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!