CSIS-AM-20-162

September 2020

Assessing the Reliability of the **Future Years Defense Program and Building a Forecast**

Author

Andrew Hunter **Greg Sanders** John Severini

Contributing Authors

Chitrakshi Bhardwaj Gabriel Coll Nidal Morrison **Cuong Nguyen** Gerhard Ottehenning

A Report of the International Security Program

CSIS CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

Assessing the Reliability of the Future Years Defense Program and Building a Forecast

Andrew Hunter – is a senior fellow in the International Security Program and director of the Defense-Industrial Initiatives Group at CSIS. From 2011 to 2014, he served as a senior executive in the Department of Defense, serving first as chief of staff to undersecretaries of defense (AT&L) Ashton B. Carter and Frank Kendall, before directing the Joint Rapid Acquisition Cell. From 2005 to 2011, Mr. Hunter served as a professional staff member of the House Armed Services Committee. Mr. Hunter holds an M.A. degree in applied economics from the Johns Hopkins University and a B.A. in social studies from Harvard University. He can be reached at CSIS, 1616 Rhode Island Ave., NW, Washington, DC 20036; PH #202-775 3128, and <u>ahunter@csis.org</u>.

Greg Sanders - is a fellow in the International Security Program and deputy director of the Defense-Industrial Initiatives Group at CSIS, where he manages a research team that analyzes data on U.S. government contract spending and other budget and acquisition issues. In support of these goals, he employs SQL Server, as well as the statistical programming language R. Sanders holds an M.A. in international studies from the University of Denver and a B.A. in government and politics, as well as a B.S. in computer science, from the University of Maryland. He can be reached at CSIS, 1616 Rhode Island Ave., NW, Washington, DC 20036; PH #202-741-3916, and gsanders@csis.org.

John Severini

Contributing Authors

Chitrakshi Bhardwaj, Gabriel Coll, Cuong Nguyen, Gerhard Ottehenning

Disclaimer

The Center for Strategic and International Studies (CSIS) does not take specific policy positions; accordingly, all views expressed in this presentation should be understood to be solely those of the author(s).

Acknowledgments

This material is based upon work supported by the Acquisition Research Program under Grant No. HQ00341910010. The views expressed in written materials or publications, and/or made by speakers, moderators, and presenters, do not necessarily reflect the official policies of the Department of Defense nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government. The study team would like to thank the range of practitioners and scholars that participated in the summer 2020 workshop reviewing preliminary results. This project was enabled by multiple efforts at CSIS beyond the study team for this paper. CSIS's Defense Budget Analysis team, especially Todd Harrison and Seamus Daniels, were generous with their insights and shared data including amalgamated version of the FY 2020 and FY 2021 FYDP and other sources used to assist in validation. Gabriel Coll led a predecessor effort at FYDP accessibility, along with Loren Lipsey and Shivani Pandya, that included the collection of some of the justification book files used in this report and contributed valuable experience. That team along with Kayla Keller and Yuanjing Han also worked to develop a tool to make the data more accessible. Kyle Libby and Simone Williams also offered valuable quality checking towards that effort that informed our understanding of the sources included in this paper. Finally, the study team would like to thank their DIIG colleagues for their review and support.

Abstract

Discerning, negotiating, and communicating priorities are necessary tasks for the U.S. defense acquisition system to effectively implement its portion of the National Defense Strategy. One of the Department of Defense's central tools for doing so is the Future Years Defense Plan (FYDP), a projection of the cost and composition of the force over the next five years. However, the publicly released FYDP suffers from important limitations: there is tension between expressing Administration preferences and accurate projection; no confidence intervals or other measures of reliability are provided; predictable budget elements have been transferred beyond the scope of the FYDP; and the detailed investment projections are challenging to gather and employ. This project works to make the FYDP more accessible and more easily evaluated. It posits two hypotheses using FY 2018 budget request data: first that FYDP projections could estimate actual 2019 spending more reliably than the President's Budget alone, and second that the reliability of projections would vary between services. The simple regression model employed found that the two year out FYDP projections significantly improved the reliability of estimates for procurement line items and RDT&E program elements.

Table of Contents

| Introduction | 5 |
|--|----|
| Background | 6 |
| Planning, Programming, Budgeting, and Execution Process | 6 |
| FYDP | 7 |
| осо | 8 |
| Procurement and RDT&E | 8 |
| Literature Review | 9 |
| Bias in Estimation | 9 |
| Enduring Budget in OCO | 13 |
| International Approaches | 14 |
| Comparative Case Study: Canada | 16 |
| The Strategic and Communication Role of the FYDP | 18 |
| Scope | 20 |
| How reliable are projections within the FYDP as an indicator for actual spending? | 20 |
| Which Military Dpartments have the most and least reliable projections? | 21 |
| Data and Methods | 21 |
| Data Sources and Structure | 21 |
| Creation of the CSIS Investment Budget Line Dataset | 22 |
| Measuring Dependent and Independent Variables | 26 |
| Dependent Variable | 27 |
| Study Variables | 28 |
| Control Variables | 29 |
| Estimating Equation | 29 |
| Results | 29 |
| How reliable are FYDP projections? | 29 |
| Which services and budget categories have the most and least reliable projections? | 33 |
| Modeling FY 2018 PB Estimate of 2019 Actual Spending | 36 |
| Discussion and Conclusions | 38 |
| Bibliography | 38 |

1 Introduction

For the US defense acquisition system to properly implement its portion of the National Defense Strategy, it must effectively discern, negotiate, and communicate its priorities. One of the Department of Defense's (DoD) central tools for this process is the Future Years Defense Plan (FYDP), a projection of the cost and composition of the force over the next five years.

Annually updated and submitted as part of the President's budget submission projection, the FYDP provides important insights into DoD's priorities and projections of the future both internally and externally. Internally within DoD, wherein the FYDP is constructed, the process forces the stakeholders involved to debate tradeoffs and outline their visions of the future. Externally, it lays out for Congress a vision of how U.S. national security strategies could be implemented in practice, which the legislature must then choose whether to fund or alter. It helps the U.S. defense industry understand where DoD plans to invest and thereby allows companies within the industry to align themselves with current priorities. It helps scholars identify trends and do research on major capital-intensive projects, which can be used to inform future projects, both defense and nondefense. It helps U.S. citizens identify how the government plans to spend their taxpayers' dollars. However, the FYDP has a few major drawbacks for these stakeholders that undercut its ability to communicate priorities.

The first drawback is the inherent tension between FYDP's role expressing the funding amount that the executive branch deems necessary to support the strategy and its role in creating a plan that can be implemented within the funding amount authorized and appropriated by Congress. Most years, this has meant that the administration requests and projects more funding than is ultimately provided which can undermine its role in priority setting.

A second, related, shortfall is the absence of any measure of reliability or predictive intervals for the projections. Some parts of the DoD budget are easier to predict than others, but the point estimate provided by the FYDP does not differentiate between known quantities, like the purchase of uniforms, and cutting-edge technology, like the development of a next-generation alloy. That said, by design Overseas Contingency Operations (OCO) budgets operate as a pressure valve for uncertainty by taking some of the most volatile spending out of the base budgets and FYDP and managing them through OCO methods instead. However, the intended functionality of the OCO accounts is muddled when predictable spending is moved to OCO accounts to avoid budget caps.

Third, the unclassified FYDP is released in a form that makes it straightforward to study topline spending or individual line items or programs but challenging to analyze anything in between. This is because the FYDP is released in dozens of PDFs through separate justification books, and not as a centralized database or even in summary documents. Collectively, these limitations present a higher barrier to entry to stakeholders and make it laborious for specialists and unappealing for anyone else to put investment plans in a meaningful context. Without analysis, it is difficult to tell the difference between a figure reliably reported for years and an uninformative placeholder.

To give context to these results, the background section details how the FYDP is constructed and key related concepts: OCO budgets and the two accounts that make up investment spending, Procurement and Research, Development, Test, and Engineering (RDT&E). Building off this background, the conceptual framework outlines hypotheses regarding FYDP reliability and drivers thereof.

This report tests the value of the unclassified FYDP for investment spending, RDT&E and Procurement, as a bottom up indicator of DoD priorities by comparing cumulative projections of PB 2019 spending from one, three, and five years in advance with the actual cumulative spending. In addition, the paper models whether the FY 2018 President's Budget (FY 2018 PB) was capable of meaningfully forecasting actual spending in 2019. The FYDP was largely prepared by the prior administration and not formally released that year. Moreover, the budget request

was submitted at the end of May, the "latest a budget has been submitted to Congress since the president was first required to submit budget requests in FY 1923."¹

To address these questions and to make the FYDP more accessible and more easily evaluated, the study team, in cooperation with other researchers, has systematically imported budget data with the most complete data captured from the 2013 through 2021 President's Budgets. The Data and Methods section lays out where this data comes from, explains how it was imported and validated, and then introduces the variables and a model that examines how well the FYDP from the 2018 President's Budget predicts actual spending in 2019.

The results section includes scatter plots and histograms comparing projected and actual spending for investment line items overall and broken up into military departments. The discussion and conclusion section analysis draws out larger implications.

2 Background

Planning, Programming, Budgeting, and Execution Process

The Planning, Programming, Budgeting, and Execution (PPBE) is a DoD process to allocate resources based on strategic objectives. This process was formerly called the Planning, Programming, and Budgeting System (PPBS), established by Secretary of Defense (SECDEF) Robert S. McNamara in 1961 with the goal of connecting budget allocations with specific objectives and plans. Krieg and Chu summarize six guiding principles that underlined the creation of the process:

- "Decisions should be based on explicit criteria of national interest, not on compromises among institutional forces.
- Needs and costs should be considered simultaneously.
- Major decisions should be made by choices among explicit, balanced, feasible alternatives.
- The Secretary of Defense should have an active analytic staff to provide him with relevant data and unbiased perspectives.
- A multiyear force and financial plan should project the consequences of present decisions into the future.
- Open and explicit analysis, available to all parties, should form the basis for major decisions."²

Shaping the relationship between the SECDEF and the military departments is a critical role of the FYDP and one that is not necessarily the case in the long-term defense planning of other democratic allies. As Thomas-Duerrel Young observes: "Uniquely different from other Western ministries of defense, the U.S. Department of Defense remains a confederacy of independent organizations, and critically, each with their own jealously guarded budgetary autonomy and legally-defined institutional responsibilities and functions."³

In a typical fiscal year cycle, the PPBE process starts more than two years before the expected year of budget execution (McGarry, 2020). The first phase, planning, is led by the Under Secretary of Defense for Policy. In addition, the Chairman of the Joint Chief of Staff (CJCS) also plays an important part in this process. "The phase involves reviewing the President's National Security Strategy (NSS), the SECDEF's National Defense Strategy (NDS), and the CJCS's National Military Strategy (NMS) to ensure the resulting Defense Planning Guidance (DPG) aligns with the Administration's policy goals and takes into account potential threats, force structure, readiness posture,

¹ Harrison, Todd and Seamus Daniels (2017 December). *Analysis of the FY 2018 Defense Budget*. Center for Strategic and International Studies, IV.

http://defense360.csis.org/wp-content/uploads/2017/12/FY-2018_DefenseBudgetAnalysis_FULL_WEB.pdf

² Kenneth J. Krieg, David S.C. Chu (2005) "Foreword: How Much is Enough", ix-x.

³ Thomas-Duerrel Young, "Questioning the "Sanctity" of long-term defense planning as practiced in Central and Eastern Europe," 361.

and other factors" (McGarry, 2020, pp. 1). The planning phase focuses on reviewing threats and assessing capabilities but is not constrained by expected resource levels.

The programming phase is executed by the services in coordination with the office of Cost Assessment and Program Evaluation (CAPE). The main focus of this phase is compliance with DPG. Being more constrained by resource and fiscal considerations, the programming phase is tasked with turning DPG into achievable and affordable programs. To do so, the heads of the military departments are charged with creating a Program Objective Memorandum (POM). These POMs cover five years of resource requirements and are reviewed and updated by CAPE, with any changes made via Resource Management Decisions.

The third phase of the PPBE process, budgeting, is led by the Under Secretary of Defense (Comptroller). The budgeting stage focuses on preparing an "executable and defensible budget" with input from the military services (Herbert, 2011, pp. 27). The Comptroller, under OMB guidance, reviews estimates for the FYDP's first year. Those results are then reviewed by the SECDEF with help from Comptroller analysts, and, after any changes are resolved with the military services, submitted to OMB (McGarry, 2020).

OMB works on the budget during the winter months and the President's Budget is typically transmitted to Congress in February (McGarry, 2020). Congress in turn possesses the power of the purse and may choose to change the amounts in the budget bills. The potential for disagreement between the two branches is an important limitation on projection reliability. This phase ends when the President signs the authorization and appropriation bills into law. If the appropriation bill is not signed into law by October 1, Congress may pass a continuing resolution (CR) which usually locks in pre-existing spending levels and prevents new starts. If neither the appropriation bill nor CR is enacted, the government would go into a temporary shutdown, although even in a shutdown "[normally]... DOD continues minimum essential operations based on national defense requirements" (U.S. Department of the Army, 2016, pp. 24). The occurrence of CRs and shutdowns does not bode well for effective projection, budgeting, or implementation.

The final phase, execution, is primarily focused on the implementation of the strategy using the funds made available by the budget. The program results are reviewed during this phase, "develop[ing] performance metrics, measur[ing] performance against plan, assess[ing] actual output against planned performance, and adjust[ing] resources to achieve desired performance goals (Hebert, 2011, pp. 28). These execution reviews are conducted in parallel with program and budget reviews to allow this information to feed back into prioritization and spending decisions (McGarry, 2020).

FYDP

The FYDP projects resource needs over a five-year period, which include the President's Budget and the next four years, also called the out years. The FYDP is conducted annually with a rolling five-year window. This aligns with the federal-government wide requirement in 31 U.S. Code § 1105 to report "estimated expenditures and proposed appropriation the President decides are necessary to support the Government in the fiscal year for which the budget is submitted and the 4 fiscal years after that year." The DoD's process is particularly robust and Enthoven and Smith succinctly describe its original purpose:

A decision by the Secretary of Defense to develop, procure, or operate a weapon system affects not only the current defense budget but future budgets as well, the latter far more than the former as a rule. When he decides to begin the engineering development of a new system, with procurement presumably to follow, he initiates a stream of expenditures which can eventually include development, procurement, and operating and maintenance costs of the completed system. He needs not only a record of current costs and manpower but also projections of this information far enough ahead to enable him to estimate the main consequences of today's decisions. The Five-Year Defense Plan, or FYDP, was developed to provide this record.⁴

Within the PPBE process it is typically created in the programming phase and updated to reflect decisions made during the budgeting phase. The FYPD database illustrates programs and plans by components (military service or defense agency), Major Force Programs, and appropriation titles (military personnel, procurement, RDT&E, operations and maintenance, etc.) (McGarry & Peters, 2020). The FYDP database itself is not publicly released, but, since 1989, significant portions are unclassified and published as part of the budget release (Mehta, 2020). The form of the unclassified FYDP numbers varies from between the appropriation titles, but consistently includes actual and enacted spending totals for the two years leading up to the budget in addition to the five years of projections and thus covering a total of seven years in each release. As a result, the information released from the FYDP is a vital tool to evaluate the growth or decline of various programs within DoD.

When considering the data included in the unclassified release of the FYDP, a few areas are worth special consideration. First, while OCO spending is excluded from the four years of FYDP projections past the President's Budget, it is important to understand the category of spending to make sense of the larger budget context. Second, projections are especially important for those appropriation titles with longer time frames, RDT&E and Procurement. These investment budgets are the focus of this paper and their relationship is key to understanding the lifecycle of programs working through the FYDP.

ОСО

OCO is funding set aside in the federal budget to cover marginal expenses for overseas operations including conflict and other emergencies. This source of funding helps cover the costs of equipment and maintenance over and above the course of normal operations.⁵ Given this emphasis, OCO spending is inherently volatile and dependent on the state of the larger world. As a result, the elements of the budget covered by OCO would be inherently difficult to estimate regardless of the budgetary process used. OCO funding is exempt from most of the PPBE Multi-Year Budgeting process and is excluded from the FYDP projections in the four year beyond the President's Budget. This division has the advantage of easing the process of creating a reliable FYDP by excluding some of the least predictable elements of the defense budget.

Procurement and RDT&E

Projections are especially important to investment spending, which in the U.S. system is covered under the RDT&E and Procurement appropriation titles. Taken together RDT&E and Procurement accounted for an average of over thirty percent (12 percent and 19 percent respectively) of DoD budgets between 2001 and 2017 (Mann, 2017).

Prior to moving to the procurement phase, programs are often considered in development and funded through the Research, . Spending is divided across seven Budget Activities, from Basic Research (6.1) to operational Systems Development (6.7).⁶ These budget activities largely track the lifecycle of technology maturation:

"Funding in codes 6.1 to 6.3 is referred to by DoD as the science and technology (S&T) budget. This portion of DoD RDT&E is often singled out for attention by analysts, as it is seen as the pool of knowledge necessary for the development of future military systems. In contrast, 6.4, 6.5, and 6.7 funds are focused on the application of existing scientific and technical knowledge to meet current or near-term operational

⁴ Enthoven and Smith RAND_CB403 How much is enough, 48.

⁵ Andrew Hunter (2019) outlines an example: "the base budget pays the salaries of an Army unit and its normal organizational and training expenses, but OCO pays the additional marginal cost of transporting the unit overseas for operations, the costs of fueling and resupplying the unit while deployed, and special pays associated with deploying the unit such as hostile fire/imminent danger pay."

⁶ Justin Doubleday (2020) reported earlier this year that the DoD was considering an eighth budget activity to cover software. Like 6.6 funding, this classification would be more focused on the type of work being done rather than the stage of research.

needs. The funds in 6.6 are for RDT&E management and support work in any of the other RDT&E." (Sargent, 2019, pp. 1-2)

RDT&E programs that mature past the Engineering and Manufacturing Development Phase or other items with less need for development may then be purchased under the Procurement appropriation title. The DoD procurement appropriations title "provides funds for non-construction related investment costs—the costs to acquire capital assets, such as an F-35 Joint Strike Fighter aircraft or a Virginia class submarine" (McGarry & Peters, 2020, pp. 1). Under the policy of full funding, "the total funding necessary to acquire a usable end item is approved by Congress in a single fiscal year, even though related work may span many years" (McGarry & Peters). Incremental funding is another annual option, one that divides the system's cost over multiple annual increments. However, this approach is largely limited to ships and submarines. A third approach, oriented towards savings per unit, is Multiyear Procurement (MYP). The prerequisites for Congress authorizing MYP include "significant savings, stable funding and design, and other standards" (McGarry & Peters, 2020, pp. 2). Thus, MYP is at the opposite end of the predictability scale from OCO spending, as the costs and funding sources should both be well understood. The government gives up flexibility under MYP by making advance commitments and payments in exchange for lower total costs.

This discussion of full funding, incremental funding, and MYP is focused at the system level, but another way to break up procurement over multiple years is <u>by procuring individual to procurement</u> components of a larger system rather than the entire platform. Advanced procurement applies to MYP but also is relevant to a larger swath of component purchases for major programs that may not meet the MYP criteria (McGarry & Peters, 2020). While the specifics vary, FYDP projects are especially relevant to, and should be informed by, any alternative to single year full funding.

3 Literature Review

There are a multitude of challenges in defense planning even within the base budget. The United States, despite its resources and robust analytical staff, faces more difficult challenges than those of many of its peer countries. First, the United States is a Presidential system with projections prepared by the executive branch, but funding authority rests with Congress. A projection process could be designed that does more to incorporate Congressional opinions into the planning process, but the role of two co-equal branches of government means that some degree of uncertainty for both the topline funding and for individual projects of interest to the legislature is irreducible.

Second, the United States is the global leader in defense research and development, and as Light, Leonard, Pollak, Smith, and Wallace (2017) find, "there is a considerable amount of cost and schedule growth risk facing all [Major Defense Acquisition Programs] at [Milestone] B" when the DoD commits to significant development spending.⁷ While the RAND corporation authors did believe that further advances in estimating were worth investing in, they recommend that their modeling approach would be suitable to "assessing overall portfolio risk and informing risk mitigation planning."⁸

Bias in Estimation

However, while there are complexities adding uncertainty, analysts have found that projections have patterns in mistakes. Over the past decades, a diverse group of scholars have found that project estimates tend to expect projects will be cheaper than they are in reality, resulting in some combination of cost growth of scope decreases. In 1980 Franklin C. Spinney argued that "There is a systematic tendency to underestimate future costs;"in 2006 Arena, Leonard, Murray, and Younossi found a "systematic bias towards underestimating the costs" of MDAP programs with development estimates (MS B) faced an average of 46 percent cost growth and production (MS C)

 ⁷ Light, Leonard, Pollak, Smith, and Wallace (2017) Quantifying Cost and Schedule Uncertainty for Major Defense Acquisition Programs (MDAPs). Santa Monica, CA: RAND Corporation., 44 https://www.rand.org/pubs/research_reports/RR1723.html
 ⁸ Light, Leonard, Pollak, Smith, and Wallace (2017), 45.

estimates faced 16 percent.⁹ In 2017 Light, Leonard, Pollak, Smith, and Wallace (2017) find that for MDAPs that "MS B [developmental] estimates tend to be optimistic, with most programs likely to experience some level of cost and schedule growth."¹⁰

Bias suggests that better estimates are possible. Indeed, McNicol (2004) found that the addition of reforms introducing independent cost estimates were associated with lower levels of cost growth.¹¹ Focusing on independent cost estimates, McNicol, Tyson, Hiller, Could and Minix found that independent estimates correctly identified underestimates in military service estimates 80 percent of the time though they had a 25 percent "false alarm" rate in a different subset of cases.¹² While uncertainty in defense projection is a constant around the world, other countries have also been better able to avoid optimistic bias. Ethan B. Kapstein and Jean-Michel Oudot in 2009 found that after two decades of reform, in France "the cost overruns that result tend to be relatively minor in scope; on the order of 5-10 percent per weapons platform, versus an average overrun of 26 percent per platform in the United States."¹³ Canada provides another example of more effective estimates, as discussed in the International Approaches section below.

The prior set of acquisition reforms, including the Weapon System Acquisition Reform Act of 2009 and the Pentagon driven Better Buying Power initiatives were both focused on reducing cost growth. It will still be years until it is possible to evaluate whether programs instituted after these reforms did systematically better to control cost, but McCormick, Cohen, Sanders, and Hunter in 2019 summarized other reporting and conducted an analysis on contract cost growth and found "these results and those of the GAO, the evidence suggests that improvements in cost control under BBP were real and cost control efforts should be maintained even as attention moves on to other acquisition priorities."¹⁴

David McNicol, who has long studies the issue of cost growth, is more skeptical of the affect of acquisition reform efforts after the aforementioned introduction of independent cost estimates. He does offer a possible apologia for optimistic bias, based on his finding unit cost estimates are shaped less by acquisition reform and instead are better when budgets are rising than when they fall:

"This explanation also is more informative in that it draws attention to the possibility that "unreasonably" optimistic MS B baselines perhaps were a rational response to the circumstances in which they arose. Instances of extremely high cost growth probably cannot be waved away on that basis. It is necessary, however, to be careful about the extent to which the DoD acquisition process creates problems—cost growth, schedule slips, performance shortfalls—and the extent to which it provides reasonable accommodation to inconsistencies between funding and force structure and missions."¹⁵

 ⁹ Arena, Mark V., Robert S. Leonard, Sheila E. Murray, and Obaid Younossi. Historical Cost Growth of Completed Weapon System Programs. Santa Monica, CA: RAND Corporation, 2006, xi. <u>https://www.rand.org/pubs/technical_reports/TR343.html</u>
 ¹⁰ Light, Leonard, Pollak, Smith, and Wallace (2017), 40-41.

¹¹ McNicol, David, (2004) "Cost Growth in Major Weapon Procurement Programs," P-3832. Alexandria, VA: Institute for Defense Analysis, 47.

¹² McNicol, D., Tyson, K., Hiller, J., Cloud, H. and Minix, J. (2005) The Accuracy of Independent Estimates of the Procurement Costs of Major Systems. P-3989. Alexandria, VA: Institute for Defense Analysis, 27. https://www.ida.org/research-andpublications/publications/all/t/th/the-accuracy-of-independent-estimates-of-the-procurement-costs-of-major-systems

¹³ They attribute the success to "hard budget constraints" and acquisition approaches that share some elements in common with reform efforts underway in the United States when the paper was being published. Ethan B. Kapstein and Jean-Michel Oudot (2009 "Reforming Defense Procurement: Lessons from France", 2. Hard financial constraints imply a willingness to sacrifice other goals of the defense acquisition system to manage costs.

¹⁴ McCormick, R., Cohen, S., Sanders, G., Hunter, A. (2019) "Acquisition Trends, 2018: Defense Contract Spending Bounces Back." Washington, D.C.: CSIS, 92. https://www.csis.org/analysis/acquisition-trends-2018-defense-contract-spending-bouncesback

¹⁵ McNicol, David (2018) Acquisition Policy, Cost Growth, and Cancellations of Major Defense Acquisition Programs. R-8396. Institute for Defense Analysis. <u>https://www.ida.org/research-and-publications/publications/all/a/ac/acquisition-policy-cost-growth-and-cancellations-of-major-defense-acquisition-programs</u>, 59.

In terms of reconciling the inconsistencies between funding and strategy, Steven Aftergood observed that one benefit of the publication of the unclassified FYDP, even with all the aforementioned flaws, is that it enables independent analysis:

Without an unclassified FYDP, Congress and the public would be deprived of unclassified analyses like <u>"Long-Term Implications of the 2020 Future Years Defense Program"</u> produced last year by the Congressional Budget Office. Other public reporting by <u>GAO</u>, <u>CRS</u>, the <u>news media</u> and <u>independent</u> <u>analysts</u> concerning the FYDP and future defense spending would also be undermined.¹⁶

In analyzing PB 2021, David Author and Matthew Woodward draw on the FYDP and project specific reporting. Employing historical factors developed from studies by the RAND Corporation and the Institute for Defense Analysis including those referenced above, they and find that "[u]sing the resulting cost estimates instead of DoD's cost estimates raises total projected acquisition costs by 3.5 percent over the FYDP period and by 6.1 percent over the 2026–2035 period."¹⁷

The relationship between these project estimates and the larger FYDP projections is complex with estimated cost influencing budget requests and a sense of total available funds influencing what the budget is able to fund and at what level. For example, the CBO estimates explore the cost implications of the President's Budget by keeping present plans constant. In practice, MDAPs and other budget lines can be descoped, slowed down, or cancelled outright. There are also further steps that could be taken to insulate the overall budget from uncertainty or bias in individual programs, for example, Leonard, Murray, and Younossi recommend as a possible risk mitigation measure that "funds might be held in reserve to address anticipated further growth within an MDAP portfolio (rather than for any single program)."¹⁸

Even without insulation in place, topline FYDP projections should not be treated simply as a sum of composite budget lines but as a consequential form of estimating in its own right. As Todd Harrison and Seamus Daniels note:

"Previous inflection points in the defense budget, both up and down, have been influenced by wars, shifts in strategy, changes in the threat environment, and economic conditions. . . "The conditions that led to deficit-driven defense downturns in the past—namely, a rapidly rising and record-setting deficit—are occurring at present. While there does not appear to be appetite to cut the defense budget in the remainder of FY 2020 or in FY 2021, as is evident by the inclusion of additional defense funding in stimulus bills, the political environment could shift markedly once an economic recovery is underway in FY 2022 or FY 2023."¹⁹

While the pandemic makes the present particularly challenging to forecast, the problem has never been an easy one for DoD. Writing in 1994, Kevin Lewis finds that while plans routinely expect small incremental changes but in practice changes, in topline spending or individual programs, are regularly more dramatic and often cyclical.²⁰ Leland Jordan in 2015 goes further and argues that historically most administrations project more funding than

¹⁶ Aftergood, Steven (2020, May 30) *Pentagon Asks to Keep Future Spending Secret.* Federation of American Scientists. https://fas.org/blogs/secrecy/2020/03/ndaa-fydp/

¹⁷ Woodward. M. and Arthur, D. "Long-Term Implications of the 2021 Future Years Defense Program, Washington, D.C.: Congressional Budget Office, 2019, p. 16. https://www.cbo.gov/publication/56526

¹⁸ Light, Leonard, Pollak, Smith, and Wallace (2017) Quantifying Cost and Schedule Uncertainty for Major Defense Acquisition Programs (MDAPs), 44.

¹⁹ Harrison, Todd and Seamus Daniels (2020). *Analysis of the FY 2020 Defense Budget and Its Implications for FY 2021 and Beyond*. Center for Strategic and International Studies, 57-58. <u>https://csis-prod.s3.amazonaws.com/s3fs-public/publication/HarrisonDaniels</u> FY20DBA v4.pdf?FA48YSbbC3Vj.aH7pKhpdIbU56OQ9QUZ

²⁰ Lewis, Kevin (1994) "The Discipline Gap and Other Reasons for Humility and Realism in Defense Planning" in New Challenges for Defense Planning: Rethinking How Much Is Enough. Santa Monica, CA: RAND Corporation, 110-113. https://www.rand.org/pubs/monograph reports/MR400.html.

materializes, showing "systematic fiscal optimism."²¹ Jordan analyzed budgets from 1975 to 1995 and discovered that 70% of the projections exceeded the appropriated amounts.²²

This optimism is a problem for the acquisition system because it undercuts what Enthoven and Smith describe as key to the FYDP's value: the way the FYPD forces "the Secretary to make controversial decisions explicitly."²³ Single year budgeting can allow for the buildup of a larger portfolio of programs than can be affordably built and maintained means that when something goes wrong, there is not a reserve of resources to address the problem. Effective projection should force difficult choices in the present, before expensive commitments have been made, rather in the future, where cost overruns or budget shortfalls may terminate programs experiencing difficulty or raiding funds from programs that are presently successful. This can also become a problem for the defense industry. Firms investing based on soaring spending projections may quickly go into debt and vendors that lose faith based on projections that rarely materialize, may be unprepared to meet demand when higher budgets do come. Leland concludes that "those administrations having demonstrated the greatest bias in their real growth projections also most seriously handicapped program managers."²⁴

Harrison and Daniels have conducted a more recent analysis as seen in Figure 1, which shows in black actual base budget authority going back to the projections and actuals for fiscal year 1980 (FY80). The FYDP projections are shown by dashed lines, with their color varying by the administration responsible for putting together the FYDP.²⁵ When the dashed line is above the black line, the FYDP is estimating more resources than will actually be available. There is a continued tendency for administrations to overshoot or undershoot subsequent funding, although sequestration and the budget control have increased the importance of OCO spending to topline discussions.

²¹ Jordan, Leland G. (2015). Systematic Fiscal Optimism in Defense Planning. *Defense ARJ*, 22(3), 274. <u>https://apps.dtic.mil/dtic/tr/fulltext/u2/1001751.pdf</u>

²² Jordan, "Systematic Fiscal Optimism in Defense Planning", 2015, 282-283. Jordan found optimistic bias in topline estimates under President Reagan, President George H.W. Bush, President Clinton. President Ford was the exception with projections largely tracking with actual spending while President Carter projected below what would be subsequently appropriated.
²³ Enthoven and Smith RAND CB403 How much is enough, 52.

²⁴ Jordan, "Systematic Fiscal Optimism in Defense Planning," 2015, 288

²⁵ The process of preparing the FYDP takes long enough that when a new administration takes office in January after an election year, much of the work has been done by the prior administration, albeit with possible input from the new administration's transition team after the prior year's November election.



Figure 1 Topline DoD FYDP Projections Reported by DoD Greenbook

Source: Harrison and Daniels, 2020, pp. 11

Enduring Budget in OCO

As discussed in the background section, OCO budgets acknowledge the existence of uncertainty due to external events such as wars or other cases where events beyond the control of DoD require rapid responses. However, the distinction between the base budget and OCO has been undercut by budgetary maneuvering since the passage of the Budget Control Act of 2011. In an attempt to reduce federal budget deficits, the BCA established spending limits on discretionary budget authority, applying to both defense and non-defense programs. These limits do not apply to the OCO budget. Subsequently, the executive and legislative branches have evaded the caps by transferring some predictable enduring spending out of the base budget and OCO spending. This approach has been criticized by some defense experts and government officials from both parties. Katherine Blakeley and Lawrence Korb (2014) from the Center for American Progress voice their concerns that "[f]inancially, the free flow of war funding has decimated any pretense of fiscal discipline at the Pentagon... Unclear budget guidance and poor financial management have allowed DoD to pay for substantial enduring costs with war funding rather than the base budget, further muddying the waters." (Blakeley and Korb, 2014, o. 28). Acting White House Chief of Staff Mick Mulvaney has also criticized in strong terms the "use of OCO funding for base budget requirements" (McGarry & Epstein, 2019, pp. 9).

The Congressional Budget Office (CBO) estimates that from 2006 to 2018, more than \$50 billion in OCO funding per year (in 2019 dollars), on average, has gone toward the costs of enduring activities rather than the temporary costs of overseas operations. This is particularly transparent in the FY 2020 President's Budget, the last dashed cyan line on Figure 1, which is below current levels for two years and then leaps above current levels for the remaining three years. As Harrison and Daniels report, "the request shifted entire categories of funding, such as Army Ammunition Procurement, from the base budget into OCO" (2020, pp. 4). With the coming expiration of the BCA caps in FY 2021, the FYDP released with the FY 2020 budget request then shifts all the money back to base spending after the expiration. The zig-zag line of the FY 2020 FYDP does not inspire trust in the quality of the projections for the years where OCO spending is implied but not reported.

The insertion of enduring items within the OCO budget does not necessarily undermine the value of OCO in those cases where it is still used as intended. Andrew Hunter (2019) defends OCO by noting that in the last decade, a major portion of OCO funds was used to support the operations of Afghan security forces. As the actual size,

operational employment, and equipage of these forces has changed rapidly, OCO funds were extremely useful to sustain the mission. Due to the fact that the levels are not planned out a full five years in advance, OCO allows the performance of missions that might be practically impossible otherwise. Moreover, the origin of OCO was an attempt to bring more oversight and transparency to wartime emergency supplemental bills. Senator McCain, objecting to the inclusion of non-emergency procurement in a supplemental bill, "demanded that DoD submit its request for war funding along with the regular budget so that it might receive a similar level of congressional scrutiny as the base budget and so that it would be less easy to embellish with non-war-related funding" (Hunter, 2019). The shift of predictable spending from the FYDP to OCO accounts is contrary to the spirit of those reforms and a failure to communicate executive branch priorities to Congress, industry, and other observers.

International Approaches

While the U.S. defense budget and its supporting industry are on such a scale to make them sui generis, there are a range of international approaches to defense planning that make it easier to understand what aspects of the FYDP reflect widespread approaches and what aspects are differentiating characteristics of the U.S. approach.

For decades, the Organization for Economic Cooperation and Development (OECD) has surveyed its member states on their multi-year budgeting practices. The latest survey, undertaken between December of 2017 and July of 2018, is summarized in figure 2 below. An observation from the 2014 survey still holds true: "the dominant trend among OECD countries is multi-year expenditure ceilings to cover between three and four years." In the latest survey, only Canada, Iceland, and the Republic of Korea joined the United States in employing a five-year window.



Figure 2 Multi-Year Budgeting Practices in OECD Countries

Source: OECD, Budgeting and Public Expenditures in OECD countries 2019.

While the time horizons for other developed nations are often shorter, Todor Tagarev (2016) finds that "[m]any NATO members and [Partnership for Peace] countries, influence by the experience of the United States since the early 1960s, have implemented resource management systems in which plans are linked to budgets through *programs.*"²⁶ The structure of Defense program vary between countries but they share the trait of attempting to capture the range of expenses that relate to a capability and of tying longer term polices to shorter term budgets.

In contrast to Tagarev, Thomas-Duerrel Young is skeptical of the way other nations have followed the U.S. model and provides a critical review of Eastern European planning efforts, not for insufficient adherence to the U.S. model but instead for overly hewing to NATO's gospel of long-term defense planning. Citing examples of

²⁶ Todor Tagarev (2006) "Introduction to Program-Based Defense Resource Management" The Quarterly Journal. Connections: The Quarterly Journal 5, no. 1, 56. http://procon.bg/ru/system/files/05.1.05_tagarev_0.pdf

"unexecutable" plans across multiple nations Thomas-Duerrel Young argues that "focusing on producing the "plan", officials have abjured what should be their first responsibility as planners: to make defense fit the existing budget."²⁷

Comparative Case Study: Canada

The distinctive features of the FYDP can be more clearly illustrated when contrasting its implementation in both the United States and Canada, the two of which share an interlinked defense industrial sector, a comparable budget window, and a commitment to transparency. Although both the U.S. and Canada are considered federalist countries, there are fundamental differences between their two systems, including their form of government: the U.S. employs a presidential system and Canada embraces a parliamentary one.²⁸ Despite this difference, both the U.S. Congress and the Canadian parliament play major roles in their respective budgeting processes (McCormack, 2007, pp. 4).

Even so, the difference in government systems means that while the ruling coalition may change, Canada does not experience divided government as the Prime Minister will be chosen from the ruling party or coalition of parties. This increases the centrality of the executive branch, as the OECD notes "[i]n general, new budgeting procedures are introduced by Cabinet decision, rather than by new statutes or major modifications to existing statutes."²⁹ In theory, this closer alignment between branches should simplify the projection process by narrowing the gap between executive branch requests and legislative branch supports. That said, multi-year planning in a parliamentary system can be constrained by the potential unpredictability of its executive branch. Because parliament can call for new elections at any given time (for instance, during a motion of no confidence), Prime Ministers and their cabinets may not stay in office over fixed multi-year periods.³⁰

Like the U.S., Canada is one of the OECD countries that practices multi-year budgeting. The fiscal year in Canada starts on April 1 and ends on March 31. The Department of National Defence (DnD) largest accounts are Operating Costs and Capital, \$CN 15.8 billion and \$CN 3.8 billion respectively in 2019-2020 and the DnD itself is the largest federal government department. The budget process uses different terminology and concepts than the U.S. process, though the need to both plan in advance and adjust to changing circumstances is reflected is common to both systems:

"DND funding is comprised of multi-year reference levels approved by Treasury Board each year and accessed through the Parliamentary Estimates process... The Department receives an appropriation from Parliament on an annual basis. The cash budget is approved initially through the Main Estimates and can be revised up to three times per year through the Supplementary Estimates."³¹

Focusing on the capital budget, the Canadian procurement process centers around a network of several key ministries, all of which have separate policy objectives. The Department of National Defense (DND) may set its own goals for specific programs, but it cannot formally begin any processes without coming to an agreement with both Canada's central procurement agent (Public Services and procurement Canada; previously the Public Works and

https://doi.org/10.1787/9789264307957-en

²⁷ Thomas-Duerrel Young, "Questioning the "Sanctity" of long-term defense planning as practiced in Central and Eastern Europe," 364.

²⁸ Both governments have bicameral legislatures; in Canada, however, members of the Senate are not elected officials, but are instead appointed by the Prime Minister.

²⁹ OECD (2019), Budgeting and Public Expenditures in OECD Countries 2019, OECD Publishing, Paris, 162.

³⁰ For a discussion of "a procurement process is vulnerable to uncertainties due to electoral cycles" both due to changes in defence minister and strategic priorities, see Berkok, U. (2009) Canadian defence procurement. defence procurement and industry policy: a small country perspective, 225.

³¹ (2020 April, 3) "Defence Budget" Defence 101. Government of Canada. https://www.canada.ca/en/department-nationaldefence/corporate/reports-publications/transition-materials/defence-101/2020/03/defence-101/defence-budget.html

Government Services Canada (PWGSC) and industrial department (Innovation, Science, and Economic Development).³²

The Canadian government's 2017 white paper summarizes strengths and weaknesses of this procurement system:

"While 90 percent of projects are delivered within their planned scope and budget, defence procurement has undoubtedly faced challenges. In particular, a small segment of complex, high-value equipment projects have faced significant challenges. Cumbersome decisionmaking and approval processes have introduced undue delays. Accountability among departments has been diffuse and at times unclear. . . Early cost estimates have sometimes proven problematic, creating financial pressures and compromises to the final capability delivered. And perhaps most challenging, 70 percent of all projects have not been delivered on time."

The 2018-2019 DND self-reports 100% of projects staying within scope and budget for the past three years.³³ Under an internal audit using another measure for cost growth, for projects that began between January 2005 and a May 2013 audit report the cost of those that began in the definition phase increased by 3.6 percent and for those in the implementation phase increased by 1.6 percent, although "some projects were de-scoped to remain within budget."³⁴ The concerns about project deliveries are a longstanding phenomenon, with critics such as Dave Perry arguing that whatever advantages it brings comes at the expense of the DND's ability to effectively meet deadlines or maintain developmental flexibility over longer periods of time. As evidence, Perry notes that for the defence capital program, from 2007/2008 to 2013/2014, "an average of 23 percent of the available Vote 5 [Capital] money supplied by Parliament, (a combined \$7.2 billion) was not spent as intended."³⁵ The internal audit similarly "identified the following root causes for cost growth: project delay, scope increase, inflation, and delays in engagement with industry."³⁶

Although arguments for a singular defense procurement organization have been occasionally floated since the end of the second World War, the post-2001 era has seen a significant Canadian commitment to support the war in Afghanistan and renewed debates as to the efficacy of the current setup and whether the prevailing decentralized model should be reevaluated. In 2008, the Harper government began the implementation of a new set of policies intended to modernize the military known as the Canada First Defense Strategy (CFDS). Five years later, a 2013 report headed by the PWSGC's Tom Jenkins recommended fundamentally reforming the connection between the procurement process and the country's industrial sectors in light of the obstacles being encountered when trying to achieve some of the CFDS's main objectives.³⁷ Heavily influenced by this report, the 2014 Defense Procurement Strategy (DPS) was announced the subsequent year, with its primary goals functionally pushing for centralization

³² Reflecting Canada's confederal structure, Ugurhan Berkok notes the range of actors with influence on procurement matters can extend to regional economic development. Both national and regional actors will often push for non-defense economic development objectives, potentially raising the costs of acquired systems. Berkok, U. (2009). Canadian defence procurement. *defence procurement and industry policy: a small country perspective*,211-213, 220.

³³ Sajjan Harjit (2019). "Department of Defence and the Canadian Armed Forces's Departmental Results Report," DND, p.47. Note that the choice to suspend acquisition of the F-35 may remove one program previously subject to cost estimate disputes from this sample.

³⁴ Chief Review Services (2013) Internal Audit of Capital Project Cost Estimation, Department of National Defence, 13.. https://www.canada.ca/en/department-national-defence/corporate/reports-publications/audit-evaluation/internal-auditcapital-project-cost-estimation.html#find-1

 ³⁵ Perry, D. (2015). Putting the 'armed' Back Into the Canadian Armed Forces: Improving the Defence Procurement in Canada.
 Conference of Defence Associations Institute, 5-6. https://www.macdonaldlaurier.ca/files/pdf/MLIdefenceprocurement.pdf
 ³⁶ Chief Review Services (2013), 13.

³⁷ Jenkins, T. (2013). *Canada First-Leveraging Defence Procurement Through Key Industrial Capabilities: Report of the Special Adviser to the Minister of Public Works and Government Services*. Public Works and Government Services Canada. https://www.tpsgc-pwgsc.gc.ca/app-acq/documents/eam-Imp-eng.pdf

of the procurement process around the DND itself.³⁸ This newly proposed framework has been continually iterated upon by subsequent governments in the years since, with the CFDS being wholly replaced by the similarly reformoriented Strong, Secure, Engaged (SSE) program of 2017. In 2019 the ruling Liberal party proposed to consolidate defense procurement and the longstanding interest in procurement reform over the course of different programs and governments seems to indicate that reducing project delays is still being seen as a keenly important priority.³⁹

Both the challenges and successes of the Canadian system are multicasual, but for the purposes of this study, the reliability of their estimates is noteworthy. The DnD is responsible for initial cost estimates, though other parts of government, such as the Office of the Parliamentary Budget Officer, do at times produce independent estimates using DnD provided data.⁴⁰ The aforementioned comparative accuracy of Canadian estimates can be explained in part by Canada having an easier challenge. The Canadian system is primarily oriented towards procurement, not development, including procurement of U.S. systems. There is less uncertainty in program costs once they reach that phase and the U.S. Foreign Military Sales system is conservative in its estimation to ensure that purchasers have resources on hand to pay for the entire cost. Nonetheless, the Chief Review Service (2013) compared Canadian projects to similar counterparts in the United Kingdom and the United States and even after accounting for fewer deliverables, found that "the median unit cost increase of deliverables of the Canadian projects from Project Approval (Definition) was 12.4 percent, the cost growth was higher for both the U.S. projects at 13.5 percent."⁴¹

Thus Canadian estimates are a bit more accurate than those of the United State, but Canada is notably more capable of keeping projects within initial budget. One key contributor to the ability to stay within budget ceilings is that, as reported by the Chief Review Service (2013), estimates at the Project Approval (Definition) and the Project Approval (Implement Phases) have confidence intervals of +/- 25 percent and +/- 15 percent respectively.⁴² When DND is less confident in an estimate, "contingency funding" can be included in the project budget and "normally range from 10 to 15 percent for Major Crown Projects and may be as high as 20 percent for complex projects."⁴³ While these contingency funds often go unspent, the Canadian example indicates an approach of conservative estimation that can keep projects within budget ceilings even when point estimates prove incorrect. However, the successes of the system must be contrasted with the systems difficulty in project timeliness and tendency towards protected negotiations.

The Strategic and Communication Role of the FYDP

Canada's comparative success at budget estimation combined with its difficulty in starting new projects is a reflection of the described by Enthoven and Smith outlined in the quote at the start of the background section. The FYDP helps the executive branch make strategic choices and communicates them to internal and external audiences. Critically the FYDP is released as part of the President's Budget and thus linked to the negotiation between branches of the U.S. government. The numbers are provided to justify the funding requests the executive

³⁸ Auger, M. (2016). The Evolution of Defence Procurement in Canada. Library of Parliament= Bibliothèque du Parlement. https://lop.parl.ca/sites/PublicWebsite/default/en_CA/ResearchPublications/201609E#a28

³⁹ Jeffrey Collins (2019) "Defence Procurement Canada: Opportunities and Constraints" Canadian Global Affairs Institute. <u>https://www.cgai.ca/defence_procurement_canada_opportunities_and_constraints</u>

⁴⁰ Yalkin, Tolga R. and Weltman, Peter (2011, March 10) "An estimate of the fiscal impact of Canada's proposed acquisition of the F-35 Lighting II Joint Strike Fighter" Office of the Parliamentary Budget Officer. http://www.pbo-

dpb.gc.ca/files/Fublications/F-35_Cost_Estimate_EN.pdf In this case the Auditor General seconded a need for a new estimate. (2012). "Chapter 2—

Replacing Canada's Fighter Jets." 2012 Spring Report of the Auditor General of Canada, Auditor General of Canada, Rcommendation 2.77. https://www.oag-

bvg.gc.ca/internet/English/parl_oag_201204_02_e_36466.html#hd3d

⁴¹ Chief of Review Service (2013). p. 19.

⁴² Chief of Review Service (2013). p. 5.

⁴³ Chief of Review Service (2013). p. 5.

branch is making to Congress. As Todd Harrison and Seamus Daniels put it "The FYDP is therefore best understood as a statement of policy rather than a prediction of where the budget is headed. It is an indication, with considerable detail, of the Defense Department's priorities and trade-offs among modernization, force structure, and readiness."⁴⁴

This emphasis on policy and strategy provides another set of criteria that can be used to evaluate the effectiveness of the FYDP. Young in particular is critical of long term, as practiced in the United States and in Eastern European countries following the U.S. example. While other authors have emphasized the benefits of budgetary stability making efficient choices, Young believes this fails to acknowledge the way adversaries can unpredictably shape choices.⁴⁵ Instead Young highlights two tasks that defense planners can achieve: "to produce costed priorities" and "creating an understanding of future financial projection of current obligations." ⁴⁶ For Young, the quality of cost estimates matters for costing priorities but the value of the FYDP is not to provide reliable predictions. Instead "the utility of these financial projections should be judged by how much flexibility they can provide ministers and senior defense officials to change the way money is being spent to produce relevant defense outcomes." ⁴⁷

While strategic flexibility is inherently appealing, it does still face challenges noted by Kevin Lewis in his argument for defense planning humility. Lewis observes a range of cyclical factors in the defense budget and a shrinking portion of the budget going to combatant forces. He cautions "we should expect inefficiencies, and substantially more negative effects on capability than might be expected from these inefficiencies, because of the increasing role of "fixed-cost" overhead items."⁴⁸ That said, the larger emphasis on the FYDP's role in achieving flexibility harks back to Enthoven and Smith who reject the idea that long-range plans limit the President's ability to implement a new strategy and argue "an organization's flexibility to move in a new direction is greatly reduced if it lacks a clear picture of the direction in which it has been heading."⁴⁹

Young's emphasis on the strategic flexibility provided by the FYDP presents a challenge, while there has been extensive research comparing projections to actual costs, measuring the agility of the acquisition enterprise is a less intuitive problem. Will Domke in 1984 provided one possible answer by analyzing how the Defense Budget responded to presidential priorities going back to the Eisenhower by analyzing the winners and losers among DoD funding accounts. He found that the balance between Military Departments change most at the start of an administration and more popular President's have greater influence.⁵⁰

Picking up on Domke's approach, Travis Sharp argues that the current debate on gaps between strategy and resources "generally [does] not provide, however, is any objective sense of whether DoD is doing relatively better or worse aligning resources with strategy."⁵¹ Sharp considers three diagnostic tests including a Winners test that

⁴⁴ Harrison, Todd and Seamus Daniels (2020). *Analysis of the FY 2020 Defense Budget and Its Implications for FY 2021 and Beyond*, 11

⁴⁵ Young specifically queries "if the U.S. FYDP constitutes a best practice, then one has the challenge of explaining why, for instance, it had not predicted the need for up-armoured vehicles (i.e. MRAPs) in Afghanistan and Iraq and that they had to be procured out-of-cycle" ("Questioning the "Sanctity" of long-term defense planning as practiced in Central and Eastern Europe," 366). An apologia for the U.S. system may note that OCO spending does not inherently represent a failure of the FYDP process, but that the OCO budget is instead a tool well suited to respond to unexpectedly effective or resilient enemy approaches.
⁴⁶ Thomas-Duerrel Young, "Questioning the "Sanctity" of long-term defense planning as practiced in Central and Eastern Europe," 366).

⁴⁷ Thomas-Duerrel Young, "Questioning the "Sanctity" of long-term defense planning as practiced in Central and Eastern Europe," 370.

⁴⁸ Lewis, "The Discipline Gap and Other Reasons for Humility in Defense Planning", 132.

⁴⁹ Enthoven and Smith "How much is enough," 50.

⁵⁰ Domke, William K. "Waste, Weapons, and Resolve: Defense Posture and Politics in the Defense Budget." *Policy Sciences* 16, no. 4 (1984): 389. Accessed September 29, 2020. <u>http://www.jstor.org/stable/4531999</u>. Domke does not include FYDP data in his analysis.

⁵¹ Sharp, Travis "Did Dollars Follow Strategy: Analysis of the 2020 Defense Budget Request" CSBA (2019), 9. https://csbaonline.org/research/publications/did-dollars-follow-strategy-a-review-of-the-fy-2020-defense-budget

looks at whether spending has shifted into areas identified by the National Defense Strategy as priorities.⁵² He looks at the FYDP's Major Program Categories and investment areas, with the latter only available through the President's Budget. In both cases Sharp finds that the 2020 President's Budget failed this test.⁵³

The ability to conduct these tests is a reminder of the final important virtue of the FYDP is that of transparency. Thomas Mahnken, president of the Center for Strategic and Budgetary Assessment and former deputy assistant secretary for policy planning noted that the Pentagon has long protected certain aspects of the budget from scrutiny but made the case that this has been successfully balanced with disclosure and that "there is also a compelling case for disclosing how the Defense Department plans to spend its resources and whether its budget is aligned with its strategy. . . Transparency ultimately helps the Defense Department make the case for the resources it needs in Congress as well as the public at large."⁵⁴

This unclassified use of the FYDP is a further expression of one of the ideas that underlay its creation, as detailed by Enthoven and Smith:

"A final basic idea underlying PPBS was that of open and explicit analysis; that is, each analysis should be made available to all interested parties, so that they can examine the calculations, data, and assumptions and retrace the steps leading to the conclusions. Indeed, all calculations, data, and assumptions should be described in an analysis in such a way that they can be checked, tested, criticized, debated, discussed, and possibly refuted by interested parties." ⁵⁵

In the earlier editions of their book, the FYDP was entirely classified, but the use today of the unclassified portions of the FYDP lives up to this ideal, albeit with a broader definition of interested parties.

4 Scope

The objective of this research is to assess the reliability of the FYDP as a planning tool and to develop publicly accessible analytical tools for defense spending that take the underlying reliability into account. This paper seeks to answer the following questions:

- How reliable are projections within the FYDP as an indicator for actual spending?
- Which Military Departments have the most and least reliable projections?

The hypotheses below are intended to add rigor and structure to the exploration of FYDP reliability and to address the first and second research questions, respectively.

How reliable are projections within the FYDP as an indicator for actual spending?

Creating and implementing a strategy is a process of matching means to ends. The planning stage of the PPBE incorporates the opinions of combatant commanders and service leadership outlining what resources they will need to meet the strategy laid out by executive branch civilian leadership. This planning is shaped by intentions but requires the setting of priorities and choosing what balance of investments will be best able to meet the strategy's future needs. The programming process, which produces the FYDP, seeks to further tether these priorities to real world constraints and to identify where further trade-offs are necessary to reconcile plans across the department. This priority setting exercise means that even when systemic biases are present, the process of

⁵² "Instead, a winner is best defined as one that receives the largest increase in its proportional share of DoD's total spending compared to what DoD's prior outyear plans forecasted"

⁵³ Sharp, Travis, 24.

⁵⁴ Mehta, Aaron (2020, April 4) *Pentagon denies it seeks to hide future budget information*. DefenseNews <u>https://www.defensenews.com/pentagon/2020/04/03/pentagon-denies-it-seeks-to-hide-future-budget-information/</u> 55 p. uk. 2010 July 2010 Jul

⁵⁵ Enthoven and Smith "How much is enough," 45.

FYDP creation should identify which procurement line items and RDT&E programs are more or less important to the DoD. This means that the bottom-up detailed information in the FYDP can be valuable even if the topline spending estimate is off the mark. Moreover, the longer cycle times of investment spending means that even a critical new technology will be ramping up spending over multiple years or a somewhat outmoded platform may be allowed time to complete its current block of production. Thus, at the margins, each year of the FYDP should reveal information about prioritization above and beyond that in the President's Budget and therefore add value at the margins to forecasting efforts.

 $H_{1A:}$ A future years defense projection of an increase (decrease) for a procurement line item directly predicts an increase (decrease) in actual spending for that line item.

H_{1B:} A future years defense projection of an increase (decrease) for an RDT&E program element directly predicts an increase (decrease) in actual spending for that program element.

Which Military Dpartments have the most and least reliable projections?

The planning process that feeds into the FYDP is filtered through the separate processes of the military departments which have their own distinct leadership, history, bureaucratic incentives, and approaches. These differences were further shaped by the different time horizons of the platforms they focus on. In particular, the thirty year shipbuilding plans reflect the Navy's orientation emphasis on a smaller number of complex and high value platforms where new technology may be more readily inserted from ship to ship and limited shipyard space strictly contains the number of units that can be under production simultaneously.

Likewise, the military departments often put different weights on the importance of readiness, force structure, and investments and thus may prioritize other funding accounts over investment spending as well as the associated analytical processes and personnel used for projection. As a result, the reliability of projects should vary between Military Departments. For this hypothesis, the categorical variable of "Military Department" acts as a mediating variable for the reliability of FYDP projections.

5 Data and Methods

Data Sources and Structure

DoD investment budget data is reported in multiple forms and through multiple sources. At the aggregate level, the OMB Public Budget Database covers budget authority for the president's budget and the next four years across the federal government. The DoD Comptroller site also provides the DoD Greenbook, which provides decades of history on a variety of different measures of spending, again aggregated at a high level. The DoD Greenbook provides topline Procurement and RDT&E for the overall DoD as well as the Military Departments. The OMB data offers greater granularity, particularly for Procurement, including main account codes such as "Procurement of Weapons and Tracked Combat Vehicles, Army" or "NATO cooperative defense programs".⁵⁶

The most pertinent limitation of OMB and DoD Greenbook is that they do not provide detail on priorities within those broad accounts. These sources can inform a truck manufacturer of the Army Procurement topline, but not the anticipated need for various categories of cargo transports. Likewise, an analyst that finds certain areas of technology to be promising or overhyped cannot easily observe changes in patterns of investment. For major weapon systems, there are multiple government sources that overcome this limitation: DoD provides a document on Program Acquisition Costs by Weapon System, and Selected Acquisition Reports.⁵⁷ GAO produces its Weapon

⁵⁶ The study team recommends Todd Harrison and Seamus Daniels (2020) as an accessible walkthrough of the data in these and other topline reports.

⁵⁷ Department of Defense. (2018). *Department of Defense Selected Acquisition Reports (SARs) – December 2018*. <u>https://www.defense.gov/Newsroom/Releases/Release/Article/1923492/department-of-defense-selected-acquisition-reports-</u>

System Annual Assessment that casts an often skeptical eye on DoD cost and schedule estimates.⁵⁸ However, while this reporting on Major Defense Acquisition Platforms does cover a significant portion of DoD investment spending, it does not include many smaller items or early-stage research, which is not associated with specific programs.

This project focuses on budget lines rather than major major programs for both financial and policy reasons. Todd Harrison reports that as of the FY 2016 President's Budget, "these smaller programs account for an average of 57 percent of the total acquisition budget over the FYDP" (2016, pp. 24). Moreover, during the study period DoD is increasingly experimenting with alternate channels, including mid-tier acquisition and other transaction authority, responding to pressure from the Executive Branch and Congress to pursue speed and innovation.

To better understand this era of reduced reliance on the major weapons system pipeline, this report chooses to focus its attention on procurement line items and RDT&E program elements. This is not the most detailed level of analysis available, however it has the advantage of being a unit of analysis of interest to a broad community of practitioners and observers while also being available from multiple sources. The first pair of these sources are the P-1s for Procurement and R-1s for RDT&E. These budget documents are provided as spreadsheets that cover the entire DoD enterprise, going back to the 1998 President's Budget.⁵⁹ However, these detailed and convenient documents do not include projections past the President's Budget, and are missing the bulk of the FYDP.

Instead, for detailed future year projections at the procurement line item (P-40) and RDT&E program element (R-2) level, its necessary to turn to the Justification Books. In these documents, the military departments and agencies lay out their spending request and describe what is being bought as well as providing program management details. A major challenge for open source researchers is that these documents are provided as large PDFs, not in the machine-readable format of the sources discussed earlier. Moreover, for investment spending alone, each President's Budget is accompanied by dozens of these files, splitting the information based on organization and funding account.

Creation of the CSIS Investment Budget Line Dataset

This project overcomes the limitations of the justification books and does so with the benefit of two external sources. First, a predecessor FYDP analysis led by Gabriel Coll bulk laid the foundation for this project by downloading many of the justification books, converting, exporting them to rough spreadsheets, and then cleaning them up and combining that spreadsheet information.⁶⁰ A parallel effort by CSIS's Defense Budget Analysis (DBA) group greatly accelerated this effort with the discovery the justification books, starting with the FY 2013 FYDP, have included XML encoded spreadsheets that do not require the intensive data cleaning effort necessary when scraping PDFs. This past and parallel work assisted in the creation of the dataset, and, in the DBA case, provided a valuable source for cross checking totals and budget line classifications.

To allow for cross-comparisons, the team has imported R-1s and P-1s files from the comptroller website for the budget requests from FY 2011 to FY 2021, the study team often had multiple choices as to which file, and in some cases, which columns to include. As a rule, if appended reports were available, the most recent one was preferred for this report. Similarly, the reporting of enacted budgets in multiple years, typically those with continuing resolutions or multiple supplementals, include variant columns. In those cases, whether the columns covered base,

<u>sars-december-2018/</u>. Thanks to Freedom of Information Act requests, many of the underlying documents, rich with qualitative and quantitative data, are available through the DoD FOIA reading room: Executive Services Directorate. (n.d.). *Selected Acquisition Reports (SARs)*. Department of Defense. <u>https://www.esd.whs.mil/FOIA/Reading-Room/Reading-Room-List 2/Selected Acquisition Reports/</u>.

⁵⁸ See Oakley, Shelby S. (2019). Weapon System Annual Assessments: Limited Use of Knowledge-Based Practices Continues to Undercut DoD's Investments (GAO-19-336SP). Government Accountability Office.

⁵⁹ The budget materials may be found at the DoD Comptroller's website: DoD Comptroller. (2020). *DoD Budget Request*. <u>https://comptroller.defense.gov/Budget-Materials/</u>.

⁶⁰ The acknowledgments at the start of this paper recognize the team behind this predecessor effort.

supplemental, or total spending, the column that included continuing resolutions or more recent enactments was chosen.

Where possible, greater granularity was also preserved. For all of the years in this report, there was OCO spending and for enacted and president's budget years, the dataset was constructed to allow differentiation between base and OCO spending. The level of granularity available varies based on the source. The reported actual values in both sources are always aggregated, and the FYDP reporting also aggregates the enacted amounts. The enacted amounts for the R-1s and P-1s show the greatest variation between years, reflecting the range of divergences from usual order as budgets were enacted late, requiring continuing resolutions, and sometimes split spending across multiple supplemental bills.

Much of the effort of the dataset focused on the creation of unique identifiers, called CSIS budget line keys, that ease the process of tracking a budget line across different sources and over time. While many of the same columns are available in the R-1s and R-2s as well as in the P-1 and P-40s, there are discrepancies in the labels used, which are more challenging in early years and with procurement data in particular. For example, in some years and for some agencies, the line number, that is to say the order in which it is presented in that PB, is the same as a line item. This causes problems P-1s and P-40s include slightly different budget lines, leading to misalignment, and because order regularly changes from year-to-year. Moreover line item standards change over time, with many budget lines change from having a six-character line number to a ten character one while still having the same broad topical focus. For procurement, for the analysis in this paper, all cost types pertaining to a single program are combined.⁶¹

A related challenge is that in a typical year, a few score PEs and LIs will cease to receive funding or be tracked for the first time. This is a natural outflow of changes in strategy, priorities, and technology. However, some of the time, a new PE or LI does mean a genuinely new project but may instead reflect a change in identifiers for an existing budget line of greater magnitude than just an increase in the length of the identifier used. This may be a matter of a change of agency, such as the move from the Defense Health Agency to the Defense Health Program or a reclassification such as a Major Force Program shift in the lead up to the creation of Space Force.

The ties between the original sources and their unique identifiers are recorded within the dataset's repository for transparency and reproducibility reasons. The study team has taken the additional step of classifying budget lines based on our confidence that potential confounding factors have been removed.

Figure 3 shows a summary of the R-1 and R-2 portions of the dataset by confidence level in the left and right columns respectively. The top row shows the total actual spending associated with the projections of each PB. For the R-1, this includes only the actual spending for the budget year in question. The R-2s show substantially more spending because their projection window goes out five years, the President's Budget and Four out years. For those cases where projections go beyond PB 2019, the latest year where actual spend data was available, the spending amount projected is shown in grey.

Focusing on those projections for which we know the actual spending, there are five confidence levels of interest:

- No FYDP Expected: This covers budget lines, such as classified spending, that are typically not included in the unclassified FYDP as reported by the R-2s or P-40s. They account for tens of billions of dollars of P-1 and R-1 spending annually, but are beyond the scope of this research efffort.
- Unanticipated Budget Line: This covers budget lines that did not yet exist when the PB was published. They may be a genuinely new budget line, or perhaps a transfer whose predecessor was not identified by

⁶¹ In the early years of the dataset, Advanced Procurement LIs sometimes had a different line item than the rest of a program. These divergent advanced procurement budget lines have been combined by the study team with the rest of their program under the same CSIS budget line key.

the study team. Unanticipated Budget Line shrinks in absolute terms because it is a lagging indicator and only available once actual spend figures are known.

- Not Confident: This category covers budget lines that are present in one source but not in the other. For procurement, this also covers cases in which there is a cost type under the budget line, for example, advanced procurement, that is not present in the other source and there is a net difference in projection of more than \$2,000 between the two sources.
- Semi-confident: This category covers budget lines that are present in both sources but that have one of a range of known discrepancies. This includes having a difference projection between the two sources of more than \$2,000. In the case where there is a missing cost type, budget lines are reported as semi-confident if there is a net difference between the two projections of \$2,000 or less but the spending occurs in different years depending on the source. Finally, if a budget line projected spending in a future year, but the budget line was not reported in the year in question, it is marked as semi-confident. Some of these semi-confident rows may be transfers to Unanticipated Budget Lines, but the study team has not yet identified the connection.
- Confident: This is the category that is the basis of modeling in this paper. To be classified as confident, the total projections have to be within \$2,000 of one another and, if there is a missing cost type, they must also be within \$2,000 of one another in every year of the projection. Budget Line may end during the projection period and still be classified as confident, but only if the ending was anticipated by the President's budget. These strict criteria are intended to limit the sample to those budget lines that are genuinely starting and ending rather than having overlooked connections.

Turning again to Figure 3, the larger light blue blocks present in the R-2 column indicate that there is a substantial amount spent by PEs that the FYDP does not see coming years in advance. The lower row of Figure 3 shows the same budget lines but using the metric of the count of lines rather than the amount of actual spending. Both Unanticipated Budget Lines and Semi-Confident PEs have a larger share of the count of budget lines than of the funds. This suggests that unanticipated line items often start small.

and sources.



Figure 3: RDT&E Actual Spending and Budget Line Count by Level of Confidence)

The Procurement dataset is summarized in Figure 4 and, by the prevalence of red and yellow, shows that there are more disagreements between the P-1s and P-40s than the R-1s and R-2s. There were multiple contributors to these problems. First, there were gaps in reporting on the P-40 side some of which the study team has already overcome through manual transcrbing of PDF files that were missing XML files.⁶², Second, there was a technical challenging importing the PB2016 P-1s as the latest data was split across two updates files. Third, reporting of cost types, particularly reductions to adjust for prior year past procurement, were sometimes missing from P-40 reporting and merit closer examination. Finally, while for RDT&E, the unanticipated budget lines clearly outnumber and out spend the semi-confident lines, that is not true for procurement. This suggests that a fair portion of these unanticipated budget lines may in fact be more cross-year transfers within procurement with corresponding line items in the semi-confident category.

Source: R-1s, R-2s, CSIS Analysis

⁶² The most problematic absence in dollar terms had been the Navy Shipbuilding and Construction account.



Figure 4: Procurement Actual Spending and Budget Line Count by Level of Confidence)

Measuring Dependent and Independent Variables

Table 1 illustrates all FYDP reporting across nine years 2013 - 2021. The light blue arrows show the connection between the first two years of the FYDP and the subsequent reporting year for the corresponding actual data. The reported actual spending always lags behind the FYDP by two years, and so it is necessary to wait three years to find out if the second year of the FYDP was accurate and a full six years to find out if the fifth year of the FYDP projections were on target.

| | Reporting P-40s, R | | s, and R-2s) | | R-2s) | | | |
|-----|--------------------|-----------|--------------|-------------|----------------------|----------------------|----------------------|----------------------|
| | Year | Actual | Enacted | President's | 2 nd Year | 3 rd Year | 4 th Year | 5 th Year |
| | | Spending | Budget | Budget | of FYDP | of FYDP | of FYDP | of FYDP |
| t | PB13 | FY 2011 | FY 2012 | FY 2013 | FY 2014 | FY 2015 | FY 2016 | FY 2017 |
| t+1 | PB14 | FY 2012 | FY 2013 | FY 2014 | FY 2015 | FY 2016 | FY 2017 | FY 2018 |
| t+2 | PB15 | FY 2013 ← | FY 2014 | FY 2015 | FY 2016 | FY 2017 | FY 2018 | FY 2019 |
| t+3 | PB16 | FY 2014 🧲 | FY 2015 | FY 2016 | FY 2017 | FY 2018 | FY 2019 | FY 2020 |
| t+4 | PB17 | FY 2015 | FY 2016 | FY 2017 | FY 2018 | FY 2019 | FY 2020 | FY 2021 |
| t+5 | PB18 | FY 2016 | FY 2017 | FY 2018 | FY 2019 | FY 2020 | FY 2021 | FY 2022 |
| t+6 | PB19 | FY 2017 | FY 2018 | FY 2019 | FY 2020 | FY 2021 | FY 2022 | FY 2023 |
| t+7 | PB20 | FY 2018 | FY 2019 | FY 2020 | FY 2021 | FY 2022 | FY 2023 | FY 2024 |
| t+8 | PB21 | FY 2019 | FY 2020 | FY 2021 | FY 2022 | FY 2023 | FY 2024 | FY 2025 |

Table 1 Guide to FY Reporting available from FY 2015 FYDP to the FY 2021 FYDP

All dollar in this report amounts are current dollars. Although changes in spending timing (delays, shifts, etc.) occur throughout the time period, much of the FYDP estimations themselves have inflation concerns built in to their reasoning.

For statistical purposes, values have been aggregated into a cumulative expenditure; this means that both the Actual Spend and Project Budget variables consolidate not just all of their individual fiscal years in question, but also the actual and projected budgets for FYDP out years as well. This amalgamates multiple spending years into single variables rather than strictly individual ones.

There are several tradeoffs to this methodology. The analysis is exchanging year-specific sensitivity for a more robust measurement better representative of the discrepancies being assessed over the time periods in question. This sacrifices the ability to directly compare the predictive strength of different FYDP years, i.e. testing the extent to which the second and third out years are more reliable than the fourth and fifth out years. Likewise, this approach makes it harder to account for any single year having an abnormal occurrence, such as sequestration. In return, the variables being compared serve as a much more accurate representation of what is being estimated in total. From an industrial standpoint, this serves as a more natural portrayal of how spending is being looked at overall.⁶³

In order to properly capture these changes, we analyzed both PB13 and PB18. PB18 is analyzed using multivariable modeling and specifically looks at only the 2nd year of the FYDP projections from PB18 to the actual spending reported in FY2019.

Dependent Variable

Total Spent [Log (FYDP2_ActCml+1) i, t+3] is the actual spending amount reported three years after PB in question. So an estimate relying on numbers reported in the PB18 will look to the actual FY2019 spend (reported in the PB21). The variable is measured in current thousands of dollars. It is incremented (histograms on left) by one and then logged (histograms on right).



⁶³ The study team also tested the possibility of including separate study variables for each year of the FYDP, but that approach often ran afoul of tests for multicollinearity.

Procurement



Study Variables

Projected Budget [Log (ProjectCml+1)_{i,t}] is the cumulative FYDP projection for a given program element or line number from the PB in question, inclusive of any OCO spending. For PB18 this means the total president's budget for FY2018 plus the second out year projection for FY2019. The variable is measured in current thousands of dollars. It is incremented by one and then logged.

Projected Base [Log (BaseCml+1)_{i,t}] is the cumulative FYDP projection for a given program element or line number from the PB in question, for PB18 this means the president's base budget for FY2018 plus the second out year projection for FY2019. The variable is measured in current thousands of dollars. It is incremented by one and then logged.

Military Department [MilDep], DoD Component [DoDComp], and Defense Organization

[DefenseOrganization] are categorical variables with values denoting the service in charge of the program element or line item at the time of the FYDP estimate. This is reported by a column in the P-40s, but in the R-2s it can be derived from the title of the justification book in question (as they are divided into organizational groupings).

- Military Department [MilDep] uses four values: Navy (baseline), Air Force, Army, and Other DoD. The Navy was chosen as the baseline service based on the large number of entries compared to all other categories.
- Defense Organization [DefenseOrganization] uses twenty-two values, all of which directly correspond to the organizations coded within DoD budget documentation (AF, ARMY, CPBD, DARPA, DCAA, etc.). Of these three categorical variables, this had the largest frequency range the Navy, Air Force, and Army all had over 400 budget lines, whereas many DoD organizations were comparatively smaller, with ten organizations having less than 5 total budget lines each.

Account [AccountDSI] refers to the top-level code listed for an entire appropriation account. This number is a combination of the Main Account Number and the Defense Service Identifier (DSI), which is generally represented as a single letter (A for Army, N for Navy, etc.)

Budget Activity [BudgetActivity] identifies the specific Budget Activity (BA) being referenced for any given program element or line item. A budget activity is a category within each appropriation and fund account that identifies the purposes, projects, or types of activities financed by the appropriation or fund.

Control Variables

Actual Spend [Log (FYPD2_BaseCml+1)] represents the actual spending for a given program element or line item from the PB in question. For instance, in PB18 this would be referring to FY16. This variable is intended to capture the level of inertia between actual spending in an account, regardless of projections.

IsFYDP2Zero indicates if a program element or line item from PB2018 receives zero funding in a later FYDP estimate. By itself, an estimate of zero spending should correlate with a lower level of spending. However, the frequency of high levels of spending for budget lines that had been zeroed out suggests that projections of no spending should be treated with greater skepticism than those for merely low levels of spending.

IsFYDP5Zero indicates the same as IsFYDP2Zero, but for PB2015.

Estimating Equation

The change in line item or program element spending is estimated using ordinary least squares regression.

Equation 1 RDT&E Program Element Estimating Equation

Estimated Actual_{*i*,t+3} =

Estimated Actual_{i,t+3} = $\alpha_{j[i]}^{BA}$ + β_1 Intercept + β_2 Log(FYDP2 + 1)_{i,t} + (β_3 Army_{i,t} + β_4 Air Force_{i,t} + β_5 Other DoD_{i,t}) + β_6 Log(Actual + 1)_{i,t} + β_7 Log(PB Base + 1)_{i,t} + β_8 Log(PB OCO + 1)_{i,t} + ϵ_i , for i = 1 to 877, t = FY2018,

$$a_i^{BA} \sim N(\mu_{\alpha}, \sigma_{\alpha}^2)$$
, for $j = 1$ to 7;

Equation 1 Procurement Line Item Estimating Equation

 $\begin{array}{l} \textit{Estimated Actual}_{i,t+3} = \alpha_{j[i]}^{Account} + \alpha_{k[i]}^{MDAP} + \beta_{1}\textit{Intercept} + \beta_{2}\textit{Log}(FYDP2 + 1)_{i,t} + \\ (\beta_{3}\textit{Army}_{i,t} + \beta_{4}\textit{Air Force}_{i,t} + \beta_{5}\textit{Other DoD}_{i,t}) + \beta_{6}\textit{Log}(\textit{Actual} + 1)_{i,t} + \\ \beta_{7}\textit{Log}(PB\textit{ Base} + 1)_{i,t} + \beta_{8}\textit{Log}(PB\textit{ OCO} + 1)_{i,t} + \epsilon_{i}, \textit{ for } i = 1 \textit{ to } 762, t = FY2018, \end{array}$

$$a_j^{Account} \sim N(\mu_{\alpha}, \sigma_{\alpha}^2), \text{ for } j = 1 \text{ to } 16;$$

 $a_k^{MDAP} \sim N(\mu_{\alpha}, \sigma_{\alpha}^2), \text{ for } k = 1 \text{ to } 62;$

6 Results

How reliable are FYDP projections?

A linear regression of the FY2019 Actual spend to its FYDP projections shows that there are statistically significant correlations across all periods, including PB2015 and PB2016. This is true when including all significant outliers. The coefficients are all positive, with the PB2019 thru FY2019 coefficient being the highest, indicating that the explanatory value of the projections are best one year out. That said, the lowest coefficient is for PB2017 thru FY2019, likely due to temporal or data quality factors specific to that year, with the PB2015 and PB2016 cumulative FYDPs each having more explanatory value. As shown in Table 2, these correlations are significant at the 0.001 level.

Table 2: Univariate Estimates of Cumulative Spending Through 2019 Spending for Budget Lines at All Confidence Levels

| | PB2019 thru FY2019 | PB2018 thru FY2019 | PB2017 thru FY2019 | PB2016 thru FY2019 | PB2015 thru FY2019 |
|------------------------|-----------------------|--------------------------|--------------------------|-----------------------|-----------------------|
| (Intercept) | 4.32*** | 8.03*** | 10.06*** | 8.21*** | 5.98*** |
| | (0.23) | (0.25) | (0.25) | (0.28) | (0.27) |
| log(ProjectionCml + 1) | 0.73*** | 0.52*** | 0.40*** | 0.45*** | 0.55*** |
| | (0.01) | (0.02) | (0.02) | (0.02) | (0.02) |
| AIC | 10404.52 | 12229.02 | 13278.71 | 14654.73 | 15658.89 |
| BIC | 10421.24 | 12245.88 | 13295.65 | 14671.79 | 15676.10 |
| Log Likelihood | -5199.26 | -6111.51 | -6636.36 | -7324.37 | -7826.45 |
| Deviance | 23682.78 | 48021.43 | 70456.37 | 105470.10 | 124710.87 |
| Num. obs. | 1949 | 2038 | 2088 | 2181 | 2290 |

[°] p < 0.001; ^{**}p < 0.01; ^{*}p < 0.05; [°]p < 0.1.

Figures 5 show cumulative FY2019 actual spending plotted against FYDP projections starting in 2015.⁶⁴ The Y-Axis is the Total Spend, the logged set of actual expenditures fully realized as outlays over given years. The X-Axis is the Projected Budget, the logged set of prior expenditure estimates for all budget lines expected to exist in those future actual spending years. The model confirms the hypothesis that there is a significant relationship between projected and actual spending, with this correlation lower for the out years.

Both RDT&E and Procurement showed similar results, though Procurement is proportionally more affected by OCO spending during PB years. Outliers lying on the axes are the result of projected items that ended up having no actual spend (x-axis) or items appearing with actual spend that projected no spending or were unanticipated (y-axis). When limiting the sample to those budget lines where the study team is more confident of the quality of the the data, as shown in the second row of Figure 5, the quality of the relationship strengthens. While the President's Budget still has the highest coefficient in Table 3, the out years, no longer as concerned with unanticipated budget lines or those that end earlier than expected, perform much better..

Table 3: Univariate Estimates of Cumulative Spending Through 2019 Spending for Confident Budget

Lines

| | PB2019 thru FY2019 | PB2018 thru FY2019 | PB2017 thru FY2019 | PB2016 thru FY2019 | PB2015 thru FY2019 |
|-------------|--------------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| (Intercept) | 3.48*** | 4.77*** | 5.30*** | 5.35*** | 4.43*** |

⁶⁴ Due to the extreme range in spending across different items, all figures will employ logarithmically transformed axes for statistical purposes. Logging both sides results in more normalized sets here, rather than the heavily skewed nominal sets.

| | (0.19) | (0.20) | (0.21) | (0.27) | (0.21) |
|------------------------|----------|----------|----------|----------|----------|
| log(ProjectionCml + 1) | 0.79*** | 0.74*** | 0.72*** | 0.72*** | 0.77*** |
| | (0.01) | (0.01) | (0.01) | (0.02) | (0.01) |
| AIC | 8316.99 | 8673.78 | 7942.08 | 4891.04 | 7430.24 |
| BIC | 8333.48 | 8690.21 | 7958.15 | 4905.42 | 7446.05 |
| Log Likelihood | -4155.50 | -4333.89 | -3968.04 | -2442.52 | -3712.12 |
| Deviance | 10623.97 | 13968.02 | 14599.51 | 12545.50 | 14834.82 |
| Num. obs. | 1802 | 1767 | 1565 | 891 | 1435 |

p < 0.001; **p < 0.01; *p < 0.05; p < 0.1.

Include outlier-excluded scatter plots here? The new regression has even stronger correlation, with Insert R2 and p values for different comparisons. Potentially include confidence intervals again.

Figure 5 Scatter Plot showing FYDP projections versus FY2019 Actual (Investment)



Figures 6 show histograms of the investment scatter plot budget lines segmented into three grouped time periods for visualization purposes. The X-Axis denotes the relative difference between projected expenditure on any single budget line and the actual spend on that same budget line in FY2019. Relative difference is so that the scale can

included cases where the projected or the actual spending figures are zero.⁶⁵ For those points in the center of each histogram, the difference between actual and projected values is small.

The Y-Axis denotes the nominal amount of budget lines with that level of difference. The chart is horizontally symmetrical around 0, with the left side of the chart indicating budget line items where the projected value was much higher than what actually ended up being spent and the right side of the chart indicating budget line items where the estimated value was much lower than what actually ended up being spent. The groupings all display normal distribution, with the same previously mentioned 0 spend outliers showing on the edges, the left edge representing cases where spending was projected but none took place and the right edge representing cases where zero dollars were projected, or the budget line was unanticipated by the PB, but spending took place nonetheless.

Blue dashed lines indicate the X-Axis median. As hypothesized, the closer the estimation years get to the actual years they're estimating for, the more accurate and center-oriented the estimations end up being. There is a distinct drop off in overall accuracy from the PB year estimation to the out years estimations, however, and this results in significantly wider histogram distributions.





⁶⁵ Relative difference, for the purposes of these charts, is calculated by dividing the numerator of (Cumulative Actual Spending - Cumulative Projected Spending) by the denominator of (Cumulative Actual Spending - Cumulative Projected Spending) / 2. Relative difference is not defined in those cases where both the actual spending and the projected spending values are zero.

Which services and budget categories have the most and least reliable projections?

Figure 7 shows the breakdown of projected vs. actual spending across Military Departments [MilDep]. Both the X and Y axes are identical to those in the previous scatterplot, figure 5. In general, this cross-section indicates that while projections across the board are in general accurate and all data continues to heavily cluster along the line of best fit. Smaller deviations can be seen from the Army and Air Force, both of whom have the most direct impact on what skewness that does exist. This is especially true for Procurement, where the inaccuracy of projections in FY2015 are especially notable.



Figure 7 Scatter Plot showing FYDP Projections versus FY2019 Actuals broken down across Military Departments for all confidence levels and for only confident budget lines

Figures 8 shows histograms similar to above, with the X-Axis denoting the difference between projected and actual spend, and the Y-Axis denoting the amount of budget lines with that difference. Again, blue dashed lines indicate the X-Axis medians. Edge outliers are especially noticeable here, with 2015 showing a high number of 0 spend outliers. Indeed, these histograms most directly show how how a larger number of unanticipated budget lines crop up as additional years of actual results become available. PB2019 presently only has 102 outliers, but with an extra two years of results that number increases to 396 for PB2017 and further rises to 483 for PB2015.



Figure 8 Histogram showing budget line distribution by difference between FY2019 actual and FYDP projection and broken down by Military Department for all confidence levels

Note: The relative difference is undefined when both cumulative actual and cumulative projected spending are zero.

Modeling FY 2018 PB Estimate of 2019 Actual Spending

The results of both the RDT&E and Procurement models are shown in Table 2 below. Because the R-2 model did not encounter multicollinearity challenges, it is able to include both the 2nd year out base prediction and the PB base prediction in the same model. That was not possible for the P-40 sample, and thus two models are presented, a control model in the middle that only includes information available in the PB and a study model on the right that includes the 2nd year out projections.

In the R-2 model in the second column the study variable, Δ FYDP2('19 vs. '18 PB Base), shows a direct relationship between the FYDP2 estimate and actual spending and is significant at the 0.1 percent level, affirming H_{1A}. All of the budget variables are denominated in thousands of dollars. The coefficient of 0.61 means that for a \$100k increase (decrease) in between the PB base and the FYDP2 projection, the actual amount spent is estimated to increase (decrease) by \$61k. The difference between the PB base and the actual spend is also significant at the 0.1 percent level and has a greater magnitude coefficient. A \$100k increase (decrease) in PB base request over the actual spend is estimated to increase (decrease) the actual spending by \$84k. While the coefficient for the PB OCO request is slightly larger than that for study variable, the standard error for PB OCO is larger still, suggesting considerable volatility.

| | R-2 Model | P-40 Model |
|--|---------------------|---------------------|
| (Intercept) | 3.59 ^{***} | 4.11 ^{***} |
| (| (0.28) | (0.37) |
| Study Variables - FYDP | (0:20) | (0.07) |
| log(FYDP2+1) | 0.23*** | 0.25*** |
| | (0.03) | (0.03) |
| Study Variables - Mil. Dept. (Baseline=Navy) | | |
| Army | 0.06 | 0.72* |
| | (0.15) | (0.30) |
| Air Force | 0.35* | 0.16 |
| | (0.15) | (0.30) |
| Other DoD | 0.41** | 0.85 |
| | (0.16) | (0.49) |
| Controls | | |
| log(Actual+1) | 0.07*** | 0.09*** |
| | (0.01) | (0.02) |
| log(PB Base+1) | 0.35*** | 0.27*** |
| | (0.04) | (0.04) |
| log(PB OCO+1) | 0.06 | 0.12*** |
| | (0.04) | (0.02) |
| AIC | 3321.37 | 3179.68 |
| BIC | 3369.13 | 3230.67 |
| Log Likelihood | -1650.68 | -1578.84 |
| Num. obs. | 877 | 762 |
| Num. groups: BudgetActivity | 7 | |
| Var: BudgetActivity (Intercept) | 0.06 | |
| Var: Residual | 2.44 | 3.50 |
| Num. groups: MDAP | | 62 |
| Num. groups: AccountDSI | | 16 |
| Var: MDAP (Intercept) | | 0.22 |
| Var: AccountDSI (Intercept) | | 0.13 |

Table 4 Estimating Actual 2019 Spending

****p < 0.001, **p < 0.01, *p < 0.05, p < 0.1.

The P-40 model was unable to include Δ FYDP2('19 vs. '18 PB Base) and Δ PB Base('18 vs. '16 Actual) in the same model due to multicollinearity issues. To test the hypothesis, the study team tested two different models, one with only the PB requests and one with Δ FYDP2('19 vs. '16 Actual) replacing both Δ FYDP2('19 vs. '18 PB Base) and Δ PB Base('18 vs. '16 Actual). Both the Akaike information criterion (AIC) and Bayesian information criteria (BIC) metrics prefer the model with the lower score, which in this case is favorable for the P-40 study model over the P-40 control model. In addition, the 0.85 coefficient for Δ FYDP2('19 vs. '16 Actual) is notably larger than for the control parameter, although both are significant at the 0.1 percent level. This support for a direct relationship between FYDP2 projections and actual spending affirms H_{1B}. In the study model, a \$100k increase in the difference between the FYDP2 estimate and the three years earlier actual spend is estimates \$85k more in spending. For this model, the PB OCO parameter is also significant with a \$100k PB OCO request in 2018 estimating \$80 more in actual spending in 2019.

7 Discussion and Conclusions

Both the RDT&E and procurement variants of hypothesis 1 were upheld as the FYDP2 projections for 2019 were able to significantly estimate actual spending, even after controlling for the base and OCO estimates in the FY 2018 PB. This positive result was achieved even though the acting DoD comptroller noted that the OMB reported FYDP topline was not created in a manner that inspired confidence: "what OMB has provided to date for the Defense Department is a flat top line beyond F.Y. '18, which is simply the F.Y. '18 number that is extrapolated and inflated across the out years" (Department of Defense, 2017). Similarly, the support for Hypothesis 1 cannot be attributed to the role of top-level leadership as the same press conference warned that "[t]he secretary has not spent any time at all looking at anything beyond F.Y. '18 to date" due to an ongoing process up updating the larger strategy" (Department of Defense, 2017).⁶⁶

This result underlines the value of making the detailed FYDP RDT&E and procurement projections accessible, as they contain information that cannot be derived simply from looking at topline numbers or listening to leadership briefings. This result does not, in and of itself, contradict critiques of the FYDP as on average only of three to four fifths of the estimated changes materialized. Moreover, an alternate interpretation of the absence of high level attention is not that it risks undercutting the FYDP process but that the Office of Secretary of Defense letting the services run the PPBE process may have actually been more effective at capturing their preferences in a way that may not be replicated in other administrations. Nonetheless, the significance of the model suggests that the priority signal in the FYDP is worth listening to and that further efforts to increase its transparency and adjust for its weaknesses would improve communications to policymakers, analysts, and industry.

⁶⁶ That said, John Roth went on to specify that "the goal was that new defense strategy would be in -- would inform anything beyond F.Y. '19 in terms of the F.Y. '19 to 23 program, for example" (Department of Defense, 2017). Thus, further analysis on the significance of a changing strategy with regards the predictive value of the FYDP projections released in FY will have to await the availability of actual FY 2020 spending numbers.

• Bibliography

Blakelye, Katherine and Lawrence Korb (2014). *The War Chest: War Funding and the End of the War in Afghanistan*. Center for American Progress. <u>https://cdn.americanprogress.org/wp-content/uploads/2014/10/WarChest-report.pdf</u>

Doubleday, Justin. (2020, February 25). DOD proposes new pilot programs to fund software through single budget activity. Inside Defense. <u>https://insidedefense.com/daily-news/dod-proposes-new-pilot-programs-fund-software-through-single-budget-activity</u>

Enthoven, A. C., & Smith, K. W. (2005). How much is enough?: Shaping the defense program, 1961-1969. Rand Corporation.

Harrison, Todd. (2016). Defense Modernization Plans through the 2020s: Addressing the Bow Wave" CSIS January 2016 https://www.csis.org/analysis/defense-modernization-plans-through-2020s

Harrison, Todd and Seamus Daniels (2017 December). Analysis of the FY 2018 Defense Budget. Center for Strategic and International Studies.

http://defense360.csis.org/wp-content/uploads/2017/12/FY-2018_DefenseBudgetAnalysis_FULL_WEB.pdf

Harrison, Todd and Seamus Daniels (2020). *Analysis of the FY 2020 Defense Budget and Its Implications for FY 2021 and Beyond*. Center for Strategic and International Studies. <u>https://csis-prod.s3.amazonaws.com/s3fs-public/publication/HarrisonDaniels_FY20DBA_v4.pdf?FA48YSbbC3Vj.aH7pKhpdIbU560Q9QUZ</u>

Hebert, Barbara. (2011, August 29). Appropriations, Congress & Budget Execution, "Color of Money 101" [PowerPoint Slides]. Defense Acquisition University. <u>https://apps.dtic.mil/dtic/tr/fulltext/u2/a557591.pdf</u>

Hunter, Andrew (2019, December 13). *Bad Idea: Blaming OCO for Our Defense Budget Blunders*. . Center for Strategic and International Studies. <u>https://defense360.csis.org/bad-idea-blaming-oco-for-our-defense-budget-blunders/</u>

Jordan, Leland G. (2015). Systematic Fiscal Optimism in Defense Planning. Defense ARJ, 22(3), 272-292

Mann, Christopher T. (2017). *Defense Primer: The National Defense Budget Function (050)*. (CRS Report No. IF10831). Congressional Research Service. <u>https://fas.org/sgp/crs/natsec/IF10618.pdf</u>

McCormack, Lee. "Performance Budgeting in Canada." *OECD Journal on Budgeting Volume* 7, no. 4 (2007). http://www.oecd.org/canada/43411424.pdf

McGarry, Brendan W. (2020). *Defense Primer: Planning, Programming, Budgeting and Execution (PPBE) Process* (CRS Report No. IF10429). Congressional Research Service. <u>https://fas.org/sgp/crs/natsec/IF10429.pdf</u>

McGarry, Brendan W. and Epstein, Susan B. (2020). *Overseas Contingency Operations Funding: Background and Status* (CRS Report No. R44519). Congressional Research Service. https://crsreports.congress.gov/product/pdf/R/R44519/10

McGarry, Brendan W. and Peters, Hedei M. (2019). *Defense Primer: Future Years Defense Program (FYDP)* (CRS Report No. IF10831). Congressional Research Service. <u>https://fas.org/sgp/crs/natsec/IF10831.pdf</u>

McGarry, Brendan W. and Peters, Hedei M. (2020). *Defense Primer: Procurement* CRS, (CRS Report No. IF10599). Congressional Research Service. <u>https://crsreports.congress.gov/product/pdf/IF/IF10599</u>

Mehta, Aaron (2020, April 4) *Pentagon denies it seeks to hide future budget information*. DefenseNews <u>https://www.defensenews.com/pentagon/2020/04/03/pentagon-denies-it-seeks-to-hide-future-budget-information/</u>

Department of Defense (2017, May 23). Department of Defense News Briefing on the President's Fiscal Year 2018 Defense Budget [Interview Transcript]. <u>https://www.defense.gov/News/Transcripts/Transcript-</u> View/Article/1191830/department-of-defense-news-briefing-on-the-presidents-fiscal-year-2018-defense/

Sargent, John F. (2020) *Defense Primer: RDT&E* (CRS Report No. IF10553). Congressional Research Service. <u>https://crsreports.congress.gov/product/pdf/IF/IF10553</u>

Tate, David M. (2016, April). *Acquisition Cycle Time: Defining the Problem*. (IDA Document NS D-5762 H 16-000350). Institute for Defense Analysis. <u>https://www.tandfonline.com/doi/abs/10.1080/14702436.2018.1497445</u>

U.S. Department of the Army. (2016). *Planning, Programming, Budgeting, and Execution* (AR 1-1). <u>https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/r1-1_FINAL.pdf</u>