

AY 2021-2022

THE NEGATIVE IMPACTS OF THE BENCHMARKS

SENIOR ACQUISITION COURSE

COL JEFFREY R. LAFLEUR

COL Jesse R. Marsalis

SEMINAR 04 AY22

COL Nancy E. Blacker, Primary Faculty Advisor

**The Dwight D. Eisenhower School
for National Security and Resource Strategy
National Defense University
Fort McNair, Washington, D.C. 20319-5062**

The views expressed in this paper are those of the author and do not reflect the official policy or position of the National Defense University, the Department of Defense, or the U.S. Government

Introduction

A little-known set of metrics within the Defense Acquisition System (DAS) is detrimental to our nation's national security. Uncodified in any Department of Defense (DoD) Instruction (DoDI), Directive, Manual, or Regulation, not even the Defense Acquisition University (DAU) has historic records on these metrics. While there is no published record of these benchmarks completing a DoD staffing and approval process, leaders in Congress, the DoD, and the services each use them to judge every acquisition effort within the Defense Acquisition Enterprise (DAE). The mysterious metric is the “Office of the Under Secretary of Defense Comptroller (OUSD(C)) Rule-of-Thumb Acquisition Obligation and Expenditure Rates,” also referred to as ‘benchmarks’, and they are slowly suffocating the DoD’s ability to innovate and deliver quality capabilities by incentivizing behaviors that prioritize fast spending over capability delivery (See Appendix A for the current benchmarks).¹

To be clear, the author does not question the value of performance metrics to assist with congressional oversight, to help align acquisition efforts with national defense priorities, or serve as an early warning for programmatic delay. Instead, the author asserts that the current benchmarks—as they are currently structured—are overly simplistic, unrealistic, and not in line with the realities of the budget execution process or modern-day capability development. While Brown et al. (2015), Gallagher and Lee (1996), and Lee et al. (1993) each present data that reinforce how an S-curve model is more realistic than a straight-line model, such as the current benchmarks follow, this author will instead seek to describe the undesirable behaviors that the unrealistic benchmarks create.² As Harvard University professor Robert D. Behn warned, “If you measure it, people will do it... [and] although performance measures shape behavior, they may shape behavior in both desirable and undesirable ways.”³

In presenting the case, the author will begin by providing a theory of why the benchmarks were developed, describe how they are unrealistic, and give a few examples of how they drive behaviors within the DAE. To illustrate the skewed nature of the benchmarks, the author will present the central findings obtained from 70 anonymous interviews with professionals in the grades of GS13/O-4 to GS15/O6 from across the DoD. The author selected a broad range of functional areas (program managers, contracting officers, financial managers, and professors) from across the DAE to ensure a fuller perspective of the benchmarks. In addition, the author integrates information gleaned from congressional staff, senior DoD leaders, and professionals from the defense industry. Finally, the author will provide a set of more realistic benchmarks that the DoD can apply, which will help alleviate the less than efficient behaviors.

Background on the Benchmarks

While the author attempted to obtain the original documents that established the benchmarks, the search turned up empty after engaging every office that one would expect to have the original documents on file. Fortunately, the author connected with several leaders familiar with how the OUSD(C) developed the benchmarks that provided helpful background information. One of the leaders advised that the benchmarks were developed over three decades ago and speculated that they “could have been a result of a tasker to get a handle on budget execution [obligation and expenditure].” Another leader supported this comment by stating that the “[OUSD(C)] came up with these OSD goals [benchmarks] in an effort to pressurize the system to get people to check their accounts.” The two comments corroborate the author's understanding that the benchmark's creation was to provide a target spending rate for the DAE to follow to help lessen the billions of dollars that expire each year before the DoD can use the

funding to acquire national defense priorities. Other senior leaders noted that the OUSD(C) developed the benchmarks as a *guideline only*, not as a directive (emphasis added).

The OUSD(C) provides a well-documented account of expired funding in the Financial Section of the DoD Annual Financial Report (AFR).⁴ For instance, the DoD had \$37.1 billion expire in 2016, including \$4.1 billion of procurement (PROC) funding and \$2.3 billion of research, development, test, and engineering (RDT&E) funding.⁵ To look at it another way, that is \$6.4 billion in investment funding that our nation paid interest on but ostensibly never led to any equipment or new capability. In comparison, \$20.8 billion expired in 2021 and consisted of \$3.3 billion of PROC and \$2.0 billion of RDT&E, or \$5.3 billion in investment funds.⁶ In those five short years (2016 to 2021), the DoD decreased the amount of total expired funding by 44 percent, including a 17 percent reduction in expired investment dollars (See Appendix B for a graphical representation of this change). While this example does not explain the 30-plus year age of the benchmarks, it does provide evidence for how the recent aggressiveness to meet the benchmarks has improved budget execution and decreased the expiration of appropriated dollars. The author acknowledges the importance of executing DoD appropriations as efficiently as possible, but an aggressive focus on spending fast does not necessarily lead to spending well.

While the OUSD(C) would not provide the author with the information used to develop the original monthly benchmarks, a leader did advise that the metrics were constructed by averaging historical obligation and expenditure metrics. The same leader acknowledged that the benchmarks had not changed much since their origination and stated that “in any given fiscal year, about 1000 lines are below metric [benchmark] and 1500 are on or above [benchmark].” When the author questioned if the OUSD(C) revisits the validity of the benchmarks, the leader noted that “[OUSD(C)] goes back and looks at performance against those metrics [benchmarks]

to see if [OUSD(C)] should change them, but the metrics [benchmarks] are pretty sound.”

Several other professionals called attention to the issue that will occur if the benchmarks are created based on historical data, then used to incentivize behavior, and subsequently validated against the same data. As one professional noted, “If it’s [benchmarks] at all based on historic data... it will influence variability...i.e., [program managers] will execute to meet that number for the purpose of meeting that number versus execute to the mission and then measuring the outcome post hoc. So, it can artificially influence the average and reduce variability based on real events as teams work to overcome them [benchmarks] simply to meet the rule-of-thumb.” In short, the way the OUSD(C) has established the benchmarks and then continued to grade acquisition efforts by the same criteria has created a bias that continues to confirm the standards. Since the OUSD(C) developed and realigns funding based on the benchmarks, the entire DoD prioritizes efforts against the benchmarks. In the world of performance measurements, it's like a self-licking ice cream cone!

The Benchmarks are Unrealistic

To the untrained eye, the structure of the monthly benchmarks appears supportable. What's to question about a set of spending metrics that start small in October and grow in a cumulative fashion and predictable straight-line slope until they reach their end of life? While past studies—such as the Brown et al. study identified earlier in this paper—have called attention to the realistic nature of S-curve models, the DoD continues to base the benchmarks on an unrealistic straight-line model.⁷ Unfamiliar to many in the DoD is the knowledge that the benchmarks disregard the reality of the budget execution process and the ambiguity that is natural to capability development. The author noted that most of those interviewed recognize the

inherent delays, including the distribution of obligation authority, the time to award a contract, and the expenditure process.

Time to receive obligation authority

Before a program office can obligate funding, it must first have the obligation authority, but the benchmarks do not account for those typical delays that occur in receipt of obligation authority. Instead, the benchmarks assume possession of the obligation authority on the first day of the fiscal year (FY) (October 1), no matter when an office receives the authority. To understand such delays, one only needs to consider the distribution of the first tranche of FY 2022 continuing resolution (CR) funds. While the OUSD(C) distributed the initial CR obligation authority to the services and fourth estate components on October 1, 2021, the pace of distribution appears to vary significantly following that day. One service budget office advised the author that it received its obligation authority on October 4, one business day later, and a fourth estate component reported not seeing its first obligation authority until October 14.

Unfortunately, the timeline worsens the closer you get to the lower-level program offices that execute the acquisition effort. One financial manager advised that their office didn't receive CR obligation authority until October 26, three weeks into the FY. After speaking to dozens of professionals, the author found that most offices had to wait three to four weeks to receive their obligation authority. Some fourth estate offices even advised that they have had to wait 45 days in previous years. The author speculates that most of the delays in obligation authority reaching the lower-level program offices are due to the limited amount of CR authority (CRA) that services are allowed to spend based on the short-term CR. For instance, the first CR of FY 2022 was signed into law on September 30, 2021, but ended on December 3, 2021, allowing the DoD access to one-sixth of the amount of funding it spent in the previous FY. As a result of the

limited amount of time that services had to obligate funding, leaders were under pressure to provide the CRA to those that could spend the fastest or risk losing that obligation authority.

Time to obligate

Like the obligation authority disbursement, the realities of the contract obligation process are also not integrated into the benchmarks as they provide little consideration for proper contract negotiations, contracting officer capacity, or the benefits gained through a delayed award. A 2013 DAU study outlined two of the most significant reasons for acquisition offices not meeting the benchmarks as obligation delays, specifically, delays in negotiations and awards.⁸ To assess how the benchmarks do not align with the obligation processes, one only needs to consider the Procurement Action Lead Time (PALT), a measurement commonly employed in contracting. The DAU defines PALT as "the amount of time required to complete the actions leading to contract award" and includes the time it takes to consolidate the required documents, issue a solicitation, receive proposals, select the winning bid, negotiate, and award the contract.⁹ While there is no published standard DoD PALT, like an infantry battalion might maintain a physical fitness standard, lower level contracting commands commonly establish PALT standards to ensure they manage expectations for customers and employees.

One example PALT model presents a standard of 45 days for awarding a \$1 million task order on an *existing contract*, while a *new competitive contract* of the same amount is estimated to take as long as 150 days (emphasis added).¹⁰ The same website publicized that the organization's average PALT is 174 days, or almost six months to award. The lengthy period it takes to award many contracts calls attention to the complexity of research, analysis, and coordination that contracting officers must conduct before obligating the government. Additionally, whichever PALT that a contracting officer is operating under, a majority still fall

outside the published monthly benchmarks. For instance, considering the time to execute the simplest of the previous examples (the \$1 million priced contract), it would take between two to six weeks to receive the obligation authority and an additional 45 days to award for a best-case timeline of two months and the worst case of three months. An experienced contracting officer noted a similar time stating, “I don't see how [contracting officers] can award anything in the first quarter unless they are just executing options.... You can award some simplified buys and continue previous service contracts, but you will see that for most organizations, just doing a new task order or delivery order can take 60 to 90 days.”

Time to expend

Without question, the most prevalent issue raised by the majority of those interviewed is the misalignment of the RDT&E monthly benchmarks to the reality of the budget execution and capability development process. Since RDT&E appropriations fund exploratory events such as prototypes, software development, and test activities, they commonly result in expenditure profiles that fall outside the benchmark guidelines. While efforts that pay for labor will realize faster and more frequent expenditures, those efforts contingent on a development, deliverable, or test event will not register an expenditure until well after the item is delivered or until the milestone is complete.

As most acquisition professionals will understand, just because an effort does not align with the benchmarks does not mean that the effort is premature or that the program is not progressing well. In many cases, a program that does not align with the benchmarks implies a delay outside the program manager's control. For instance, any significant slip in a test event can interrupt the program office's ability to utilize the test resource (e.g., range, lab, or aircraft) for some time, leading to delays in expenditure. Unfortunately, to comptrollers and congressional

staff, expenditures that do not align with the benchmarks present evidence that the technology is not ready or the program office did not plan properly. In many cases, such delays can even get a program the infamous “early to need” cut, even though such delays are uncontrollable and aligned against an unrealistic benchmark. While some will say that the comptroller and congressional staff will consider such challenges before rephasing funding, the author posits that the issue resides in the fact that their analysis is flawed because they use a set of unrealistic benchmarks.

Before a program office can realize an expenditure, it must first maneuver the complex invoicing process. While the complexities of the invoicing process are relatively unknown or ignored by many that prioritize the benchmarks, they play a considerable part in expenditure delays and are out of the program office’s control. For example, in the case of a prototype development, the equipment must be accepted by the government representative, the vendor must submit an invoice (i.e., an electronic payment request for the work completed), and a receiving report must be provided to the paying agency before the payment can be made to the vendor. Any interruption in any one of the previous steps, no matter how slight or justifiable, will have a cascading impact on the expenditure rate of the program. In discussing the invoicing process, one senior program manager stated, “the vendor executes invoices every 30 or 45 days, which is outside my control.” Emphasizing the lack of influence that program managers have over the invoicing process specifically, one financial manager noted that “it’s not like he [program manager] can knock on the vendor's door and say, 'Where’s the invoice?’” However, the author spoke with several acquisition officers who advised that they attempted just that with limited success. Another financial manager introduced an additional complexity that is important

to note, stating that “industry has five years to invoice, and some don’t want to invoice until its shareholder time.”

But they’re just guidelines!

The author discovered that the OUSD(C) originally intended for the benchmarks to serve as a guideline and not as a mandate for spending speed. However, anyone that believes that authorities treat the benchmarks as guidelines has not taken the opportunity to understand how the metrics drive behaviors throughout the acquisition workforce. As Harvard professor Robert D. Behn noted, “[o]ften, such requirements are described only as guidelines.... Do not be fooled. These guidelines are really requirements, and these requirements are designed to control.”¹¹ While the originators of the benchmarks might have intended them to serve as a guideline, a majority of the 70 professionals the author spoke with confirmed that the benchmarks incentivize—if not control—the behavior of the entire workforce. In this case, it is essential to appreciate the warning that Jerry Z. Muller noted in his book, *The Tyranny of Metrics*, when he wrote that “any measure used for control is unreliable.”¹² The author asserts that the benchmarks are used for control and are unreliable.

The benchmarks drive behaviors

Campbell’s Law

When authority places significant attention on a metric, the authority can create unintended incentives that drive behaviors that can have detrimental impacts. The consequences that result from focusing on such metrics were studied by the American social psychologist David T. Campbell in 1975 and resulted in a commonly used principle. Campbell’s Law states that “the more any quantitative social indicator is used for social decision-making, the more

subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.”¹³ An easier to understand description of Campbell's Law, originated by Marilyn Strathern, states that “when a measure becomes a target, it ceases to be a good measure.”¹⁴

The principle captured by Campbell's Law provides a warning of the damage that can occur when leadership applies significant weight to a metric, especially one as unrealistic as the monthly benchmarks. The author first experienced the negative impact that the benchmarks created in 2013 when the organization he was a member of had a significant amount of funding rephased based on not meeting the monthly benchmarks. What made the loss of funding hard to stomach is the fact that it occurred shortly after the publication of a 2012 OSD memorandum, cosigned by OUSD(C) (HON Robert F. Hale) and Under Secretary of Defense for Acquisition, Technology and Logistics (HON Frank Kendall). The memorandum stated that the "late obligation of funds should not be presumed to imply that the funds are not needed or that future budgets should be reduced unless there is other evidence to support that conclusion.”¹⁵ In the author's case, instead of an evidence-based decrement, authorities took the funding through a simple “peanut butter spread” reduction across the enterprise based on how organizations ranked against the monthly benchmarks. The author's command learned the lesson, and behaviors quickly changed, as leaders prioritized the benchmarks above everything else, including operational needs. The author was surprised sometime later when he saw the customer direct a shift of a higher priority effort to execute later in the year to support improved monthly benchmarks. In short, the aggressive focus on the monthly benchmarks had taken a stranglehold as the customer started prioritizing them to protect funding above the pursuit of capabilities

deemed critical for its warfighters. The benchmarks became the target, not the delivery of capability.

Pressure to meet the Benchmarks

While the OUSD(C) developed benchmarks were intended to serve as a guideline, they have become the leading standard by which the DAE measures the status of acquisition efforts, even above capability delivery. Throughout the anonymous interviews, the author discovered a troubling pressure within acquisition culture to aggressively pursue the monthly benchmarks. This pressure appears to be exacerbated by the fact that the OUSD(C) provides the benchmarks to the very institution that provides funding to the DoD, the congressional appropriations committees. In turn, the appropriation committees use the benchmarks as a valid performance measurement to assess the status of acquisition efforts. Of course, since congressional staff and the OUSD(C) use the benchmarks to realign funding, the services and components amplify the metric with their PEOs. Naturally, since the PEOs are under scrutiny by the benchmarks, they heavily prioritize their use to ensure their subordinate offices target them, and the process repeats itself.

The two most influential users of the monthly benchmarks are the OUSD(C) and congressional committee staff. The OUSD(C) values the simplicity of the monthly benchmarks as a guideline in the development of spending plans and to use them as an early warning of issues with programs. In addition, the OUSD(C) uses the benchmarks at decisive points to rephase funding from those programs that do not align with the benchmarks. Congressional committee staff, however, utilize the benchmarks to a lesser degree. Professional Staff Members (PSM) interviewed advised the author that, while they receive and monitor monthly accounting reports provided by OSD, they underscored that they do not use the benchmarks as an “exact

dial.” The PSMs further advised that “If OSD is looking at the [monthly benchmarks] strictly... [or] if slave to the [monthly] rates [benchmarks]... *that’s a problem* (emphasis added).”

Unfortunately, because we are operating with unrealistic benchmarks that leaders use and emphasize, the PSM’s concern about strict adherence to the monthly benchmarks is already a reality. Recalling the PSM’s concern, *we have a problem* (emphasis added).

One of the first motivators present within the DAE culture that quickly emerged during the interviews is that most declared they wanted to be *green* in the benchmarks (emphasis added). In explaining the focus on the benchmarks, one program manager said, "If you're green, no one bothers you!" Another program manager said, "leaders don't want to have to answer questions of appropriators..., so they set an extraordinarily high focus on the monthly rates so that they are not scrutinized." The benchmarks are so engrained in the culture that the acquisition education system has integrated them into their course. In addition to being a part of many DAU courses, one acquisition officer advised that “[service system coordinators] are literally trained to focus the forms [PROC and RDT&E documents] against the monthly rates [benchmarks], so they don't lose funding.”

The benchmarks drive behaviors.

As highlighted in the previous section, there is significant pressure to meet the monthly benchmarks, incentivizing the acquisition community to behave less than efficiently. One prevalent behavior that the author discovered is to select a capability developer based on how fast an organization can spend rather than on the innovative capabilities they can deliver. When speaking about how the benchmarks drive the selection of contractors, one acquisition official said that small businesses and new entrants are too slow to expend, so he prefers to “pick proven vendors.” Another program manager advised that many “government providers are prioritized

because they can obligate immediately over industry many times because they can execute faster, even when it isn't a better product." While Congress and OSD have provided many tools to reach the more innovative companies, such as Mid-Tier Acquisition and Other Transaction Authorities, those avenues do not excuse the DAE from meeting the monthly benchmarks. The damage that such pressure creates in trying to stay innovative while also trying to abide by designated metrics was noted by Muller when he wrote, "trying to force people to conform their work to preestablished numerical goals tends to stifle innovation and creativity."¹⁶

Another widespread behavior that the author discovered is how the benchmarks drive the DAE to select specific contract types based on the speed of expenditure rather than on which is the best business decision for the government. For instance, when it might be possible to delay a contract award to provide the government with a better value, program offices are incentivized to make contract award decisions based solely on the monthly benchmarks. As one program manager noted, "it might make sense to do an indefinite delivery/indefinite quantity [contract] but choose a stand-alone [contract] because it's faster." In another case, a contracting officer said, "I have been in programs where options were exercised that might not have been exercised if we were not under those rates [benchmarks]." The benchmarks remove any ability of the acquisition community to make the best business decisions while incentivizing them to make bad ones.

The author also found that most acquisition professionals interviewed construct RDT&E contracts to expend rapidly to ensure that their office gets as close to the benchmarks as possible and as early as possible. Program offices use such techniques when they are under constant scrutiny to *get green* in the monthly benchmarks (emphasis added). One common example of such a technique is integrating performance-based payments (PBP) into a contract. Many acquisition professionals use PBPs regularly to expend large amounts of funding early in the

effort during such milestones as a program kickoff, preliminary design review, or critical design review. In most cases, PBPs are the only realistic way developmental efforts can meet the expenditure benchmarks since the benchmarks are out of phase with the apportionment, obligation, and expenditure process. Considering the combined 60 to 90 days that it takes to receive an obligation authority, award a contract, and register an expenditure, many acquisition offices will have to expend up to 18.3 percent for the first milestone if they want to keep pace with the benchmarks. Unfortunately, chasing unrealistic benchmarks through PBPs has negative implications. As one experienced acquisition official highlighted, “you can give a lot of money at the beginning, but you are giving the contractor more power and taking on more risk by doing so.”

Another typical benchmark-driven behavior is accepting deliverables with deficiencies so that the office does not delay its expenditures. As noted earlier in this paper, before an expenditure can register, a government representative must first accept the equipment, but the slightest delay will harm the program’s benchmarks. The author learned through the anonymous interviews that many in the DAE will accept the hardware and correct the deficiencies later, costing the government additional time, money, and potentially delay innovations. Defense industry professionals interviewed also confirmed that they use this technique to their advantage, recognizing the pressure that program offices are under to meet the benchmarks. As one program manager advised, the benchmarks “encourage [the] government to accept more risk, instead of pushing the vendor to fix the development effort. We accept the deliverable and plan to just fix it down the road.” This same behavior is also realized during test events, as described by one program manager that said, “often in a situation where our test results are not as good as we expected, [we] would like to continue to test, but if we do, money is going to get taken.” While

this gamesmanship might keep the effort on schedule to meet the benchmarks, it costs the government more money in the long run.

Conclusion

When you consider how the benchmarks are misaligned from the reality of the budget execution process, recognize their lack of attention for how modern-day capability development occurs, and understand the bias in how the benchmarks are validated, it becomes clear that they are an unreliable performance measurement. While modernizing the benchmarks is not a radical change for the OUSD(C) to implement, the long-term positive influence of such an update on the DoD's ability to innovate and deliver quality capabilities is incalculable. To that end, the author recommends that the OUSD(C) immediately replace the current benchmarks with a version that accounts for delays in obligation authority disbursement, contract award timelines, and the realities of the expenditure process.

To assist in the recommended change, the author provides a more realistic set of benchmarks which include adjustments in both obligation and expenditure for PROC and RDT&E (Appendix C). The benchmarks provided in this paper more closely reflect the reality while affording both congressional staff and OUSD(C) a modern-day performance metric that they can use in assessing acquisition efforts. It should be noted that these benchmarks provide a slight adjustment to the obligation benchmarks while offering a significant adjustment to RDT&E expenditures. The significant change to the expenditure benchmarks over the obligations acknowledges the misalignment of the current benchmarks and where they are causing most of the damage to innovative capability development and quality of deliverables.

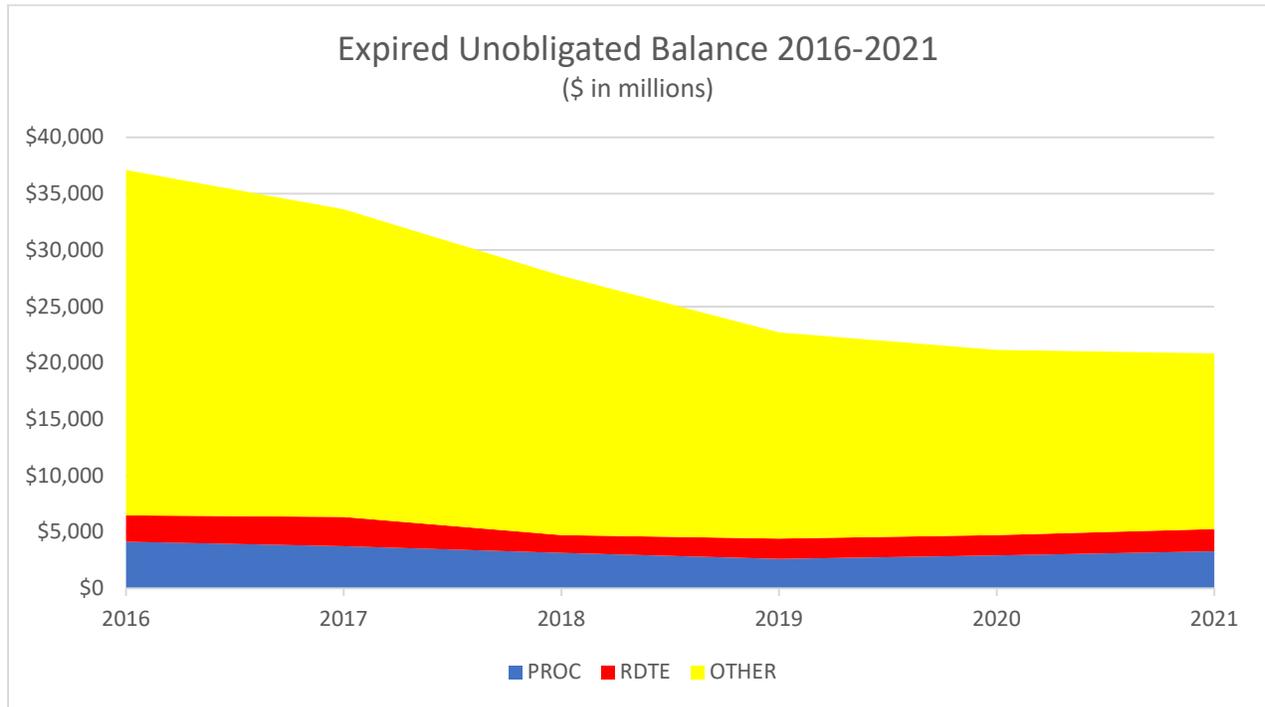
While some will suggest that modernizing the benchmark profile will create an increased amount of expired funds, the author believes that a modernized set of benchmarks—combined

with an increased leader focus across the whole life of the appropriations—will have the opposite effect. Allowing the acquisition workforce to use more of the life allowed by the congressionally appropriated multi-year funding, rather than restricting it to an unrealistic set of benchmarks, will enable the DAE to perform their duty while making the best business decisions for our nation. Finally, and most importantly, when the benchmarks cease to be the measure that acquisition efforts are measured by, the DAE will turn back to what matters for the defense of our nation, the delivery of innovative capabilities to the warfighter.

Appendix A

Current OUSD(C) Rule-of-Thumb Acquisition Obligation and Expenditure Rates

		RDT&E		Procurement		O&M		MILCON	
	Month	Obl.	Exp.	Obl.	Exp.	Obl.	Exp.	Obl.	Exp.
First Year of Availability	Oct	7.5%	4.6%	6.7%	N/A	8.3%	6.3%	5.4%	1.2%
	Nov	15.0%	9.2%	13.3%	N/A	16.7%	12.5%	10.8%	2.3%
	Dec	22.5%	13.8%	20.0%	N/A	25.0%	18.8%	16.3%	3.5%
	Jan	30.0%	18.3%	26.7%	N/A	33.3%	25.0%	21.7%	4.7%
	Feb	37.5%	22.9%	33.3%	N/A	41.7%	31.3%	27.1%	5.8%
	Mar	45.0%	27.5%	40.0%	N/A	50.0%	37.5%	32.5%	7.0%
	Apr	52.5%	32.1%	46.7%	N/A	58.3%	43.8%	37.9%	8.2%
	May	60.0%	36.7%	53.3%	N/A	66.7%	50.0%	43.3%	9.3%
	Jun	67.5%	41.3%	60.0%	N/A	75.0%	56.3%	48.8%	10.5%
Second Year of Availability	Jul	75.0%	45.8%	66.7%	N/A	83.3%	62.5%	54.2%	11.7%
	Aug	82.5%	50.4%	73.3%	N/A	91.7%	68.8%	59.6%	12.8%
	Sep	90.0%	55.0%	80.0%	N/A	100.0%	75.0%	65.0%	14.0%
	Oct	90.8%	57.9%	80.8%	N/A	100.0%	77.1%	67.1%	18.1%
	Nov	91.7%	60.8%	81.7%	N/A	100.0%	79.2%	69.2%	22.2%
	Dec	92.5%	63.8%	82.5%	N/A	100.0%	81.3%	71.3%	26.3%
	Jan	93.3%	66.7%	83.3%	N/A	100.0%	83.3%	73.3%	30.3%
	Feb	94.2%	69.6%	84.2%	N/A	100.0%	85.4%	75.4%	34.4%
	Mar	95.0%	72.5%	85.0%	N/A	100.0%	87.5%	77.5%	38.5%
Third Year of Availability	Apr	95.8%	75.4%	85.8%	N/A	100.0%	89.6%	79.6%	42.6%
	May	96.7%	78.3%	86.7%	N/A	100.0%	91.7%	81.7%	46.7%
	Jun	97.5%	81.3%	87.5%	N/A	100.0%	93.8%	83.8%	50.8%
	Jul	98.3%	84.2%	88.3%	N/A	100.0%	95.8%	85.8%	54.8%
	Aug	99.2%	87.1%	89.2%	N/A	100.0%	97.9%	87.9%	58.9%
	Sep	100.0%	90.0%	90.0%	N/A	100.0%	100.0%	90.0%	63.0%
	Oct	100.0%	90.8%	90.8%	N/A	100.0%	100.0%	90.4%	65.5%
	Nov	100.0%	91.7%	91.7%	N/A	100.0%	100.0%	90.8%	68.1%
	Dec	100.0%	92.5%	92.5%	N/A	100.0%	100.0%	91.3%	70.6%
Fourth Year of Availability	Jan	100.0%	93.3%	93.3%	N/A	100.0%	100.0%	91.7%	73.2%
	Feb	100.0%	94.2%	94.2%	N/A	100.0%	100.0%	92.1%	75.7%
	Mar	100.0%	95.0%	95.0%	N/A	100.0%	100.0%	92.5%	78.3%
	Apr	100.0%	95.8%	95.8%	N/A	100.0%	100.0%	92.9%	80.8%
	May	100.0%	96.7%	96.7%	N/A	100.0%	100.0%	93.3%	83.3%
	Jun	100.0%	97.5%	97.5%	N/A	100.0%	100.0%	93.8%	85.9%
	Jul	100.0%	98.3%	98.3%	N/A	100.0%	100.0%	94.2%	88.4%
	Aug	100.0%	99.2%	99.2%	N/A	100.0%	100.0%	94.6%	91.0%
	Sep	100.0%	100.0%	100.0%	N/A	100.0%	100.0%	95.0%	93.5%

Appendix B

Appendix C

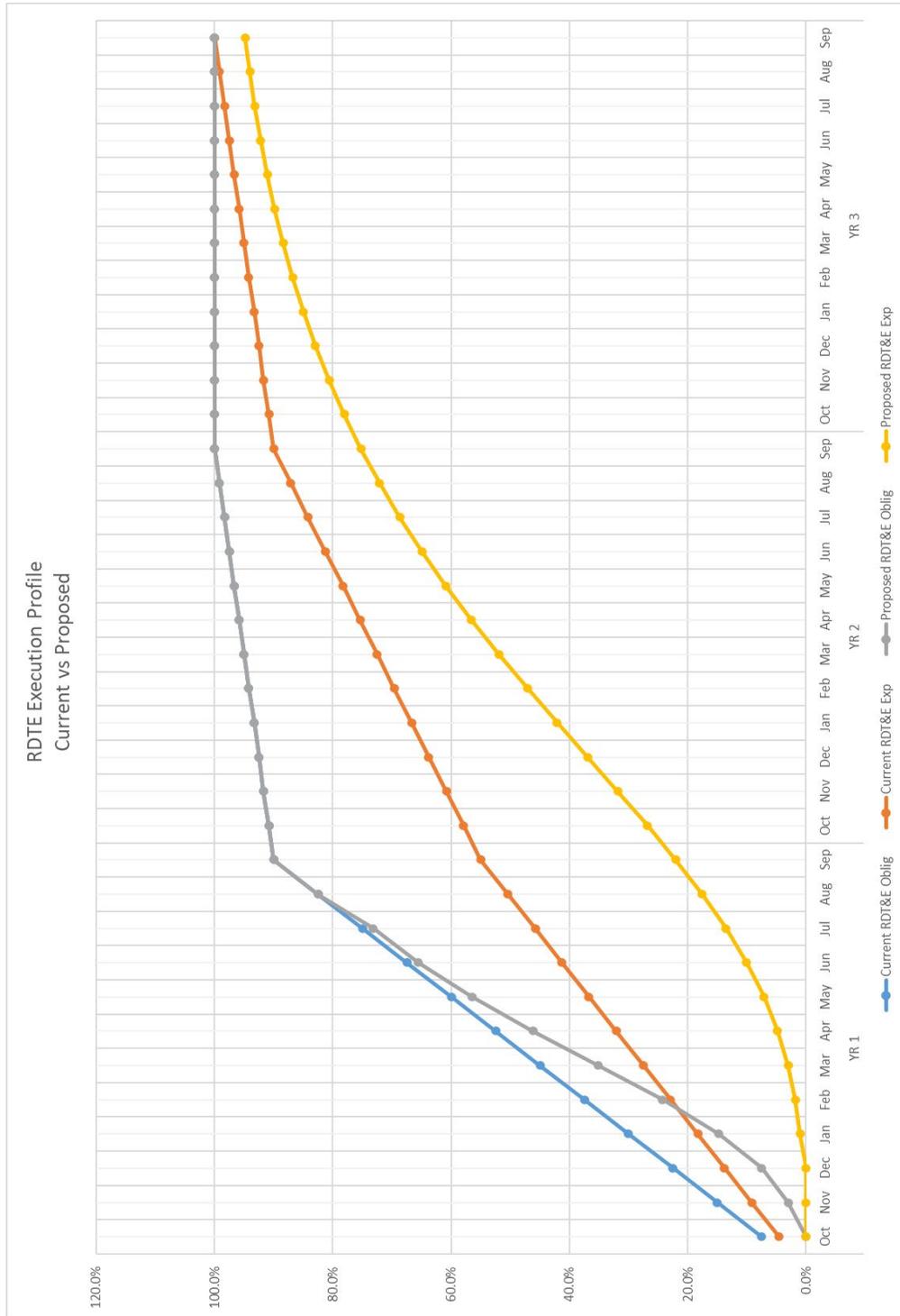
Current versus Proposed

OUSD(C) Rule-of-Thumb Acquisition Obligation and Expenditure Rates

Year of Avail	Month	Current RDT&E		Proposed RDT&E		Current PROC		Proposed PROC	
		Oblig	Exp	Oblig	Exp	Oblig	Exp	Oblig	Exp
YR 1	Oct	7.5%	4.6%	0.0%	0.0%	6.7%		0.0%	
	Nov	15.0%	9.2%	3.0%	0.0%	13.3%		2.0%	
	Dec	22.5%	13.8%	7.5%	0.0%	20.0%		5.3%	
	Jan	30.0%	18.3%	14.7%	1.0%	26.7%		11.1%	
	Feb	37.5%	22.9%	24.3%	1.8%	33.3%		19.2%	
	Mar	45.0%	27.5%	35.1%	3.1%	40.0%		29.0%	
	Apr	52.5%	32.1%	46.2%	4.8%	46.7%		39.5%	
	May	60.0%	36.7%	56.5%	7.1%	53.3%		49.8%	
	Jun	67.5%	41.3%	65.6%	10.1%	60.0%		59.3%	
	Jul	75.0%	45.8%	73.2%	13.6%	66.7%		67.6%	
	Aug	82.5%	50.4%	82.5%	17.6%	73.3%		74.5%	
Sep	90.0%	55.0%	90.0%	22.1%	80.0%		80.2%		
YR 2	Oct	90.8%	57.9%	90.8%	26.8%	80.8%		80.8%	
	Nov	91.7%	60.8%	91.7%	31.9%	81.7%		81.7%	
	Dec	92.5%	63.8%	92.5%	37.0%	82.5%		82.5%	
	Jan	93.3%	66.7%	93.3%	42.1%	83.3%		83.3%	
	Feb	94.2%	69.6%	94.2%	47.1%	84.2%		84.2%	
	Mar	95.0%	72.5%	95.0%	51.9%	85.0%		85.0%	
	Apr	95.8%	75.4%	95.8%	56.5%	85.8%		85.8%	
	May	96.7%	78.3%	96.7%	60.9%	86.7%		86.7%	
	Jun	97.5%	81.3%	97.5%	64.9%	87.5%		87.5%	
	Jul	98.3%	84.2%	98.3%	68.7%	88.3%		88.3%	
	Aug	99.2%	87.1%	99.2%	72.1%	89.2%		89.2%	
Sep	100.0%	90.0%	100.0%	75.3%	90.0%		90.0%		
YR 3	Oct	100.0%	90.8%	100.0%	78.1%	90.8%		90.8%	
	Nov	100.0%	91.7%	100.0%	80.6%	91.7%		91.7%	
	Dec	100.0%	92.5%	100.0%	82.9%	92.5%		92.5%	
	Jan	100.0%	93.3%	100.0%	85.0%	93.3%		93.3%	
	Feb	100.0%	94.2%	100.0%	86.8%	94.2%		94.2%	
	Mar	100.0%	95.0%	100.0%	88.4%	95.0%		95.0%	
	Apr	100.0%	95.8%	100.0%	89.8%	95.8%		95.8%	
	May	100.0%	96.7%	100.0%	91.1%	96.7%		96.7%	
	Jun	100.0%	97.5%	100.0%	92.2%	97.5%		97.5%	
	Jul	100.0%	98.3%	100.0%	93.2%	98.3%		98.3%	
	Aug	100.0%	99.2%	100.0%	94.0%	99.2%		99.2%	
Sep	100.0%	100.0%	100.0%	94.8%	100.0%		100.0%		

Appendix D

Current versus Proposed Benchmarks



 Endnotes

- ¹ Roberta Tomasini, “OUSD Rule-of-Thumb Acquisition Obligation and Expenditure Rates” (Defense Acquisition University, May 2017), [https://www.dau.edu/tools/Lists/DAUTools/Attachments/292/OSD%20\(C\)%20Color%20Rule-of-Thumb%20Acq%20Obligation%20and%20Expenditure%20Rates.pdf](https://www.dau.edu/tools/Lists/DAUTools/Attachments/292/OSD%20(C)%20Color%20Rule-of-Thumb%20Acq%20Obligation%20and%20Expenditure%20Rates.pdf).
- ² Gregory E. Brown et al., “Time Phasing Aircraft R&D Using the Weibull and Beta Distributions,” *Journal of Cost Analysis and Parametrics* 8, no. 3 (2015): 150–64, <https://doi.org/10.1080/1941658X.2015.1096219>; Mark A. Gallagher and David A. Lee, “Final-Cost Estimates for Research & Development Programs Conditioned on Realized Costs,” *Military Operations Research* 2, no. 2 (1996): 51–65; David A. Lee, Michael R. Hogue, and Mark A. Gallagher, “Determining a Budget Profile from a Development Cost Estimate” (Washington, DC: Office of the Secretary of Defense, Director of Program Analysis and Evaluation, September 1, 1993), <https://apps.dtic.mil/sti/citations/ADA275864>.
- ³ Robert D. Behn, “Why Measure Performance? Different Purposes Require Different Measures,” *Public Administration Review* 63, no. 5 (October 2003): 599, <https://doi.org/10.1111/1540-6210.00322>.
- ⁴ “2021 DoD Agency Financial Report (AFR) / DoD Performance and Accountability Report (PAR),” Under Secretary of Defense (Comptroller), February 9, 2022, <https://comptroller.defense.gov/ODCFO/afr/>.
- ⁵ “United States Department of Defense Agency Financial Report for Fiscal Year 2016: Financial Section” (Washington, DC: Office of the Deputy Chief Financial Officer, 2016), 137, https://comptroller.defense.gov/Portals/45/Documents/afr/fy2016/3-Financial_Section_Final20170511_HH.pdf.
- ⁶ “Financial Section and Message from the Under Secretary of Defense (Comptroller) / Chief Financial Officer” (Washington, DC: Office of the Deputy Chief Financial Officer, 2021), 236, <https://comptroller.defense.gov/odcfo/afr2021.aspx>.
- ⁷ Brown et al., “Time Phasing Aircraft R&D Using the Weibull and Beta Distributions.”
- ⁸ Robert L Tremaine and Donna J Seligman, “OSD’s Obligation and Expenditure Rate Goals,” *Defense AT&L* July-August 2013 (2013): 17.
- ⁹ “DAU Acquipedia: Procurement Lead Time,” Defense Acquisition University, Procurement Lead Time, accessed February 21, 2022, <https://www.dau.edu/acquipedia/pages/ArticleContent.aspx?itemid=348>.
- ¹⁰ James E. (MG Simpson USA), “FY18 Procurement Action Lead Time (PALT) Metric” (U.S. Army Contracting Command, March 12, 2018), 1, https://acc.army.mil/contractingcenters/acc-orl/palt_files/paltsession/ACCFY18PALTMemo12Mar18.pdf.
- ¹¹ Behn, “Why Measure Performance? Different Purposes Require Different Measures,” 594.
- ¹² Jerry Z. Muller, *The Tyranny of Metrics* (Princeton: Princeton University Press, 2018), 20.
- ¹³ Muller, 19.
- ¹⁴ Marilyn Strathern, “‘Improving Ratings’: Audit in the British University System,” *European Review* 5, no. 3 (1997): 308, [https://doi.org/10.1002/\(SICI\)1234-981X\(199707\)5:3<305::AID-EURO184>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1234-981X(199707)5:3<305::AID-EURO184>3.0.CO;2-4).
- ¹⁵ Robert F. Hale and Frank Kendall, “Department of Defense Management of Unobligated Funds; Obligation Rate Tenets” (Department of Defense, September 10, 2012), 2,

https://www.acq.osd.mil/fo/docs/OSD%20Memo_DoD%20Mgt%20of%20Unobligated%20Funds_Obligation%20Rate%20Tenets_10Sep12.pdf.

¹⁶ Muller, *The Tyranny of Metrics*, 20.