SYM-AM-18-107

PROCEEDINGS
OF THE
FIFTEENTH ANNUAL ACQUISITION RESEARCH SYMPOSIUM

THURSDAY SESSIONS
VOLUME II

Acquisition Research:
Creating Synergy for Informed Change

May 9–10, 2018

March 30, 2018

Approved for public release; distribution is unlimited.

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

Towards Game Theoretic Models for Agile Acquisition

Scott Rosen—is the Chief Engineer for the Operations Research Department of the MITRE Corporation with more than 10 years of application experience. He holds a PhD and MS in Industrial Engineering and Operations research from The Pennsylvania State University and a BS in Industrial and Systems Engineering from Lehigh University. [srosen@mitre.org]

Kelly Horinek—is a Principal Analyst in the Center for Acquisition and Management Sciences Division of The MITRE Corporation in McLean, VA. She holds a BBA in Marketing from Texas A&M University, and is a Certified Professional Contracts Manager and a former Contracting Officer. [khorinek@mitre.org]

Alexander Odeh—is Lead Economist and Business Analyst in the Center for Acquisition and Management Sciences Division of The MITRE Corporation in Chantilly, VA. He holds a bachelor's degree in Economics from McGill University and a master's degree in Public Policy from George Washington University. [aodeh@mitre.org]

Les Servi—is a Group Leader in the Operations Research Department of The MITRE Corporation. He holds a PhD and MS in Engineering from Harvard University, and an ScM and ScB in Applied Mathematics from Brown University. He is a Fellow of INFORMS, a former editor of three journals, and a past member of two Defense Science Board studies. [lservi@mitre.org]

Andreas Tolk—is Technology Integrator for the Modeling, Simulation, Experimentation, and Analytics Division of The MITRE Corporation in Hampton, VA. He holds a PhD and MS in Computer Science from the University of the Federal Armed Forces, Munich, Germany. He is a Fellow of the Society for Modeling and Simulation and senior member of ACM and IEEE. [atolk@mitre.org]

Abstract

Game Theory has been applied to drive strategies for contract negotiations in the business world. This paper investigates the idea of applying game-theoretic utility models and strategies to provide a means to illuminate better contracting tradeoffs for the government. This additional insight is intended to provide strategies that move potential contractors into the government's preferred negotiation point and expedite the decision-making process in acquisition. The case studies presented in this paper focus on developing accurate utility functions that would enable such a game theory framework.

Introduction

Game theory has been a dominant research paradigm for studying conflict, bargaining, and negotiations for more than 50 years. It is widely applied throughout business to develop strategies that reflect priorities and tradeoffs. The government has an opportunity to leverage game theory in the federal acquisition system to improve outcomes and increase the agility of government acquisitions. As programs become more technical and complex, game theory can help decision makers identify strategies and leverage information to make data-driven decisions that reflect government priorities and tradeoffs. This paper explores the applicability of game theory to the federal acquisition process and provides a framework to help decision makers identify critical attributes and develop implementable negotiation strategies.

Industry plays a high-stakes game of survival as they act as a testbed for new technologies, and refine processes and strategies in their efforts to compete. Business, competitive, and technological pressures on industry necessarily drive a rapid pace of change and decision making. Industry has successfully applied game theory to develop business strategies that reflect for strategic decisions and negotiations in the business world. The government experiences similar decisions as competing missions, interests, and strategies actively shape acquisitions and budgets. Game theory provides proven insights and approaches that help both industry and the government develop baselines and strategies to create mutually beneficial solutions.

The federal acquisition process is governed by a system of clearly defined rules and regulations codified in the Federal Acquisition Regulation (FAR; 2017). The codification, publication, and adherence to a uniform acquisition system establishes a common understanding or common knowledge of the rules of engagement. Common knowledge, a central tenet of game theory, encourages industry to develop and execute rational business strategies that differentiate solutions and reflect tailored cost, schedule and performance tradeoffs. This creates the framework for achieving best-value through competing strategies and decisions. As a rulebook, the FAR ensures fairness and transparency in the acquisition process for all players, including both industry and the government. Structure and process of federal acquisitions seem to be well positioned to leverage game theory.

In game theory, successful negotiations require clear communication of the attributes. Federal source selections adhere to this principle by the mandatory disclosure of evaluation criteria as key discriminators or attributes. By advertising its source selection criteria and relative order of importance, the government signals its tradeoff considerations. Industry acts as players in the game by tailoring and offering solutions to the government to meet these considerations. One of the initial applications of game theoretic concepts is therefore helping the government identify and develop the key attributes or criteria as well as their importance relative to each other. To support such applications, game theory assumes rationality and that players attempt to maximize the outcome or their utility. Defining the utility for the government is one of the main challenges, as rarely all criteria can be measured in monetary values. Industry seeks to maximize their expected outcome or utility (e.g., profits, market share, etc.) by tailoring solutions that reflect the government's evaluation criteria. Similarly, the government maximizes its expected outcome or utility of industry solutions through its best-value tradeoff considerations. The government can leverage game theory to develop and execute negotiation strategies to improve decision making under uncertainty and contract performance.

Table 1 summarizes some of the similarities between game theory and the government source selection process and shows how the government inherently implements several key aspects of game theory.

Table 1. Similarities of Game Theory and Government Source Selection

Game Theory Principles	Government Source Selection
Players know the "rules of the game"	Federal acquisition process governed by well-defined rules and regulations codified in the Federal Acquisition Regulation (FAR)
Requires clear communication of attributes, priorities and outcomes	Mandatory disclosure of evaluation criteria that will be used to evaluate the proposal and their relative importance
Players are rational and seek to maximize their expected outcome or utility	Government maximizes expected outcome or utility through best-value tradeoff

Applying a game theory approach allows the government to objectively manage risk without compromising its cost, schedule, and performance tradeoffs. It provides insight into decision attributes and negotiating strategies that move industry into a preferred negotiation point while also considering the government's best value tradeoff constraints.

Quantitative Decision Support in Acquisition

The current federal acquisition system typically follows a structured and serial process outlined in the FAR to guide the government in navigating the complexities and cumbersome nature of the source selection process. This regimented system ensures both a fair and transparent acquisition and compliance with oversight stakeholders and the FAR. Unfortunately, these standard acquisition practices can result in the government spending too much time and money in negotiating and establishing contracts. Fortunately, FAR 1.102, Statement of Guiding Principles for the Federal Acquisition System, outlines an opportunity to introduce the agility and efficiencies of game theory by allowing strategies, practices, or procedures that are in the best interests of the government that are not specifically limited or prohibited by the FAR, Executive Order, or regulation. By enhancing negotiations and discussions, game theory complements the fairness and transparency provisions of the FAR. Game theory and utility theory offer an innovative and agile approach for driving bids closer to the desired attributes of most value, thereby saving time and money through improved negotiations and contract outcomes. FAR 1.102 grants government agencies tremendous acquisition flexibility, but the risk-averse nature of some agencies and acquisition professionals may limit their desire to leverage this flexibility. We are pursuing with this research a quantitative decision support framework that is intended to reduce risk and introduce acquisition agility and efficiencies for government and offerors. This game theory approach is consistent with FAR 15.305, Proposal Evaluation, as evaluations may be conducted using any rating method or combination of methods, including color or adjectival ratings, numerical weights, and ordinal rankings. Additionally, this approach is consistent with the April 1, 2016, Department of Defense (DoD) Source Selection Procedures (Section 2.3: Develop the Request for Proposals) where evaluation criteria may be quantitative, qualitative, or a combination of both. Although numerical or percentage weighting of the relative importance of evaluation criteria may not be used in DoD, assigning quantifiable or value tradeoffs in evaluating an offeror’s proposal is allowable and harmonious with the game theory approach.

Applicability of Game and Utility Theory for Acquisition Support

The application of game theory and utility theory can help facilitate the decision-making process in acquisitions. Utility functions provide a framework to translate player preferences into mathematical functions to which standard optimization techniques can be applied. Over a finite set of tradeoffs, there is a utility function that represents a rational preference ordering. This allows decision-makers the insight to tradeoff attributes and criteria based on their expected utility and affords decision makers a framework to make decisions among several alternatives that may result in several possible uncertain outcomes (e.g., what is the likelihood that Proposal X best matches the cost/schedule/performance tradeoff).

While government solicitations or Request for Proposal (RFP) identify the evaluation criteria and relative importance, the qualitative and subjective nature of evaluations may result in suboptimal tradeoff and selection analysis. By providing a mathematical framework, utility models have the capability to generate initial bids that reflect the government’s underlying preferences, help to compare and balance alternatives, and are transparent to everyone partaking in the bidding process. They have the potential of providing additional insight to objectively evaluate the government’s critical attributes and tradeoff considerations of the desired product. Moreover, through integrating the government’s utility function with cost constraints established by industry, a process for generating better initial bids can be established. One high level framework, originally suggested in Simon and Melese (2011), presents such a concept and is based on executing the following high level technical steps:

- Formulate the government’s preferences that are modeled as utility functions parameterized by critical non-cost attributes, criteria or discriminators;
- Publicize government utility functions in RFP for industry to formulate and submit cost functions parameterized by the critical non-cost attributes, criteria or discriminators;
- Assess the uncertainty distribution or likelihood of success of various solutions; and
- Evaluate and optimize government objective function subject to industry provided cost functions and government utility functions to select preferred alternative subject to uncertainty distribution.

A wealth of research questions, however, present themselves when further designing and implementing this type of a framework in a real-world acquisition scenario, the first of which pertains to the type of utility model and the accurate and efficient calibration of the model for use as the foundation of such a framework. This paper focuses on utility models that can be used in this framework.

Survey of Game Theory Literature Relevant to Acquisition

Every scientific endeavor begins with a review of solutions and constraints that help to better define the research to be conducted. Therefore, in the early stages of the underlying research, our team conducted a literature survey. While we originally searched for applications of game theory in the context of procurement and related government activities, we realized quickly that we needed to address two additional topics as well.

As game theory and the optimization using game theoretic approaches relies heavily on the underlying value function, we extended our search to include utility theory with emphasis on utility functions that not only measure monetary values, but also express priorities and intangible preferences.

These led us to the third topic, multi-criteria decision making, as these methods are needed to support the utility-value functions that then can be used to optimize decisions and strategies using game theoretic approaches. The next three subsections will present a small subset of the literature evaluated by us with focus on the approaches we utilized in our study. This selection is neither complete nor exclusive.

Literature on Multi-Criteria Decision-Making

One of the most challenging tasks in acquisition is the selection of the criteria that have to be evaluated to reflect the preferences of the government. Wallenius et al. (2008) provide a good overview of the various methods that are currently in use, placing them into a historical perspective as well, as the same author group also conducted a similar review in 1992. Their overview is written from a management perspective. A more engineering leaning perspective is given among others by Parnell, Driscoll, and Henderson (2011).

Velasquez and Hester (2013) conducted a literature review and analysis of multi-criteria methods. They observe that outranking methods, which were prevalent in early approaches, were overtaken by value measurement approaches. Further they show that deficiencies can be overcome by combined approaches, although this requires a clear understanding of the advantages and disadvantages of the individual approaches, which are captured in a summarizing table in their conclusion section. Their paper assessed the more common methods of Multi-Criteria Decision-Making in order to benefit practitioners to choose a method for solving a specific problem, and they state clearly that this can only be the first step in selecting the right approach.

Agarwal et al. (2011) provide an alternative viewpoint on the selection of the best Multi-Criteria Decision-Making method by focusing on the proper evaluation and selection of suppliers, which is highly relevant in acquisition as well. An additional insight provided by them is the need to evaluate the suppliers based on the inputs of the strategic, functional and operational levels. They present that the

> implication of lean manufacturing and popularly used JIT approach has forced the researchers to shift the focus from the efficiency based model to quality based approach. The single criterion approach of the lowest cost supplier is no more accepted in this challenging and continuously changing environment. Thus, price or cost shifted down the line with respect to its importance in evaluating the suppliers, while the quality and delivery performance climbed up the hierarchy. (Agarwal et al., 2011, p. 808)

This insight is relevant for the government as well and needs to be addressed in the selection of the appropriate methods.

Both recent literature reviews show that there is no universal best solution, but that the selection of the best method is determined by the problem and may even require the use of problem-specific hybrid solution that require an in-depth knowledge of the problem as well as of the tools.

Literature on Utility Theory and Utility Functions

While the literature highlighted in the previous section focuses on the aggregation of multiple criteria in support of decision making, the references given in this section were evaluated regarding the definition of utility functions to reflect the preferences of the decision makers. Slantchev (2012) provides sound definitions of preferences and utilities to support decision making, including those to be made under uncertainty. As he is writing for political scientists, explanations and examples are easy to follow and do not require an in-depth education in gem-theoretic mathematical foundations.

If data is available that reflects preferences of earlier decision-making processes for either side of the negotiating partners, the methods and algorithms described by Afrait (1967) are still relevant. We assumed to be able to find more on the application of big data methods in support of the definition of utility functions, but it seems that this is still a topic of ongoing research and not predominant methods emerged so far.

An interesting variant for multi-issue closed negotiations addressing multi-time as well as multi-lateral negotiation strategies is described by Matsune and Fujita (2017), who developed not only the concept, but demonstrated it in an agent-based simulation environment. Theoretically, nothing speaks against applying these ideas for acquisition specific challenges as well, but we did not see any applications in this domain within our survey. What makes the application described in this paper so interesting is the ability to learn the opponents' utility information from observing their bidding choices within a strategy.

While the mathematics behind utility theory and utility functions is well understood, how to elicit the knowledge about their preferences from decision makers is still a challenge in itself. Our survey did not reveal any predominant strategy.

Literature on Game Theoretic Application for Government Solution

Obviously, every game theoretic insight can be applied to support government solutions better, but two of the evaluated papers deserve special attention, as they directly apply game theory to acquisition and government decision making.

Levenson (2014) provides an overview of the constraints of DoD procurement, showing why typical solutions from commercial markets are often not applicable and lead to undesired and unforeseen results. He describes the effects of fixed price and competitive price contracts and comes to the conclusion that

> only when one or more competitors offer innovations that truly reduce the costs of development and production does the government substantially benefit from competition over sole-source procurement without the adverse side effects of cost overruns. Distinguishing between true innovation and optimistic cost estimating, however, can pose a challenge for DoD acquisition officials. (Levenson, 2014, p. 437)

Blott et al. (2015) compiled a set of auction and game theory based recommendations for DoD acquisitions by synthesizing literature into specific military acquisition categories: procurement with unknown cost and no risk, items with known costs and existent but understood stochastic risk, and items with unknown costs and/or unknown stochastic risk. Some examples further evaluate if multiple competing vendors participate, and if the lot to be procured from several bidders, potentially at different stages of the project.

In summary, the literature survey provided sufficient examples of successful applications, but also the need for continuous research, in particular on how to elicit preferences and utilities from decision makers and apply these methods in a multiple criteria environment under the special constrains of acquisition.

Optimization With Game Theory

Before we go into the details of specific research conducted, the following section shall give an overview of the general concepts that will be addressed in the Selected Approaches section and the Case Studies and Results section.

The following optimization problem drives the application of the utility model. A decision maker has to choose from a set of solutions provided by vendors. The solution is defined by a set of weighted attributes. Furthermore, each vendor is involved with the mathematical optimization that is specific to their own individual cost constraints.

$$\text{Max } V(x) = \sum w_i v_i(x_i)$$

subject to: $\sum c_i x_i < B$

where: x_i = is the level for attribute i

w_i = the weight for attribute i

v_i = the single attribute value function for attribute i

c_i = the offeror cost function for attribute i

B = budget constraint for maximizing utility i

Solving this optimization formulation allows for vendors to generate bids reflecting their specific cost constraints. This yields initial bids that are more consistent with the government's preferences based on the levels of the key non-cost attributes of interest. The solutions to the above optimization formulation allow for stronger initial bids by the interested vendors. These solutions aren't necessarily final solutions or final bids but are an efficient means to getting the bidder close to what the government is looking for. This can more efficiently set up the next stage of proposal tweaking and negotiation on both sides. The rigor of this approach also allows for unambiguous documentation of the negotiation, selection, and provides means for repetition and further evaluation.

Moreover, RFP language can make it difficult for potential bidders to extract out what the most important attributes are for the government (e.g., when too many attributes are included and the evaluation criteria are unclear). Using this quantitative mathematical programming formulation instead allows for bidders to move directly towards those key attributes through an automated means.

To show the applicability of game-based approaches as discussed in these introductory sections, we selected three approaches to evaluate in more detail, which is the topic of the rest of this paper.

Selected Approaches

After conducting the literature survey, we applied three candidate approaches in our research. We selected them due to their perceived potential in being implemented in an acquisition procedure:

- Best/Worst Method (Rezaei, 2015),
- Multi Swing Method (Schmidt, 2017), and
- Functional Dependency Network Analysis (Garvey, Pinto, & Santos, 2014).

Beside their potential for application in an acquisition setting, all three approaches have calibration procedures that are not overly burdensome to the decision maker. They also complement each other. Testing of these methods, as discussed in the Case Studies and Results section, will further reveal the features of the acquisition scenarios where each approach does well. In order to conduct our research, we applied all three approaches to a small decision problem that involved just five attributes to get decision maker accustomed to the steps and procedures needed to be conducted. Furthermore, we applied the Best/Worst method to a larger, 20 attribute problem. These test cases are discussed in more detail in the Case Studies and Results section.

Initial collaborations and discussion with a government sponsor identified three best practices or considerations that impacted our utility function assessment procedure and resulted in the application of multiple assessment techniques.

The first is that the level of effort in developing the assessment procedure must be commensurate with the size, scope, and complexity of the acquisition. A day-long interview process to fit a model may be realistic for a highly complex multi-billion-dollar Acquisition Category (ACAT) 1 program but is not realistic for all acquisitions, and surely not for supporting a smaller research effort like ours. In contrast, a one-hour discussion may be sufficient for many complex acquisitions. Decision makers must balance competing objectives and may not have the luxury of time or resources to support a lengthy and involved process to support the development and calibration of assessment procedure. A long and drawn out initial assessment procedure may result in fatigue and complacency, which may lead to inconsistencies in preference articulation.

The second consideration is that assessment procedures must be adaptable so that they can be effectively applied to decision makers who are either more quantitative or qualitative in nature. However, our research showed that most acquisition professionals are comfortable with relative importance and prefer qualitative descriptions of their preferences. Introducing descriptive adjectives in place of numerical values, in many questions, can help alleviate this issue. Finally, there are many acquisition situations where there is a large attribute set that influence the decision. The size of this attribute set can be overwhelming for any decision maker. Therefore, preference modeling methods must be able to screen out attributes of minimal significance to isolate the critical non-cost attributes and the critical tradeoffs between those attributes. This supports an acquisition best practice of focusing on critical non-cost attributes to avoid diluting the importance of key discriminators.

Overview of Best/Worst Method and Extensions

The Best/Worst method originates from Rezaei (2015) and this research has extended the approach to work more smoothly for cases where there are a large number of attributes at hand and when the attributes are binary in nature (result in either a 0/1 or yes/no value). One of the Best/Worst method's features is its ability to perform calibration in a short series of questions. Moreover, these questions have the ability to be phrased to not be overly burdensome to the decision maker. From our observations, having simple and clear acquisition questions to identify key discriminators facilitates the acquisition and conforms to best practices.

Consistent with source selection practices, the procedure for the Best/Worst method starts with selecting the attributes or discriminators that effect the decision. Then feasible ranges are assigned for each of these attributes. The next step is the assignment of weights for each attribute reflecting the preferences and importance. This applies specifically to the attribute to identify key discriminators and does not apply numerical weights to proposals in the source evaluation process. This step begins with selecting the most important attribute as well as selecting the least important attribute. From there, comparisons are made to understand the relative importance of the most important attribute to each of the other attributes. In a similar manner, comparisons are then made to assess the relative importance of the least important attribute to each of the other attributes.

The question phrasing to the decision maker is the key to getting this approach to work effectively. The decision maker needs to be directly asked how much more important is the most important attribute for each of the other attributes individually. Mapping qualitative scales to numerical scales was shown to work well in our studies for preserving rank order. For instance, levels, such as, "just as important," "slightly more important," "more important," "significantly more important," and "extremely more important" were applied with good success while being mapped on a scale of 1–5.

The end goal of the Best/Worst assessment procedure is to obtain a preference function in the form: $V(x) = w_1 v_1(x_1) + w_2 v_2(x_2) + \dots \dots w_n v_n(x_n)$. The Best/Worst procedure primarily focuses on the weights. Suggestions in this paper for extending to the assessment of the single attribute utility functions $v_1(x_1)$ focus on fitting a function across sample points for each individual attribute. Sampling can be effective with just four points on the utility curve. When doing a qualitative mapping, those points can be referenced as the min, midpoint, target, and max. On a [0,1] scale those reference points were mapped to values of 0, 0.5, 0.75, and 1 respectively. The qualitative assessment questions can first focus on the target. Here the question is asked, "What is the value of this attribute that you would really want to have?" Then the level representing satisfactory for the attribute is assessed: "What level for this attribute is acceptable and would not hinder my use? It can be considered being like a minimum requirement that is not ideal but gets the job done." Then the maximum level for the attribute can be assessed: "What is the level for the most functionality that you could possible handle need—any more wouldn't make life any better." Finally, the minimum level for the attribute is assessed: "What is the maximum attribute level where there is zero utility or where you would have absolutely no use for this product if this attribute was at this level."

The Best/Worst method was extended to a large number of attributes (>20). In acquisitions we observed with our government sponsors, the number of attributes was typically quite large. The Government Acquisition Case Study With Large Number of Attributes section provides details on the application of the Best/Worst method extension to a government acquisition study. For large number of attributes, the procedure was updated in the following manner:

1. First do pairwise comparisons across adjacent pairs of attributes start at attribute #1 and then work down the attribute list.
2. Bin the attributes based on whether the attributes were more important than two attributes, one attributes, or no attributes. End up with three bins: prime, mid, low.
3. Reassess attributes in each bin to make sure they are in the right place.
    a. Ask for best and worst for each bin.
    b. Do pairwise comparison of best in mid and low bin with worst in the higher-level bin.
    c. Repeat 3A and 3B until no more changes are made.
4. Identify the attributes for inclusion into the Best/Worst method
    a. Take all attributes in prime bin.
    b. Take best and worst in mid bin.
    c. Take best and worst in low bin.
5. Best/Worst method is then implemented on attributes in the prime bin.
6. Best/worst method is then implemented on all other attributes kept above.
7. Ask the level of difference between the worst attribute in prime and the best in mid. This level of difference then becomes the difference level for the weights in prime bin and the weights in the remaining bins and the weights are then scaled accordingly.

After these assessment procedures are made the weights for the preference function can be solved through the optimization outlined in Rezaei (2015). The pairwise comparisons given at the beginning of the assessment procedure can also be used to solve for the weights more effectively as well as for validation of the results.

Multi-Swing Rollup Method

The Multi-Swing Rollup Method (MSRM) was developed by MITRE as a new aggregation method for multi-attribute decision problems (Schmidt, 2017). As discussed in the Literature on Multi-Criteria Decision-Making section, rolling up multiple values into one representative value is a general challenge, as already discussed in our literature survey. The MSRM is using a generalized addition tallying organization (GATO) approach. While the classical approach uses the weighted sum of the contributing decision factors, MSRM/GATO uses a non-linear combination in areas in which the simple addition leads to counterintuitive results.

MSRM starts with the definition of multi-swing tables to collect data and combine getting weights and utility functions in one user-driven process. These multi-swing tables are then multiplicatively rolled up and calibrated to fit to a percentage scale. The four steps of the methods are as follows:

1. selecting and quantifying the metrics for each contribution;
2. defining a scale for quantifying the overall score;
3. constructing the multi-swing tables for each contribution;
4. constructing and calibrating the rollup function.

Selecting and quantifying the metrics for each contribution starts with identifying the qualities the user is interested in. The result is a quality tree that identifies the contributions and the metrics used to quantify them. Examples are the resolution (metrics) for the display (contribution), or the battery capacity in minutes (metrics) that keeps the device functional (contribution). These examples will become clearer with the application presented in the Cell Phone Example section.

Defining a scale for quantifying the overall score of the attributes (not numerical scoring of proposals) ensures consistency when assessing the overall value increase or decrease when evaluating the individual contributions. MSRM recommends using a mapping of generally understood expressions to numerical values, such as ideal = 100%, very good = 90%, good = 70%, indifferent = 50%, poor = 30%, very poor = 10%, and not acceptable = 0%. The scale does not have to be symmetric. It is more important that it reflects the weighting priorities and preferences of the user.

Constructing the multi-swing tables is conducted for each contribution, starting with defining baseline with typical and acceptable values for each contribution. For each contribution, we define next a set of swing scores that can be better or worse than the baseline. For each contribution, a set of swing scores spanning all values that can occur in the selection process are collected and the swing rows constructed. If the value of a contribution is a show stopper, e.g., the battery life is too short to support operational use of the item, it is marked as such. In acquisition settings, every attribute that falls under a minimal value becomes a show stopper.

The baseline and all swing rows are then captured in one multi swing table. In this table, in each row only one of the values is changed in comparison with the baseline, so that a comparison with the baseline can be used to access an overall score using the expressions identifies in step 2 of the MSRM. While the baseline may be seen as good, a less screen resolution may decrease the value to indifferent, poor, or may even become a show stopper, while longer life span may result in a very good overall value.

Constructing and calibrating the rollup function uses the multi-swing tables as its foundation, As each row in the multi-swing table captures how the overall value changes when we swing one contribution at a time, a multiplicative roll-up approach can now be applied to compute how the value changes when several of such changes occur at the time. If, for example, the resolution decreases, resulting in a change value decrease of 20%, and the battery life gets shorter as well, decreasing the value by 10%, then the occurrence of both changes should result in a decrease of 28%. The idea is to multiply the individual effects to generate the combined effects.

While the approach naturally results in the elimination of all show stoppers (as the multiplicative approach results in a zero whenever one of the contributions is not acceptable), the positive results can multiply up to more than 100%, which can be addressed using rescaled proportion retention multipliers that ensure that no combination exceeds the 100% limits.

One of the remaining challenges is the combinatorial explosion with the increasing number of contributions. Our initial application shown in the Case Studies and Results section was limited to five attributes, but still required more than 45 minutes to build all multi-swing tables. On the positive side, the method allows the linear integration of new contributions after the initial set-up: a new attribute can be integrated without having to change the trade-offs between the already existing attributes.

Functional Dependency for Network Analysis

The last approach utilized in our research was originally developed for a systems engineering setting, but due to its general applicability, we decided to include it in our evaluation. The application of the Functional Dependency for Network Analysis (FDNA) methodology involves

1. data gathering,
2. preference inference,
3. quantifying accuracy, and
4. making predictions.

The data gathering step involves constructing an experimental design to capture data on the different attributes of the product in accordance to the decision maker's input. The preference inference step involves proposing specific preference models and using the gathered data to infer the defining parameters which are most consistent with the data. The quantifying accuracy step involves the application of cross-validation to assess the accuracy of the fitted preference model. Finally, the making prediction step entails converting a test case to the form selected in the first step and make predictions using the parameters inferred in the second step.

For data gathering with FDNA it is necessary to create a dictionary of qualitative descriptions of product attributes and an assigned numerical representation to each. In the acquisition setting, the dictionaries are highly reusable according to our experiences, although no study has been conducted to verify this observation. As a general practical matter, many spirals of potential dictionaries should be generated and tested to ensure that the definitions are neither to narrow or too broad so that the decision maker who is being models will be assigning a broad ranges of numerical preference scores to the anticipated set of optional designs. Then a set of optional designs of interest can be generated. Assuming the absence of *a priori* knowledge of the decision maker's preference, the designs are randomly generated.

Motivated by the work of Garvey and Pinto (2009) and Servi and Garvey (2017) two different preference models are included in the approach[1]:

$$f_s(P_s,\gamma) = \min_i[\ P_s^i + \beta_i] \tag{1}$$

or

$$f_s(P_s,\gamma) = \alpha_0 + \sum_i \alpha_i\ \ P_s^i \tag{2}$$

[1] If there were a larger amount of experimental data, it would have been more desirable to use the precise FDNA model documented in the references. Due to the limited size of the data, two different aspects of the FDNA were used for this analysis.

where P_s^i is the numerical level of preference of the ith characteristic of the sth experiment.

With the decision maker's evaluation of different attribute combinations, the values of β_i or α_i computed using the training data, which are rows comprising all attribute values and the resulting evaluation by the decision maker, it is possible to estimate the preferences of the decision maker.

For the case study discussed in the following section, the accuracy using (1) was found to be superior to that using (2), it is recommended that when predicting the comparative preferences of the decision maker to two alternatives, the prediction is made using only (1) and, in the case of a tie, (2) is used to break the tie.

Case Studies and Results

The model assessment procedures are applied and tested on two case studies to test their applicability in the acquisition setting. The first case study, described previously, is a cell phone purchasing example, easily understood by everyone, and was used for internal testing for all three selected approaches. The second case study, described in the section titled Government Acquisition Case Study With Large Number of Attributes, is a real-life acquisition scenario that added complexities not existing when methods like these are tested and presented in literature. We only discuss the Best/Worst method example exemplifying the challenges.

Cell Phone Example

A case study for buying a cell phone was first used to work through the question phrasings of each method in a simpler environment. The test subject or decision maker was given the scenario of purchasing a new smartphone, such as an iPhone or Samsung Galaxy. The decision maker was made known that there are dozens of alternatives to choose from. They are then made to envision that there are five main attributes that will affect their decision as to which smartphone to buy. Here are the five attributes that were told that effected their decision:

A. Weight [0–5 pounds]

B. Lifespan [0–10 years]

C. Screen resolution [0–4,000p]

D. Processing speed [0–10x]

E. Storage amount [0–500 Gigs]

Included above are the ranges of values that each attribute can take. The ranges are meant to exceed what is true in reality. For example, one cannot obviously have a cell phone of no weight and there are no phones in the market that weigh 5 pounds. After reading through the example and taking the role of the purchaser, a series of questions was conducted about their preferences in accordance to the assessment procedures for the three preference functions tested.

Application of Best/Worst Method

The application of the Best/Worst method began with the assessment of the weights for the five attributes of interest. The test subject was asked a series of trade-off questions and identified processing speed as the most important attribute and lifespan as the least important attribute. Through a series of questions relating the level of importance of each attribute with respect to processing speed the following vector was obtained: A_b = [5,4,1,1,2]. The numerical values in this vector specify how much more important processing speed was with respect to weight, lifespan, screen resolution, and storage amount. The scale for importance is on a range of 1–5. So, the first value of 5 represents processing speed being extremely more important than weight. As indicated by the 4 assigned to the second slot, processing speed is considered significantly more important than lifespan. The remaining values show that processing speed is equally important to screen resolution, to itself, and more important than storage amount. These are the same mappings to numerical values introduced in the Selected Approaches section.

After this the same questions regarding relative attribute importance were asked with respect to the attribute noted as the least important. This resulted in the following vector A_w = [1,1,4,5,3]. As shown in Rezaei (2015). These two vectors representing relative importance between each attribute and the best and worst attribute, respectively can be used to perform a least squares approximation to solve for the weights. The following numerical weights were obtained: [0.04, 0.12, 0.28, 0.39, 0.17].

The final step was to solve for the single attribute value functions pertaining to each attribute. Here the test subject was asked for each attribute to specify the minimum, midpoint, target, and maximum values for each of the five attributes and the question wordings introduced in the Overview of Best/Worst Method and Extensions section were applied. Table 2 presents the values obtained for the min, midpoint, target, and max for each attribute.

Table 2. Subject Responses to Min, Midpoint, Target, and Max Levels for Each Attribute

	Min	**Midpoint**	**Target**	**Max**
Weight	1 lb.	0.75 lb.	0.5 lb.	0.33 lb.
Lifespan	1 year	3 years	4 years	6 years
Screen resolution	400p	720p	1080p	2000p
Processing Speed	0.5x	1x	3x	5x
Storage Amount	64 gigs	128 gigs	250 gigs	500 gigs

The values in Table 2 are used to solve for the single attribute value functions $v_i(x_i)$ for all five attributes in this case study. For this case study, a simple second order polynomial was applied for fitting these single attribute value functions and the method of least squares was used for fitting. The weights for all five attributes can then be integrated into the single attribute value functions to obtain the following function for the preference model:

$$V(x) = w_1 v_1(x_1) + w_2 v_2(x_2) + w_3 v_3(x_3) + w_4 v_4(x_4) + w_5 v_5(x_5) \quad (3)$$

In order to test the accuracy of the Best/Worst method, a series of comparisons across six purchasing alternatives was performed. The following pairwise comparisons across all combinations of these purchasing options below were performed by the decision maker.

A. [2, 2, 2000, 0.75, 256]

B. [0.5, 5, 720, 2, 128]

C. [4, 1, 4000, 1, 64]

D. [1, 4, 720, 1, 256]

E. [2, 3, 1080, 2, 256]

F. [0.5, 3, 4,000, 4, 64]

The results of these pairwise comparisons can be applied to generate a ranking. The rankings obtained here are compared to rankings generated through the preference model sampled under these same alternatives. In addition, the proportion of the pairwise comparisons that are consistent between the decision maker and model was measured. There were 15 different combinations of pairwise combinations resulting from the six scenarios above. The preference model resulting from the calibration involving the Best/Worst method resulted in consistency amongst all 15 pairwise comparisons. That meant when the subject specified, for example, that alternative B was more preferable than alternative A, that the preference model outputted a larger value when inputting in the attribute levels for alternative B than when inputting the attribute levels for alternative A. As naturally follows, the rankings for all six alternatives were consistent as well. The case study demonstrated promise in the Best/Worst method to generate an accurate model in a short amount of time.

The entire assessment procedure was done in roughly 45 minutes for this scenario involving five attributes. The process can be supported by tools, and our research resulted in the definition of new tool support that is currently prototypically developed to support data collection, calibration, and presentation of the results. The acquisition professional doesn't have to provide all the details captured in this section, but should provide the comparisons and evaluate the rankings.

Application of MSRM Method

The Multi-Swing Rollup Method (MSRM) was applied in the same setting as the Best/Worst method, using the same experts to conduct the experiments. Using the same attributes as enumerated in Table 2, we defined one positive and one negative swing for each attribute, as shown in this Table 3.

Table 3. Attributes and Swing States (Green Variations Are Positive, Red Are Negative)

	Baseline	Variations	
(W) Weight	0.5	0.33	0.75
(LS) Lifespan	2	1	4
(SR) Screen resolution	1080	720	2000
(PS) Processing speed	4	2	5
(SA) Storage amount	256	128	500

Next, we defined the utility factor terms to be used to rate the comparisons between the baseline and the swings. In the discussion with the experts and decision makers, we ended up with a table that showed the semantic equivalencies between different families of terms describing comparisons, status descriptions, and grades, which the group was comfortable with (see Table 4).

Table 4. Utility Value Terms

Utility value

significantly better	ideal	A
much better	very good	A/B
better	good	B
little bit better	above average	B/C
average/mid-point	average	C
little bit worse	below average	C/D
worse	poor	D
much worse	very poor	E
showstopper	showstopper	F

Having five attributes with two swing states each does result in eleven entries. Using the utility terms, each entry was compared individually with the baseline to identify the overall change in utility by changing one attribute. Table 5 shows the individual utility contributions in percent that resulted from our discussions with the experts.

Table 5. Utility Value Changes in Percent Relative to the Baseline

	Weight	Lifespan	Screen resolution	Processing speed	Storage amount
negative	-6.7	-26.7	-26.7	-46.7	-20.0
positive	13.3	20.0	33.3	33.3	26.7

Using this information, the full multi-swing rollup table with all 243 entries can be created. The resulting table contains all possible combinations of multi-swings plus the baseline. The resulting overall utility is calculated by multiplying the individual changes. When ordering the table, the entry with all negatives obviously is the lowest, and the entry with all positives the highest, but all possible permutations in between are listed as well, showing the ranking of all alternatives, including the selected subset used in the Best/Worst method. As we derived the same ranking, this should at least consistency in the evaluation, no matter which of the first two approaches was used.

The assessment procedure conducted with our decision makers was shorter than for the Best/Worst method, but only because several of the results could be reused. In an internal comparison with in-house experts, the amount of time needed for the first two methods was approximately the same for the cell phone example.

Application of FDNA

In a previous section, we introduced the two different preference models that were motivated by the work of Garvey and Pinto (2009) and Servi and Garvey (2017). First, the term needs to be defined. The dictionary shown in Table 6 is comparable to the terms defined in Table 5 for the utility terms used in the MSRM.

Table 6. A Dictionary Assigning a Numerical Preference Level to the Preference Level of Qualitative Characteristics of iPhone

	(W) Weight	(LS) Lifespan	(SR) Screen resolution	(PS) Processing speed	(SA) Storage amount
0 - Crummy	heavy	1	good images	email, word OK	some added apps, best photos
1 - OK	not heavy	2	good for printing	very good non-games, slow video	apps and photos
2 - Good	light	3	great images, good enlarging	good for video	huge for apps and photos, some videos
3 - Great	ultra light	3+	very good for enlarging	everything great	virtually unlimited including videos

Next, we generated possible solutions for the five attributes important for the selection of the cell phone: weight, lifespan, screen resolution, processing speed, and storage amount. Table 7 shows the 27 generated cases, using the index numbers defined in the dictionary to specify the solution. The decision maker that graded the various solutions as captured in the column “evaluator.”

Table 7. The Experimental Training Data

No.	Weight	Lifespan	Screen resolution	Processing speed	Storage amount	Evaluator
1	3	3	0	0	0	**0**
2	2	2	3	2	2	**3**
3	2	2	3	2	2	**3**
4	3	3	1	0	0	**0**
5	2	2	2	1	3	**2**
6	0	1	1	2	3	**1**
7	2	1	0	3	0	**0**
8	1	2	2	1	0	**0**
9	3	2	1	2	2	**2**
10	2	1	2	2	3	**3**
11	3	1	2	0	1	**0**
12	2	1	0	3	0	**0**
13	3	1	1	1	1	**1**
14	1	0	3	2	3	**2**
15	1	3	0	3	1	**1**
16	2	2	0	2	1	**1**
17	1	1	2	2	2	**2**
18	0	3	0	0	1	**0**
19	0	1	1	2	2	**1**
20	1	3	0	1	3	**1**
21	1	0	2	2	2	**2**
22	1	0	1	1	3	**1**
23	1	1	1	2	2	**2**
24	2	0	3	3	3	**3**
25	3	3	3	2	1	**2**
26	2	2	2	3	3	**3**
27	1	0	1	3	1	**1**

Given the data in the Cell Phone Example section, it is possible to exhaustively search for the integer values of β_i most consistent with the data in terms of the mean sum of squares error, $\alpha_0 = 1, \alpha_1 = 8, \alpha_2 = 6, \alpha_3 = 0, and\ \alpha_4 = 0$ as well as analytically solving for the values of α_i most consistent with the data, ($\beta_0 = -1.6204, \beta_1 = 0.2101, \beta_2 = 0.2219, \beta_3 = 0.4290, \beta_4 = .4375, \text{and } \beta_5 = 0.5724$). This leads to mean sum of square error of 0.19 when using equation (1) and a worse mean sum of square error of 0.44 when using equation (2).

For FDNA, however, the more precise approach to quantifying the error is using the method of cross-validation. Here, the values of β_i or α_i are computed using a random set of 8/9 of the data in Table 7 and then the accuracy of the prediction is computing using the 1/9 of the data not trained on. This was repeated numerous times. This lead to the conclusion that the mean sum of square when using equation (1) was 0.19 (and a standard deviation of 0.24) and when using equation (2) was 0.54 (with a standard deviation of 0.23). The conclusion, for this data, is that equation (1) leads to a superior model of this decision maker, which means that $f_s(P_s, \gamma) = \min_i[\ P_s^i + \beta_i]$ is the better model to capture preferences.

These examples given in the Cell Phone Example section exemplify the different application possibilities of the three approaches as well as their strengths and mutual support, reemphasizing the need for having a toolbox of different solutions in support of acquisition decisions.

Government Acquisition Case Study With Large Number of Attributes

An additional case study was also performed with a government sponsor involving an acquisition scenario consisting of 20 attributes to show the scalability of approaches. This is a more challenging case study in that the decision maker must go through a lengthy assessment procedure to make comparisons among a very large set of attributes. Another twist to this problem was that these attributes were binary in nature. Each attribute had an objective value and the government was only interested in if the attribute exceeded that value. So, each attribute has two levels (0, 1) to represent whether it met the objective or not. A new modification of the Best/Worst method was applied as presented in the Overview of Best/Worth Method and Extensions section to handle this scenario of having a large set of attributes. The types of attributes cannot be discussed in this paper, but the implementation of the procedure can be discussed. The first portion of the assessment procedure involved doing pairwise comparisons across the adjacent pairs of attributes starting with the first attribute. The resulting table of the results to these pairwise comparisons is shown to help further explain the approach (see Table 8).

Table 8. Pairwise Comparisons Across Adjacent Attributes

Attribute	1st Comparison	2nd Comparison
1	1 > 2	1 > 20
2	2 < 1	2 > 3
3	3 < 2	3 < 4
4	4 > 3	4 > 5
5	5 < 4	5 > 6
6	6 < 5	6 < 7
7	7 > 6	7 < 8
8	8 > 7	8 > 9
9	9 < 8	9 > 10
10	10 < 9	10 < 11
11	11 > 10	11 > 12
12	12 < 11	12 < 13
13	13 > 12	13 > 14
14	14 < 13	14 < 15
15	15 > 14	15 > 16
16	16 < 15	16 > 17
17	17 < 16	17 > 18
18	18 < 17	18 > 19
19	19 < 18	20 < 19
20	20 < 19	20 < 1

After this initial pairwise comparison is done, the attributes are binned based on whether they were more important than two attributes, one attributes, or no attributes. This results three bins, which are named prime, mid, low, respectively. The resulting bins are shown in Table 9 to further exemplify the approach.

Table 9. Binning of Attributes After Initial Pairwise Comparisons

Prime Bucket	Mid Bucket	Low Bucket
Attribute 1	Attribute 2	Attribute 3
Attribute 4	Attribute 11	Attribute 6
Attribute 8	Attribute 7	Attribute 10
Attribute 5	Attribute 9	Attribute 12
Attribute 13	Attribute 16	Attribute 14
Attribute 15	Attribute 17	Attribute 19
	Attribute 18	
	Attribute 20	

The next step is to reassess the attributes in each bin to make sure they are allocated properly. This is done through asking the test subject to first identify the most important and least important attribute in each bin. Then pairwise comparisons are done between the most important attribute in each bin and the least important attribute in each bin. After any reassignments are made, the test subject is then asked to identify again the most important and least important attribute in each bin. If there are any changes to these assignments, the process is repeated and the most important attribute in each bin is compared again with the least important attribute in each bin. The process repeats until no attributes can be exchanged between bins in this manner. The final binning of attributes for this case study is shown in Table 10.

Table 10. Final Binning of Attributes After Extra Validation Questions

Prime Bucket	Mid Bucket	Low Bucket
Attribute 1	Attribute 2	Attribute 18
Attribute 4	Attribute 11	Attribute 3
Attribute 16	Attribute 7	Attribute 6
Attribute 5	Attribute 9	Attribute 10
Attribute 13	Attribute 8	Attribute 12
Attribute 15	Attribute 17	Attribute 14
	Attribute 20	Attribute 19

At this point the Best/Worst method can be executed across a subset of these attributes. The first step in doing this is to identify the most important and least important attribute in each of the three bins. Note that due to the initial pairwise comparisons being made between adjacent attributes it is not necessary to compare every attribute with the most important attribute and every attribute with the least important attribute. In the prime bin, each attribute is compared with the most important and least important attribute. Then for the mid bin, the only the most important and least important attribute are compared with the least important attribute in the prime bin. Likewise, these attributes are compared with the most important attribute in the low bin. Then finally, the most important and least important attribute of the low bin are compared to the least important attribute in the mid bin. These measures of relative importance are again done on a scale of 1–5. After all of these comparisons are made, there is enough information to perform a least squares estimation to approximate the weights for all 20 attributes.

The nice feature about having attributes with a binary value is that it is not necessary to assess a single attribute utility function for each attribute. If the attribute meets the threshold then the utility is mapped to a value of 1 and if it does not meet the threshold it is mapped to a value of 0. Therefore, the weights can be used with 0, 1 terms for the attribute values directly to result in this equation for the preference model: $V = \sum_i w_i t_i$ where t_i = 1 if the value for attribute i meets its threshold level for the objective.

The most notable result of this new assessment procedure was that the interaction time with the test subject to train this preference model with 20 attributes was done in less than one hour. For this amount of time, it is expected that the decision maker can stay engaged for the duration of the assessment procedure and remain accurate and limit inconsistencies. Another promising feature of this method is that the initial pairwise comparisons can be held out from training of the model and used for validation. When reserving eight for validation, the model resulted in matching the decision maker in seven out of eight (87.5%) of the test cases.

As we did not conduct any further experiments on the degree to which these results are scalable or generally applicable in other domains as well, we do not want to oversell any results, but the observed trends do make sense within our experience. Due to time and budget constraints, we did not conduct similar experiments using the MSRM or FDNA methods.

Conclusions

This paper suggests the idea of applying game-theoretic models as a foundation for a quantitative decision-making framework in support of acquisition. Through successful calibration of utility functions, we suggest there is a strong potential to develop a framework that can more effectively illuminate strategies that move industry into the government's preferred negotiation point and expedite the decision-making process in acquisition.

The case studies presented in this paper focus on the potential for developing accurate utility functions that would enable such a game theory framework. The government's utility functions, representing their level of preference for attribute levels involved in a proposal, are the cornerstone for enabling such a decision support framework to be utilized effectively and accurately. In this paper, we examined three potential utility function calibration procedures from literature and adapt these procedures to examine the strengths and weaknesses of each approach. The three methods were applied in a case study involving the acquisition or purchasing of a cell phone. All three reveled benefits for different acquisition scenarios. The Best/Worst method showed robustness in handling a small or large number of attributes effectively. The MSRM method demonstrating the ability to capture sharp drops in utility in individual attributes. This is an important feature when some attributes present thresholds where the entire product becomes unusable by the government. The FDNA method showed the ability to work effectively when the decision maker is more qualitative in nature than quantitative.

The Best/Worst method was also extended to be more applicable to acquisition scenarios involving a large number of attributes (20 plus attributes). These situations are more common than the small number of attribute examples that are often provided in the toy examples in the literature. This paper provided an efficient procedure for screening and implementing the Best/Worst method when the attribute set is large. It was tested on a real-world government acquisition example and was shown to be able to calibrate an accurate preference function in under one hour of decision maker engagement time.

Future work involves integrating these utility function assessment procedures into a decision support framework that can enable potential bidders to maximize the fitted utility function with respect to their own specific cost functions, which are parameterized by the same attributes as the utility function. The sampling procedure of these cost functions along with the best optimization algorithm to apply is another area of research needing attention. The optimization algorithm must have the ability to generate solutions in near real-time in order for this decision support framework to be usable and effective.

References

Afriat, S. N. (1967). The construction of utility functions from expenditure data. *International Economic Review, 8*(1), 67–77.

Agarwal, P., Sahai, M., Mishra, V., Bag, M., & Singh, V. (2011). A review of multi-criteria decision making techniques for supplier evaluation and selection. *International Journal of Industrial Engineering Computations, 2*(4), 801–810.

Blott, J., Boardman, N., Cady, A., Elliott, J., Griffin, W., Mastronardi, N., & Quinn, P. (2015). *Auction and game theory based recommendations for DOD acquisitions* (No. USAFA-AM-14-189). Colorado Springs, CO: Air Force Academy.

Federal Acquisition Regulation (FAR), 48 C.F.R. ch. 1 (2018). Retrieved March 2018 from https://www.gpo.gov/fdsys/pkg/CFR-2017-title48-vol1/xml/CFR-2017-title48-vol1-sec1-101.xml

Garvey, P. R., & Pinto, C. A. (2009, June). Introduction to functional dependency network analysis. In *The MITRE Corporation and Old Dominion, Second International Symposium on Engineering Systems* (Vol. 5). MIT, Cambridge, Massachusetts.

Garvey, P. R., Pinto, C. A., & Santos, J. R. (2014). Modelling and measuring the operability of interdependent systems and systems of systems: Advances in methods and applications. *International Journal of System of Systems Engineering, 5*(1), 1–24.

Levenson, W. J. (2014). DoD acquisition—To compete or not compete: The placebo of competition. *Defense ARJ, 21*(1), 416–440.

Matsune, T., & Fujita, K. (2017, July). Weighting estimation methods for opponents' utility functions using boosting in multi-time negotiations. In *Proceedings of the IEEE International Conference on Agents (ICA)* (pp. 27–32). IEEE.

Parnell, G. S., Driscoll, P. J., & Henderson, D. L. (Eds.). (2011). *Decision making in systems engineering and management* (Vol. 81). Hoboken, NJ: John Wiley & Sons.

Rezaei, J. (2015). Best-worst multi-criteria decision-making method. *Omega, 53*, 49–57.

Schmidt, B. K. (2017). *Generalized Addition Tallying Operation (GATO).* Bedford, MA: The MITRE Corporation.

Servi, L. D., & Garvey, P. R. (2017). Deriving global criticality conditions from local dependencies using functional dependency network analysis (FDNA). *Systems Engineering, 20*(4), 297–306.

Simon, J., & Melese, F. (2011). A multiattribute sealed-bid procurement auction with multiple budgets for government vendor selection. *Decision Analysis, 8*(3), 170–179.

Slantchev, B. L. (2012). *Game theory: preferences and expected utility* [Technical report, Political Science Courses]. San Diego, CA: University of California.

Velasquez, M., & Hester, P. T. (2013). An analysis of multi-criteria decision making methods. *International Journal of Operations Research, 10*(2), 56–66.

Wallenius, J., Dyer, J. S., Fishburn, P. C., Steuer, R. E., Zionts, S., & Deb, K. (2008). Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead. *Management Science, 54*(7), 1336–1349.

Disclaimer

This work was supported by the internal MITRE Innovation Program in support of Agile Connected Government. The views, opinions and/or findings contained in this report are those of The MITRE Corporation and should not be construed as an official government position, policy, or decision, unless designated by other documentation. This work has been approved for public release with unlimited distribution; Case Nr. 18-1141.

ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL
555 DYER ROAD, INGERSOLL HALL
MONTEREY, CA 93943

www.acquisitionresearch.net