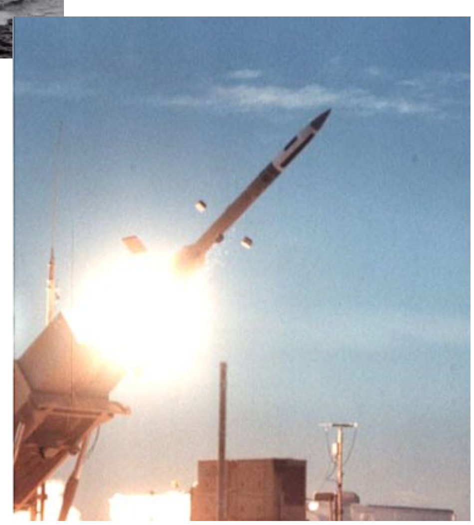
Design Build Concurrency: Cost Implications

Donald Birchler • Eric Groo • Gary Christle

CRM D0020008.A4/1REV
May 2009



CNA
ANALYSIS & SOLUTIONS

Approved for distribution: May 2009

Jino Choi
Cost Acquisition Team
Resource Analysis Division

Contents

Summary

This paper explores the relation between concurrency and cost growth in large weapon programs. Intuitively, developing or designing a weapon system while in production has potential to increase program risk, and is sometimes cited as a reason for cost growth.

Despite its appeal, solid research had yet to substantiate the hypothesis that concurrency contributes to cost growth. Our literature review discovered few articles that address this issue, and those that did were inconclusive. Thus, at the direction of the Deputy Assistant Secretary of the Navy for Management & Budget (DASN M&B), we looked at major programs across all services and ranging from ships to aircraft to see if there was any strong relation between concurrency and cost growth.

To measure this relationship, we defined concurrency as the proportion of Research Development Test and Evaluation (RDT&E) appropriations that are authorized during the same years that procurement appropriations are authorized. We calculated several metrics based on data extracted from Selected Acquisition Reports for a sample of mature acquisition programs. After adjusting for program quantity changes, and other distorting effects, we calculated cost growth in RDT&E and Procurement accounts.

Next, based on the initial program cost estimates, we calculated planned concurrency. This measurement gauged technical risk planned into each program from the outset. We repeated this calculation on the final program appropriation profiles to gauge the actual level of concurrency experienced by the program. Finally, to determine how much the program deviated from the initial plan, we subtracted the planned concurrency from the actual concurrency—we dubbed this difference “unplanned concurrency.” We used these metrics to test for relationships between cost growth and concurrency.

Our results indicated that concurrency, as defined above, does not *per se* predict cost growth—there appears to be no global relationship relating the two variables. Using classical regression techniques, we modeled the relation between the cost growth and concurrency metrics with a quadratic equation. This was intended to reflect the common notion that there could be too much or too little concurrency. We found no evidence in support of this relationship. To accommodate alternate, heretofore unarticulated relations between cost growth and concurrency we also used a smooth curving technique to "allow the data to speak for itself." These experiments also failed to discover any strongly consistent patterns in the data although in one case (planned concurrency versus procurement cost growth) it did highlight a breakpoint indicating that programs with low planned concurrency faired worse than those with high planned concurrency. In sum, we concluded that neither planned nor actual concurrency of RDT&E and procurement funding were good predictors of cost growth.

As with prior studies of the subject, our results are sensitive to the definition of concurrency, and are affected by the lack of more specific data on the timing of production relative to relevant RDT&E expenditures. We discuss these issues, and other details in the body of this report.

Introduction

Issue

Typically, Navy programs experience some level of concurrency. That is, production of the weapon system happens while some portions of the design are still being completed. Many people within the acquisition community argue that high levels of design/build concurrency ultimately lead to cost growth, as it implicitly creates a greater level of risk. For example, the 6 February 2006 memorandum “Design/Build Concurrency” from the Assistant Secretary of the Navy, Research, Development, and Acquisition (ASN-RDA) identified the high degree of concurrency in the Littoral Combat Ship (LCS) as being a large contributor to the program’s overall cost growth.

The GAO periodically publishes their “Assessment of Major Weapons Programs,” which measures a program’s level of knowledge in critical technologies, design, and production. As their report suggests, “If a program is not attaining this level of knowledge, it incurs increased risk of technical problems, accompanied by cost and schedule growth. If a program is falling short in one element, like technology maturity, it is harder to attain knowledge in succeeding elements.”[1]

To gain more insight into concurrency and how it affects program costs and schedule, we conducted interviews with several Navy and DoD acquisition officials (see Appendix B for a list of interview questions). The results of the interviews show that this simplistic view of concurrency is not commonly shared. The following discussion encapsulates the consensus view shared by most of these officials.

---

[1] GAO Report to Congressional Committees, *Assessment of Major Weapons Programs,* GAO, 2003.

The design/production sequence

Production cannot start without design, design cannot proceed without firm specifications and essential information[2] for the components needed for production, and specifications cannot exist without clear operating requirements for the end product.

In a perfect world,

- The requirements, concept of operations (CONOPs) and substantial prior development would be completed before the release of the design Request for Proposal (RFP);
- 100 per cent of the design would be complete before the release of the production RFP; and
- Material/components needed to construct the first production item would never be late because steps would have been taken to assure all material and components had already been available before production started. If these conditions existed, we would say the program has zero overlap, or concurrency, and production risk would be considered to be very low.

In this "perfect" scenario, requirements and specifications would not change once design started, design products (engineering drawings essential for production) would not change once production started, and production would flow smoothly without delays caused by late software or hardware.

Why concurrency exists

Unfortunately, this zero-risk approach to production planning is virtually impossible to achieve, and even if it were, many would argue that it is not desirable. The Japanese, for example, pioneered the "just-in-time" inventory strategy, where materials essential for production are not only unavailable before production start, they are deliberately fabricated and delivered at the last possible moment to

2. Size, weight, footprint, power requirements, computing power, etc.

reduce costly in-process inventory. And most planners conclude that there is no good reason to delay starting production in one area (for which design is complete), just because designs are not yet complete for other (unrelated) areas.

What other reasons are there to inject plans with some design or production concurrency to accelerate delivery, despite potential increases in risk of cost growth?

1. Urgent need for the product

The need for early introduction of vital operational capabilities has long been the major factor driving accelerated production of military items. For these items, the adverse consequences of not having the item available for warfighting are considered greater than the possible adverse consequences (usually cost) of starting design before all information is available, or starting production before design is complete.

2. Maintaining the Industrial Base

"Carrying the infrastructure" is another justification that acquisition decision makers use to gain approval for plans that include concurrency. They recognize there are risks involved in proceeding before all necessary design information is available, but they reason that the costs of idling design or production teams can cause a serious loss of talent (and sometimes loss of special tools or production capacity) that cannot be recovered. So they choose to start design and/or construction earlier than they might otherwise, accepting the risk of increased costs due to concurrency as an acceptable alternative to the costs of a design/production gap that would require rehiring and retraining technical personnel or restarting the production line.

In the area of submarine construction, there is considerable support for this approach, and the advantages have been quantified and affirmed by RAND in a 2007[3] report . The report summarizes the issue by stating:

3. Rand National Defense Institute, John F. Schank, et al, *Sustaining U.S. Nuclear Submarine Design Capabilities*, Rand Corporation, 2007

> The least-cost number of designers and engineers to sustain varies based on assumptions concerning the start of the new design, the duration and level of the effort, and which shipyards conduct the design. But, "doing something is always better than "doing nothing."... Dropping below 50 percent of the future peak workload seems to have the greatest impact on future cost and schedule.

3. Avoiding obsolescence

In programs requiring cutting edge technology, such as combat aircraft, missiles, or electronic countermeasures, waiting to go into production until all design and tests are completed could introduce the additional risk of obsolescence. Unless orders are placed promptly, a vendor may not be able to provide certain parts that may become unavailable after the testing was completed. One aircraft executive we interviewed said that when production is delayed, new components often have to be developed to keep up with the threat, or to substitute for items no longer in production. Ironically, in such cases, delaying production to decrease risk could actually increase risk if the delay causes a shift to a newer, less proven component.

4. Reducing exposure to requirements changes

Some believe that reducing the time from start of design to start of production reduces exposure to requirements "creep."

5. Underestimating the risks of concurrency

Concurrency in some programs has been allowed simply because risks have been underestimated. Risk management typically requires the evaluation of two factors to determine risk severity. The first factor is the probability that an "undesired event" will occur, and the second is the impact on cost, schedule, or performance if the event does occur. When planners estimate that design material or hardware for certain portions of a weapon system will be available just before its production need date, an undesired event would be that the material is unavailable at the time it is needed for production, and the estimated impact of the event would be disruption to the production sequence. Planners often consider a "worst case" estimate for the impact of such delays as day-for-day (the end product delivers one day late for every day the design material or hardware is late) unless "workaround" plans, which in themselves can be costly, are implemented. But the cost and schedule impact can be far greater than day for day. For example, some shipyards require ice-free conditions to deliver; if design material delays preclude ship delivery before navigable waters ice up, the impact of late material for just a few days or weeks could cause a ship delivery delay of as much as two or three months. Planners familiar with this type of production facility should foresee such a disproportionate impact, but other effects of design delays are less predictable. For example, contractors have in the past filed claims for "cross-contract" impact, where a delay on one contract could be shown to impact progress on another unrelated contract at the same site because the delay caused overlaps of facilities or workforce demands between the two contracts. And some design material delays impact more systems than might reasonably be foreseen, such as combat system designs that not only reveal a need for more computing power but for more physical space and cooling capacity as well, thereby causing a number of redesigns in separate areas.

Thus concurrency is sometimes recognized in advance and accepted because its probability is misjudged to be small, or the possible adverse effects are thought to be minimal. This can be a consequence of inexperience or lack of diligence on the part of program managers or planners, or perhaps a lack of acceptance of risk management effectiveness, but it is not unusual. A survey of 300 DoD and industry risk analysis professionals[4] indicated

- 27 percent of analyses perform the risk assessment separately from the cost estimate
- 38 percent of cost risk analyst analysts have received no training, either formal or informal
- 44 percent of risk ranges are intuitive judgments, without historically or guided-survey
- 26 percent of program managers do not accept risk assessment at all, not even slightly

---

4. 1998 *U.S. Aerospace Cost Risk Analysis Survey, 2000,* as reported by Capt. Vince Sipple, USAF, Maj Edward "Tony" White, USAF, and Maj Michael Greiner, USAF, in *Defense Acquisition Review Journal* January-April 2004. *Surveying Cost Growth,*

- 18 percent of unfavorable assessments are ignored, as managers stay the course.

6. Not recognizing concurrency

It is sometimes the case that some data may be necessary for design, but at the outset it is not known to be needed. In these situations, concurrency is an "unknown unknown" that does not become evident until well after the program starts, perhaps too late to resolve without program delays.

The Program Executive Officers (PEOs) and Program Managers (PMs) who we interviewed all said that there is always some concurrency in major programs—much of which is actually an integral part of the plan (see reasons 1-3 above). The challenge is to properly identify it and manage it properly.

Analyzing concurrency

In our earlier analysis,[5] we found that most prior reports on the subject defined "concurrency" as "the overlap between completion of development of an item and the start of production" of that item. With this definition, development funding that overlaps with production funding for the first authorized quantity is identified as concurrency. While this makes some sense in the abstract, it might overstate the existence of concurrency, especially in shipbuilding, because it does not discriminate between portions of the item that are still under development and portions that have completed development. Certainly ship mast production should not proceed before the configuration of the antennae for the masts are well defined, but it could proceed if the only development still ongoing for the ship was limited to a few engine room components.

Moreover, concurrency definitions based upon overlap of hardware development and production can understate program risks because they overlook other types of concurrency that can cause problems,

---

5. Memorandum for Deputy Assistant Secretary of the Navy, Management and Budget, subject: Scientific Analyst Note – Design/build Concurrency, 17 April 2007.

such as concurrency of facility construction with production (if the plant is not ready, production can't start), or concurrency of design tools development with production (introduction of computer-aided design has been known to cause many production delays of the hardware being designed).

Given these shortcomings, it is no surprise that previous studies using this definition of concurrency found no statistical relationship between concurrency and cost growth,[6] since programs having or not having concurrency may not have been not properly identified as such.

Some prior studies have used more specific definitions, such as GAO's analysis[7] of a limited number of ongoing programs. GAO defined "a non-concurrent system" as "one in which planned operational testing has occurred before the production decision," and a "highly concurrent system as one in which little or no operational testing has occurred before the production decision." This more specific definition might be more likely to capture "actual" concurrency that could affect cost growth, but we did not adopt this definition, in part, because it relies on specific schedule or milestone dates, such as Independent Operational Test and Evaluation (IOT&E) start and completion dates, that we could not obtain across all programs.

- Instead, our earlier study used two measures to quantify the amount of concurrency in a program. One used funding overlap between RDT&E and procurement funding for the first authorized production item[8] as a concurrency measure, and the other used schedule overlap. In both cases, we divided the concurrent portion by the total R&D (either

---

6. RAND PROJECT AIR FORCE, *Historical Cost Growth of Completed Weapon System Programs,* Mark V. Arena et al, RAND Corporation, 2006.
7. U.S. General Accounting Office, *Weapon Systems: Concurrency in the Acquisition Process,* Statement of Frank C. Conahan, Assistant Comptroller General, National Security and International affairs Division, before the Committee on Armed Services, United States Senate, May 17, 1990
8. Includes procurement funding for long lead and other items that might be funded before the first production article(s) is authorized.

funding amount or number of years). They result in a ratio that ranges from zero to one; zero indicating no concurrency and one indicating 100% concurrency. Zero concurrency implies that there is no overlap between RDT&E and the procurement funding streams, and 100% concurrency implies that the RDT&E years (and the associated funding streams) are wholly concurrent with the procurement years (and the associated funding streams). Both of these measures rely on data that are readily available from the Selected Acquisition Reports (SARs) contained in the Defense Acquisition Management Information Retrieval (DAMIR) system, a definite plus for any study with limited time or resources. We note however, that some programs depend upon significant items that are concurrently developed and funded for other programs. This concurrency problem can be significant for some programs and is not apparent from the SAR data. We address this subject further in the section on "Definitions."

Contribution to the research

To the best of our knowledge, our analysis is the first time that the affect of concurrency of RDT&E and production funding was examined statistically to see how it contributes to cost growth. Furthermore, we divide program cost growth into two elements; procurement cost growth and RDT&E cost growth. While programs do experience other forms of cost growth, this represents the vast majority of funds used for weapon development and production. In addition, we examined planned, actual, and unplanned concurrency (the difference between the actual and planned levels of concurrency) to see if any of these have a strong correlation to either form of cost growth. Finally, we use more than one statistical method to see if there is any correlation between cost growth and concurrency.

Methodology

Introduction

Our selection of methods and data preparation were driven by two primary questions:[9]

- Relative to cost growth, is there an ideal amount of concurrency that should be programmed for large acquisitions?
- If there is no 'ideal,' what is the relationship, if any, between cost growth and concurrency?

These questions suggested a hybrid approach, employing traditional statistics and hypothesis testing methods as well as more modern methods of data exploration. First, using an explicit global function, we used a linear regression to fit a basic function to the data. Accepting the restrictive assumption imposed by this method allowed various hypotheses to be tested based on the condition of the 'fit.' Second, we used locally weighted scatterplot smoothing (LOESS), a non-parametric regression method, to allow the data to express itself, restricting only the 'roughness' of potential functions. We assessed the 'fit' of this second approach with a bootstrapping technique[10].

For measures of cost growth and concurrency, we gathered data from Selected Acquisition Reports (SAR), published annually by all major defense acquisition programs (MDAP) and provided to Congress since the late 60's. To insure that we were using completed cost growth profiles, we sampled from mature programs, defined as programs that had begun Initial Operating Capacity (IOC), con-

[9] The rational for these questions is discussed in more detail in the "Results" section.

[10] Bradley Efron and Robert J. Tibshirani, *An Introduction to the Bootstrap*, Washington D.C.: Chapman & Hall, 1998.

tained in DAMIR. Of these, after discarding programs for which we were unable to locate initial baseline cost estimates, we were left with an initial set of 43 programs. For these complete programs, we used the procurement and RDT&E acquisition profiles to calculate cost growth and concurrency.

Cost growth is attributable to a plethora of random events and systematic changes that arise between the first estimate and final spending profile for a given program. This makes it a 'noisy' process, and is the primary reason why program cost estimation is so difficult. To facilitate making inference about concurrency, or to reduce the unrelated noise, we chose to directly control for a few known, significant influences. First, though cost growth is measured as a proportion, to control for changes in base years between SARs, we rebaselined all the reported costs in constant 2009 dollars. Second, we used a learning curve adjustment to control for the systemic affects on cost of quantity changes. This adjustment required stable associations between procurement costs and units for programs between the first and final cost profiles. This requirement, and other issues discussed below, reduced the data set to 28 programs suitable for analysis. Due to information and funding limitations we did not adjust for scope changes. This likely leads to an overstatement of cost growth for some programs, but it was unavoidable.

Definitions

Our definitions for concurrency and cost growth are as follows:

1. Concurrency is the proportion of RDT&E appropriations that are authorized during the same years that procurement appropriations are authorized. This proportion is further restricted to the first 95 percent of total RDT&E.

2. Cost growth is, after adjusting for quantity changes and inflation, the proportional increase of the final cost to the initial cost estimate.

These definitions reflect our objective of determining the relationship between concurrency and cost growth. For many programs, prior to the obligation of funds for the first lot or unit, significant monies are expended from the procurement appropriation. This advanced procurement is usually associated with a variety of activities, including the pre-purchase of materials. However, there was little reason to believe that these pre-production preparations were subject to the conflicts or synergies arising from *concurrent* design and build phases. So, to focus on the period most likely to be influenced by concurrency, we used the year of the first batch as the indicator when the program would be expected to start production, or when they started cutting metal.

The additional restriction, of only considering the first 95 percent of the total RDT&E appropriation for the concurrency calculation, was motivated by similar considerations. For most programs, RDT&E monies, after an intensive period associated with the primary development effort, continue throughout the life of the program, albeit at a much reduced rate. This is usually due to the ongoing need for updates and modifications, but has little bearing on concurrency issues. We were satisfied, after a little experimentation, that a 95 percent cutoff removed this tail for most programs. The main reason for this is that cost growth is cost growth, unlike the case of concurrency where some of the overlap between profiles should not be counted as real concurrency. This is a simplification, as in the case of procurement costs that extend for years after funds for the final lot or unit are obligated—cost growth in these dollars is unlikely to be related to concurrency.

With our definition, any development year that overlaps with any procurement authorization year assumes a concurrency of development activity with production activity. This tends to overstate the existence of concurrency because the overlapping development work could be (and frequently is) for application to later production quantities and therefore not concurrent with current production at all. In fact, many long term shipbuilding and aircraft programs have years of on-going RDT&E funding concurrent with annual production authorizations, yet most is for production quantities planned for many years in the future.

Even if RDT&E in a given year is targeted for an item authorized in that same year for production, our definition may still overstate the existence of concurrency because it does not discriminate between portions of the item that are still under development and portions that have completed development.[11]

Although our application of the SAR funding data probably overstates concurrency above, it also misses concurrency in related programs that can have a significant effect on costs and schedule of an item, such as:

- Concurrency of facility construction with production (if the plant is not ready, production can't start)
- Concurrency of design *tools* development with production (introduction of computer-aided design before it is well proven).
- Concurrency of ship production with development of items designated for the ship, but being developed under other programs (such as Radar, Sonar etc. being developed for more than one platform).

The definition for cost growth may initially appear to be overly broad, allowing for costs completely unrelated to concurrency to be included. However, adjustments for these would have been much more complex, requiring symmetric changes in both the initial estimates and final profiles tailored to each program. Out of concern that this process would become somewhat ad hoc, limited by information available on the first estimate profiles, and very costly, we concluded that it was of limited value, and left the definition broadly defined with adjustments made for quantity and inflation only (see the discussion below for more detail).

Quantity Adjustment

A dramatic change in quantity will obviously affect the total cost of an acquisition program. Consequently, it seems reasonable to measure cost growth by tracking increases in average procurement

[11] For example, ship mast production should not proceed before antenna development is complete, but it could proceed if the only development still ongoing for the ship was limited to a few engine room components.

unit costs. This approach, though appealing, suffers two significant shortcomings when dealing with large procurements:

1. Fixed costs don't adjust proportionally with unit changes. If a facility for building tanks is running at half capacity, the costs of the facility remain the same. These costs are not totally independent. For instance, if production quadruples more manufacturing facilities will need to be opened, but the relationship is certainly not linear, eluding accurate representation with an average. Similarly, the costs of RDT&E for a tank remain the same regardless of the number of final units produced. These issues become particularly significant when the number of units is small, as in the case of ship or satellite programs.

2. The marginal costs of production drop with quantity. With each additional unit, workers become more efficient, manufacturing processes are refined, cheaper suppliers are discovered, and quality control improves. This process is referred to as learning and is incorporated into every baseline cost estimate—our strategy was to reverse-engineer these estimates to accommodate different quantities.

To meaningfully discuss cost growth, the above issues, incidental to changes in quantity, have to be controlled for in the initial estimate or final program cost profile. We elected to adjust the initial estimates for procurement to reflect the total procurement quantities reported in the final SAR. The primary reason for this approach was that initial estimates are, unsurprisingly, less noisy than the actual performance of a program, and as such are more tractable for statistical analysis. Most programs experience significant deviations from the plan, so an adjustment to the final profile would have to control for overhead reburdening *and* learning. In addition, technical upgrades occur regularly in many programs, resulting in substantial deviations from planned unit costs that require statistical control. Using the initial estimate avoids most of these issues as it provides a stable frame of reference for a single, homogeneous production line. As mentioned above, RDT&E isn't affected by quantity changes, so we directly compared the initial and final RDT&E profiles, as per the definition.

We didn't directly adjust for the first issue above, or the impact of overhead reburdening on the initial estimate. Procurement is a combination of support costs and flyaway costs, further broken into recurring and non-recurring costs. Though we could have used these data to estimate and adjust for fixed costs, we were concerned about their reliability, particularly from the first reported cost estimates.[12] For this reason, our data gathering process didn't capture this information, limiting itself to total procurement costs and RDT&E. This likely had the effect of exaggerating our adjustments for quantity, biasing downward our estimates of cost growth.

Learning Curve

The learning curve is one of the most common models in cost analysis. We specified the learning curve so that the marginal cost of the Nth unit in a procurement is a power function of N, with parameters T_1 and b

$$MC(N) = T_1 N^b$$

where T_1 is a scale parameter and b is interpreted as the rate of learning. The intuition of this formulation is best understood by looking at the learning slope, or the ratio of the marginal costs of the 2N and Nth units

$$\rho = MC(2N)/MC(N) = 2^b$$

This ratio indicates that costs decline at a constant rate, determined by ρ, proportional to every doubling in quantity. For example, if learning slope was .9, the 10th unit in a production line would cost 90 percent of price paid for the 5th. Furthermore, the 20th unit in the same line would cost 90 percent of the 10th, or 81 percent of the 5th. T_1 is approximately the cost of the first unit.

[12] Hough, Paul G. "Pitfalls in Calculating Cost Growth from Selected Acquisition Reports," Rand Corporation 1992.

Lot Midpoint

A difficulty with applying this model is that it is specified in units, whereas our program cost profiles were enumerated in yearly totals, or by lots. This problem is common to learning curves estimation, as most procurement data comes in lots. The common solution is to find a typical unit, or lot midpoint within each group to use for estimating the model above.

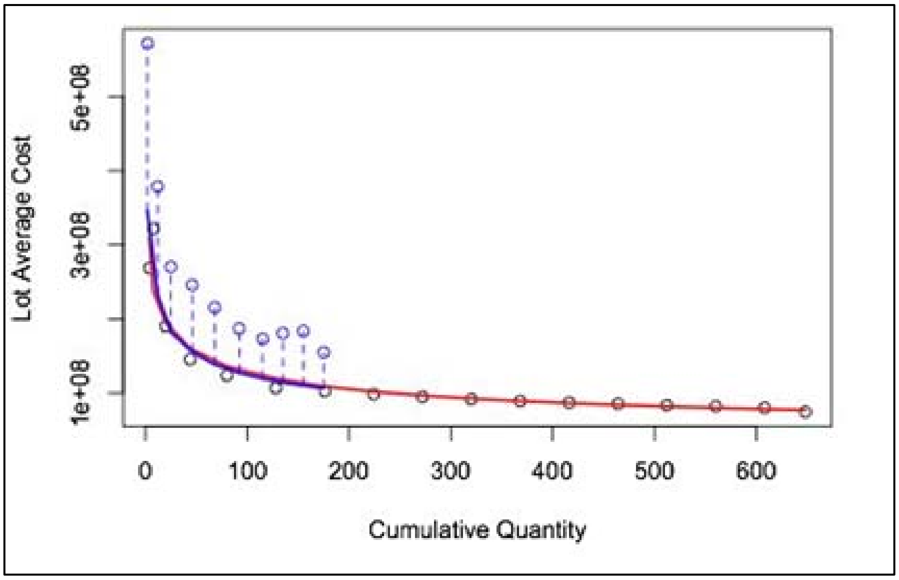
Notionally, the average lot cost, or the total cost of procurement for each year, divided by the total procurement quantity for each year, could be paired with the lot midpoint for this typical unit. As before, the problem with using an average for this value is that the learning curve is *non-linear*, so the average cost for the first half of the lot will be *higher* than the average cost for the second half. The result is that the true typical midpoint will come before the average unit in the lot. So, by picking the middle unit for the regression, the curve will be biased downward.

Complicating this approach, the lot midpoint for each year is *co-determined* by the value of b. To overcome this issue, we estimated the lot midpoint for each lot and b *simultaneously* using Non-linear least squares (NLS). The details of this estimating procedure, and the specific functional form of the lot midpoint calculation, are complicated, and are more fully explained in the appendices.

Figure 1 illustrates the learning curve adjustment for the F-22 program. The black circles correspond to the quantities and costs reported in the original program cost estimate. Notice that the dots curve sharply downward, but then flatten out as the total quantity increases. This pattern corresponds to an anticipated initial period of intensive learning, as production deals with low-hanging fruit, which progressively tapers as it becomes more difficult to optimize.

The red line is the estimated learning curve, calculated with the lot-midpoint adjustment discussed above. What is most striking about the line is how closely it appears to fit the data, without any additional modification. This was the case with the majority of programs. Exceptions included programs, such as JPATS and AIM-9X, which had an initial lot with an extreme lot average cost or, like the Bradley Upgrade, had an outlying final lot. In these cases, we simply excluded these values to achieve a better fit for the purposes of the learning curve estimate.[13] In some cases, like JASSM and FMTV, where the learning curve was untenable even excluding outliers we simply took the average unit cost for the program.

Figure 1. Learning curve adjustment illustration



The blue circles correspond to the quantities and costs reported in the final SAR profile for the F-22 program. As is usually the case, the final unit cost for the program was significantly higher, though the lot average costs decline at a rate reflective of the original estimate. The blue curve is the result of entering the successive lots in the final procurement profile into the learning curve fit to the original estimate. This curve graphically illustrates the quantity-adjusted initial estimate that we used as the basis for our cost growth calculation.

Figure 1 also illustrates cost growth. Per our definition, cost growth is the difference between the adjusted initial estimate and the final reported cost profile for a program. This is literally the area demarcated by the vertical dashed lines. Taking average unit costs to cal-

[13] This likely contributed to a mild overstatement of cost growth as we didn't filter the final profiles in this manner. This bias, however, would be small as it would affect only a limited number of units in any program.

culate cost growth, in comparison, would be the same as drawing a horizontal line.

Inflation

Inflation is a familiar concept, and is summarily described as the increasing costs of goods and services over time. The rate of inflation varies over time, and over the particular type of goods and services purchased. So, for a real comparison of costs to occur, we had to first adjust reported costs so that they reflected the same underlying values of goods and services.

Fortunately, the DoD annually updates and publishes official rates for historical and prospective inflation in the National Defense Budget Estimates, commonly referred to as "The Green Book."[14] We used the Procurement and RDT&E indices from the DoD deflators—*outlays by Title* table from the 2009 Green Book to rescale all of our data to 2009 dollars.

Looking forward, the procurement deflators table from the Green Book only extends to 2013. For several programs, production is anticipated to continue beyond this date so we extended the tables to appropriately deflate out-year costs. Fortunately, the procurement indices model out-year inflation at a fixed rate of 2 percent per year; so, for 2014 and beyond, we simply extended the table by the same rate.

Data

We gathered data from SARs based on the procurement and RDT&E profiles available on DAMIR. We reviewed this list and selected programs based on their maturity and availability of data.

[14] Quoting from the 2009 Greenbook, "DoD arrives at the figures in this book using inflation rates published by the Office of Management and Budget (OMB) as a baseline. OMB typically bases their rates on Gross Domestic Product (GDP) composite rates, accounting for non-pay factors only. DoD, however, includes pay, fuel, and medical accrual factors in its composite rates. In addition, outlay rates are factored into the final DoD inflation rates."

Regarding maturity, two factors persuaded us that programs past IOC would be very unlikely to deviate from the remaining program plan. First, passing IOC usually indicates that the major developmental and production hurdles to a system have been overcome, resulting in fewer unpredictable deviations from the plan. So, we assumed that the current estimate, extending in some cases out into the future, represented the final actual costs for the program.

On some occasions, we needed to drop new lines of production from final SARs that were absent from the first SAR. This was necessary to make apples-to-apples comparisons, particularly when controlling for quantity changes, due to the tendency of some programs to add on additional, heterogeneous lines of production to existing procurement programs.

An illustrative case is the V-22. The initial estimate for the program was essentially for a single airframe for use by the Marines. During the course of the program, the Air Force Special Forces ordered a modified version of the airframe. This new line of production, however, cost dramatically more than the Marine edition, presumably because of modifications and enhancements necessitated by the requirements for Special Forces operations. This growth in unit costs was obviously due to scope changes and not incidental to changes in program quantity or concurrency, our primary controls in this study. Fortunately, funding for the additional RDT&E and Procurement costs associated with these units was entirely funded out of Air Force appropriations, making it relatively easy to exclude it from the cost growth and concurrency calculations. For other programs with similar issues, where the distinction was less apparent, we reviewed Presidential Budget Exhibits and other publicly available budget justification materials for information to tease new subprograms away from historical program plans. This was not always possible, leading us to drop several programs from the analysis. The specific justifications for excluding each of these messy programs are contained in the appendix .

Concerns

There is one notable sampling issue highlighted by the SADARM program, the one significant outlier in our data. In selecting mature programs, we risked running afoul of sample selection bias. Basically, by filtering for maturity, we excluded programs that failed to reach maturity, potentially due to cost overruns induced by concurrency. This situation is suggested by the fact that our sample includes only one program—SADARM—that experienced a large reduction in its scope and procurement profile after its IOC. Most notably, SADARM appears to be subject to different processes than the rest of the data. If concurrency plays a significant role in the premature cancellation of programs, either through cost growth or some other mechanism, the results of our analysis would be biased in favor of programs that didn't incur any ill effects due to concurrency.

This is actually a systemic problem, as failed programs may be related to concurrency in a totally different way than programs that reach maturity. To fully examine these possibilities would require that we gather an entirely new data set, conditioned on premature cancellation.

An interesting corollary to the above, is that our conclusions about cost growth and concurrency are consequently only relevant to mature programs. Attempting to make inference about two distinct processes (i.e. cost growth versus concurrency for mature versus canceled programs) without controls in the statistical procedure necessarily blurs the results. So, instead of learning if a particular relationship exists in the data, without a good control, regression statistics would indicate that the relationship doesn't fit *all* of the data very well. By implicitly excluding a class of programs that categorically behave differently, we are consequently able to make stronger inference about the existing sample of mature programs.

The learning curve adjustment, ironically, was another potential source of bias due to the restrictions imposed on the data. This was partly alluded to in the earlier narrative about the 'Special Forces Edition' of the V-22. Essentially, to make an adjustment for learning due to quantity changes, we had to be reasonably sure that we were adjusting the same things. For example, it would be difficult to adjust expectations regarding the marginal unit cost of producing tanks at the beginning of a program using the combined number of tanks and radios produced at the end. This forced us to drop several programs that experienced quantity changes but didn't have sufficient documentation to disentangle the costs from fundamentally different units.

Another concession made to the learning curve calculation was for programs with multi-year procurement cycles. A learning curve estimate is based on per-unit costs, not costs directly, so we were required to associate all procurement costs related to quantity changes with units. This was challenging because:

1. Most programs had significant amounts of pre- and post-procurement funding. This refers to the non-aligned funding preceding and following the main block of funding aligned with quantities.

2. Multi-year delivery contracts spread funding over several years. This often had the effect of obscuring unit costs for a given year.

3. There were often gaps between years of funding with aligned quantities. This is similar to (2) in that it obscured the specific costs for each lot, but it required that we re-associate all of the funding for these years against quantities in different years.

These issues might have been partially resolved had we collected flyaway in addition to aggregate procurement costs. As we explained before, we chose not to gather these data because of concerns with the reporting in the initial estimates. Consequently, we fit learning curves to cumulative flyaway and support costs, both presumably affected in different rates by quantity changes. Our assumption is that these differences would average out, and be effectively filtered out as noise by our statistical methods. Unfortunately, the above issues precluded the use of the learning curve, so we had to manually address each of them in turn.

For all of the above, our primary criterion was average unit cost per lot. Our approach was to quickly review each initial program estimate (as we didn't adjust the final costs), cutting away the head and tail of funding that appeared to be unassociated with the first and final lots. We then refined this initial cut by plotting the average unit cost per lot on a graph and aggregating years of unaligned funding until the costs between lots were comparable. When this approach seemed inadequate, we would, as mentioned previously in regards to heterogeneous lines of production, review other public budget documentation for more information. When this failed, we would exclude the program from analysis. Details on these aggregations and exclusions are contained in the appendix.

As a final caveat, following the learning curve estimate and adjustment, we chose to exclude funds that were not aligned with quantity. In other words, when we calculated cost growth, the unaligned costs contained in the head and tail of each profile was missing. Two reasons motivated this decision; one, we had no basis for adjusting these costs for quantity changes; and two, in the initial estimates, the proportion of these costs relative to total program costs was typically very small. The effect of this exclusion was that cost growth was slightly overstated for many of the programs, as these missing costs would have brought initial estimate and final profiles 'closer' together. Regarding the estimates, this likely contributed added additional noise, as the overstatement wouldn't be distributed equally among programs—it would be directly proportional to the amount of pre- and post-procurement funding that appeared in the original cost estimate for each program, essentially a random number.

This page intentionally left blank.

Results

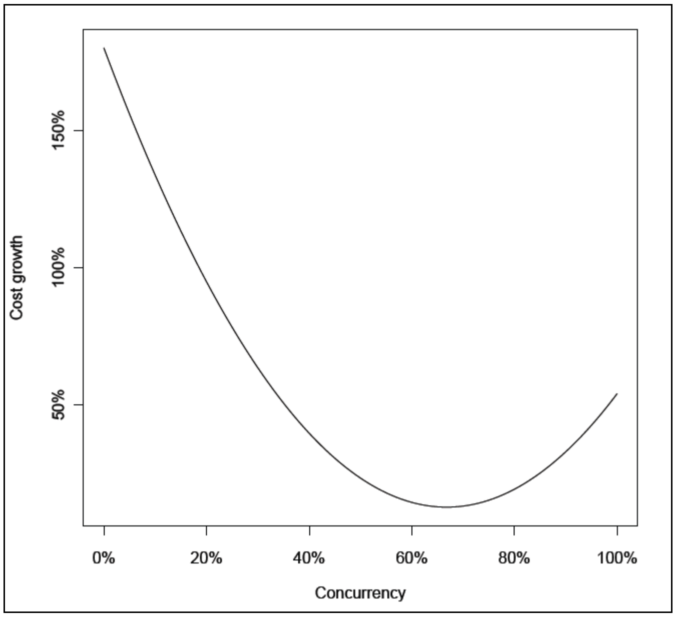
General hypothesis

For this research, we broke program cost growth into two elements, procurement cost growth and RDT&E cost growth. Although programs can experience cost growth in other areas (e.g. military construction) these two elements represent the vast majority of program spending and are the most common areas in which cost growth occurs.

For each element of cost growth, we wanted to see if there was any correlation to concurrency, as measured by the percent of RDT&E spending that occurs when procurement spending is happening at the same time. As mentioned in the methodology section of this report, we calculated concurrency in three ways. First, using the first published SAR, we determined what the planned concurrency was. Similarly, looking at the last SAR we calculated what the actual concurrency was. Subtracting planned from actual concurrency gave us our measure of unplanned concurrency. Thus, for each element of cost growth, we looked for correlations with three different measures of concurrency.

Based upon the feedback that we received from various Navy and DoD acquisition officials, we decided that a good starting hypothesis is that, for all three measures of concurrency, too little is bad for a program as serial design and production yields a longer duration (and thus more cost) before fielding of the weapon. Too much concurrency is also bad as it accepts too much technical risk. Thus, some moderate level of concurrency would be the optimal in the sense that it minimizes cost growth. This Goldilocks rule (not too cold, not too hot, but somewhere in the middle) might yield a curve similar to that shown in figure 2.

Figure 2. Hypothetical quadratic relation between cost growth and concurrency



The logic behind this approach for planned concurrency is relatively simple. Program managers plan for a certain level of funding concurrency. If they plan for too much, they may accept too much risk that could yield cost growth. In short, they have planned for failure. On the other hand, too little funding concurrency forces them to create completely serial development/design and production processes which prolongs program duration and also creates cost growth. In sum, it is the planned level of concurrency itself that forces managers to make decisions that ultimately lead to cost growth if either too much or too little concurrency is accepted.

The logic for the actual concurrency follows along similar lines. Program managers may or may not have planned for concurrency but events led to the situation where some level of concurrency occurred which, if too high or too low, led to excessive cost growth. Again, the assumption is that some intermediate level of actual concurrency would be the optimum.

The logic for the difference between planned and actual concurrency, i.e. unplanned concurrency, is less clear. One could argue that any movement in actual concurrency towards some middle level would be associated with less cost growth while movements to more extreme actual concurrency would yield more cost growth. This argument does not lend itself to a U-shaped curve.

On the other hand, one could argue that decreasing actual concurrency (i.e. negative unplanned concurrency) implies unanticipated serial design/build processes. Conversely, steep increases in concurrency (or positive unplanned cost growth) may imply excessive unplanned risk which, in turn led to cost growth. This second argument would follow a U-shaped curve with the global minimum being close to zero unplanned concurrency (i.e. what actually happened in terms of concurrency was approximately what was planned for). For the sake of this paper, we adopt this argument as the most likely outcome given our interview responses (see earlier discussion).

In all cases, this simple rule can be represented by a quadratic function. Thus, for both forms of cost growth and all three concurrencies, we will estimate the following function using Ordinary Least Squares (OLS):

$$(1)\quad CostGrowth = \beta_0 + \beta_1 Concurrency + \beta_2 Concurrency^2 + \varepsilon$$

However, there is a possibility in each case that our hypothesis will be rejected. That is, the model shown above does not accurately explain the variability in cost growth. One could argue that the relation between concurrency and cost growth is not as simple as quadratic formula implies and that some other higher order polynomial more accurately reflects the relationship between the two. Thus, for those cases where the quadratic model is not a strong fit, we use a curve smoothing algorithm to look for a better fit. We then bootstrap these results to see how well conditioned the data are.

Procurement cost growth

Planned concurrency

Our first model explored the relation between planned concurrency and procurement cost growth. Using OLS, we estimated the following quadratic equation:

(2) $procCG_i = \beta_0 + \beta_1 PC + \beta_2 PC^2 + \varepsilon_t$

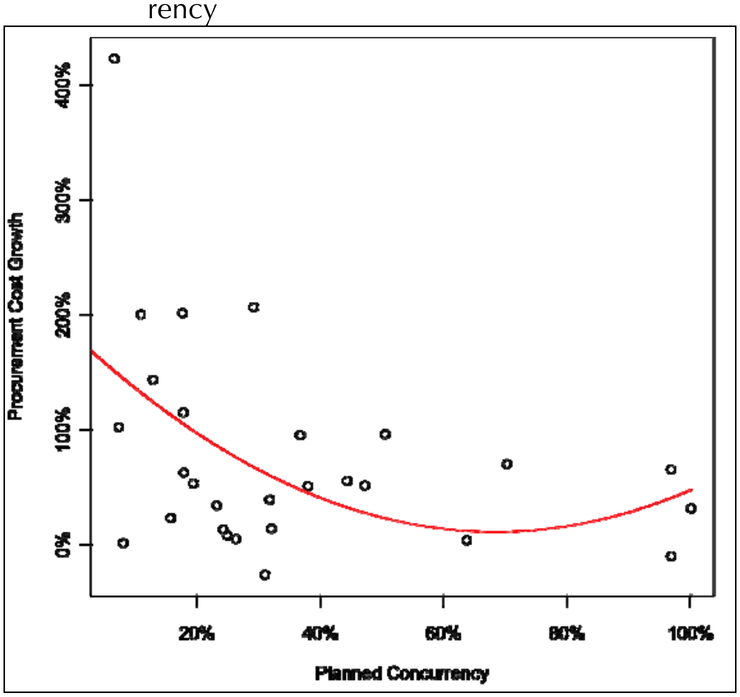
where $procCG_i$ is the cost growth in the procurement accounts for program i, PC_i is planned concurrency, and ε_i is a normally distributed disturbance term.

Table 1. OLS results—Procurement cost growth vs. planned concurrency

	Estimate	Std Error	P value
Intercept	1.825	0.484	0.001
Concurrency	-5.014	0.456	0.052
Concurrency squared	3.667	2.273	0.119
Adjusted R squared		*0.137*	

From Table 1 we see that while two of the parameters are statistically significant at the .10 level and the fitted line does give us a U-shaped curve (see figure 3), the adjusted R-squared is very low, which forces us to conclude that the quadratic model has little prediction powers of procurement cost growth.

Figure 3. Fitted curve—Procurement cost growth vs. planned concurrency

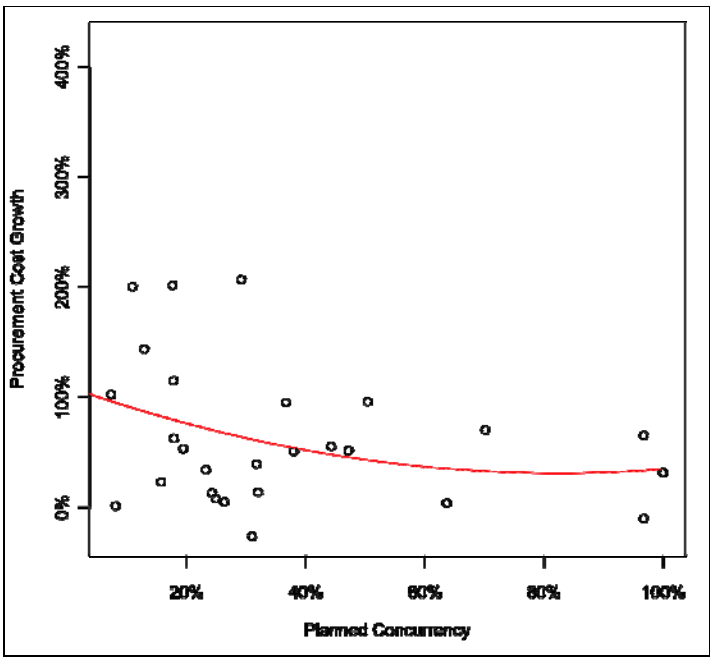


Note that much of the curvature in the model comes from one outlier. To see how well the model improves without this data-point, we ran the same model excluding the outlier. This resulted in no appreciable improvement to the model at all and slightly less curvature.

Table 2. OLS results—Procurement cost growth vs. planned concurrency (outlier excluded)

	Estimate	Std Error	P value
Intercept	1.097	0.390	0.010
Concurrency	-1.907	1.934	0.334
Concurrency squared	1.157	1.764	0.518
Adjusted R squared		*0.021*	

Figure 4. Fitted curve—Procurement cost growth vs. planned concurrency (outlier excluded)



Finally, in order to see if some other possible polynomial relation was evident, we ran the LOESS smooth curving routine on all of the data including the outlier. We then bootstrapped the 90 percent inter-quartile range to see how well conditioned the data are to the original curve. If the data are well conditioned, the smoothed curves generated by the repeated sampling should be very similar to the original and the confidence intervals defined should be fairly tight around the original curve.[15] The results of these exercises can be seen in the figures below.

Figure 5. LOESS curve smoothing - Procurement cost growth vs. planned concurrency

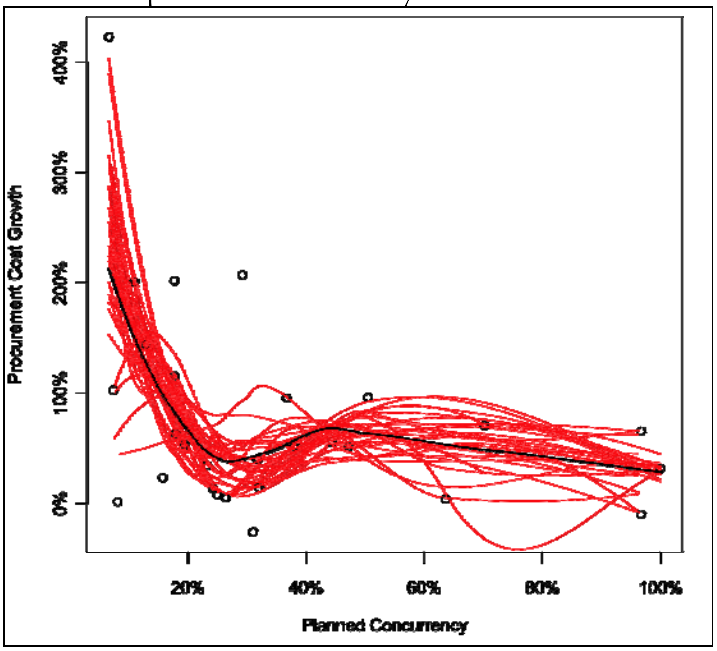


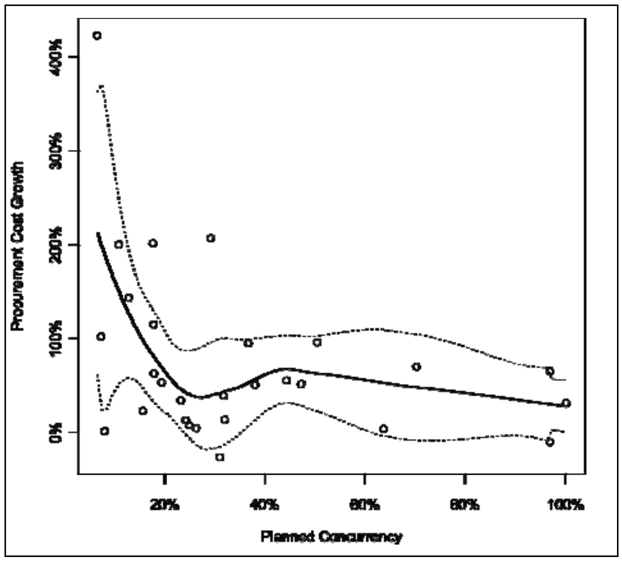
Figure 5 shows the results of 30 bootstrap replications. The black line is the smoothed curve produced by the original data. The red lines are the results of repeating the same exercise on the re-sampled data. Note that in the original curve, there are two inflection points, indicating the possibility that a third degree polynomial would be a better fit than our quadratic model. However, the bootstrap results do not consistently display this same overall shape, indicating that the data only weakly conform to a fourth degree polynomial.

To assess the fit of the curve, we replicated this experiment 2000 times to approximate the empirical distribution of the estimator. As we can see from figure 6, below, the interval is extremely wide. For

[15] The LOESS bootstrap method is non-parametric implying that we make no assumptions about the structure of the error term. However, we measure tightness by creating an interval around the original curve that includes 90% of the bootstrapped curves. These curves approximate the 5th and 95th percentiles of the true underlying distribution.

example, if a program had planned concurrency of .2, then, within the 90 percent inter-quantile range, the procurement cost growth for that program could easily range from 50 percent to over 100 percent.

Figure 6. Confidence interval - Procurement cost growth vs. planned concurrency



To ensure that the outlier was not a significant factor in these results, we ran the same experiment excluding this data-point. This did not improve the results in any discernable way (see figures 7 and 8 below).

Figure 7. LOESS curve smoothing - Procurement cost growth vs. planned concurrency (outlier excluded)

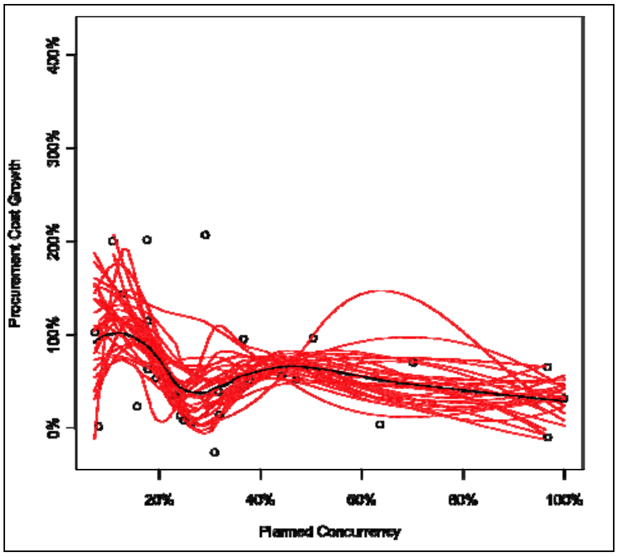
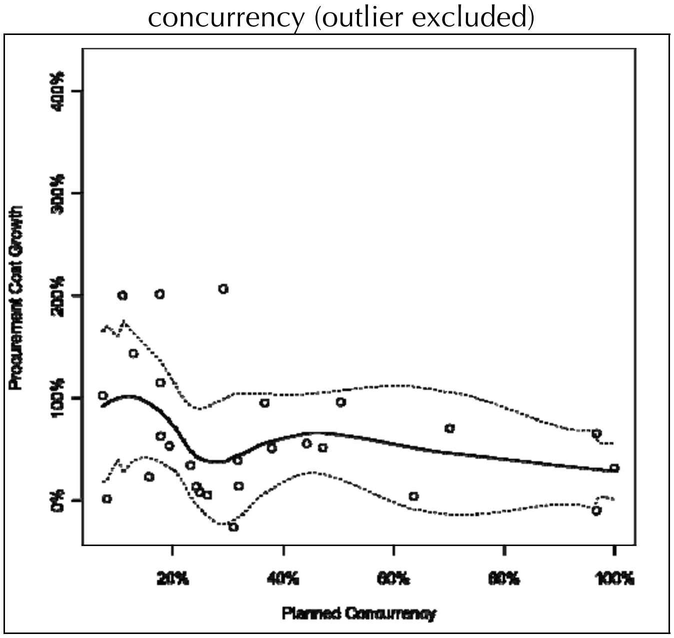


Figure 8. Confidence interval - Procurement cost growth vs. planned concurrency (outlier excluded)



In spite of the fact that the data only weakly indicate a third degree polynomial, we do see a pattern in the data that suggests that low levels of planned concurrency is more problematic than higher levels of concurrency. Again turning to the data without the outlier, we calculated the mean cost growth in procurement for those program with concurrency levels under 30 percent and compared it to the means for those programs with concurrency over 30 percent. Those under 30 percent experienced, on average, approximately 110 percent cost growth while those over 30 percent experienced an average cost growth of approximately 50 percent. This difference was statistically different at the 95 percent confidence level.

In sum, we adopted several statistical methods to see if there was any strong relation between planned concurrency and procurement cost growth. We specifically rejected the notion that planned concurrency has a quadratic relation to procurement cost growth. Furthermore, we found no other polynomial relationship that was strongly consistent with the data. Thus, our conclusion is that planned concurrency of RDT&E and production funding is a poor predictor of procurement cost growth. However, the LOESS technique did highlight a breakpoint of about 30 percent concurrency where higher concurrent programs experienced less procurement cost growth than those programs with lower concurrency. That is, lower concurrency is more problematic than high concurrency. This phenomena was not apparent in any of the other relationships that we examined.

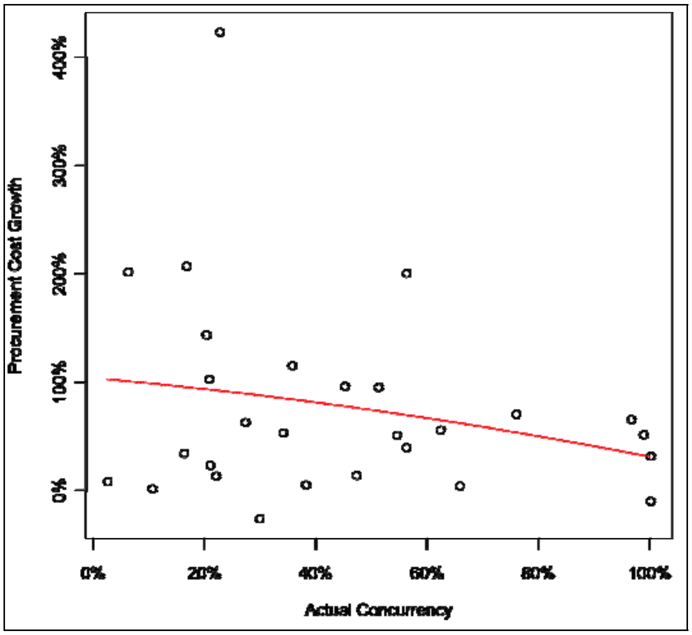
Actual concurrency

We next turn our attention to procurement growth as a function of actual concurrency. Table 3, below, shows the results of estimating the quadratic model using OLS. As in the case with planned concurrency, only the intercept parameter, β_0, is significant at the .01 level. The model as a whole has an adjusted R-squared of -0.01889 indicating that the model has little explanatory power. Note also that the fitted line is concave which is the exact opposite of what our hypothesis was (i.e. a U-shaped curve).

Table 3. OLS results—Procurement cost growth vs. actual concurrency

	Estimate	Std Error	P value
Intercept	1.037	0.530	0.062
Concurrency	-0.453	2.399	0.852
Concurrency squared	-0.275	2.168	0.900
Adjusted R squared		*-0.019*	

Figure 9. Fitted curve—Procurement cost growth vs. actual concurrency

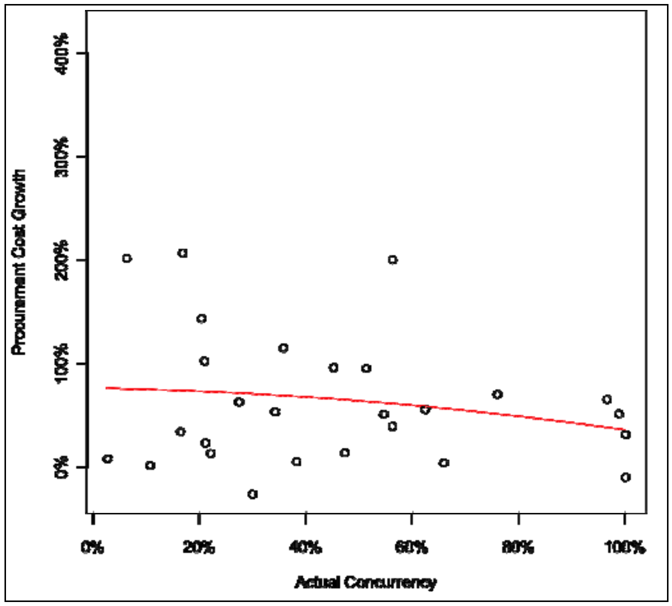


We also note the existence of an outlier that could exhibit a fairly large effect on the model. To account for this possibility, we ran the same OLS model again without this outlier. The results are show below.

Table 4. OLS results—Procurement cost growth vs. actual concurrency (outlier excluded)

	Estimate	Std Error	P value
Intercept	0.762	0.376	0.054
Concurrency	-0.086	1.684	0.960
Concurrency squared	-0.319	1.521	0.836
Adjusted R squared	*-0.040*		

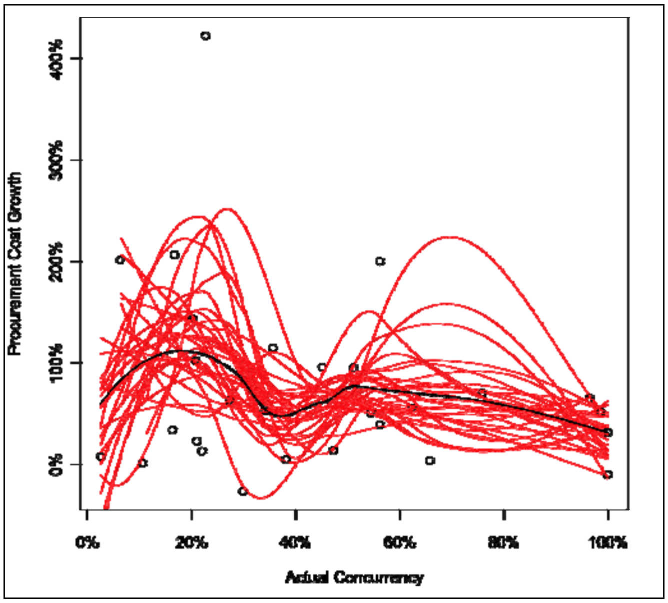
Figure 10. Fitted curve—Procurement cost growth vs. actual concurrency (outlier excluded)



Using these data, the model still performed poorly with only the intercept being significant at the .10 level. Further, the fitted line was still concave.

Using the LOESS smooth curving method, we examined the data to see if other polynomial functions could possibly explain the data better than a simple quadratic function.

Figure 11. LOESS curve smoothing - Procurement cost growth vs. actual concurrency



The original smoothed curve using all of the data is the black line (see figure 11 above). The red lines are the results of repeating the same exercise 30 times on the re-sampled data. Note that there are three inflection points in the original estimated line indicating the possibility that a fourth degree polynomial would be a better fit than our quadratic model. However, the bootstrap results do not consistently exhibit this same overall shape indicating that the data do not strongly conform to a fourth degree polynomial.

We repeated this exercise 3000 times to create a confidence interval around the original curve. As in the case for planned concurrency, the confidence interval is very wide, indicating that actual concurrency is also a poor predictor of procurement cost growth.

To ensure that the outlier was not a significant factor in these results, we ran the same experiment excluding this data-point. This did not improve the results in any discernable way (see figures below).

Figure 12. Confidence interval - Procurement cost growth vs. actual concurrency

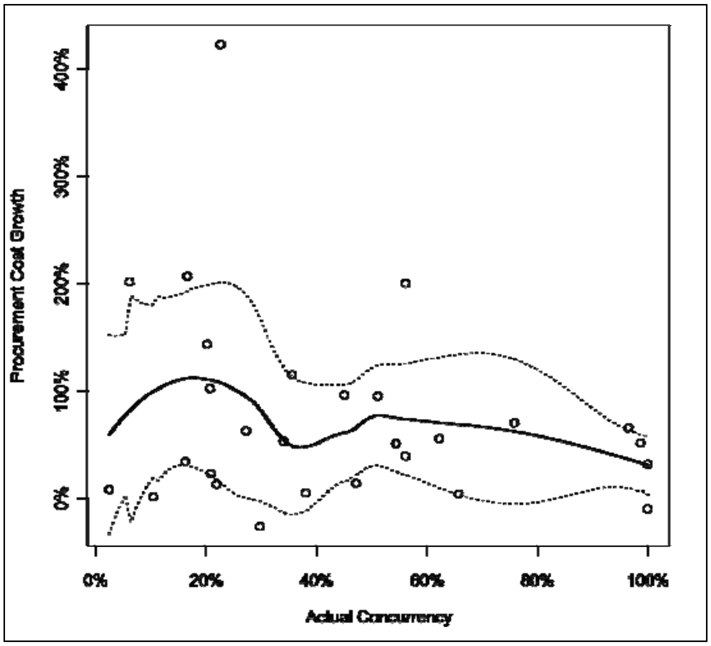


Figure 13. LOESS curve smoothing - Procurement cost growth vs. actual concurrency (outlier excluded)

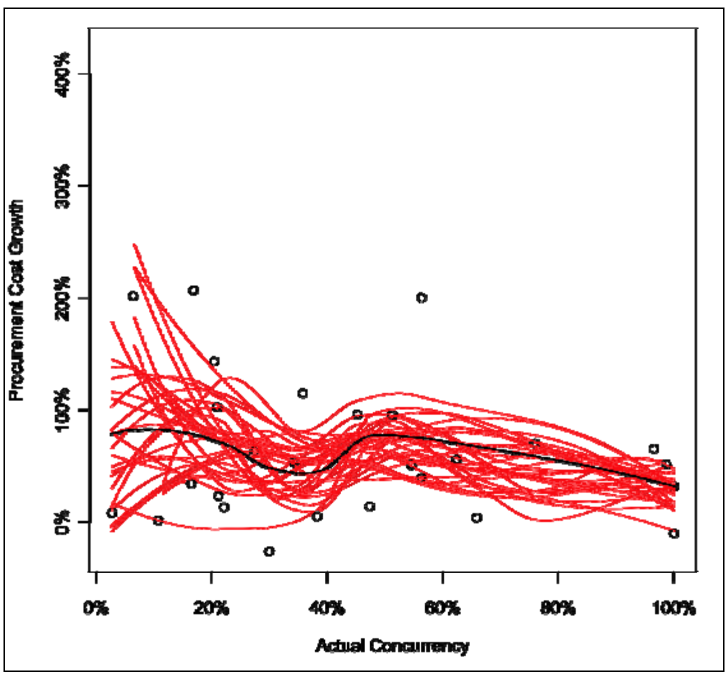
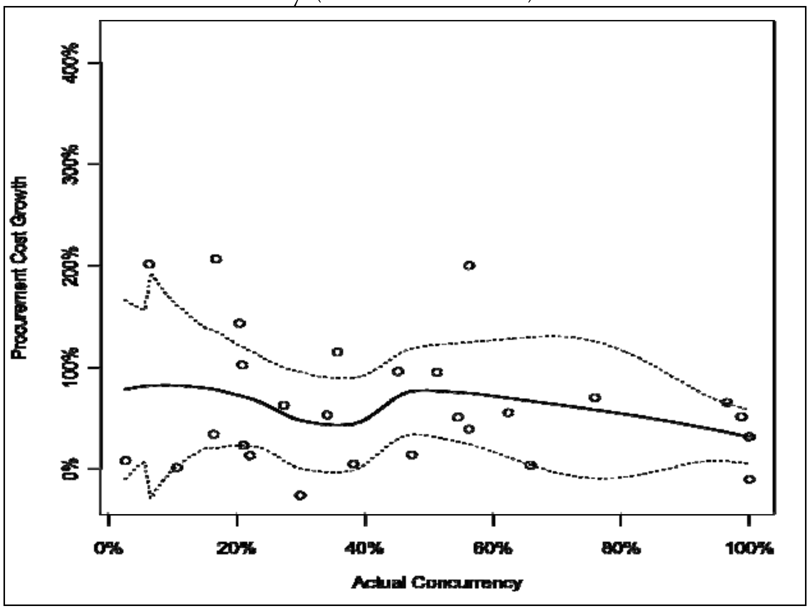


Figure 14. Confidence interval - Procurement cost growth vs. actual concurrency (outlier excluded)



Again, we used several statistical methods to discover any relation between actual concurrency and procurement cost growth. We specifically reject the notion that actual concurrency has a quadratic relation to procurement cost growth and find no other polynomial relationship that was consistent with the data. Thus, our conclusion is that actual concurrency of RDT&E and production funding is not a strong predictor of procurement cost growth either.

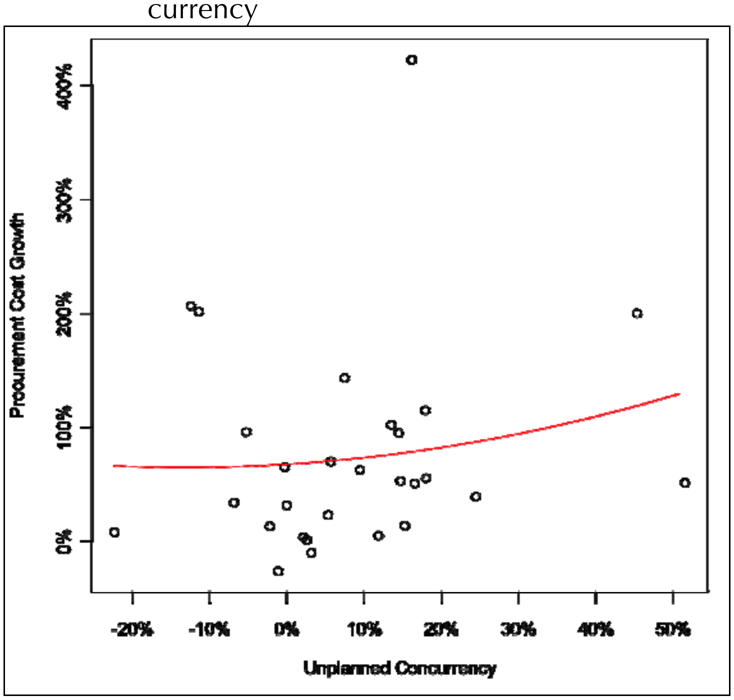
Unplanned concurrency

Finally, we examine the relation between procurement growth and unplanned concurrency. The results of estimating the quadratic model using OLS are shown below. As in the case with planned concurrency, only the intercept parameter, β_0, is significant at the 0.05 level. The model as a whole has an adjusted R-squared of -0.051, indicating that the model has little explanatory power. Note also that the fitted line is only slightly convex.

Table 5. OLS results—Procurement cost growth vs. unplanned concurrency

	Estimate	Std Error	P value
Intercept	0.678	0.207	0.003
Concurrency	0.417	1.804	0.819
Concurrency squared	1.576	4.663	0.738
Adjusted R squared		*-0.051*	

Figure 15. Fitted curve—Procurement cost growth vs. unplanned concurrency

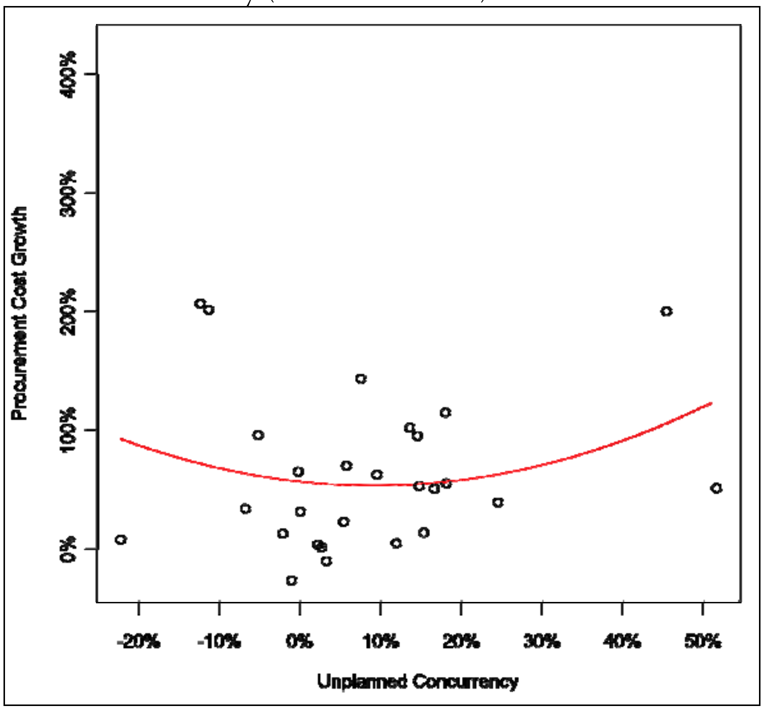


We again have an outlier that, in theory, could exhibit a fairly large effect on the model. To account for this possibility, we ran the same OLS model again without this outlier. The results are shown below.

Table 6. OLS results—Procurement cost growth vs. unplanned concurrency (outlier excluded)

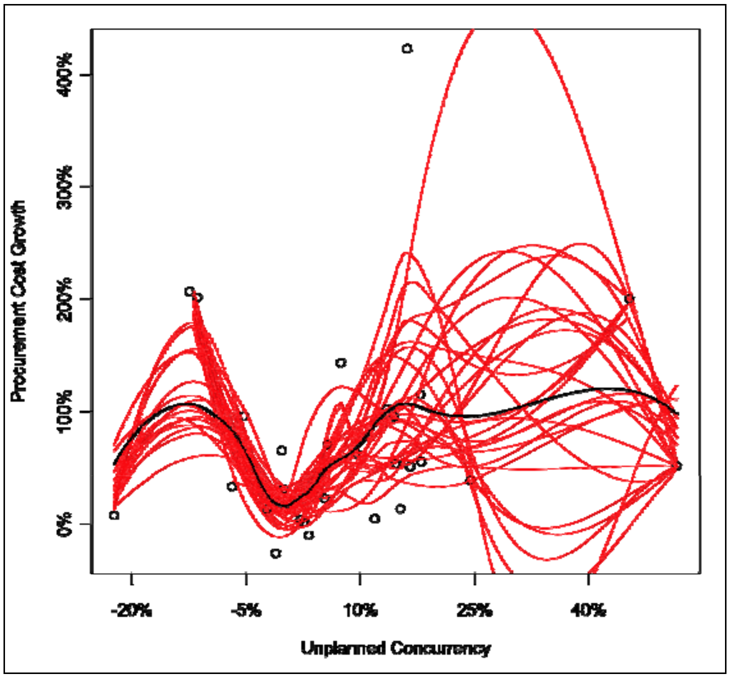
	Estimate	Std Error	P value
Intercept	0.569	0.142	0.001
Concurrency	-0.728	1.241	0.563
Concurrency squared	3.993	3.193	0.223
Adjusted R squared		*-0.004*	

Figure 16. Fitted curve—Procurement cost growth vs. unplanned concurrency (outlier excluded)



Using these data, the model still performed poorly with only the intercept being significant at the .05 level although the fitted curve generally showed more convexity.

Figure 17. LOESS curve smoothing - Procurement cost growth vs. unplanned concurrency

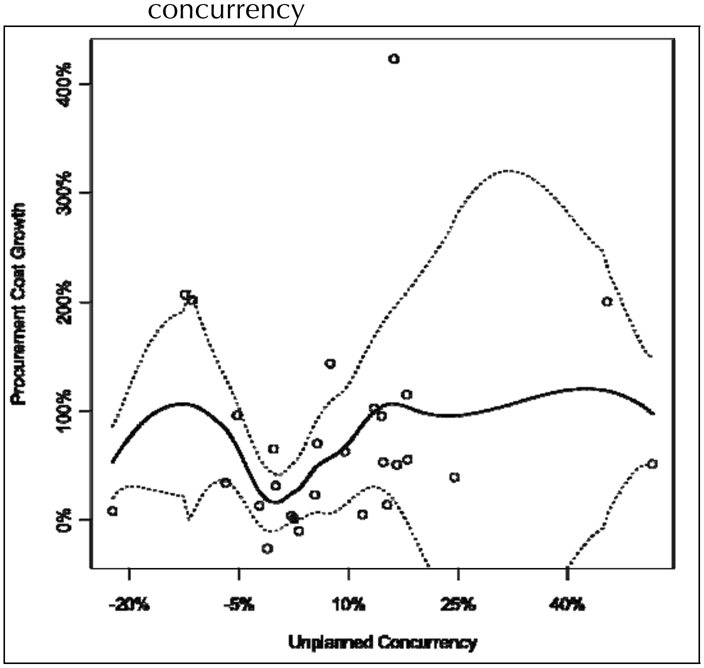


Using the LOESS smooth curving method, we examined the data to see if other polynomial functions could possibly explain them data better than a simple quadratic function.

The original smoothed curve using all of the data is the black line (see figure 17 above). The red lines are the results of repeating the same exercise 30 times on the re-sampled data. Note that for unplanned concurrency between -0.2 and 0.1, there is a modestly consistent U-shape with a local minimum at zero. However, for those values greater than 0.1, the data do not support any strong relation.

This can be seen even more clearly in figure 18 where the variation around zero unplanned concurrency is tighter than it is further from zero. This is possible evidence that the real concurrency problem is not how much concurrency is planned or realized, but how far away from the original plan a program deviates. This result is very local and the variance gets very large the further away we get from zero unplanned concurrency.

Figure 18. Confidence interval - Procurement cost growth vs. actual concurrency



To ensure that the outlier was not a significant factor in these results, we ran the same experiment excluding this data-point. This did not improve the results in any discernable way (see figures 19 and 20 below).

Figure 19. LOESS curve smoothing - Procurement cost growth vs. planned concurrency (outlier excluded)

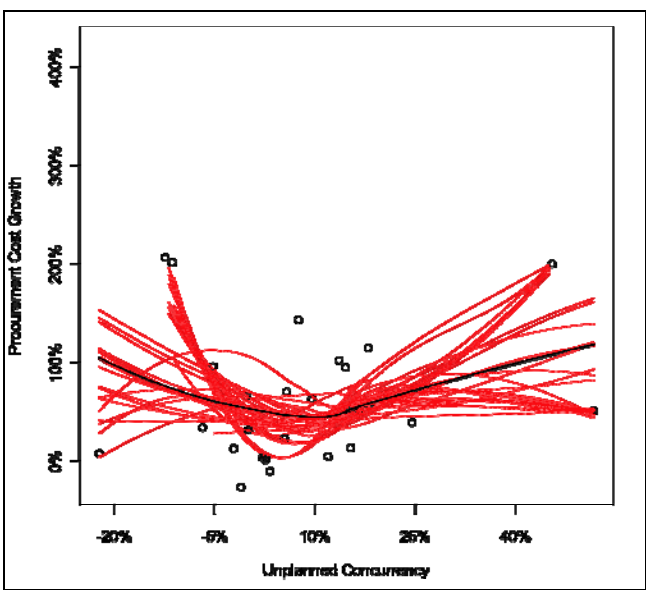
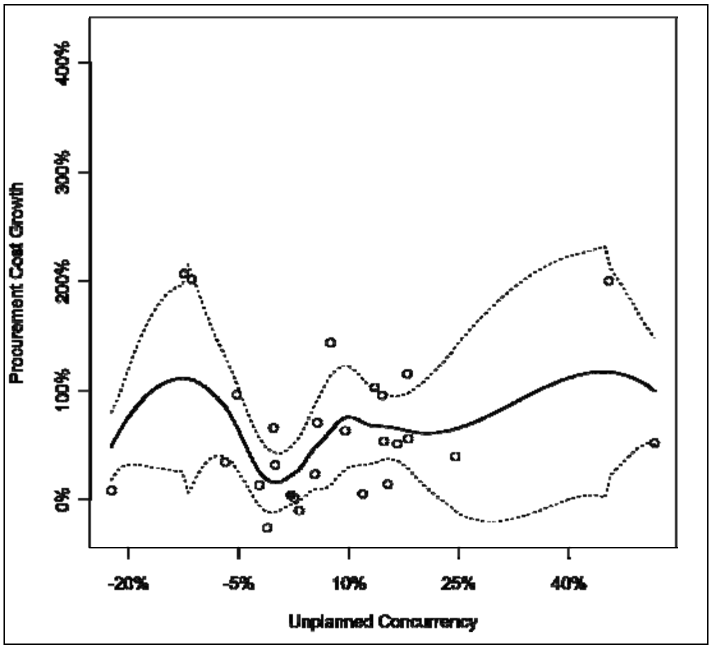


Figure 20. Confidence interval - Procurement cost growth vs. planned concurrency (outlier excluded)



Again, we used several statistical methods to explore any strong relation between actual concurrency and procurement cost growth. We specifically reject the notion that unplanned concurrency has a quadratic relation to procurement cost growth and find no other polynomial relationship that was consistent with the data. Thus, our conclusion is that unplanned concurrency of RDT&E and production funding is not a strong predictor of procurement cost growth either.

Conclusions for procurement cost growth

In all cases, we reject the hypothesis that procurement cost growth is related to any measure of concurrency in a way described by a quadratic function. The idea that too little or too much unplanned concurrency is specifically rejected.

We also found no other polynomial relation that accurately supports the data. While using the LOESS curve smoothing routine on all forms of concurrency did lead to an initial curve, suggesting some other possible polynomial relation, further re-sampling of the data created curves that were far different. However, there is slight evidence that small (-5 to 10 percent) deviations from planned concurrency can contribute to cost growth. As programs deviate significantly from planned concurrency the data do not support any specific relationship.

In addition, the 90 percent interquantile range from our bootstrap experiments were extremely large. Thus, even if we accepted the implied curvature, the predictive power of the model for any of the concurrency measures was extremely low. In sum, we found that there is little if any explanatory power of concurrency on procurement cost growth.

RDT&E cost growth

Planned concurrency

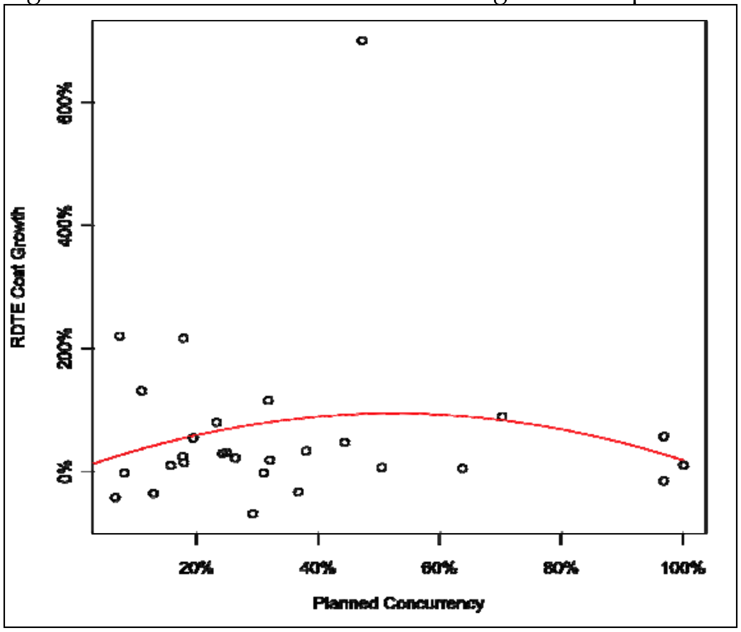
We now turn our attention to cost growth in a program's RDT&E budget. As in the case with procurement cost growth, we begin by exploring the possibility of a quadratic relation between this form of cost growth and planned concurrency. The OLS results for the quadratic equation are shown below.

Table 7. OLS results—RDT&E cost growth vs. planned concurrency

	Estimate	Std Error	P value
Intercept	0.024	0.818	0.977
Concurrency	3.507	4.155	0.407
Concurrency squared	-3.352	3.844	0.391
Adjusted R squared		*-0.048*	

As table 7 shows, none of the parameters for the quadratic model are statistically significant and the adjusted R-squared is negative indicating a very poor fit. In addition, the fitted line is concave which is the opposite of what our hypothesis states.

Figure 21. Fitted curve—RDT&E cost growth vs. planned concurrency.



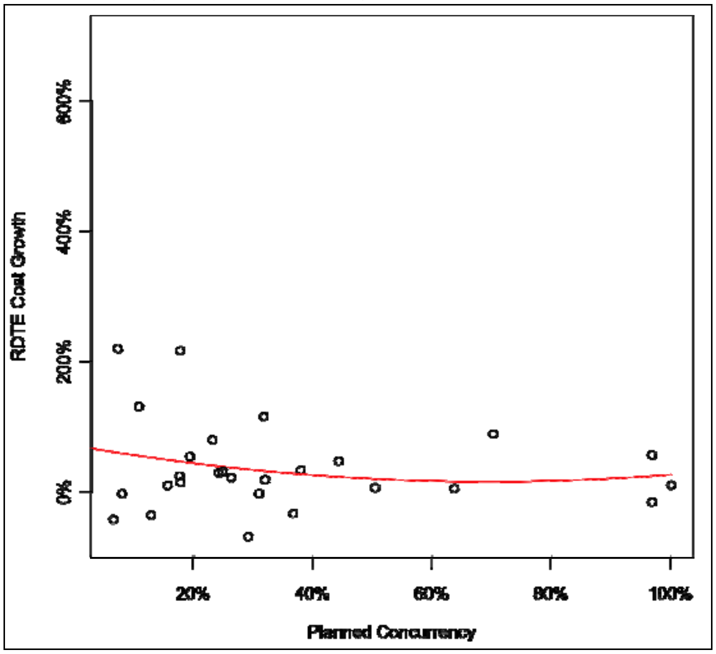
Note, however, that much of the curvature in this model comes from the existence of an outlier whose cost growth is greater than 700 percent (see figure 21 above). To see how the model behaves in the absence of the outlier, we ran the regression again without this data point.

Table 8. OLS results—RDT&E cost growth vs. planned concurrency (outlier excluded)

	Estimate	Std Error	P value
Intercept	0.713	0.403	0.090
Concurrency	-1.599	2.088	0.451
Concurrency squared	1.150	1.926	0.556
Adjusted R squared		*-0.042*	

The model performed equally poor with the exclusion of the outlier though the fitted curve did become convex as our hypothesis would suggest. Even so, the adjusted R-squared is still very low, indicating that the model as a whole still performs very poorly.

Figure 22. Fitted curve—RDT&E cost growth vs. planned concurrency (outlier excluded)



Clearly the quadratic model does not support any relation between planned concurrency and RDT&E cost growth. As such, we again turn to the LOESS smooth curving technique to explore other possible polynomial relations.

Figure 23 shows the results of 30 bootstrap replications. Note that the original line (drawn in black) has two local inflection points indicating the possibility of a third degree polynomial. However, the 30 bootstrapped lines do not share the same curvature. Thus, the data are not well conditioned to this polynomial relation.

Figure 23. LOESS curve smoothing—RDT&E cost growth vs. planned concurrency

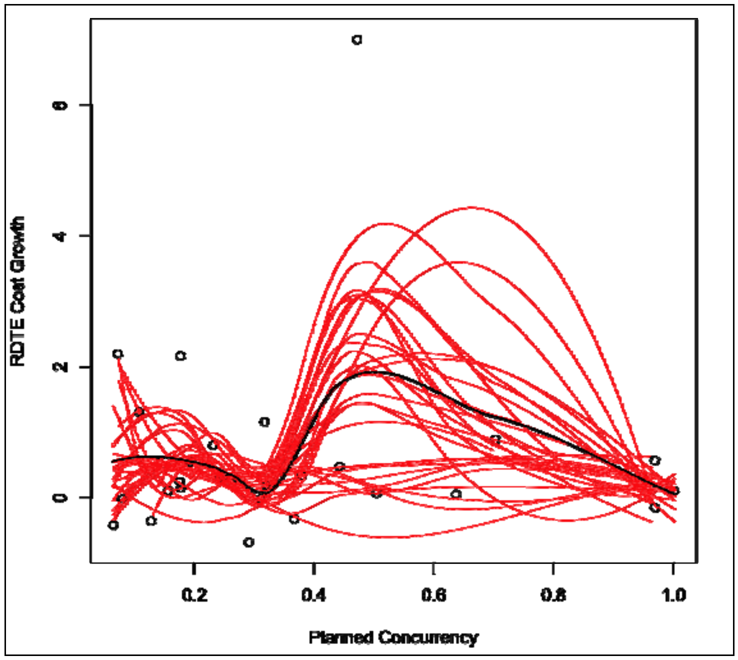
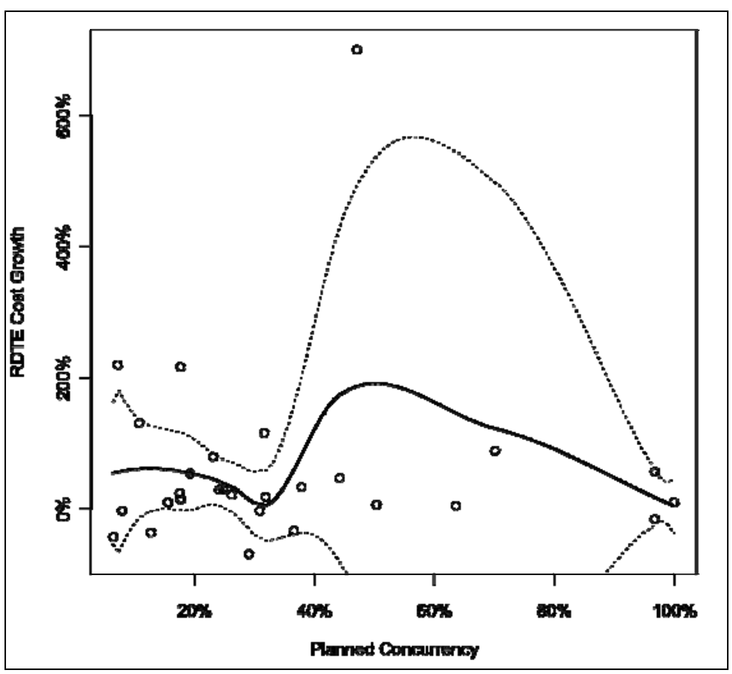


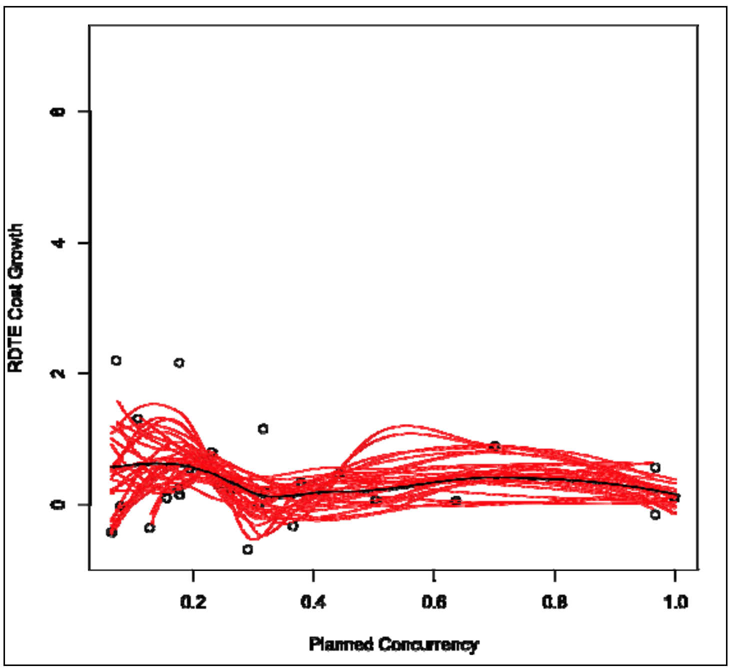
Figure 24. Confidence interval—RDT&E cost growth vs. planned concurrency



Furthermore, the 90 percent inter-quantile range is very wide, especially when planned concurrency exceeds 50 percent.[16] For example, the narrowest portion of the interval is when planned concurrency is about 30 percent. However, at that level of planned concurrency, the level of RDT&E cost growth ranges from 0 to 70 percent.

We again note the presence of the same outlier which clearly affects the LOESS results when planned concurrency is greater than 0.4. To see how the data behave without this outlier, we ran the bootstrap experiment again without this data point.

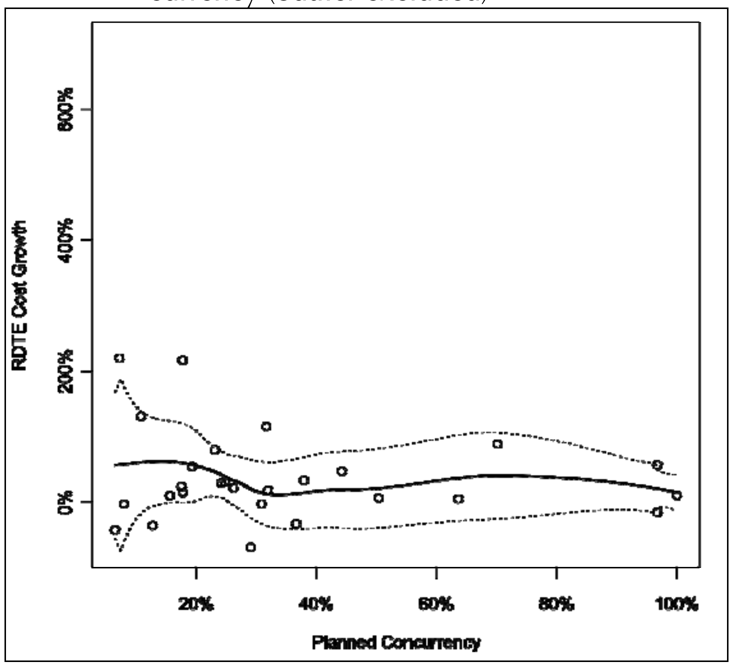
Figure 25. Confidence interval—RDT&E cost growth vs. planned concurrency (outlier excluded)



Looking at the 30 bootstrap curves versus the original line, we do see some improvement in the stability of the curvature. Nonetheless, the data are still not well behaved and many of the curves generated had curvatures vastly different than the original.

[16] The same caveat about the bootstrapped inter-quantile range mentioned above apply here as well.

Figure 26. Confidence interval—RDT&E cost growth vs. planned concurrency (outlier excluded)



The two standard deviation confidence interval also improved somewhat, especially at planned concurrency levels greater than 0.4. Even so, the confidence interval is still very wide and has little predictive power.

In sum, we reject the hypothesis that planned concurrency has a quadratic relation to RDT&E cost growth. Furthermore, there is no polynomial relationship that explains the data in any strong way. We conclude that planned concurrency of RDT&E and production funding has little or no explanatory power for RDT&E cost growth.

Actual concurrency

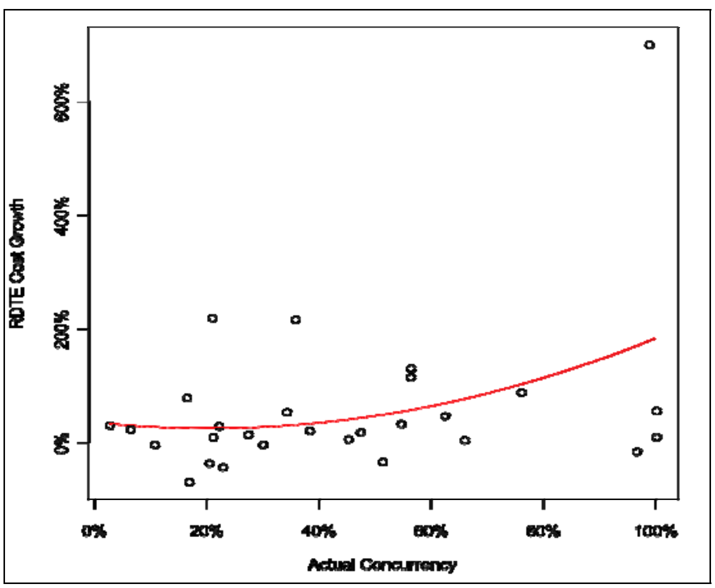
We next explore the relation between actual concurrency and RDT&E cost growth. We again begin by examining the possibility of a quadratic relation between this form of cost growth and actual concurrency. The OLS results for the quadratic equation are shown below.

As shown in table 9, none of the parameters for the quadratic model are significant and the adjusted R-squared is very low, indicating a very poor fit. The fitted line does exhibit the convexity that our hypothesis suggests, but we note that the outlier may have a large influence on the model. In addition, this was first model that exhibited some form of heteroskedasticity as measured by the Breusch-Pagan statistic. This typically implies that the reported standard errors are too small, leading to an overestimate of the statistical significance of the parameters.

Table 9. OLS results—RDT&E cost growth vs. actual concurrency

	Estimate	Std Error	P value
Intercept	0.364	0.778	0.644
Concurrency	-1.015	3.518	0.775
Concurrency squared	2.506	3.180	0.438
Adjusted R squared		*0.070*	
Breusch-Pagan		*8.264*	

Figure 27. Fitted curve—RDT&E cost growth vs. actual concurrency



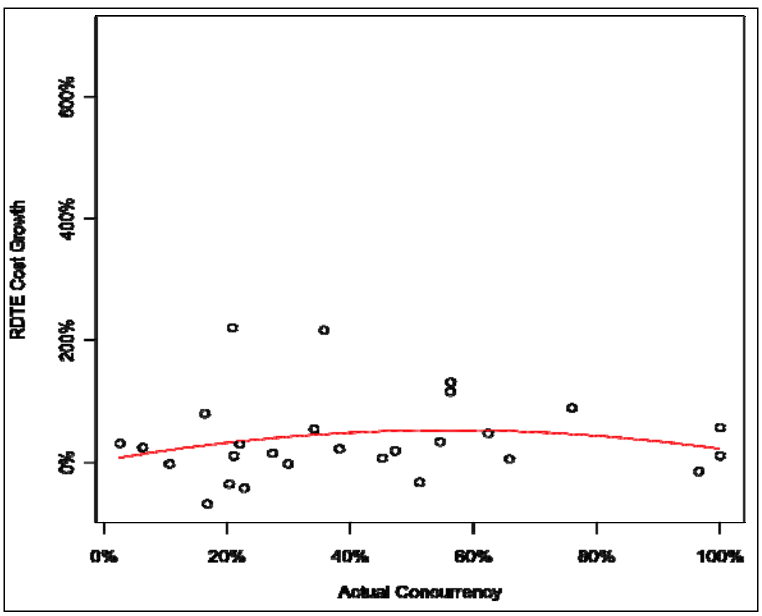
We ran the same model without the outlier to see how well it performs without this data point. While the heteroskedasticity disappears, the overall model actually performs worse according to the adjusted R-squared statistic. Furthermore, none of the estimated parameters are statistically significant.

Table 10. OLS results—RDT&E cost growth vs. actual concurrency (outlier excluded)

	Estimate	Std Error	P value
Intercept	0.023	0.402	0.944
Concurrency	1.773	1.839	0.345
Concurrency squared	-1.585	1.706	0.362
Adjusted R squared		*-0.043*	
Breusch-Pagan		*1.210*	

Also note that the fitted curve becomes concave which is the opposite of our hypothesis. This clearly indicates that the convex curvature of the fitted line using all the data was a result of the very large influence of the outlier. The removal of this outlier resulted in essentially a flat line near zero cost growth (see figure 28 below.)

Figure 28. Fitted curve—RDT&E cost growth vs. actual concurrency (outlier excluded)



To explore other polynomial relations between actual concurrency and RDT&E cost growth, we again use the LOESS smooth curving technique used in the previous cases.

Figure 29 shows the results of 30 bootstrap replications. Note that the original line (drawn in black) is very flat but does have two local inflection points indicating the possibility of a third degree polynomial. However, the 30 bootstrapped lines do not share the same curvature. Furthermore, the outlier clearly adds a great deal of variation in the curves for concurrency levels greater than 0.8. Thus, the data are not well conditioned to a third degree polynomial relation.

Figure 29. LOESS curve smoothing—RDT&E cost growth vs. actual concurrency

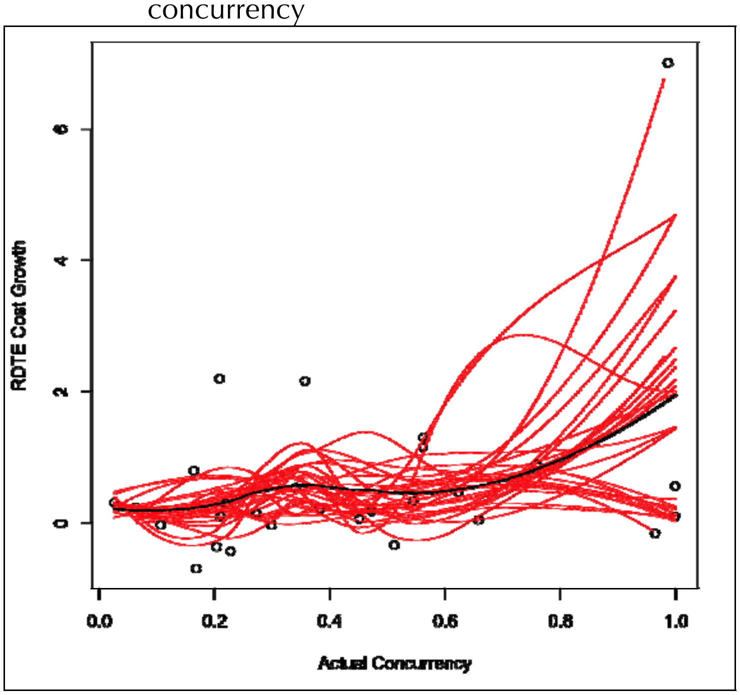
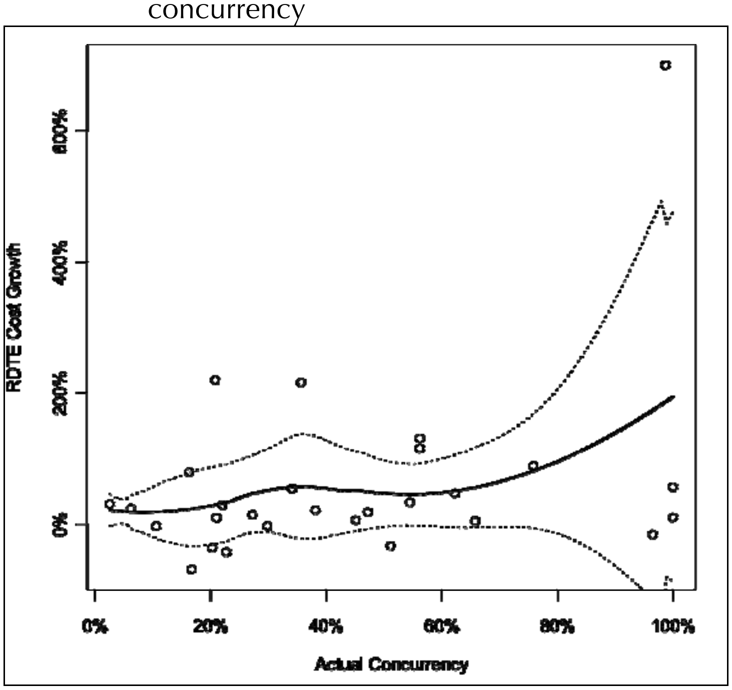
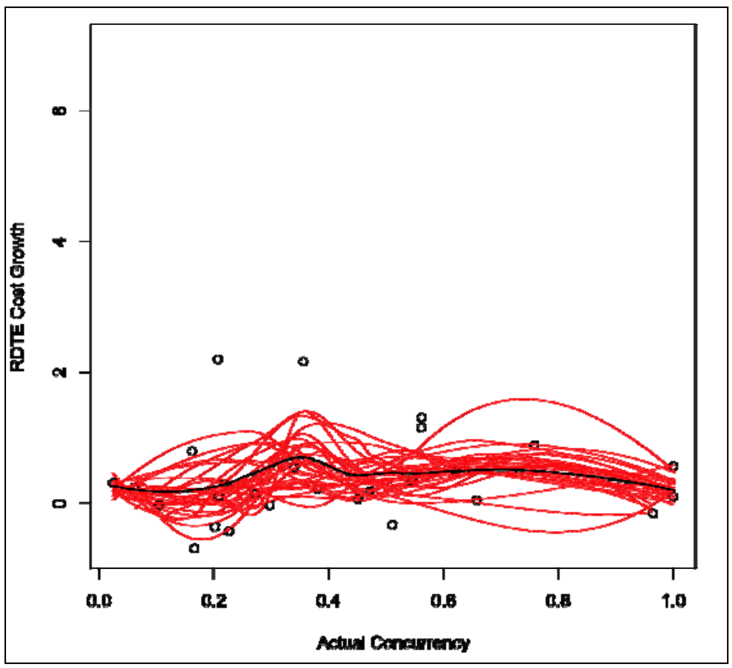


Figure 30. Confidence interval—RDT&E cost growth vs. actual concurrency



The increased variance of RDT&E at very high levels of concurrency is evident in figure 30. However, much of this is an artifact of the outlier. For actual concurrency values less than 60 percent, the variance level of the fitted models are relatively stable except for very small levels. For actually concurrency levels below 5 percent, the variances are correspondingly small. In summary, using all the data, we arrive at a curve that is not stable and whose variance is, for most of the range of concurrency, extremely wide.

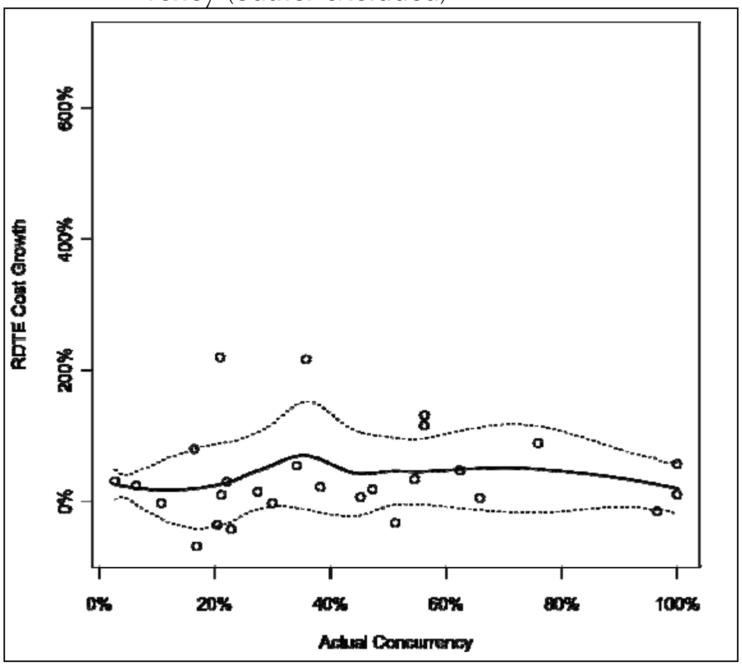
Figure 31. LOESS curve smoothing—RDT&E cost growth vs. actual concurrency (outlier excluded)



The presence of the outlier again forces us to examine how well the data behave in its absence. figure 31 shows the results of 30 bootstrap replications. Despite the excluded outlier, the data are still ill behaved in terms of curvature. The original curve has three inflection points, suggesting a fourth degree polynomial. However, the curves defined by the 30 bootstrapped samples do not consistently exhibit this same curvature. Thus, the data are not well conditioned to a fourth degree polynomial relation.

The variance displayed by the data, with the outlier excluded is much improved at high levels of concurrency. In addition, for most of the range of concurrency, the variance is relatively stable. However, it is still very large.

Figure 32. Confidence interval—RDT&E cost growth vs. actual concurrency (outlier excluded)



In sum, actual concurrency of RDT&E and production funding is not a strong predictor of RDT&E cost growth. The quadratic relation is specifically rejected even with the outlier. In addition, the LOESS smooth curving technique did not yield any other strong polynomial relation between concurrency and cost growth either.

Unplanned concurrency

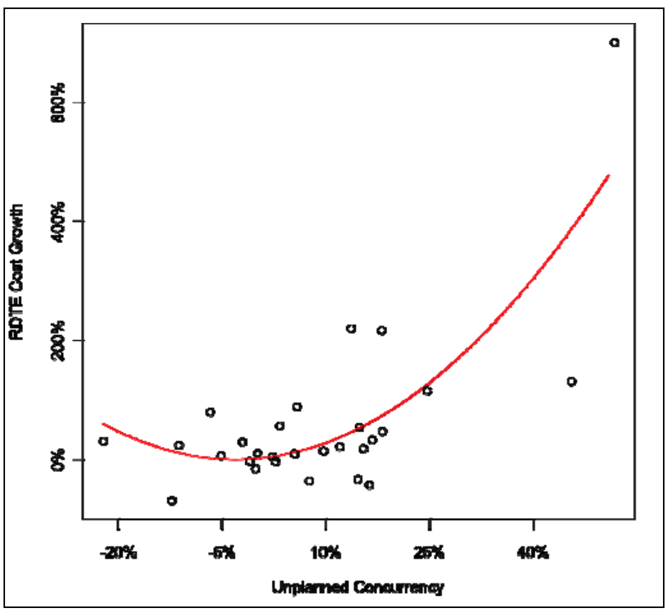
In our final experiment, we explore the relation between unplanned concurrency and RDT&E cost growth. We again begin by examining the possibility of a quadratic relation between this form of cost growth and unplanned concurrency. The OLS results for the quadratic equation are shown below.

Table 11. OLS results—RDT&E cost growth vs. unplanned concurrency

	Estimate	Std Error	P value
Intercept	0.022	0.198	0.913
Concurrency	1.054	1.728	0.547
Concurrency squared	16.401	4.466	0.001
Adjusted R squared	*0.591*		
Breusch-Pagan	*19.498*		

Of all the models presented so far, this one performs the best. The parameter B_2 is highly significant and the adjusted R-squared is 0.591, indicating that the model does a reasonable job of explaining the variance in the expected cost growth. However, we also note that the Bruesch-Pagan test strongly indicates heteroskedasticity in the error term.

Figure 33. Fitted curve—RDT&E cost growth vs. actual concurrency



Looking at the fitted curve, we note strong convexity with the minimum at approximately zero unplanned concurrency. However, the outlier again exhibits a strong influence on the model and may be the cause of heteroskedasticity.

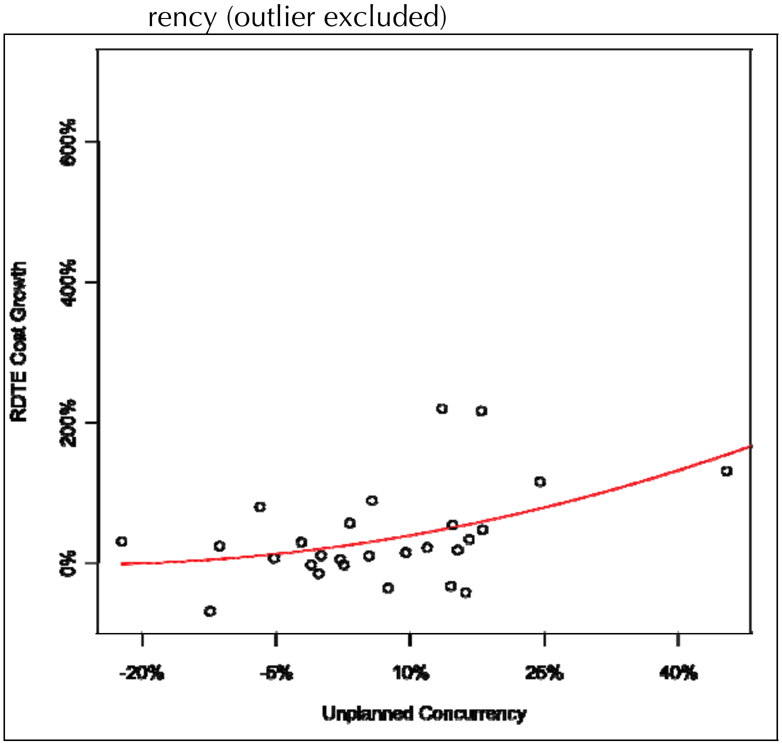
Table 12. OLS results—RDT&E cost growth vs. unplanned concurrency (outlier excluded)

	Estimate	Std Error	P value
Intercept	0.201	0.144	0.177
Concurrency	1.623	1.228	0.199
Concurrency squared	2.955	4.120	0.480
Adjusted R squared		*0.130*	
Breusch-Pagan		*1.495*	

We ran the same model without the outlier to see how well it performed without this data point. The results are presented in table 12. While the heteroskedasticity disappears, the overall model actually performs worse according to the adjusted R-squared statistic. Furthermore, none of the estimated parameters are statistically significant.

On the other hand, the fitted curve does remain convex without the outlier, which is consistent with our hypothesis. However, the fitted curve is very flat and could just as easily be explained by a simple linear relation. In sum, without the outlier, the model loses much of its predictive power and the quadratic relation is weaker.

Figure 34. Fitted curve—RDT&E cost growth vs. unplanned concurrency (outlier excluded)



To explore other polynomial relations between unplanned concurrency and RDT&E cost growth, we again use the LOESS smooth curving technique used in the previous cases.

Figure 35 shows the results of 30 bootstrap replications. Interestingly, the curve defined by the original data is very close to the fitted curve defined by the estimated parameters in table 11. And many of the bootstrapped curves also display this curvature, especially in higher levels of concurrency. However, there is still a strong tendency in the data to produce other curves which have more inflection points, especially at lower levels of concurrency.

Figure 35. LOESS curve smoothing—RDT&E cost growth vs. unplanned concurrency

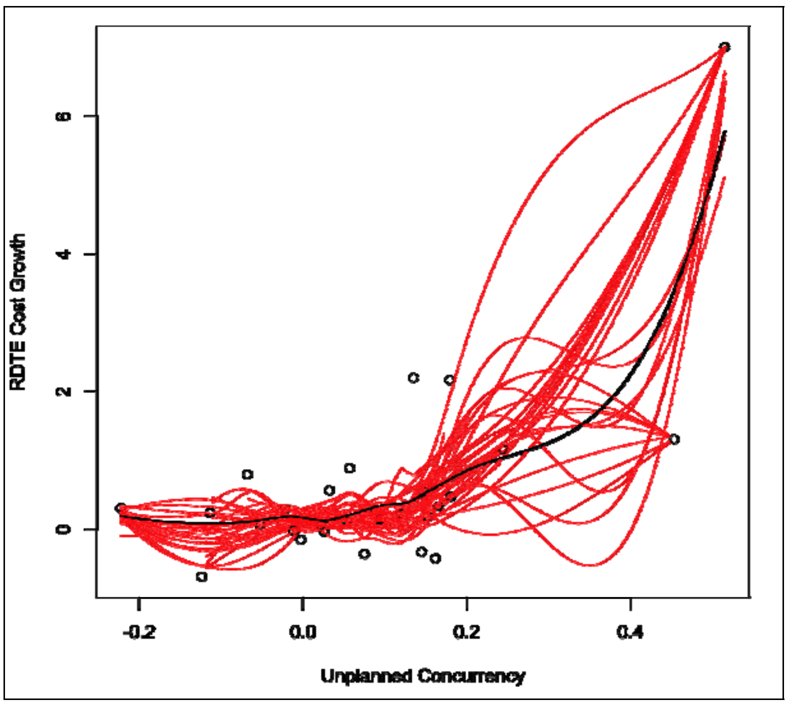
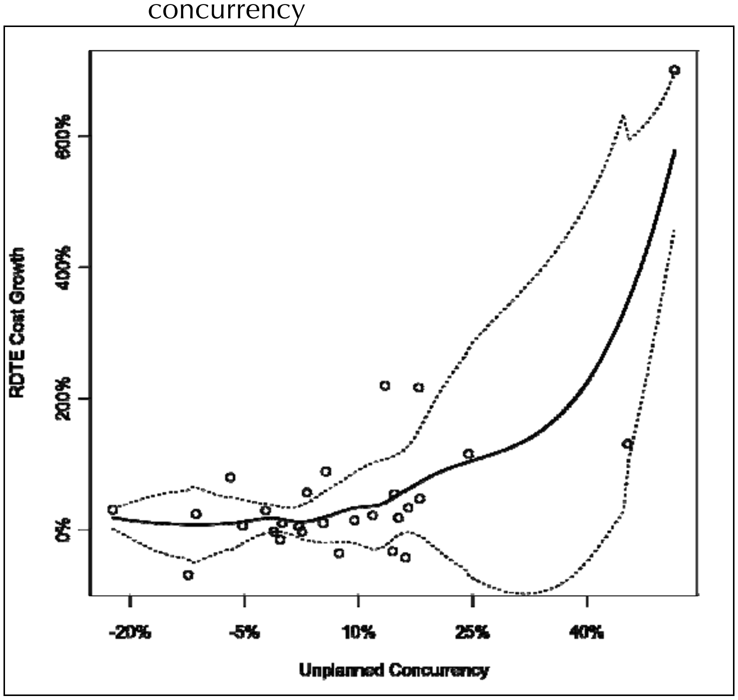
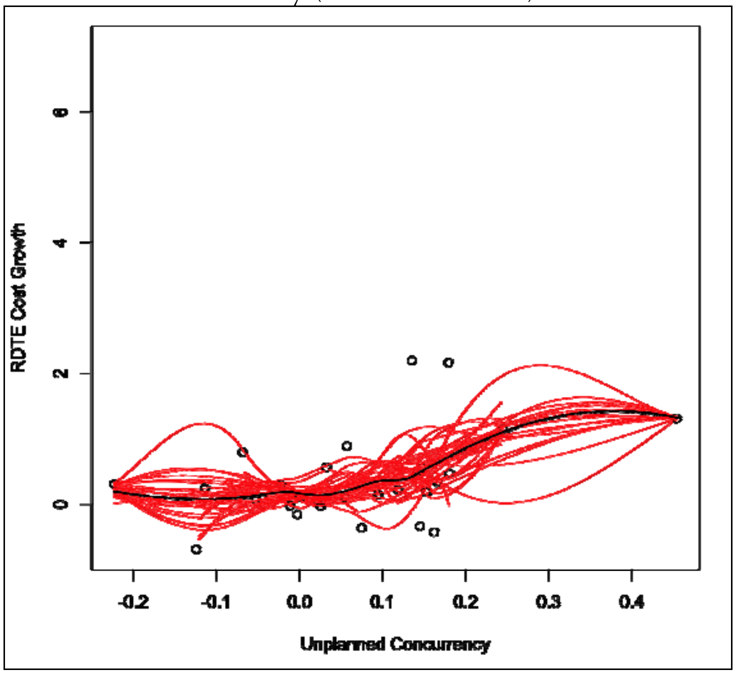


Figure 36. Confidence interval—RDT&E cost growth vs. unplanned concurrency



This behavior is clearer when looking at the two standard deviation confidence interval in figure 36. At negative values for unplanned concurrency, the variance is fairly small and at zero unplanned concurrency, it is actually narrow and centered around zero cost growth. Thus, we conclude that the data are not very well conditioned overall and that the relatively stable behavior in the curvature at high levels of concurrence is largely attributable to the outlier.

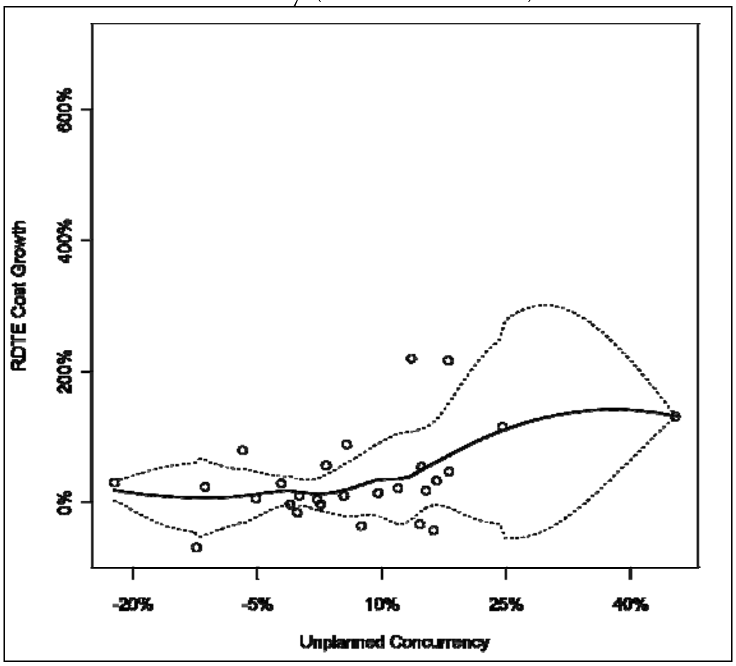
Figure 37. LOESS curve smoothing—RDT&E cost growth vs. unplanned concurrency (outlier excluded)



The presence of the outlier again forces us to examine how well the data behave in its absence. figure 37 shows the results of 30 bootstrap replications. Despite the excluded outlier, the data still displays an inconsistent curvature. Furthermore, the original curve actually demonstrates some downward movement in higher levels of concurrency, which does not make much intuitive sense. Thus, the data, even without the outlier, are not well conditioned to any particular polynomial relation.

The variance displayed by the data with the outlier excluded is improved at high levels of concurrency but is still very high except at unplanned concurrency near zero. For example, the range of cost growth for unplanned concurrency equal to 0.2 is approximately -0.5 to 1.5.

Figure 38. Confidence interval—RDT&E cost growth vs. unplanned concurrency (outlier excluded)



In sum, unplanned concurrency is not a strong predictor of RDT&E cost growth. The quadratic relation is specifically rejected even with the outlier. In addition, the LOESS smooth curving technique did not yield any other strong polynomial relation between concurrency and cost growth either.

Conclusions for RDT&E cost growth

In all cases, we reject the hypothesis that RDT&E cost growth is related to any measure of concurrency of RDT&E and production funding in a way described by a quadratic function. The idea that too little or too much concurrency is a problem is specifically rejected.

We also found no other polynomial relation that accurately supports the data. While using the LOESS curve smoothing routine on all forms of concurrency did lead to an initial curve, suggesting some other possible polynomial relation, further resampling of the data created curves that were far different.

In addition, the two standard deviation confidence intervals from our bootstrap experiments were extremely large. Thus, even if we accepted the implied curvature, the predictive power of the model for any of the concurrency measures was extremely low. In sum, all measures of concurrency had little, if any, relation to RDT&E cost growth.

This page intentionally left blank.

Conclusion

After interviewing several DoD and Navy officials about concurrency, we formulated the hypothesis that too little or too much concurrency caused programs to experience some form of cost growth either in their procurement funding or their RDT&E funding. We formulated this hypothesis as a quadratic response function.

Our results indicate that RDT&E/production funding concurrency does not predict cost growth. We used OLS to assess the hypothesized quadratic relation using the data collected from SARs. In all cases, the quadratic relation was not statistically significant. In addition, using the LOESS curve smoothing method, we allowed for other polynomials to be expressed by the data. Only the LOESS results for unplanned concurrency exhibited some modest relationship but only across a very small range of concurrency. All the other LOESS exercises exhibited no statistical relation, reflecting the OLS results.

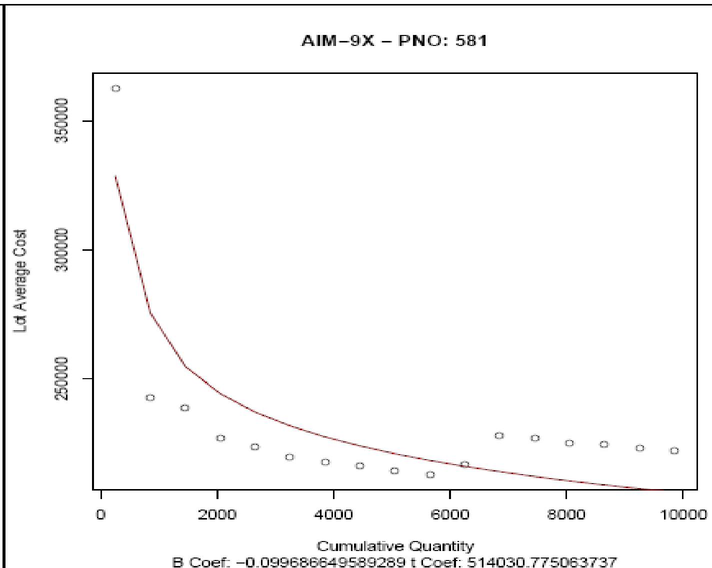
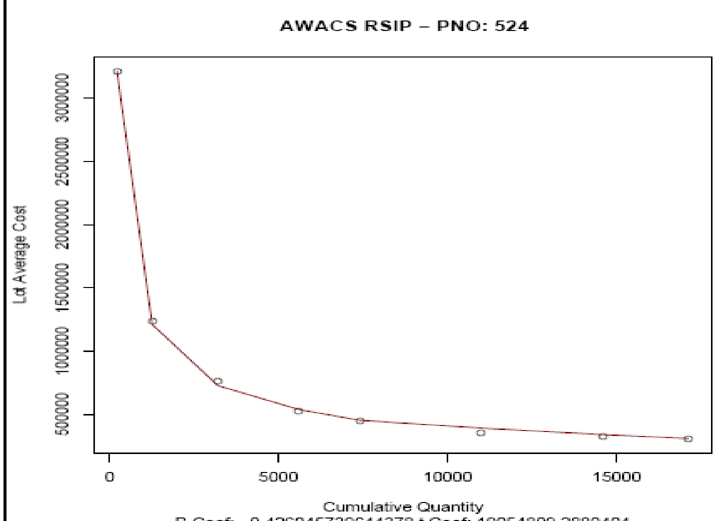
This is a useful finding for acquisition planning purposes, because it indicates that there is no valid reason to avoid funding RDT&E *and* production for the same program in the same year or years.

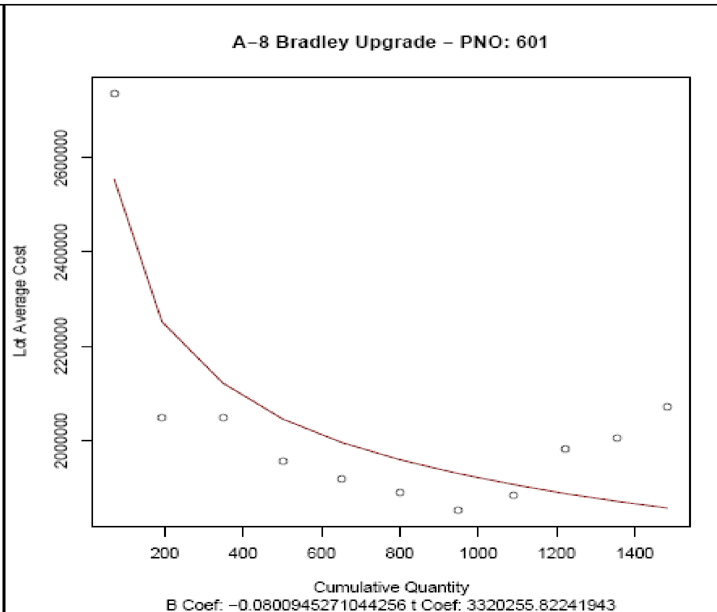
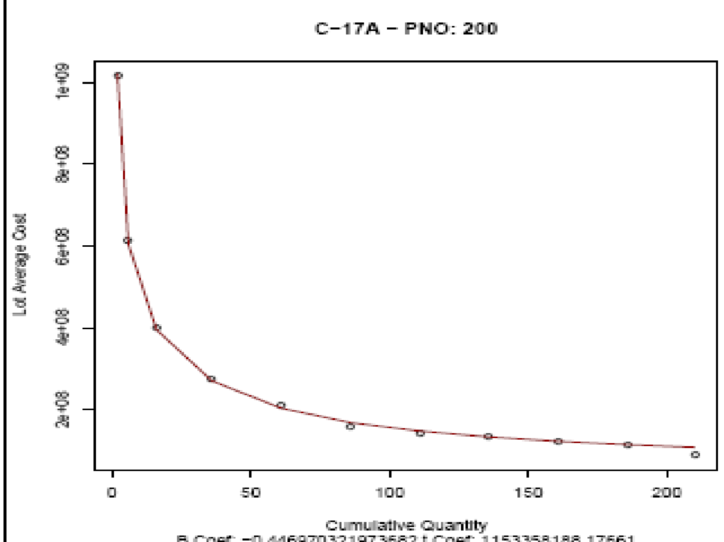
We also considered other definitions of concurrency based on the literature and the conducted interviews. However, data availability and other issues prevented us from assessing them. In addition, SAR data limitations and our models unavoidably yielded results that, in some cases, understate actual concurrency of development/design and production, and in other cases overstate it. Similarly, as noted in the paper, cost growth was overstated for some programs, and understated in others. In spite of this, we are persuaded that our measures for funding year concurrency are a reasonable proxy for development/production concurrency, and our results generally apply.

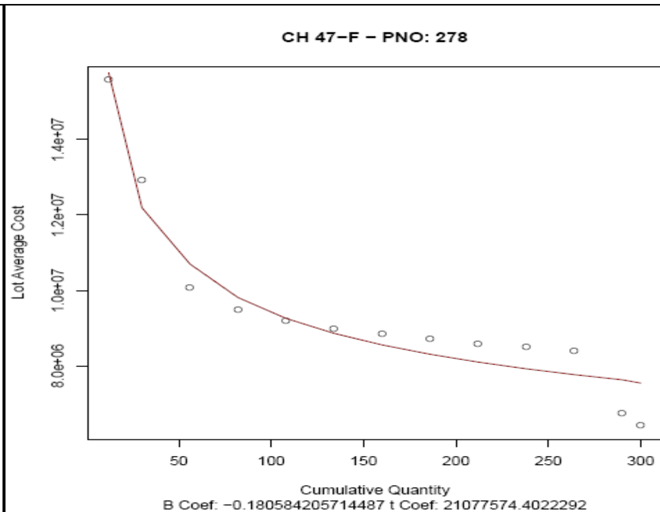
This page intentionally left blank.

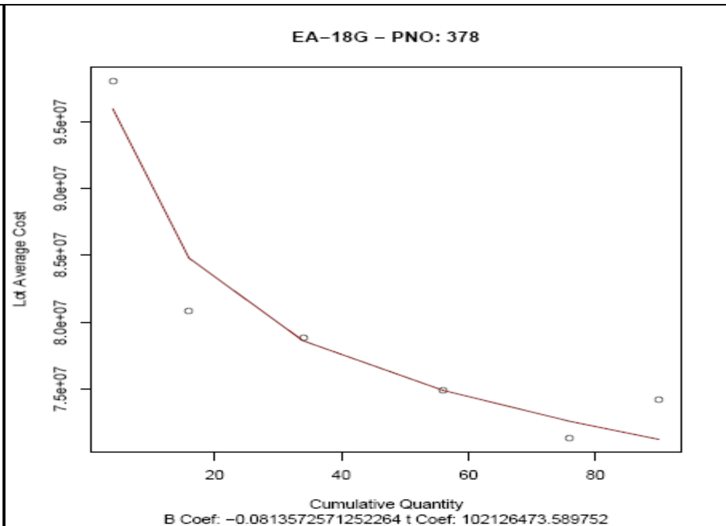
Appendix A: Data Adjustments

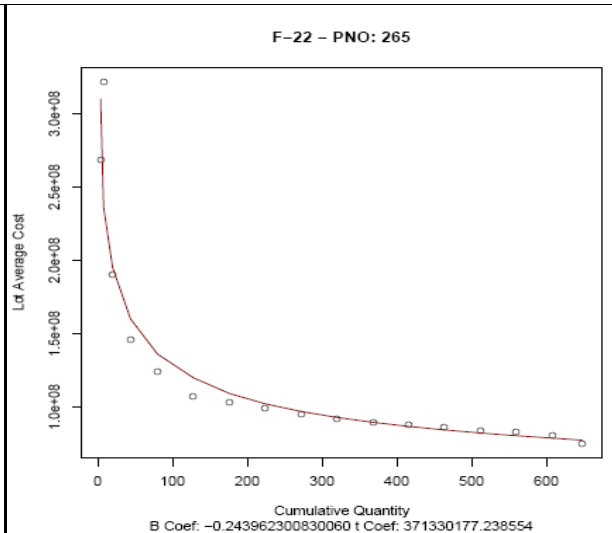
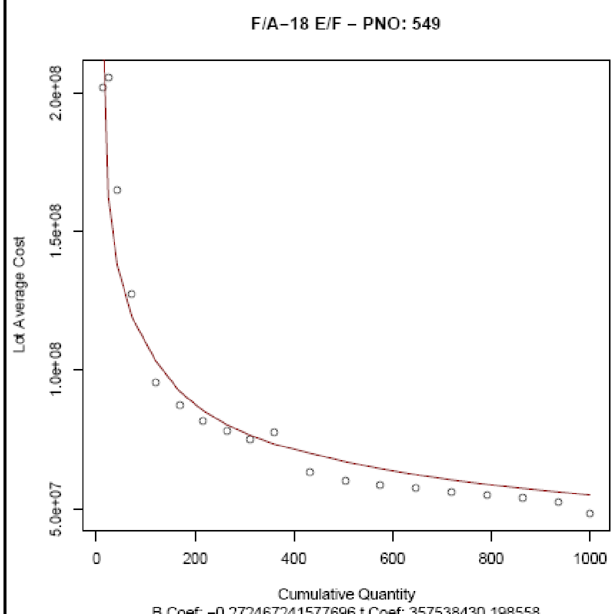
Program	PNO	Notes	Modi-fied	Dropped
1 AESA	330			✓
1.1		Final SAR doesn't contain procurement profile because AESA was apparently rolled up into the F-18 E/F and EA-18G programs.		
1.2		Budget exhibits for the F-18 program contained a line item for 'radar upgrade' that included several programs; however, there was not enough information to pull out the specific costs for the AESA.		
1.3		F-18 R-2 in 2008 Navy Budget contains a partial accounting for some costs. See: Exhibit R-2a, RDTEN Program Element Number and Name:0204136N F/A-18 SQUADRONS Project Number and Name: 2065 F/A-18 RADAR Upgrade		
2 AFATDS	526			✓
2.1		The program was re-baselined several times, preventing comparison of the first and last SARs		
2.2		In the last SAR, pre-98' costs, Block I, are apparently treated as sunk and not accounted. Post-98', Block 2 (future efforts), separated from program along with TCU RLCU Programs. There are no distinctions in the first few SARs that would allow for a reasonable separation of costs into these 'Blocks'		
3 AIM-9X	581			

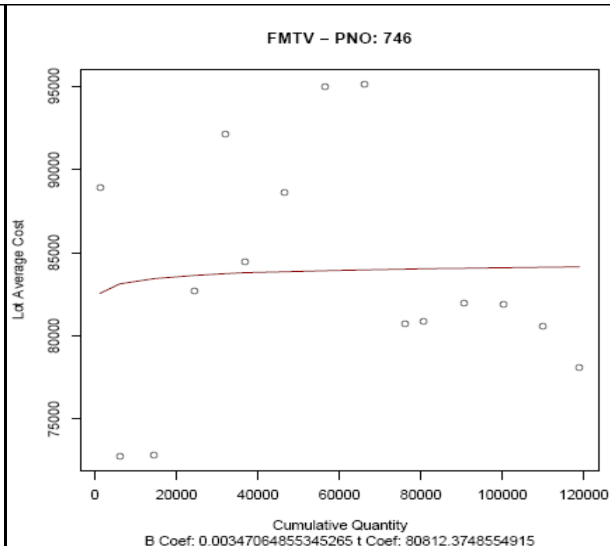
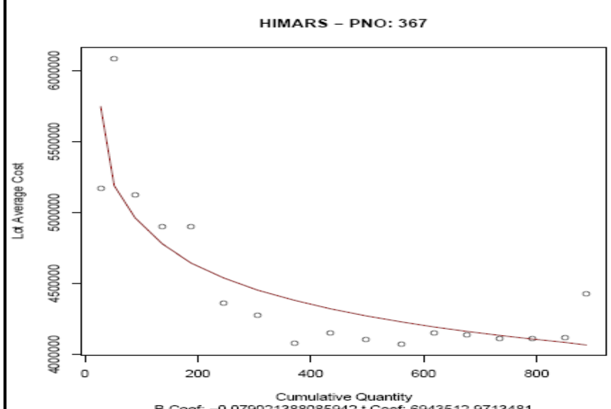
3.1				
4 AWACS RSIP	524			
4.1		We dropped the years of procurement with no associated quantity, in last SAR,		
4.2				
5 Bradley Upgrade	601			
5.1		We dropped the 'tail', or trailing years of procurement with no associated quantity, in last and first SAR,		

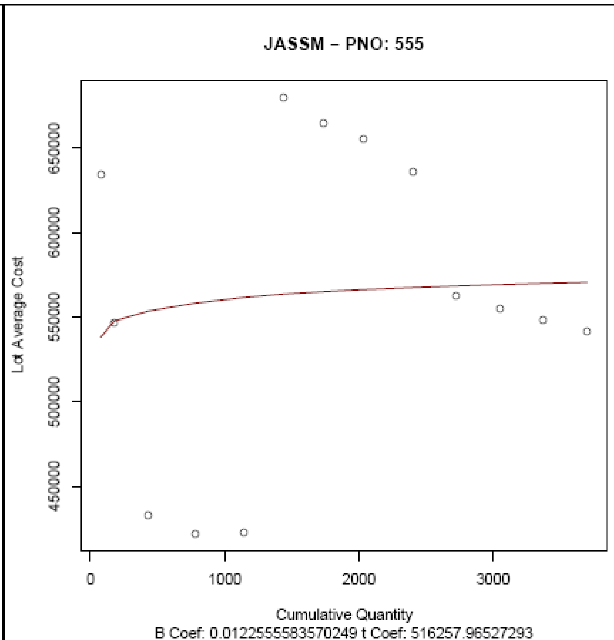
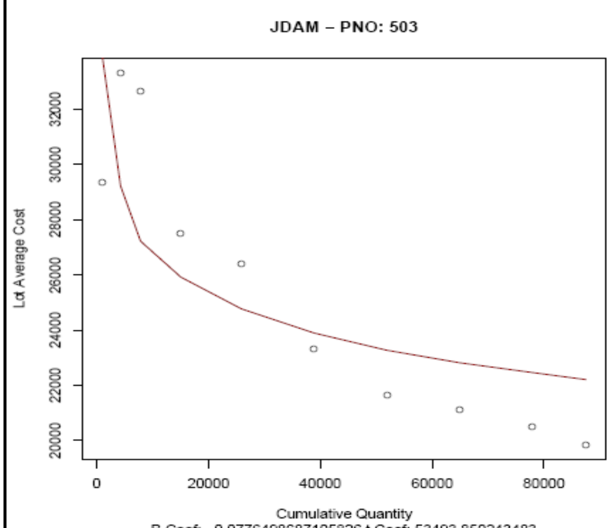
5.2				
6 C-17A	200			
6.1		We dropped the years of procurement with no associated quantity, in last SAR,		
6.2		In last SAR, there is a substantive 'tail', extending from 2008 through 2013. The Air Force budget for the program primarily attributes these to spares, secondarily to training costs.		
6.3				
7 CH-47F	278			
7.1		We dropped the years of procurement with no associated quantity, in the first and last SARs.		

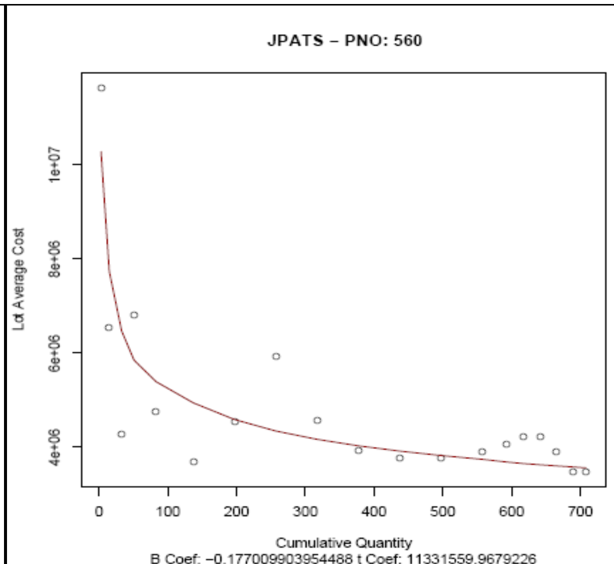
7.2				
8 DMSP	203			✓
8.1		We couldn't match current satellites with the original funding stream. The Air Force Budget Exhibits do not go back into the 1980's.		
8.2		Note from the last SAR: FY86 recurring amount is for primary and mission sensors for the development spacecraft (S-15). The amount shown for non-recurring cost is associated with the Federally Funded Research and Development Center(FFRDC) support. Funding does not match the budget documentation because the SAR is limited to DMSP Blocks 5D-2 Improved and 5D-3 (Satellites 11-20).		
9 EA-18G	378			
9.1		We dropped the years of procurement with no associated quantity, in the first and last SARs.		

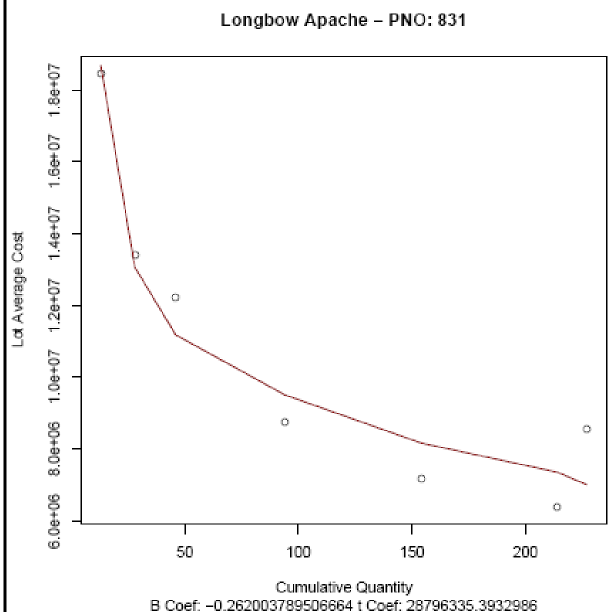
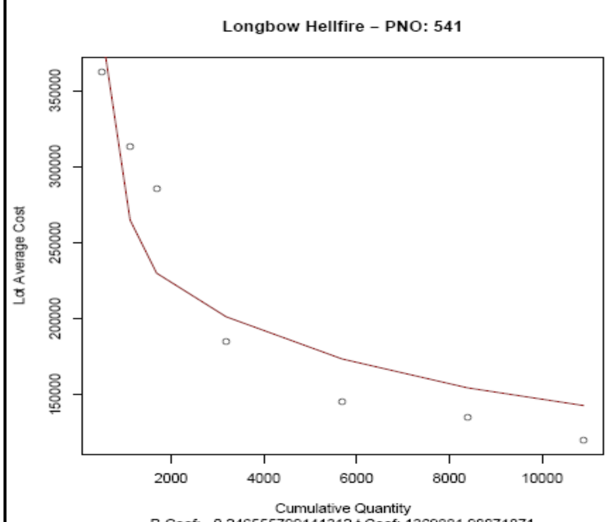
9.2				
10 F-22	265			
10.1		We dropped the years of procurement with no associated quantity, in the first and last SARs		
10.2		Because of inconsistencies in unit cost in the last SAR, we rolled the procurement from 97'-98' into 99', and 00' into 01'. These years had no associated quantities and so were likely advanced procurement for subsequent years.		
10.3		For learning curve adjustment, we dropped the 57-3011 appropriation, Ammunition, from the last SAR because there was no unit association, and Ammunition wasn't planned in the first SAR.		
10.4		Average unit cost inflation appears attributable to ongoing modernization. See Exhibit R-2 in the 2003 and 2002 Air Force Budgets (PE 0604239F)		

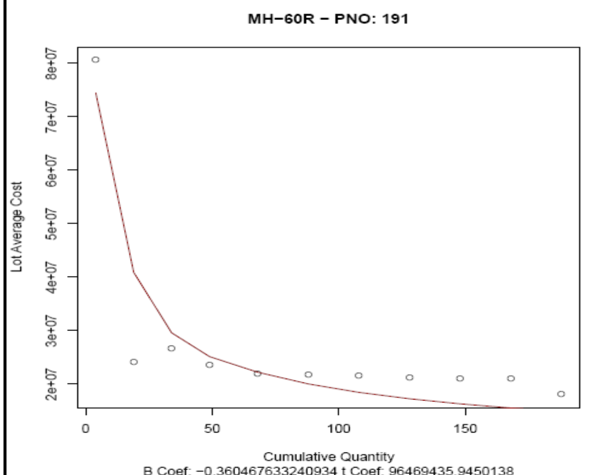
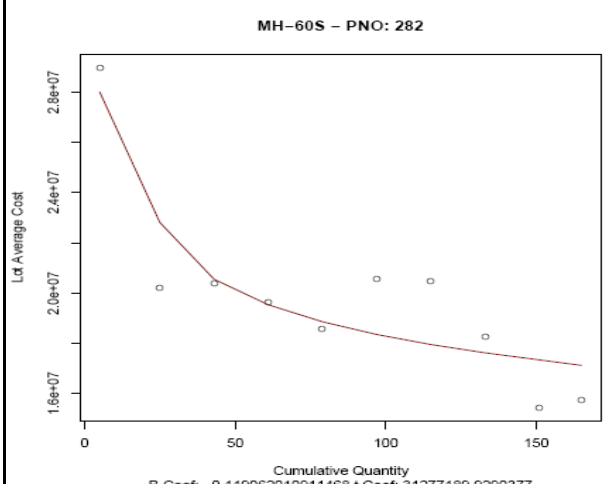
10.5				
11 F/A-18 E/F	549			
11.1		We dropped the years of procurement with no associated quantity, in the first and last SARs.		
11.2				
11.3				
12 FMTV	746			
12.1		We dropped the final year of procurement in last SAR.		

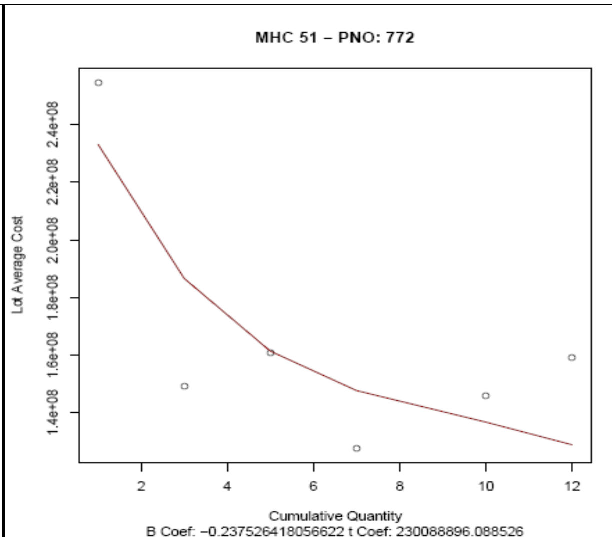
12.2				
13 HIMARS	367			
13.1		We dropped the final two years of procurement in last SAR.		
				
14 JASSM	555			
14.1		We dropped the first year of procurement in both SARs.		

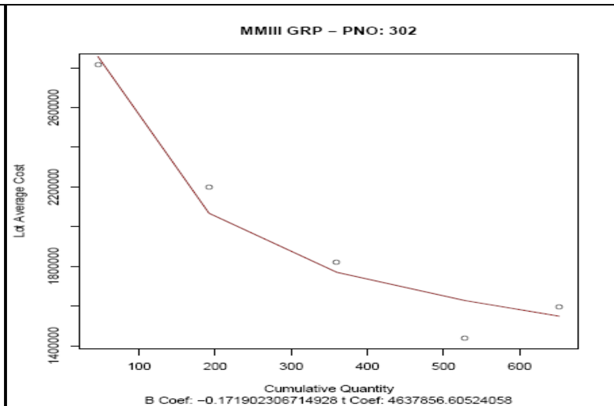
14.2				
15 JDAM	503			
		Transcription from first SAR mislabeled AF procurement as RDTE - corrected. Base year was also corrected in several instances from 93' to 95'		
15.1		We dropped 01' from procurement in the last SAR because there was no associated quantity.		
15.2				
16 JPATS	560			
16.1		We dropped the years of procurement with no associated quantity, in the first and last SARs.		

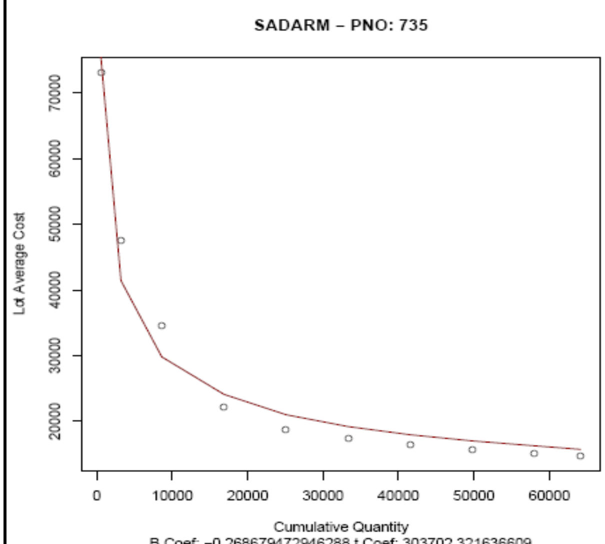
16.2				
17 JSIPS (CIGS)	572			✓
17.1		A significant proportion of costs appeared to be for modifying and integrating existing systems, not new units.		
17.2		Based on large differences in average annual unit costs, we concluded that the units produced for each service were not comparable. And without data to separate shared costs, no per-service learning curve adjustment was possible.		
17.3		Large changes in quantity made direct comparison of the first and last SARs untenable.		
18 LHD-1	217			
18.1		There were large differences in the unit costs between ships. We concluded that this was the result of substantively different configurations - resulting in various final costs. So we were not able to do a quantity adjustment.		
18.2		We concluded that direct comparison between the first and last SARs was inappropriate:The original planned quantity was 3; this was changed 8 units by the final SAR.		
19 Longbow Apache	831			
19.1		We dropped the years of procurement with no associated quantity in the last SAR		

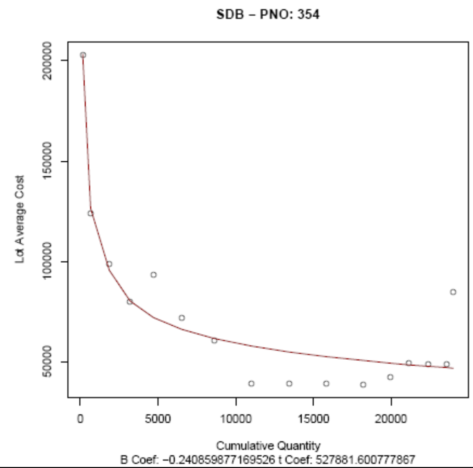
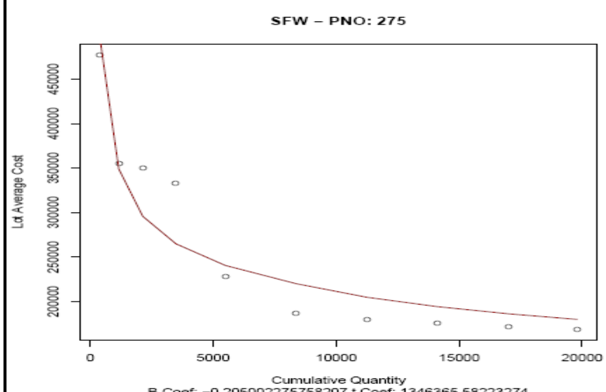
19.2				
20 Longbow Hell-fire	541			
20.1		We dropped the years of procurement with no associated quantity, in the first and last SARs.		
20.2				
21 MH-60R	191		✓	
21.1		We dropped the years of procurement with no associated quantity, in the first and last SARs.		
21.2		The first year of procurement in the final SAR appeared to be associated with RDT&E, so we dropped it from the learning		

		curve calculation		
21.3		Added first year of procurement to RDT&E, based on a review of the Navy's 2000-2003 Presidential Budgets submissions		
21.4				
22 MH-60S	282			
22.1		We dropped the years of procurement with no associated quantity in the last SAR		
22.2				
23 MHC 51	772			
23.1		We dropped the years of procurement with no associated quantity, in the first and last SARs		

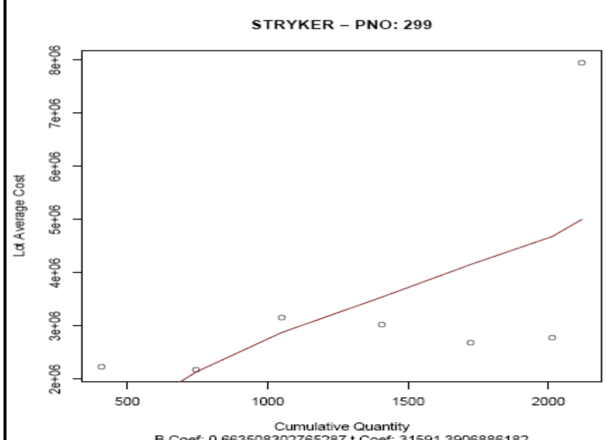
23.2				
24 MIDS	554			✓
24.1		This was an international program with many variants. And it evolved overtime to include more services, each with an apparently unique variant. So, we concluded that comparison with the original SAR was untenable, and we chose not to extract the original estimates for each variant from the historical record		
24.2		The Navy's Budget indicates that these units are purchased by Airframe program, e.g. 11 were purchased by the Prowler program in 07'. There is also indication of several variants, including the MIDS-LVT and MIDS-JTRS.		
25 MM III GRP	302			
		Following transcription, the program name was incorrectly associated with PNO 401. We manually extracted data for this program from DAMIR, and converted it to BY93 dollars.		
25.1		We dropped the years of procurement with no associated quantity, in the first and last SARs		

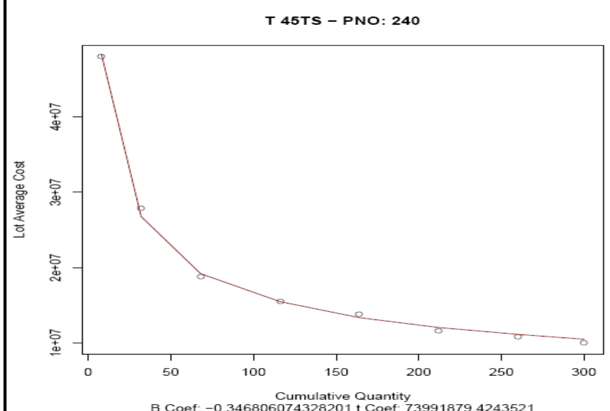
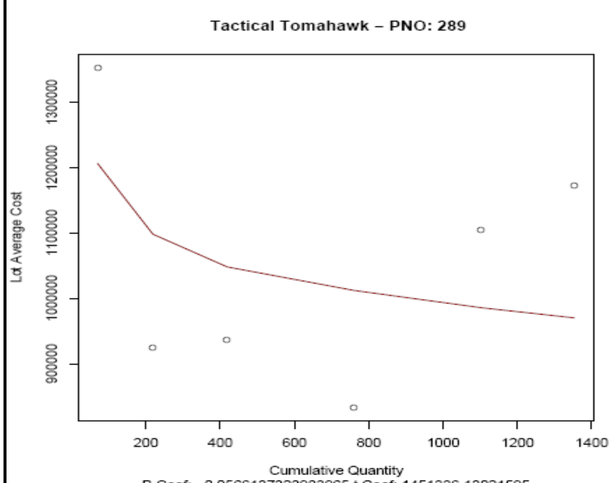
25.2				
26 NAS	537			
26.1		We dropped this program because we couldn't adjust for large quantity differences between the first and last SAR. This procurement funded several, very distinct, subprograms and didn't provide enough information to disentangle the funding		
26.2		The subprograms should be identified as sub-activities or sub-items, because they are so different. For example, Tower automation systems, DAAS, and Airspace Management were all part of the procurement		
27 NESP	551			✓
27.1		We concluded that the program variants were too different to adjust in the aggregate. And the three variations didn't clearly map between the last and first SARs, so a learning curve adjustment wasn't possible		
27.2		GlobalSecurity.org has a helpful summary of the programGlobalsecurity.org		
28 PAC-III	148			✓
28.1		In the first SAR, we found a significant amount of advanced procurement in 95' and 96' with no associated quantities. Unable to locate budget documentation that indicated how we might allocate the money, we elected to drop the program		
28.2		The 94' SAR indicated that the Army Procurement Appropriation was associated with group support equipment, specifying radios and launchers. The final SAR		

<table>
<tr><td></td><td></td><td>lacked this distinction and associated nearly every year of funding with a procurement quantity</td><td></td><td></td></tr>
<tr><td>29 SADARM</td><td>735</td><td></td><td>✓</td><td></td></tr>
<tr><td>29.1</td><td></td><td>The final SAR didn't include the MLRS Rockets sub-program, so we dropped them from the first SAR for all purposes. Otherwise, we thought the learning curve adjustment would still be appropriate, though quantities differ by several orders of magnitude.</td><td></td><td></td></tr>
<tr><td>29.2</td><td></td><td>We dropped the final two years of procurement in the last SAR for the learning curve adjustment. Neither year had an associated quantity</td><td></td><td></td></tr>
<tr><td>29.3</td><td></td><td>
</td><td></td><td></td></tr>
<tr><td>30 SDB</td><td>354</td><td></td><td></td><td></td></tr>
<tr><td colspan="4">Changed Date to 12/30/03</td><td></td></tr>
</table>

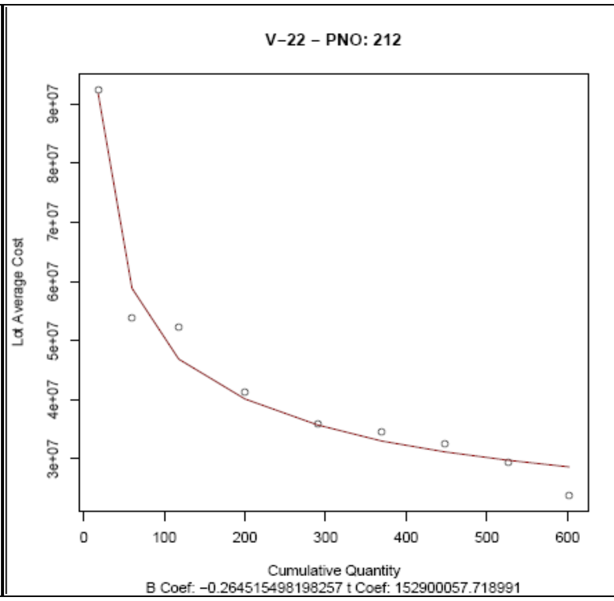
				
31 SFW	275			
31.1				
32 SM-2	234			✓
32.1		The program changed dramatically between the first and last SAR, moving from two sub-programs to four different 'blocks'. We dropped this program because we couldn't locate data to map the costs and quantities between SARs		
32.2		There were fantastic changes in quantity between SARs, so direct comparison was infeasible		
32.3		There were large discrepancies regarding the start of funding for procurement and RDT&E between SARs. We inferred that this was due to an exclusion of early 'blocks' from the baseline for the program		
33 SMART-T	568			✓
33.1		This program had a complicated history,		

		involving multiple services, each with a unique variant of the system, that made the learning curve adjustment untenable. There were also large changes in unit costs that prevented direct comparison of the SARs		
34 SSGN	337			✓
34.1		After reviewing the documentation, we decided that this was primarily a Upgrade/Refuel program, and not a genuine procurement. As such, we thought it couldn't be reasonably compared to the other programs		
34.2		There were substantial additions for SOCOM upgrades that appeared to be unrelated to the purported aims of the procurement. The SAR referred to this as 'Strategic Platform Support Equipment'		
34.3		There was insufficient information in the SAR and DoN Budgets to clearly specify which costs were associated with each unit		
35 SSN-21	258			
35.1		There was no change in the number of units between the first and last SAR, so a learning curve adjustment wasn't necessary. The profiles were directly compared for the concurrency and cost growth calculations		
35.2		The final SAR included a 1810 account that we guessed was associated with the AN/BSY suite. This was absent from the first SAR, but required no adjustment for the above reason		
36 SSN 774	516			
		The transcribed date in the first SAR was incorrect, we corrected it to 9/29/1995		
36.1		There was no change in the number of units between the first and last SAR, so a learning curve adjustment wasn't necessary. The profiles were directly compared for the concurrency and cost growth calculations		
37 Strategic Sea-	565			✓

lift				
		This is the extreme example, but we need to confirm if we want to include modification/upgrade programs in addition to (mostly) new procurement		
37.1		There was no change in the number of units between the first and last SAR, so a learning curve adjustment wasn't necessary. The profiles were directly compared for the concurrency and cost growth calculations		
37.2		The SAR clearly indicates that this is an upgrade program, and not a genuine procurement. As such, we thought it couldn't be reasonably compared to the other programs		
38 STRYKER	299			
		The date for the original SAR was transcribed incorrectly, as were several years of procurement. We corrected these errors.		
38.1		We dropped the years of procurement with no associated quantity, in the first and last SARs		
				
39 T-45S	240		✓	
39.1		In the final SAR, we allocated the procurement costs for 90' and 91' to 89'. We chose this grouping because unit costs for 89' were abnormally low, and the total for 90' and 91' was large		
39.2		Exempting 90' and 91', we dropped the years of procurement with no associated quantity, in the first and last SARs		

39.3		There was a strange unit cost inflation from 02' to 08'. The Navy Budget Exhibits indicated that this was primarily due to increases in per airframe costs. We inferred that this was due to large fixed costs at there was a concurrent decrease in quantities produced.		
39.4				
40 Tactical Tomahawk	289			
40.1				
41 Titan IV	270			
41.1		We found that we couldn't clearly aggregate unit costs because the Titan IV program appear to absorb costs from numerous other missile programs (Titan II, IUS, EELV, DSP, etc...). Other costs, primarily Launch Base Operations, didn't appear in the SARs until 2001, further complicating comparison.		
41.2		There were large changes in the total num-		

		ber of units, preventing direct comparison between SARs.	
41.3		According to the SARs, the NRO funds 50% of missile procurement for the Titan IV program, but no corresponding account appears in the SARs.	
42 Trident II	178		
42.1		The program appeared to be used for life extension for different things, for different boats every year.	
42.2		In the last SAR, there was a large spike in RDT&E for 07'. The DoN Exhibit R-2, RDT&E Technology Application, # PE0101221N contained lots of small, similar developmental items that didn't clearly map to the program.	
42.3		In the last SAR, the procurement budget extends far past the final unit. This was attributable to a D-5 life extension that included holds and hull component upgrades for four C-4 boats.	
43 V-22	212		✓
		We concluded that the Air Force variant should be dropped entirely from the analysis	
43.1		The Air Force and DoD funded special forces variants of the V-22 with significantly higher unit costs. We chose to exclude the appropriations associated with this variant (3010, 3600, 0300, 0400) for the learning curve adjustment.	
43.2		Note from the last SAR: Quantities for the CV-22 are shown under appropriation 3010. In accordance with the approved program plan, the Air Force is funding the majority of the procurement cost for the CV-22. USSOCOM is funding delta costs above the baseline (MV-22) aircraft for Special Operations Forces (SOF) unique equipment As of September 28, 2005, aircraft modification requirements are reflected in the total program procurement costs.	

43.3		V–22 – PNO: 212 Lot Average Cost Cumulative Quantity B Coef: −0.264515498198257 t Coef: 152900057.718991	

Appendix B: Concurrency Questions

1. In most reports on the subject, "concurrency" is defined as "the overlap between completion of development and the start of production."
   Sometimes the definition is more specific, such as CBO's definition, which states "a *non-concurrent* system is one in which planned operational testing has occurred before the production decision" and a "*highly concurrent* system is one in which little or no operational testing has occurred before the production decision." Others have referred to ship "concurrent design and construction" as concurrency.
   - Which of these, if any, makes more sense to you? Are there other definitions you would suggest, and how would you measure it?

2. In your experience, do you think there is an increased technical and/or cost risk associated with
   - Starting weapon system production before its design is complete?
   - Starting production before development or operational testing for major components of the platform (ship/aircraft) is complete?

3. In cost estimation or cost risk analysis, how would you take into account the unknowns of starting production or production design before major components have completed development?

4. Do you know of any reports on the subject of concurrency?

5. Do you have any examples in the programs that you reviewed where concurrent development slowed or otherwise adversely affected the program, or, conversely, helped make the program's cost, performance, or schedule?

This page intentionally left blank.

Bibliography

Mark V. Arena et al, RAND PROJECT AIR FORCE, *Historical Cost Growth of Completed Weapon System Programs,* RAND Corporation, 2006.

Christle, Gary, et. al, "Scientific Analyst Note – Design/Build Concurrency," CNA Memorandum for the Deputy Assistant Secretary of the Navy, April 2007.

Efron, Bradley and Robert J. Tibshari, "An Introduction to the Bootstrap," Chapman and Hall, 1993.

GAO Report to Congressional Committees, "Assessment of Major Weapons Programs," GAO, 2003.

Hough, Paul G. "Pitfalls in Calculating Cost Growth from Selected Acquisition Reports," Rand Corporation 1992.

Memorandum for Deputy Assistant Secretary of the Navy, Management and Budget, subject: Scientific Analyst Note – Design/build Concurrency, 17 April 2007.

Office of the Under Secretary of Defense (Comptroller), "National Defense Budget Estimates for FY2009," March 2008.

Schank, John F. et al, "Sustaining U.S. Nuclear Submarine Design Capabilities," Rand Corporation, 2007.

Sipple, Vince et al, *Defense Acquisition Review Journal* January-April 2004. *Surveying Cost Growth,*

U.S. General Accounting Office, *Weapon Systems: Concurrency in the Acquisition Process,* Statement of Frank C. Conahan, Assistant Comptroller General, National Security and International affairs Division, before the Committee on Armed Services, United States Senate, May 17, 1990

This page intentionally left blank.

List of figures

This page intentionally left blank.

List of tables

This page intentionally left blank.

CRM D0020008.A4/1REV