CSIS-CM-19-198

August 2019

Measuring Service Contract Performance: Preliminary Findings on Effects of Service Complexity, Managerial Capacity, and Prior Relationship

Project Director:
Andrew P. Hunter

Authors:
Gregory Sanders

Jonathan Roberts

Justin Graham

Lindsay McDonald

Contributing Authors:
Zach Huitink

Xin Yuan

Robert Karlén

Xinyi Wang

A Report of the
CSIS Defense-Industrial Initiatives Group

CSIS | CENTER FOR STRATEGIC & INTERNATIONAL STUDIES

About CSIS

Established in Washington, D.C., over 50 years ago, the Center for Strategic and International Studies (CSIS) is a bipartisan, nonprofit policy research organization dedicated to providing strategic insights and policy solutions to help decisionmakers chart a course toward a better world.

In late 2015, Thomas J. Pritzker was named chairman of the CSIS Board of Trustees. Mr. Pritzker succeeded former U.S. senator Sam Nunn (D-GA), who chaired the CSIS Board of Trustees from 1999 to 2015. CSIS is led by John J. Hamre, who has served as president and chief executive officer since 2000.

Founded in 1962 by David M. Abshire and Admiral Arleigh Burke, CSIS is one of the world's preeminent international policy institutions focused on defense and security; regional study; and transnational challenges ranging from energy and trade to global development and economic integration. For the past seven years consecutively, CSIS has been named the world's number one think tank for international security by the University of Pennsylvania's "Go To Think Tank Index."

The Center's over 220 full-time staff and large network of affiliated scholars conduct research and analysis and develop policy initiatives that look to the future and anticipate change. CSIS is regularly called upon by Congress, the executive branch, the media, and others to explain the day's events and offer recommendations to improve U.S. strategy.

CSIS does not take specific policy positions; accordingly, all views expressed herein should be understood to be solely those of the author(s).

Acknowledgments

This material is based upon work supported by the Naval Postgraduate School Acquisition Research Program under Grant No. HQ0034-18-1-0009. 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 Naval Postgraduate School nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government. The study team would like to thank Madeleine Gagne for her support of this paper.

Abstract

Services contracts have a distinct set of challenges relating to the uncertainty and the challenges of measuring performance. Past researchers identified three overarching characteristics of interest: service contract complexity, contract management capacity, and the relationship between the buyer and the contractor. Researchers have often turned to surveys of government contracting personnel to take on the challenge of measuring service contract performance. This report takes a large-dataset, quantitative approach to looking at service contract outcomes derived from information in the publicly available Federal Procurement Data System.

The report found that the relationship between vendor and customer, as measured by the number of how many of the past seven years involved interaction, had the most consistent positive effects on performance across multiple metrics. Secondarily, the service code invoice derived from the Inventory of Contracted Services, was used as a proxy for service contract complexity and estimated likelihood and sizes of ceiling breaches as well as a greater likelihood that all options on a contract would be exercised, among contracts with some exercised options.

Table of Contents

Tables of Figures and Tables

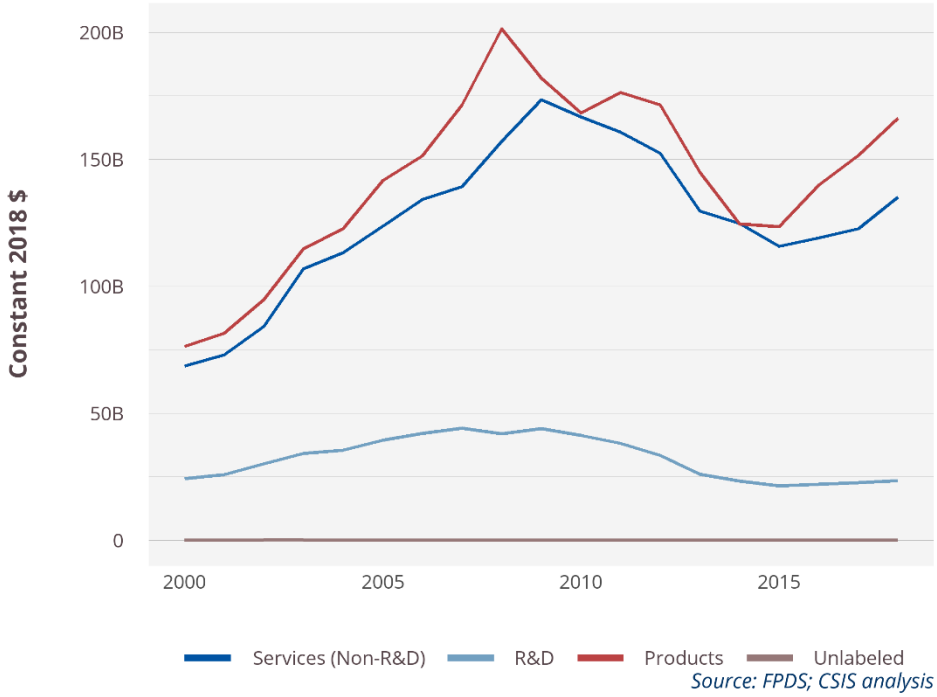
Table of Figures

Table of Tables

1 Introduction

Services contracts have long been understood to be distinct in key ways from their product counterparts in ways that add ambiguity and their own sets of challenges. Products are countable or otherwise objectively measurable, and while testing to see whether they meet requirements can prove complicated and controversial, there is at least a common item being argued over and measured. Service contracts inherently put more attention on the qualitative aspects of labor. Simple service contracts, like transportation or custodial services, have straightforward results to evaluate but can nonetheless introduce a host of concerns if, for example, taking place in contingency environments such as Afghanistan. Even familiar services like construction often must be evaluated not just on the quality of the final product but also the creation process, which is often not contained in an easily measurable outcome and can bring a host of disruptions. The most challenging services can be those that do something new or ill-defined, where trying to put all the details in the contract at the outset might not only be an exercise in futility, but actively counterproductive. In such situations, the buyer and contractor have to solve problems together that were not fully anticipated when the contract was initiated.

Figure 1 Defense Contract Obligations by Area, FY2000-FY2018



Source: FPDS; CSIS analysis

This report addresses three broad categories of characteristics that distinguish services contracting: service contract complexity, contract management capacity, and the relationship between buyer and vendor. These characteristics are highly relevant to government contracting but are also key concerns for service contracting in the private sector. Figure 1 shows the extent of DoD service contracting compared to products and R&D. While services contracts obligations have not regained their Iraq war and pre-budget cap peaks, they nonetheless constitute more $115 billion in annual obligations for every year since 2015.

For the U.S. federal government, and in particular the DoD, services constitute a significant portion of contract spending but are often a lower priority than contracting for products or R&D from a regulatory

and policy perspective. As Figure 1 shows, services by themselves are consistently secondary to product acquisition but the two have obligations on the same order of magnitude. This observation about a lack of policy focus comes not just from critics in Congress, which has a range of concerns about services contracts, but also from the DoD itself, where improving services acquisition tradecraft was a prominent part of the Better Buying Power initiatives. Some of the history of these acquisition reform efforts can be seen in Rhys McCormick (2015) "Measuring the Outcomes of Acquisition Reform by Major DoD Components" but suffice to say the problems of services contracting have long been a known issue.

While the prioritization of major defense acquisition programs over services acquisition is specific to DoD, the challenges of services contracting are universally acknowledged: in the private sector, by sectors and levels of government, and in nonprofit organizations. Trends in the wider economy also suggest that services may be further growing in importance, as software as a service and space launch services compete with acquisition approaches that would treat rockets or software packages as products. This study takes a new quantitative look at services contract performance by employing the Federal Procurement Data System, an open source transaction database with records of over a million service contracts within the past decade. This large dataset approach builds on past research regarding the public and private sector that often relied on surveys with smaller samples or case studies.

1.1 Scope

The research project seeks to answer the following questions:

Under what circumstances are services contracts likely to encounter challenges, as measured by terminations and cost ceiling breaches, or prosper, as measured by the exercising of options?

What services contracting policy choices influence these outcomes, for better or worse?

This iteration of the study focuses specifically on DoD contracts within a 2008 to 2015 study period. Past CSIS work with the contract dataset have covered both products and services contracts, but to better focus on the challenges of services contracting, this effort has focused on three especially relevant categories of characteristics:

First, service contract complexity, which as discussed above can indicate the difficulty in specifying or monitoring the performance of a service. A more complex service may require technical expertise to create contracts or evaluate outcomes and on top of that may need a collaborative partnership between the government and a contractor to have a chance at success.

Second, contract management capacity covers that capabilities and approaches by the contracting organization are associated with high service performance. Overcoming the principal-agent problem, that is the sometimes divergent interests between the government and contractor, including contract formation, monitoring, and enforcement where necessary but also mastery of approaches to encourage cooperative problem solving.

Third, the relationship between the government and the contractor, as touched on in the prior to categories, is central to service contract success. With complex service contracts, fully specifying what's necessary and then enforcing contract terms is not enough and excessive reliance on contracting

provisions can be counterproductive. However, while cooperation in the contracting relationship is important, a larger environment of competition is key to ensuring proper vendor incentives.

To measure contracting outcomes, this study builds on an approach developed in past CSIS studies that looked at the terminations and ceiling breaches. This contract centric approach is expanded in two ways. First, the look at ceiling breaches now considers both whether they occur and a new measure of their magnitude for contracts that experience a breach. Second, this paper adds a measure for positive performance by looking at exercised options. Options are portions of a contract that are specified but not awarded when the contract starts, leaving the government the choice of whether to invoke them during the tenure of the contract. Options are invoked in lower performance situations as well, but as a general rule, a contract with more options invoked is one where the customer desires more from the contractor and the contractor is happier to have the additional revenue, or at least not to have the transaction costs of winning a new contract or task order to receive that revenue.

2 Literature Review

This literature review will clearly delineate the different aspects of services contract management in several sections. In general, this includes service contract complexity, contract management capacity, and customer-vendor relationships.

There are incredibly few comprehensive definitions for contract management. For a broad definition, contract management may be defined as "all activities performed by the government . . . that are relevant to contracts with private or nonprofit organizations . . . such as writing or creating the Invitation to Bid or Request for Proposal, devising a rating system for bid responses, rating the bid responses, awarding the contract, additional negotiations leading to a signed contract, and contract administration" (Lawther 2002).

2.1 Service Contract Complexity

It stands to reason that the relative complexity of a given contract is a determinant of the level of cost (in labor, funds, or both) required by the government to effectively manage it, and in this vein, the literature delineates between low-, mid-, and high-complexity. For low-complexity contracts, “specificity and monitoring are simple and undemanding” (Joaquin and Greitens, 809). “Under mid- complexity scenarios, requests for proposals are more detailed and specific, and managers need to possess more technical expertise.” (Joaquin and Greitens 2012, 809). For highly complex services, and when understanding of the service delivery means are not clear, the agency and the contractor should enter into a true public-private partnership and recognize that the service to be provided will evolve in a dynamic manner, echoing studies by Fernandez (2007, 2009) and Joaquin and Greitens (2012, 809)

High levels of task complexity and uncertainty at the federal level are associated with those tasks where the government requires a definitively new service or capability. From the defense perspective, this could be new software architecture, an emerging hardware technology, or specified research and development. Such high-complexity contracts complicate the ability of contract managers to write contractual requirements that are comprehensive and highly detailed, which most literature has presumed is a necessity in successful contracting. The results are mixed on the need for specificity as

Brown and Potoski (2003, 2006) find support for this in overall contracting while Fernandez (2007, 2009) determines that asset specificity is unrelated to service contracting success. Moreover, even moderate levels of complexity and uncertainty increase the likelihood that some of the contract requirements will be erroneous. A successful contract, then, may require that the principal and agent interact frequently to make "necessary adjustments in work processes, performance standards, quantities, and prices" and fill in the gaps in the contract (Fernandez 2007, 1127). One additional consideration is that, contracting for management tasks can carry a large amount of risk, specifically that the government will enter into a monopoly relationship with the vendor (Brown and Potoski 2006, 327). This could come about if the government no longer retains sufficient management technical capacity to make key decisions without its contracting partner.

2.2 Contract-Management Capacity

While there are various definitions of contract-management capacity in the literature, many of them do not create a complete picture of the actual scope of managing contracts. The all-encompassing definition, as is required here, is provided by Brown and Potoski (2003):

> "Contracting is not a one-size-fits-all proposition. The success or failure of any alternative service-delivery arrangement likely depends on how well governments can manage the entire contract process, from assessing the feasibility of contracting through implementation to monitoring and evaluation-activities that require strong government contracting capacity. Governments investing in contract-management capacity may be better positioned to harness the promise of effective contracting while avoiding its pitfalls" (Brown and Potoski 2003, 153).

Governments invest in contract-management capacity because contracting is a complex process, fraught with potential problems and pitfalls. In fact, "governments can respond to poor conditions by investing in the managerial capacity to identify suitable situations for contracting, negotiate strong contracts, and monitor vendor performance." (Brown and Potoski 2003, 162) Romzek and Johnston (2002) determine seven factors that positively influence service-contracting effectiveness: competition among providers, resource adequacy, planning for performance measurement, training for state contract managers, evaluation of contractor staffing capacity, evaluation of contractor financial management capacity, and theoretical rationale for reform. These and other responsibilities of the government as contract managers have been discussed, but they often fall into general bins. The overall literature expresses a range of opinions on the bins that explain contract-management capacity. This is largely due to their different definitions and measures. However, the most popular systems come from two sets of authors. Brown and Potoski (2003) determine 3 subfields of contract-management capacity: assessment, implementation, and evaluation capacities. Yang, Hsieh, and Li (2009) build on this model by adding another measure to Brown and Potoski's and renaming it. Therefore, formulation capacity for them is the same as implementation capacity for Brown and Potoski while Yang et al determines implementation capacity to be the ability of the government to create and sustain a public-private partnership for contracts. This is an important delineation because many findings detail the effects of public-private partnership ability which is the capacity of the government to manage relationships and, as such, falls under contract-management capacity.

Diving into the separate bins, contract assessment is first in the process. Yang et al find that increased assessment capacity is positively associated with cost reduction, but it is not associated with efficiency increase or quality improvement. (Yang, Hsieh, Li 2009) Additionally, Fernandez (2007), using substantively weighted least squares to statistically identify the top performers, finds that ex ante evaluation (an assessment responsibility) is a significant determinant of the most successful contracts. Moreover, Yang et al find a time component to management capacities. For agenda setting, the "process during which the values and preferences of stakeholders are manifested and compromised," the impact on cost reduction decreases as time passes for assessment. (Yang, Hsieh, Li 2009, 687) Another aspect of assessment is the determination of asset specificity from the outset. Planning asset-specific requirements for either end product or production tasks requires significant expertise and insight. Brown and Potoski (2006) state that "managers need to master the complex production process in order to ensure that production tasks integrate with other processes required to deliver the service." However, Fernandez (2007, 2009) finds under many different statistical methods (OLS, SWLS, 2SLS) that although it is common for government managers to develop design specifications and hand it to industry to build, this is a retrograde approach, which "makes it impossible to hold contractors responsible for delivering solutions that work, because as long as what's delivered meets the specifications, it's the government's fault if the products don't work". Interestingly enough, Fernandez (2007) finds that task uncertainty is a very significant factor in diminishing contract performance. This differentiation indicates that defining the asset as specifically as possible does not necessarily define what the government wants contractors to accomplish in the contract. Additionally, this suggests another side where contract specification capacity is not the problem, but rather where a cultural shift away from strictly measurable design specification into more of a capability-based contract could produce more efficient and higher quality products.

As for implementation (formulation) capacities, the research is fairly anemic. These responsibilities vary but generally fall under "setting a fair bidding process, identifying the best-fit contractor, and reaching an excellent contract" (Yang, Hsieh, Li 2009, 683). Yang et al (2009) find that an increase in formulation capacity does not affect contract performance. This is speculation, but it could be because much of this is process dependent, and much of where the human capital of capacity comes into effect is in the agenda setting stage, where a high-level of skill and expertise is needed, whereas implementation capacity, required for creating the actual final contract, could be most affected by personnel numbers. The agenda setting stage requires a few highly capable people for negotiation and technical requirements creation while the act of formulation of the actual contract could require a larger number of less skilled workers. Nonetheless, both stages could have the same overall contract-management capacity. One illuminating point by Fernandez (2009, 86) regarding the system rather than the workers indicates that "ongoing competition between contractors during the implementation phase, rather than ex ante competition during the bidding phase, appears to be the form of competition that improves overall contracting performance."

As for public-private governance capability, the literature supports the importance of the contract enhancing capacities of government and the private sector working together during the contracting period to increase the quality of the service. Speaking to the ability of the government in this respect, Yang et al provides the most evidence. Their use of "implementation phase" refers to "government agencies' active, ongoing involvement in or support for the contractor's operation." Put another way,

the key implementation question for the government is "how can we help the contractor succeed" (p. 684)? They found that "implementation capacity has a curvilinear relationship with cost saving" with moderate levels having the most benefits (p. 691). They further find that the role of time increases the importance of implementation capacity:

> However, the function also shows that time has a magnifying effect. The impact of implementation capacity on cost reduction turns positive after the interaction efforts gain steam, and then, the impact of implementation activities on efficiency increases accelerates in that the benefits multiply as time passes, indicating that efforts to develop collaboration and mutual support will have long-term advantages (Yang, Hsieh, Li 2009, 691).

Their results suggest that the government's ability to influence mechanisms such as trust, parallel expectations, and joint action hold significant potential to improve contracting performance. Additionally, Fernandez (2009) finds evidence that joint problem-solving efforts are positively correlated with overall contracting performance.

In the last bin we have evaluation or monitoring capacity, which is the ability of the government to monitor a contractor's performance and/or enforce the contract. Yang, Hsieh, Li (2009) suggest that the impact of the evaluation capacity depends on its strength: a strong evaluation system is beneficial, but a weak evaluation system does no good. They also show that a "strong evaluation capacity may promote cost reduction and efficiency increases but may not help improve quality" (Yang, Hsieh, Li 2009, 691). However, they also find that the benefits of evaluation activities decrease over time. This suggests that a contract needs more evaluation in the beginning, but that once the expectations are clearly established, things run much more smoothly. On the other hand, Fernandez (2007, 2009) do not show any significance of the impact of monitoring activities in either scope or intensity while the enforcement mechanism is mixed. Fernandez's (2007) findings indicate that the most successful contractual relationships perform at higher levels when public managers make periodic use of the 'stick' to enforce the contract:

> "In fact, among the high performing cases, tactics such as imposing financial penalties and threatening to terminate the contract seem to enhance contracting performance more than alternative means for resolving disputes, such as negotiation and mediation, since the coefficient for reliance on alternative means for resolving disputes is not statistically significant" (Fernandez 2007, 1135).

Then, Fernandez (2009) finds the complete opposite for services for the exact same dataset. This would seem completely contradictory, but Fernandez (2007) utilized SWLS to identify the top performers and then conducted an OLS analysis of the whole sample and a WLS analysis with the weights going to the high-performers. Negotiation and mediation seem to work for the overall sample, but when compared to the most successful contracts, legal enforcement and threats win the day. This heavily indicates the need to differentiate between the different types of contracts.

As for services specifically, much of the earlier literature evaluates service contracting as augmenting contract-management capacity. Service delivery contracting includes producing the service but can also include delegating to vendors management responsibilities, such as monitoring outcomes. "All service delivery management need not occur within government, though effective contracting clearly requires

that governments maintain some contract-management capacity. For example, even though governments can transfer some monitoring responsibilities to vendors, they likely still need to monitor their vendors' performance to some degree" (Brown and Potoski 2006, 324). Alternatively by contracting for management responsibilities contracts and introducing third party verifiers, governments may accumulate more monitoring than they would have been able to conduct on their own. (Brown and Potoski 2006) In the case of easy-to-measure services, contract managers can focus more on outcome monitoring and less on the actual production of the service. In such cases, external monitoring becomes an attractive option, contract managers can quickly check the vendors' intensive reports against their own outcome observations. Unfortunately, these cases are only available when services are easily monitorable with specific requirements. Otherwise, delegating complex monitoring to the vendor is obviously easily susceptible to the agent opportunism problem (Brown and Potoski 2006).

Lastly, it is important to make the distinction between in-house contract-management capacity and overall government capacity. While some cases of governance "may indeed see the abdication of management responsibilities, cutting management staff and activities does not necessarily translate into reducing management capacity" (Brown and Potoski 2006, 325). Earlier literature indicated that government outsourcing the capacity to deliver the service diminished their direct capacity to manage the service. Yet, governments can, in fact "buy" management services to adequately address their own deficits in capacity (Brown and Potoski 2006, 324). This is corroborated in a U.S. Government Accountability Office (2009) report on the defense acquisition workforce. For example, at the national level, many federal agencies now employ third-party evaluators to assess the quality of production activities for which they have contracted (e.g., information technology), a practice often referred to as "independent verification and validation." Therefore, while the government may have reduced their in-house capacity, the overall capacity remains the same or even increases at lower costs. The tradeoff is that contracting and other forms of alternative service delivery do not eliminate the need for contract-management capacity, but instead create an imperative for new types of management capacities. These problems may be more likely to occur in cases such as:

- Limited or no competition among potential vendors
- Contracted products and services that are difficult to specify and describe in written contracts
- Vendors that have special knowledge or skills about the product that is unavailable to public managers
- Public managers that have a hard time monitoring vendor performance once the contract has been let. (Brown and Potoski 2003, 154)

2.3 Customer-Vendor Relationship

As has been mentioned above in the capacity section, trust, joint problem-solving, and public-private partnerships have a huge impact on contracting performance. The earlier literature showed theoretical divides on the principal-agent problem and how government and the private sector should behave. Johnston and Romzek (1999) conclude that the agent's (contractor's) response to the principal's monitoring system depends on many factors such as the reliability and credibility of the system as well as the principal's willingness to enforce punishment. This game theory understanding of the principal-agent problem is complicated in government contracting as elected officials and networks of contractors add multiple layers of accountability. Additionally, "scholars have found that the overreliance on legal

means of conflict resolution may evoke conflict, opportunism, and defensive behavior" (Yang, Hsieh, Li 2009, 686). Brown and Potoski (2006) provide evidence that longer contracts may also begin to mirror monopoly relationships, exposing governments to the risk that vendors will shirk their responsibilities. Moreover, "governments that entered into longer contracts spot checked vendor performance only 68 times a year on average, compared to 95 times a year, on average, for governments operating with short term contracts." (Brown and Potoski 2006, 336)

As the literature matured and developed ways to measure the relationship of the government and contractors, the tone adapted. Fernandez (2007) found that the effect of joint problem-solving on contracting performance is greater among the most successful contractual relationships than in the average case. Since contract managers work more closely with the contractor's staff to solve issues that arise, the level of contracting performance tends to increase. In a later study, Fernandez found that trust has a positive independent effect on overall contracting performance and is the largest coefficient in his model. (Fernandez 2009) However, some of the literature on trust suggests the possibility of an endogenous relationship between trust and performance (Fernandez 2009, 86). Fernandez then conducted a 2SLS regression to account for endogeneity. He determined that contract duration does not appear to interact with trust. That is, "the effect of trust on contracting performance does not increase as the relationship evolves over time." (Fernandez 2009, 87) Additionally, he discovered that monitoring activities and trust do not serve as substitutes. Going deeper into the model as it is of great interest to this study, "factor analysis was used to develop multi-dimensional measures of communication, and joint problem-solving efforts after contract award. Since contracts of longer duration can facilitate learning and allow the parties more time to iron out the kinks in service delivery, the model also includes a measure of the duration of the contract, measured in months." (Fernandez 2007, 1127)

Finally, there are some scattered findings throughout the literature on the effect of management responsibilities on relationships. When contracts specify in great detail how a service should be delivered, the contractor may have less incentive to innovate. Additionally, "lengthy negotiations can damage the relationship between partners and inhibit their adaptation to unanticipated situations" (Yang, Hsieh, Li 2009, 686). Yang et al argue that "information searching, contract negotiation, and contract writing" give rise to transaction costs that can offset their cost-saving benefits, and that overuse of contracts for enforcement can curtail the development of trust and collaboration. (Yang, Hsieh, Li 2009, 690) As time goes on, what is more important is to develop authentic partnerships between the government and the contractor so that information can be shared and collaboration achieved. (Yang, Hsieh, Li 2009, 693) As for efficiency, the overuse of contracts for enforcement may decrease efficiency, but, developing authentic partnerships during the implementation phase counteracts this, and the effect increases over time. Yang et al go so far as to suggest that "the best contracting strategy for government is to depart from pure contracting and shift to a collaborative model such as public-private partnerships" (Yang, Hsieh, Li 2009, 692). Mentioning sub-relationships, Fernandez (2007) also examined the use of subcontractors because "arrangements involving multiple subcontractors imposes additional burdens on the prime contractor, including higher coordination costs, the likelihood of delays, and sometimes even conflict over the choice of goals and means, all of which ultimately weaken performance" (Fernandez 2007, 1129). The use of multiple subcontractors was significant in the overall OLS sample but was not significant in the case of high performers. This

indicates once again that it is paramount to find the distinctions between types of contracts as different types of contracts may have different mechanisms to develop trust.

3 Conceptual Framework and Hypothesis

This paper posits and tests a conceptual argument linking three categories of characteristics with services contract performance: first, service complexity; second, contract-management capacity on the part of the buyer; third, the strength of the relationship between the buyer and the contractor. By specifying all three characteristics, the argument captures the inherit challenges of services contracting, those most under the control of the buyer, and those most of interest to individual vendors. FPDS does not contain direct measures of these variables, and so the paper introduces proxies for each under the relevant hypotheses.

3.1.1 Service Complexity

The complexity of the underlying service can introduce challenges in two broad ways. First, it may raise the technical expertise required from acquisition officials. A simple service, such as lawn mowing, can be easily specified and overseen while a more complicated service, like maintaining aircraft, requires a higher level of understanding and assurance, as important problems might not be immediately visible. The second aspect of complexity is the challenge of specifying the service in clear and comprehensive terms. When acquiring new services or ones that otherwise involve significant uncertainty, acquisition officials and contractors cannot simply rely on the initial performance work statement to deliver a successful outcome but will have the flexibly to incorporate changing conditions or new information. This greater requirement for partnership asks more of both buyer and vendor and leaves much room for disagreement and conflicting interests. In both cases, this complexity makes the work more demanding and thus, all else equal, raises the risks of negative contracting outcomes.

> H_1: As service complexity increases (decreases), the likelihood of cost ceiling breaches and terminations increases (decreases) and the likelihood of exercised options decreases (increases)

The paper employs two labor-based measures to attempt to capture service complexity. Service contracting inherently emphasizes labor and measures of pay, and number of employees is a metric that can be relevant across disparate forms of services contracting.

The first measure is the average salary for the North American Industrial Classification System (NAICS) detailed industry that the contract is classified under. Higher salaries may have multiple sources, but one of them is the difficulty of the work and the experience and education required.

> H_{1A}: As average salary increases (decreases), the likelihood of cost ceiling breaches and terminations increases (decreases) and the likelihood of exercised options decreases (increases)

The second measure is more services contracting specific: average cost per employee. At this stage of the research, the average cost is calculated based on averages for the given product or service code, though the study team hopes to incorporate direct contract level measures where available in future iterations. It employs an existing government metric, called the invoice rate, that approximates how much the government is charged annually for each comparable full-time employee supporting a service

contract. A service contract with a large number of lower-paid staff would have a lower invoice rate while one that employed a small number of experts or that had extensive capital costs would have a higher invoice rate. Similarly, a service contract that was just making contracting personnel directly available to the buyer in government facilities and using government equipment would, all else equal, have a lower invoice rate than a than one that also promised a full package of services and charged overhead for the infrastructure in place to help deliver them. As with average salary, this hypothesis assumes that scarcer labor or labor acquired at a greater premium, all else equal, indicates a more complex service.

> H_{1B}: As average cost per employee increases (decreases), the likelihood of cost ceiling breaches and terminations increases (decreases) and the likelihood of exercised options decreases (increases)

3.1.2 Contract Management Capacity

Contract management capacity can manifest in a variety of forms, including assessment, contract formulation capacity, evaluation, and ability to sustain a public-private partnership. The literature affirmed the importance of this capacity, in particular for the more complex services discussed for H_1.

> H_2: As a contracting office's contract management capacity increases (decreases), the likelihood of cost ceiling breaches and terminations decreases (increases)

The first measure considered is the only one where FPDS reports on one of the capabilities discussed in the literature review: performance-based services acquisition (PBSA). Defined in FAR 37.601, PBSA tracks multiple measures relevant to public-private partnership governance including the foundation of how the contract is defined. A performance-based services acquisition "[d]escribes the requirements in terms of results required rather than the methods of performance of the work" (GSA Federal Procurement Data System, 2017, p. 52). Other characteristics included measurable performance standards, plans for monitoring, and the potential for monetary adjustments depending on the quality of the output.

> H_{2A}: As contract office usage of performance-based services acquisition increases (decreases), the likelihood of cost ceiling breaches and terminations decreases (increases) and the likelihood of exercised options increases (decreases)

For the other forms of contract-management capacity, specific measures employed by prior surveys and case studies are not available within FPDS, and headcount data for contracting officers is not publicly available at the contracting office level. To capture this important but elusive variable, this paper employs a measure that scales based on the contracting office's history. This approach assumes that the throughput with a given type of product or service code correlates with the development of technical expertise. As the prior section covered, complexity and expertise requirements can vary greatly from one category to another, and a contracting office may have high capacity in one area that would not translate to a new area.

> H_{2B}: As the share of contracting office obligations for a given service code increases (decreases), the likelihood of cost ceiling breaches and terminations decreases (increases) and the exercised options increase (decrease) for that service.

3.1.3 Extent of Prior Relationship

The importance of partnership, trust, and handling difficult problems and uncertainty together naturally lead into the last characteristic: the relationship between the contractor and buyer. The literature suggests that a perfectly written contract is no guarantee of nor substitute for effective collaboration. In the absence of data directly on trust, this hypothesis focuses on the level of interaction that provides the opportunity to build a deeper relationship.

> H_3: As the extent of the government's prior relationship with its vendor increases (decreases), the likelihood of cost ceiling breaches and terminations for that partnership decreases (increases).

The first measure is the number of past years of the relationship between the contracting office and the contractors, with a single transaction between the two in a given fiscal year enough to qualify a given year as part of a continuing relationship. The second measure is the number of actions on the vendor's contracts with that office in the prior year. Contract action counts vary wildly from contract to contract, but even if the obligated amount per action is small, they still represent more opportunities for interaction for the office and contractor.

> H_{3A}: As the number of past years a vendor has contracted with an office increases (decreases), the likelihood of cost ceiling breaches and terminations for that partnership decreases (increases).
>
> H_{3B}: As the number of contract actions a vendor has performed for an office in the past year increases (decreases), the likelihood of cost ceiling breaches and terminations for that partnership decreases (increases).

4 Data and Methods

4.1 Data Sources and Structure

The primary source of this paper is FPDS, which is the transaction database for U.S. government contracts including military and civilian as well as products and services. With some exclusions, such as classified contracts, the U.S. postal service, and the Defense Commissary Agency, U.S. federal government contracts above a $3,500 threshold are reported into FPDS. Services contracts are delineated using the product or service codes included in FPDS, and include R&D contracting for the purposes of this report. The study team maintains their own copy of the FPDS, which has been supplemented by the ad hoc search tool and information from various data dictionaries. This and past contract datasets are freely available for download for other researchers.

FPDS data has been supplemented using the Services Contract Inventory mandated by the 2010 Consolidated Appropriations Act (U.S. General Services Administration, 2019). The study team continues working on importing and matching contracts from both the civilian agency data held by the GSA and the separate DoD dataset. At this stage in the research, the analysis relies not on the contract inventory itself, which is only available for larger contracts in the first place, but on the invoice rates derived for product and service codes through the work of the U.S. Army. Those invoice rates are used on an annual

basis to estimate the number of comparative full-time employees for contracts in the inventory that lack more detailed data. They are broken out for both Overseas Contingency Operations, which are of special interest because they imply coverage of contractors supporting military operations overseas including those directly present in Iraq and Afghanistan.

This report uses a unit of analysis of individual service contracts and task orders.[1] These are identified in FPDS through the unique combination of a procurement identifier and, for task orders, a parent procurement identifier. The dataset is made up of completed contracts and task orders for services contracts for the Department of Defense, completed between fiscal years 2008 and 2015.[2] Many of the variables in the dataset have been built up and tuned over three CSIS reports on Fixed-Price contracts, industrial consolidation and competition, and crisis-funded contracts (Sanders, 2015; Sanders and Huitink, 2019; Sanders and Hunter, 2017). Services contracts are less numerically prevalent than their products counterparts, but still constitute 1.3 million contracts and task orders. At this stage of the research, 24.2 percent by count and 21.0 percent by value obligated are eliminated from the sample because of missing data. The study team believes that recent upgrades to USAspending.gov may enable a reduction in this missing data rate.

The exercised options outcome variable focuses on a narrower subset of contracts and task orders, namely those with unexercised options as of their initial transaction. This reduces the count tenfold, only 103 thousand contracts and task orders qualify. However, excluding these contracts from the options exercised sample is important because the choice of whether or not to include options in a contract is a contract formulation decision and not a direct reflection of performance on a given contract. The importance of contracts and task orders with options is affirmed by their value, they account for 23.7 percent of the total services dataset. Their missing data rate is similar to the overall dataset with a reversal between the metrics, data is missing for 20.1 percent of contracts and task orders by count and 23.7 percent by value. Henceforth in this study, for simplicity, both contract awards and task orders will be referred to simply as contracts, except in those cases where the distinction matters.

4.1.1 Inventory of Contracted Services

The Inventory of Contracted Services (ICS) is mandated across the federal government and has an obvious value to this project above and beyond the inclusion of the invoice rate variable. By statute, DoD has a separate inventory process from the standard GSA process that includes extracting contract data from the Federal Procurement Data System (FPDS) and the System for Award Management (SAM). The study team has analyzed ICS data from both DoD and GSA to better understand service contract complexity and found that each source has its own set of challenges. Generally, since FY 2012 DoD ICS

[1] The overarching contracts for indefinite delivery vehicles are also reported in FPDS but are not included in the dataset because the study team instead chose task orders as the unit of analysis.

[2] Completion is measured by having no modifications involving new obligations or any form of increased ceilings since the end of calendar year 2016, not counting terminations and close outs. In addition, to be counted as completed, contracts that were not terminated or closed out must have a current completion date before the start of calendar year 2017.

data includes comparable contractor full-time equivalents (CFTE) related information with clear ICS guidance and information available. GSA ICS data by comparison is easier to import because it is not spread across many somewhat inconsistent Excel Tabs. However, GSA ICS relies on supplemental documents for explanation and has not published these supporting documents at all for some years since the start of the ICS, posing difficulties in cross-checking and reference.

The main challenge in importing the data was the inconsistent format in which the data was reported and published, which potentially complicated the consolidation process before import to the CSIS database system, especially for validating data type, generating unique identifiers, etc. Together with widespread underreporting, especially in later years, utilizing contract level reporting proved impractical. Nonetheless, the study team was successful in importing and using key DoD guidance documents to extract annual service code-level invoice rates.

4.2 Measures of Dependent and Independent Variables

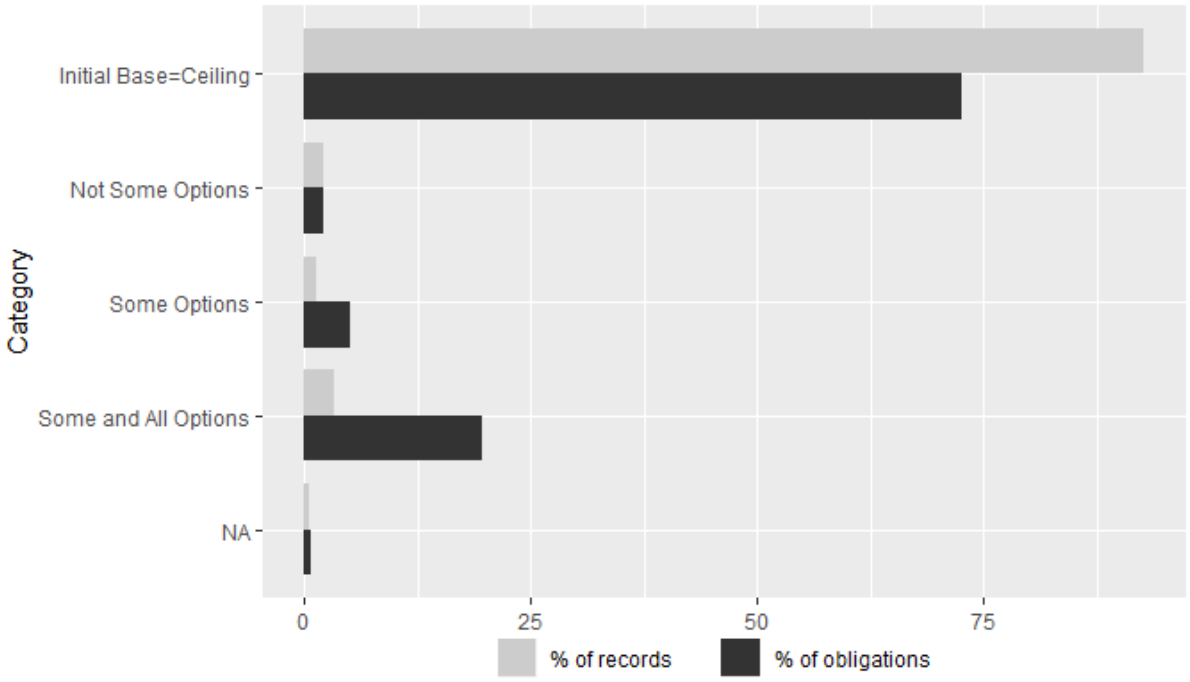
4.2.1 Dependent Variables

This paper includes three performance outcomes measured across five dependent variables: the likelihood that some options will be exercised and, in that case, that all options will be exercised; the likelihood that a ceiling breach will occur and, if a breach does occur, what will be the breach size.

Exercised Options

Exercised Options is tracked by a pair of metrics and, in contrast to terminations and ceiling breaches, is a positive measure of contract performance. They reflect that the buyer has chosen to acquire additional services within the scope of the original contract and is willing to pay a higher price as a result. One common source of options is multiple year contracts where the original “base” contract only covers the first year. Both government and contractor may assume that this extension will take place with a high degree of confidence, but in strictly legalistic terms the buyer is under no obligation to continue and may unilaterally allow the contract to end without the liability that may be incurred in a termination. The vast majority of contracts do not have options available to exercise at contract start, but those that do account for a size portion of spending, as is shown in Figure 2 below. In keeping with the premise of that exercised options signify a greater desire by the government to employ the vehicle and more revenue for the contractor, note that the some options as well as some and all options columns show substantially more spending than those contracts with available options that were left unexercised.

Figure 2 Distribution of Contracts by Availability of Options and Extent Exercised



The first metric is some exercised options. This metric is a binary variable that is true if the contract has passed one of two tests. First, if it has at least one transaction that meets all three of the following criteria:

1)The reason for modification is an exercised option, a supplemental agreement for work within scope, or a funding only action (collectively "steady scope" modifications).
2) The base and *exercised* options value of the contract increases as part of the transaction.
3) The base and *all* options value of the contract does not increase as part of the transaction.

The study team used this conservative definition in order to ensure that exercised options were clearly differentiated from cost overruns. As a secondary threshold, contracts with exercised options growth that exceeded the total ceiling growth, after setting aside close out and termination transactions, were counted as exercising some options. Thus by these two criteria, a contract could qualify if it had a exercised option with an appropriate reason for modification not accompanied by ceiling growth or if its exercised option growth was clearly greater than its ceiling growth.

The second metric is a binary variable that tracks whether all options were exercised. A contract can qualify by meeting either of two tests. First, if the base value of the contract plus exercised options meeting the three prong criteria described above exceed the initial contract ceiling, after accounting for administrative reductions.[3] Second, if the base and all exercised options value exceeds both the initial

[3] Specifically, if the sum of administrative and steady scope base and exercised options modifications is less than zero than that the exercised options measure is reduced by that amount.

contract ceiling and the cumulative contract ceiling, after setting aside transactions for terminations and close out. This metric rounds up to allow for some imprecision in the raw data, any contract that reaches 95 percent of the gap between the initial base and the relevant ceilings also is treated as having exercised all options. As can be seen in Table 1, two-thirds of contracts with available options exercise some options and three-quarters of contracts exercising some options go on to exercise all options.

Table 1 Frequency of Some Options, All Options, Ceiling Breaches, and Terminations

Variable	Comparison Group	Value	% of records	% of $s
Some Options	Contracts with Available Options	0 (No Qualifying Exercised Options)	32.05%	8.14%
		1 (Some Options)	67.89%	91.83%
		Not Available	0.06%	0.03%
All Options	Contracts Exercising Some Options	0 (Only Some Options)	26.79%	79.54%
		1 (All Exercised Options)	72.10%	79.54%
		Not Available	1.11%	0.29%
Ceiling Breach	Full Dataset	0 (None)	93.96%	76.60%
		1 (Ceiling Breach)	6.04%	23.40%
Terminations	Full Dataset	0 (Unterminated)	98.12%	97.54%
		1 (Partial or Complete Termination)	1.88%	2.46%

Note: Not Available data includes contracts for which exercised option or ceiling values could not be effectively calculated, for example steady scope and administrative reductions exceed increases.

Ceiling Breaches

Ceiling Breaches are tracked by a pair of measure that attempts to track the risk of cost increases. The measure for breach likelihood is a binary variable that is true if a contract's cost ceiling has increased as part of a change order or definitize change order (collectively "change orders"). Each transaction in the database tracks the change in obligated amount, contract ceiling, and base and exercised options as well as the reason for that change. As shown in Table 1, over 6 percent of services contracts experience a ceiling breach, a notably higher rate than the "1.04 percent" for the full DoD contracts dataset (Sanders and Huitink 2019, 19).

The measure for breach size is a continuous variable tracking the cost ceiling increase, measured in 2018 constant dollars based on the contract's starting fiscal year. For the breach size measure, the sample is limited only to those contracts that have experienced a breach, so a contract characteristic could estimate that a lower likelihood of breach but that a greater magnitude should one occur. While this is a smaller sample, over a fifth of all obligations in the dataset went to contracts experiencing breaches. As shown in Table 2 both the average and median breach size were around ten thousand dollars but the upper bound for DoD services in the period was well over one hundred million dollars in ceiling increase.

Both measures focus on change orders, rather than modifications for additional work, because the combination of a change order and an increase in ceiling suggests an unanticipated development that will cost the acquirer more money. These variables intentionally measures ceiling change, not obligations, because ceiling changes typically do identify the reason for the change but many obligations

are funding only actions that do not differentiate between whether the spending supports a prior change order, an exercised option, or other forms of new work. In those cases where there appears to have been a correction made via an administrative or steady scope ceiling modifications the ceiling breach size measure is reduced accordingly.[4]

Table 2 Ceiling Breach Size for Breached Contracts

Variable Name	Min	Max	Median	Geometric Mean	% of records NA	% of Obligation to NA records
Ceiling Breach Size	0	$367,080,377.48	$10,583.14	$9,750.76	3.55%	5.34%

Note: NA refers to not available data, for example because negative steady scope or administrative adjustments exceeded the initial increase in ceiling size.

Terminations

Terminations are measured using a binary variable that is set to true if a contract has experienced a partial or complete termination at any point in its lifespan. This includes terminations for default and convenience (partial or complete) as well as terminations for cause and legal contract cancellations.[5] Perhaps unintuitively, this can include both a traditional cancellation of a major weapon system and the cancellation and reassignment of a contract due to a bid protest. The overwhelming majority of terminated contracts have no obligations after their final termination date. However, among contracts with higher ceilings, substantial spending after termination is common and in line with Federal Acquisition Regulations Part 49, which allows for settlement spending, which may explain most of the post-termination spending in the data. The study team found no clear dividing line between partial and complete terminations and choose to include all instances of terminations whether they occurred at the start or end of a contract.

As shown in Table 1, both terminations are rare, although roughly twice as common as in the overall DoD dataset where "0.91 percent of contracts experienced at least one partial or complete termination"

[4] The specific process largely aligns with the one used for exercised options. If the sum of administrative modifications and steady scope modifications is less than zero than that the ceiling breach measure is reduced by that amount. However, if the reduction is made due to a change order, as part of a close out or termination, or for other reasons, the ceiling breach amount is not adjusted. The study team reviewed transaction records and contract totals and believes that administrative and steady scope net reductions are most likely to capture corrections or updates. This metric intentionally does reduce the size of the ceiling breach in response to descoping change orders or due to descoping or ceiling reductions as part a termination or close out.

[5] The reason for modification field, the source for termination labels, does differentiate between different types of terminations but groups partial and complete terminations together within those categories. Due to the rarity of the more severe forms of termination, that is terminations for default or cause, the study team did not differentiate between different degrees of terminations.

(Sanders and Huitink 2019, 19). Perhaps surprisingly, overlap between terminations and ceiling breaches is small.

4.2.2 Study Independent Variables

Service Contract Complexity

Detailed Industry Salary: Each contract in FPDS is labeled by its NAICS Detailed Industry category, the most granular level available in NAICS codes. The U.S. Economic Census provides enough data to calculate average wage, although it is only available every five years and thus has a variable lag based of one to five years based on the time since the last census.

Service Code Invoice rate: What is the average annual charge rate for comparable full-time employees? The invoice rate is available through the Service Contract Inventory and is dependent on U.S. Army calculations at the individual product or service code level or for the broad service category.[6] When the invoice rate for a specific product or service code is available for the prior fiscal year, that factor is used. When the invoice rate is available for a code but not for the prior year, the average across all years is imputed. For those codes with no reported invoice rates, instead the broad service code is used for that year if available, and an average of the invoice rate for all available years is used otherwise.

Among the most used service codes, the broad patterns in invoice rates align with expectations: "Operation/Dining Facilities" and "Custodial – Janitorial Services" have the lowest invoice rates. By comparison more complex services like and "Architect-Engineering Services" are fairly high.

Table 3 shows the descriptive statistics for these variables, which are logged and rescaled in the model.

Table 3 Detailed Industry Salary and Service Code Invoice Rate

Variable Name	Min	Max	Median	Geometric Mean	1 unit below	1 unit above	% of records NA	% of Obligation to NA records
Detailed Industry Salary	$9,996	$278,829	$64,591	$60,479	$31,677	$115,469	5.03%	0.62%
Service Code Invoice Rate	$7,370	$1,908,520	$167,919	$170,757	$68,483	$425,773	0.03%	0.00%

Contract Management Capacity

***Office Performance-Based*:** What share of office obligations for a given office were for Performance Based Services Contracting in the prior year?

[6] Product or service codes have four characters. Services codes start with a letter while product codes start with a number. The broad services category e.g. the letter Y for construction or the letter D for automated data processing, refers to the first letter of the services code.

***Office Service Code Experience*:** For any given contract, what percentage of obligations for the office went to contracts with the same product or service code over the past seven years? Table 4 shows the descriptive statistics for these variables, which are rescaled in the model.

Table 4 Office Performance-Based, Service Code Experience Descriptive Statistics, and Paired Years

Variable Name	Min	Max	Median	Arithmetic Mean	1 unit below	1 unit above	% of records NA	% of Obligation to NA records
Off. Perf.-Based	0%	100%	27.8%	33.9%	-26.1% *	93.9%	0.10%	0.00%
Off Serv. Code Exp.	0%	100%	1.9%	14.0%	-37.7% *	63.0%	0.10%	0.00%
Paired Years	0	7	4	3.49	-1.31 *	8.32 *	0.25%	0.52%

* 1 unit below values are less than minimal value for variable.

Extent of Prior Relationship

***Paired Years*:** For any given contract's vendor and office pairing, how many of the past seven years involved interaction between the vendor and the office? For a new relationship, this value would be zero. Table 5 shows the descriptive statics for this variable, which is rescaled in the model.

***Paired Actions*:** For any given contract's vendor and office pairing, how many contracting actions did the vendor perform for that office across all contracts in the prior year? Table 5 shows the descriptive statistics for this variable, which is incremented by 1 to make zeros eligible for logarithmic transformation and is then logged and rescaled.

Table 5 Paired Years and Actions Descriptive Statistics

Variable Name	Min	Max	Median	Geometric Mean	1 unit below	1 unit above	% of records NA	% of Obligation to NA records
Paired Actions	1**	7,806,579	27	34	0.5 *	2,355	0.25%	0.52%

* 1 unit below values are less than minimal value for variable. ** True minimum value is 0.

4.3 Empirical Approach

The study team has created five statistical models, one per dependent variable metric. The binary outcome metric, that is some options, all options exercised, ceiling breaches, and terminations, use a maximum likelihood logit multilevel model. The single continuous metric, the size of ceiling breaches, an ordinary least squares multilevel regression is used. The samples vary between models based on relevance: the some option models is limited to defense service contracts with initial options to exercise; the all options sample is limited to those contracts that exercise some options; terminations and ceiling breach likelihood use the entire 1 million record defense services sample; and finally the ceiling breach size model is limited to a sample of those contracts that experienced a ceiling breach. The varying models are used to allow the all options model to build off the some options results and the breach size model to build off the build likelihood results.

4.3.1 Summary of Included Control Variables

The term multilevel refers to the modeling approach that captures both differences between individual contracts and larger group. This approach adopts techniques employed by Andrew Gelman and Jennifer Hill (2017) and Nicolas Sommet and Davide Morselli (2017) that allow for a different intercept for each of the sectors, acquirers, places of performance, and start fiscal years of each contract. Acquirers refers to the hierarchical description of each customer: the individual of contracting ***Office*** (level 2) and the ***Agency*** they report to (level 3) Sectors refers to the hierarchical description of the type of service: from top to bottom, the *service area* (level 2), NAICS ***Detailed Industry*** (level 3), and NAICS ***Subsector*** (level 4).[7] ***Place of Performance*** and ***Start Fiscal Year*** are both captured by level 2 variables without further hierarchy. The multilevel groupings employed in this model and their respective counts are shown in Table B-3 and Table B-4 in Appendix B:.

The more traditional level 1 inputs, in addition to the study variables discussed in the prior section, include three categories of inputs as controls.

Subsector-Level and Detailed Industry Variables

The first category of variables draw on defense and overall U.S. economy NAICS economic sector data. NAICS codes are hierarchal, detailed industries have six-digit codes and the first three digits of that code refer to the parent subsector.

Subsector / Detailed Industry Herfindahl-Hirschman Index (HHI): Measures the level of consolidation within a given subsector or detailed industry (used at both levels). This is calculated by squaring the percentage share of obligations going to each contractor and summing them up. This approach gives the greatest weight to the shares of the most prominent vendors and a higher value represents greater consolidation.

Subsector / Detailed Industry DoD: U.S.: As the name suggests, this variable is a ratio of the size of the DoD sector, measured in obligations, versus the comparable U.S. sector, measured in dollars of revenue. A larger ratio indicates a more defense-oriented sector and a smaller ratio indicates a proportionately larger commercial, non-government, and civil sector. The ratio is logged, rescaled, and capped at 1.

Office and Vendor-Office Pair Variables

The first category focus on the contracting office as well as the pairing of the contract's vendor and office and are new to this paper.

Office Obligations: How much the office spent, in logged 2018 constant 2018 obligations over the prior seven fiscal years. One dollar is added this total, to allow offices with no spending history to be logged.

Office Focus: How concentrated the office's contract counts are in a small number of detailed industries. It uses the HH index transformation but tracking the number of contracts in each detailed industry

[7] The service areas were originally developed for CSIS services reports (Berteau et al, 2013) although in those reports the categories of facility based and construction were combined.

rather than the dollars to each vendor. This value is then logged and rescaled with a high value representing most contracts coming from a small number of codes.

Paired Share: What portion of the contracting spending of the office has this vendor won? It is calculated as the percentage of office obligations over the prior seven years that went to the vendor in question, rescaled. A high value in this variable may reflect vendor lock.

Contract-Level Variables

There are three continuous scope variables that measure different aspects of the scale of the contract: ***Initial Base*** in 2018 dollars based on contract start year, logged and rescaled; ***Initial Ceiling:Base*** is the ratio of the initial base to the initial ceiling, the difference between the two being the sum of options that could be exercised; and ***Planned Duration*** is the maximum length of contract as reported at contract start, logged and then rescaled.

Competition: Whether a contract was competed and if so how many offers were received. There is a baseline of no competition and three alternatives: competition, available for competition but receiving only 1 offer, 2-4 offers, 5+ offers.

Vehicle: The mechanism by which the service contract was awarded. There is a baseline of definitive contracts and purchase orders, but also four types of indefinite delivery vehicles: Single-Award IDCs (S-IDC); Multi-award IDCs (M-IDCs); Federal Supply Schedule or Government-Wide Acquisition Contract (FSS-GWAC); and Blank Purchase Agreement or Basic Ordering Agreement (BPA-BOA).

Pricing: The method of payment for the vendor. A firm-fixed price is used as a baseline with six alternatives handled by dummy variables: incentive fee contracts (whether fixed price or cost-based), combination; combination or other contracts which include multiple types, time and materials, labor hours, or fixed price: level of effort (T&M/LH/FP:LoE); other fixed price including all types of fixed price not covered by earlier categories; whether the contract began as an undefinitized contract award (UCA); and other cost-based covering all types of cost-based contracts not covered by earlier categories.

Crisis Funding: Whether the funding for a service was drawn from an emergency account. The baseline is drawing from non-emergency accounts with three alternatives for Overseas Contingency Operations (OCO), disaster response, and the American Recovery and Reinvestment Act (Recovery Act).

5 Results

5.1 Options

Table 6 shows the results of the logit models for (a) whether some options are exercised in a given contract, and (b) whether all options are exercised in a given contract. In Table 6 and subsequent model tables, the sign of the coefficient estimates whether the variable has a positive or negative association with the outcome measure. Likewise, because the variables are recentered, their relative influence can be judged by comparing the absolute value of different coefficients, with larger values estimating a stronger relationship. However, we would encourage the reader to interpret effect magnitude for a

given variable by looking at the logit coefficients after they are transformed into odds-ratio form, as is done for all study variables in Table 7, and for all variables in Table C-1 and

variable	OR	2.50%	97.50%
(Intercept)	1.05	0.68	1.63
Log(Det. Ind. Salary)	1.01	0.90	1.12
Log(Serv. Code Invoice Rate)	0.94	0.89	1.00
Office Perf.-Based %	0.85	0.77	0.93
Office Serv. Code Exp. %	0.83	0.75	0.93
Paired Years	1.30	1.24	1.37
Log(Paired Actions)	0.93	0.85	1.02
Log(Init. Base)	1.25	1.17	1.34
Log(Init. Ceiling:Base)	1.35	1.31	1.39
Log(Planned Dur.)	1.33	1.25	1.42
Comp=1 offer	1.01	0.95	1.08
Comp=2-4 offers	0.97	0.92	1.02
Comp=5+ offers	0.94	0.88	1.00
Vehicle=S-IDC	0.50	0.46	0.54
Vehicle=M-IDC	0.79	0.73	0.86
Vehicle=FSS/GWAC	0.84	0.78	0.91
Vehicle=BPA/BOA	0.75	0.64	0.87
Pricing=Other Fixed-Price	0.50	0.36	0.69
Pricing=Incentive Fee	0.71	0.39	1.27
Pricing=Comb. or Other	1.43	1.21	1.69
Pricing=Other Cost-Based	1.28	1.14	1.45
Pricing=T&M/LH/FP:LoE	1.06	0.93	1.20
Pricing=UCA	0.91	0.73	1.15
Crisis=Recovery Act	0.66	0.48	0.91
Crisis=Disaster	0.73	0.40	1.33
Crisis=OCO	0.71	0.56	0.89
Log(Det. Ind. HHI)	0.88	0.80	0.96
Log(Det. Ind. DoD:US)	0.99	0.91	1.07
Log(Subsector HHI)	0.98	0.88	1.09
Log(Subsector DoD:US)	1.03	0.88	1.21
Log(Paired Share %)	0.98	0.89	1.09
Log(Office Obl.)	1.22	1.11	1.35
Log(Office Focus)	0.72	0.65	0.81
Office Perf.-Based %:Log(Paired Share %)	2.00	1.69	2.36
Paired Share %:Log(Office Obligations)	1.14	0.91	1.43
Paired Years:Pricing=Other Fixed-Price	0.91	0.44	1.91

Paired Years:Pricing=Incentive Fee	4.72	1.75	12.77
Paired Years:Pricing=Comb./Other	0.64	0.45	0.90
Paired Years:Pricing=Other Cost-Based	1.23	1.01	1.48
Paired Years:Pricing=T&M/LH/FP:LoE	1.09	0.86	1.37
Paired Years:Pricing=UCA	1.17	0.76	1.80
Office Perf.-Based %:Log(Paired Actions)	0.54	0.49	0.61
Log(Init. Base):Log(Init. Ceiling:Base)	1.30	1.25	1.35
Office Serv. Code Exp. %:Log(Office Obl.)	2.79	2.05	3.79

Table C-2 in Appendix D: Odds Ratios. An odds ratio of one indicates there is no relationship between a given variable and the outcome in question, while a ratio of greater than one suggests a positive relationship, and a ratio of less than one suggests a negative relationship.

5.1.1 Services Complexity Variables

Hypothesis	Some Options Model	All Options Model
H1B	No support	Mixed/Limited Supprot
H2A	Mixed/limited support	Mixed/limited Support
H2B	Supported	Supported

Only one of the services complexity variables were statistically significant in either model, meaning **we find limited support for H1,** that is to say that attempt to use detailed industry salary as a proxy for service complexity found results but that the log of the service code invoice rate did estimate lower likelihood of exercising all options (significant at the 5 percent leve)**.**

- Serv. Code Invoice Rate
    - Individually estimates a lower likelihood of exercising all options.
    - When interacting with pricing options, three of the cost-based options, other-cost based, UCA, and combination/other, estimated significant positive association with the likelihood of all options being exercised (significant at the 1 percent, 0.1 percent, and 0.05 levels, respectively).

5.1.2 Office Capacity Variables

All Options

- Office Performance-Based Percentage
    - No individually statistically significant effect.
    - The interaction term between office use of performance-based contracting and paired contract actions estimates a positive, statistically significant relationship with the probability that all options are exercised for a given level of paired actions.
    - Thus, for any level of logged, centered, standardized contract actions for a given office-vendor pair greater than zero, there is a positive relationship between office use of performance-based contracting and the probability that all options are exercised.

- **Thus, H2A supported for all options model** *when logged, centered, standardized contract actions for a given office-vendor pair are greater than zero*.

- Office Service Code Experience Percentage
  - Statistically significant effect: a higher level of office service code experience estimates a greater probability that all options in a given contract are exercised. For every unit increase in our standardized measure of office performance-based percentage, the odds that all options are exercised in a contract rise by a factor of 1.20, as can be seen in the odds ratio listed for this variable in Table 7. **H2B supported for all options model.**

Table 6 Some and All Options

	Some Options Exercised Model	**All Options Exercised Model**
(Intercept)	0.05 (0.22)	**1.15** (0.15)***
Study Variables		
Log(Det. Ind. Salary)	0.01 (0.06)	0.08 (0.06)
Log(Serv. Code Invoice Rate)	-0.06 (0.03).	**-0.07** (0.03)*
Office Perf.-Based %	**-0.16** (0.05)***	0.00 (0.05)
Office Serv. Code Exp. %	**-0.18** (0.06)**	**0.18** (0.06)**
Paired Years	**0.26** (0.02)***	**0.14** (0.03)***
Log(Paired Actions)	-0.07 (0.04)	**-0.22** (0.05)***
Contract Characteristics		
Log(Init. Base)	**0.23** (0.04)***	**-0.14** (0.03)***
Log(Init. Ceiling:Base)	**0.30** (0.01)***	**-0.80** (0.02)***
Log(Planned Dur.)	**0.29** (0.03)***	**-0.45** (0.04)***
Comp=1 offer	0.01 (0.03)	0.00 (0.03)
Comp=2-4 offers	-0.03 (0.03)	-0.04 (0.03)
Comp=5+ offers	**-0.07** (0.03)*	**-0.08** (0.03)*
Vehicle=S-IDC	**-0.69** (0.04)***	**-0.20** (0.05)***
Vehicle=M-IDC	**-0.23** (0.04)***	**-0.13** (0.04)**
Vehicle=FSS/GWAC	**-0.17** (0.04)***	-0.06 (0.04)
Vehicle=BPA/BOA	**-0.29** (0.08)***	**0.25** (0.10)*
Pricing=Other Fixed-Price	**-0.70** (0.17)***	**-0.50** (0.22)*
Pricing=Incentive Fee	-0.34 (0.30)	0.58 (0.31).
Pricing=Comb. or Other	**0.36** (0.09)***	**-0.14** (0.07)*
Pricing=Other Cost-Based	**0.25** (0.06)***	**0.21** (0.06)***
Pricing=T&M/LH/FP:LoE	0.06 (0.07)	-0.03 (0.07)
Pricing=UCA	-0.09 (0.12)	-0.09 (0.15)
Crisis=Recovery Act	**-0.42** (0.16)*	-0.24 (0.22)
Crisis=Disaster	-0.32 (0.31)	**-1.46** (0.42)***
Crisis=OCO	**-0.34** (0.12)**	**-0.26** (0.12)*

NAICS/Office Characteristics		
Log(Subsector HHI)	-0.02 (0.05)	-0.05 (0.06)
Log(Subsector DoD:US)	0.03 (0.08)	-0.02 (0.08)
Log(Det. Ind. HHI)	**-0.13** (0.05)**	-0.09 (0.05).
Log(Det. Ind. DoD:US)	-0.01 (0.04)	0.04 (0.05)
Log(Office Obl.)	**0.20** (0.05)***	**-0.09** (0.03)**
Log(Office Focus)	**-0.32** (0.06)***	**-0.13** (0.06)*
Log(Paired Share %)	-0.02 (0.05)	-0.07 (0.06)
Interactions		
Log(Serv. Code Invoice Rate):Pricing=Other Fixed-Price		-0.23 (0.61)
Log(Serv. Code Invoice Rate):Pricing=T&M/LH/FP:LoE		0.17 (0.20)
Log(Serv. Code Invoice Rate):Pricing=Incentive Fee		-0.60 (0.70)
Log(Serv. Code Invoice Rate):Pricing=Other Cost-Based		**0.43** (0.14)**
Log(Serv. Code Invoice Rate):Pricing=UCA		**1.42** (0.37)***
Log(Serv. Code Invoice Rate):Pricing=Comb./Other		**0.40** (0.20)*
Office Serv. Code Exp. %:Log(Office Obl.)	**1.03** (0.16)***	
Office Serv. Code Exp. %:Log(Office Focus)		-0.10 (0.11)
Office Perf.-Based %:Log(Paired Actions)	**-0.61** (0.06)***	
Office Perf.-Based %:Log(Paired Share %)	**0.69** (0.08)***	0.15 (0.09)
Paired Years:Pricing=Other Fixed-Price	-0.09 (0.37)	
Paired Years:Pricing=T&M/LH/FP:LoE	0.08 (0.12)	
Paired Years:Pricing=Incentive Fee	**1.55** (0.51)**	
Paired Years:Pricing=Other Cost-Based	**0.20** (0.10)*	
Paired Years:Pricing=UCA	0.16 (0.22)	
Paired Years:Pricing=Comb./Other	**-0.45** (0.18)*	
Log(Init. Base):Log(Init. Ceiling:Base)	**0.26** (0.02)***	
Paired Share %:Log(Office Obligations)	0.13 (0.12)	
Log(Office Obligations):Log(Office Focus)		-0.00 (0.07)
AIC	75772.83	59424.62
BIC	76242.82	59858.10
Log Likelihood	-37835.41	-29663.31
Num. obs.	74274	51357
Var: ServArea:(NAICS6:NAICS3) (Intercept)	0.14	0.03
Var: Office:Agency (Intercept)	0.54	0.22
Var: NAICS6:NAICS3 (Intercept)	0.04	0.09
Var: Place (Intercept)	0.27	0.07

Var: NAICS3 (Intercept)	0.03	0.02
Var: Agency (Intercept)	0.00	0.06
Var: StartFY (Intercept)	0.29	0.03

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$. Numerical inputs are rescaled.

Some Options

- Office Performance-Based Percentage
    - Negative, statistically significant relationship with the probability that some options are exercised when considered individually.
    - Interaction term between log(paired actions) and office performance-based percentage estimates a statistically significant, negative relationship between office use of performance-based contracting and the probability that some options are exercised, for a given level of log(paired actions).
    - Interaction term between office performance-based percentage and the share of office obligations filled by the vendor in question (Paired Share %) estimates a statistically significant, positive relationship between office use of performance-based contracting and the probability that some options are exercised, for a given level of obligations shared between the contractor-vendor pair.
    - Given that all the variables in question are each centered and standardized, we can compare the magnitude of their effects relatively accurately through naïve comparison of regression coefficients, which indicates that **H2A is substantiated** *in our some options model for offices with a low level of contract actions between the office and a given vendor, and where said vendor nonetheless makes up a high percentage of all obligations of said office*, a situation which, while possible, seems to be in tension with itself. *For offices with high levels of log(paired actions) and low levels of paired share percentage*, **H2A is not upheld**. Exact results will vary through the combination of the values of the specific variables contained in our interaction terms.
- Office Service Code Experience Percentage
    - Negative, statistically significant relationship with the probability that some options are exercised when considered by itself.
    - Interaction term between log(office obligations) and office service code experience percentage estimates a highly significant, large magnitude positive relationship between the level of office service code experience and the probability that some options are exercised, while at a given level of office obligations.
    - For most values of log(office obligations) the interaction effect will outweigh the effect of office service code experience on its own, and thus that office service code experience has a positive relationship with the probability that some options are exercised, **substantiating H2B in our some options model**. For reference, the median value of our logged, standardized, and centered measure of office obligations is 0.128, which results, at this median value, of a net regression coefficient for office service code experience percentage in our model of 0.08, a net positive relationship with some options probability.

5.1.3 Past Relationship Variables

All Options

- Paired years
    - Paired years demonstrates a positive, statistically significant relationship with the probability that all options are exercised. For a one unit increase in our standardized measure of paired years, the odds that all options will be exercised in a given contract rise by a factor of 1.16. Thus, **H3A is supported by the all options model.**

Hypothesis	Some Options Model	All Options Model
H3A	Mostly Supported	Supported
H3B	Limited Support	Limited Support

- Paired actions
    - Paired actions estimates a negative, significant relationship with the probability that all options are exercised.
    - The interaction term between office use of performance-based contracting and paired actions, as discussed above, estimates a positive, significant relationship between paired actions and the probability that all options are exercised, for a given level of office use of performance-based contracting.
    - If we perform a basic comparison of the magnitude of the coefficients on the two effects discussed in the bullet points above, we see that for an office with a very high level of use of performance-based contracting, the interaction term will overwhelm the effect of paired actions on its own, and we see a positive relationship between paired actions and all options probability. For lower levels of office use of performance-based contracting, a negative relationship will hold. **H3B is thus supported only under limited circumstances.**

Some Options

- Paired years
    - Paired years demonstrates a statistically significant, positive relationship with the probability that some options are exercised when considered on its own.
    - Several categories of our categorical variable for contracting pricing system show statistically significant interactions with paired years. For each of these, we can compare the magnitude of effects, and determine when a positive relationship exists between paired years and the probability that some options are exercised.
        - When contract pricing is incentive-based, paired years has a very high magnitude, significant, positive relationship with some options probability – in this case, a one-unit increase in our standardized measure of paired years leads to one of the largest magnitude results reported in this section.
        - When contract pricing is in the other cost-based category, we still see a positive relationship between paired years, and some options probability. This result is also only significant at the 0.05 level, while the above result is significant at the 0.01 level.

Table 7: Options Study Variable Odds Ratios	Variable	Odds Ratio	95% Confidence Interval	
			Lower Bound	Upper Bound
Some Options Exercised Model	Log(Det. Ind. Salary)	1.08	0.97	1.21
	Log(Serv. Code Invoice Rate)	0.94	0.88	1.00
	Office Perf.-Based %	1.00	0.91	1.10
	Office Serv. Code Exp. %	1.19	1.06	1.34
	Paired Years	1.16	1.10	1.22
	Log(Paired Actions)	0.80	0.73	0.88
All Options Exercised Model	Log(Det. Ind. Salary)	1.08	0.97	1.21
	Log(Serv. Code Invoice Rate)	0.94	0.88	1.00
	Office Perf.-Based %	1.00	0.91	1.10
	Office Serv. Code Exp. %	1.19	1.06	1.34
	Paired Years	1.16	1.10	1.22
	Log(Paired Actions)	0.80	0.73	0.88

        - When contract pricing falls into our combined/other category, the relationship between paired years and some options probability flips to negative, but remains statistically significant at the 0.05 level.
    - Thus, **H3A is largely substantiated in the some options model**, *with the exception of when contract pricing falls into our combined/other category.*

- Paired actions
    - Log(paired actions) has no statistically significant relationship with some options probability when considered on its own.
    - However, as discussed above, the interaction term between office use of performance-based contracts and paired actions estimates a statistically significant, negative relationship between paired actions and the probability that some options are exercised, for a given level of office use of performance-based contracting.
    - cp_OffPSC7 is our indicator variable for office use of performance-based contracts. It has been standardized and re-centered, with a minimum value of -0.57, and a median of -0.07. Since our measure of office use of performance-based contracts used in the logit model can be negative, this means contracts through offices with a particularly low level of performance-based contracting will demonstrate a positive relationship between paired actions between said office and their vendor, and the probability that some contract options are exercised.

- This means that **we find support for H3B only in limited circumstances**, as this relationship will be negative for offices with higher levels of use of performance-based contracting.

5.1.4 Further Discussion

While the reader will be left to pick through our control variables largely on their own, our contract scope variables are worth commenting on briefly. In the some options model, initial base, initial ceiling to base ratio, and initial planned contract duration all have positive, highly significant relationships with the probability that some options will be exercised. In the all options model, we observe precisely the opposite relationship, where the levels of initial base, initial ceiling to base ratio, and initial planned contract duration are correlated with a lower probability that all options will be exercised.

The two above multilevel models includes varying intercepts for the government customer, section of the economy, contract start fiscal year, and country of performance. The variances are listed at the bottom of Table 6, where a great value indicates greater variation in the intercepts for different categories within the variable, i.e. some levels have higher magnitude positive intercepts and some have lower negative intercepts. In both of these models, but particularly in the some options model, there is a large degree of variance based on the same government purchasing entity, implying that a sizeable portion of our overall variance, especially in the some options model, is driven by differences between contracting offices/agencies, rather than overall trends in the data. This suggests that while our results hold, much of the variation in whether some or all options of a contract are exercised is dependent on contracting office/agency specific factors. In the some options model, contract start fiscal year and place of performance also account for a fairly large portion of the variance, but have a much smaller presence in the all options model. In fact, across the board, overall variance in the some options model is much more driven by variance between classes than overall variance is in the all options model. Service area multi-level controls pick up a lot of variance within the some options model, but are much less relevant in the all options model, following the same general trend described above.

5.2 Ceiling Breaches

Table 8: Ceiling Breaches

	Likelihood (Logit)	**Size Given Breach (Regression)**
(Intercept)	**-4.55** (0.18)***	**8.82** (0.13)***
Study Variables		
Log(Det. Ind. Salary)	-0.00 (0.06)	-0.02 (0.06)
Log(Serv. Code Invoice Rate)	**0.06** (0.01)***	**0.09** (0.02)***
Office Perf.-Based %	**0.07** (0.03)**	-0.06 (0.04)
Office Serv. Code Exp. %	**0.19** (0.02)***	**0.13** (0.03)***
Paired Years	**-0.08** (0.01)***	0.03 (0.02)

Log(Paired Actions)	**0.44** (0.02)***	-0.01 (0.03)
Contract Characteristics		
Log(Init. Base)	**1.35** (0.01)***	**2.65** (0.02)***
Log(Init. Ceiling:Base)	**0.25** (0.01)***	**0.25** (0.02)***
Log(Planned Dur.)	**0.28** (0.01)***	**0.05** (0.02)*
Comp=1 offer	-0.03 (0.02).	-0.05 (0.03).
Comp=2-4 offers	**0.11** (0.01)***	**-0.11** (0.02)***
Comp=5+ offers	**0.18** (0.01)***	**-0.07** (0.02)**
Vehicle=S-IDC	**-0.48** (0.02)***	-0.01 (0.02)
Vehicle=M-IDC	**-0.18** (0.02)***	-0.02 (0.03)
Vehicle=FSS/GWAC	-0.01 (0.03)	**0.12** (0.05)*
Vehicle=BPA/BOA	**-0.29** (0.03)***	**-0.13** (0.06)*
Pricing=Other Fixed-Price	**-0.40** (0.09)***	**-0.36** (0.14)*
Pricing=Incentive Fee	**2.39** (0.07)***	**0.81** (0.12)***
Pricing=Comb. or Other	**0.26** (0.05)***	**0.43** (0.07)***
Pricing=Other Cost-Based	**-0.09** (0.03)**	**0.81** (0.05)***
Pricing=T&M/LH/FP:LoE	**0.11** (0.04)*	**0.66** (0.07)***
Pricing=UCA	0.07 (0.04).	**0.37** (0.07)***
Crisis=Recovery Act	**0.12** (0.04)**	-0.06 (0.06)
Crisis=Disaster	0.07 (0.09)	**0.39** (0.13)**
Crisis=OCO	**-0.10** (0.05)*	0.08 (0.08)
NAICS/Office Characteristics		
Log(Subsector DoD:US)	**-0.46** (0.07)***	**0.29** (0.08)***
Log(Det. Ind. HHI)	-0.02 (0.02)	**0.06** (0.03)*
Log(Det. Ind. DoD:US)	**0.07** (0.03)*	0.01 (0.03)
Log(Office Obl.)	**0.04** (0.02)*	**0.06** (0.02)*
Log(Office Focus)	**-0.36** (0.04)***	
Log(Paired Share %)	**-0.24** (0.03)***	0.01 (0.04)
Interactions		
Office Perf.-Based %:Log(Paired Actions)	**0.38** (0.03)***	**-0.20** (0.04)***
Office Perf.-Based %:Log(Planned Dur.)	**-0.12** (0.02)***	
AIC	347679.35	242103.91
BIC	348163.79	242464.78
Log Likelihood	-173798.68	-121011.95
Num. obs.	1000000	61184
Var: ServArea:(NAICS6:NAICS3) (Intercept)	0.21	0.14
Var: Office:Agency (Intercept)	1.54	0.15
Var: NAICS6:NAICS3 (Intercept)	0.15	0.04
Var: Place (Intercept)	0.19	0.11

Var: NAICS3 (Intercept)	0.20	0.07
Var: Agency (Intercept)	0.16	0.06
Var: StartFY (Intercept)	0.02	0.01
Var: Residual		2.98

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$. Numerical inputs are rescaled.

shows the results for the logit model of ceiling breach likelihood in the left column and the ordinary least squares regression model for ceiling breach size among contracts experiencing a breach in the right column. While the breach size regression coefficients can be interpreted directly as estimates of the influence of variables on contract size, the breach likelihood logit coefficient are more straightforward to understand after being transformed into odds ratios as is done for the study variables in Table 9 below or for all variables in

Table C-3 in Appendix C: Model Odds Ratios.

Table 8: Ceiling Breaches

	Likelihood (Logit)	**Size Given Breach (Regression)**
(Intercept)	**-4.55** (0.18)***	**8.82** (0.13)***
Study Variables		
Log(Det. Ind. Salary)	-0.00 (0.06)	-0.02 (0.06)
Log(Serv. Code Invoice Rate)	**0.06** (0.01)***	**0.09** (0.02)***
Office Perf.-Based %	**0.07** (0.03)**	-0.06 (0.04)
Office Serv. Code Exp. %	**0.19** (0.02)***	**0.13** (0.03)***
Paired Years	**-0.08** (0.01)***	0.03 (0.02)
Log(Paired Actions)	**0.44** (0.02)***	-0.01 (0.03)
Contract Characteristics		
Log(Init. Base)	**1.35** (0.01)***	**2.65** (0.02)***
Log(Init. Ceiling:Base)	**0.25** (0.01)***	**0.25** (0.02)***
Log(Planned Dur.)	**0.28** (0.01)***	**0.05** (0.02)*
Comp=1 offer	-0.03 (0.02).	-0.05 (0.03).
Comp=2-4 offers	**0.11** (0.01)***	**-0.11** (0.02)***
Comp=5+ offers	**0.18** (0.01)***	**-0.07** (0.02)**
Vehicle=S-IDC	**-0.48** (0.02)***	-0.01 (0.02)
Vehicle=M-IDC	**-0.18** (0.02)***	-0.02 (0.03)
Vehicle=FSS/GWAC	-0.01 (0.03)	**0.12** (0.05)*
Vehicle=BPA/BOA	**-0.29** (0.03)***	**-0.13** (0.06)*
Pricing=Other Fixed-Price	**-0.40** (0.09)***	**-0.36** (0.14)*
Pricing=Incentive Fee	**2.39** (0.07)***	**0.81** (0.12)***
Pricing=Comb. or Other	**0.26** (0.05)***	**0.43** (0.07)***
Pricing=Other Cost-Based	**-0.09** (0.03)**	**0.81** (0.05)***
Pricing=T&M/LH/FP:LoE	**0.11** (0.04)*	**0.66** (0.07)***
Pricing=UCA	0.07 (0.04).	**0.37** (0.07)***
Crisis=Recovery Act	**0.12** (0.04)**	-0.06 (0.06)
Crisis=Disaster	0.07 (0.09)	**0.39** (0.13)**
Crisis=OCO	**-0.10** (0.05)*	0.08 (0.08)
NAICS/Office Characteristics		
Log(Subsector DoD:US)	**-0.46** (0.07)***	**0.29** (0.08)***
Log(Det. Ind. HHI)	-0.02 (0.02)	**0.06** (0.03)*
Log(Det. Ind. DoD:US)	**0.07** (0.03)*	0.01 (0.03)
Log(Office Obl.)	**0.04** (0.02)*	**0.06** (0.02)*
Log(Office Focus)	**-0.36** (0.04)***	

Log(Paired Share %)	**-0.24** (0.03)***	0.01 (0.04)
Interactions		
Office Perf.-Based %:Log(Paired Actions)	**0.38** (0.03)***	**-0.20** (0.04)***
Office Perf.-Based %:Log(Planned Dur.)	**-0.12** (0.02)***	
AIC	347679.35	242103.91
BIC	348163.79	242464.78
Log Likelihood	-173798.68	-121011.95
Num. obs.	1000000	61184
Var: ServArea:(NAICS6:NAICS3) (Intercept)	0.21	0.14
Var: Office:Agency (Intercept)	1.54	0.15
Var: NAICS6:NAICS3 (Intercept)	0.15	0.04
Var: Place (Intercept)	0.19	0.11
Var: NAICS3 (Intercept)	0.20	0.07
Var: Agency (Intercept)	0.16	0.06
Var: StartFY (Intercept)	0.02	0.01
Var: Residual		2.98

***p < 0.001, **p < 0.01, *p < 0.05, ·p < 0.1. Numerical inputs are rescaled.

For service contract complexity, only the log(service invoice rate) for the size model was significant (0.001 level) with a one unit increase in the recentered service invoice rate estimating an 9 percent increase in breach size in line with the predictions of H_1.

For H_2, office capacity, the results were more complicated. The percentage of office obligations that were performance-based services was not significant in its own right, but it did have significant interactions. For the likelihood model, the interaction of office performance-based percentage and the log(planned duration) estimated a lower risk of breach. The interaction of performance based office percentage and another study variable, the log(paired actions) was also significant and are covered below.

None of the office capacity, H_2, variables supported the hypotheses as both were significant with the opposite of the expected sign. Office service experience estimating that a one unit increase in the recentered variable would be associated with a 1.21 times higher likelihood of ceiling breach (significant at the 0.01 level) and with a one unit increase in the recentered variable estimating a 13 percent increase in breach size (significant at the 0.001 level).

Finally, the hypothesis 3, past relationship, variables were only significant for breach likelihood and not breach size. A one unit increase in recentered paired years estimated a 0.92 times lower likelihood of ceiling breaches (significant at the 0.05 level). Contrary to expectation, a one unit shift in recentered log(paired actions) estimated a 1.56 times greater likelihood of ceiling breaches (significant at the 0.001 level). The interaction between paired actions and the percent of performance-based contracting for the office was significant for both models but with opposite signs. The interaction term between our variable referencing the extent of the relationship between the office and the vendor and our variable

accounting for office use of performance-based contracting estimated a higher likelihood of breaches, but the size model found that a one unit increase of these two recentered variables together estimated a 20 percent decrease in breach size (significance of 0.1 percent in both cases).

Table 9 Ceiling Breach Study Variables Odds Ratio		95 Percent Conf. Interval	
Variable	Odds Ratio	Lower Bound	Upper Bound
Log(Det. Ind. Salary)	1.00	0.88	1.12
Log(Service Invoice Rate)	1.06	1.03	1.09
Office Perf.-Based %	1.08	1.02	1.13
Office Service Exp. %	1.21	1.15	1.27
Paired Years	0.92	0.90	0.95
Log(Paired Actions)	1.56	1.50	1.62

Moving to the contract level control variables, for multi-offer competition, breach likelihood and breach size estimated in opposite direction. For likelihood, competition with 2-4 offers and with 5+ offers both estimated a respective greater likelihood of ceiling breaches (a ratio of 1.11 and 1.19 at the 5 percent and 0.1 percent significance levels respectively). For breach size, the results were the opposite. Competition with 2-4 offers estimated 11 percent smaller breaches and competition with 5+ offers estimated 7 percent smaller breaches (significant at the 0.001 and 0.001 levels, respectively).

The results were more consistent for the scope variables log(initial base), log(ratio base to ceiling) and log(days). The initial base size of the contract is a powerful predictor of breach likelihood and a one unit increase in recentered base size estimates 3.83 times increase in likelihood and a 265 percent increase in breach size (significance 0.1 percent in both cases). The ratio of base to ceiling has a lower magnitude, but also positively estimates the frequency of contract breaches and size (again, significance at the 0.001 level in both cases). Duration is also associated with more risks, but much more strongly in the frequency model, where it has a ratio of 1.32, slightly larger than the 1.28 ratio for base to ceiling (0.1 percent significance in both cases). In the size model, duration likewise has a positive estimate, but has one of the lowest coefficients of significant variables and is only significant at the 0.05 level.

Contract vehicles proved more influential estimators for ceiling breach frequency, with single-award IDCs, multiple award IDCs, and BPA/BOA all estimating a lower chance of a breach occurring (each significant at the 0.001 level). For breach size, BPA/BOAs estimated smaller breaches while FSS/GWAC vehicles estimated larger breaches, though in both cases this effect was only significant at the 0.05 level. Pricing mechanism goes the other way and is a weaker estimator for likelihood than for size. Incentive fee as well as combination and other contract estimate a greater chance of breach, with ceiling breaches estimated as having a ratio of 10.9, the largest coefficient in that model (significant at the 0.001 level). For size all pricing types other than the base of firm fixed price and other fixed price estimate larger breaches, with incentive fee and other cost-based contracts both estimating 81 percent larger breaches respectively (the coefficients for all pricing mechanisms were significant at the 0.001 level). Finally, for the last of the contract level variables, ARRA contracts had 1.13 times greater likelihood of breaches

while OCO contracts had a smaller likelihood (significant at the 0.05 level). Meanwhile, disaster relief contracts estimated 39 percent larger breaches (significant at the 0.01 level).

For the controls based on NAICS sector statistics, the results for the frequency model and the size model did not align. Log(Detailed Industry HHI) was not significant for likelihood but did estimate larger breaches (significance level of 5 percent). The log(Subsector Ratio) having a proportionally larger defense sector estimated a lower likelihood of breaches but those breaches that did occur were estimated to be larger (significance levels of 1 percent and 0.1 percent respectively. In a difference with the other models, Subsector HHI was not included in this model because variance inflation factor checks indicated that inclusion of the variable raised risks of multicollinearity unless one of the NAICS measures were removed.

The last category of level one control variables are the office characteristics. In the likelihood model, a higher percent of an office's market held by a single vendor was associated with a 0.79 times lower chance of a ceiling breaching occurring (significant at 0.001 level). For both the likelihood and size models, a larger office volume of office obligations estimated greater risks. The Log(Office Focus) estimated that more focused offices were less likely to experience a breach (significant at the 0.001 level). However, the breach size model was unable to include Log(Office Focus) because of a high correlation with the Office Service Experience study variable.

Turning to multilevel variables which have different intercepts for acquirers (levels two and three), sector (levels two through four), place of performance (level two) and start fiscal year(two), contracting offices (level 2) and their parent agencies (level 3) explain a fair amount of variance. Contracting offices in particular explain the most variance in each model, and in the likelihood model, explain more variance than the rest of the level 2-4 variables put together. The hierarchical chain of service area (level 2), NAICS detailed industry (level 3), and NAICS subsector (level 4) explain about a third of the office and agency variance in the likelihood model but explain a greater proportion of the size model. For both models, place of performance (level 2) was overshadowed by the collective office and service type variables, but still accounted for considerably greater variance than the start year of the contract (level 2).

5.3 Terminations

Table 10 shows our multilevel model for contract terminations, and the corresponding estimates for regression coefficients. As described in section 5.1, this is a logit model, and so for the convenience of the reader in interpreting effect magnitudes, the odds ratios for study variables are listed in Table 11. The odds ratios for all variables are listed in Table XY in Appendix D: Odds Ratios.

Services Complexity

No service contract complexity variable is statistically significant on its own. That said, the interaction between average salary and our contract pricing system categorical variable is significant for several pricing categories. When contract pricing system falls into the time and materials/labor hours/fixed price: level of effort category, our model estimates a negative, statistically significant, relationship between average salary and the probability of the contract being terminated. This effect

has a relatively strong magnitude, as a one unit increase in our measure of average salary when contract pricing falls into this category is associated with a 45% fall in the odds of termination.

However, when contract pricing falls into our other cost-based category, which encompasses almost all cost-plus contracts, there is a positive relationship estimated between average salary and the probability of termination occurring, though this effect is only statistically significant at the 0.05 level. Thus, we find evidence to support H1A only when contract pricing falls into the other cost-based category, and in no other circumstances. We find no evidence to support H1B.

Office Capacity

Only office service code experience displays any sort of statistically significant relationship with termination probability. Here, a higher level of office experience with a given service code is associated with a higher probability of contract termination occurring, exactly the opposite of what was expected in H2B. For context, a one unit increase in our standardized measure of office service code experience leads to a 36% increase in the odds of termination occurring, a result which is significant at the 0.1% level. We find no evidence to support H2A.

Past Relationship

Again, only one variable displays any sort of statistically significant relationship with termination probability, but this category of study variables remains closer in line with our expectations. Here, paired years demonstrates a significant, negative relationship with the probability of termination – specifically, a one unit increase in our standardized measure of paired years leads to a 34% fall in the odds of termination occurring, which provides evidence in favor of H3A. We find no evidence to support H3B.

Table 10 Terminations Model

	Termination
(Intercept)	**-4.45** (0.15)***
Study Variables	
Log(Det. Ind. Salary)	-0.05 (0.06)
Log(Serv. Code Invoice Rate)	-0.02 (0.02)
Office Perf.-Based %	0.03 (0.04)
Office Serv. Code Exp. %	**0.31** (0.04)***
Paired Years	**-0.28** (0.02)***
Log(Paired Actions)	-0.04 (0.03)
Contract Characteristics	
Log(Init. Base)	**0.29** (0.02)***
Log(Init. Ceiling:Base)	**0.50** (0.01)***
Log(Planned Dur.)	**0.90** (0.02)***
Comp=1 offer	**0.30** (0.03)***
Comp=2-4 offers	**0.37** (0.03)***
Comp=5+ offers	**0.71** (0.03)***

Vehicle=S-IDC	**-0.67** (0.03)***
Vehicle=M-IDC	**-0.41** (0.03)***
Vehicle=FSS/GWAC	**-0.18** (0.04)***
Vehicle=BPA/BOA	**-0.97** (0.06)***
Pricing=Other Fixed-Price	**-0.95** (0.10)***
Pricing=Incentive Fee	**-1.02** (0.39)**
Pricing=Comb. or Other	-0.16 (0.10)
Pricing=Other Cost-Based	**-0.37** (0.08)***
Pricing=T&M/LH/FP:LoE	**-0.45** (0.09)***
Pricing=UCA	**-0.72** (0.14)***
Crisis=Recovery Act	**-0.32** (0.13)*
Crisis=Disaster	**0.48** (0.19)*
Crisis=OCO	-0.07 (0.08)
NAICS Characteristics	
Log(Subsector HHI)	0.03 (0.04)
Log(Subsector DoD:US)	**0.24** (0.09)**
Log(Det. Ind. HHI)	0.04 (0.04)
Log(Det. Ind. DoD:US)	0.02 (0.04)
Office Characteristics	
Log(Office Obl.)	0.03 (0.03)
Log(Office Focus)	**-0.24** (0.05)***
Log(Paired Share %)	**-0.13** (0.04)***
Interactions	
Paired Years:Pricing=Other Fixed-Price	0.09 (0.18)
Paired Years:Pricing=T&M/LH/FP:LoE	**-0.60** (0.18)***
Paired Years:Pricing=Incentive Fee	-0.09 (0.65)
Paired Years:Pricing=Other Cost-Based	**0.30** (0.13)*
Paired Years:Pricing=UCA	-0.51 (0.26).
Paired Years:Pricing=Comb./Other	0.32 (0.23)
AIC	142606.35
BIC	143149.86
Log Likelihood	-71257.17
Num. obs.	1000000
Var: CrisisProductOrServiceArea:(NAICS:NAICS3) (Intercept)	0.18
Var: Office:Agency (Intercept)	0.66
Var: NAICS:NAICS3 (Intercept)	0.11
Var: PlaceCountryISO3 (Intercept)	0.39
Var: NAICS3 (Intercept)	0.04
Var: Agency (Intercept)	0.06

Var: StartFY (Intercept)	0.01

***p < 0.001, **p < 0.01, *p < 0.05, ·p < 0.1. Numerical inputs are rescaled.

Further Discussion

There are several interesting relationships exhibited amongst our control variables that are worth commenting on. Our contract-scope controls, initial base, initial ceiling to base ratio, and initial planned contract duration, all estimate positive, statistically significant relationships with the probability of contract termination – i.e. contracts with larger initial purchases, longer planned durations, or elevated amounts of options relative to their base size are all more likely to be terminated. As expected, contracts priced using incentives are significantly less likely to experience terminations – a contract with incentive pricing has 64% lower odds of termination than a contract using firm-fixed pricing. More surprisingly though, UCA pricing systems, as well as those classified as T&M/LH/FP:LoE, and other cost-based pricing systems, are also less likely to experience terminations (though the effect is of lower magnitude). Crisis funding appropriated through the ARRA is associated with a lower probability of termination, while crisis funding for disasters is correlated with a higher probability of termination, though these results are only significant at the 5% level.

Table 11 Ceiling Breach Study Variables Odds Ratio		95 Percent Conf. Interval	
Variable	Odds Ratio	Lower Bound	Upper Bound
Log(Det. Ind. Salary)	0.95	0.85	1.07
Log(Service Invoice Rate)	0.98	0.93	1.03
Office Perf.-Based %	1.03	0.95	1.12
Office Service Exp. %	1.36	1.25	1.48
Paired Years	0.76	0.73	0.79
Log(Paired Actions)	0.96	0.91	1.02

The multilevel model includes varying intercepts for place of performance, government customer, sector, and contract start fiscal year. The variance for each multilevel category are included at the bottom of Table 10. A larger variance corresponds with a greater range in the intercepts for that multilevel variable, which is to say that the differences between categories within the variable can matter more. The variance for acquirer (office and agency) is quite elevated, suggesting a large portion of the variety in termination outcomes comes from variation between offices. While not as high, the cumulative variance for sector-based variables is also elevated, as is the variance for place of performance, again implying that a substantial portion of overall variation in our data is driven by differences between categories, and not within them.

6 Discussion

6.1 Service Contract Complexity

The results supported the idea that, when looking at exercising all options and ceiling breaches, service code invoice rate might be a proxy for service contract complexity $H_{1B.}$ This finding was significant at the

0.001 level and held for both ceiling breach likelihood and the average size of the resulting breaches. However, no such support was found for H_{1A} which examined the average U.S. salary paid in the detailed industry in question. For options exercised and terminations, neither variable was significant at the 0.05 level.[8]

When considering this difference, it's helpful to review the underlying characteristics that service contract complexity is intended to measure. Joaquin and Greitens put forth that lower complexity services may prove easier to set and measure performance. This could include lower cost categories such as "Operation/Dining Facilities" and "Custodial – Janitorial Services." By comparison, medium and higher complexity categories can be technically demanding to monitor, for example "architecture-engineering services" or may represent a greater extent of partnership. The difficulty of evaluating these more complex contracts may explain why the ceiling breach measures proved to be the only performance measure with a significant association with invoice rate. Measurement challenges might make it harder to determine whether an option should be exercised or whether a termination was appropriate, but even in those cases the government can track changes to cost ceilings.

However, the rationales discussed by Joaquin and Greitens could apply to both detailed industry salary and service code invoice rate. Some possible explanations are less theoretically interesting, for example service codes, as measure specific to government contracting, may have smaller bands of service contract complexity than detailed industry. Another complicated factor is that detailed industries in many cases include both provision of not just services but also products or R&D. That said, study team examined the possibility of greater reliance on service codes rather than NAICS codes, but did not find that service codes performed better on their own.

A more interesting possibility might be that the non-labor parts of the invoice rate, from overhead to capital expenses, may capture something important about service complexity. A high wage contract may be a form of labor substitution, provisioning a highly experienced or educated worker who nonetheless works under the direct supervision of the government customer. The government may simply be getting what it pays for in terms of salary and not taking on higher risk despite bringing on contractors from an expensive detailed industry.

When considering why only ceiling breaches had a significant relationship, Brown and Potoski's research may point to a possible explanation, as they emphasize the specific risk of contracting for management tasks specifically, rather than emphasizing risk for all highly paid contracts, as they fear it can lead to a monopoly relationship between government customer and vendor. In such a relationship, the government may feel obligated to exercise options and avoid terminations while nonetheless choosing to employ costly change orders when a contract as originally specified does not meet present needs. However, the controls that more directly measured monopoly risks, the HHI indices, the paired share of obligations, and the presence of competition, often estimated a lower likelihood of breaches even if they also at times significantly estimated that these breaches would be higher. This finding suggests that further deeper study specifically on the small subset of contracts included within the services contract

[8] At the 0.10 significance level, the service code invoice rate does predict a lower likelihood of any options being exercised, but this result does not meet the significance standard of this report.

inventory may be worthwhile. Service code invoice rate can differentiate between higher charging and lower charging service categories, but not between comparatively expensive and low-cost vendors within the same category. That analysis would face data quality challenges. Unfortunately, this report's investigation did reveal oddities in the later years, with multiple categories suddenly declining in 2016 and with some categories collapsing their rate in 2017. Sometimes the end of a single large contract can do a great deal to explain fluctuations, as the study team found with Waste Treatment & Storage Facilities, but that explanation did not hold in other cases. Because this variable is defined to use lagged data, the 2016 and 2017 invoice rates are not included in any of the statistics, that said, these challenges could be an obstacle to future research. In addition, any contract-level examination of rate does run the risk of a reverse causation challenge as a contract that is performing poorly may have fluctuations in staffing levels as a result of problems and not only as source of them. Finally, future researchers may benefit from examining whether product or service codes can better identify some potential risk categories, such as managerial contracts, than NAICS detailed industries.

Finally, one of the contract-level controls included in the study may also act as a proxy for service contract complexity. Specifically, the ratio of initial contract ceiling to base significantly estimates changes in performance measures for all the dependent variables. As was discussed in section 4.2.1, the majority of contracts have no options to exercise but those with options represent a disproportionate share of obligations. This measure may relate to service contract complexity because the greater flexibility the unexercised options give the contracting office may also reflect greater uncertainty about how the contract will be executed. A greater ratio of initial ceiling to base, that is to say a contract with more room to expand, did positively estimate a greater likelihood of ceiling breaches and breach sizes, although base size was the more influential factor. For options, in line with expectations, a greater ratio estimated a higher chance that some options would be executed but a lower chance that all would be. However, the most intriguing result was that despite the flexibility offered by a large ratio, the risk of termination went up substantially, and is estimated to have a higher magnitude than the influence of the size of the base itself, though a lower magnitude than the planned maximum duration of the contract. While factors other than service contract complexity do drive up the ratio, for example limitations on multiyear funding for Operations and Maintenance contracts, but nonetheless practitioners and future researchers should consider whether mechanisms the literature found appropriate to mitigate the risks of complex services, for example a closer partnership, may be more important in situations where there is a large proportion of options to be exercised. On the other hand, it may also be worth examining if more options might increase the complexity of some contracts that might otherwise be simpler.

6.2 Contract-Management Capacity

The results largely did not support, and were often significant and contrary to, the expectations of H_2. The sole supportive finding is that for contracts with some options exercised, a greater degree of office service code experience estimated a higher likelihood that all options would be exercised (significant at the 0.001 level). However, for all other dependent variables, more office service code experience significantly estimated worse performance (i.e. lower likelihood of some options exercised, more likely and larger ceiling breaches, and more likely terminations, all significant at the 0.001 level).

In the case of the proportion of office performance-based contracting for three of the dependent variables there were no significant results and for likelihood of some options and ceiling breaches, the results were significant and opposed to the hypothesis: a greater proportion of office performance-based contracting estimated a lower likelihood that some options would be exercised (significant at the 0.001 level) and a higher likelihood of ceiling breaches (significant at the 0.05 level). However, the interactions do show mixed result of that uphold some of the findings.

The interactions explored all focused on the proportion of office performance-based contracting, and thus overlap with the literature on government-vendor relationship as well as the relationship history results discussed in the next section. Yang, Hsieh, and Li center their definition of implementation capacity around the customer's ability to create public-private partnerships. Performance-based services contracting is one means of facilitating such partnerships and thus the interactions tested were combinations under which partnership may prove more likely to be successful. For the options models, the interaction between percentages of office performance-based contracting and the paired share of office obligations sought to examine whether performance-based approaches may mitigate the potential risks of reliance on a particular vendor. While not significant for the all options model, the interaction did estimate a higher likelihood that some options would be exercised (0.001 significance level). Another tested interaction for the ceiling breach likelihood model was the relationship between office performance-based contracting percentage and the planned duration of the contract. Past CSIS has work has found that performance-based logistics contracts may be more effective if they include the potential for a longer duration. Indeed, the interaction did estimate a lower likelihood of ceiling breach interactions (significant at the 0.001 level). The remaining interactions with the number of paired contract actions study variable had a significant positive or negative influence on performance, depending on the model, and are discussed in the next section.

While the performance-based contracting interaction terms have some results in line with theory, overall it does not appear that the two proxy variables used in H_2 were successful in capturing the larger phenomenon of office capacity. For office performance-based contracting percentage, the problem may be that the reputation of performance-based contracting is so strong that offices that only weakly adhere to its tenants may overclaim their extent of performance-based contracting. For the related category of performance-based logistics, past CIS work has found the definitions used can vary widely and often can be over expansive.

The explanation may simply be that even a rough proxy for office capacity may require personnel, or at least office specific budget information, that is not available through FPDS or other open source reporting. The models do indicate that there are office characteristics not captured in the level 1 controls. In each of the five models the varying intercepts for each contracting office explains more variance than the remainder of the higher-level terms included in the model.

For the terminations model specifically, more promising than either of the two office capacity study variables was the contracting office focus. This measure of the distribution of contract counts between different NAICS sectors found that an office more focused on a small number of sectors estimated a lower risk of termination (significant at 0.001 level). However, for the options models, a focused office was less likely to exercise some options and less likely to go on to exercise all options once some options had been exercised (0.001 and 0.05 significance level respectively).

For future researchers, the best case would be greater access to statistics on contracting offices. Absent that, possible areas of interest may be to focus specifically on contracts for managerial capacity which were discussed both as being a risk area in service contract complexity but also raised by Brown and Potoski as a way to increase office capacity. Another approach would build on the fact that many offices rely on a small number of indefinite delivery contracts and examining the relationship of offices to vehicles as well as considering the inclusion of specific contract vehicles as a level of analysis, may capture some of the unexplained office variance in the model. Finally, this dataset may be usefully combined with more survey-based approaches, such as those performed by Apte and Rendon (2013) to calibrate whether this dataset can replicate the importance of office capacity in those cases where more detailed survey measures are available.

6.3 Customer-Vendor Relationships

The history between a vendor and contracting office proved the most reliable estimator of the study variables, supporting H_{3A}. The count of years out of the past seven that the buyer had the vendor under contract estimated a greater likelihood of some and all options being exercised, a lower likelihood of ceiling breach, and a lower likelihood of termination (significance level 0.001 in all cases) For average contract breach size, no significant relationship was found. The paired history does not directly measure trust but does establish the existence of a relationship. Yang, Hsie, and Li emphasize the importance of authentic partnerships rather than relying on overusing contract mechanisms for enforcement. In the best case, this sustained partnership may be an opportunity to develop trust and, as Fernandez covers, a relationship that engages in joint problem solving.

The study team had limited the number of years considered for both practical and theoretical reasons. The practical reason is that FPDS data is most readily available starting in fiscal year 2000 and limiting this and other longer-term variables to seven years allows for an equivalent period regardless of contract start year. The theoretical justification is Fernandez finding that the influence of trust does not increase over longer relationships, with reference to longer duration contracts. Seven years, particularly given that the rotation of contracting office personnel typically happens on a faster cycle, gives enough time for a trusting relationship to be established.

However, the frequency of buyer-vendor interactions, as measured by the logged number of contract actions between the buyer and vendor, was not nearly as effective at estimating positive performance. For the majority of the dependent variables, there was no significant relationship between logged paired actions and performance outcomes. Worse, for all options, more logged paired actions estimated a lower likelihood that all options would be exercised and for a greater likelihood a ceiling breach would occur. The simplest explanation for this may be that the frequency of contract actions may tell us more about the structure of a contract. For example, whether equipment is separately rented each time it is used or whether it is made consistently available over the period of a year and only leads to a contract action when it is maintained by the vendor. While contract actions are opportunities for monitoring or simply working more closely, they may in practice describe a relationship that is more literally transactional. Yang, Hsieh, and Li warn that highly specified contracts decrease the contractor's drive for innovation and a greater number of contract actions may indicate such greater specificity.

For the options and ceiling breach models, the office-performance based percentage was interacted with the log of the number of paired contract actions. This combination was meant to test whether the frequency of connections between the two partners would be mutually reinforcing with performance-based contracting oriented offices. While the interaction term did find that ceiling breachesThe results were mixed. Contrary to expectations, the interaction term estimating a lower likelihood that some options would be exercised and a higher likelihood of ceiling breaches (significant at the 0.001 level). Thus greater number of paired contract actions for offices that use more performance based contracting are associated with smaller ceiling breaches.

The contrast between log paired actions and logged paired history suggests that sometimes it is the existence of a relationship, and not its depth, that may be the more important estimator of the potential for trust. That said, the finding raises the question “what role does past obligations play” and whether vendor lock might explain the government’s eagerness to exercise some or all options and hesitation to terminate contracts.

7 Conclusions

Between service contract complexity, contract-management capacity, and buyer-vendor relationship, this report found a most straightforward relationship between the buyer-vendor relationship and contract performance. Pair history estimated better performance across both of the study team's established contract performance measures: likelihood of ceiling breach and terminations; and also two thirds of the new measures: likelihood of exercising some or all options, but not average contract size. While there were findings of interest for the other variables, the consistency of paired history has important implications beyond the metrics contained in this report. As past CSIS work by Cohen (2018) has found, the number of new entrants into DoD contract has been declining, which echoes a larger national trend of decline in new business starts (Casselman, 2017). DoD seeks a competitive marketplace in which new and non-traditional entrants can bring ideas from the larger commercial and research spheres to support the department's mission, but contracting officers may also be motivated in seeking the assurance of a known partner, especially when facing a consolidated marketplace. The same dynamics may contribute to a less welcoming environment for start-ups in an increasingly services-oriented national economy.

The literature's complementary emphasis on the importance of cooperative partnership and the benefit of competition suggest that while recruiting first time vendors is a worthy objective, practitioners and researchers should also pay attention to developing means to build trust and cooperative problem-solving relationships with new vendors and in taking advantage of the strengths, while minimizing the risks, of longstanding relationships. Possible avenues for future research and effort include vendor mentorship programs, programs that work with non-traditional vendors even after they have won their first contract such as Small Business Innovative Research, and the role of geography and colocation as a means of building relationships.

The complex findings on office capacity are also an indicator of a bigger picture data gap. Information on civilian and military acquisition personnel is collected but not available to open source researchers and in many cases is closely held even for those within the government. This data gap makes it harder to provide evidence of where investing in the workforce have paid off, limiting the ability to respond to persistent calls from Congress for headquarters reductions which may lead to suboptimal management of the contracting workforce. Perhaps widespread reporting on training, average tenure, and budgets, could respect the practical considerations of contracting offices in protecting their data while still allowing large dataset analysis. All of these would of course still be proxies, best able to capture whether a particular office is under resourced or under trained, rather than whether it was deploying its resources well. However, the ability to do such analysis would be a definite step forward.

The high variance in estimated performance levels between different contracting offices suggests a substantial marginal return to identifying quality office capacity measures. In the absence of new data, those doing the hard work of examining contracting office staffing and sharing such information with other practitioners through venues like the Acquisition Research Symposium are to be commended for addressing a clear need.

Where office capacity data does exist, such as the performance-based services contracting measures, it may be undermined by the temptation to broadly define what qualifies as a contracting technique, i.e.

performance-based contracting, with support from the literature and formal policy. The stronger results in interactions suggest that relying on multiple pieces of evidence to confirm performance-based approaches may be wise when seeking to measure the extent of implementation and whether these approaches are working.

Finally, the study found that service code invoice rates may have some value estimating the implications of service contract complexity on the likelihood that all options would be exercised as well as the likelihood of ceiling breach and size. The absence of any support from the some options or terminations models suggests that contract cost does reflect service contract complexity but that there are other aspects of complexity that cannot be detected by price alone. Other variables such as the ratio of initial cost ceiling to base may help fill this void, but some low hanging fruit in reporting, such as consistently reporting the solicitation date to FPDS or including easy to collect specificity measures like length of contract can give a clearer picture of the services contracting challenge the federal government is taking on.

References

Apte, U. and Rendon, R. (2013). Services Supply Chain in the Department of Defense: Defining and Measuring Success of Services Contracts in the U.S. Navy. Monterey: Naval Postgraduate School. Retrieved from http://www.acquisitionresearch.net/publications/detail/1164/

Berteau, D. et al., (2013) Structure and Dynamics of the U.S. Federal Services Industrial Base, 2000–2012, Washington D.C: Center for Strategic and International Studies,. 2013. https://csis-prod.s3.amazonaws.com/s3fs-public/legacy_files/files/publication/130827_Sanders_FederalServicesContractingTrends2000-2012_Web.pdf

Casselman, Ben (2017, Sept 20) A Start-Up Slump Is a Drag on the Economy. Big Business May Be to Blame. New York Times. Retrieved from https://www.nytimes.com/2017/09/20/business/economy/startup-business.html

Cohen, S. (November 2018) New Entrants and Small Business Graduation in the Market for Federal Contracts, Washington D.C: Center for Strategic and International Studies. Retrieved from https://www.csis.org/analysis/new-entrants-and-small-business-graduation-market-federal-contracts

Brown, T. L., & Potoski, M. (2003, March 04). Managing contract performance: A transaction costs approach. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/pam.10117

Brown, T., & Potoski, M. (2006, March 02). Contracting for management: Assessing management capacity under alternative service delivery arrangements. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/pam.20175

Fernandez, Sergio. 2007. “What Works Best When Contracting for Services? An Analysis of Contracting Performance at the Local Level.” Public Administration, 85: 1119-1140.

Fernandez, Sergio. 2009. “Understanding Contracting Performance: An Empirical Analysis.” Administration and Society, 41: 67–100.

GSA Federal Procurement Data System-Next Generation (FPDS-NG) Data Element Dictionary. (2017, November 10

Sanders, G. (2015). Avoiding Terminations, Single-Offer Competition, and Costly Changes with Fixed-Price Contracts. Retrieved from https://www.csis.org/analysis/avoiding-terminations-single-offer-competition-and-costly-changes-fixed-price-contracts

Sanders, G., Huitink, Z. (2019). Evaluating Consolidation and the Threat of Monopolies within Industrial Sectors. Retrieved from https://www.csis.org/analysis/evaluating-consolidation-and-threat-monopolies-within-industrial-sectors

Joaquin, M. E., & Greitens, T. J. (2012, August 23). Contract Management Capacity Breakdown? An Analysis of U.S. Local Governments. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6210.2012.02587.x

McCormick, R., Hunter, A. P., Sanders, G., Cohen, S., & McQuade, M. R. (2018, October 05). Measuring the Outcomes of Acquisition Reform by Major DoD Components. Retrieved from https://www.csis.org/analysis/measuring-outcomes-acquisition-reform-major-dod-components

Romzek, B. S., & Johnston, J. M. (2002, July 01). Effective Contract Implementation and Management: A Preliminary Model. Retrieved from https://academic.oup.com/jpart/article-abstract/12/3/423/909448?redirectedFrom=fulltext

Sanders, G., & Hunter, A. P. (2017). Overseas Contingency Operations Contracts After Iraq: Enabling Financial Management Research and Transparency Through Contract Labeling. Retrieved from https://www.researchsymposium.com/conf/app/researchsymposium/unsecured/file/145/SYM-AM-17-051-005_Sanders.pdf

U. S. General Services Administration. (2019). Service Contract Inventory. Retrieved from https://www.acquisition.gov/service-contract-inventory

U.S. Government Accountability Office, Additional Actions and Data Are Needed to Effectively Manage and Oversee DOD's Acquisition Workforce, GAO-09-342 (Washington, DC, March 2009). Retrieved from https://www.gao.gov/assets/290/287498.pdf

Yang, K., Hsieh, Y. & Li, S. (2009) Contracting Capacity and Perceived Contracting Performance: Nonlinear Effects and the Role of Time, Public Administration Review, 69(4), pp. 681-696. Retrieved from https://www.jstor.org/stable/27697914?seq=1#page_scan_tab_contents

Software Employed

Bache, Stefan Milton and Hadley Wickham (2014). magrittr: A Forward-Pipe Operator for R. R package version 1.5. Retrieved from https://CRAN.R-project.org/package=magrittr

Bates, Douglas, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Gelman, Andrew and Yu-Sung Su (2018). arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package version 1.10-1. Retrieved from https://CRAN.R-project.org/package=arm

Grolemund, Garrett and Hadley Wickham (2011). Dates and Times Made Easy with lubridate. Journal of Statistical Software, 40(3), 1-25. Retrieved from http://www.jstatsoft.org/v40/i03/

Harrell, Frank E. Jr, with contributions from Charles Dupont and many others. (2019). Hmisc: Harrell Miscellaneous. R package version 4.2-0. Retrieved from https://CRAN.R-project.org/package=Hmisc

Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.1. https://CRAN.R-project.org/package=stargazer

Leifeld, Philip (2013). texreg: Conversion of Statistical Model Output in R to LaTeX and HTML Tables. Journal of Statistical Software, 55(8), 1-24. Retrieved from http://www.jstatsoft.org/v55/i08/

Nash, John C. and Ravi Varadhan (2011). Unifying Optimization Algorithms to Aid Software System Users: optimx for R. Journal of Statistical Software, 43(9), 1-14. Retrieved from http://www.jstatsoft.org/v43/i09/

Nash, John C. (2014). On Best Practice Optimization Methods in R. Journal of Statistical Software, 60(2), 1-14. Retrieved from http://www.jstatsoft.org/v60/i02/.

R Core Team (2017). foreign: Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', ... R package version 0.8-70. Retrieved from https://CRAN.R-project.org/package=foreign

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/.

Wickham, Hadley (2007). Reshaping Data with the reshape Package. Journal of Statistical Software, 21(12), 1-20. Retrieved from http://www.jstatsoft.org/v21/i12/

Wickham, Hadley (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickham, Hadley (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. Retrieved from https://CRAN.R-project.org/package=tidyverse

Wickham, Hadley, Jim Hester and Romain Francois (2018). readr: Read Rectangular Text Data. R package version 1.3.1. Retrieved from https://CRAN.R-project.org/package=readr

Wickham, Hadley (2018). scales: Scale Functions for Visualization. R package version 1.0.0. Retrieved from https://CRAN.R-project.org/package=scales

Wickham, Hadley, Romain François, Lionel Henry and Kirill Müller (2019). dplyr: A Grammar of Data Manipulation. R package version 0.8.3. https://CRAN.R-project.org/package=dplyr

Wickham, Hadley (2019). forcats: Tools for Working with Categorical Variables (Factors). R package version 0.4.0. Retrieved from https://CRAN.R-project.org/package=forcats

Wickham, Hadley (2019). stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.4.0. Retrieved from https://CRAN.R-project.org/package=stringr

Xie, Yihui (2015) Dynamic Documents with R and knitr. 2nd edition. Chapman and Hall/CRC. ISBN 978-1498716963

Xie, Yihui (2014) knitr: A Comprehensive Tool for Reproducible Research in R. In Victoria Stodden, Friedrich Leisch and Roger D. Peng, editors, Implementing Reproducible Computational Research. Chapman and Hall/CRC. ISBN 978-1466561595

Appendix A: About the Authors

Project Director:
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.

Authors:
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.

Email: GSanders@csis.org; Address: 1616 Rhode Island Ave NW, 20036, Washington, DC; Ph: 202.741.3916; Fax: 202.775.3199

Jonathan Roberts - is currently an Army ROTC cadet pursuing an M.A degree in applied economics at Georgetown University. He is also a graduate of Louisiana Tech University with a B.S. in mechanical engineering and a B.A. in political science. His research interests include economic statecraft in the era of great power competition and the economics of technological innovation.

Justin Graham is pursuing an M.A. in Security Studies at Georgetown University, and holds a B.A. (with honors) in Philosophy, Politics, and Economics from the University of Oxford. His work at CSIS focuses on machine learning and econometric modeling. Outside of CSIS, his research interests include technological innovation, counter-terrorism, and great power competition.

Lindsay McDonald is a rising senior at Oberlin College pursuing a BA in politics and economics, with a focus on quantitative analysis of public policy. Her work at CSIS included both model creation and extensive qualitative research on topics such as defense market consolidation, service contract performance, and foreign military sales. Her own research addresses political mobilization, transnational movements, and their influence on policy outcomes.

Contributing Authors:
Zach Huitink – is a postdoctoral fellow at Syracuse University and an externally affiliated researcher working with the Defense-Industrial Initiatives Group at CSIS. His research and analysis focuses on national security, government contracting, and public-private partnerships. Huitink holds a Ph.D. from the Department of Public Administration and International Affairs at Syracuse University's Maxwell School.

Xin Yuan is an intern with the Defense-Industrial Initiatives Group (DIIG) at CSIS. She mainly works on service contract inventory data consolidation, processing and visualization. She holds a B.S in economics with concentration of public finance & taxation from Tianjin University of Finance and Economics and a M.S in business analytics from George Washington University.

Robert Karlén graduated from the University of Washington with a BA in political science and international security. At CSIS, Robert helped with research and writing for various projects, including artificial intelligence, contracting performance, and Future Vertical Lift. Outside of CSIS, his work focuses on defense innovation and contracting.

Xinyi Wang was a research intern with the Defense-Industrial Initiatives Group (DIIG) at CSIS. Her work here including exploratory statistics and econometrics modeling for multiple quantitative research topics, like monopolies in defense industry, etc. She holds a B.A. in mathematical economics from Shanghai University of Finance and Economics, and a M.S in business analytics from George Washington University.

Appendix B: Variable Summary

The prefix of each variable indicator describes the structure of the data and the transformations performed on it: “c” means centered and rescaled; “l” means logarithmically transformed, adding 1 if zeroes are present in the data; “b” means binary data; “n” means general numerical data; “r” means a ratio where one indicator was divided by another; “p” means percentage, similar to ratio but typically capped at 100 percent and is always calculated from variables using a common measure. Categorical variables and their individual levels are not given a prefix.

Table B-1: Dependent Variables

Name	Variable Indicator	Type	Description
Some Option Exercised	b_SomeOpt	Binary	Whether the contract had at least one option exercised via a modification that did not increase scope. The sample for this for this model is restricted to those contracts with an initial base less than their initial ceiling.
All Options Exercised	b_AllOpt	Binary	Whether the vast majority, over 95 percent by value, of a contract’s options were exercised. The sample for the model using this variable is limited to contracts that exercised some options in order to more clearly illustrate the relationship between independent variables and all options being exercised.
Any Ceiling Breach	b_CBre	Binary	Whether a contract had its cost ceiling increased via a change order.
Ceiling Breach Size	ln_CBre	Continuous	The increased contract ceiling due to change orders, logged, in 2018 constant dollars based on contract start fiscal year. The dataset for these models is limited to contracts with any ceiling breaches, in order to avoid clustering at zero-based problems.
Any Termination	b_Term	Binary	Whether contracts experienced a partial or complete termination

Note: See section 4.2.1 Dependent Variables for additional detail and descriptive statistics.

Table B-2: Study Variables			
Name	Variable Indicator	Type	Description
Service Contract Complexity			
Detailed Industry Salary	cln_US6sal	Continuous	Average U.S. wage in the contract's detailed industry in the calendar year prior to contract start. Deflated to 2018 constant dollars based on the prior calendar year, logged, and rescaled.
Service Code Invoice Rate	cln_PSCrate	Continuous	Average annual invoice rate for the contract's product or service code in the fiscal year prior to contract start. In 2018 constant dollars based on the start fiscal year, logged and rescaled.
Contract Management Capacity			
Office Performance-Based	cp_OffPerf7	Continuous	Percentage share of office 2018 constant obligations that were for Performance-Based Service Contracts in the prior 7 fiscal years, rescaled.
Office Service Code Experience	cp_OffPSC7	Continuous	Percentage of office 2018 constant obligations that went to contracts with the same product or service code as the product in question over the last 7 fiscal years, rescaled.
Extent of Prior Relationship			
Paired Years	cn_PairHist7	Continuous	How many of the past 7 fiscal years involved obligations being made from the office to the vendor, recalled.
Paired Actions	cln_PairCA	Continuous	How many contracting actions the contract's vendor performed for the office in question in the prior fiscal year, logged and rescaled. 1 action was added to all office-vendor pairings to allow for logarithmic transformation even when there was no previous acts.

Note: See section 4.2.2 Study Independent Variables for additional detail and descriptive statistics.

Table B-3: Level 2, 3, and 4 and Variables Included in the Model				
Name	Variable Indicator	Level	Type	Description

Sector				
Subsector	NAICS3	4	Categorical	NAICS Subsector Code (3-digit code) with 82 subsectors within the defense services sample.
Detailed Industry	NAICS6	3	Categorical	NAICS Detailed Industry Code (6-digit code) in a given Subsector with 878 detailed industries within the defense services sample.
Service Area	ServArea	2	Categorical	Service Area is based on Product or service codes groupings (Construction; Equipment Related; Information Communications Technology, Facility-Related; Professional, Administrative, and Management) with each service area in a given NAICS Detailed Industry treated as a different group.
Acquirer				
Agency	Agency	3	Categorical	Contracting Agency Code with 26 agencies within the defense services sample.
Office	Office	2	Categorical	Contracting Office Code within a given Agency with 1,092 offices within the defense services sample.
Place	Place	2	Categorical	Country in which the contract was performed with 186 countries within the defense services sample.
Start Fiscal Year	StartFY	2	Categorical	Fiscal Year the contract was signed in with 9 groups from 2008 to 2015.

Table B-4: Number of Total Observations and Groups in Each Higher-Level Variables by Sample

Name	Population	Overall Sample	Options Samples		Breached Size Sample
			Some	All	
Contracts/Task Orders	1,241,449	1,000,000	74,274	51,357	61,184
Sector: Subsector	108	82	78	73	75
Sector: Detailed Industry	1,059	878	624	573	557
Sector: Service Area	3,825	3,242	1,508	1,327	1,292
Acquirer: Agency	28	26	25	25	24
Acquirer: Office	1,228	1,092	758	676	656
Place	197	186	121	103	123
Start Fiscal Year	9	9	9	9	9

Note: Some of the gap between the population and overall sample results from an uneven distribution of missing data, e.g. most NAICS codes that changed between 2007 and 2012 were set as missing.

Table B-5: Office and Vendor-Office Pair Level 1 Control Variables

Name	Variable Indicator	Type	Description

Office Obligations	cln_OffObl7	Continuous	Total office constant 2018 obligations in the prior 7 fiscal years, logged and rescaled. 1 dollar is added to all values to allow for logarithmic transformation of offices with no prior spending.
Paired Share	cp_PairObl7	Continuous	Percentage of an office's constant 2018 obligations that are accounted for by this vendor, rescaled.
Office Focus	cln_OffFocus	Continuous	Determines the distribution of office attention, as measured in contract counts processed through the Herfindahl-Hirschman Index, across different detailed industries, logged and then rescaled.

Note: The study variables Office Performance-Based and Office Service Code Experience vary based on office and year like the office focus variable in this table. The study variables Paired Years and Paired Actions vary with each combination of office, vendor, and year like the paired share variable in this table.

Table B-6: Subsector and Detailed Industry Level-1 Control Variables			
Name	Variable Indicator	Type	Description
Subsector HHI	cln_Def3HHI	Continuous	A measure of industrial concentration (Herfindahl-Hirschman Index) of subsector industries, logged and rescaled.
Detailed Industry HHI	cln_Def6HHI	Continuous	A measure of industrial concentration (Herfindahl-Hirschman Index) of detailed industries, logged and rescaled.
Subsector DoD:US	clr_Def3toUS	Continuous	Ratio of defense obligations to U.S. wide revenues in same NAICS subsector, logged and rescaled. Draws from economic census in calendar '07 (signed year '08-'12) and '12 (signed year '13-'15). Capped at 1.
Detailed Industry DoD:US	clr_Def6toUS	Continuous	Ratio of defense obligations to U.S. wide revenues in same NAICS detailed industry, logged and rescaled. Draws from economic census in calendar '07 (signed year '08-'12) and '12 (signed year '13-'15). Capped at 1.

Note: The Detailed Industry Salary study variable varies based on detailed industry and five year block, like the detailed industry DoD:US variable in this table.

Table B-4: Initial Contract Scope Level-1 Control Vehicles			
Name	Variable Indicator	Type	Description

Initial Base	cln_Base	Continuous	The initial contract base which excludes options that have not been exercised at contract start. Deflated in 2018 dollars based on contract start fiscal year, cost ceiling, logged and rescaled.
Initial Ceiling: Base	clr_Ceil2Base	Continuous	The initial ratio of the ceiling to the base, logged and then rescaled
Planned Duration	cln_Days	Continuous	The initial maximum duration of the contract in days, logged and rescaled

Table B-4: Description Independent Contract Level 1 Categorical Variables			
Name	Variable Indicator	Type	Description
Competition	Comp	Categorical	The level of competition for any given contract.
No Competition	(Baseline)	Binary	Whether the contract did not use competitive procedures.
1 Offer	1Offr	Binary	Whether contracts received 1 offer
2-4 Offers	2-4Offr	Binary	Whether contracts received 2, 3 or 4 offers.
5+ Offers	5plusOffr	Binary	Whether contracts received 5 or more offers.

Table B-4: Description Independent Contract Level 1 Categorical Variables			
Name	Variable Indicator	Type	Description
Vehicle	Veh	Categorical	Contract Vehicle: The omitted baseline is definitive awards and purchase orders.
Definitive / Purchase Order	(Baseline)	Categorical	Does the contract have no parent vehicle (I.e. definitive award or purchase Order)
Single-Award	SIDV	Binary	Is the contract vehicle is a single-award indefinite delivery contract.
Multi-Award	MIDV	Binary	Is the contract vehicle is a multiple-award indefinite delivery contract.
FSS-GWAC	FSS-GWAC	Binary	Is the contract vehicle is a Federal Supply Schedule or Government-Wide Acquisition Contract.
BPA-BOA	BPA-BOA	Binary	Whether the contract vehicle is a Blank Purchase Agreement or Basic Ordering Agreement.

Table B-4: Description Independent Contract Level 1 Categorical Variables			
Name	Variable Indicator	Type	Description
Pricing	Pricing	Categorical	The means of calculating vendor payments. The omitted baseline is firm fixed price contracts.
Firm Fixed-Price	(Baseline)	Binary	Firm-Fixed Price Contract
Other Fixed-Price	Other_FP	Binary	Fixed-price redetermination, fixed-price award fee, and fixed-price economic price adjustment
Incentive Free	Incentive	Binary	Including fixed-fee incentive fee, cost plus incentive fee, and cost sharing
Combination or Other	Comb-Other	Binary	Covers contracts using multiple pricing mechanisms or unusual and unclassified types
Other Cost-Based	Other_CB	Binary	All types of cost-based contracts, excluding incentive fee
Time & Materials / Labor Hours/ FFP:LoE	TM-LH-FPLOE	Binary	Time and materials, labor hours, and fixed-price level of effort contracts respectively

Table B-4: Description Independent Contract Level 1 Categorical Variables			
Name	Variable Indicator	Type	Description
Crisis Funding	Crisis	Categorical	Whether the funding came from emergency accounts
Not Crisis	(Baseline)	Binary	Whether CSIS classified the contract as crisis-funded.
Recovery Act	ARRA	Binary	Whether the funds come from the American Recovery and Reinvestment Act
Disaster	DIS	Binary	Whether the funds come from discretionary spending
OCO	OCO	Binary	Whether the funds come from the Overseas Contingency Operations fund

Appendix C: Model Odds Ratios

Table C-1 Some Option Exercised Odds Ratios

variable	OR	2.50%	97.50%
(Intercept)	1.05	0.68	1.63
Log(Det. Ind. Salary)	1.01	0.90	1.12
Log(Serv. Code Invoice Rate)	0.94	0.89	1.00
Office Perf.-Based %	0.85	0.77	0.93
Office Serv. Code Exp. %	0.83	0.75	0.93
Paired Years	1.30	1.24	1.37
Log(Paired Actions)	0.93	0.85	1.02
Log(Init. Base)	1.25	1.17	1.34
Log(Init. Ceiling:Base)	1.35	1.31	1.39
Log(Planned Dur.)	1.33	1.25	1.42
Comp=1 offer	1.01	0.95	1.08
Comp=2-4 offers	0.97	0.92	1.02
Comp=5+ offers	0.94	0.88	1.00
Vehicle=S-IDC	0.50	0.46	0.54
Vehicle=M-IDC	0.79	0.73	0.86
Vehicle=FSS/GWAC	0.84	0.78	0.91
Vehicle=BPA/BOA	0.75	0.64	0.87
Pricing=Other Fixed-Price	0.50	0.36	0.69
Pricing=Incentive Fee	0.71	0.39	1.27
Pricing=Comb. or Other	1.43	1.21	1.69
Pricing=Other Cost-Based	1.28	1.14	1.45
Pricing=T&M/LH/FP:LoE	1.06	0.93	1.20
Pricing=UCA	0.91	0.73	1.15
Crisis=Recovery Act	0.66	0.48	0.91
Crisis=Disaster	0.73	0.40	1.33
Crisis=OCO	0.71	0.56	0.89
Log(Det. Ind. HHI)	0.88	0.80	0.96
Log(Det. Ind. DoD:US)	0.99	0.91	1.07
Log(Subsector HHI)	0.98	0.88	1.09
Log(Subsector DoD:US)	1.03	0.88	1.21
Log(Paired Share %)	0.98	0.89	1.09
Log(Office Obl.)	1.22	1.11	1.35
Log(Office Focus)	0.72	0.65	0.81
Office Perf.-Based %:Log(Paired Share %)	2.00	1.69	2.36
Paired Share %:Log(Office Obligations)	1.14	0.91	1.43
Paired Years:Pricing=Other Fixed-Price	0.91	0.44	1.91

Paired Years:Pricing=Incentive Fee	4.72	1.75	12.77
Paired Years:Pricing=Comb./Other	0.64	0.45	0.90
Paired Years:Pricing=Other Cost-Based	1.23	1.01	1.48
Paired Years:Pricing=T&M/LH/FP:LoE	1.09	0.86	1.37
Paired Years:Pricing=UCA	1.17	0.76	1.80
Office Perf.-Based %:Log(Paired Actions)	0.54	0.49	0.61
Log(Init. Base):Log(Init. Ceiling:Base)	1.30	1.25	1.35
Office Serv. Code Exp. %:Log(Office Obl.)	2.79	2.05	3.79

Table C-2 All Options Exercised Odds Ratios

variable	OR	2.50%	97.50%
(Intercept)	3.15	2.37	4.20
Log(Det. Ind. Salary)	1.08	0.97	1.21
Log(Serv. Code Invoice Rate)	0.94	0.88	1.00
Office Perf.-Based %	1.00	0.91	1.10
Office Serv. Code Exp. %	1.19	1.06	1.34
Paired Years	1.16	1.10	1.22
Log(Paired Actions)	0.80	0.73	0.88
Log(Init. Base)	0.87	0.82	0.92
Log(Init. Ceiling:Base)	0.45	0.43	0.46
Log(Planned Dur.)	0.64	0.59	0.70
Comp=1 offer	1.00	0.94	1.07
Comp=2-4 offers	0.96	0.90	1.01
Comp=5+ offers	0.92	0.86	0.99
Vehicle=S-IDC	0.82	0.75	0.90
Vehicle=M-IDC	0.88	0.81	0.96
Vehicle=FSS/GWAC	0.94	0.87	1.02
Vehicle=BPA/BOA	1.28	1.06	1.56
Pricing=Other Fixed-Price	0.61	0.40	0.93
Pricing=Incentive Fee	1.79	0.97	3.30
Pricing=Comb. or Other	0.87	0.75	0.99
Pricing=Other Cost-Based	1.24	1.09	1.41
Pricing=T&M/LH/FP:LoE	0.97	0.84	1.11
Pricing=UCA	0.91	0.68	1.21
Crisis=Recovery Act	0.79	0.51	1.21
Crisis=Disaster	0.23	0.10	0.53
Crisis=OCO	0.77	0.60	0.98
Log(Det. Ind. HHI)	0.92	0.83	1.00
Log(Det. Ind. DoD:US)	1.04	0.94	1.15

Log(Subsector HHI)	0.95	0.85	1.07
Log(Subsector DoD:US)	0.98	0.84	1.15
Log(Paired Share %)	0.93	0.83	1.05
Log(Office Obl.)	0.92	0.86	0.97
Log(Office Focus)	0.88	0.78	0.98
Office Perf.-Based %:Log(Paired Share %)	1.16	0.96	1.39
Office Serv. Code Exp. %:Log(Office Focus)	0.91	0.73	1.13
Log(Serv. Code Invoice Rate):Pricing=Other Fixed-Price	0.79	0.24	2.61
Log(Serv. Code Invoice Rate):Pricing=Incentive Fee	0.55	0.14	2.17
Log(Serv. Code Invoice Rate):Pricing=Comb./Other	1.49	1.01	2.22
Log(Serv. Code Invoice Rate):Pricing=Other Cost-Based	1.54	1.17	2.03
Log(Serv. Code Invoice Rate):Pricing=T&M/LH/FP:LoE	1.19	0.80	1.76
Log(Serv. Code Invoice Rate):Pricing=UCA	4.15	2.01	8.57
Log(Office Obligations):Log(Office Focus)	1.00	0.86	1.15

Table C-3 Ceiling Breach Odds Ratios

variable	OR	2.50%	97.50%	
(Intercept)		0.01	0.01	0.01
Log(Det. Ind. Salary)		1.00	0.88	1.12
Log(Serv. Code Invoice Rate)		1.06	1.03	1.09
Office Perf.-Based %		1.08	1.02	1.13
Office Serv. Code Exp. %		1.21	1.15	1.27
Paired Years		0.92	0.90	0.95
Log(Paired Actions)		1.56	1.50	1.62
Log(Init. Base)		3.84	3.74	3.94
Log(Init. Ceiling:Base)		1.28	1.25	1.31
Log(Planned Dur.)		1.32	1.29	1.36
Comp=1 offer		0.97	0.93	1.00
Comp=2-4 offers		1.11	1.08	1.14
Comp=5+ offers		1.19	1.16	1.23
Vehicle=S-IDC		0.62	0.60	0.64
Vehicle=M-IDC		0.83	0.80	0.87
Vehicle=FSS/GWAC		0.99	0.93	1.05
Vehicle=BPA/BOA		0.75	0.70	0.80
Pricing=Other Fixed-Price		0.67	0.57	0.80
Pricing=Incentive Fee		10.87	9.40	12.56
Pricing=Comb. or Other		1.30	1.18	1.42
Pricing=Other Cost-Based		0.91	0.85	0.97
Pricing=T&M/LH/FP:LoE		1.12	1.03	1.21
Pricing=UCA		1.07	0.99	1.17
Crisis=Recovery Act		1.13	1.04	1.23
Crisis=Disaster		1.07	0.89	1.28
Crisis=OCO		0.90	0.82	1.00
Log(Det. Ind. HHI)		0.98	0.94	1.01
Log(Det. Ind. DoD:US)		1.07	1.01	1.13
Log(Subsector DoD:US)		0.63	0.55	0.72
Log(Paired Share %)		0.79	0.75	0.83
Log(Office Obl.)		1.04	1.00	1.08
Log(Office Focus)		0.70	0.64	0.76
Office Perf.-Based %:Log(Paired Actions)		1.47	1.40	1.54
Office Perf.-Based %:Log(Planned Dur.)		0.89	0.85	0.93

Table C-4 Terminations Odds Ratios

variable	OR	2.50%	97.50%
(Intercept)	0.01	0.01	0.02
Log(Det. Ind. Salary)	0.95	0.85	1.07
Log(Serv. Code Invoice Rate)	0.98	0.93	1.03
Office Perf.-Based %	1.03	0.95	1.12
Office Serv. Code Exp. %	1.36	1.25	1.48
Paired Years	0.76	0.73	0.79
Log(Paired Actions)	0.96	0.91	1.02
Log(Init. Base)	1.34	1.28	1.40
Log(Planned Dur.)	2.46	2.35	2.58
Log(Init. Ceiling:Base)	1.65	1.62	1.69
Comp=1 offer	1.34	1.27	1.42
Comp=2-4 offers	1.45	1.38	1.52
Comp=5+ offers	2.04	1.94	2.15
Vehicle=S-IDC	0.51	0.48	0.54
Vehicle=M-IDC	0.66	0.62	0.71
Vehicle=FSS/GWAC	0.83	0.76	0.91
Vehicle=BPA/BOA	0.38	0.34	0.43
Pricing=Other Fixed-Price	0.39	0.32	0.47
Pricing=Incentive Fee	0.36	0.17	0.77
Pricing=Comb. or Other	0.85	0.69	1.05
Pricing=Other Cost-Based	0.69	0.60	0.80
Pricing=T&M/LH/FP:LoE	0.64	0.54	0.76
Pricing=UCA	0.49	0.37	0.64
Crisis=Recovery Act	0.73	0.57	0.93
Crisis=Disaster	1.61	1.11	2.34
Crisis=OCO	0.93	0.80	1.08
Log(Det. Ind. HHI)	1.04	0.96	1.13
Log(Det. Ind. DoD:US)	1.02	0.95	1.10
Log(Subsector HHI)	1.03	0.95	1.12
Log(Subsector DoD:US)	1.27	1.06	1.52
Log(Paired Share %)	0.88	0.82	0.95
Log(Office Obl.)	1.03	0.97	1.09
Log(Office Focus)	0.79	0.71	0.87
Paired Years:Pricing=Other Fixed-Price	1.09	0.77	1.55
Paired Years:Pricing=Incentive Fee	0.91	0.25	3.28
Paired Years:Pricing=Comb./Other	1.38	0.88	2.17
Paired Years:Pricing=Other Cost-Based	1.35	1.04	1.75
Paired Years:Pricing=T&M/LH/FP:LoE	0.55	0.39	0.77
Paired Years:Pricing=UCA	0.60	0.36	1.00

Appendix D: Model Diagnostics

Figure 3 Fitted and Residual Plots for Exercised Options

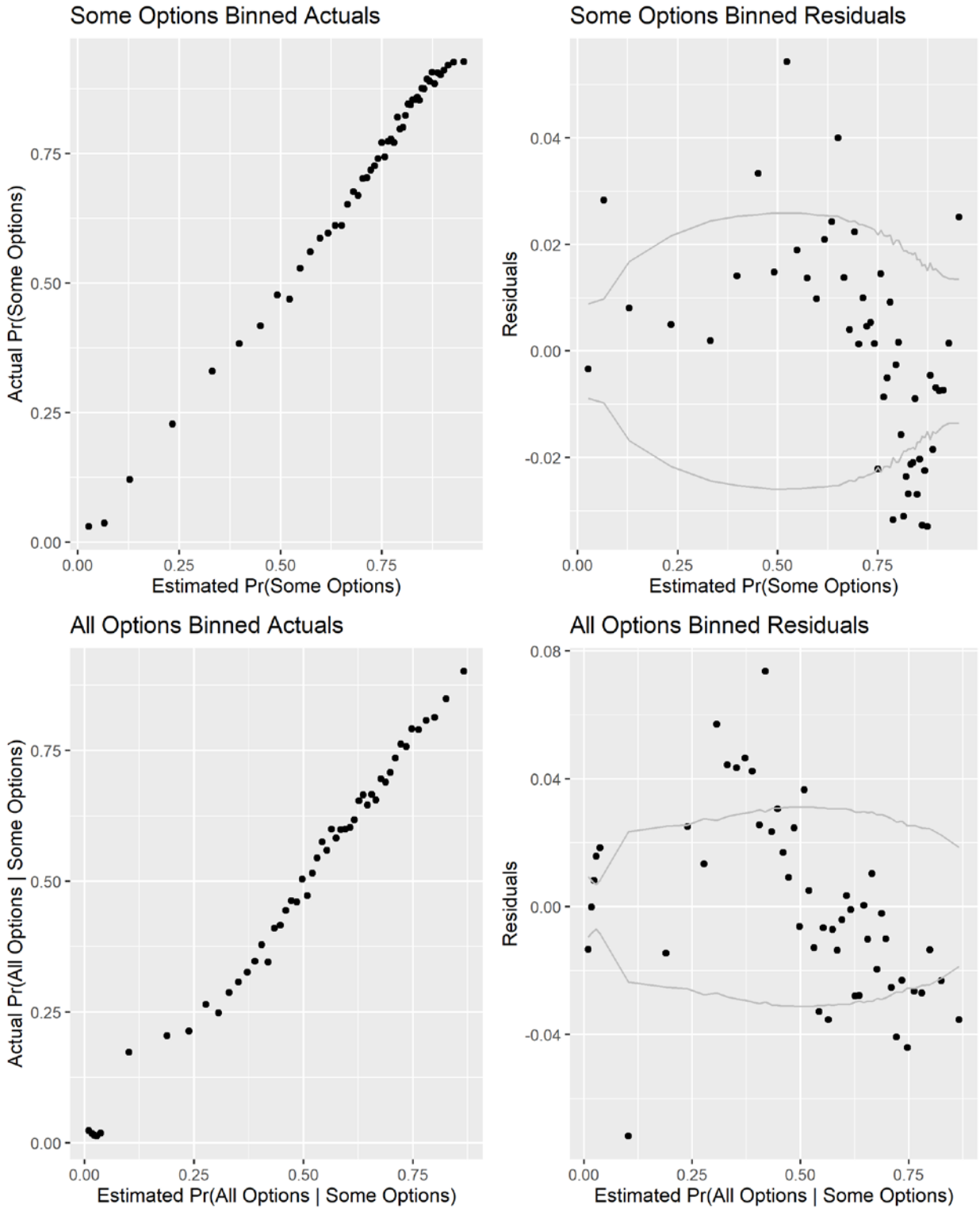


Figure 4 Fitted and Residual Plots for Ceiling Breaches

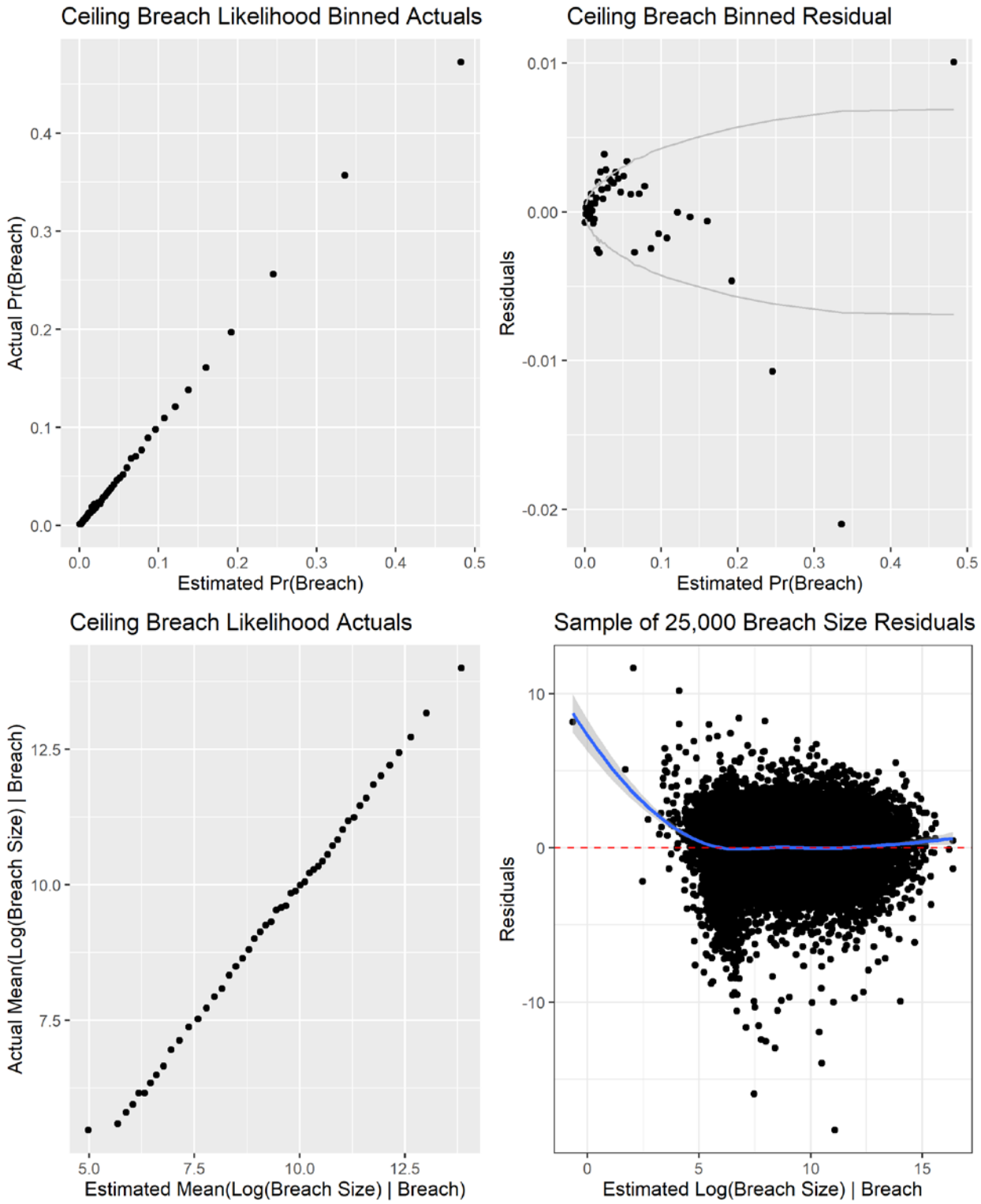


Figure 5 Fitted and Residual Plots for Terminations