Contract Incentives Under Uncertainty: Data-Driven Contract Geometry Best Practices

Peter Braxton, Kevin Hetrick, Kyle Webb, Ian Whitehead-Scanlon, Derreck Ross

Abstract
With the current emphasis on Incentive Contracting, the cost analyst plays a vital role in establishing target cost, fee, and shareline. Drawing from the Joint Contract Price Database, this paper examines actual contract geometries across a wide range of programs and their effectiveness in encouraging cost management. Using the published Risk-Based ROS methodology, it provides a framework for implementing incentive arrangements for both competitive and negotiated procurements.

Keywords
Cost Management, Data-Driven, DOD/MOD, Government, Methods, Program Management, Risk, Statistics, Contracts, Contract Management, Incentive Contracting, Incentive Fee, CPIF, FPIF

"Getting Defense Acquisition Right": The Role of Contract Incentives
Former Secretary of Defense (SECDEF) Ashton Carter and Under Secretary of Defense for Acquisition, Technology and Logistics (USD(AT&L)) Frank Kendall in their Better Buying Power (BBP) initiative stressed the importance of both incentives in general and contract incentives in particular. Many of Dr. Kendall's related papers and articles were recently published in book form by Defense Acquisition University (DAU)[1].

Better Buying Power (BBP) and Incentives: "The Threat of Competition"
The basic principle of incentives is the establishment of one or mechanism that will structurally encourage and reinforce the desired behavior. One essential BBP mechanism is that of competition, or "the threat of competition." If programs can avoid a comfortable sole source environment, it is thought that the "pressure" or threat of losing work will lead to more prudent management, more dedicated "varsity team" staffing, and other actions that help contain costs. While this may often be true, "winner-take-all" may also lead to rash decisions and desperate approaches (typically, "buying in"). Instances of markedly different outcomes can be found, even in the same realm. For example, in the annals of U.S. Navy amphibious ships, the LSD 41 class achieved negative cost growth thanks to competitive bidding amongst six shipyards[2], whereas LPD 17 became one of the notorious examples of cost growth[3].

---

[1] *Getting Defense Acquisition Right*, The Honorable Frank Kendall, Defense Acquisition University Press, 2017.
[2] "The End of S-Curve Alchemy: Gold from a New SAR Database," Todd Andrews, Jeff Pincus, Brian Flynn, ICEAA, 2017.
[3] "Lessons To-be-Learned: Causes of Cost Growth in LPD 17," Dr. Tzee-nan K. Lo, Dr. David L. McNicol, Department of Defense Cost Analysis Symposium (DoDCAS), 2006.

Competitive environment and overarching Acquisition Strategy are beyond the scope of this paper, though it behooves us to keep in mind how they may affect the initial baseline on particular contracts. Our focus is on Incentive Contracts, also known as incentive-type contracts. These are defined in the Federal Acquisition Regulation (FAR) to include Fixed Price Incentive (FPI) contracts, with both Firm and Successive targets; Cost Plus Incentive Fee (CPIF) contracts; and Cost Plus Award Fee (CPAF) contracts. Because the latter employs subjective scoring against established criteria by an Award Fee (AF) panel, it is more challenging to address in research such as this. We will primarily treat FPIF and CPIF in contrast to other (non-incentive) contract types.

FPIF and CPIF contracts both provide a cost incentive, the severity of which is dependent in large part on the share ratio established to distribute cost risk equitably between the Government and Contractor. Share ratio is generally specified by two numbers that sum to 100%, separated with a slash, with the Government share always coming first. For example, a 70/30 share ratio denotes a Government share of 70% and a Contract share of 30%. This means that the Government will "cover" 70 cents on the dollar of any overrun, increasing its price, while the Contractor will "eat" the remaining 30 cents on the dollar, reducing its profit. Often a *different* share ratio is specified for below target cost, but for the sake of explanation, the same 70/30 share ratio would result in the Government's reaping the benefit of 70 cents on the dollar for underruns, reducing its price, which the Contracts reaps the remaining 30 cents on the dollar, increasing its profit.

The combination of "over" and "under" share ratios, along with other relevant contract geometry parameters, define what is known as a shareline. It is common to draw graphs of both price (the Government's primary concern) and profit (the Contractor's primary concern) as a function of final cost. This is done in Module 14 Contract Pricing of ICEAA's Cost Estimating Body of Knowledge (CEBoK®). The example shown below is from the Contract Incentive Impact Tool (CIIT). DAU has a similar Excel workbook that can graph sharelines.

FPI								
Cost	**$M**	**Profit/Fee**	**$M**	**%**	**Price**	**$M**	**%**	
Target Cost	$ 10.0	Target Profit/Fee	$ 1.0	10.0%	Target Price	$ 11.0		
		Margin (ROS)	$ 1.0	9.1%				
PTA	$ 12.9				Ceiling Price	$ 13.0	130.0%	*applies to FPI only*
RIE Low	$ 9.2	Min Fee	$ 0.3	3.0%				*applies to CPIF only*
RIE High	$ 12.3	Max Fee	$ 1.5	15.0%				*applies to CPIF only*
		Share Ratio						
		Under Gov Share		40%	Under Cont Share		60%	*applies to FPI and CPIF*
		Over Gov Share		70%	Over Cont Share		30%	*applies to FPI and CPIF*

Figure 1. Contract Incentive Impact Tool (CIIT) Sample Inputs

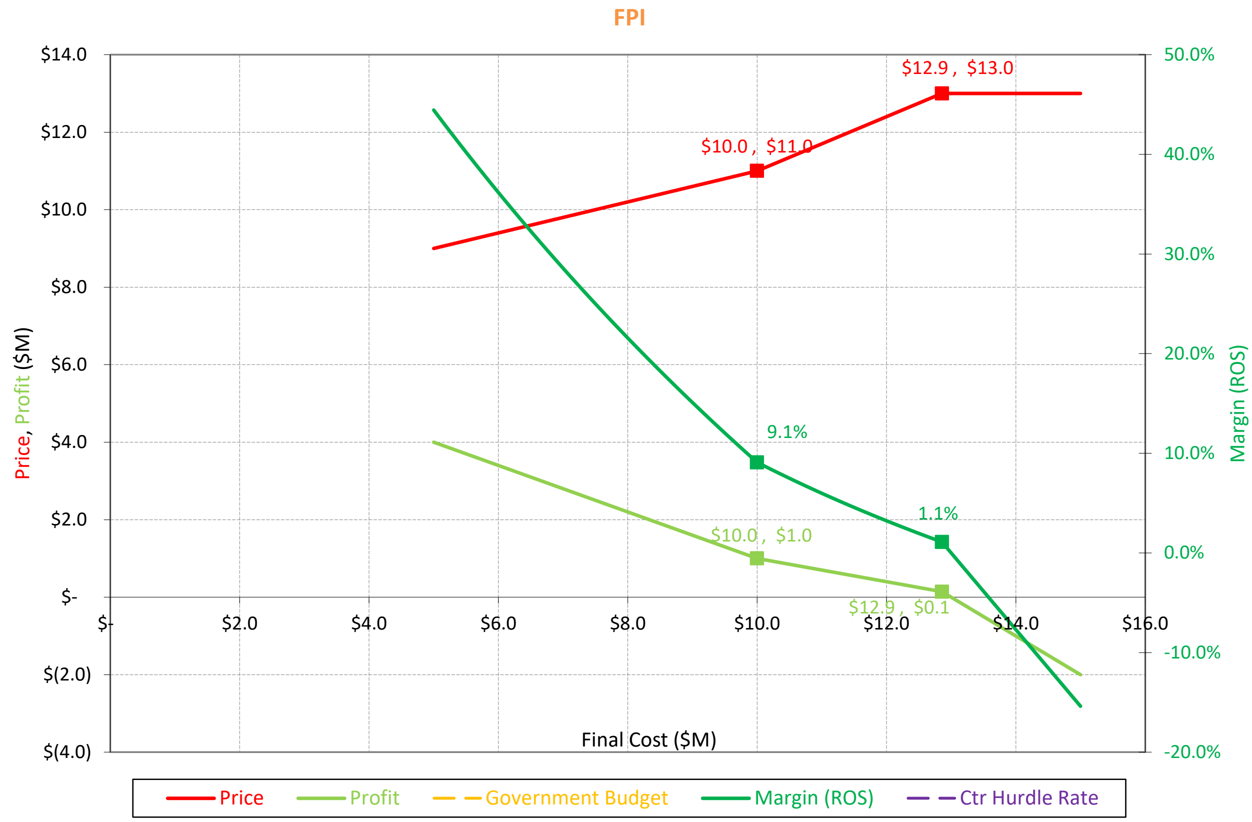

Figure 2. FPI Shareline Example

Other contract incentives are possible, including schedule and performance incentives. Schedule incentives tie additional fee payments to on-time (or early) delivery. Performance incentives tie additional fee payments to achievement of threshold (or objective) performance targets, such as a certain sortie rate in aircraft carrier design. (Note that these performance incentives are often somewhat problematic, as they rely on simulation results and not operational testing. We will focus on cost incentives, for two primary reasons. First, they are by far the most common, since the FAR requires at least one cost incentive whenever incentives are used.[4] Second, since they are the results of the piecewise linear functions that capture the behavior of the price and profit sharelines in a mathematically precise manner, they lend themselves to systematic quantitative analysis.

Contract Incentives Common Misconceptions

There are several misconceptions about contract types and contract incentives.

First, "contract type" is itself a misnomer, because a contract that is all one contract type is the exception not the rule. Contract types actually apply at the Contract Line Item Number (CLIN) level, and it is common for a contract (or delivery order) to have multiple CLINs across multiple contract type.

"Incentive-type contracts" is also a misnomer, since *all* contract types provide an incentive to control costs, albeit to varying degrees. In fact, the only contract type that would provide the opposite

[4] Federal Acquisition Regulation (FAR), Subpart 16.4—Incentive Contracts, 16.401 General, https://www.acquisition.gov/?q=/browse/far/16.

incentive – to overrun *ad infinitum* – is expressly forbidden by the FAR: a Cost Plus contract with a fixed *percentage* fee[5]. The fixed fee in CPFF refers to a fixed *dollar amount* of fee, so that while this amount is often initially determined as a percentage of target cost, it is thereafter fixed (barring any contract modifications). This means that while profit dollars are constant, the margin percentage, commonly known as Return on Sales (ROS) because it is expressed as a fraction of total sales (Cost + Profit = Revenue), is actually decreasing as final cost increases. This is shown in the below example from CIIT. (To make this clearer, the author has proposed a simple notation for "universal contact type," consisting of a series of break points and intervening share ratios[6]. This would accommodate all the major FAR-specific contract types, including both CPFF – a 100/0 share ratio – and Firm Fixed Price (FFP) – a 0/100 share ratio.)

Figure 2. CPFF Contract Geometry Showing Cost Incentive

As this CPFF example illustrates, the omnipresent cost incentive of FAR-approved contract types is tantamount to the fact that since (1) the price to the government, representing revenue to the contractor, is monotonically non-decreasing, and (2) the profit to the contractor is monotonically non-increasing, the margin percent (hereinafter "ROS") must actually be strictly decreasing as final cost increases, which is the exact cost control incentive desired. While the contractor's goals of increasing total revenue and increasing total profit may seem to be at odds with each other in the context of a single contract, the latter tends to win out in response to demands for profitability on the part of shareholders of publicly-traded corporations, while the former is more often addressed at a portfolio level (i.e., more contracts and more diverse contracts).

Note that a significant burden falls on the government in contract administration. In particular, the procuring contracting officer (PCO) must take care to use contract modifications (hereinafter "mods") that increase target cost, and with it target profit and target price, only to add genuinely new work and *not* as "aiding and abetting" cost overruns. (This is a loose analogue of the caution to program managers in Earned Value Management (EVM) not to use Management Reserve (MR) to "cover" cost overruns on a particular control account.) This is because such mods provide additional revenue without any degradation in ROS, which while fair and equitable for new work, essentially violates the spirit of the FAR's prohibition on cost-plus-fixed-percentage-fee contract types for existing work.

It is straightforward to demonstrate the effectiveness cost incentives *absent contract mods* using shareline graphs produced by a tool like CIIT (or "homegrown" in Excel), but mods – like any kind of rebaseline – essentially reset the playing field. Similar to Terms and Conditions (hereinafter "Ts & Cs"), as noted in a previous paper[7], mods can essentially take cost growth *off the shareline*. To further explore how this happens, we wish to explore how this has manifested on real program in the past. If only we had an extensive historical contracts database….

---

[5] FAR 16.102 Policies (c), https://www.acquisition.gov/?q=/browse/far/16.
[6] "Risk-Based Return On Sales (ROS) As a Tool For Complex Contract Negotiations," Peter J. Braxton, Richard L. Coleman, DoDCAS, 2010.
[7] "Risk-Based Return On Sales (ROS) for Proposals with Mitigating Terms and Conditions," P.J. Braxton, R.L. Coleman, E.R. Druker, B.L. Cullis, C.M. Kanick, A.V. Bapat, J.M. Callahan, B.P. Caccavale, SCEA/ISPA, 2009.

Utilizing the Full Spectrum of Contract Incentives

In order to utilize the full spectrum of contract incentives, we wish to understand when it is appropriate to use each contract type – as noted above, often simultaneously across multiple CLINs of the same contract – and how we can most effectively establish the contract geometry for each. As we build this Contract Risk Management Framework, we will tap into a powerful existing database that can serve as a foundation for data-driven contract management, but first we need to look at the all-too-frequent disconnects between the disciplines of cost and risk analysis on the one hand and contract management on the other.

Disconnects Between Cost and Risk Analysis and Contract Management

Contract type and contract geometry have been largely ignored in cost and risk analysis (with notable exceptions[8]), assumed away in historical cost growth analysis, and generally relegated to the realm of sensitivity analysis (as opposed to risk analysis) in Contract Management. We briefly treat each of these disconnects before attempting to address them in our Contract Risk Management Framework.

Profit/Fee in Cost Analysis

Cost analysts generally estimate "at cost" and then tack on a typical fee. Often this is some sort of a perceived or rule-of-thumb "maximum" fee, such as 15% for Development efforts. The oft-quoted 6% Architectural and Engineering (A&E), 10% Production and Services, and 15% R&D fee numbers are actually only specified in the FAR as maximum values for the CPFF contract type[9], though they can certainly be used as guidelines for target fee for other contract types.

The problem with this approach is that it improperly models percentage fee or profit as invariant with cost in typical risk analyses performed via Monte Carlo simulation. Ideally, contract geometry would be incorporated where known (or reasonably estimated) for major contracts, in which case cost risk and uncertainty are run through the "filter" of contract geometry to determine the resultant distribution of both Price and Margin. This implementation would be straightforward for both Monte Carlo and enhanced Scenario-Based Method (eSBM) approaches[10]. To further refine, we'd want to appropriately model the separation of *on-shareline* growth and *off-shareline* growth (as the result of mods and/or Ts & Cs), but this requires delving into our historical database first.

Profit/Fee in Historical Cost Growth Analysis

The previous section begs the question of where the cost risk distributions come from in the first place. Ideally, they are from historical cost growth analysis, including the update of research using Selected Acquisition Report (SAR) data sponsored by the Naval Center for Cost Analysis (NCCA) and presented at this very same conference ("The End of S-Curve Alchemy: Gold from a New SAR Database," Todd Andrews, Jeff Pincus, Brian Flynn, ICEAA, 2017). But this analysis, whether done at the program level (using SARs, e.g., as in the aforementioned paper) or contract/CLIN level (as in our contracts database) is essentially price growth, not cost growth. The effects of contract geometry are essentially "buried" in each data point, and previous work has demonstrated that this contract geometry often produces an

---

[8] "Risk-Based Return On Sales (ROS) As a Tool For Complex Contract Negotiations," Peter J. Braxton, Richard L. Coleman, DoDCAS, 2010.

[9] FAR 15.404-4 Profit (c)(4)(i), https://www.acquisition.gov/?q=/browse/far/15.

[10] Paul R. Garvey , Brian Flynn, Peter Braxton & Richard Lee (2012): Enhanced Scenario-Based Method for Cost Risk Analysis: Theory, Application, and Implementation, Journal of Cost Analysis and Parametrics, 5:2, 98-142.

astoundingly counterintuitive distribution of ROS (and similarly price) for the "mainstream" contract types.

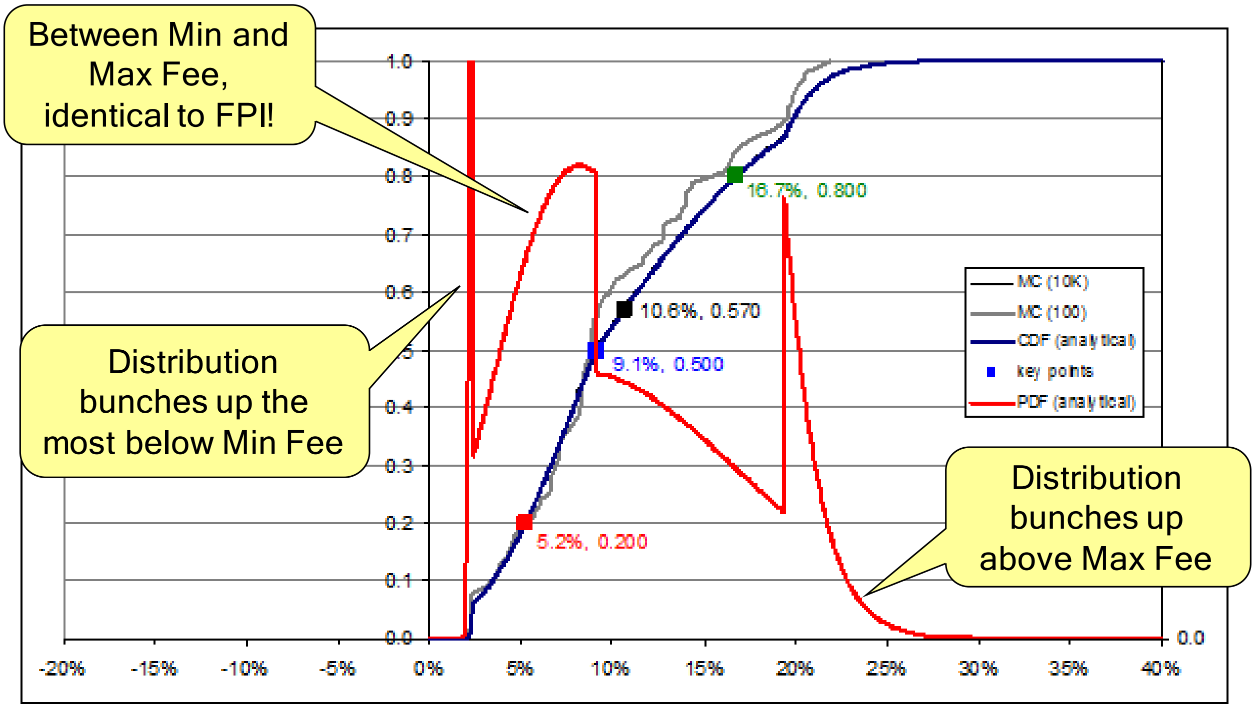

Figure 3. ROS Distribution for CPIF

In order to get to true cost growth, we need both to separate out cost and price, and to separate out cost growth into on- and off-shareline. This paper lays out a path for doing so as we continue to enhance our contracts database over the coming year.

Schedule Growth is also important and is supported by our contracts database (slips in final delivery over time) but beyond the scope of this paper.

Profit/Fee in Contract Management

In the realm of Contract Management, profit and fee are generally the subject of negotiations (especially in a sole source environment) and the application of "weighted guidelines." The focus is to avoid excessive profit, and rightfully so, but this is often considered in isolation from "most likely cost." Ideally, the negotiation of contract geometry, including profit/fee, and the reconciliation of cost risk analysis would occur in conjunction with each other and not in isolation. (This approach is outlined in the Contracts Risk paper given multiple times to both the National Contract Management Association (NCMA) and International Cost Estimating and Analysis Association (ICEAA).) The Government coming to a greater realization that there is an inherent trade-off for the corporations between target fee (or more importantly target margin) and risk position of the bid. They may accept a riskier position as long as the target fee is lucrative enough to guard against erosion in the case of an overrun. Conversely, they may accept a lower target fee in return for a more conservative target cost.

Contracts Database (KDB) As an Invaluable Resource for Data-Driven Contract Management

The sharing of lessons learned in Contract Management is often restricted to within an individual buying command. Broader efforts at information sharing such as CBAR (cite), sponsored by the Defense Contract Management Association (DCMA), have focused on areas such as contractor rates. The Contract Price and Schedule Database, or Contracts Database for short (hereinafter "KDB") is unique in providing CLIN-level price and schedule data, tracking value changes by every single mod for a broad collection of contracts exceeding half a trillion dollars in total value.

KDB Scope and Current Status

The thought leadership for KDB has come in large part from the Air Force Cost Analysis Agency (AFCAA), who have been stalwart sponsors over the years. Guidance and funding have also been provided by Marine Corps Program Executive Office Land Systems (PEO LS), the Office of the Deputy Assistant Secretary of the Army for Cost and Economics (ODASA-CE), and NCCA. Recently, KDB was brought under the aegis of the Cost Assessment Data Enterprise (CADE) initiative (http://cade.osd.mil/) to further ensure alignment with broader cost community data needs.

The below Galaxy Chart captures a snapshot of KDB, providing number of contracts by Service, Commodity, and Phase. (Note that for our purposes, each individual delivery order, or DO, is counted as its own contract.)

Refer to the paper presented previously at ICEAA by Brian Octeau for more background on KDB.

Current efforts to greatly accelerate the pace of data entry into KDB are discussed in the Automation section below.

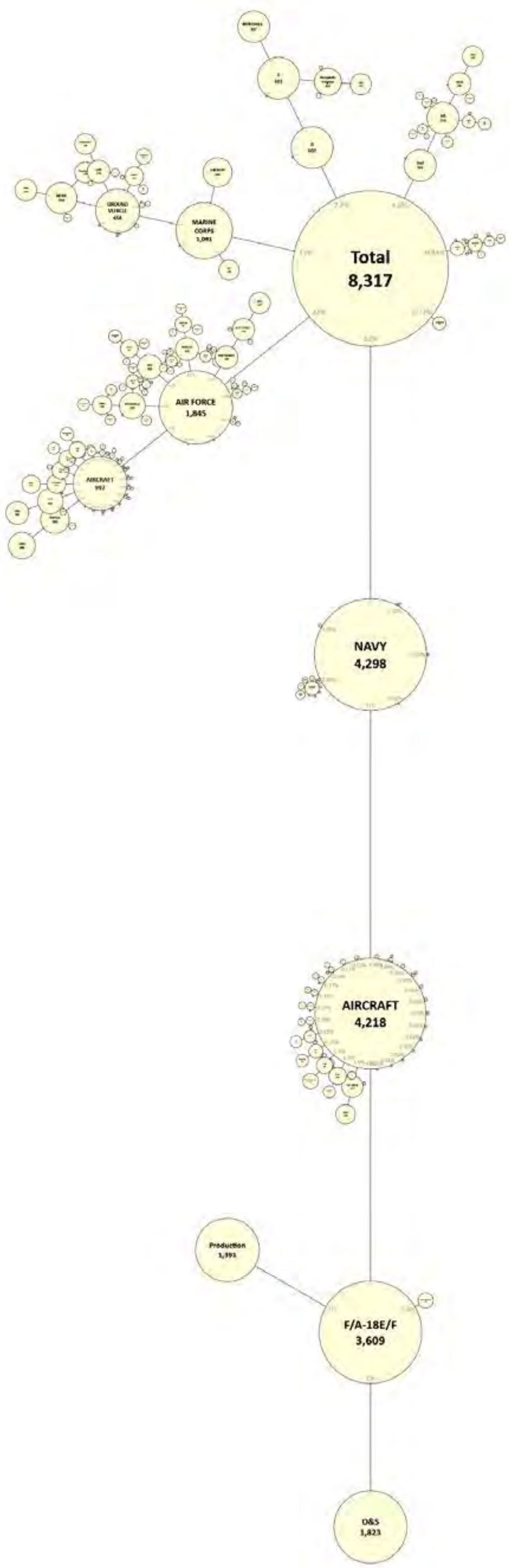
Total
8,317
MARINE
CORPS
1,091
GROUND
VEHICLE
454
AIR FORCE
1,845
AIRCRAFT
992
NAVY
4,298
AIRCRAFT
4,218
Production
1,391
F/A-18E/F
3,609
O&S
1,823

Figure 4. KDB Contracts by Service, Commodity, and Phase

CLIN-Level Cost, Fee, Price, and Schedule Data (Including Growth!)

For each contract in KDB, including all delivery orders (DO's) associated with -D- and -G- contracts, all CLINs and Mods are enumerated, and every price and schedule change to every CLIN as a result of each mod captured. Since we want to capture total contract value, KDB primarily stores value changes at price, but a recent effort has begun to break out price into cost and profit/fee where possible (e.g., CPFF, CPAF, CPIF, FPIF). For the FFP contract type in particular, the contracts and mods do not break out cost and profit. The cost community continues to rely on Contractor Cost Data Reports (CCDRs) for insight into profitability on FFP contracts. Under the auspices of CADE, profitability analysis is a key area of promise for data fusion between KDB and CCDRs.

One of the key analytical steps in populating KDB is the classification of each mod as to the reason for the price change. The below table shows the definitions of the various Growth Categories.

Table 1. KDB Growth Categories

Baseline	for all modifications where ***anticipated* changes** impact the contract value.
Cost	for all modification where ***unanticipated* price changes** <u>with *no* change in scope</u> impact the contract value
Technical	for all modifications where ***unanticipated* scope changes** impact the contract value
Schedule	for all modifications where ***unanticipated* schedule changes** impact the contract value
Administrative	for all modifications that <u>do not change the overall contract price</u>

In view of these definitions, it is hypothesized that Cost and Schedule mods constitute on-shareline growth, whereas Baseline and Technical mods constitute off-shareline growth, as discussed below.

Benchmarks for Contract Type and Contract Geometry Parameters

As noted above, "contract type" is misnomer, since it applies at the CLIN level, not the contract level. Therefore, we must mine KDB at the CLIN level to obtain useful data to inform selecting a contract type and choosing the associated contract geometry parameters. (See next the section for a discussion of contract "texture," an attempt to characterize the overall CLIN structure and distribution of contract types by CLIN.)

Table 2. Contract Types in KDB

Contract Type	Count	Value
CPAF	632	$ 51,012,918,376.09
CPIF	665	$ 24,081,165,007.56
FPIF	1,296	$ 70,482,041,186.76
FFP	40,446	$ 167,832,774,252.10
COST & CPFF	6,949	$ 29,755,157,311.01
Other	6,559	$ 52,434,424,590.58
Total	**56,547**	**$ 395,598,480,724.10**

Once the backfill effort is complete, we will be able to mine KDB for the historical distribution of basic contract geometry parameters including: target fee percent; ceiling price percent; share ratios (over and under); min and max fee percents; and base fee and award fee percents. Since these are not yet in KDB in a form that can be queried, our analysts have to go back to the original source documents (contracts and mods) to extract this information.

A key desired computation is the initial and final ROS – or equivalently Return on Cost (ROC) – for all CLINs. Among other things, this would provide significant insight into historical on- and off-shareline growth, since the former degrades profitability but the latter typically does not.

Contract "Texture" and Categorization: Initial Approaches and Results

We wanted to develop a characterization of contract "texture" that takes in account the following broadly defined attributes: *scope* (total $ value); *complexity* (number and variety of CLINs, mixture of programs and/or phases across CLINs); *volatility* (number and frequency of mods); and *homogeneity/heterogeneity* (variety of contract types across CLINs). (DO-type contracts represent an added layer of complication.)

The initial steps toward identifying different contract textures were to first create a latent variable which would incorporate quantitative data from the contracts database. Mahalanobis distance is a reliable metric which normalizes data to calculate relative distances between points. For a vector of observations, x, the formula below shows how the Mahalanobis distance, $D_M(x)$, would be calculated:

$$D_M(x) = \sqrt{(x-\mu)'S^{-1}(x-\mu)}$$

where μ is the vector of means and S is the covariance matrix. Notice that S^{-1} is the inversion of the covariance matrix and that this rescales the sum of squared differences between x and μ.

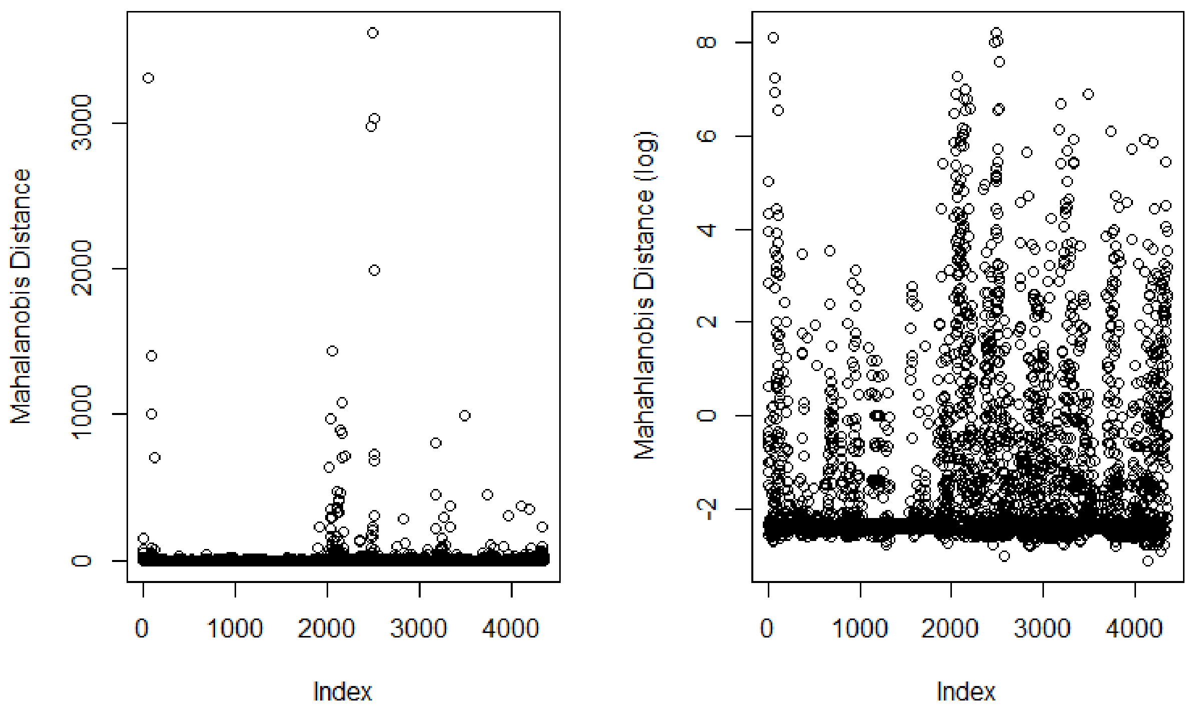

Figure 5. Log Transformation on Mahalanobis Distance

The Mahalanobis metric utilized variables Mod Count, CPIF Count, FPIF Count, FFP Count, COST&CPFF Count, Other Count, CPIF $, FPIF $, FFP $, COST&CPFF $, and Other $. CPAF Count and CPAF $ were excluded in this approach because of their sparsity in the data and the added complexity when the inverted covariance matrix was computed. Figure 1 shows that the log transformation was imposed to reduce the multiplicative scale of these distances to an additive scale.

The next step was to implement a K-means clustering algorithm on our latent variable, Mahalanobis Distance (log). K-means is an unsupervised method of classification in which each observation belongs to a cluster with the nearest mean. The number of clusters are pre-determined to be “k” the algorithm to assign observations to one of the k clusters takes a nearest neighbor classifier approach where the objective function is shown for our vector of observations, x.

$$\arg\min_{s} \sum_{i=1}^{k} \sum_{x \text{ in } S_i} ||x - \mathrm{m}_i||^2$$

In this algorithm, s is the collection of all points for each of the k clusters and m_i is the mean of the points in s_i.

For our analysis we considered clusters of size three and seven. Using this metric to separate our data in multiple levels makes clustering by categorical variables easier and more representative of similar data points. A sample of initial results is shown below.

Figure 6. Contract Texture Clusters by Phase

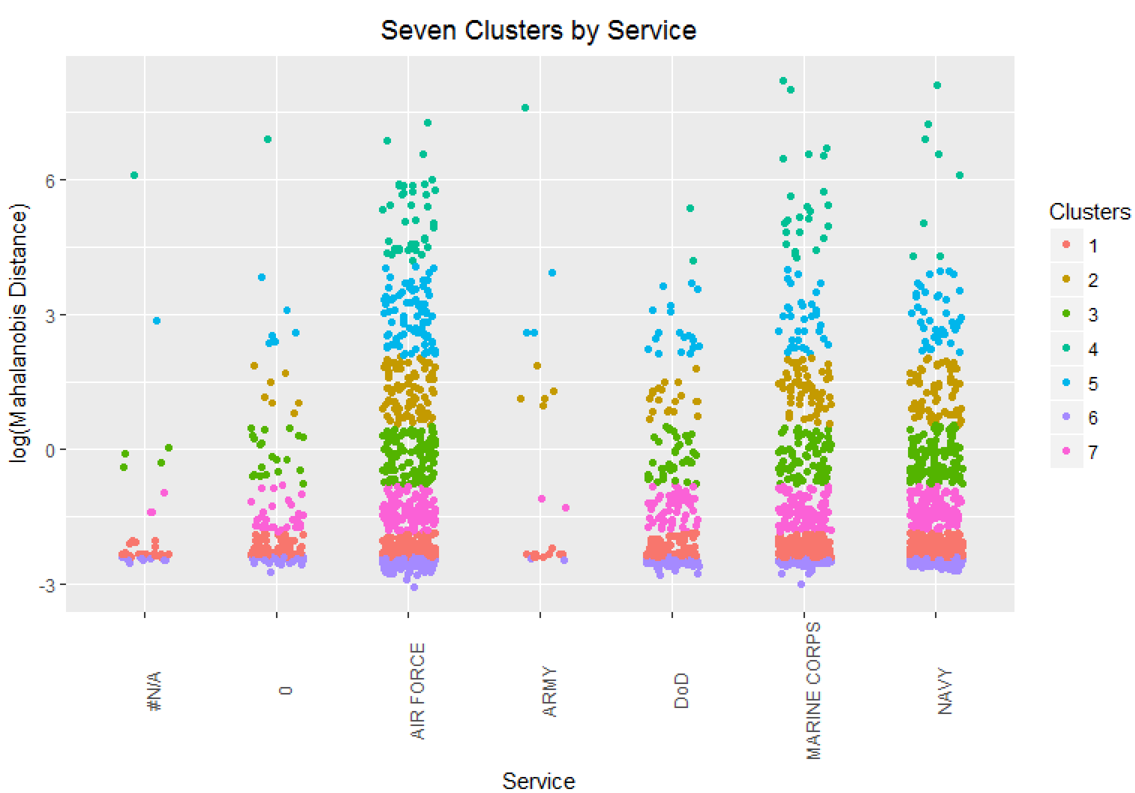

Figure 7. Contract Texture Clusters by Service

Coming Full Circle: Reconciling Historical Cost Growth Analysis and Contract Risk Modeling

There are two major tasks ahead if we wish to tap the full potential of KDB for the confluence of cost analysis, risk analysis, and contract management. The first is to increase the fidelity of historical cost growth analysis by taking into account the interaction of cost growth and contractual mechanisms, specifically mods and contract geometry.

On-Shareline Cost Growth and Implications for Price and Margin

For purposes of this discussion, all of the "Big Four" contract types (FFP, FPI, CPIF, and CPFF) have sharelines. The FFP is essentially a 0/100 share ratio, with the Contractor bearing all of the cost risk. Conversely, CPFF is essentially a 100/0 share ratio, with the Government bearing all of the cost risk. (See discussion of Universal Contract Type in Risk-Based ROS paper.)

"On-shareline" cost growth by definition occurs on one or more existing CLINs, where in a contract mod revises the projected final cost without modifying the underlying contract geometry. This causes revisions in both the price to the government and the contractor's margin, as the cost change (stereotypically an overrun) "runs up" the shareline. This type of cost growth should correspond to the COST and SCHEDULE mods within KDB, but this has yet to be comprehensively verified.

Off-Shareline Cost Growth and Implications for Price and Margin

Conversely, "off-shareline" cost growth occurs when "new work" is added to some combination of existing and new CLINs in a "profit-neutral" manner. That is, the contract geometry is modified to add both cost and fee in the same proportion as in the original CLIN(s). This type of cost growth should correspond to the BASELINE and TECHNICAL mods within KDB, but this has yet to be comprehensively verified.

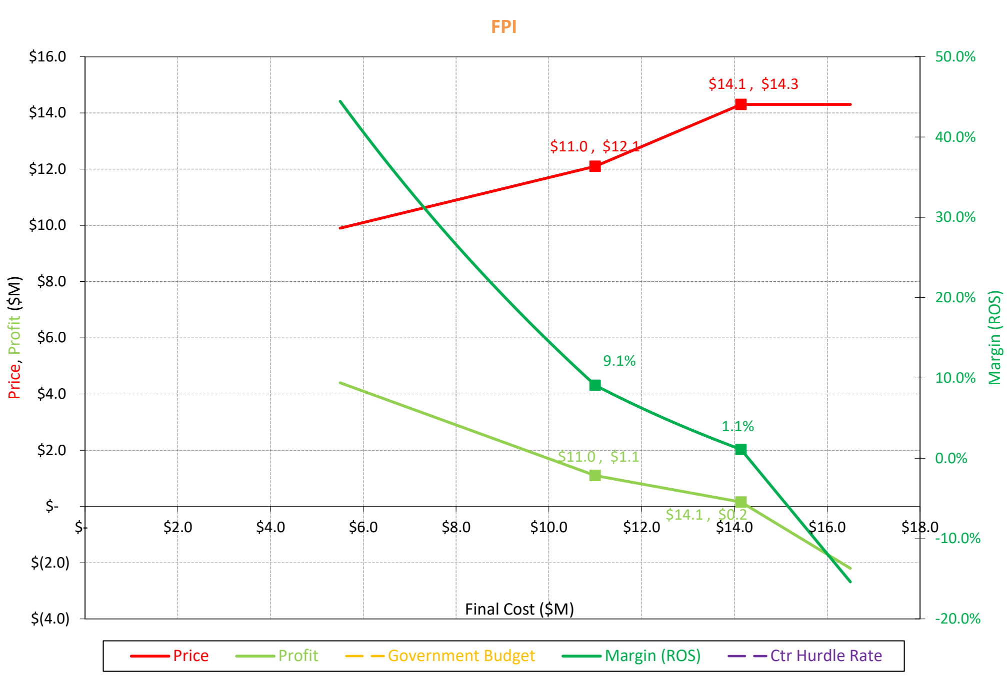

Figure 8. FPIF Example with New Work

The impact of off-shareline growth is illustrated by adding $1M in additional work to our example FPIF contract at the same 10% target fee.

Analysis of Historical On- and Off-Shareline Cost Growth

We will begin with a couple of case studies using the KDB data, and as the cost/fee breakout backfill effort proceeds, we will be able to provide comprehensive summaries of on- and off-shareline mods and their impact on ROS and final price.

Assessing Appropriateness of Contract Types and Contract Geometry

Ultimately, we wish to be able to make recommendations about which contract types to implement in which situations, as well as recommendations on the associated contract geometry parameters.

Historical Cost Growth by Contract Type

KDB has an associated Visual Analysis Tool (VAT), which can easily produce aggregate historical cost growth metrics by contract, program, contractor, etc. We can also use the KDB Pivot Tool to generate historical cost growth by contract type. While such summaries are useful and informative, it would be simplistic to conclude that one contract type is inherently better than another for all situations. The challenge in disentangling the historical data is that we would expect *a priori* that each contract type was sometimes applied appropriately and other times applied inappropriately.

Ecclesiastes: To Every Contract Type, There Is a Season

To every thing there is a season, and a time to every purpose under the heaven:
A time to be born, and a time to die; a time to plant, a time to reap that which is planted;
A time to kill, and a time to heal; a time to break down, and a time to build up;
A time to weep, and a time to laugh; a time to mourn, and a time to dance;
A time to cast away stones, and a time to gather stones together;
A time to embrace, and a time to refrain from embracing;
A time to get, and a time to lose; a time to keep, and a time to cast away;
A time to rend, and a time to sew; a time to keep silence, and a time to speak;
A time to love, and a time to hate; a time of war, and a time of peace.

Our thesis, which largely accords with conventional wisdom, is that the greater risk is applying an inappropriate contract type to a given program phase and scope of work. Diagnosing this with greater fidelity requires a mixture of better metadata (see CADE Integration below) and expert judgment. That being said, we should be beware of “hindsight is 20/20” retroactive diagnoses that rationalize what may or may not have gone better had a different course of action had been chosen. (Like cost analysis, contract management is an “unrepeatable experiment.”)

Next Steps for Implementing the Contract Risk Modeling Framework

This paper lays out a path for additional data collection, fusion, and analysis that will enable KDB to reach its full potential in underpinning the data-driven Contract Risk Modeling Framework.

Expanding Contracts Database Through Automation

With any statistical data analysis, accruing additional data points will generally strengthen the results by tightening confidence intervals through the “power of n,” for example. Because data collection priorities have been directed by the study sponsors over the years, KDB is particularly robust for certain programs and system types (also called "commodities"), such as fixed-wing aircraft, missiles and munitions, and wheeled and tracked vehicles. By contrast, KDB has more sporadic data for some system types, such as electronics, and almost none for others, such as ships. For the current effort, space systems are among those being added to the database. By more comprehensively populating other system types, not only do we increase the overall number of usable data points in the database, but we enable comparison between system types to determine whether there are any statistically significant differences between them.

The Promise of Automation. To date, the population of KDB has been a manually intensive process. Individual PDF files – contracts and mods – must be downloaded from the Electronic Document Access (EDA) system and reviewed by analysts to extract all the required data. Some automation steps have been taken in the past, including developing a web crawler to systematically download all the files for a given contract from EDA, and utilizing the Federal Procurement Data System (FPDS) to provide cross-checks for total price change by mod, but these provided only marginal improvements in overall productivity.

The current Automation process uses Python scripts to transfer the contents of each computer-generated PDF file (contract or mode) into a text file (.txt). The text file can then be "cleaned" and parsed to extract the requisite information and data. The scripts work in "batches" to process a group of documents and populate the data into an Excel file, which can then directly feed the master Microsoft Access database. In addition to Python, the process uses other scripts written in R, and Jupiter to link the two platforms in order to collect the proper data. The benefits of this tailored approach are that we will be able to greatly increase the speed and accuracy of data acquisition. Note that the analyst role is transformed from one of (largely rote) data entry to a combination of programming (to define and implement the automation) and quality assurance (to check and augment the results of the automation).

We will continue to refine and implement this automation process throughout calendar year 2017. This approach leverages computer-generated PDF files, but it would be even better to "cut out the middleman" of the PDF and tap directly into the data fields that were used to generate the contract. The different buying commands within DoD use a myriad of different contract writing tools, but the Defense Procurement and Acquisition Policy (DPAP) organization has worked with the community to develop a Procurement Data Standard (PDS)[11]. The PDS specifies the exact XML schema for all information needed to generate a valid contract or mod. When PDS information is directly available for a given contract, from EDA or otherwise, we should be able to develop a repeatable mechanism to ingest the requisite data into KDB even more efficiently. Metadata tags such as classifying mods by Growth Category and mapping CLINs to work breakdown structure (WBS) for the nonce remains the purview of the proverbial "analyst in the loop."

It remains to be verified whether PDS is sufficient to support direct insight into contract geometry. Put another way, it makes a different whether the contract geometry parameters are captured in individual

[11] "Procurement Data Standard and Other Enterprise Initiatives," Defense Procurement and Acquisition Policy (DPAP), http://www.acq.osd.mil/dpap/pdi/eb/procurement_data_standard.html.

data fields (preferred) or whether the contract geometry is simply described as part of a free-form text field (still requires manual data extraction).

Leveraging Data Fusion Via CADE Integration

As noted in the brief discussion of CCDRs and the insight they provide into profitability on FFP contracts, there is tremendous possibility in data fusion between KDB and the current contents of the CADE back-end relational database, particularly cost data from CCDRs; EVM data from Contract Performance Reports (CPRs) and Integrated Program Management Reports (IPMRs); and budget, quantity, contract funding, and contextual major program event data integrated from the Defense Acquisition Management Information Retrieval (DAMIR) system (http://www.acq.osd.mil/damir/).

Revisiting and Reconciling Program and Contract Cost Growth

Some studies in the past (cite NAVAIR Coleman, Summerville, et al.) have noted that historical cost growth analyses using SARs and those using contracts produce similar results. While these claims ring true, it bears a closer and more detailed look. It should be relatively straightforward to map (by contract number and CLIN) KDB contracts to CSDR and EVM reporting contracts in CADE, which should include all "major contracts" noted in the SARs (also available in CADE). On a program-by-program basis, and in aggregate (broken out by Service, Commodity, etc.), we should be able to compare the latest SAR cost growth numbers (see the aforementioned Andrews, et al.) with those from KDB. The hope is that this will yield not only analytical insights but also lesson learned for a tighter integration of cost and risk analysis and contract management functions going forward.

Supplementing Contracts Data with CSDR Data

While the SAR data in CADE present the greatest opportunity for cross-checking historical cost and schedule growth metrics with KDB, the Cost and Software Data Reporting (CSDR) data – CCDRs in particular – present the greatest opportunity for supplementing the cost and profit/fee breakout in KDB, particularly for FFP efforts.

Refining Contract Texture Metrics and Contract Categorization

We cost and risk analysts worship in the church of historical data while still acknowledging potential concerns such as the proverbial self-fulfilling prophecy (as captured in the "should cost" vs. "will cost" duality of BBP) and "driving by the rearview mirror." That being said, we will endeavor to use an artful combination of automation, systematic data analysis, and targeted subject matter expert (SME) insight to try to untangle appropriate vs. inappropriate use of contract types in the mass of historical KDB data. As we refine our characterization of contract texture, we can develop rubrics that can serve as guides to encourage appropriate use of contract types and contract geometries, and conversely, to provide warnings of potential inappropriate use.

Summary

This paper, building on both previous papers and the labor of love that is KDB, has laid the groundwork for an integrated Contract Risk Management Framework. By reconciling traditionally disparate cost analysis, risk analysis, and contract management viewpoints, and by leveraging the tremendous wealth of data in the ever-expanding KDB and its “sister” data source in CADE, we will be able to provide powerful tools to PCOs and cost analysts alike to ensure that we negotiate contracts using the right contract type and the right contract geometry at the right time. This will result in value for money for the U.S. taxpayer; equitable (but not excessive) profits for U.S. corporations and their shareholders; stable programs for the acquisition workforce; and effective delivery of capability to the warfighter to preserve freedom and democracy around the globe!