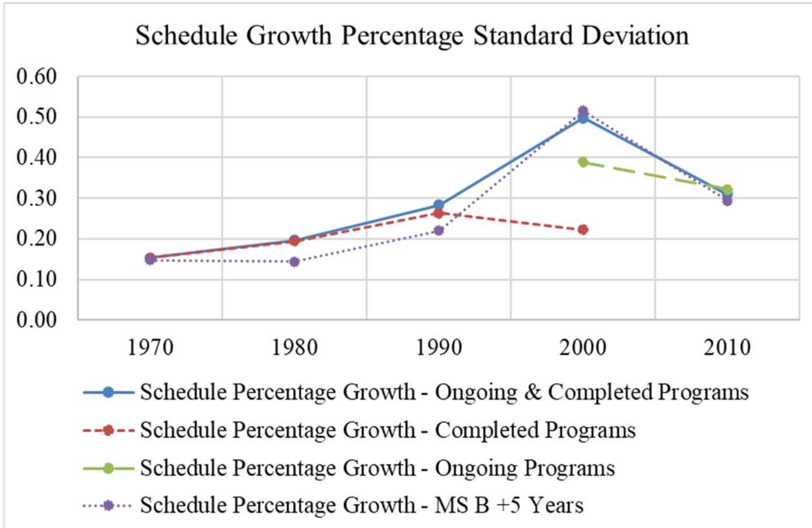
Steel - Dwass - MS B +5 Years- Overall CGF								
Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
2000	1980	5.1853	5.1646	1.0040	0.8537	0.04	-0.1	0.23
1990	1980	3.6405	4.8673	0.7479	0.9451	0.06	-0.16	0.48
2010	1980	-0.5843	4.7542	-0.1229	0.9999	0	-0.18	0.19
2000	1990	-0.7035	5.0649	-0.1389	0.9999	-0.01	-0.33	0.17
2010	1990	-3.2717	4.5894	-0.7129	0.9536	-0.05	-0.37	0.15
2010	2000	-6.0041	4.9771	-1.2063	0.7477	-0.06	-0.2	0.08
1990	1970	-8.4832	4.5574	-1.8614	0.3384	-0.2	-0.84	0.1
2000	1970	-9.2857	4.9573	-1.8731	0.3319	-0.23	-0.8	0.05
1980	1970	-10.1641	4.7272	-2.1501	0.1990	-0.365	-0.9	0.03
2010	1970	-10.8454	4.3931	-2.4688	0.0978	-0.24	-0.88	0.02

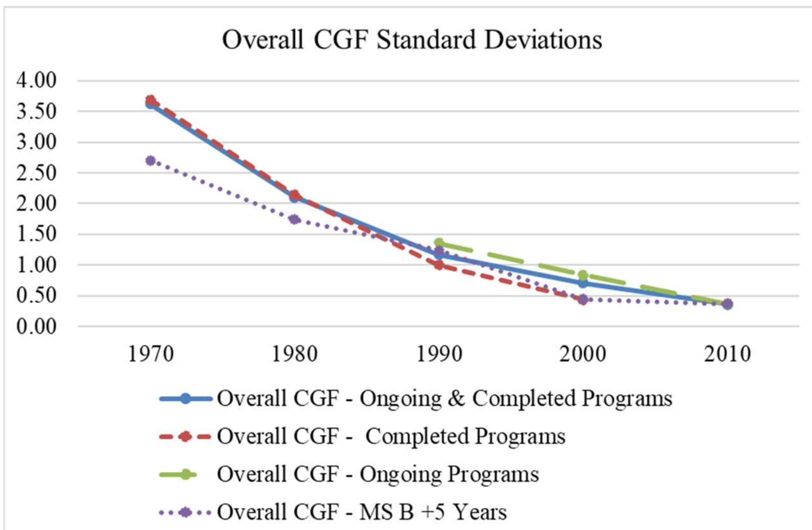
Level	- Level	Score Mean Difference	Std Err Dif	Z	p-Value	Hodges-Lehmann	Lower CL	Upper CL
1990	1980	0.7820	4.6518	0.1681	0.8665	0.01	-0.18	0.19
1980	1970	0.3536	4.4064	0.0803	0.9360	0	-0.2	0.25
1990	1970	0.1719	4.5005	0.0382	0.9695	0	-0.29	0.22
2000	1970	-3.2389	4.7332	-0.6843	0.4938	-0.05	-0.33	0.07
2000	1980	-4.1295	4.8473	-0.8519	0.3943	-0.05	-0.22	0.07
2000	1990	-7.0490	4.9068	-1.4366	0.1508	-0.1	-0.23	0.04
2010	1970	-7.8141	4.2371	-1.8442	0.0652	-0.14	-0.44	0
2010	1980	-8.5579	4.4378	-1.9284	0.0538	-0.13	-0.31	0
2010	2000	-9.5663	4.7527	-2.0128	0.0441	-0.07	-0.13	0
2010	1990	-11.3805	4.5277	-2.5135	0.0120	-0.17	-0.31	-0.04

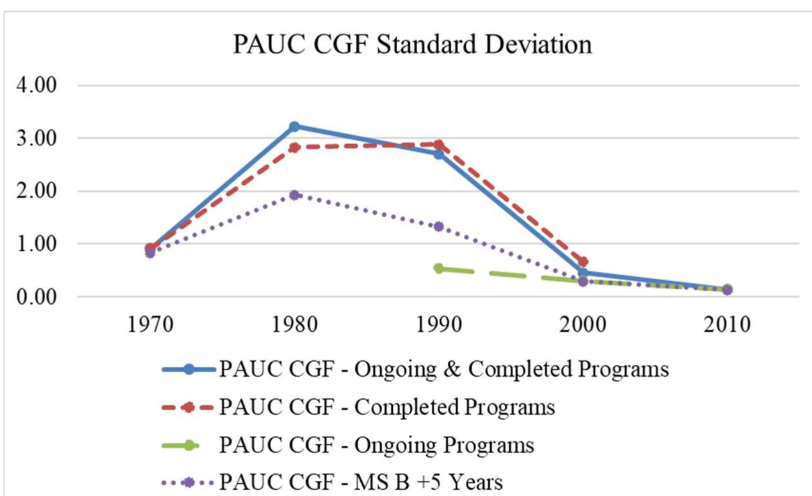
Appendix C – Standard Deviations - Master Table and Charts

	Schedule Percentage Growth - Ongoing & Completed Programs	Overall CGF - Ongoing & Completed Programs	PAUC CGF - Ongoing & Completed Programs	Schedule Percentage Growth - Completed Programs	Overall CGF - Completed Programs	PAUC CGF - Completed Programs	Schedule Percentage Growth - Ongoing Programs	Overall CGF - Ongoing Programs	PAUC CGF - Ongoing Programs	Schedule Percentage Growth - MS B +5 Years	Overall CGF - MS B +5 Years	PAUC CGF - MS B +5 Years
1970	0.15	3.62	0.91	0.15	3.69	0.91				0.15	2.70	0.83
1980	0.20	2.11	3.22	0.19	2.15	2.83				0.14	1.74	1.93
1990	0.28	1.17	2.70	0.26	1.00	2.88		1.35	0.53	0.22	1.24	1.33
2000	0.50	0.71	0.46	0.22	0.44	0.66	0.39	0.84	0.30	0.52	0.45	0.30
2010	0.31	0.37	0.14				0.32	0.37	0.14	0.29	0.37	0.13


Standard Deviation Chart
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
1970
1980
1990
2000
Schedule Percentage Growth - Ongoing & Completed Programs
Overall CGF - Ongoing & Completed Programs
PAUC CGF - Ongoing & Completed Programs
Schedule Percentage Growth - Completed Programs
Overall CGF - Completed Programs
PAUC CGF - Completed Programs
Schedule Percentage Growth - Ongoing Programs
Overall CGF - Ongoing Programs
PAUC CGF - Ongoing Programs
Schedule Percentage Growth - MS B +5 Years
Overall CGF - MS B +5 Years
PAUC CGF - MS B +5 Years

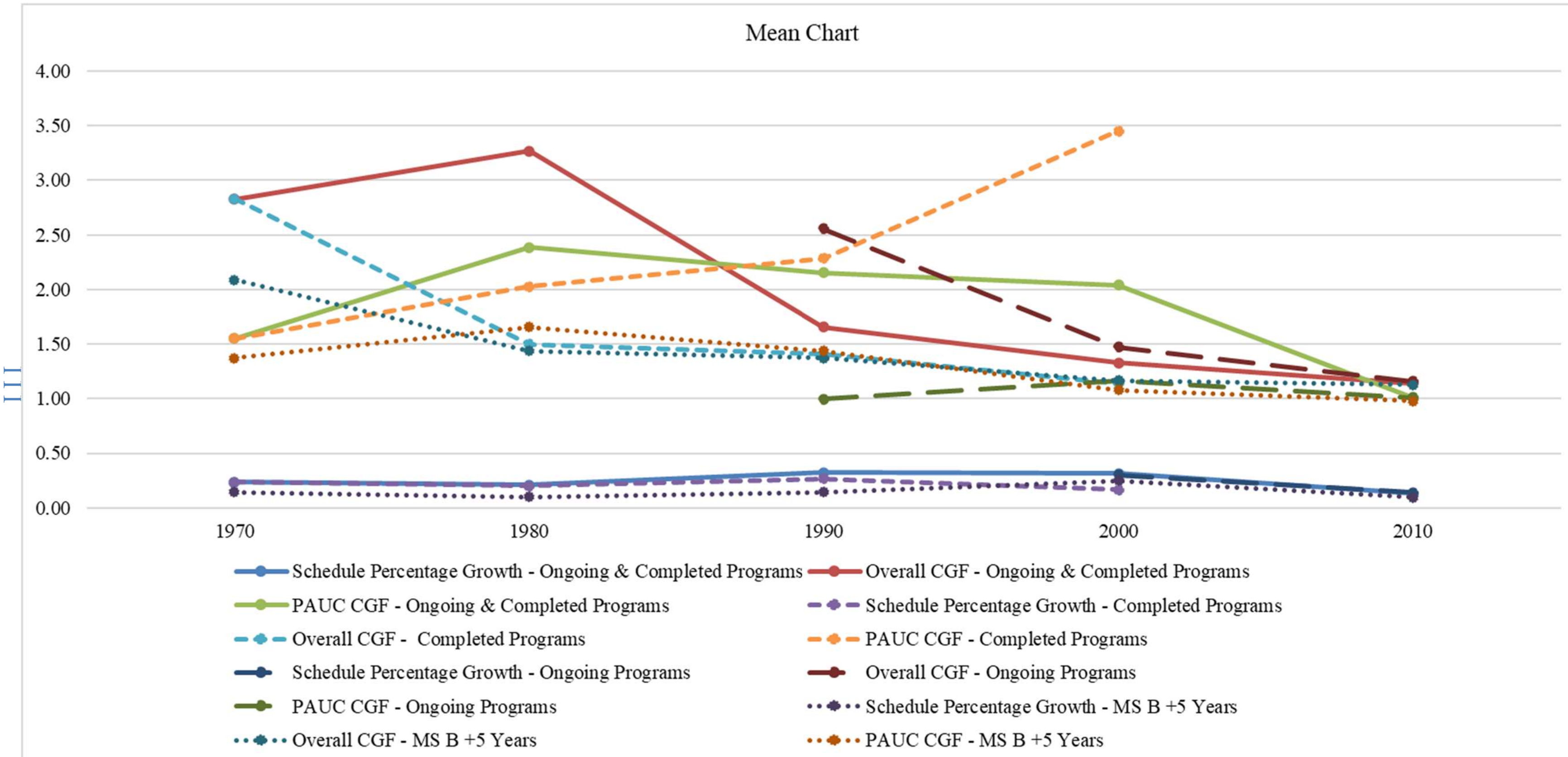

Schedule Growth Percentage Standard Deviation
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1970
1980
1990
2000
2010
Schedule Percentage Growth - Ongoing & Completed Programs
Schedule Percentage Growth - Completed Programs
Schedule Percentage Growth - Ongoing Programs
Schedule Percentage Growth - MS B +5 Years

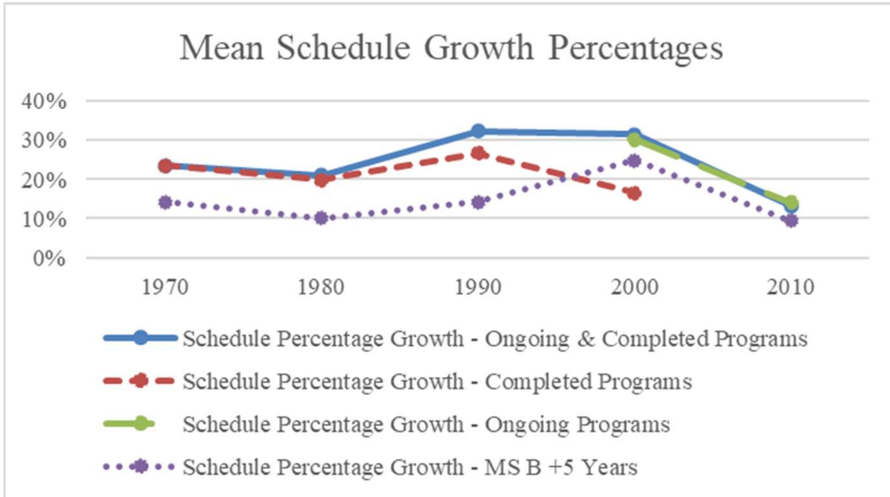

Overall CGF Standard Deviations
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
1970
1980
1990
2000
2010
Overall CGF - Ongoing & Completed Programs
Overall CGF - Completed Programs
Overall CGF - Ongoing Programs
Overall CGF - MS B +5 Years

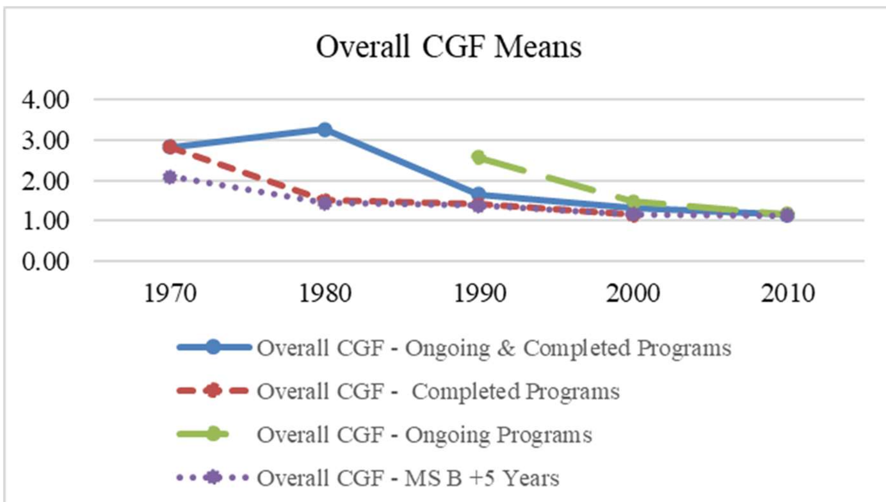

PAUC CGF Standard Deviation
4.00
3.00
2.00
1.00
0.00
1970
1980
1990
2000
2010
PAUC CGF - Ongoing & Completed Programs
PAUC CGF - Completed Programs
PAUC CGF - Ongoing Programs
PAUC CGF - MS B +5 Years

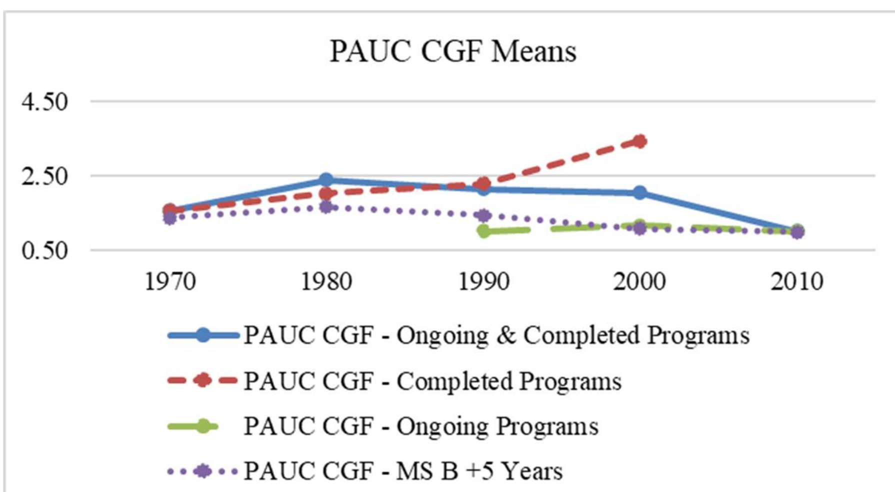
Appendix D – Means - Master Table and Charts

Means	Schedule Percentage Growth - Ongoing & Completed Programs	Overall CGF - Ongoing & Completed Programs	PAUC CGF - Ongoing & Completed Programs	Schedule Percentage Growth - Completed Programs	Overall CGF - Completed Programs	PAUC CGF - Completed Programs	Schedule Percentage Growth - Ongoing Programs	Overall CGF - Ongoing Programs	PAUC CGF - Ongoing Programs	Schedule Percentage Growth - MS B +5 Years	Overall CGF - MS B +5 Years	PAUC CGF - MS B +5 Years
1970	24%	2.83	1.55	24%	2.83	1.55				14%	2.09	1.37
1980	21%	3.27	2.39	20%	1.50	2.03				10%	1.44	1.66
1990	32%	1.66	2.15	27%	1.41	2.29		2.56	1.00	14%	1.37	1.44
2000	32%	1.33	2.04	17%	1.14	3.45	30%	1.47	1.16	25%	1.16	1.08
2010	13%	1.14	1.01				14%	1.16	1.01	9%	1.13	0.98


Mean Chart
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
1970
1980
1990
2000
2010
Schedule Percentage Growth - Ongoing & Completed Programs
Overall CGF - Ongoing & Completed Programs
PAUC CGF - Ongoing & Completed Programs
Schedule Percentage Growth - Completed Programs
Overall CGF - Completed Programs
PAUC CGF - Completed Programs
Schedule Percentage Growth - Ongoing Programs
Overall CGF - Ongoing Programs
PAUC CGF - Ongoing Programs
Schedule Percentage Growth - MS B +5 Years
Overall CGF - MS B +5 Years
PAUC CGF - MS B +5 Years

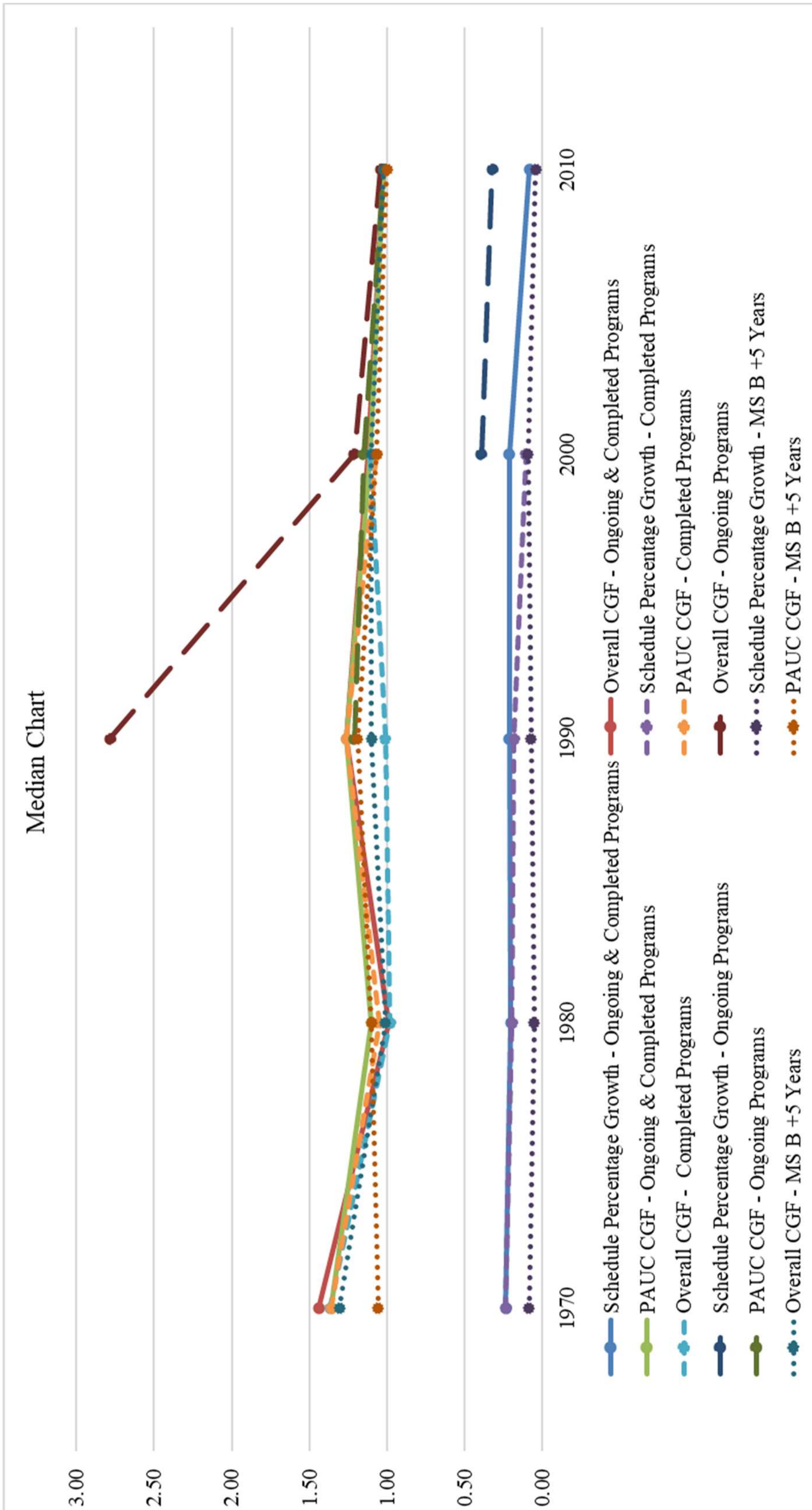

Mean Schedule Growth Percentages
40%
30%
20%
10%
0%
1970
1980
1990
2000
2010
Schedule Percentage Growth - Ongoing & Completed Programs
Schedule Percentage Growth - Completed Programs
Schedule Percentage Growth - Ongoing Programs
Schedule Percentage Growth - MS B +5 Years


Overall CGF Means
4.00
3.00
2.00
1.00
0.00
1970
1980
1990
2000
2010
Overall CGF - Ongoing & Completed Programs
Overall CGF - Completed Programs
Overall CGF - Ongoing Programs
Overall CGF - MS B +5 Years

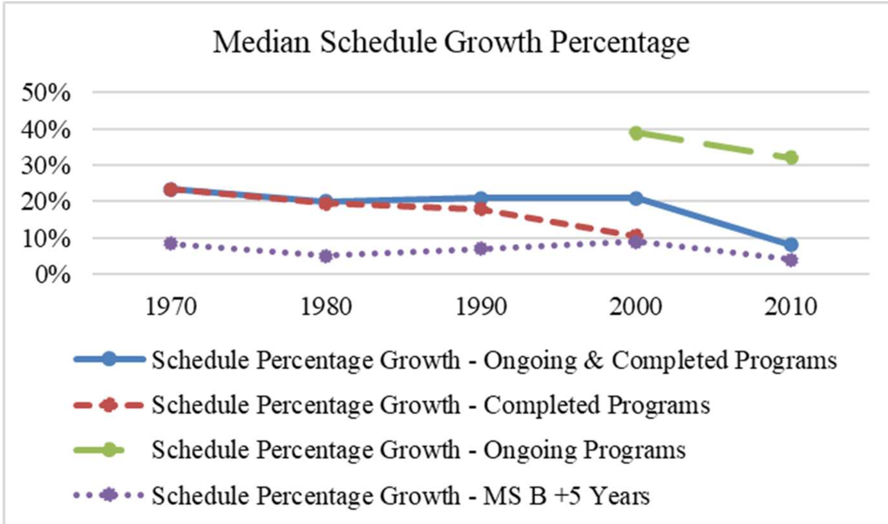

PAUC CGF Means
4.50
2.50
0.50
1970
1980
1990
2000
2010
PAUC CGF - Ongoing & Completed Programs
PAUC CGF - Completed Programs
PAUC CGF - Ongoing Programs
PAUC CGF - MS B +5 Years

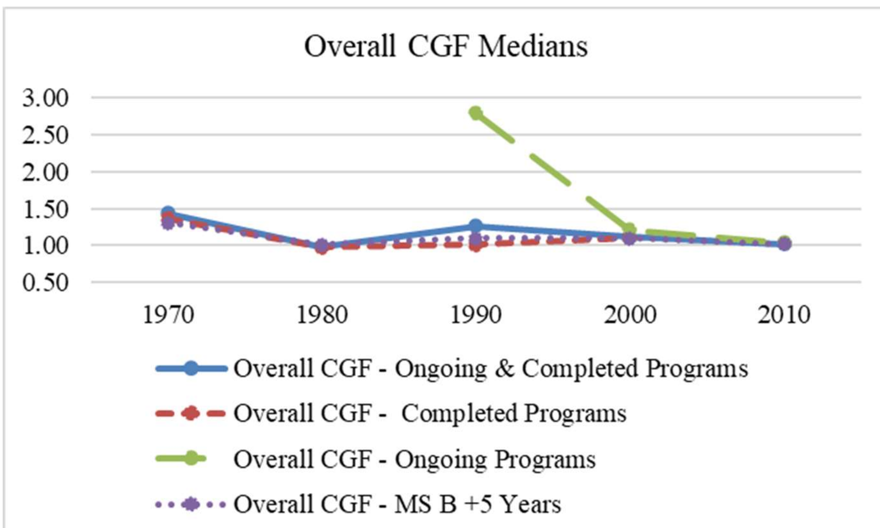
Appendix E – Medians - Master Table and Charts

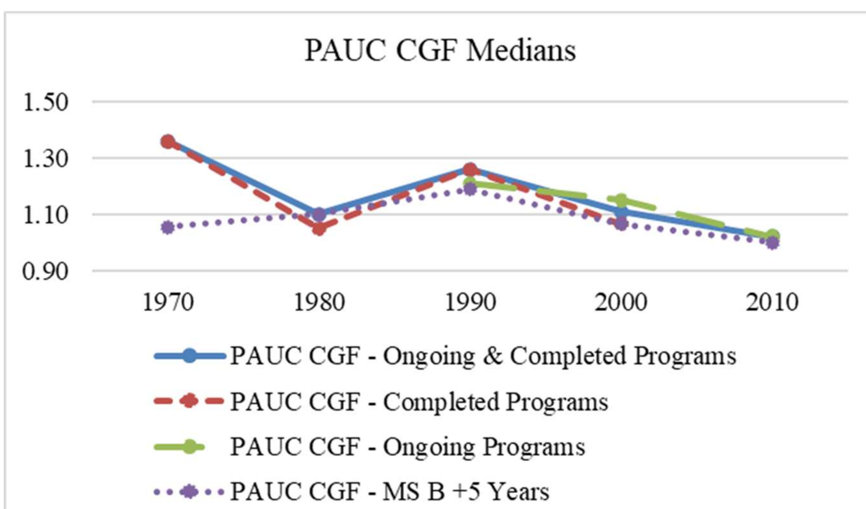
	Schedule Percentage Growth - Ongoing & Completed Programs	Overall CGF - Ongoing & Completed Programs	PAUC CGF - Ongoing & Completed Programs	Schedule Percentage Growth - Completed Programs	Overall CGF - Completed Programs	PAUC CGF - Completed Programs	Schedule Percentage Growth - Ongoing Programs	Overall CGF - Ongoing Programs	PAUC CGF - Ongoing Programs	Schedule Percentage Growth - MS B +5 Years	Overall CGF - MS B +5 Years	PAUC CGF - MS B +5 Years
1970	0.24	1.44	1.36	0.24	1.37	1.36				0.09	1.31	1.06
1980	0.20	0.99	1.10	0.20	0.98	1.05				0.05	1.01	1.10
1990	0.21	1.26	1.26	0.18	1.01	1.26		2.79	1.21	0.07	1.10	1.19
2000	0.21	1.12	1.11	0.11	1.11	1.07	0.39	1.21	1.15	0.09	1.10	1.07
2010	0.08	1.02	1.02				0.32	1.04	1.02	0.04	1.02	1.00



Median Chart
3.00
2.50
2.00
1.50
1.00
0.50
0.00
1970
1980
1990
2000
2010
Schedule Percentage Growth - Ongoing & Completed Programs
Overall CGF - Ongoing & Completed Programs
PAUC CGF - Ongoing & Completed Programs
Schedule Percentage Growth - Completed Programs
Overall CGF - Completed Programs
PAUC CGF - Completed Programs
Schedule Percentage Growth - Ongoing Programs
Overall CGF - Ongoing Programs
PAUC CGF - Ongoing Programs
Schedule Percentage Growth - MS B +5 Years
Overall CGF - MS B +5 Years
PAUC CGF - MS B +5 Years


Median Schedule Growth Percentage
50%
40%
30%
20%
10%
0%
1970
1980
1990
2000
2010
Schedule Percentage Growth - Ongoing & Completed Programs
Schedule Percentage Growth - Completed Programs
Schedule Percentage Growth - Ongoing Programs
Schedule Percentage Growth - MS B +5 Years


Overall CGF Medians
3.00
2.50
2.00
1.50
1.00
0.50
1970
1980
1990
2000
2010
Overall CGF - Ongoing & Completed Programs
Overall CGF - Completed Programs
Overall CGF - Ongoing Programs
Overall CGF - MS B +5 Years


PAUC CGF Medians
1.50
1.30
1.10
0.90
1970
1980
1990
2000
2010
PAUC CGF - Ongoing & Completed Programs
PAUC CGF - Completed Programs
PAUC CGF - Ongoing Programs
PAUC CGF - MS B +5 Years

References

AcqNotes LLC. (2021, March 23). *Acquisition Process Overview*. AcqNotes. https://acqnotes.com/acqnote/acquisitions/acquisition-process-overview

Arena, M. V., Leonard, R. S., Murray, S. E., & Younossi, O. (2006). *Historical cost growth of completed weapon system programs* (Report No. TR-343). Santa Monica, CA: RAND. https://apps.dtic.mil/sti/pdfs/ADA455020.pdf

Banford, A., & Weidman, D. (2014). *Analyzing the Effects of the Weapon Systems Acquisition Reform Act*. [Master's Thesis, Naval Postgraduate School] https://calhoun.nps.edu/bitstream/handle/10945/42579/14Jun_Banford_Weidman.pdf?sequence=1&isAllowed=y

Brown, M. B., & Forsythe, A. B. (1974). Robust Tests for the Equality of Variances. *Journal of the American Statistical Association*, *69*(346), 364–367. https://doi.org/10.1080/01621459.1974.10482955

Cross, S. (2006). *Data Analysis and its Impact on Predicting Schedule & Cost Risk Data Analysis and its Impact on Predicting Schedule & Cost Risk*. [Master's Thesis, Air Force Institute of Technology] https://scholar.afit.edu/cgi/viewcontent.cgi?article=4333&context=etd

Defense Acquisition University. (n.d.-a). *Defense Acquisition Glossary*. Www.dau.edu. Retrieved July 19, 2021, from https://www.dau.edu/glossary/Pages/Glossary.aspx#

Defense Acquisition University. (n.d.-b) *DAG-CH-1-Program-Management.* Www.dau.edu. https://www.dau.edu/pdfviewer?Guidebooks/DAG/DAG-CH-1-Program-Management.pdf

Department Of Defense. (2020, September 9). *The Defense Acquisition System* (DoD Directive 5000.01). Office of the Under Secretary of Defense for Acquisition and Sustainment, Department of Defense. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/500001p.pdf?ver=2020-09-09-160307-310

Dwyer, M., Tidwell, B., Blivas, A., & Hunter, A. (2020, April 8). *Cycle Times and Cycles of Acquisition Reform*. Acquisition Research Program, Naval Postgraduate School. https://dair.nps.edu/bitstream/123456789/4183/1/SYM-AM-20-046.pdf

Fox, J. R., Allen, D., Moody, W., & Shiman, P. (2012). *Defense Acquisition Reform 1960-2009: An Elusive Goal.* U.S. Government Printing Office. https://history.defense.gov/Portals/70/Documents/acquisition_pub/CMH_Pub_51-3-1.pdf.

Government Accountability Office. (2012, December). *WEAPONS ACQUISITION REFORM. Reform Act Is Helping DOD Acquisition Programs Reduce Risk, but Implementation Challenges Remain* (GAO Publication GAO-13-103). United States Government Accountability Office. https://www.gao.gov/assets/gao-13-103.pdf

Government Accountability Office. (2021, June). *Weapon Systems Annual Assessment Updated Program Oversight Approach Needed Report to Congressional Committees United States Government Accountability Office* (GAO Publication GAO-21-222). United States Government Accountability Office. https://www.gao.gov/assets/gao-21-222.pdf

Hough, P. (1992). *RAND Pitfalls in Calculating Cost Growth ·from Selected Acquisition*

Reports (Report No. N-3136-AF). Santa Monica, CA; RAND. https://apps.dtic.mil/sti/pdfs/ADA596255.pdf

House of Representatives, Congress. (2011, December 30). *10 U.S.C. 2430 - Major defense acquisition program defined.* Www.govinfo.gov; Government Publishing Office. https://www.govinfo.gov/app/details/USCODE-2011-title10/USCODE-2011-title10-subtitleA-partIV-chap144-sec2430

Julian. (2014, August 13). *Re: Does JMP use Bonferroni correction for ANOVA and Kruskal-Wallis?* [Online forum post] JMP User Community. https://community.jmp.com/t5/Discussions/Does-JMP-use-Bonferroni-correction-for-ANOVA-and-Krusal-Wallis/m-p/9186#M9107

Kozlak, S., White, E., Ritschel, J., Lucas, B., & Seibel, M. (2017). Analyzing Cost Growth at Program Stages for DoD Aircraft. *Defense Acquisition Research Journal*, *24*(3), 386–407. https://doi.org/10.22594/dau.16-763.24.03

Laerd Statistics. (2014). *Kruskal-Wallis H Test in SPSS Statistics | Procedure, output and interpretation of the output using a relevant example.* Lund Research Ltd. https://statistics.laerd.com/spss-tutorials/kruskal-wallis-h-test-using-spss-statistics.php

Light, T., Leonard, R. S., Pollak, J., Smith, M. L., & Wallace, A. (2017). *Quantifying Cost and Schedule Uncertainty for Major Defense Acquisition Programs (MDAPs)* (Report No. RR1723). Santa Monica, CA; RAND. https://www.rand.org/pubs/research_reports/RR1723.html

McNicol, D., Tate, David M., Burns, Sarah K., & Wu, Linda, (2016). *Further Evidence on the Effect of Acquisition Policy and Process on Cost Growth of Major Defense*

Acquisition Programs (Revised) (IDA Publication No. P-5330). Institute for Defense Analysis. https://apps.dtic.mil/sti/pdfs/AD1013803.pdf

McNicol, D (2018). *Acquisition Policy, Cost Growth, and Cancellations of Major Defense Acquisition Programs* (IDA Publication No. R-8396). Institute for Defense Analysis. https://apps.dtic.mil/sti/pdfs/AD1122550.pdf

Office of the Deputy Assistant Secretary of the Army. (2018). *U.S. Army Cost Benefit Analysis Guide*. https://www.asafm.army.mil/Portals/72/Documents/Offices/CE/US%20Army%20Cost%20Benefit%20Analysis.pdf

Office of the Under Secretary of Defense (Comptroller). (2020a, January 30). *Program Acquisition Cost by Weapon System* (OUSD Report No. B-5152D55). Office of the Under Secretary of Defense (Comptroller), Department of Defense. https://comptroller.defense.gov/Portals/45/Documents/defbudget/fy2021/fy2021_Weapons.pdf

Office of the Under Secretary of Defense (Comptroller). (2020b). *PROCUREMENT PROGRAMS (P-1)*. Office of the Under Secretary of Defense (Comptroller), Department of Defense. https://comptroller.defense.gov/Portals/45/Documents/defbudget/fy2021/fy2021_p1.pdf

OSD CAPE. (2019, February 14). *CADE Selected Acquisition Report (SAR) Database Now Available in CADE - Applications*. Cade.osd.mil. https://cade.osd.mil/News/blog

Ritschel, J., White, E., Lucas, B., & Mrla, D. (2019). The Impact of WSARA on the Cost of Air Force Weapon Systems. *Journal of Public Procurement*, *19*(1), 2–14. https://doi.org/10.1108/JOPP-03-2019-022

Stephanie, G. (2015, March 11). *Brown-Forsythe Test: Definition*. Statistics How To. Retrieved December 15, 2021, from https://www.statisticshowto.com/brown-forsythe-test/

Weapon Systems Acquisition Reform Act of 2009, P.L. No. 111 - 23 (2009). https://www.congress.gov/111/plaws/publ23/PLAW-111publ23.pdf

Younossi, O., Arena, M. V., Leonard, R. S., Roll, C. R., Jain, A., & Sollinger, J. M. (2007). Is Weapon System Cost Growth Increasing? A Quantitative Assessment of Completed and Ongoing Programs (Report No. MG-588-AF). Santa Monica, CA; RAND. https://www.rand.org/pubs/monographs/MG588.html

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 074-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
22-03-2022	Master's Thesis	October 2020-March 2022

TITLE AND SUBTITLE	
ANALYSIS OF COST AND SCHEDULE ESTIMATION TRENDS FOR MAJOR DEFENSE ACQUISITION PROGRAMS	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
AUTHOR(S)	**5d. PROJECT NUMBER**
Jones, Sammantha J.	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-7765	AFIT-ENV-MS-22-M-215
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Life Cycle Management Cost Staff Shawn Valentine, Civ, USAF 1865 Fourth Street, Building 14, Area B Wright-Patterson AFB, OH 45433-7765 937-656-5504 ATTN: Shawn Valentine (shawn.valentine@us.af.mil)	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFLCMC/FZCE **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRUBTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This study uses both descriptive and inferential techniques to investigate trends in cost and schedule estimates created by the Department of Defense for their Major Defense Acquisition Programs (MDAPs) throughout the last five decades. For schedule growth percentages, we did not identify any statistically significant trend regarding increasing or decreasing schedule changes or variances of schedule estimates throughout the five decades analyzed. Examining the overall cost growth of MDAPs, a statistically significant difference between the Cost Growth Factors (CGFs) calculated between the 1990s and the 2010s was found, with the 2010s exhibiting lower CGFs. A downward trend in the variances of overall CGFs throughout the decades was identified in three out of four of the analyses conducted. Lastly, differences were detected in the program acquisition unit cost (PAUC) CGFs between the 1990s and the 2010s, and slight differences between the 1970s and the 2010s. Additionally, PAUC CGFs also displayed differences in variances. Overall, we found no trends identifying that the DoD is improving accuracy of cost or schedule estimates for MDAPs; however, we did identify a statistically consistent reduction in the variances of overall CGFs for MDAPs through the five decades. This finding appears to be the first documented case known to us.

15. SUBJECT TERMS
DoD acquisitions, schedule growth, cost growth, PAUC growth

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. Edward D. White III, AFIT/ENC
U	U	U	UU	135	**19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-3636, ext. 4540. e-mail: Edward.White@afit.edu

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