SYM-AM-18-039

PROCEEDINGS
OF THE
FIFTEENTH ANNUAL ACQUISITION RESEARCH SYMPOSIUM

WEDNESDAY SESSIONS
VOLUME I

Acquisition Research:
Creating Synergy for Informed Change

May 9–10, 2018

Published April 30, 2018

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL

Further Evidence on Program Duration and Unit Cost Growth

David McNicol—joined the Department of Defense (DoD) in 1982. From 1988 until 2002, he was a Deputy Director of Program Analysis and Evaluation (PA&E). Earlier, Dr. McNicol taught at the University of Pennsylvania and the California Institute of Technology. He holds a BA in Economics from Harvard and an MS in Management and PhD in Economics and Finance from MIT. Employed at the Institute for Defense Analyses (IDA) since his retirement from the DoD, he became Director of the Cost Analysis and Research Division in 2006. Still at IDA, Dr. McNicol stepped down in 2012 to return to his previous role as a Research Staff Member. [dmcnicol@ida.org]

Abstract

David McNicol, in "Post-Milestone B Funding Climate and Cost Growth in Major Defense Acquisition Programs," in *Proceedings of the 14th Annual Acquisition Research Symposium, Vol. 1*, explored the association between Program Acquisition Unit Cost (PAUC) growth of Major Defense Acquisition Programs (MDAPs) and funding climates post-Milestone (MS) B. A strong positive association was found for MDAPs that passed MS B in a bust funding climate; the association was weak for programs that passed MS B in boom climates. This paper uses four alternative regression equations to extend these results. In each case, the same pattern of results appears—MDAPs that passed MS B in a bust climate had significantly higher growth than those that passed MS B in a boom climate, the 1969 Packard reforms reduced average PAUC growth, and the reduction persisted through the end of the study (FY 2009)—but changes to the acquisition process after the Packard reforms through 2009 did not further reduce average PAUC growth. The lower PAUC growth after the Packard reforms probably was due mainly to more realistic MS B baselines. This pattern does not depend on the inclusion of post-MS B funding climate and program duration, although those factors have significant effects.

Introduction[1]

McNicol (2017b) explored the association between growth in the unit costs of Major Defense Acquisition Programs (MDAPs) and the funding climates programs experienced after passing Milestone (MS) B.[2] While this topic arose by serendipity, a little reflection establishes that it is plausible to expect average cost growth to be higher for MDAPs that entered a boom climate sometime after passing MS B than it is for those that did not. MDAPs that passed MS B in bust climates probably are especially influenced by a post-MS

---

[1] This paper draws on Chapters 2 and, especially, 3 of McNicol (2018), which is a synthesis of a series of papers on the association of funding climate, acquisition policy, and other factors on cost growth of major acquisition programs.

[2] While the label MS B is used here for all time periods, through the years there have been changes in the labels used for milestones and, to some extent, in their definitions. During FY 1966–1969, there were two milestones in the OSD-level acquisition process, neither of which had a name. Reforms instituted early in FY 1970 provided for three milestones, labeled MS I, MS II, and MS III. In 1987, MS IV and MS V were added. By 1991, MS IV had been eliminated and what had been MS V became MS IV. MS IV had been eliminated by 1996. In 2000, the milestones were changed to MS A, MS B, and MS C, and the definition of MS B modestly changed. See McNicol (2018), Chapter 5.

B boom. Some of these programs presumably had unrealistic baselines and would find a post-MS B boom climate a good time to "get well." Even programs that established realistic baselines at MS B might tend to be less capable than the service wanted and good candidates for adding capability when funding constraints were relaxed.

A strong positive association was found between MDAPs that passed MS B in a bust funding climate and subsequently entered a boom funding climate. The association was much weaker for programs that passed MS B in boom climates. This paper uses a series of four models to extend those results by incorporating acquisition policy variables and program duration. The section titled Cost Growth Due to Program Changes also briefly examines the extent to which PAUC growth post-MS B reflects costs due to subsequent decisions to acquire capabilities beyond those of the MS B baseline.

The Models

The first of the models (Model 1) relates cost growth only to funding climate and acquisition policy configuration.[3] Next (Model 2), post-MS B funding climate is introduced into the model. Two models (Models 3 and 4) that in different ways include both post-MS B funding climate and program duration are then presented. We begin by pointing to important features common to the four models.

Framework of the Models

The topic requires distinguishing between bust and boom funding climates. During the 45 years (fiscal year [FY] 1965–FY 2009) covered by this study, there were two complete bust-boom cycles in Department of Defense (DoD) procurement funding: (1) the bust climate for modernization of weapon systems that began in the mid-1960s and lasted until the Carter-Reagan buildup of the early to mid-1980s, and (2) the long post-Cold War bust climate followed by the post-9/11 boom.

Where a bust funding climate may provide an upward pull on cost growth, acquisition policy and process can be expected to provide a restraining push. For that reason, it is necessary also to recognize changes over time in acquisition policy and process configurations. Five policy and process configurations are distinguished:

1. McNamara-Clifford (FY 1964–FY 1969)
2. Defense Systems Acquisition Review Council (DSARC; FY 1970–FY 1982)
3. Post-Carlucci DSARC (P-C DSARC; FY 1983–FY 1989)
4. Defense Acquisition Board (DAB; FY 1990–FY 1993 and FY 2001–FY 2009)
5. Acquisition Reform (AR; FY 1994–FY 2000)

Policy and *process* tend to be intertwined; process typically is required to implement policy, and the most successful and durable policies are those embedded in process. For this reason, and to avoid constant repetition of "process and policy," the term *acquisition policy* is used here in a broad sense to encompass both policy on particular topics (for

[3] The most developed explanations of funding climate and acquisition policy configuration are provided in McNicol (2018), Chapter 1.

example, contract types) and the Office of the Secretary of Defense (OSD)-level oversight process (for example, definition of the milestones).

Finally, a measure of cost growth is required. The measure used is based on Program Acquisition Unit Cost (PAUC). PAUC is the sum of Research, Development, Test and Evaluation cost and procurement cost, divided by the number of units acquired. For this paper, PAUC growth is computed by comparing the MS B baseline value of PAUC in program base-year dollars—which can be thought of as a goal or a prediction—to the actual PAUC reported in the program's last Selected Acquisition Report (SAR) in program base-year dollars and adjusted to the MS B baseline quantity. Appendix B of McNicol (2017a) describes the conventions used in assembling the database, the sources of the data used, and the quantity adjustment computations. The unit cost growth estimates were updated to the December 2015 SARs. Only completed programs (defined as programs with an end date of FY 2016 or earlier) are used in this analysis because some costs associated with a program may not be fully reflected in its SAR until the program is completed.

To be clear, in what follows, the term *PAUC growth* means PAUC growth as defined previously, that is, growth from MS B through the end of procurement, adjusted to the MS B quantity.

Model 1—The Baseline Model

The "baseline model" is the following assumed relationship:

$$PAUC_i = a_0 + a_1Climate_i + a_2DSARC_i + a_3P\text{-}CDSARC_i + a_4DAB_i + a_5AR_i + e_i \quad (1)$$

The subscript i denotes the ith MDAP in the sample. This model provides a baseline in that it includes as independent variables only funding climate and acquisition policy configuration.

Climate is a categorical variable[4]; it takes on a value of zero for MDAPs that passed MS B in bust climates and 1 for those that passed in boom climates. The intercept term a_0 is assumed to measure primarily the climate effect.[5] For programs that passed MS B in a bust climate, a_0 is the intercept; for those that passed in a boom climate, the intercept is $a_0 + a_1$. The expectation is that the estimate of a_1 is negative; that is, that MDAPs that passed MS B in a boom climate on average have lower PAUC growth.

The model includes a categorical variable for each of the four acquisition policy configurations. These variables have a value of 1 for the years of the period in question (e.g., FY 1994–FY 2000 for AR), and zero for other years. For technical reasons, one of a set of categorical variables always must be omitted (or the constant term constrained to zero). The selection of the omitted variable is arbitrary insofar as the statistics are concerned; the McNamara-Clifford period was chosen because that is convenient for the exposition. The estimated coefficient of each of the acquisition policy categorical variables is the difference between average PAUC growth in that bin and average PAUC growth in

[4] These are often referred to as "dummy variables" but are more descriptively called categorical variables or indicator variables.
[5] The estimated coefficient of Climate also includes the average net effect of any relevant variables not included in this model and the effect on the estimated intercept of any non-linearity in the response of PAUC growth to the model's explanatory variables.

McNamara-Clifford. That difference is statistically significant if the estimated coefficient of the acquisition policy period categorical variable is statistically significant.[6]

Finally, the error term e_i represents myriad unpredictable factors that influence PAUC growth; it is assumed to be a normally distributed random variable with a mean of zero and constant variance. The coefficients of the model are estimated using ordinary least squares (OLS; also known as multiple regression, linear regression, and least squares regression).[7] The results are presented in Table 1. We use the p-value to characterize statistical significance and refer to any estimate with a p-value of no more than 0.10 as "statistically significant." A p-value of 0.10 means that there is an (estimated) one chance in 10 that the observed estimate would occur by chance even if the true value of the coefficient were zero.

---

[6] Note that for all of the observations of the McNamara-Clifford period, PAUCj = $a_0 + e_j$, and since it is assumed that $E(e_j) = 0$, $E(PAUC_j) = a_0$. If the underlying model is correct and the assumptions of OLS are satisfied, the estimated value of the intercept (denoted $\hat{a}_0$))" is an unbiased estimate of a_0 and of the sample value of the average PAUC growth of the McNamara-Clifford period. Similarly, the expected value of the intercept and the average PAUC growth for the ith acquisition policy bin is "$a_0 + a_i$," and the difference between that and the average for the reference group is $a_0 - (a_0 + a_i) = -a_i$. Hence, if $\hat{a}_i$ is statistically significantly different from zero, the average PAUC growth for acquisition policy configuration i is significantly different from average PAUC growth for McNamara-Clifford. The burden of the assumptions is lightened by the fact that, in this context, "just about" counts. For example, no great harm is done if $E(e_j)$ is small rather than zero.

[7] Readers unfamiliar with this technique can find an explanation in any introductory econometrics text, in many introductory statistics texts, or on the internet. For example, *TIBC Statistica*, http://www.statsoft.com/Textbook/Multiple-Regression; Penn State Eberly College of Science, STAT 501: Regression Methods, https://onlinecourses.science.psu.edu/stat501/node/283; John H. McDonald, *Handbook of Biological Statistics*, http://www.biostathandbook.com/multipleregression.html; and David M. Lane, "Introduction to Multiple Regression," Chapter 14, in *Online Statistics Education: An Interactive Multimedia Course of Study*, http://onlinestatbook.com/2/regression/multiple_regression.html.

Table 1. Estimated Parameters of the Basic Model of PAUC Growth

	Coefficient	p-value
Intercept	86.7%***	< 0.001
Funding Climate		
Climate	-37.3%***	< 0.001
Acquisition Policy		
DSARC	-47.1%***	< 0.001
P-C DSARC	-42.3%***	0.003
DAB	-47.0%***	< 0.001
AR	-65.4%***	< 0.001

* Statistically significant at less than the 10% level.

** Statistically significant at less than the 5% level.

*** Statistically significant at less than the 1% level.

R-Square = 0.25 F = 9.720 (P < 0.001) N= 149. Estimated by OLS. With the Bonferroni correction, Wald's test for the equality of the estimated coefficients the categorical variables for acquisition policy periods yields F = 1.02, p = 0.999.

The dataset used to estimate the model in Table 1 omits four extremely long duration programs. Each of these "programs" is actually a series of modifications and upgrades of an initial program reported on the SAR of the original program. Also excluded are three programs from the early 1980s boom period that were acquired using variants of Total Package Procurement (TPP). These observations were excluded for reasons stated below.

The criteria typically used to judge regression equations readily accept the results in Table 1:

- The intercept and the estimated coefficient of each of the independent variables have the expected signs.
- Their magnitudes are reasonable.[8]
- The intercept and the estimated coefficients of the independent variables are highly significant.
- The estimated equation as a whole is highly significant.
- The proportion of the variation in sample PAUC growth captured by the estimated equation is towards the upper end of what can be expected for panel data.

---

[8] Evaluations of the reasonableness of the estimated coefficients of the acquisition policy periods must weigh the Climate effect by the proportion of the acquisition policy period spent in a boom climate. This was, for example, zero for AR and 0.154 (=2/13) for DSARC.

In addition, the overall features of the results are consistent with what would be expected from the history of OSD-level oversight of MDAPs over the relevant period (FY 1965–FY 2009).[9]

Four important conclusions are implied by the estimates in Table 1:

- The highly significant negative coefficient of Climate implies that the average PAUC growth of programs that passed MS B in a boom climate was significantly less than that of programs that passed in a bust climate.
- The 1969 Packard reforms of the acquisition process (which define the DSARC bin) resulted in a significant reduction in average PAUC growth compared to that of the preceding McNamara-Clifford period.
- The other three acquisition policy configurations (P-C DSARC, DAB, and AR) also had average PAUC growth significantly lower than that of McNamara-Clifford.
- Average PAUC growth in the four post-McNamara-Clifford acquisition policy bins did not differ significantly from one another.[10]

In brief—funding climate has the expected association with PAUC growth, the 1969 Packard reforms reduced average PAUC growth, and the reduction persisted through the end of the study period (FY 2009), but changes to the acquisition process after the Packard reforms through FY 2009 were not associated with further reductions in average PAUC growth.

The regression in Table 1 contains a remarkable feature. Ordinarily, when outliers (of the dependent variable) are removed from the dataset, the test statistics of the regression improve. This is not the case for the baseline model. If PAUC and duration outliers and programs procured with TPP are removed from the dataset, three of the four estimated coefficients of acquisition policy bins (including that for DSARC) are not statistically different from zero. The point is that the results are driven by the extreme values of PAUC growth. That is to say, the 1969 Packard reforms were effective because they reduced the frequency of MDAPs with extremely high PAUC growth.

Table 2 provides data that can be used to directly test this interpretation. The striking feature of these data is the paucity of outliers after the introduction of the Packard reforms in 1969. The PAUC growth of three of the 16 programs of the McNamara-Clifford years was large enough (at least 134%) to qualify as an outlier[11]; of the 94 MDAPs that passed MS B during the other four periods, only two had PAUC growth of at least 134%. This difference is

---

[9] See McNicol (2018), Chapter 5.

[10] This statement rests on the results of Wald's test with the Bonferroni correction. Wald's test, as used here, tests whether, considered jointly, any of $\hat{a}_1$, $\hat{a}_2$, $\hat{a}_3$, and $\hat{a}_4$ are significantly different from the others. The Bonferroni correction effectively increases the critical value used to judge statistical significance to recognize that in multiple comparisons there is a considerable probability of a significant difference arising by chance even if the underlying population values are identical.

[11] We use the word *outlier* here as defined by John Tukey: observations 1.5 times the Inter Quartile Range above the third quartile or below the first quartile. None of the outliers had exceptionally low PAUC growth.

statistically significant.[12] Similar differences were not found for PAUC growth of at least 50% and at least 100%.[13] It appears then that the Packard reforms worked mainly by reducing the frequency of very high cost growth programs rather than by reducing cost growth on programs generally.

Table 2. Average PAUC Growth by Acquisition Policy Configuration and the Number of High Cost Growth MDAPs in Each Cohort, Bust Funding Climates for Completed Programs

Acquisition Policy Configuration	Period (FY)	Average PAUC Growth*	≥ 50%	≥ 100%	≥ 134%
McNamara-Clifford	1964–1969	74% (16)	9	4	3
DSARC	1970–1980	37% (49)	18	4	0
Post-Carlucci DSARC	1987–1989	34% (11)	2	2	1
DAB	1990–1993 2001–2002	40% (15)	5	1	0
Acquisition Reform (AR)	1994–2000	31% (19)	5	1	1

Model 2—The Basic Model Plus Boom Effects

Model 2 is prompted by the conjecture that boom climates facilitate PAUC growth of ongoing programs that enter them. If this is so, average PAUC growth presumably will be higher for MDAPs that entered a boom climate sometime after passing MS B than it will for those that did not. This would in particular be expected of programs that passed MS B in a bust climate, but it might also be true of programs that passed in a boom climate.

A two-part naming convention is used to label programs that encountered a boom climate post-MS B and those that did not. The first part of the label gives the funding climate prevailing when the program passed MS B—bust or boom. The second part—0, 1, or 2—denotes the number of boom climates the program entered post-MS B. Programs that passed MS B in a bust climate and were completed entirely within that bust climate will be referred to as Bust0—Bust because they passed MS B in a bust funding climate and zero because they were completed without entering a boom climate. Programs that passed MS B in a bust period and continued into a subsequent boom period make up Bust1 or, for the few programs that extended into two boom periods, Bust2. Programs that passed MS B in boom climates are, similarly, denoted Boom0 or Boom1. (There are no programs in Boom2 as of this writing because programs that passed MS B during the Carter-Reagan boom climate had only one subsequent boom climate they could enter, the post 9/11 boom.)

These definitions capture just one feature of the post-MS B funding climates experienced by programs. They exclude other features that possibly are relevant. For example, they do not take into account the time spent in different funding climates or

[12] Fisher's Exact Test (FET); p = 0.021. Application of FET to the five bins of the 134% column of Table 2 yields p = 0.016.
[13] FET; p = 0.297 and p = 0.271 for 50% and 100%, respectively.

transitions from boom to bust of programs that passed MS B in boom periods. These simple definitions do nonetheless provide a way to examine whether boom effects are visible in the data.

We start with average PAUC growth for Bust0, Bust1, and Bust2 presented in Table 3 for the post-McNamara-Clifford portion of the first bust-boom cycle period[14] and the entire bust portion of the second cycle. Recall that only data for completed programs are used. In both periods, average PAUC growth for the treatment group (Bust1) is higher than it is for the control group (Bust0)—42% compared to 18% for the first period and 44% compared to 12% for the second. These differences are statistically significant.[15] For programs that passed MS B in a bust period, subsequent entry into a boom period is then associated with higher PAUC growth. PAUC growth for Bust2 is higher than that of Bust0 but less than that of Bust1.[16] Average PAUC growth for Bust2, however, is not significantly different from that for either Bust0 or Bust1.

Table 3. Average PAUC Growth by the Number of Boom Periods Experienced for Completed MDAPs That Passed MS B in Post–McNamara-Clifford Bust Climates

Bin	1st Bust Period FY 1970–FY 1980	2nd Bust Period FY 1987–FY 2002
Bust0	18% (7)	12% (10)
Bust1	42% (38)	44% (35)
Bust2	28% (4)	none

The next step is to include boom effects in the baseline model. We define two variables, C_{boom}Bust and C_{boom}Boom. C_{boom}Bust is 1 for all programs in Bust1 and Bust2; these programs passed MS B in a bust and then experienced one or two boom periods post-MS B. For all other programs, C_{boom}Bust is zero. Similarly, C_{boom}Boom is 1 for programs in Boom1 and zero for all other programs; these programs passed MS B in the Carter-Reagan boom and then experienced the 9/11 boom period post-MS B. The results are presented in Table 4.

---

[14] Average PAUC growth for Bust0 programs is 87% and that for Bust1 is 34%. The anomaly here is not the average PAUC growth for Bust1—which is in line with the averages for the other bust periods—but the exceptionally high cost growth of Bust0.

[15] Kolmogorov-Smirnov (K-S) and Anderson-Darling (A-D) find the PAUC growth data in each of the three bins of the first bust period to be consistent with a normal distribution. An F-test found the variances for Bust0 and Bust1to be significantly different. A two-tailed t-test assuming unequal sample variances found the means of Bust1 and Bust0 for the first period to be significantly different ($p = 0.003$). K-S finds the distribution of PAUC growth for Bust1 of the second bust period to be non-normal. The means of Bust0 and Bust 1 for the second bust period are significantly different by the Man-Whitney U test: $p = 0.018$, $U = 97.5$, $n_1 = 35$, $n_2 = 10$.

[16] The programs in Bust2 are the CVN 68, with a PAUC growth of 7%; the NAVSTAR GPS (85%); ATCCS-MCS (-34%); and the UH-60A (54%). A two-tailed t-test, with unequal variances as appropriate, found the mean of Bust2 not to be significantly different from that of Bust0 ($p = 0.732$) or Bust1 ($p = 0.440$).

Table 4. Estimated Coefficients and p-Values for a Model That Includes the Effects of Climate at MS B and Post-MS B

	Coefficients	p-value
Intercept	81.5%***	< 0.001
Funding Climate		
Climate	-26.6%**	0.027
C_{boom}Bust	18.0%*	0.099
C_{boom}Boom	8.8%%	0.702
Acquisition Policy		
DSARC	-56.5%***	< 0.001
P-C DSARC	-47.1%***	0.001
DAB	-55.9% **	< 0.001
AR	-77.1%***	< 0.001

* Statistically significant at less than the 10% level.

** Statistically significant at less than the 5% level.

***Statistically significant at less than the 1% level.

R-Square = 0.27 F = 7.420 (P < 0.001) N= 149. Estimated using OLS. Boom2 programs and the three mid-1980s MDAPs acquired using TPP-like contracts are omitted. With the Bonferroni correction, Wald's test for the equality of the estimated coefficients of the categorical variables for acquisition policy periods yields F = 1.47 (p =0.90).

Three MDAPs in Boom1 were acquired using a TPP contract. These three programs have much higher PAUC growth than Boom0 programs—because they were acquired using a TPP contract, not because they passed into a boom period post-MS B. For that reason, they are dropped from the sample.[17]

The estimated coefficient of C_{boom}Bust (18.0%) is marginally significant. We do, then, see evidence of a boom effect. The estimated coefficient of C_{boom}Boom (8.8%) is smaller and not significant. As with the previous model, estimates imply that the Packard reforms of 1969 resulted in a decrease in PAUC growth; that decrease persisted, but subsequent changes in acquisition policy apparently did not result in further significant decreases in PAUC growth.

Model 3—Program Duration

The longer a program's duration, the greater its chance of moving into a boom funding climate. For that reason alone, longer duration presumably is associated with higher PAUC growth.

Figure 1 provides some evidence on the premise of the discussion. It plots the average PAUC growth for the bust and boom bins along with the corresponding average program duration (defined as the number of years from MS B through the end of the

[17] See McNicol (2018) Chapter 3, Section B.

acquisition phase).[18] The prefixes 1st and 2nd indicate the bust-boom cycle—FY 1965–FY 1986 (1st)[19] and FY 1987–FY 2009 (2nd). Programs that passed MS B in bust climates added more PAUC per year of duration than did programs that passed MS B in a boom climate.

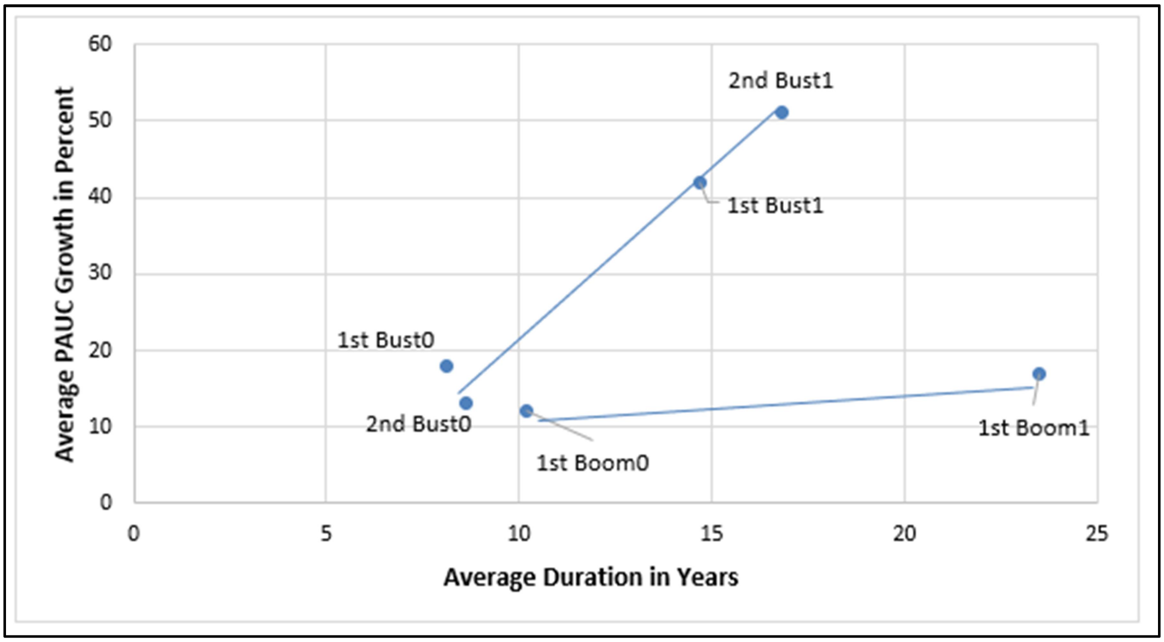

Figure 1. Average PAUC Growth and Average Program Duration for Boom and Bust Periods

A plausible approach to separating the boom effect from a duration effect is to enter into the model a variable defined as the number of years spent in boom climates (T_{boom}) and another variable that is the number of years spent in bust climates (T_{bust}). Very simple definitions of T_{boom} and T_{bust} were adopted:

- T_{boom} = number of years in boom climates post-MS B
- T_{bust} = number of years in bust climates post-MS B

Note that this definition counts a year during which the program was in Engineering and Manufacturing Development the same as a year in which the program was in Low Rate Initial Production (LRIP) or full rate production. There are several alternatives to this definition. For example, the duration variables might be defined as the years in boom and bust climates, respectively, after the program enters LRIP.

Setting aside for the moment the categorical variables for the acquisition policy configurations, the core model considered is shown in Equation 2:

$$PAUC_i = a + bClimate + cT_{boom,i} + dT_{bust,i} + v_i. \qquad (2)$$

---

[18] The end of the acquisition phase was defined as the final year in which substantial procurement funding was obligated, as reported in the program's final SAR.
[19] The McNamara-Clifford period was excluded from the first cycle. See footnote 14

where v_i is the error term. Note that c and d are measured in units of percentage points per year; they are the rates at which programs' PAUC growth increases per year in boom and bust climates, respectively.

We expect the estimated coefficient of Climate to be negative, implying that programs that passed MS B in boom climates have lower PAUC growth than those that passed in bust climates. This specification also allows for climate effects in that the estimates of c and d may be different. In particular, we would expect the estimate of c to be larger than that of d—that is, that PAUC growth accumulates more rapidly in boom than in bust years. Estimates of the parameters of this model expanded to include the categorical variables for the acquisition policy configurations are presented in Table 5.

Table 5. Estimated Coefficients and p-Values for a Model That Includes the Effects of Post-MS B Funding Climate and Duration

	Coefficients	p-value
Intercept	76.5%***	< 0.001
Funding Climate		
Climate	-40.2%***	< 0.001
T_{boom}	4.2%/yr***	0.008
T_{bust}	0.22%/yr	0.804
Acquisition Policy		
DSARC	-57.5%***	< 0.001
P-C DSARC	-44.5%***	0.002
DAB	-57.5%***	< 0.001
AR	-83.1%***	< 0.001

* Statistically significant at less than the 10% level.

** Statistically significant at less than the 5% level.

***Statistically significant at less than the 1% level.

R-Square = 0.30 F = 8.471 (P < 0.001) N= 149. Estimated using OLS. Boom2 programs and the three mid-1980s MDAPs acquired using TPP-like contracts are omitted. With the Bonferroni correction, Wald's test for the equality of the estimated coefficients of the categorical variables for acquisition policy periods yields F= 2.18, p = 0.3724.

All of the estimated coefficients have the expected signs, and all except that for T_{bust} are statistically significant. Like the estimates for Models 1 and 2, the estimates for Model 3 imply that the Packard reforms of 1969 resulted in a decrease in PAUC growth; that decrease persisted, but subsequent changes in acquisition policy apparently did not result in further significant decreases in PAUC growth. The new result provided by Model 3 is that PAUC growth on average increases by 4.2 percentage points (the estimated coefficient of T_{boom}) for each year spent in a boom climate. Note that Model 3 assumes that the effect on PAUC growth of a boom year is the same for programs that passed MS B in boom climates as it is for those that passed in bust climates, which probably is not the case.

Model 4—Alternative Representation of Climate Effects

There is a way to overcome this limitation of Model 3. The alternative uses what are called slope categorical variables, one for boom years (T_{boom}*Climate) and one for bust years (T_{bust}*Climate). In this approach, climate effects are captured in the estimated coefficients of T_{boom}, T_{bust}, and the slope categorical variables. As is illustrated later, introduction of these variables allows the regression to pick different rates of cost accumulation for MDAPs that passed MS B in boom climates than for those that passed in bust climates. We expect that MDAPs that passed MS B in boom years accumulate less PAUC growth in both bust and boom years than MDAPs that passed MS B in bust years. The estimated coefficients for T_{boom}*Climate and T_{bust}*Climate are then expected to be negative.

Table 6 presents the estimated coefficients and p-values for this alternative. Once again, the estimates imply that the Packard reforms of 1969 resulted in a decrease in PAUC growth; that decrease persisted, but subsequent changes in acquisition policy apparently did not result in further significant decreases in PAUC growth. These estimates, however, shed considerable light on why programs that passed MS B in bust climates on average had higher cost growth than those that passed in boom climates. To see how requires being clear about the estimated rates at which programs accumulated cost growth over time.

The estimated coefficient of T_{boom} (4.8%/yr) is the rate at which a program that passed MS B in a bust climate accumulates cost in boom years. The rate for programs that passed MS B in boom climates is much lower, and in fact, negative: -0.1%/yr. This is the sum of the estimated coefficient for T_{boom} and the coefficient for T_{boom}*Climate (4.8%/yr – 4.9%/yr = -0.1%/yr). The uncertainties in the estimates are such, however, that the estimated rate could about as easily be 0.1%/yr as 0.1%/yr.[20] The point here is that programs that passed MS B in bust climates evidentially accumulate PAUC growth much more rapidly when they encounter a boom period than do programs that passed MS B in a boom climate.

---

[20] The negative estimates are not unreasonable *a priori.* Some programs that pass MS B in a boom climate may be used as "banks"—that is, relatively safe places to hold out-year claims on funding. A program used as a bank will show negative quantity normalized PAUC growth when "withdrawals" are made. It may be relevant in this regard that about one-third of the programs that passed MS B in boom climates show negative PAUC growth.

Table 6. Estimated Coefficients and p-Values for a Model That Includes the Effects of Post-MS B Funding Climate and Duration

	Coefficients	p-value
Intercept	64.2%***	< 0.001
Funding Climate		
T_{boom}	4.8%/yr***	0.006
T_{boom}* Climate	-4.9%/yr*	0.095
T_{bust}	1.2%/yr	0.232
T_{bust}* Climate	-1.8%/yr	0.230
Acquisition Policy		
DSARC	-57.0%***	< 0.001
P-C DSARC	-48.1%***	< 0.001
DAB	-60.4%***	< 0.001
AR	-82.9%***	< 0.001

* Statistically significant at less than the 10% level.

** Statistically significant at less than the 5% level.

*** Statistically significant at less than the 1% level.

R-Square = 0.29 F = 5.690 (P < 0.001) N= 149. Estimated using OLS. Boom2 programs and the three mid-1980s MDAPs acquired using TPP-like contracts omitted. With the Bonferroni correction, Wald's test for the equality of the estimated coefficients of the categorical variables for acquisition policy periods yields F = 1.95 p = 0.4948.

Programs that passed MS B in a bust climate accumulate cost in subsequent bust years at an estimated rate of 1.2 percentage points per year. This estimate is not statistically significant. The corresponding rate for programs that passed MS B in a boom climate is -0.6%/yr (= 1.2%/yr – 1.8%/yr). The estimated coefficients for T_{bust} and T_{bust}*Climate also are not statistically significant. The estimates, then, do not say much about the rate at which PAUC growth accumulates in bust years.

Cost Growth Due to Program Changes

The duration variables of Model 4 direct attention to the question of the extent to which cost growth of programs that passed MS B in bust climates is due mainly to unrealistic MS B baselines. At one extreme, most programs that pass MS B may have unrealistic MS B baselines and use entry into a boom period as a chance to "get well." At the other extreme, the tendency in bust periods may be to approve austere programs. When these programs enter a boom climate, they are expanded to acquire capabilities beyond those in their MS B baseline, which is to say that PAUC growth may be largely a matter of program changes.

Selected Acquisition Report: Global Broadcast System (DoD, 2003) provides an example of a program whose content was increased early in the post-9/11 boom. (It passed MS B early in FY 1998 and is accordingly a Bust1 program. GBS was ongoing at the end of FY 2016.) According to the report,

> The current GBS architecture is based on Asynchronous Transfer Mode (ATM) technology. … In December 2002, DoD directed GBS's migration to a more sustainable commercial and standards-based open architecture, based

upon the Internet Protocol (IP). Also, the GBS program received FY03 Iraqi Freedom Funds (IFF) supplemental funding for IP Acceleration of production units to replace deployed ATM units. Based upon extensive warfighter inputs, the accelerated IP production effort included design and development of a new, single case version of the Receive Suite (88XR) for the Army, Navy, and Marine Corps.

Space Based Infrared Satellite-High (SBIRS-High), like GBS, is a Bust1 program. It passed MS B early in FY 1997. As of the December 2015 SARs, funding for the Baseline SBIRS-High program was expected to end in FY 2018. SBIRS-High is a useful contrast to GBS. A large portion of the growth in SBIRS-High unit procurement cost for the baseline program—roughly one-third—occurred before FY 2003, while most of the other two-thirds occurred during FY 2003–FY 2009. This increase was not driven by increased capability, however, but by the unrealistic cost estimate in the MS B SBIRS-High baseline (See Porter et al., 2009; Younossi et al., 2008; Kim et al., 2015).

In the GBS example, it seems clear that capabilities beyond those in the MS B baseline were added to the program. While unit cost did increase, that was a matter of paying more for more. For SBIRS-High, in contrast, it appears that the advent of a boom funding climate provided a program experiencing severe problems an opportunity to “get well.” In effect, what otherwise would have been capability shortfalls were converted into cost growth and, relative to MS B, the DoD eventually paid more for the MS B SBIRS-High capability than had been anticipated.

As these examples indicate, the boom effect in general results from acquisition of capability beyond that in the MS B baseline and unrealistic assumptions in the MS B baseline. In examples, the PAUC growth associated with the boom climate mainly appeared in the SARs for the boom years. While we have no examples to offer, PAUC growth for Bust1 and Boom1 programs also occurs between MS B and the subsequent boom. Again, this growth can reflect either acquisition of capability beyond the MS B level or recognition that the cost of acquiring the MS B capability is higher than anticipated.

During a period of nearly 20 years starting in 1989, the Office of Program Analysis and Evaluation (PA&E), predecessor of the Office of Cost Assessment and Program Evaluation (CAPE), funded development of a database that separated cost growth due to program changes[21] from cost growth due to what PA&E called “mistakes.”[22] The data in Table 7 are drawn from the version of the PA&E database updated through the December 2002 SARs.[23]

---

[21] A major difficulty in separating program changes from Errors of Inception is ambiguity in statements of capabilities to be acquired. Those responsible for compiling the PA&E database were well aware of this problem.

[22] In about 2010, the Office of Program Assessments and Root Cause Analyses (PARCA) defined top-level proximate causes of cost growth. These included both Errors of Inception and Errors of Execution. As defined earlier by PA&E, the “mistakes” category is the sum of Errors of Inception and Errors of Execution. See McCrillis (2003).

[23] This is the database used in McNicol (2004).

Table 7. PAUC Growth Due to Errors and Program Changes

Cycle	Period (Fiscal Years)	Number of MDAPs that Passed MS B	Errors†	Program Changes‡	Total	Program Changes as a Percent of Total
Boom	1981–1986	35	4%	11%	14%#	79%
	1970–1980	42	24%	14%	38%	37%
Bust	1987–1997	46	21%	10%	31%	32%
	Combined bust	88	22%	12%	34%	35%

† Errors of Inception plus Errors of Execution.

‡ Changes made as a result of decisions to alter from the MS B baseline the capabilities the program is to acquire.

Components do not add to the total because of rounding error.

In the boom climate FY 1981–FY 1986, program changes were almost 80% of the total PAUC growth. In the bust periods, however, PAUC growth due to program changes was about one-third of the total. These data imply that the higher PAUC growth of programs that passed MS B in bust climates is primarily due to errors.

This analysis can be carried forward another step. The most interesting number in Table 7 for this purpose is the 4% for errors in the boom period FY 1981–FY 1986. This number is the sum of Errors of Inception and Errors of Execution. It is reasonable to assume, however, that Errors of Inception are on average small for programs that passed MS B in a boom period. Pushing that assumption to its limit, we have an estimate for Errors of Execution for the programs for the first boom period of 4%. Unfortunately, comparable data for the second boom period (FY 2003–FY 2009) are not available, so we have no check on how representative this estimate is; it is the only estimate we have of the average Errors of Execution for a substantial number of programs. If it is accepted as representative, the data in Table 7 imply that the average PAUC growth of MDAPs that passed MS B in bust climates due to unrealistic MS B baselines is about 17% to 20%.

Conclusion

Each of the models yielded the same pattern of results:

- MDAPs that passed MS B in a bust climate on average had significantly higher PAUC growth than those that passed MS B in a boom climate;
- The 1969 Packard reforms reduced average PAUC growth;
- The reduction persisted through the 1980s, 1990s, and 2000s;
- Changes to the acquisition process after the 1969 Packard reforms are not associated with further reductions in average PAUC growth.

Incorporation of the boom effect and program duration in the models does not provide further policy insights. These factors were significant, however, and must be considered in future analyses of MDAP cost growth.

The PA&E data on PAUC growth due to program changes suggest that the lower PAUC growth after the Packard reforms probably was due mainly to adoption of more realistic MS B baseline lines. We also find in those data an indication that cost growth baked into the MS B baselines—that is, Errors of Inception—are several times larger than Errors of Execution. That conclusion, however, amounts to less than it might seem to at first glance. The classic Error of Inception occurs when the DoD contracts for a Lincoln and budgets for a Ford. Eventually, additional funding must be added to the budget to buy the Lincoln. The DoD must make the necessary budgetary adjustments within a given top line—usually within funding for acquisitions. These adjustments include such measures as stretches, delays, cancellations, and descoping of programs. It is the cost increase imposed by these adjustments, rather than the difference between the cost of a Lincoln and a Ford, that is the relevant cost of Errors of Inception.[24]

References

DoD. (2003). *Selected acquisition report: Global broadcast system.* Defense Acquisition Management Information Retrieval (DAMIR) System.

Kim, Y., Axelband, E., Doll, A., Eisman, M., Hura, M., Keating, E. G., Libicki, M. C., & Shelton, W. (2015). Acqu*isition of space systems, Volume 7: Past problems and future challenges* (RAND MG-1171/7-OSD). Santa Monica, CA: RAND. Retrieved from http://www.rand.org /pubs/monographs/MG1171z7.html

McCrillis, J. (2003, January). *Cost growth of major defense programs*. Briefing presented at the Department of Defense Cost Analysis Symposium, Williamsburg, VA.

McNicol, D. L. (2004). *Cost growth in major weapon procurement programs* (2nd ed.). Alexandria, VA: Institute for Defense Analyses.

McNicol, D. L. (2017a). *Post-Milestone B funding climate and cost growth in major defense acquisition programs* (IDA Paper P-8091). Alexandria, VA: Institute for Defense Analyses.

McNicol, D. L. (2017b). Post-Milestone B funding climate and cost growth in major defense acquisition programs. In *Proceedings of the 14th Annual Acquisition Research Symposium* (Vol. 1, pp. 86–97). Retrieved from http://acqnotes.com/wp-content/uploads/2017/08/SYM-AM-17-034_Wednesday-Vol-1_5-1-2017.pdf

McNicol, D. L. (2018, forthcoming). *Acquisition policy, funding climate, and cost growth of major defense acquisition programs* (IDA Report R-8396). Alexandria, VA: Institute for Defense Analyses.

O'Neil, W. D., & Porter, G. H. (2011). *What to buy? The role of director of defense research and engineering (DDR&E)—Lessons from the 1970s* (IDA Paper P 4675). Alexandria, VA: Institute for Defense Analyses.

Porter, G., Gladstone, B., Gordon, C. V., Karvonides, N., Kneece, Jr., R. R., Mandelbaum, J., & O'Neil, W. D. (2009). *The major causes of cost growth in defense acquisition: Volume I—Executive summary* (IDA Paper P 4531). Alexandria, VA: Institute for Defense Analyses.

Younossi, O., Lorell, M. A., Brancato, K., Cook, C. R., Eisman, M., Fox, B., … Sollinger, J. M. (2008). *Improving the cost estimation of space systems: Past lessons and future recommendations* (MG-690-AF). Santa Monica, CA: RAND. Retrieved from http://www.rand.org/pubs/monographsMG690.html

[24] The magnitude of Errors of Inception is extremely difficult to measure. The only published attempt to do so seems to be McNicol (2004), pp. 9–10 and Appendix B. This computation recognizes some considerations in addition to stretches, which adds to the complexity.

Acknowledgments

David M. Tate (the dean of the reviewers and co-author of one of the papers in the cost growth series), David A. Sparrow, Daniel Cuda, Prashant Patel, and Philip Lurie, all of IDA, provided insightful comments on successive drafts. I am similarly grateful to Gregory Davis for his support as critic and task leader for the past two years. Brian Gladstone and Sarah Burns (co-author of one of the papers), of IDA, and Mark Husband, of Defense Acquisition University, provided valuable comments on the first three papers of the series. Linda Wu, formerly of IDA, managed data acquisition and the database. More recently, J. M. Breuer provided helpful assistance with the statistical work. The research presented in this paper was sponsored by the Director, PARCA.

ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL
555 DYER ROAD, INGERSOLL HALL
MONTEREY, CA 93943

www.acquisitionresearch.net